Rebuilding the Post-Pandemic Economy

Foreword by **HENRY M. PAULSON, JR.** and **ERSKINE BOWLES**

Edited by **MELISSA S. KEARNEY** and **AMY GANZ**

ASPEN ECONOMIC STRATEGY GROUP
aspen institute

The Aspen Institute
2300 N Street, N.W.
Suite 700
Washington, DC 20037

Published in the United States of America in 2021 by The Aspen Institute

Printed in the United States of America
Publication Number: 21/002

Rebuilding the Post-Pandemic Economy

Foreword by **HENRY M. PAULSON, JR.**
and **ERSKINE BOWLES**

Edited by **MELISSA S. KEARNEY**
and **AMY GANZ**

ASPEN ECONOMIC STRATEGY GROUP
aspen institute

DECEMBER 2021

Acknowledgements

We are grateful to the many individuals who were instrumental in the production of this volume. Members of the Aspen Economic Strategy Group provided thoughtful discussions that motivated the topics addressed in the volume. We thank the authors for their contributions to this volume and their commitment to informing the public policy conversation with rigorous research.

We also thank the individuals and organizations whose generous support makes the work of the Aspen Economic Strategy Group possible. Many staff invested countless hours in producing this volume: Emily Vincent provided superb project management; Krista Martin of Krista Martin Design contributed cover art and formatted the charts and graphs throughout the book. Kate Wheeling provided excellent proofreading and editing assistance. We are also grateful for the help of Joan Burch, Kelley Folino, Andy Morimoto, and Sogand Sepassi.

Henry M. Paulson, Jr.
Co-Chair

Erskine Bowles
Co-Chair

Melissa S. Kearney
Director

Amy Ganz
Deputy Director

Table of Contents

Foreword

By Erskine Bowles and Henry M. Paulson, Jr.

After suffering the worst economic shock since the Great Depression last year, the American economy is recovering in fits and starts. While many businesses are reopening their doors and thriving, others are struggling with tenuous demand, supply constraints, and higher labor costs. Americans are traveling, dining out, and resuming other activities that weren't possible before vaccines. Yet, remaining uncertainty about the course of the virus continues to hamper a full return to normal activity.

The COVID-19 pandemic reinforced and exacerbated many of the biggest structural economic challenges in our society. It precipitated the largest economic relief and stimulus spending in US history and rewrote the playbook for responding to future economic crises. The pandemic also transformed the way that millions of Americans live and work, with automation, e-commerce, and telework all playing a bigger role.

The pandemic and its aftershocks reignited not only the perennial debates about the appropriate role and size of government, but also present new and urgent questions about how the post-pandemic economy will take shape.

What are some initial lessons we can take away from the novel government programs that were deployed to provide economic relief and stimulus? What kinds of investments do we need to make to our infrastructure so that it is once again the envy of the world? After a year of widespread school closures, what have we learned about the role of K-12 education in perpetuating or reducing social and economic inequities? And how should American trade policies evolve to promote economic recovery and strengthen America's role in the global economy?

None of the answers to these questions are predetermined. The choices and actions of policymakers in Washington and around the country can and will make a difference.

This book is an attempt to elucidate some of the challenges and opportunities of the post-pandemic economy. At its core, it underscores that the challenge for economic policymakers is not simply to return to the pre-pandemic economy—rather, it is to rebuild an economy that is more prosperous, dynamic, fair, and resilient to future shocks in the post-pandemic era. We hope that the non-partisan, evidenced-based research and recommendations contained in this volume are helpful towards this end.

Introduction

Melissa S. Kearney* and Amy Ganz**

September 2021

The COVID-19 pandemic plunged the US economy into recession, challenged the survival of millions of businesses, and threatened the economic security of American households. The recession officially lasted only two months, ending in April 2020, but looming economic challenges remain and the path of the post-pandemic recovery is uncertain. The US labor market recovery is slow, global supply chains are disrupted, the pace of vaccination in the United States has stalled, and emerging variants of the virus threaten a return to pre-pandemic normalcy.

The pandemic also ushered in major changes to the US economy, many of which may persist even after the pandemic recedes. The sudden shift to working from home, changes in consumers' preferences and habits, and the acceleration of technology adoption by businesses may have lasting effects on economic growth and inequality—for better or worse. Widespread school closures and the shift to remote instruction has impeded the educational and social development of US children and exacerbated already large disparities in learning, with potential long-term negative consequences.

At the same time, the pandemic and accompanying economic crisis prompted an unprecedented US policy response. Congress authorized trillions of dollars in spending to support businesses and households, staving off business failures and bolstering household income and savings. With aggregate demand now increasing, the US economy faces a new set of challenges—among them a higher inflation forecast driven by both demand and supply factors. The Biden administration and congressional Democrats are calling for trillions of dollars in federal spending on an ambitious package of health, education, early childhood, and climate initiatives, in addition to the $1 trillion bipartisan infrastructure package passed by the Senate in August 2021. Critics worry about the size and scope of the package, as well as the prospect of further deficit spending and higher taxes.

Amidst these domestic challenges, the geopolitical landscape facing the United States continues to shift, in particular with China's rapid ascendance as a major

* Aspen Economic Strategy Group; University of Maryland

** Aspen Economic Strategy Group

economic rival able to wield greater economic and political influence across Asia and the much of the world. The need to maintain American competitiveness in this changing global context highlights the need for well-designed investments at home, in infrastructure, in human capital, and in basic science and technology.

Rebuilding the Post-Pandemic Economy considers several current, major economic challenges facing the nation. Its eight chapters served as background reading materials for the Aspen Economic Strategy Group annual meeting in July 2021 and are now published as a resource for broader policy audiences.

Part 1 consists of five chapters that focus on various elements of the US economic recovery following the Covid-19 pandemic. Chapter 1 highlights productivity gains that could result from the sudden shift to working from home if US households were to have universal access to reliable, high-speed internet. Chapter 2 discusses lessons learned from the novel business recovery programs introduced during the pandemic and their applications for "garden variety" recessions. Chapter 3 presents strategies for preventing long-term unemployment and assisting workers whose jobs have been permanently lost as a result of sectoral reallocation. Chapter 4 discusses the underlying causes of longstanding inequities in the US K-12 education system, which were laid bare by the pandemic, and promising avenues for systemic improvement. Chapter 5 addresses the current state of American trade policy, including reforms to promote American geopolitical interests and economic recovery.

Part 2 consists of three chapters that focus on the US infrastructure agenda. Chapter 6 addresses the economics of infrastructure investment, emphasizing the central role of cost-benefit analysis in selecting projects. Chapter 7 focuses on federal regulatory reforms and infrastructure investments necessary to support the US economy's transition to clean energy sources. Chapter 8 makes the case for greater federal investment in research and development (R&D) based on the extremely high social return on such investments and their role in promoting broader innovation and prosperity.

Part I: The Post-Pandemic Economic Recovery

The unprecedented economic shock caused by the onset of the COVID-19 pandemic in the United States was met with an unprecedented policy response. Congress appropriated more than $5.5 trillion in stimulus and relief funding in a series of legislation listed in Table 1. In addition to these spending measures, a federal eviction moratorium was enacted in March 2020 and subsequently extended by the Biden administration through July 2021. The Federal Reserve also took exceptional steps to stabilize markets by directly purchasing corporate bonds, deploying lending programs to stabilize small- and medium-sized businesses, and providing liquidity to the municipal bond market. In addition, the Fed deployed its usual recession-fighting tools, such as cutting interest rates, relaxing regulations to promote liquidity, and reviving several Great Recession-era lending programs.[1] These measures generally succeeded in stabilizing markets by shoring up business and household balance sheets.

Table 1: Fiscal stimulus and relief measures appropriated in response to COVID-19

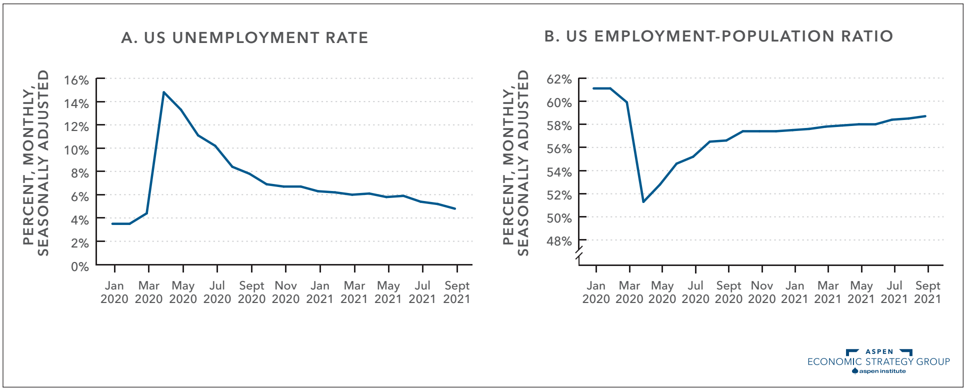
DATE ENACTED	PROGRAM	AMOUNT (in billions $)
March 6, 2020	Coronavirus Preparedness and Response Supplemental Appropriations Act, 2020	$8.3
March 18, 2020	Families First Coronavirus Response Act	$3.4
March 27, 2020	CARES Act	$2,300
April 24, 2020	Supplemental stimulus	$484
December 27, 2020	Consolidated Appropriations Act	$868
March 11, 2021	American Rescue Plan	$1,900
	TOTAL	**$5,564**

Nearly two years after the onset of the COVID-19 pandemic, the US labor market is improving, but still far from a full recovery. The official unemployment rate peaked at 14.8% in April 2020 and has since recovered to 4.8% in September 2021 (shown in Figure 1a).[2] The employment-population ratio has climbed from its low of 51.3% in April 2020 to 58.7% in September 2021, but it remains well below the pre-pandemic level of 61% (shown in Figure 1b). As of August 2021, the economy has restored roughly 18 million jobs, but is still missing an estimated 7 million jobs relative to pre-pandemic trends.

1 See Clarida, Duygan-Bump, and Scotti (2021).

2 Estimates that account for misclassification and the drop in labor force participation put that number a bit higher.

Figure 1: Unemployment rate and employment-population ratio, January 2020–August 2021



Source: Bureau of Labor Statistics (BLS) via Federal Reserve Economic Data (FRED).

The labor market recovery appears to be hindered by a slower than expected rate of transition of unemployed workers into jobs. Though aggregate spending has increased and employers are posting job openings, there are several potential reasons why workers might be hesitant to return to work. These include fears about contracting the COVID virus, potential disincentive effects of enhanced unemployment benefits[3], ongoing childcare challenges affecting parents[4], the financial ability to delay a return to work on account of saving up income over this past year, along with a reexamination of work and life goals after an unprecedented set of experiences. The stalled pace of the labor market recovery raises questions about how long and what it will take to restore employment to pre-pandemic levels.

The pandemic also led consumers and workers to change their habits and routines and businesses to alter their operations. What changes will stick and become permanent and to what economic effect? How is work going to change after the pandemic? What policies would support rather than impede an efficient reallocation process among firms and workers? What should be done to support workers at risk of long-term unemployment and joblessness?

---

3 A recent examination of this explanation by a team of researchers using JPMorgan Chase Institute data finds that the disincentive effects of pandemic UI benefits are empirically small in magnitude (Ganong et al. 2021).

4 A recent paper by Furman, Kearney, and Powell (2021) casts doubt on this explanation, finding that employment challenges particular to parents of young children cannot explain a meaningful share of the aggregate employment decline observed in 2021.

The work-from-home trend and worker productivity

The pandemic changed the way that Americans work, commute, and engage with their physical workplace. Many of these changes are unlikely to fully reverse after the pandemic is over. In Chapter 1, "Internet Access and its Implications for Productivity, Inequality, and Resilience," Jose Maria Barrero (Instituto Tecnológico Autónomo de México), Nicholas Bloom (Stanford University), and Steven J. Davis (University of Chicago) present novel survey data on work from home trends and productivity and simulate how improvements in internet connectivity could lead to improvements in economic productivity.

The authors have been fielding an original survey on work arrangements, productivity, and attitudes toward work from home since May 2020, collecting to date 43,000 responses from working-age Americans. Survey respondents report higher productivity when working from home during the pandemic as compared to when working on employer premises before the pandemic. In previous research, the authors combined this survey data with information about employer plans regarding post-pandemic work arrangements to predict what a re-optimization of work arrangements post-pandemic would look like.[5] They estimated that one-fifth of paid workdays will be supplied from home in the post-pandemic economy and more than a quarter of workdays on an earnings-weighted basis. They estimate that that re-optimization could be expected to boost productivity by close to 5%, largely through saved commuting time.

In this chapter, the authors augment those previous findings with a consideration of how improved internet connectivity would enhance the productivity boost coming from the post-pandemic re-optimization of work arrangements. They estimate that a move to universal access to high-quality home internet connection would boost labor productivity by 1.1%, which implies an increase in GDP flows of $160 billion per year. They additionally estimate that the earnings gains to workers from the associated productivity enhancements would be nearly uniform across income and demographic groups, meaning that productivity improvements would not come at the expense of widening inequality.

The data and analyses of this chapter are highly relevant to ongoing considerations about investments in internet technology. As the nation debates the size and scope of significant new infrastructure spending, the findings suggest a substantial economic payoff for investments that improve internet connectivity for households across the country. However, the authors' analysis does not inform questions about the relative

5 Barrero, Bloom, and Davis (2021)

costs and benefits of such investments in different locations and settings. The need for project-by-project cost-benefit analysis of specific infrastructure projects is an issue that is highlighted in Chapter 6 by Edward Glaeser and James Poterba, described below.

Supporting small businesses through the recovery and future recessions

The Paycheck Protection Program (PPP) was designed to preserve workers' relationships with their employers by providing forgivable loans to small- to mid-sized employers who maintained payrolls for a specified length of time. The goal of the program was to minimize costly and unnecessary labor market separations during a slowdown in economic activity, resulting from public health concerns and mandated shutdowns. A second program, the Main Street Lending Program (MSLP), made government-backed loans available to small and medium sized businesses and were administered by existing lenders who were required to bear a small share of the loan risk. As of September 2021, the PPP provided over 11 million loans valued at nearly $800 billion, with roughly $530 billion forgiven so far.[6] The MSLP provided a much smaller number of loans—1,830—valued at a combined $17.5 billion.[7]

Chapter 2, "Business Continuity Insurance in the Next Disaster," by Samuel Hanson (Harvard Business School), Adi Sunderam (Harvard Business School), and Eric Zwick (Booth School of Business at The University of Chicago) reflects on lessons from the PPP and the MSLP and considers the policy case for small business support during economic and non-economic crises, as well as during the post-pandemic recovery. Their chapter first highlights the unique features of how the COVID pandemic affected small businesses, including the dramatic scale of revenue losses, the large number of firms simultaneously affected, and the lengthy duration of the crisis. They observe that there was a strong economic case for supporting small businesses during this national disaster, beyond social insurance paid directly to workers and households. They additionally observe that many of the market failures that justify business support during the pandemic—such as frictions in capital and labor markets and nominal rigidities in contracts—also justify business support during typical economic recessions, though at lower levels of generosity.

The authors compare the design and implementation of the PPP and the MSLP programs, noting that the "softer" loan terms of the PPP (*i.e.*, more lenient repayment terms) led to a much higher disbursement rate. The PPP disbursed nearly 80% of its allocated funds, as compared to only 3% for the MSLP. The authors' primary

6 SBA (2021)

7 Bräuning and Paligorova (2021)

critique of the PPP's design is that it was not well targeted to businesses experiencing hardship, thus forgivable loans were granted to businesses that did not experience particular hardship or need government support to weather the crisis.

The authors draw on the lessons of the pandemic experience, the PPP, and the MSLP to make recommendations for the design of policy to support businesses during future crises. Their chapter describes a new program concept called Business Continuity Insurance, which they had proposed in a previous paper coauthored with Jeremy Stein.[8] The design of this proposed policy takes seriously the challenge of targeting business support toward firms with the highest private benefit and social insurance value relative to program cost—namely, firms whose operations are severely affected by a current shock, that are unable to smooth the shock on their own, and for which bankruptcies would create substantial spillovers. In practice, the authors recognize that to minimize administrative burden and maximize take-up, such a program should use relatively simplistic targeting that exploits information already available to the government.

Finally, the authors argue that the justification for government small-business support during the economic recovery is considerably weaker than it was during the crisis, when firms experienced dramatically reduced revenue, still had to cover recurring obligations, and access to private market liquidity support was insufficient. They caution against the use of traditional fiscal policy levers targeting firms, such as investment or payroll tax credits, which are not necessarily well targeted. They also express skepticism about proactive policy approaches to encourage reallocation given the uncertainty about how much reallocation is optimal and what the nature of any such reallocation in the economy should be.

Helping the long-term unemployed

The job losses induced by the COVID-19 pandemic were concentrated among workers without college degrees, Blacks, Hispanics, younger individuals, and women. The economic effects of this pandemic could very well exacerbate the relative disadvantage of these groups of workers and worsen inequality. Furthermore, evidence from previous recessions indicates that those who lose their job during a recession and those who remain unemployed for a long time are at risk of experiencing long-term losses in earnings.[9] It will be important to support the return to work of vulnerable workers at risk of long-term unemployment and associated consequences.

8 Hanson et al. (2020)

9 Davis and von Wachter (2011)

Till von Wachter (University of California, Los Angeles) offers answers to the question of what can be done to spur reemployment during the recovery and to assist those workers at risk of long-term unemployment and economic hardship. In Chapter 3, "Data-Driven Opportunities to Scale Reemployment Opportunities and Social Insurance for Unemployed Workers During the Recovery," he proposes that existing programs and services can be effectively scaled to help avoid hardship while further speeding reemployment and assist workers in obtaining better-paying jobs.

First, drawing lessons from a successful outreach program by the California Employment Development Department, von Wachter highlights how existing data residing within state agencies can be used to target eligible participants with information about unemployment insurance (UI), the Supplemental Nutrition Assistance Program (SNAP), Medicaid, Temporary Assistance for Needy Families (TANF), and other programs while also predicting recipients who are at greatest risk of benefit exhaustion. Second, von Wachter suggests improvements to the Short Time Compensation (STC) program, including better marketing to employers, establishing a federal subsidy for firms to use STC rather than laying off workers, and allowing participating workers and firms to utilize subsidized training programs. Third, von Wachter reinforces the need to improve the UI system by automatically triggering benefit increases and extensions and relaxing eligibility requirements during economic downturns. This is a policy idea that has broad support among policy economists and was also put forward in the 2020 AESG task force report on promoting economic recovery after COVID-19.[10]

Finally, von Wachter calls for modernizing reporting requirements by states for UI and upgrading systems to allow for better targeting, improved data for evaluation, and effective real-time decision making by policymakers. The COVID-19 pandemic highlighted the inadequacies of many states' UI data management systems during crises, which in many cases led to significant delays in benefit payments. The changes proposed by von Wachter would make the UI system more resilient and put the government in a better position to support workers and household balance sheets during future crises.

Disparities in K-12 education in the United States

The pandemic situation and the associated school closures laid bare longstanding disparities in US K-12 educational outcomes. Many observers worried that the rapid switch to remote learning would exacerbate socioeconomic gaps in student performance, with children from lower-income families less likely to thrive doing

10 Furman et al. (2020)

school from home. Early data confirms these worries. Rebuilding after the pandemic will require addressing learning deficits acquired over the past year, while at the same time, reducing longstanding disparities in educational experiences. In Chapter 4, “Addressing Inequities in the US K-12 Education System,” authors Nora Gordon (Georgetown University) and Sarah Reber (University of California, Los Angeles) examine the root causes of persistent inequities in the K-12 education system, disparities in schooling outcomes, and ways the system could be amended to improve student performance and address persistent racial and class gaps in outcomes.

Gordon and Reber argue that making equitable progress will require shoring up fundamentals throughout the K-12 system so that schools and teachers are better prepared to serve all students well. The authors provide a comprehensive picture of the sources of inequities in the K-12 education system and highlight promising levers for intervention.

The authors encourage policymakers to return to the fundamental inputs of schools: staff, peers, curriculum, and the environment—including physical infrastructure—in which students learn. They caution against becoming distracted by trendy new ideas and educational buzzwords. Although the K-12 system is a foundational pillar of economic opportunity in the United States, the authors readily acknowledge that it alone is not responsible for all social and economic disparities. The authors emphasize that discrimination in housing and labor markets, policing, environmental hazards, and neighborhood violence are all stressors that significantly inhibit student learning.

Multilateral economic cooperation and US trade policy

International engagement and trade are critical for advancing US geopolitical interests and promoting economic growth. Our nation’s approach to international engagement and trade also has important implications for the US economic recovery after COVID-19. Following the isolationist stance of the Trump administration and its upending of the rules-based trade system, the Biden administration faces the challenges of rebuilding multilateral cooperation with allies, defining the United States’ trade relationship with China, and supporting political constituencies most harmed in the past by free trade.

In Chapter 5, “America and International Trade Cooperation,” Chad Bown (Peterson Institute for International Economics) discusses ways the Biden administration might improve US trade policy with respect to China and its western allies. Bown observes that thus far, the Biden administration is taking a “worker-centric” trade approach, but has yet to articulate a specific policy toward China and other trade partners. Bown suggests that clarity is needed around whether the United States’ stance toward China has permanently changed to a “noncooperative” relationship, which would require abandoning the rules-based system that has been in place since 1947 and negotiating new rules. Alternatively, the United States could maintain a "cooperative" relationship with China and thus stay within existing international trade rules, while tweaking existing agreements.

Bown discusses the US trade system through the framework of cooperation versus noncooperation and optimal policy responses within each of the two scenarios. Bown also highlights policies that would be beneficial regardless of which stance the United States takes, including adjusting the US tariffs unilaterally imposed on China (to the detriment of US producers), patching up relations with western allies, and fixing the WTOs dispute settlement system. Finally, Bown describes areas of mutual interest for the United States, its allies, and China including climate change and global health.

Part II. The US Infrastructure Agenda

Bipartisan support for new infrastructure spending reflects an emerging consensus that infrastructure investments would enhance American economic competitiveness and increase the economy's productive capacity. Sound investments also have the potential to accelerate the US economy's transition to sustainable energy sources, as well as address underlying sources of domestic inequality. As the debate about the size and scope of new investments progresses in Washington, Part II considers what types of infrastructure projects are most likely to foster economic growth and widespread prosperity, the role of cost-benefit analysis in driving project selection, what types of financing mechanisms ought to be pursued, and how to avoid unnecessary or wasteful spending.

The economics of infrastructure

Chapter 6, "Economic Perspectives on Infrastructure Investment," by Edward Glaeser (Harvard University) and James Poterba (MIT), highlights policy lessons from the voluminous research literature on the economics of infrastructure projects. The authors focus primarily on traditional infrastructure projects that involve fixed capital investments associated with the movement of goods or people (as with bridges and roads) and electric or digital content (as with the electricity grid, broadband, and fiber optic cables.)

The authors stipulate that the optimal level of infrastructure investment should be determined by a project-based consideration of the costs of acquiring infrastructure capital with the benefits of using it. They contrast this economic approach with the "engineering" approach of defining infrastructure need without a consideration of marginal costs and marginal benefits. They propose that the United States should rely on cost-benefit analysis in determining which projects to undertake and they suggest the creation of a nonpartisan federal agency to perform such analyses.

The authors highlight wide variance in the benefits and costs of individual investment proposals, even within categories such as roads and bridges. They also challenge the prevailing narrative of America's "crumbling infrastructure," by highlighting that interstate highways today are smoother, fewer bridges are structurally deficient, and dam collapses are less frequent than in the past.

Glaeser and Poterba note that American infrastructure costs are very high by international standards—and hence the optimal amount of infrastructure is likely to be lower relative to other countries. The authors suggest several strategies that could help to control costs, such as improving procurement practices and project management. Maintaining existing infrastructure rather than building new projects is often more cost effective, despite the "ribbon-cutting bias" of many politicians.

Finally, the authors discuss various mechanisms for financing infrastructure spending. They highlight the advantages of user fees, which, along with congestion pricing, can improve the efficiency of infrastructure use. They recognize the potential for user fees to burden low-income users, but note that some user fees are progressive, such as airport fees, and that policymakers could offset the impact of user fees on low-income users with targeted rebates or vouchers.

The role of infrastructure in the transition to clean energy

Chapter 7, "Challenges of a Clean Energy Transition and Implications for Energy Infrastructure Policy," by Severin Borenstein (Berkeley Haas School of Business) and Ryan Kellogg (University of Chicago Harris School of Public Policy), discusses the major barriers that policy needs to overcome in order to successfully execute a transition to a low-carbon energy system at reasonable cost.

The authors emphasize that new infrastructure investment will be needed to support the transition to a low-carbon energy system. They explain that a core problem with transitioning to a clean energy system is that wind and solar generation are intermittent. Provision of reliable, zero-carbon emission supply will therefore require combining wind and solar resources with investments in dispatchable zero-emission sources (such as nuclear or hydroelectric sources), long-distance transmission, demand flexibility, and storage technologies. Given uncertainties about technological progress, the authors argue that broad incentives that do not discriminate across zero-emission resources—such as carbon pricing, clean energy standards, or clean energy subsidies—will be essential for directing capital toward cost-effective investments in clean energy infrastructure.

The authors recognize that incentives for clean energy infrastructure will be insufficient to meet the climate challenge. The authors also call for new investments in research and development, noting that rapid innovation is needed. They claim that exporting new technologies that reduce carbon emissions will be essential for the United States to help combat the *global* climate crisis. New charging infrastructure will be necessary for the transition to electric vehicles. However, the authors recommend flexible approaches to government support and suggest that the private sector should take the lead in technological development, charger siting decisions, and business model experimentation.

Managing electricity transmission, storage, and demand flexibility will require new infrastructure and regulatory reforms. A central problem for reliable energy supply is coordinating electricity demand volatility with intermittent generation sources. The authors identify investment in new devices that can time-shift electricity demand to respond to dynamic price signals as a promising approach to addressing this challenge. Regulatory hurdles must also be removed to enable long-distance interstate electricity transmission. Specifically, those seeking to build transmission lines must currently obtain permissions from local authorities along the route to do so. The authors suggest that allowing the Federal Energy Regulatory Commission to have authority over rights-of-way and use of eminent domain, as it does for natural gas transmission, would alleviate this problem.

Borenstein and Kellogg suggest that reforms are also needed in wholesale and retail electricity markets, which are poorly suited for the intermittent nature of renewable energy supply. Traditional pricing in which costs are levied per kilowatt hour do not fit an energy market increasingly reliant on renewables, which requires changing output rapidly depending on demand. However, controlling costs while maintaining the reliability of electricity is a central tension. In addition, many states have financed climate policies by pricing electricity above social marginal cost, which both discourages electricity adoption and is regressive. Finally, the authors discuss the need to decommission legacy fossil fuel infrastructure, which also offers an opportunity to employ displaced oil and gas workers.

Public investment in research and development

US public funding for research and development is at a 60-year low.[11] In Chapter 8, “Science and Innovation: The Under-Fueled Engine of Prosperity,” Ben Jones (Kellogg School of Management at Northwestern University), makes the case for why the United States should invest more public funding in basic science and technology research.

11 See the policy memo “14 Facts about US Investments in Infrastructure and R&D,” by AESG staff Amy Ganz and Emily Vincent.

Jones explores the government's important role in supporting science and innovation in the national interest. The United States massively underinvests in science and innovation at its own peril, Jones argues, since such developments drive long-run productivity growth which are key to rising standards of living, advancements in health, and the ability to respond to crises. Jones quantifies the social returns to R&D investment: for every $1 that is invested, society reaps at least $5 in return. Despite very high expected returns, the economic uncertainty and political risk inherent to federal investment in basic research have fueled skepticism about allocating scarce public funding to R&D. For every CRISPR breakthrough, there is also a Solyndra anecdote.

A second crucial driver of innovation are policies that advance human capital. Jones highlights recent research showing that children exposed to innovators and entrepreneurs are more likely to become one themselves. In particular, girls who move to regions with higher shares of female inventors are more likely to become inventors. Importing talent through immigration is also critical. US immigrants patent more often than native-born Americans and make up a disproportionate share of the science and engineering workforce. Immigrants also account for a disproportionate share of entrepreneurs and are more likely to start companies of all sizes, including high-growth start-ups.

References

Barrero, Jose Maria, Nicholas Bloom, Steven J. Davis, and Brent Meyer. 2021. "COVID-19 Is a Persistent Reallocation Shock." SSRN Scholarly Paper ID 3763443. Rochester, NY: Social Science Research Network. February: https://doi.org/10.2139/ssrn.3763443.

Barrero, Jose Maria, Nicholas Bloom, and Steven J. Davis. 2021. "Why Working from Home Will Stick." National Bureau of Economic Research Working Paper #28731.

Bräuning, Falk, and Teodora Paligorova. 2021. "'Uptake of the Main Street Lending Program,' FEDS Notes." Washington: Board of Governors of the Federal Reserve System (U.S.). https://doi.org/10.17016/2380-7172.2897.

Clarida, Richard, Burcu Duygan-Bump, and Chiara Scotti. 2021. "The COVID-19 crisis and the Federal Reserve's policy response" in Monetary Policy and Central Banking in the Covid Era ed. William English, Kristin Forbes, and Angel Ubide, CEPR Press.

Davis, Steven J. and Till von Wachter. 2011. "Recessions and the Costs of Job Loss." *Brookings Papers on Economic Activity* (2).

Furman, Jason, Timothy Geithner, Glenn Hubbard, and Melissa S. Kearney. 2020. "Promoting Economic Recovery After COVID-19." Washington, D.C.: Aspen Economic Strategy Group. https://www.economicstrategygroup.org/wp-content/uploads/2020/11/Promoting-Economic-Recovery-After-COVID-0615-FINAL.pdf.

Furman, Jason, Melissa Kearney, and Wilson Powell. 2021. "The Role of Childcare Challenges in the US Jobs Market Recovery During the COVID-19 Pandemic," National Bureau of Economic Research Working Paper #28934, June.

Ganong, Peter, Fiona Greig, Max Liebeskind, Pascal Noel, Daniel M. Sullivan, and Joseph Vavra, 2021. "Spending and Job Search Impacts of Expanded Unemployment Benefits: Evidence from Administrative Micro Data," February 11 mimeo, https://conference.nber.org/conf_papers/f152625.pdf, accessed July 22, 2021.

Ganz, Amy, and Melissa S. Kearney. 2019. "Introduction: The Need to Expand Economic Opportunity." In *Expanding Economic Opportunity for More Americans.* Washington, D.C.: Aspen Economic Strategy Group. https://www.economicstrategygroup.org/publication/the-need-to-expand-economic-opportunity/.

Ganz, Amy and Emily Vincent. 2021. *14 Facts about US Investments in Infrastructure and R&D.* Washington, D.C.: Aspen Economic Strategy Group, July: https://www.economicstrategygroup.org/publication/14-facts-about-us-investments/.

Goolsbee, Austan, Amy Ganz, Glenn Hubbard, and Melissa S. Kearney. 2019. "A Policy Agenda to Develop Human Capital for the Modern Economy," Washington, D.C.: Aspen Economic Strategy Group. https://www.economicstrategygroup.org/publication/a-policy-agenda-to-develop-human-capital-for-the-modern-economy/.

Hanson, Samuel, Jeremy Stein, Adi Sunderam, and Eric Zwick. 2020. "Business Continuity Insurance and Loans: Keeping America's Lights on During the Pandemic." University of Chicago Becker Friedman Institute White Paper, April: Business Continuity Insurance and Loans: Keeping America's Lights On During The Pandemic.

Huff Stevens, Ann. 2019. "What Works in Career and Technical Education (CTE)? A Review of Evidence and Suggested Policy Directions." Washington, D.C.: Aspen Economic Strategy Group. https://www.economicstrategygroup.org/publication/what-works-in-career-and-technical-education-cte-a-review-of-evidence-and-suggested-policy-directions/.

Lerman, Robert I. 2019. "Scaling Apprenticeship to Increase Human Capital." Washington, D.C.: Aspen Economic Strategy Group. https://www.economicstrategygroup.org/publication/scaling-apprenticeship-to-increase-human-capital/.

Pew Research Center. 2021. "Fact Sheet: Demographics of Internet and Home Broadband Usage in the United States." Pew Research Center. https://www.pewresearch.org/internet/fact-sheet/internet-broadband/.

SBA. 2021. "PPP Data." U.S. Small Business Administration (SBA). https://www.sba.gov/funding-programs/loans/covid-19-relief-options/paycheck-protection-program/ppp-data#section-header-11.

PART I

THE POST-PANDEMIC ECONOMIC RECOVERY

Internet Access and its Implications for Productivity, Inequality, and Resilience

AUTHORS

Jose Maria Barrero,* Nicholas Bloom, and Steven J. Davis*****

ABSTRACT

About one-fifth of paid workdays will be supplied from home in the post-pandemic economy, and more than one-fourth on an earnings-weighted basis. In view of this projection, we consider some implications of home internet access quality, exploiting data from the new Survey of Working Arrangements and Attitudes. Moving to high-quality, fully reliable home internet service for all Americans ("universal access") would raise earnings-weighted labor productivity by an estimated 1.1% in the coming years. The implied output gains are $160 billion per year, or $4 trillion when capitalized at a 4% rate. Estimated flow output payoffs to universal access are nearly three times as large in economic disasters like the COVID-19 pandemic. Our survey data also say that subjective well-being was higher during the pandemic for people with better home internet service conditional on age, employment status, earnings, working arrangements, and other controls. In short, universal access would raise productivity, and it would promote greater economic and social resilience during future disasters that inhibit travel and in-person interactions.

* Instituto Tecnológico Autónomo de México. Email: jose.barrero@itam.mx

** Stanford University. Email: nbloom@stanford.edu

*** University of Chicago Booth School of Business and the Hoover Institution. Email: Steven.Davis@chicagobooth.edu

Acknowledgements: Prepared for the Economic Strategy Group at the Aspen Institute. Thanks to Melissa Kearney and Amy Ganz for comments on an earlier draft. We gratefully acknowledge Stanford University, the University of Chicago Booth School of Business, Asociación Mexicana de Cultura A.C., the Stanford Institute for Human-Centered Artificial Intelligence, the W.E. Upjohn Institute for Employment Research, and Toulouse Network for Information Technology for financial support.

1. Introduction

Americans supplied half or more of paid workdays from home in spring 2020, 10 times the pre-COVID share. They continued to supply more than 40% of workdays from home through spring 2021.[1] We explain why the shift to working from home (WFH) will endure in Barrero, Bloom, and Davis (2021b). Our analysis and forward-looking survey evidence suggest that WFH will settle at about one-fifth of paid workdays in the coming years, and at higher levels for the well-educated and highly compensated. This abrupt, enormous, persistent shift to remote work has been greatly enabled by home internet access and a host of complementary technologies.

These remarkable developments prompt several questions that we tackle in this chapter. First, if everyone had high-quality, fully reliable home internet service in the coming years, how much would it boost productivity? Second, how much did subpar internet service degrade productivity during the pandemic? Third, would universal access to high-quality, reliable internet service materially increase WFH in the coming years and, if so, by how much? Fourth, would universal access raise or lower earnings inequality? As we explain, the implications for earnings inequality are unclear a priori, even as to the sign of any effect. Fifth, video conversations and virtual meetings yield some of the emotional and psychological benefits that humans normally enjoy in person, raising another question: How do subjective and objective indicators of well-being relate to internet access quality during the pandemic, a time of pervasive (physical) social distancing?

To address these questions, we tap multiple waves of data from the Survey of Working Arrangements and Attitudes (SWAA), an original cross-sectional survey of our design. We have fielded the SWAA monthly since May 2020, thus far collecting 43,000 responses from working-age Americans who earned at least $20,000 in 2019. The survey asks about working arrangements during the pandemic, internet access quality, productivity, subjective well-being, employer plans about the extent of WFH *after* the pandemic ends, and more. The SWAA asks explicitly about working "full days at home." Thus, our WFH measures do not encompass workdays split between home and office or work at satellite business facilities.

Many SWAA respondents report higher productivity when WFH during the pandemic than when working on employer premises before the pandemic. Using SWAA data on the relative productivity of WFH, employer plans about who will work from home in the post-pandemic economy and commuting times, Barrero, Bloom, and Davis (2021b, hereafter "BBD") estimate that a re-optimization of working arrangements

1 Barrero, Bloom, and Davis (2020b), Bick et al. (2020), Brynjolfsson et al. (2020) and Ozimek (2020) provide evidence on the extent of working from home in the spring of 2020. Barrero, Bloom, and Davis (2021b) provide evidence on its evolution at a monthly frequency through April 2021.

in the post-pandemic economy will boost productivity by 4.6% relative to the pre-pandemic situation.[2] The main source of this productivity gain is the savings in commuting time afforded by WFH. The productivity boost reflects a combination of higher productivity when WFH for some workers *and* the selected nature of who works from home in the post-pandemic economy.

The productivity projections in our earlier work are conditioned on the prevailing level of internet access quality, which varies considerably across households. In this chapter, we instead estimate the productivity effects of universal access to high-quality internet service. We approach the matter in two ways: First, using responses to the SWAA question, "How much would your efficiency working from home increase if you had perfect high-speed internet?" Second, using regression models that relate SWAA data on the relative productivity of WFH to internet access quality. Under both approaches, we exploit SWAA data on employer plans for who will work from home in the post-pandemic economy, and how much.

To preview our main results, we estimate that universal access to high-quality home internet service (hereafter, "universal access") would raise earnings-weighted productivity in the post-pandemic economy by 1.1%. To obtain this figure, we combine employer plans for who will work from home, and how much, with self-assessed productivity effects of universal access. For many workers, the implied productivity effect is zero—either because they don't plan on WFH in the post-pandemic economy, already have high-quality home internet service, or don't anticipate a productivity effect in any event. However, some workers who plan on WFH in the post-pandemic economy also expect a productivity gain from better internet service.

Given an aggregate output elasticity with respect to labor services of two-thirds, a 1.1% boost in labor productivity implies flow GDP gains of $160 billion per year, or a present value gain of $4 trillion at a 4% discount rate. Our data also suggest that better home internet access increases the propensity to work from home. Universal access would, according to our estimate, raise the extent of WFH in the post-pandemic economy by about seven-tenths of a percentage point. When we account for this effect, it slightly raises our estimate for the earnings-weighted productivity benefits of moving to universal access.

To assess the distributional consequences of universal access, we adopt the benchmark assumption that earnings are proportional to productivity in the cross section. Under this assumption, the SWAA data let us estimate the impact of universal

2 This estimate is a projection of how the pandemic will affect future productivity through the pandemic-induced re-optimization of working arrangements. See BBD and Erdziek (2021) on how the pandemic drives a re-optimization of working arrangements. To be sure, the pandemic may have additional productivity consequences, positive and negative, through other channels.

access on the earnings distribution. Two basic effects are in play: On average, lower-income workers have home internet service of lower quality and lesser reliability. In isolation, this fact suggests that moving to universal access would reduce inequality. However, planned levels of WFH in the post-pandemic economy rise strongly with earnings in the cross section. This effect cuts the other way. On net, we find that universal access would be of little consequence for overall earnings inequality and for the distribution of average earnings across major demographic groups.

Internet access quality was more consequential during the pandemic, because WFH accounted for such a large share of labor services. For the period from May 2020 to April 2021, we estimate that subpar internet access degraded earnings-weighted productivity by 3%. As before, our counterfactual is universal access to high-quality, fully reliable internet service. For perspective, consider the size of the U.S. output shortfall during the pandemic. Real GDP per capita was about 11% below its pre-pandemic trend in the second quarter of 2020 and about 4.5% below trend in the third quarter.[3] These figures imply that universal access would have materially moderated the U.S. output shortfall during the pandemic. They also imply that the flow payoffs to universal access are larger in other disaster scenarios (*e.g.,* another pandemic) that inhibit travel and in-person interactions but do not cut off the internet itself. Partly for this reason, we see universal access as even more valuable to society than suggested by a simple capitalization of its expected flow benefits.

More broadly, our societal experience with remote work and virtual connectivity during the COVID-19 pandemic highlights the resilience value of the internet and other technologies that facilitate commerce and socializing at a distance. These technologies enabled large sectors of the economy to function well during the pandemic, even as people sheltered in place and socially distanced. They also facilitated a swift expansion of online commerce and delivery services, making it much easier for people to socially distance and curtail the spread of the virus. Similarly, universal access would promote economic resilience in the face of future pandemics and other disasters that inhibit travel and in-person interactions.

Universal access also promotes other forms of resilience. To see this point, start with the fact that loneliness and social isolation are harmful to mental and physical health. This observation leads naturally to the conjecture that social distancing during the pandemic had negative health effects for many Americans. It also raises the question of whether internet access alleviates the harmful health effects of social distancing. In this regard, we find that subjective well-being increases with internet access quality during the pandemic conditional on work status, working arrangements, and a battery of other controls. While we do not estimate causal

3 See Figure 6 in Davis, Liu, and Sheng (2021).

effects on well-being, our evidence suggests that home internet access mitigates the negative health effects of loneliness and social isolation in a time of pervasive social distancing, and that better access does so to a greater extent.

Before proceeding, we note some limitations of our data and analysis. First, our SWAA sample may underrepresent persons who lack home internet access, leading us to misstate the impact of universal access.[4] Second, we rely on worker assessments of productivity in gauging the consequences of universal access. Employer assessments may differ. Third, our projections neglect static general equilibrium effects. For example, suppose universal access encourages more WFH and thereby drives down the cost of office space in city centers. In turn, cheaper office space could moderate the induced shift to remote work. We think equilibrium effects of this sort are likely to be quite small, especially given the very modest size of our estimate for how universal access would affect the extent of WFH. Fourth, our projections ignore dynamic effects. These could flow from technological advances that promote WFH over time (Bloom, Davis, and Zhestkova 2021) or the longer-term consequences of universal access itself, which could prompt changes in job design that facilitate remote work.[5] These dynamic effects could be important, but they are also hard to project. Finally, we are silent about costs, which are surely relevant to judgements about the desirability of moving to universal access.

2. Working arrangements, productivity, and internet access in the cross section

2.a. The Survey of Working Arrangements and Attitudes

We have fielded the SWAA since May 2020, collecting 2,500 to 5,000 responses per month. Each survey wave contains 40 to 55 questions on demographics, employment status, working arrangements, earnings, commuting, internet access, expectations and experiences related to WFH, perceptions, and more. Our focused questions and large sample size give us an unparalleled window into the WFH phenomenon during the pandemic and let us make data-based projections for the post-pandemic U.S. economy.[6]

4 BBD show that the self-assessed productivity effects of WFH align reasonably well with less subjective measures based on commuting time savings. They also find that desired and planned levels of WFH in the post-pandemic economy rise strongly with the self-assessed, relative productivity of WFH. These results give us confidence that our productivity data are meaningful.

5 As one example, doctors and patients report that the use of video conferencing to discuss test results and conduct routine follow-up consultations can be more efficient and convenient than in-person visits. Better internet access facilitates this type of remote healthcare delivery (McCollough et al., 2021), which makes it more practical for healthcare professionals to work remotely.

6 See Barrero, Bloom, and Davis (2021a) for the full set of questions. Our data are available to other researchers at www.WFHresearch.com, and we continue to field the SWAA and update the website about once a month. We do not collect personally identifiable information, do not contact respondents directly, and have no way to recontact them.

The SWAA target population covers U.S. residents, 20 to 64 years old, who earned at least $20,000 in 2019. Given these parameters, QuestionPro and Inc-Query recruit respondents on our behalf from lists of verified persons supplied by leading market research aggregators, who gather potential respondents from multiple sources. One reason to tap multiple sources is that the form of respondent compensation depends on where and how they are recruited. Some respondents receive airline miles in exchange for survey participation, for instance, while others receive cash or credits that unlock internet game features. No respondents sign up specifically for our survey.[7]

The resulting distribution of SWAA respondents appears similar to that of working-age respondents in Current Population Survey (CPS) data from 2010 to 2019, except the SWAA features notably larger shares of high earners and persons with advanced degrees. Throughout this chapter, we reweight raw SWAA responses (after dropping speeders) to match the distribution of respondents in the 2010–2019 CPS over cells defined by the cross product of four age bins, sex, six education categories, and four earnings bins. The resulting marginal distributions by age, sex, education, earnings, major industry, and Census division in the reweighted SWAA data are very similar to the corresponding CPS distributions (Figure 2 in BBD).

Respondents can, and sometimes do, take our survey using mobile devices (that do not require home internet access) or by accessing the internet outside their homes. Still, our sample may be skewed away from persons who lack home internet access. Insofar as our sample is skewed in this manner, we may understate the impact of universal access for the simple reason that it would involve a bigger change for persons who currently lack access. Other sources of sample selection may affect some of our results, as we discuss below.

2.b. The extent of working from home

Table 1 summarizes the extent of WFH before, during, and after the COVID-19 pandemic. Our post-COVID projections rely on responses to the following SWAA question:[8]

7 Following best practice for surveys of this type, we drop persons who complete the survey in less than two minutes in May, less than three minutes in the July to November 2020 waves, and less than five minutes in later waves. Given the nature and number of our survey questions, these “speeders” are unlikely to supply careful responses. After dropping speeders, which cuts the sample about 20%, median completion time ranges from three minutes and ten seconds in May 2020 to 10 minutes and 55 seconds in December 2020.

8 Before the August 2020 wave, the question specified “After COVID in 2021…” instead of 2022.

After COVID, in 2022 and later, *how often* ***is your employer planning*** *for you to work full days at home?*

- *Never*
- *About once or twice per month*
- *1, 2, 3, 4, or 5+ days per week [separate options for each]*
- *My employer has not discussed this matter with me or announced a policy about it*
- *I have no employer*

In constructing our projections, we drop persons with no employer in the survey week. We assign zeros to "Never" and "About once or twice per month," 20% for one full day per week WFH, 40% for two full days, and so on. We also assign zeros to "My employer has not discussed this matter with me ..." on the view that employers are unlikely to raise the matter with workers in jobs for which WFH is impractical or infeasible. See BBD on how we estimate the extent of WFH before and during COVID and comparisons to results from other surveys.

Table 1: Working from home before, during, and after the COVID-19 pandemic

Percent of full paid days worked from home					
	Pre-COVID	During COVID		Post-COVID	
Equal-weighted	4.8	47.4	(0.3)	21.9	(0.3)
Earnings-weighted	--	54.4	(0.3)	27.7	(0.3)
Percent of full paid days worked from home					
Education		**During COVID**		**Post-COVID**	
Less than high school		24.6	(2.8)	12.8	(2.3)
High school		32.0	(0.7)	15.5	(0.6)
1 to 3 years of college		40.4	(0.6)	18.6	(0.5)
4-year college degree		57.9	(0.5)	27.2	(0.5)
Graduate degree		63.4	(0.4)	30.7	(0.5)

Notes: The pre-COVID estimate for the extent of WFH relies on data from the 2017–2018 American Time Use Survey, as described in Barrero, Bloom, and Davis (2021b). Estimates for "During COVID" rely on data from the May 2020 through May 2021 waves of the SWAA. Estimates for "Post-COVID" rely on worker responses to questions about employer plans in the six most recent waves of the SWAA, namely December 2020 to May 2021. We re-weight raw responses in the SWAA to match the share of working-age respondents in the 2010–2019 CPS in each {age x sex x education x earnings} cell. Standard errors in parentheses.

As reported in Table 1, we project that WFH will account for 21.9% of full paid workdays in the post-pandemic economy, 27.7% on an earnings-weighted basis. The higher earnings-weighted figure reflects the strongly positive cross-sectional relationship between the extent of WFH and worker earnings. WFH also rises strongly with educational attainment, as seen in the lower panel. In contemplating these figures, recall that our target population is persons 20–64 who earned at least $20,000 in 2019. Thus, we under sample low-wage and part-time workers, who tend to be concentrated in Food Services, Retail Trade, and other industries with lesser scope for WFH. For this reason, our results may overstate the equal-weighted WFH share in the post-pandemic economy. This feature of our sample matters little for earnings-weighted results.

2.c. The relative productivity of working from home

To assess the relative productivity of WFH, the SWAA puts the following question to all persons who report WFH at some point during the pandemic:

> *How does your efficiency working from home* ***during the COVID-19 pandemic*** *compare to your efficiency working on business premises* ***before the pandemic?***
>
> – *Better—I am more efficient at home than I was working on business premises*
> – *About the same—I'm equally efficient in both places*
> – *Worse—I am less efficient at home than I was working on business premises*

For those who respond "Better" ["Worse"], we follow up with:

> *How much* ***more [less] efficient*** *have you been working from home* ***during the COVID-19 pandemic*** *than on business premises* ***before the COVID-19 pandemic?***

Response options are: *Under 5% more [less] efficient; 5 to 10% more [less] efficient; 10 to 15% more [less] efficient; 15 to 25% more [less] efficient; 25 to 35% more [less] efficient; and Over 35% more [less] efficient.*

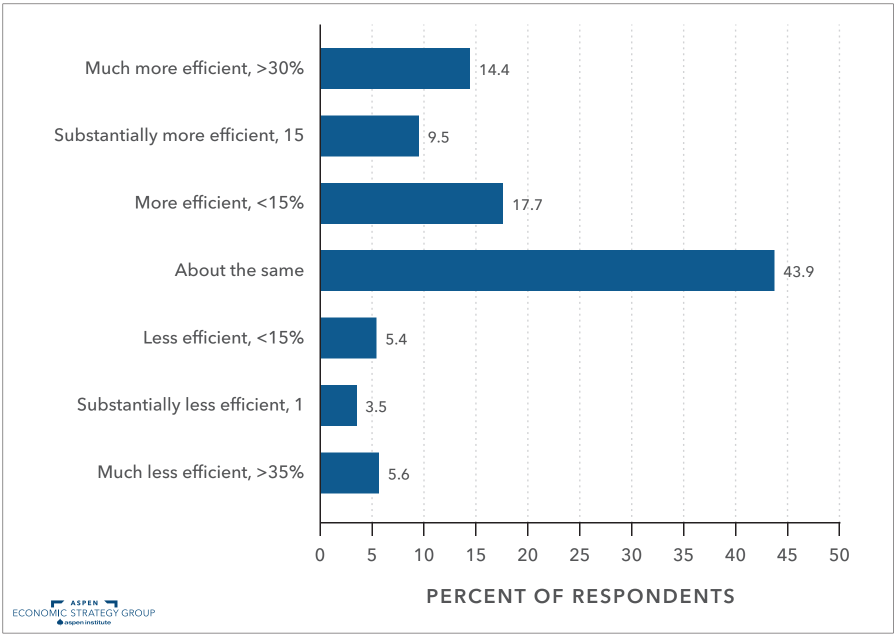
As seen in Figure 1, 44% of respondents say that WFH is about as productive as working on employer premises. The balance of the other 56% tilts toward greater productivity when WFH. That is, the average worker reports greater productivity when WFH. As shown in BBD, the planned extent of WFH in the post-pandemic economy rises strongly with the relative productivity of WFH. Thus, the productivity boost generated by a shift to WFH in the post-pandemic economy reflects a combination of higher productivity when WFH for many workers and the selected nature of who works from home in the post-pandemic economy.

Putting the various pieces of information together, BBD estimate that the post-pandemic shift to WFH will drive an earnings-weighted productivity gain of 4.6% relative to the situation with pre-pandemic working arrangements. This gain arises mainly from the savings in commuting time afforded by more WFH. Because they do not account for commuting time, conventional measures of productivity will show a smaller gain. Indeed, when BBD mimic conventional measures, they project that the re-optimization of working arrangements in the post-pandemic economy will boost measured productivity by only 1%.

2.d. The cross-sectional distribution of home internet access quality

Figure 2 displays the distribution of home internet access quality based on responses to the question, "How reliable is your internet connection?" About 41% of SWAA respondents say they have "perfect" home internet service that "works 100% of the time." Another 43% say their home internet service works "90% of the time," 12% say it works "70% to 80% of the time," 2% say less than "70% of the time," and 2.5% have no home internet connection.

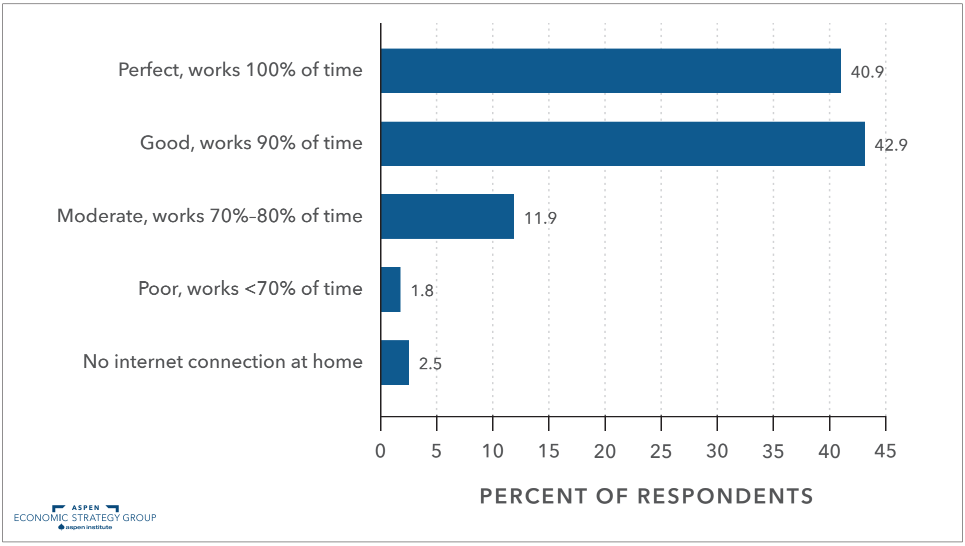
Figure 1: Efficiency of WFH vs. working on business premises



Source: Responses to the question, "How does your efficiency working from home **during the COVID-19 pandemic** compare to your efficiency working on business premises **before the pandemic?"**

Notes: Data are from 38,250 survey responses collected from August 2020 to May 2021 by Inc-Query and QuestionPro. We asked a similar question in earlier waves but focus on August 2020 to May 2021 when we kept the question and response options consistent. We re-weight raw responses to match the share of working age respondents in the 2010–2019 CPS in each {age x sex x education x earnings} cell.

Figure 2: Distribution of internet quality among SWAA respondents



Source: Respondents to the question, "How reliable is your internet connection?" in the Survey of Working Arrangements and Attitudes.

Notes: Data are from 43,250 survey responses collected from May 2020 to May 2021 by Inc-Query and QuestionPro. We re-weight raw responses to match the share of working age respondents in the 2010-2019 CPS in each {age x sex x education x earnings} cell.

Table 2 provides information about how the average quality of internet access varies by demographics and other respondent characteristics. The overall average access quality is 88.9%, meaning that home internet service works about 89% of the time for the average person with WFH experience during the pandemic. Average access quality is similar for men and women and somewhat smaller for persons 50–64 years of age. Average access quality rises with education and with earnings in 2019. When we further weight responses by number of children in the household, overall average access quality is somewhat higher at 90.8%. Here, it's worth keeping in mind that households with no working parent, or no parent who earned at least $20,000 in 2019, are not in-scope for our sample. Appendix Table A.1 provides more information about how internet access quality varies with observables in the SWAA.

Table 2: Average internet access quality by group

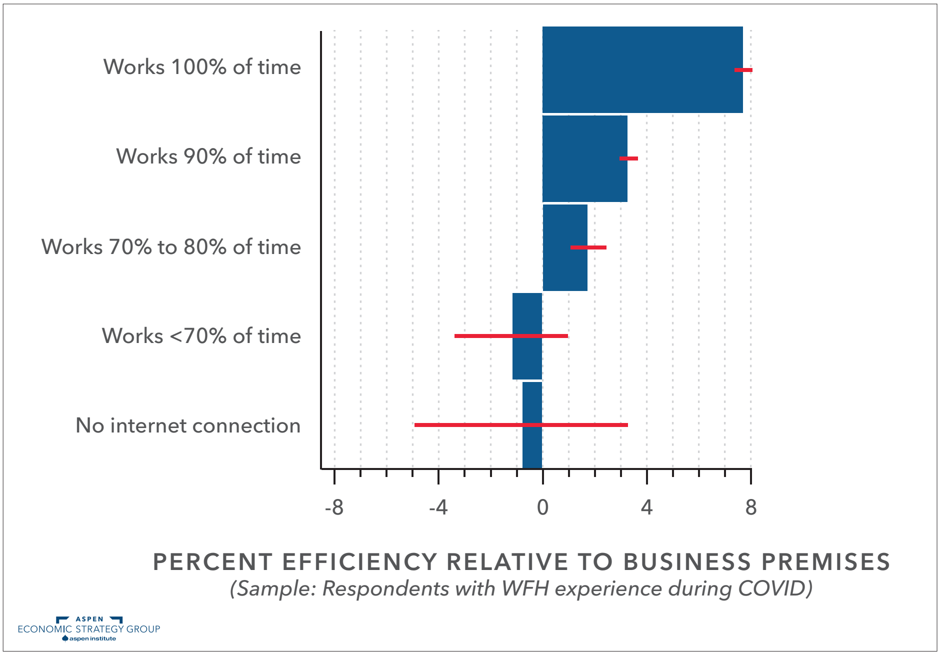
Percent of the time the internet works	Equally-weighted mean (SE)		Mean weighted by number of children (18 & under) in the household (SE)	
Overall	88.9	(0.1)	90.8	(0.1)
Women	88.7	(0.1)	89.9	(0.2)
Men	89.1	(0.1)	91.4	(0.1)
Age 20 to 29	89.0	(0.2)	89.6	(0.3)
Age 30 to 39	90.3	(0.2)	91.7	(0.2)
Age 40 to 49	89.3	(0.2)	90.9	(0.2)
Age 50 to 64	87.6	(0.2)	89.4	(0.4)
Red (Republican-leaning) State	88.6	(0.1)	89.3	(0.1)
Blue (Democrat-leaning) State	89.3	(0.1)	90.9	(0.1)
Less than high school	85.3	(1.3)	89.2	(1.6)
High school	86.8	(0.3)	89.3	(0.3)
1 to 3 years of college	88.6	(0.2)	89.7	(0.3)
4-year college degree	90.4	(0.1)	91.6	(0.2)
Graduate degree	91.6	(0.1)	93.3	(0.1)
Ann. Earnings of $20 to $50K	87.9	(0.2)	89.6	(0.3)
Ann. Earnings of $50 to $100K	89.9	(0.2)	91.4	(0.2)
Ann. Earnings of $100 to $150K	91.1	(0.2)	92.8	(0.2)
Ann. Earnings over $150K	92.0	(0.2)	94.0	(0.2)

Notes: Percent of the time that the internet works, based on responses to the question, "How reliable is your internet connection?" Data are from over 40,000 survey responses collected between May 2020 and May 2021 by Inc-Query and QuestionPro. We reweight raw responses to match the share of working age respondents in the 2010–2019 CPS in a given {age x sex x education x earnings} cell. The second and fourth columns additionally weight by the number of children present in the household.

2.e. How WFH productivity relates to home internet access quality

Figure 3 summarizes how the relative productivity of WFH relates to internet access quality in the cross section. Persons with internet access that works all the time report an average productivity difference of nearly 8% in favor of WFH. At the other end of the scale, persons with internet access that works less than 70% of the time and those who lack home internet access report average productivity differences of about 1% in favor of employer premises. Figure 3, in conjunction with Figure 2, clearly points to the potential for universal access to raise productivity for persons who work from home. Since WFH is projected to account for more than one-fourth of all earnings-weighted workdays in the post-pandemic economy, Figures 2 and 3 also imply that universal access would raise overall productivity in the economy.

Figure 3: Self-assessed efficiency while WFH by reported internet quality



Source: Responses to the following questions in the Survey of Working Arrangements and Attitudes:

"How reliable is your internet connection?"

"How does your efficiency working from home **during the COVID-19 pandemic** compare to your efficiency working on business premises **before the pandemic?**"

"How much more [less] efficient have you been working from home during the COVID-19 pandemic than on business premises before the COVID-19 pandemic?"

Notes: The red lines show 95 percent confidence intervals. Data are from 38,250 survey responses collected from August 2020 to May 2021 by Inc-Query and QuestionPro.. We re-weight raw responses to match the share of working age respondents in the 2010–2019 CPS in each {age x sex x education x earnings} cell.

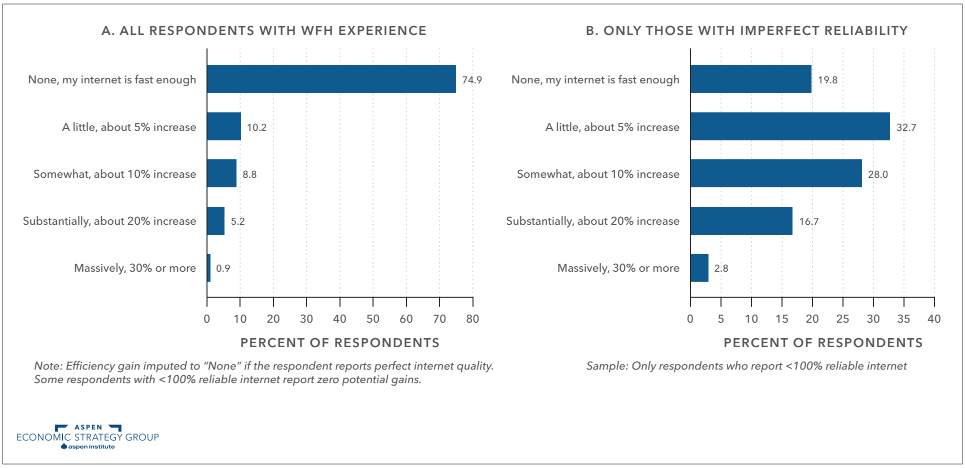
3. Projecting the effects of universal access

3.a. Direct productivity effects

We now project the effects of a hypothetical move to universal access, defined as a shift from the current access situation (summarized in Figure 2 and Table 2) to one with high-quality, fully reliable internet access in all households. In forming our productivity projections, we combine individual-level data on the planned extent of WFH in the post-pandemic economy with individual-level estimates for the productivity impact of universal access.

We estimate the individual-level productivity impacts using two distinct approaches. The first exploits responses to the following question: "How much would your efficiency working from home increase if you had perfect high-speed internet?" Responses to this question, summarized in Figure 4, elicit self-assessed causal effects of the hypothetical in question. Accordingly, we interpret the suitably aggregated responses as yielding estimated causal effects, subject to the qualifications sketched above and discussed more fully below. Survey responses to this question for persons with WFH experience during the pandemic say that universal access raises the earnings-weighted average efficiency of time spent WFH by 3.3%. Appendix Table A.2 provides information on how the self-assessed efficiency gains vary with observables.

Figure 4: "How much would your efficiency working from home increase if you had perfect high-speed internet?"



Source: Responses to the following questions in the Survey of Working Arrangements and Attitudes:

"How reliable is your internet connection?"

"How much would your efficiency working from home increase if you had perfect high-speed internet?"

Our second approach relies on regression models that relate the productivity of WFH to internet access quality in the cross section. The dependent variable is WFH efficiency during the pandemic relative to that of working on employer premises before the pandemic. Appendix Table A.3 reports our estimated regression models. Our preferred model controls for earnings, education, gender, presence of children, state of residence, and industry of the worker's current or last job. Using this model, and interpreting the coefficient on internet access quality as a causal effect, we calculate the individual-level productivity change associated with universal access.

Table 3 draws together the pieces of our analysis to report two sets of results: Estimates for the aggregate labor productivity shortfall caused by subpar internet access during the pandemic in column (1), and estimates for the aggregate labor productivity gains caused by universal access in the post-pandemic economy. The regression approach implies a productivity shortfall due to subpar internet service of 0.7% during the pandemic and a productivity gain from universal access of 0.3% in the post-pandemic economy. The approach based on self-assessed causal effects yields larger productivity consequences: a 3% shortfall during the pandemic and a 1.1% gain from universal access in the post-pandemic economy.

Table 3: Earnings-weighted productivity effects of internet access quality

	(1)	(2)	(3)
	Productivity Shortfall During COVID	Post-COVID Gains from Universal Access	
Approach to Estimating the Individual-Level Productivity Effects of Universal Access	**Productivity losses due to subpar internet access**	**Using employer plans for WFH**	**And adjusted for WFH rise induced by universal access**
Regression-imputed (simple)	-0.8	0.4	0.4
Regression-imputed (controls)	-0.7	0.3	0.4
Self-assessed causal effect	-3.0	1.1	1.1

Notes: Column (1) reports the estimated aggregate productivity shortfall during the pandemic due to subpar internet access quality by many Americans who worked from home. See Figure 2 for the distribution of subpar internet service. Column (2) reports the estimated earnings-weighted productivity gains of universal access to high-quality, fully reliable home internet service in the post-pandemic economy when using employer plans for who works from home, and how much. Column (3) also adjusts for the post-pandemic rise in WFH that we estimate, and which we report in Table 4 below. See Figure 4 for the distribution of self-assessed causal effects from gaining access to high-quality, fully reliable internet service. See Table A3 for regression models that relate WFH productivity to internet access quality. We use column (1) in Table A3 for the row titled, "Regression-imputed (simple)"; and we use column (7) in Table A3 for the row, titled "Regression-imputed (controls)". Data are from 43,250 survey responses collected from May 2020 to May 2021 by Inc-Query and QuestionPro. We re-weight individual-level data to match the share of working age respondents in the 2010–2019 CPS in each {age x sex x education x earnings} cell.

The regression approach might seem more familiar than the approach that relies on the self-assessed efficiency effects of better internet access. However, we see the self-assessment approach as superior, precisely because it relies on survey questions that seek to elicit a causal effect. In contrast, our regression-based approach to quantifying the causal productivity effects of better internet service relies on strong assumptions that might not hold.[9] Accordingly, we focus on results that rely on self-assessment effects in the rest of the chapter.

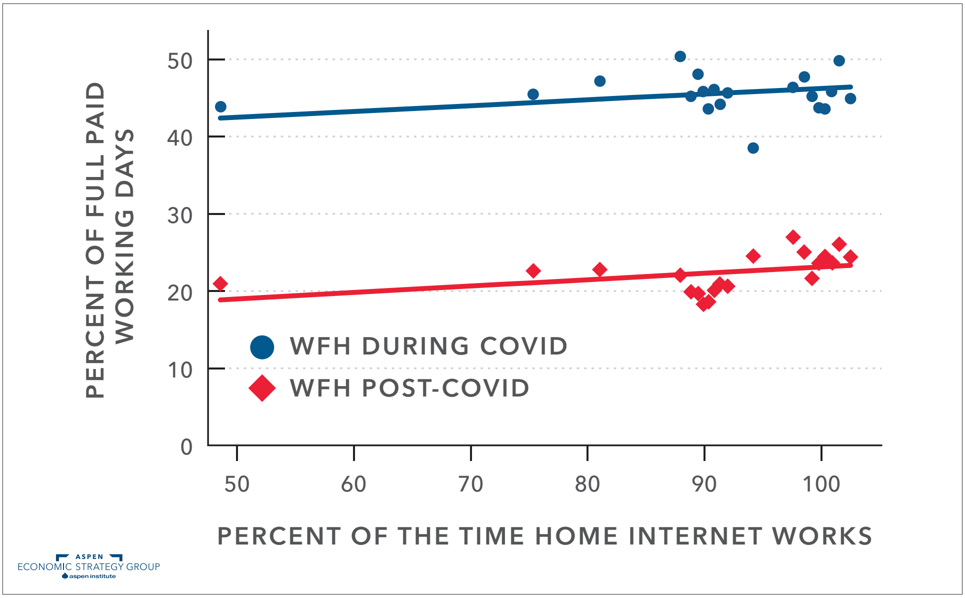
3.b. The response of WFH to universal access and knock-on productivity effects

Thus far, we have ignored any effect of home internet access quality (and availability) on the extent of WFH. If better access improves WFH efficiency, as our foregoing evidence strongly indicates, we expect universal access to increase the extent of WFH. The question is how much.

Figure 5 confirms that the extent of WFH rises with internet access quality conditional on the worker's earnings and industry of employment, which we interpret here as crude controls for the nature of the worker's job. The upper regression line and blue dots reflect data on reported levels of WFH during the pandemic as of the survey week. The lower regression line and red dots reflect employer plans for WFH in the post-pandemic economy, as reported by the worker. In both cases, a 10 percentage-point increase in the working availability of home internet access brings a 0.8 percentage-point increase in the extent of WFH. This effect is statistically significant but modest in size.

9 The slope coefficient on internet access quality has a causal interpretation only under strong assumptions, including the assumption that the variation in internet access quality is conditionally uncorrelated with omitted determinants of WFH efficiency.

Figure 5: How the incidence of WFH relates to internet access quality conditional on industry and earnings



Notes: Coef during COVID = 0.08 (0.03). Coef post-COVID = 0.08 (0.02) N = 24890. Controls for industry, survey wave FE, and 2019 earnings. 7/2020 and later survey waves.

Table 4 reports the estimated impact of universal access on the extent of WFH when we interpret the slope coefficients in these regressions as causal effects. The overall estimated impact on the extent of WFH—an increase of 0.7 percentage points—is quite modest both during and after COVID. The impact also varies little across demographic groups. Hence, when we account for the impact of universal access on the extent of WFH in the post-pandemic economy, our estimates for the aggregate labor productivity effects of universal access barely budge. This point can be seen by comparing the results in columns (2) and (3) of Table 3.

Table 4: Extra WFH induced by universal access during and after the COVID-19 pandemic

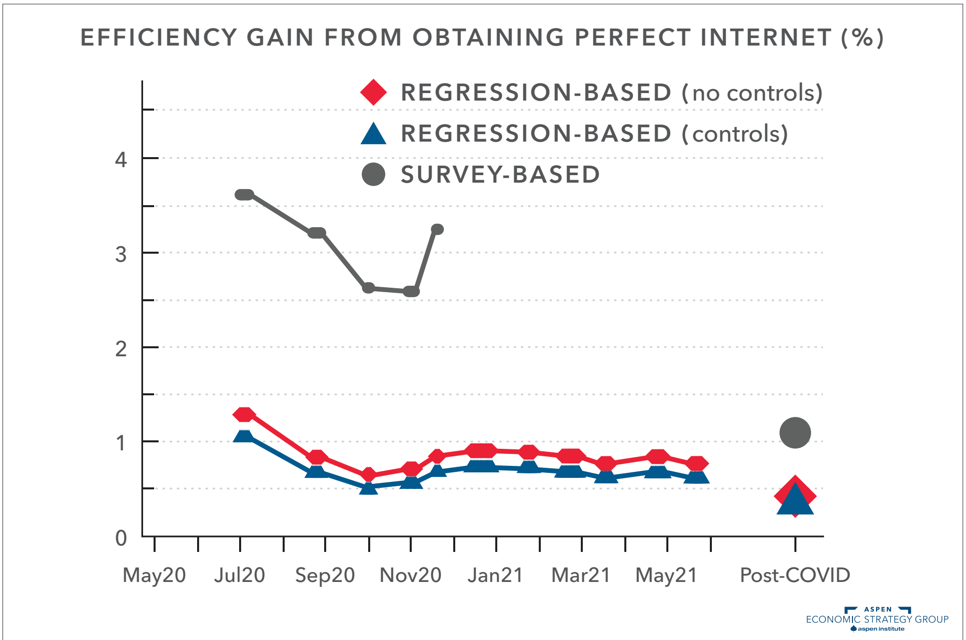
Increase in percent of full paid days WFH induced by universal access to perfect high-speed internet	During COVID	After COVID
Overall	**0.7**	**0.7**
Women	0.7	0.8
Men	0.6	0.7
Age 20 to 29	0.7	0.8
Age 30 to 39	0.6	0.7
Age 40 to 49	0.7	0.7
Age 50 to 64	0.7	0.8
Ann. Earnings of $20 to $50K	0.8	0.8
Ann. Earnings of $50 to $100K	0.6	0.7
Ann. Earnings of $100 to $150K	0.6	0.6
Ann. Earnings over $150K	0.4	0.5
Less than high school	0.8	0.8
High school	0.8	0.9
1 to 3 years of college	0.7	0.8
4-year college degree	0.6	0.7
Graduate degree	0.6	0.6

Notes: Estimated percent of full paid WFH days are 45.4 during COVID and 22.2 post-COVID. Average increase in percent of WFH days are based on a regression and imputation method that estimates the relationship between WFH days and internet quality with controls for 2019 earnings and industry of the current or most recent job. We multiply the slope coefficient by the increase in internet reliability that takes each respondent to 100% reliable internet. The sample includes respondents who are working during COVID. Data are from over 40,000 survey responses collected between May 2020 and May 2021 by Inc-Query and QuestionPro. We reweight raw responses to match the share of working age respondents in the 2010–2019 CPS in a given {age x sex x education x earnings} cell.

3.c. Assessing the estimated productivity and output effects

Figure 6 summarizes our results on the aggregate labor productivity effects of universal access. The "Post-COVID" data points are from column (2) in Table 3. The other data points trace out our estimates for the productivity shortfall caused by subpar internet access during the pandemic. The magnitude of the shortfall ranges from 2.6 to 3.6% for the self-assessment approach, fluctuating over time with the extent (and cross-sectional distribution of) WFH during the pandemic. The smaller 1.1% gain that we estimate for the causal effect of universal access in the post-COVID economy reflects a lower incidence of WFH.

Figure 6: Earnings-weighted productivity gains from universal access to high-quality, reliable home internet service



Notes: Adjusts for each respondent's amount of WFH during COVID and employer plans post-COVID. Data are from over 40,000 survey responses collected from July 2020 to May 2021 by Inc-Query and QuestionPro. We re-weight raw responses to match the share of working age respondents in the 2010– 2019 CPS in each {age x sex x education x earnings} cell.

As remarked in the introduction, we can use these labor productivity estimates and a standard aggregate production function to derive implications for aggregate output. In particular, we work with a production function that exhibits constant returns to scale and a value of two-thirds for the elasticity of output with respect to labor services. Given this production function, and holding fixed the values of non-labor inputs, a 1.1% improvement in labor productivity implies an aggregate output gain of 0.73% (= $(1.011)^{2/3}$ — 1) per period in the post-pandemic economy.[10] Similarly, a 3% labor productivity shortfall due to subpar internet access implies an aggregate output loss of 2% during the pandemic. In other words, the flow output loss during the pandemic is nearly three times as large as the projected flow benefits from universal access in the post-COVID economy. This comparison underscores the economic resilience value of universal access: the output payoff is much larger in pandemic-like disaster states when output is unusually low and the marginal value of output is unusually high.

Our estimates for the impact of universal access on productivity and output could be biased for various reasons. As explained above, we may under sample persons who lack home internet access. Moving to universal access would be a very large change for persons who currently lack access. Thus, if our sample is selected against those who currently lack access, it may downwardly bias our estimated productivity effects. Perhaps, however, under-represented persons have smaller productivity responses to better internet service because their jobs offer less scope for WFH. This effect cuts the other way. In any event, persons with no internet access tend to have low earnings and productivity. Thus, a given percentage change in their productivities would have relatively modest effects on earnings-weighted mean productivity. In light of these observations, we think sample selection against persons who lack home internet access is a small concern in our analysis of aggregate productivity and output effects.

Two other sources of bias strike us as potentially more important. First, insofar as pandemic-related stresses, the presence of kids at home due to school closures, a lack of familiarity with remote work technologies, and other forces pull down WFH productivity *during* the pandemic, our regression models understate the likely strength of any relationship between the relative efficiency of WFH to internet access quality *after* the pandemic. To a lesser extent, this point also applies to our estimated individual-level productivity effects under the self-assessment approach. For example, an initial lack of familiarity with remote work technologies may pull down the self-assessed impact of better internet service on the relative efficiency of WFH. Such transitory negative effects on WFH productivity during the pandemic lead us to understate WFH efficiency in the post-pandemic economy and, hence, to understate the productivity and output benefits of universal access in a post-pandemic setting.

10 See Kessler et al. (2021) for an assessment of how better internet access would affect output and economic development in one Tennessee county. Zuo (2021) provides evidence that subsidizing internet access raises employment and earnings among low-income Americans.

Second, we have no data on the relative efficiency of WFH for respondents with no WFH experience during the pandemic (as of the survey date). That's 43.3% of respondents on an equal-weighted basis and an estimated 34.2% on an earnings-weighted basis.[11] Thus, the Table 3 and Figure 6 estimates rest on an implicit assumption that persons with no WFH experience during the pandemic have the same average productivity responsiveness to better internet service as persons in the analysis sample. If excluded persons disproportionately hold jobs that are poorly suited for WFH, which seems likely, their exclusion leads us to overstate the effects of universal access on productivity and output.

A few additional observations are helpful in thinking about the potential effects of universal access on future productivity and output. First, we expect the structure of the economy to continue evolving in ways that expand opportunities for remote work. Examples include greater remote service delivery by health care professionals, social workers, educators, and customer-service staff in government agencies, all of which shifted to greater remote work in reaction to the pandemic. Even activities as seemingly unsuitable as operating oil and gas wells are seeing a shift to remote workers (Jiao and Tovar 2020). Looking across countries, Hatayama et al. (2020) find a strong positive relationship between GDP per capita and the extent to which jobs in the country are amenable to WFH. In light of these observations, it seems likely that the flow productivity and output benefits of universal access will rise over time.

Advances in complementary technologies is another reason to expect the flow benefits of universal access to rise over time. In this regard, Bloom, Davis, and Zhestkova (2021) find that the pandemic drove a rise in the share of new U.S. patent applications that advance technologies in support of video conferencing, telecommuting, and remote interactivity. This finding suggests that a redirection of technical change in reaction to COVID-19 and a persistent shift to WFH will raise the quality and efficiency of remote work in the future. Insofar as complementary technologies improve, universal access is likely to have larger payoffs.

3.d. Earnings inequality

Since the foregoing productivity analyses are built up from micro data, we can easily consider the implications of universal access for the distribution of individual-level productivities. If, in addition, we assume that individual earnings are proportional to productivities in the cross section, we can estimate the consequences of universal access for the distribution of earnings. Clearly, the proportionality assumption is only an approximation, but we regard it as a useful and transparent one.

11 To derive the earnings-weighted estimate, we assign respondents to the midpoints of their 2019 earnings bins (or $1 million for the top bin of $500,000 or more).

Applying this assumption, we estimate that universal access would raise earnings by 1% for persons who earned $20,000–50,000 in 2019, 1.2% for those who earned $50,000–100,000, 1.3% for those who earned $100,000–150,000, and 1.1% for those who earned more than $150,000 in 2019. These estimates reflect employer plans for who will work from home in the post-pandemic economy, how much, and self-assessed productivity effects of better internet service. The proportional earnings gains are smaller at the bottom end, because low-wage jobs offer little scope for WFH. In short, moving to universal access would not materially affect earnings inequality according to our analysis.

Table 5 reports estimated productivity effects of universal access for demographic and other groups. Here as well, our projections imply that universal access would have modest effects on the (log) earnings distribution. Using the self-assessment approach, we find the smallest estimated effect of universal access for persons who did not finish high school (0.3%) and the largest for persons with a four-year college degree (1.4%).

Table 5: Efficiency gains from universal access to high-quality internet by group

Efficiency gain from perfect high-speed internet post-COVID (adjusted for the amount of WFH post-COVID)	Self-assessed
Overall	**1.1**
Women	1.1
Men	1.0
Age 20 to 29	1.2
Age 30 to 39	1.3
Age 40 to 49	1.1
Age 50 to 64	0.7
Less than high school	0.3
High school	0.7
1 to 3 years of college	1.1
4-year college degree	1.4
Graduate degree	1.0
Ann. Earnings of $20 to $50K	1.0
Ann. Earnings of $50 to $100K	1.2
Ann. Earnings of $100 to $150K	1.3
Ann. Earnings over $150K	1.1
Goods-producing sectors	0.8
Service sectors	1.1
No children	0.9
Living with children under 18	1.2
Red (Republican-leaning) State	1.0
Blue (Democrat-leaning) State	1.1

Notes: Average WFH efficiency gain post-COVID from universal access to high-quality internet, based on responses to "How much would your efficiency working from home increase if you had perfect high-speed internet?" The sample includes respondents who responded to the self-assessment question and the question about how much their employer is planning for them to work from home, except those who said they have no employer. For each respondent we multiply the potential efficiency gain from perfect internet by the fraction of working days their employer is planning for them to be WFH post-COVID. Data are from over 40,000 survey responses collected between May 2020 and May 2021. We reweight raw responses to match the share of working age respondents in the 2010–2019 CPS in a given {age x sex x education x earnings} cell.

4. Internet access and subjective well-being during the pandemic

That loneliness is negatively associated with physical and mental health is well documented in the psychology literature. As remarked in the opening paragraph of a highly cited article by Holt-Lunstad et al. (2015), "Being socially connected is not only influential for psychological and emotional well-being but it also has a significant and positive influence on physical well-being and overall longevity." Similarly, a highly cited article by Thoits (2011) opens by remarking, "Substantial evidence has accumulated over the past few decades showing that social ties and social support are positively and causally related to mental health, physical health, and longevity…. Evidence also documents that social support buffers the harmful physical and mental health impacts of stress exposure."

This body of evidence suggest that social distancing during the pandemic and pandemic-related stresses had negative health effects for many Americans.[12] It also motivates the hypothesis that better internet access during the pandemic alleviated the harmful psychological and other health effects of social distancing and pandemic-related stresses. Consistent with this hypothesis, Wallinheimo and Evans (2021) find higher life quality and lower depression scores for middle-age and older Americans who used the internet more often during June and July 2020. These positive associations were concentrated among people who used the internet mainly for communication, while those who used it for government or health-related searches experienced more depression symptoms. Varma et al. (2021) find that younger people were particularly vulnerable to stress, anxiety, and depression during the pandemic. Suicides and internet queries about suicide fell during the pandemic, contrary to concerns when lockdowns were first implemented. See, for example, Ahmad et al. (2021), Ayers et al. (2021), and Sinyor et al. (2020). Ability to connect over the internet may be one reason why suicides did not rise during the pandemic.

Other studies point to a broader potential for internet use and social media to be sources of harmful effects on well-being. See, for example, Alcott et al. (2021), Servidio et al. (2021), and Elhai et al. (2020). We do not aim to assess the overall effects of internet usage on well-being. Our much more limited objective is to provide evidence as to whether better internet access is associated with positive effects on well-being during a period with sharply restricted in-person interactions. The effects of better internet access during normal periods may well be different. Also, in contrast to most other studies, we examine the relationship of well-being to internet *access quality*

12 Indeed, U.S. drug-overdose deaths rose nearly 30% in 2020 to a record high level (McKay 2021). The rise was especially sharp from March 2020, which coincides with the onset of the pandemic in the United States, lockdowns, and social isolation.

rather than internet *usage or usage patterns*. Access quality is arguably more exogenous with respect to well-being than usage intensity or usage patterns.

To quantify subjective well-being among SWAA respondents, we ask the following question: "Please imagine a ladder with steps numbered from zero at the bottom to ten at the top. The top of the ladder represents the best possible life for you and the bottom of the ladder represents the worst possible life for you. If the top step is ten and the bottom step is zero, on which step of the ladder do you feel you personally stand at the present time?" We multiply the responses by ten to put them on a scale that runs from zero to 100.

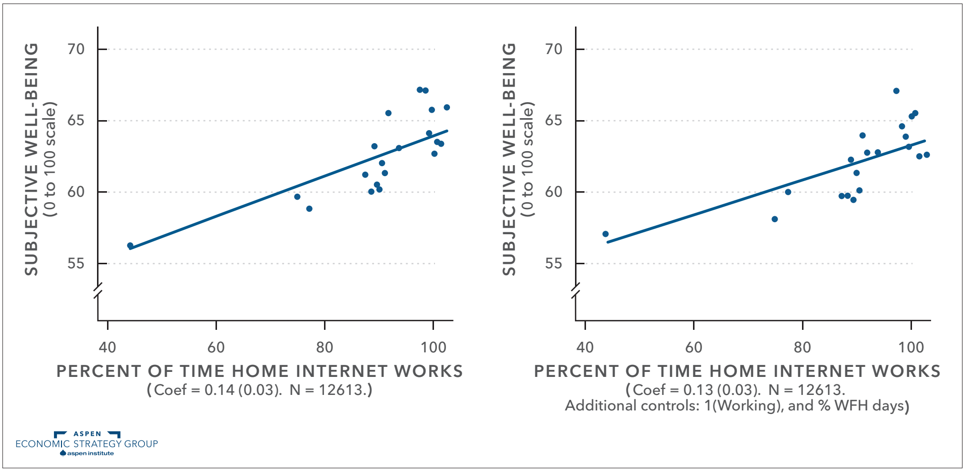
Figure 7 presents binned scatter plots of subjective well-being against the percent of time that the respondent's home internet service works. The left plot controls for gender, years of education, our four age bins, the log of 2019 earnings, whether the respondent lives with other adults, whether he or she lives with children, and whether the respondent lives with a partner. The slope coefficient of 0.14 says that a 10 percentage-point increase in internet availability is associated with a 1.4 percentage-point higher value of well-being. The right plot, which adds controls for employment status and working arrangements, yields a very similar relationship.

Appendix Table A.4 considers more flexible statistical models and extra controls. One additional result is that employed persons enjoy substantially higher well-being.[13] Controlling for other factors, the improvement in subjective well-being associated with working is more than one-third as large as the standard deviation of well-being in the sample.[14] For those who work, well-being is higher for persons who work from home a larger percentage of the week. Perhaps surprisingly, we find little evidence that the association of well-being with internet access quality is weaker for persons living with a partner or other adult.

13 In line with a large body of evidence that job loss and unemployment bring sizable declines in subjective well-being (Frey and Stutzer 2002). To see this point, recall that the SWAA sample is limited to persons who earned at least $20,000 in 2019. Thus, SWAA respondents who are jobless in the survey week recently had jobs but became unemployed or left the labor force.

14 See the coefficient of 8.2 in Column (6) of Table A.4.

Figure 7: How subjective well-being relates to internet access quality during the pandemic



Notes: Binned scatter plots of subjective well-being against internet quality. Both specifications control for gender, years of education, log(2019 earnings), age bin FE, whether living with other adults, whether living with children, and whether living with a partner. Subjective wellbeing is 10 times the response to the following question:

> "Please imagine a ladder with steps numbered from zero at the bottom to ten at the top. The top of the ladder represents the best possible life for you and the bottom of the ladder represents the worst possible life for you. If the top step is 10 and the bottom step is 0, on which step of the ladder do you feel you personally stand at the present time?"

Data are from over 40,000 survey responses collected between May 2020 and May 2021 by Inc-Query and QuestionPro. We reweight raw responses to match the share of working age respondents in the 2010–2019 CPS in a given {age x sex x education x earnings} cell. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Returning to the link between well-being and access quality, our preferred model says that moving from no home internet service to 100% reliable service is associated with a 15 percentage-point improvement in subjective well-being.[15] That's two-thirds as large as the standard deviation of well-being in the sample. Interpreted causally, this result says that universal access would materially improve well-being during pandemic-like disasters for persons who currently lack home internet service. Smaller improvements in well-being would accrue to persons who currently have subpar access.

5. Universal access as a source of economic and social resilience

By raising output in the face of infectious disease outbreaks, biological attacks, and other disaster states that involve physical social distancing, universal access to high-

15 Using column (6) in Table A.4.

quality home internet service would strengthen U.S. economic resilience. For society as a whole and for individual firms and workers, the capacity to quickly switch between production modes of roughly equal productivity is a valuable option that pays off especially in bad states of the world. Firm-level examples include contamination events, flood damage, explosions, and destructive fires that temporarily sideline the employer premises as a place of work. At the macroeconomic level, our analysis says that the output payoff to universal access during pandemic-like disasters is nearly three times as large as the payoff during normal periods. Our evidence also suggests that universal access promotes resilience by providing a ready means of engagement and socializing when circumstances compel physical distancing.

Universal access has other important benefits that we do not quantify, and that are likely to be especially valuable during pandemic-like disasters. For example, better internet access improves the ability of households to turn to online shopping and home delivery services during a pandemic-like disaster. As another example, Chiou and Tucker (2020) find that compliance with stay-at-home orders during the COVID-19 pandemic rose with access to high-speed internet service, even after controlling for household income. As a third example, better internet access promotes student engagement in remote-learning settings. Obviously, the value of remote learning is greater when a pandemic or other disaster leads to school closures. Using data from the early stages of the COVID-19 pandemic, Sen and Tucker (2020) find that areas with lower test scores, more poor children, and more non-White children faced greater problems with internet access. Thus, universal access may ameliorate the gap in learning opportunities between children from more and less advantaged families.

In sum, high-quality home internet access and complementary technologies enhance economic and social resilience in the face of pandemics and certain other disasters that inhibit travel and in-person interactions. In this regard, we note that there are sound reasons to fear that the SARS-CoV-2 virus "will ping pong back and forth across the globe for years to come," triggering recurrent outbreaks of COVID-19 (Brilliant et al. 2021). If that somber possibility materializes, the value of high-quality home internet access will be even greater than our analysis suggests. That said, we recognize that internet access is not a general-purpose source of resilience in the face of all disasters. For example, extended electricity outages over a large area would prevent most people in the area from accessing the internet to work, socialize, or study remotely. Cyberattacks that disable the electrical power grid or the internet itself would be hugely disruptive in any event, and possibly more disruptive insofar as the economy is highly adapted to remote work. As this remark suggests, widespread reliance on the internet and remote work can intensify other vulnerabilities.

6. Conclusion

The COVID-19 pandemic triggered a huge shift to working from home, and much of that shift will endure. Using our forward-looking survey data, we project that more than one-quarter of earnings-weighted workdays will be supplied from home after the pandemic ends.

Motivated by these developments, we examine data on how internet access quality affects productivity when working from home. According to our analysis, moving to high-quality, fully reliable home internet service for all Americans would raise earnings-weighted labor productivity by an estimated 1.1% in the coming years. The implied output gains are $160 billion per year, or $4 trillion when capitalized at a 4% rate. Estimated flow output payoffs to universal access are nearly three times as large in COVID-like disaster states. Better home internet service during the pandemic is also associated with greater subjective well-being, conditional on employment status, working arrangements, and a battery of other controls. The extra economic and social benefits of universal access during the pandemic underscore its resilience value in the face of disasters that inhibit travel and in-person interactions.

We express our main quantitative results as the benefits of moving to universal access, but the underlying empirical analysis rests on linear models and relationships. Thus, closing half the gap between universal access and the current household distribution of internet access quality has, according to our analysis, productivity and output effects that are half as large. This feature of our analysis simplifies a comparison of the benefits to the costs of better home internet access. There is an obvious need to quantify these costs to inform judgments about the wisdom of moving part or all the way to universal access. We hope that our work encourages a study of the cost side as well as further examinations of the benefits. We also hope to encourage additional research into sources of economic and social resilience in the face of disasters, which we see as an important but understudied topic.

References

Ahmad, Farida B. and Robert N. Anderson. 2021. “The Leading Causes of Death in the US for 2020.” *Journal of the American Medical Association* 325 (18): 1829–1830.

Alcott, Hunt, Matthew Gentzkow, and Lena Song. 2021. “Digital Addiction.” National Bureau of Economic Research Working Paper #28936.

Altig, David, Jose Maria Barrero, Nick Bloom, Steven J. Davis, Brent Meyer, Emil Mihaylov and Nick Parker. 2020a. “Firms Expect Working from Home to Triple.” *Macroblog*, 28 May.

———. 2021. “WFH Is Onstage and Here to Stay,” *Macroblog*, 25 February.

Altig, David, Jose Maria Barrero, Nick Bloom, Steven J. Davis, Brent Meyer and Nick Parker, 2020b. “Surveying Business Uncertainty.” Forthcoming, *Journal of Econometrics*.

Angelici, Marta and Paola Profeta. 2020. “Smart working: Flexibility without constraints.” CES Working Paper #8165.

Ayers John W., Adam Poliak, Derek C. Johnson, Eric C. Leas, Mark Dredze, Theodore Caputi, and Alicia L. Nobles. 2021. “Suicide-Related Internet Searches During the Early Stages of the COVID-19 Pandemic in the US.” *JAMA Network Open,* 4 (1).

Barrero, Jose Maria, Nicholas Bloom, and Steven J. Davis. 2020a. “60 Million Fewer Commuting Hours per Day: How Americans Use Time Saved by Working from Home.” *VoxEU CEPR Policy Portal,* 23 September.

———. 2020b. “COVID-19 Is Also a Reallocation Shock.” *Brookings Papers on Economic Activity*, Summer, 329–371.

———. 2021a. “Question Repository for the Survey of Working Arrangements and Attitudes.” available at www.WFHresearch.com.

———. 2021b. “Why Working from Home Will Stick.” National Bureau of Economic Research Working Paper #28731.

Bartik, Alexander W., Zoe B. Cullen, Edward L. Glaeser, Michael Luca, and Christopher T. Stanton, 2020. “What Jobs are Being Done at Home During the Covid-19 Crisis? Evidence from Firm-Level Surveys.” National Bureau of Economic Research Working Paper #27422.

Behrens, Kristian, Sergey Kichko, Jackques-Francois Thisse. 2021. “Working from Home: Too Much of a Good Thing?” CESifo Working Paper #8831.

Bick, Alexander, Adam Blandin, and Karel Mertens. 2020. “Work from Home After the COVID-19 Outbreak,” Federal Reserve Bank of Dallas Working Paper.

Bloom, Nicholas, Steven J. Davis, and Yulia Zhestkova. 2021. "COVID-19 Shifted Patent Applications toward Technologies that Support Working from Home." *American Economic Association, Papers & Proceedings,* May.

Bloom, Nicholas, James Liang, John Roberts, and Zhichun Jenny Ying. 2015. "Does working from home work? Evidence from a Chinese experiment," *Quarterly Journal of Economics,* 130 (1): 165–218.

Brilliant, Larry, Lisa Danzig, Karen Oppenheimer, Agastya Mondal, Rick Bright, and W. Ian Lipkin. 2021. “The Forever Virus: A Strategy for the Long Fight Against COVID-19.” *Foreign Affairs,* July/August.

Brynjolfsson, Erik, John J. Horton, Adam Ozimek, Daniel Rock, Garima Sharma, and Hong-Yi TuYe. 2020. "COVID-19 and Remote Work: An Early Look at US Data," National Bureau of Economic Research Working Paper #27344.

Choudhury, Prithwiraj, Cirrus Foroughi, and Barbara Zepp Larson. 2021. "Work-from-anywhere: The productivity effects of geographic flexibility," Strategic Management Journal 42 (4): 655–683.

Chiou, Lesley and Catherine Tucker. 2020. "Social Distancing, Internet Access and Inequality," National Bureau of Economic Research Working Paper #26982, April.

Cicala, Steve. 2020. "Powering work from home." National Bureau of Economic Research Working Paper #27937.

Davis, Steven J., Dingqian Liu and Xuguang Simon Sheng. 2021. "Stock Prices and Economic Activity in the Time of Coronavirus." National Bureau of Economic Research Working Paper #28320, revised.

DeFilippis, Evan, Stephen Michael Impink, Madison Singell, Jeffrey Polzer, and Raffaella Sadun, 2020. "Collaborating during coronavirus: The impact of COVID-19 on the nature of work." National Bureau of Economic Research Working Paper #27612.

Deole, Sumit, Max Deter and Yue Huang. 2021. "*Home Sweet Home:* Working from Home and Employee Performance during the COVID-19 Pandemic in the U.K." Working Paper.

Elhai, Jon D., Haibo Yang, Dean McKay, and Gordon JG Asmundson. 2020. "COVID-19 anxiety symptoms associated with problematic smartphone use severity in Chinese adults." *Journal of Affective Disorders* 274: 576–582.

Emanuel, Natalia and Emma Harrington. 2021. "'Working' remotely? Selection, treatment and the market provision of remote work." Working Paper, 9 April.

Erdsiek, Danial. 2021. "Working from Home During COVID-19 and Beyond: Survey Evidence from Employers." ZEW Mannheim, June.

Frenette, Marc, Kristyn Frank, and Zechuan Deng. 2020. "School Closures and the Online Preparedness of Children during the COVID-19 Pandemic." *Economic Insights,* Issue 2020001 (103), April.

Frey, Bruno S. Alois Stutzer, 2002. "What Can Economists Learn from Happiness Research?" *Journal of Economic Literature,* 40, no. 2 (June): 402–435.

Garbe, Amber, Uzeyir Ogurlu, Nikki Logan, and Perry Cook. 2020. "COVID-19 and Remote Learning: Experiences of Parents with Children during the Pandemic." *American Journal of Qualitative Research* 4 (3): 45–65.

Hatayama, Maho, Mariana Viollaz and Herman Winkler. 2020. "Jobs' Amenability to Working from Home: Evidence from Skill Surveys for 53 Countries." Policy Research Working Paper 9241, World Bank Group.

Holt-Lunstad, Julianne, Timoth B. Smith, Mark Baker, Tyler Harris and David Stephenson. 2015. "Loneliness and Social Isolation as Risk Factors for Mortality: A Meta-Analytic Review," *Perspectives on Pyschological Science,* 10 (2): 227–237.

Jiao, Guandong and Melisa Ramirez Tovar. 2020. "Remote Operations Evolve to Offshore Industry Challenges," *Offshore: Drilling & Completion*, 1 November. https://www.offshore-mag.com/drilling-completion/article/14184916/schlumberger-remote-operations-evolve-to-offshore-industry-challenges.

Kessler, Lawrence M., Alex Norwood, Richard Beem and William F. Fox. 2020. "How Will Expanding Broadband Internet Access Benefit Knox County, TN?" Boyd Center for Business and Economic Research, University of Tennssee, Knoxville.

McCullough, Jefrey S., Kartik K. Ganju, and Chandy Ellimoottil. 2021. "Does Telemedicine Transcend Disparities or Create a Digital Divide? Evidence from the COVID-19 Pandemic." Working paper, University of Michigan.

McKay, Betsy. 2021. "U.S. Drug-Overdose Deaths Soared Nearly 30% in 2020, Driven by Synthetic Opiods." *Wall Street Journal,* 14 July.

Möhring, Katja, Elias Naumann, Maximiliane Reifenscheid, Alexander Wenz, Tobias Rettig, Ulrich Krieger, Sabine Friedel, Marina Finkel, Carina Cornesse and Annelies G. Blom. 2021. "The COVID-19 pandemic and subjective well-being: longitudinal evidence on satisfaction with work and family," *European Societies,* 23.

Ozimek, Adam. 2020. "The future of remote work," Social Science Research Network Working Paper.

Pabilonia, Sabrina and Victoria Vernon. 2021. "Telework, wages, and time use in the United States," GLO Discussion Paper #546.

Sen, Ananya and Catherine E. Tucker. 2020. "Social Distancing and School Closures: Documenting Disparity in Internet Access among School Children." Social Science Research Network Working Paper, Available at: https://ssrn.com/abstract=3572922, April 10.

Servidio, Rocco, Maria Giuseppina Bartolo, Anna Lisa Palermiti, and Angela Costabile, 2021. "Fear of COVID-19, depression, anxiety, and their association with Internet addiction disorder in a sample of Italian students." *Journal of Affective Disorders Reports* 4 (100097).

Sinyor, Mark, Matthew J. Spittal, and Thomas Niederkrotenthaler. 2020. "Changes in Suicide and Resilience-Related Google Searches during the Early Stages of the COVID-19 Pandemic." *The Canadian Journal of Psychiatry* 65 (10): 741–743.

Steemers, Frank, Robin Erickson, Amanda Popiela and Gad Levanon. 2020. "Adapting to the Reimagined Workplace: Human Capital Responses to the COVID-19 Pandemic." Conference Board, October.

Taneja, Shivani, Paul Mizen and Nicholas Bloom. 2021. "Working from home is revolutionising the UK Labour market." *VoxEU CEPR Policy Portal,* 15 March.

Thoits, Peggy A., 2011. "Mechanisms Linking Social Ties and Support to Physical and Mental Health," *Journal of Health and Social Behavior* 52 (2): 145–161.

Varma, Prerna, Moira Junge, Hailey Meaklim, and Melinda L. Jackson. 2021. "Younger people are more vulnerable to stress, anxiety and depression during COVID-19 pandemic: A global cross-sectional survey." *Progress in Neuro-Psychopharmacology & Biological Psychiatry* 109 (110236).

Wallinheimo, Anna-Stiina and Simon L. Evans. 2021. "More Frequent Internet Use during the COVID-19 Pandemic Associates with Enhanced Quality of Life and Lower Depression Scores in Middle-Aged and Older Adults." *Healthcare*, 9 (4): 393.

Zuo, Geoge, W. 2021. "Wired and Hired: Employment Effects of Subsidized Broadband Internet for Low-Income Americans." Working paper, University of Maryland.

Appendix

Table A.1.: What predicts high-quality internet?

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dependent variable					Percent of the time home internet works					
log(Population Density of Zipcode of Residence)	1.38***	1.20***	1.11***	1.11***	1.24***	1.31***	1.26***	1.22***		
	(0.15)	(0.15)	(0.16)	(0.16)	(0.16)	(0.19)	(0.19)	(0.19)		
log(2019 Earnings)		1.42***	0.99***	0.96***	0.87***	0.92***	0.90***	0.93***		
		(0.16)	(0.16)	(0.16)	(0.16)	(0.16)	(0.16)	(0.16)		
Years of education			0.85***	0.84***	0.82***	0.83***	0.77***	0.80***		
			(0.20)	(0.20)	(0.20)	(0.20)	(0.20)	(0.20)		
1(Children under 18)				0.50*	0.54*	0.58**	0.53*	0.49*		
				(0.28)	(0.28)	(0.27)	(0.28)	(0.27)		
1(Minority)					-1.16***	-0.88***	-0.78**		-0.83***	
					(0.31)	(0.31)	(0.31)		(0.31)	
Black or African-American								-0.03		-0.17
								(0.47)		(0.45)
Hispanic or Latino/a (of any race)								0.00		-0.25
								(0.50)		(0.52)
Asian								-1.78***		-1.19**
								(0.54)		(0.55)
Native American or Alaska Native								-3.31***		-3.50***
								(1.03)		(1.08)
Native Hawaiian or Pacific Islander								-8.36***		-9.25***
								(2.05)		(2.10)
Other								-2.04**		-2.24**
								(0.92)		(0.94)
State FE						Y	Y	Y		
Industry of current/most recent job FE							Y	Y		
Dependent variable mean	90.24	90.24	90.24	90.24	90.24	90.24	90.24	90.24	90.24	90.24
Observations	30,646	30,646	30,646	30,646	30,646	30,646	30,646	30,646	30,646	30,646
R-squared	0.01	0.01	0.02	0.02	0.02	0.02	0.03	0.03	0.00	0.00

Note: We standardize continuous explanatory variables to mean zero and unit standard deviation so the coefficients reflect a one-standard deviation change.

Table A.2.: How the self-assessed efficiency gains from "perfect high-speed internet" services relate to observables in the SWAA data

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dependent variable	Efficiency increase from getting perfect internet (self-assessed)									
log(Population Density of Zipcode of Residence)	0.30***	0.26***	0.20**	0.20**	0.15	0.16	0.15	0.16		
	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.11)	(0.11)	(0.11)		
log(2019 Earnings)		0.27***	0.02	-0.03	0.01	0.01	0.02	0.03		
		(0.09)	(0.11)	(0.11)	(0.11)	(0.11)	(0.11)	(0.11)		
Years of education			0.49***	0.48***	0.49***	0.49***	0.42***	0.41***		
			(0.11)	(0.11)	(0.11)	(0.11)	(0.10)	(0.10)		
1(Children under 18)				0.49***	0.49***	0.49***	0.48***	0.47***		
				(0.18)	(0.18)	(0.17)	(0.17)	(0.17)		
1(Minority)					0.48**	0.45**	0.48**		0.43**	
					(0.20)	(0.20)	(0.20)		(0.20)	
Black or African-American								0.08		-0.08
								(0.27)		(0.27)
Hispanic or Latino/a (of any race)								0.64**		0.66**
								(0.32)		(0.32)
Asian								1.03**		1.28***
								(0.40)		(0.39)
Native American or Alaska Native								0.97		0.55
								(0.66)		(0.79)
Native Hawaiian or Pacific Islander								0.57		0.64
								(0.83)		(0.76)
Other								0.99		0.83
								(0.94)		(0.92)
State FE						Y	Y	Y		
Industry of current/most recent job FE							Y	Y		
Dependent variable mean	2.61	2.61	2.61	2.61	2.61	2.61	2.61	2.61	2.61	2.61
Observations	7,179	7,179	7,179	7,179	7,179	7,179	7,179	7,179	7,179	7,179
R-squared	0.00	0.00	0.01	0.01	0.02	0.03	0.04	0.04	0.00	0.00

Note: We standardize continuous explanatory variables to mean zero and unit standard deviation, so the coefficients reflect a one-standard deviation change.

Table A.3.: Regression models for the relative efficiency of WFH

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable				Relative efficiency of WFH			
Internet quality	2.53***	2.29***	2.25***	2.25***	2.25***	2.20***	2.04***
	(0.22)	(0.22)	(0.22)	(0.22)	(0.22)	(0.22)	(0.22)
log(Earnings)		2.24***	1.97***	1.79***	1.73***	1.72***	1.36***
		(0.19)	(0.19)	(0.19)	(0.19)	(0.19)	(0.19)
Years of education			0.57**	0.58***	0.57**	0.55**	0.69***
			(0.23)	(0.22)	(0.22)	(0.22)	(0.24)
1(Male)				0.35	-0.88*	-0.92*	-1.24**
				(0.33)	(0.48)	(0.48)	(0.49)
1(Children under 18)				1.70***	0.32	0.35	0.43
				(0.33)	(0.43)	(0.43)	(0.43)
1(Male) x 1(Children under 18)					2.51***	2.41***	1.84***
					(0.66)	(0.65)	(0.65)
State FE						Y	Y
Industry (of current or most recent job) FE							Y
Dependent variable mean	4.93	4.93	4.93	4.93	4.93	4.93	4.93
Observations	22,237	22,237	22,237	22,237	22,237	22,237	22,237
R-squared	0.01	0.02	0.03	0.03	0.03	0.04	0.06

Note: We standardize continuous explanatory variables to mean zero and unit standard deviation so the coefficients reflect a one-standard deviation change. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A.4.: Regression models for the relationship of subjective well-being to internet access quality during the pandemic

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable				**Subjective well-being (0 to 100 scale)**			
% time home internet works	0.14***	0.13***	0.12**	0.12***	0.14***	0.15***	0.12***
	(0.03)	(0.03)	(0.05)	(0.03)	(0.05)	(0.04)	(0.04)
x 1(Working)			0.02				
			(0.06)				
x 1(Working on business premises)				0.03			
				(0.02)			
x 1(Live with other adults)					-0.02		
					(0.05)		
x 1(Live with a partner)						-0.03	
						(0.05)	
x 1(Live with children under 18)							0.03
							(0.05)
1(Working)		8.22***	6.26	8.23***	8.22***	8.22***	8.22***
		(0.89)	(5.08)	(0.89)	(0.89)	(0.89)	(0.89)
% WFH days in the current week		0.03***	0.03***	0.05**	0.03***	0.03***	0.03***
		(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)
Controls for gender, years of education, log(2019 earnings), age bin FE, whether living with other adults, or with children, and whether living with a partner	Y	Y	Y	Y	Y	Y	Y
Dependent variable mean	62.6	62.6	62.6	62.6	62.6	62.6	62.6
Dependent variable SD	23.5	23.5	23.5	23.5	23.5	23.5	23.5
Observations	12,613	12,613	12,613	12,613	12,613	12,613	12,613
R-squared	0.08	0.10	0.10	0.10	0.10	0.10	0.10

Note: We regress subjective wellbeing on self-reported internet quality and the respondent's working and living situation. Subjective wellbeing is ten times the response to the following question: “Please imagine a ladder with steps numbered from zero at the bottom to ten at the top. The top of the ladder represents the best possible life for you and the bottom of the ladder represents the worst possible life for you. If the top step is 10 and the bottom step is 0, on which step of the ladder do you feel you personally stand at the present time?” Internet quality is based on a question of "How reliable is your internet connection?" Data are from over 40,000 survey responses collected between May 2020 and May 2021 by Inc-Query and QuestionPro. We reweight raw responses to match the share of working age respondents in the 2010–2019 CPS in a given {age x sex x education x earnings} cell. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Business Continuity Insurance in the Next Disaster

AUTHORS

Samuel Hanson*, Adi Sunderam, and Eric Zwick*****

ABSTRACT

This article draws lessons from the business support policies pursued in the COVID-19 pandemic to guide policy design for the next disaster. We contrast the performance of the Paycheck Protection Program to the Main Street Lending Program to illustrate how design principles—targeting, repayment terms, and deployment through the banks versus government agencies—affect policy outcomes. We develop a framework for understanding why a novel business support policy could usefully complement traditional support programs. One surprising insight that emerges from this analysis is that many of the market failures used to justify support during the pandemic also arise in "garden-variety" recessions. Given our framework, the policy case for small-business support during the recovery is considerably weaker than during the disaster, though credit policies that promote firm entry could aid the reallocation process.

* Harvard Business School
** Harvard Business School
*** National Bureau of Economic Research, Chicago Booth School of Business

Acknowledgements: Prepared for the Aspen Economic Strategy Group. We thank Harleen Kaur for excellent research assistance. The authors served as uncompensated consultants to Federal Reserve Bank of Boston President Eric Rosengren on the design of the Main Street Lending Program. Zwick has provided compensated expert testimony on behalf of a PPP loan servicer.

1. Introduction

The COVID-19 pandemic triggered an economic shock in the United States unparalleled in severity and breadth since the Great Depression. The shock was especially severe in the small-business sector, as many small firms operate in industries dependent upon in-person interaction. In response to this shock, policymakers deployed a suite of business support policies equally unprecedented in scale and generosity. Because past recessions provided little guidance for how to design support programs, a range of approaches were pursued, with varying degrees of success and efficiency. As we emerge from the crisis with fresh memories and the benefit of ongoing research, the current moment provides an opportunity to design better policy to prepare for the next disaster.

The purpose of this article is to draw some lessons from the business support policies pursued in the pandemic to guide future policy design. We first briefly describe the unique features of how the COVID-19 disaster affected small businesses, which include the scale of revenue losses, the number of firms simultaneously affected, and the duration of the crisis. We then focus on two business credit programs designed to help small, private firms—the Paycheck Protection Program (PPP) and the Main Street Lending Program (MSLP)—and describe how their designs affected take-up, policy impact, and the distribution of benefits.

The PPP was designed quite liberally, delivering forgivable loans to most small businesses with relatively few strings attached. In contrast, the MSLP was designed to protect federal funds by delivering support via low-interest loans that required banks to retain credit risk. The programs were similar on some dimensions with both offering broad eligibility and deferred repayment for borrowers. Yet they differed considerably in the extent of loan "softness," *i.e.*, the extent to which repayment would be required in the future and conditioned on future performance.

The performance of these programs illustrates how design principles affect policy outcomes. Both programs received allocations of approximately $600 billion in the spring of 2020. Ultimately, the PPP managed to disburse 80% of these funds to nearly five million borrowers in just over three months, while the MSLP expired in December 2020, having distributed just 3% of its total allocation and extending 1,800 loans.

While the PPP was quite successful in disbursing funds, there is strong reason to believe it was overly generous in not requiring repayment from firms that faced relatively little hardship. An emerging (but unsettled) consensus in the literature suggests that the funds proved inframarginal for many firms, in that their behavior would have changed little in the absence of support. Because private business ownership is so concentrated and the PPP was so large, one could therefore characterize the program as a tax rebate to top-1% owners equivalent to more than a full year of their typical business tax burden in return for a positive, but relatively small, impact on aggregate employment.

The lesson from both programs is that there is room for improvement in policy design for the next disaster. We turn to sketch a framework for understanding why a novel business support policy would usefully complement traditional social insurance programs. The question facing policymakers is whether providing financial support to small businesses during a disaster improves social welfare. In the absence of spillovers or financial frictions, the answer is no. In this benchmark world, firm failures are efficient, firm owners are diversified and prepared to bear aggregate risk, and barriers to entry are low.

The pandemic made salient the extent to which we must depart from this frictionless model. Supporting small businesses during a disaster can alleviate congestion externalities in bankruptcy courts, in asset markets when many firms would be forced to liquidate, and in the labor market during a time of mass layoffs. Support can help firms overcome nominal rigidities in capital contracts that prevent renegotiation and force owners to bear overhead costs that are ideally shared with capital providers. Support can strengthen firm balance sheets during the disaster, ensuring that firms are healthy enough to rehire workers and resume normal operations when the crisis abates. Finally, support can overcome traditional financial frictions that are not easily addressed through conventional monetary or credit policy, including by smoothing the uninsurable idiosyncratic risk borne by entrepreneurs.

One surprising insight that emerges from this analysis is that many of the market failures used to justify support during the pandemic also arise in "garden-variety" recessions, though we do not believe support during a normal recession should be as generous as during a noneconomic disaster. Support should only be deployed in circumstances where it would not be a bailout for malfeasance or poor past performance. These conditions are much more likely to be met during a noneconomic disaster. At the same time, a case can be made that a program of partial business continuity insurance during recessions could improve welfare.

The experience of the PPP and the MSLP highlight the importance and difficulty of solving the "targeting problem"—which firms should benefit? And how generous should those benefits be? A program that is too generous will be unfair and disproportionately benefit wealthy entrepreneurs. A program that is too restrictive will fail to make funds widely available in a timely fashion. We consider four dimensions of targeting, each of which were hotly debated during the pandemic: firm size, ownership, shock severity, and industry. We then describe key considerations in program implementation, focusing on the goal of helping firms cover recurring fixed, nonlabor obligations during a disaster; the timing of repayment; and whether a program is better administered by banks, the IRS, or another government agency.

Finally, we assess the extent to which our framework supports policy action to promote the economic recovery. To us, the policy case for small-business support during the recovery is considerably weaker than during the disaster. The market failures targeted by business support programs are most severe during the crisis, when firms face recurring obligations, difficulty in renegotiation, and the absence of suitable liquidity support from private markets. One area to focus on is ensuring that forgiveness grants are easy to apply for so that firms do not face surprise debts when the grace period for their PPP loans ends. Another option is to allow temporary continuation of unemployment insurance to the self-employed when they start a new firm. A final policy option would relax some of the restrictions in the SBA's subsidized loan program to support new entrants and address financial frictions and unusually large demand for loans that could aid the reallocation process.

2. Surveying the pandemic damage and policy response

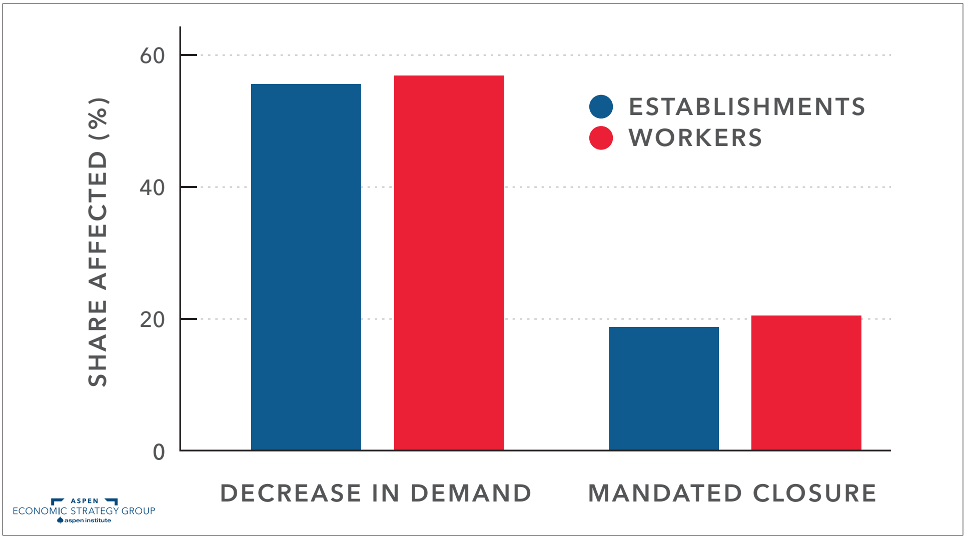
2.a. Facts on small business hardship during the pandemic

The COVID-19 pandemic triggered an economic shock unparalleled in severity and breadth across the U.S. economy since at least the Great Depression. Figure 1 presents statistics from the Bureau of Labor Statistics (BLS) Business Response Survey to the Coronavirus Pandemic. According to the survey, 1.6 million establishments experienced a government-mandated closure of their business in the spring of 2020, equal to 18.7% of all private-sector establishments, to reduce the spread of the virus. These closures affected 26 million workers, or 20.5% of private sector employment.

Beyond those firms forced to close, many more firms experienced substantial drops in demand due to the fall in mobility, stay-at-home orders, and shift to remote work for many workers. Overall, 4.7 million establishments, which account for 55.6% of all establishments, experienced a decrease in demand for products or services over this time. These establishments accounted for 72 million workers, or 56.9% of private sector employment. The duration of the economic shock was also noteworthy. Figure 2 presents data from alternative sources that reveal how long many businesses suffered. Panel A presents data from the first four phases of the Census Small Business Pulse Survey, in which respondents answer the question: "In the last week, did this business have a change in operating revenues/sales/receipts, not including any financial assistance or loans?" Panel B presents data from the Opportunity Insights

(OI) economic tracker on small-business revenue declines relative to January 2020 for firms in all industries, as well as for the food and accommodation and professional services subgroups.

Figure 1: The initial effect of the pandemic on businesses and workers

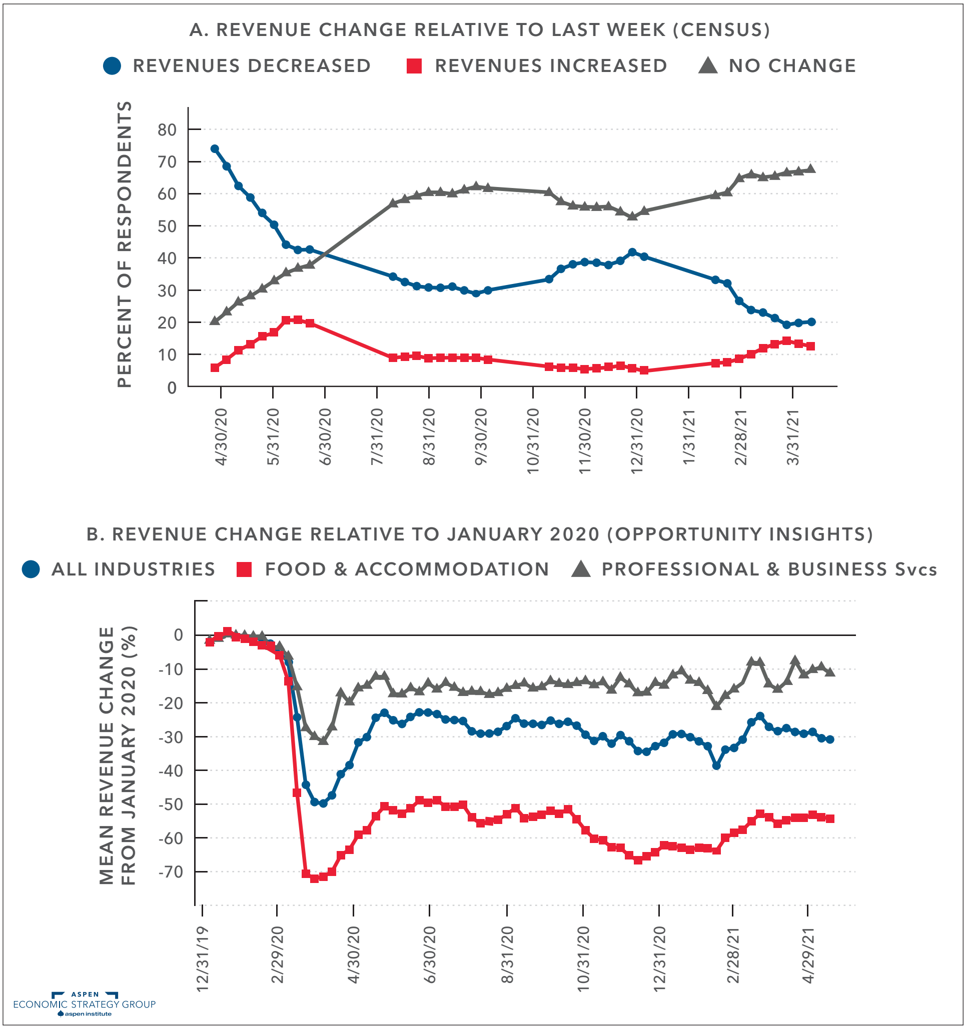


Source: BLS

Both series show a sharp, widespread decline in revenues during April and May 2020 and slow and incomplete recoveries subsequently. At the end of April, the first week of the Census survey, more than 70% of respondents report a revenue decline. This level declines to just over 40% by the beginning of June, as many states lifted stay-at-home orders and the economy partially reopened. However, the level flattens over the summer of 2020, as does the share of respondents reporting increases or no change. These patterns then partly reverse during the second wave of virus spread in the winter of 2020. Only in the spring of 2021 do we see the share of respondents reporting decreases in revenues fall to 20%.

The patterns are broadly similar for the OI data, with mean revenue declines of 50% at the trough in April 2020 and revenues remaining 30% below initial levels at the end of May 2021. The OI data also reveal the scale of heterogeneity across more and less exposed industries. Professional service firms only lost 30% of revenues on average at the trough and recovered most of this ground by May, while food and accommodation firms lost 70% of revenues at the trough and remained below 50% for the duration of the pandemic.

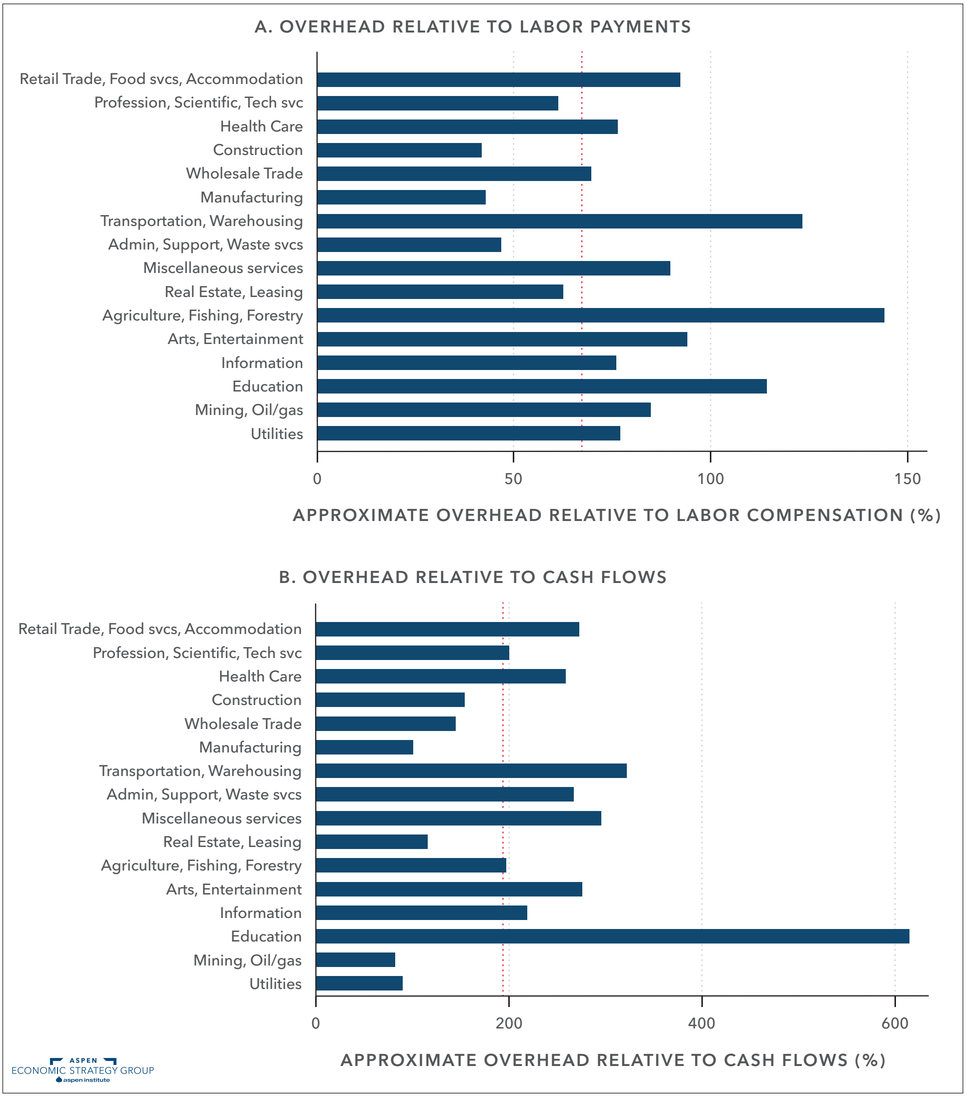
Figure 2: The duration of the shock to small business revenues



Source: Census Small Business Pulse Survey and Opportunity Insights tracker using data from Homebase.

During this time of temporary closure and revenue declines, small businesses remained obligated to cover overhead costs. Figure 3 presents an estimate of overhead costs relative to the cost of labor compensation and relative to a measure of cash flows for different industries. We sort industries in declining order of total overhead costs, as estimated in Hanson et al. (2020a). Overhead costs are measured using tax return data from a representative and weighted sample of S-corporations, which are private, closely held firms that account for a large share of employment among small businesses. We define overhead costs as revenues less the cost of goods sold, labor compensation, profits, and investment. This residual would include rent, interest payments, utilities, maintenance expenses, and local taxes. Cash flows are defined as profits plus interest and depreciation.

Figure 3: Overhead costs for small businesses are large



Source: Hanson et al. (2020a) using data from Smith et al. (2020)

On average, these firms face overhead costs equal to approximately 70% of labor payments and 200% of cash flows. Thus, even in a world with fully flexible labor expenses, sharp and persistent declines in revenues placed massive burden on these firms to meet or renegotiate these obligations or else risk permanent failure. In addition, some firms face a relatively higher burden than others. These obligations are especially important in retail trade, food and beverage services, and accommodations, sectors that also experienced worse than average economic shocks. In contrast, firms in construction and professional services face relatively light burdens due to lower capital intensity and higher labor intensity.

2.b. The impacts of business support programs

In response to the unusually sharp, severe, and widespread shock to businesses, policymakers in the United States deployed support policies equally unprecedented in scale and generosity. The two most relevant for our purposes were the PPP and the MSLP.

The intent of the PPP was to assist small firms, defined as those with fewer than 500 employees.[1] Firms were eligible for loans up to the minimum of 2.5 months of payroll in normal times and $10 million. While firms applied for PPP loans through private banks, these low-interest loans were guaranteed by the Small Business Administration (SBA). In addition, if most of the loan proceeds were used to cover eligible payroll and nonpayroll expenses, PPP loans would be forgiven. Aside from the size threshold, eligibility was defined quite broadly. Importantly, the extent of forgiveness did not depend on the severity of the shock a firm faced. Given the guarantee and the generous eligibility and forgiveness criteria, the expectation among policymakers was that most of the loans would be forgiven, thus converting to grants. To the extent the loans were not forgiven, they carried a 1% interest rate with all payments deferred for at least one year and a two-year maturity.

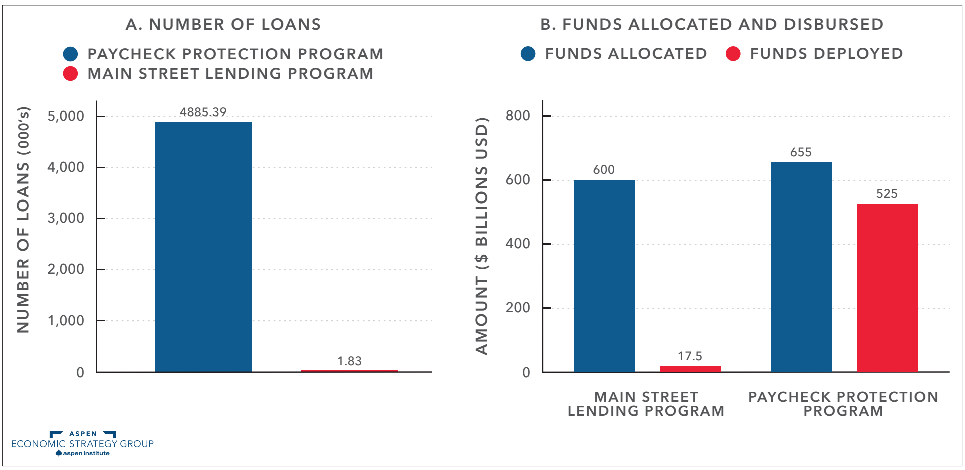
The Federal Reserve and the Department of the Treasury established the $600 billion MSLP to provide loans of up to $35 million to small- and medium-sized firms. Under the terms eventually adopted in late July, private banks made loans to qualifying firms, with the MSLP purchasing 95% of the loan and the originating bank retaining 5%. Firms were eligible for the MLSP if they satisfied size restrictions on the number of employees and revenues and also had relatively low leverage. All loans made under the program had a five-year maturity with principal payments deferred for two years and carried an interest rate of LIBOR plus 300 basis points. Firms were generally prohibited from using these loans to prepay or refinance existing debt and were subject to restrictions on executive compensation, dividends, and repurchases.

1 Firms in the Accommodations and Food Service sector (NAICS 72) could apply this threshold at the establishment level.

Three aspects of these programs deserve note. First, both programs applied quite broad targeting criteria, thus allowing participation to cover most private sector employment outside of large companies. Second, both programs featured delayed repayment, thus offering more lenient terms relative to a conventional private sector loan. Third, the programs differed considerably in the extent of loan "softness," *i.e.*, the extent to which repayment would be required in the future. The PPP's repayment terms were extremely soft—all firms that used the funds on eligible categories over an 8- or 24-week period were absolved from repayment. By contrast, the MSLP's repayment terms looked much more like traditional loans. To the extent that lenders were concerned about default, and given the cap on the loan's interest rate, loan supply under the 95% guarantee was likely to be much more conservative than under the 100% PPP guarantee. Moreover, loan demand by firms for traditional loans was likely to be much more tentative than demand for a likely-to-be-forgiven loan.

Figure 4 presents data from the two programs, pointing starkly toward how the differences across programs likely affected their distribution. Both programs received allocations of approximately $600 billion in the spring of 2020. Despite some initial hiccups, the PPP managed to disburse 80% of these funds in just over three months. In sharp contrast, the MSLP did not begin taking applications until June and expired in December 2020, having distributed just 3% of its allocation. While the PPP reached nearly five million borrowers, the MSLP issued just over 1,800 loans. Moreover, most of these funds were deployed relatively late in the pandemic in November and December of 2020. Thus, we can safely conclude the impact of the MSLP on the economy was limited.

Figure 4: Comparing pandemic credit programs, PPP vs. MLSP



Source: The PPP data from the SBA. the MSLP data from the Federal Reserve.

Given the PPP managed to deploy a remarkable amount of funds, researchers moved rapidly to evaluate the PPP as data became available. What do we know about the impacts of the program on firm behavior and survival? A full review is beyond the scope of this article, but we highlight a few points that bear on the topic of program design principles.

First, broad eligibility criteria and program generosity contributed to very strong loan demand, such that the first tranche of funds was exhausted in less than two weeks. This fact illustrates the benefits from limiting program targeting, though it also reflects the notion that most firms expected ultimate loan forgiveness.

Second, the interaction between scarce initial funds and program deployment through the banking system led some borrowers to access the program ahead of others. These borrowers tended to be larger, connected to certain overperforming banks, and more likely to be preexisting loan customers.[2] This fact points toward how program design details—specifically, whether to use private or public entities to distribute support—mediate the distribution of program resources.

Third, the program had considerable employment effects, with some uncertainty about the estimated magnitudes.[3] Even more modest estimates suggest employment impacts of several million jobs retained due to the program. However, when compared to the total size of the program, these estimates are relatively modest and point toward a substantial share of funds going to inframarginal borrowers. If the program had instead embedded soft repayment terms or conditioned forgiveness on revenue losses, as we propose below, the bang-for-buck would likely have looked much more favorable and the regressive features of the program would have been mitigated.

Finally, there is evidence that, while the funds may not have contributed to immediate employment effects, firms used the funds to strengthen their balance sheets, either by holding the loan as savings or by paying off outstanding debts, and to avoid defaults on fixed payments.[4] Thus, while the short-term employment impacts may well have been small, the longer-term impacts on firm survival could enable larger employment impacts later.

---

2 Bartik et al. (2020) and Granja et al. (2020) analyze the PPP targeting and the role of banks in mediating the distribution of funds. Humphries, Neilson, and Ulyssea (2020) shows that borrower sophistication appears to play a role in program access for small firms and sole proprietors.

3 An active literature studies the employment impacts of the PPP using alternative research designs. Autor et al. (2020), Chetty et al. (2020), and Hubbard and Strain (2020) use the 500 worker eligibility threshold to estimate employment effects, with the former two papers finding modest impacts and the latter paper finding somewhat larger impacts. Granja et al. (2020) and Faulkender, Jackman, and Miran (2020) use regional variation in program exposure generated by differences in bank performance in deploying loans, with the former paper finding modest impacts and the latter paper finding large impacts. Pardue (2020) finds that firms reduced employment after the expiration of headcount requirements needed to receive loan forgiveness.

4 Bartik et al. (2020) find that the PPP increased the firms' expected survival probability. Granja et al. (2020) find that the PPP increased firm cash holdings and reduced the probability firms missed loan and non-loan payments. Chodorow-Reich et al. (2020) find that firms used the PPP funds to pay down other loans.

More confident conclusions will take a few years to develop as more complete data arrive. In particular, we are still awaiting comprehensive data that adequately identify permanent firm failures and allow us to estimate the effects of the PPP on survival. We have also seen little analysis of either applications for loan forgiveness or the more recent tranches of PPP funding deployed in the winter and spring of 2021.

3. A framework for small business support in a disaster

During the COVID-19 pandemic, policymakers faced an unprecedented shock to small businesses with little guidance from past disasters for how to design support programs. As a result, a range of approaches were pursued, with varying degrees of success and efficiency. It is therefore useful to take a step back and reflect on whether we should prepare for the next disaster by establishing similar or better programs.

The question facing policymakers is whether providing financial support to small businesses during a disaster improves social welfare. In the absence of spillovers or financial frictions, the answer is no. In this benchmark world, firm failures are efficient, firm owners are diversified and prepared to bear aggregate risk, and barriers to entry are low. Under these circumstances, social insurance programs that target workers—such as unemployment insurance—or that provide income support to a broader group of people in need—such as economic impact payments, Medicaid, and food stamps—provide adequate support.

3.a. Rationales for small-business support

The speed, scale, and severity of the COVID-19 pandemic made salient the extent to which we must depart from the frictionless model. Consider first the case of congestion externalities. Even an economy in shock can easily absorb the resources idled by one small-business failure. When a million small businesses that collectively employ tens of millions of workers all risk simultaneous failure, strains appear elsewhere in the system.

We highlight three types of congestion externalities. First, there may be spillovers generated by congestion in the bankruptcy process. As a crisis persists, many firms will exhaust their cash reserves and become unable to service their debts and other fixed obligations. In some cases, firms will be able to work with their liability holders to voluntarily restructure their obligations. However, in many cases, businesses—especially small businesses—may be unable to renegotiate out of court. Conflicts of interest between creditors and firms, coordination failures between creditors, and other frictions are the reason bankruptcy courts exist in the first place. Eventually, concerns about firm liquidity become concerns about solvency, increasing the severity of these frictions and forcing firms to file for bankruptcy protection or permanently close.

While the U.S. bankruptcy process does an excellent job of allocating losses to liability holders and enabling large firms to restructure in normal times, it is not designed to deal with a mass wave of bankruptcies. Such a wave of business bankruptcies would create significant delays in bankruptcy court proceedings and a shortage of debtor-in-possession financing for firms operating under bankruptcy protection.

A second congestion externality arises in capital markets when a glut of firms close simultaneously. The resulting rushed business liquidations and fire sales could create large deadweight losses for society. Related to this idea is the notion that a bias toward excessive liquidation, especially for small firms, destroys franchise capital that can only be slowly rebuilt with significant start-up costs. Inefficient liquidation at substantial scale imposes negative spillovers by weakening asset values and hence the balance sheets of healthy firms.

Third, congestion in the labor market due to mass furloughs and layoffs could prevent workers from finding a new job or reentering the workforce, as well as overwhelming the unemployment insurance (UI) system. While we view the UI system's performance during the pandemic more favorably than some others in reaching an unprecedented number of beneficiaries, the pandemic did highlight a need to invest in upgrading UI systems across states to ensure timely receipt of benefits for large numbers of workers, changes to benefit formulae, and the possible extension of benefits to the self-employed.

The pandemic shutdowns came quickly, leaving many firms with significant overhead obligations—including rent, utilities, loan payments, maintenance, and employee benefits—and no cash flows to cover these costs. Nominal rigidities in capital contracts that prevent renegotiation can force owners to bear these costs and bias them toward closing permanently. This force operates not only for small businesses with loans but also for the many non-borrowers who rent real estate and equipment. When a firm closes permanently, its brand capital liquidates, as does the nexus of contracts embodied in that firm, including its relationships with customers, suppliers, workers, and capital providers. The fixed costs of startup, which create this intangible capital, are borne by the next set of owners and not fully internalized by the current owners. Such inefficient liquidation can thereby lead to an inefficiently slow recovery.

Weakened aggregate demand is another source of potential spillovers that may warrant support for small businesses. In the absence of government interventions, many businesses that are lucky enough to avoid bankruptcy during the downturn could nonetheless emerge with weakened balance sheets in an environment with low aggregate demand. This erosion in the health of firm balance sheets could greatly limit the ability of businesses to rehire workers and resume normal operations. Aid to firms in a time of crisis can help ensure that the downturn itself does not hobble the economy's productive capacity, setting the stage for swift recovery.

In addition to the various externalities associated with a mass of small-business failures, there are traditional financial frictions. In a severe downturn, particularly one triggered by shocks of noneconomic origins like natural disasters, firms that would otherwise be viable after the downturn may not be able to access capital markets to obtain the financing needed to survive the shock. For instance, in the early stages of the COVID-19 pandemic in March 2020, capital markets froze, and bank lending standards tightened substantially as banks anticipated heavy losses on existing loans. Against this backdrop, it would have been challenging (if not impossible) for many small businesses to raise incremental financing to offset their unprecedented revenue losses and help survive the pandemic. The vast majority of smaller, privately owned businesses in the United States, as well as many larger, publicly traded firms, did not have the financial reserves needed to survive such large, temporary declines in revenue. In the absence of aid, these firms would have been forced to lay off many of their employees and would have struggled to meet recurring fixed obligations.

Given the noneconomic origins of the crisis and the relative health of the financial system during the pandemic, traditional monetary and credit policy could not address the core problem facing small firms. This crisis caused increased demand for external financial support by small firms that were temporarily closed and required repayment terms that were much more flexible than traditional loan terms. This situation contrasts with a financial crisis, which can be thought of as a fall in supply of external finance driven by impaired banks. While traditional lender-of-last-resort policies focused on the financial system can help restore credit supply, these policies are less well-equipped to meet abnormally high demand and flexible repayment needs.

In the presence of these economic frictions and externalities, it is worth noting that any evaluation of government support programs needs to account for fiscal spillovers when estimating the cost of the program. For example, if there are significant employment or output effects of supporting businesses, these generate fiscal spillovers by raising tax revenue via the income, payroll, and sales tax. Such effects can meaningfully alter the perceived costs and benefits of a program relative to focusing only on the cost of the direct transfer.

A final and less appreciated departure from the standard model of firm support concerns the nature of firm ownership. In the standard model, capital is rented to firms by a large, diversified, representative firm owner. In reality, small- and medium-sized businesses are owned by one or two individuals, for whom the firm's capital accounts for a disproportionate share of their total wealth. In other words, entrepreneurs bear a dramatic amount of idiosyncratic risk entwined with the fate of their firms. This risk is uninsurable in private markets as an inescapable consequence of the same financial frictions that prevent firms from costlessly accessing external funds.

In a world with uninsurable idiosyncratic risk borne by entrepreneurs, supporting small businesses during a disaster can yield valuable social insurance benefits. Firm owners are typically not eligible for other forms of social insurance, such as UI. Moreover, UI may not provide adequate risk mitigation, given losses at the firm-level scale with firm size and may be more persistent than the typical job loss. Against this argument, one might argue that because many business owners are relatively wealthy, the consumption-smoothing benefits of supporting them are likely small. This fact highlights the importance of proper targeting when estimating a program's potential insurance value: it may be best to condition program generosity on the wealth or total income of business owners.

3.b. Garden-variety recessions are different, but. . .

One surprising insight that emerges from this analysis is that many of the market failures used to justify support during the pandemic also arise in "garden-variety" recessions. A case can therefore be made that a program of partial business support during recessions could improve welfare. At the same time, support should only be deployed in circumstances where it would not be a bailout for malfeasance or poor past performance. These conditions are more likely to be met during a noneconomic disaster.

First, garden-variety recessions are often associated with the efficient closure of low-productivity businesses (Schumpeter 1939; Foster, Grim, and Haltiwanger 2016). Current losses reflect future losses, and hence activity must be reallocated in the medium run. In garden-variety recessions, there is also a reduced need to preserve the intangible capital associated with startup costs because replacement entrants are less likely to resemble those that exit. In other words, adjustment costs are a necessary step in reallocation and are best paid when the net costs of such reallocation are low. Providing support that prevents firms from closing and labor and capital from new deployments might well postpone recovery and deter long-run productivity growth.

Second, the bankruptcy and unemployment systems work reasonably well in absence of congestion. In a garden-variety recession, firm failures are less concentrated in a narrow window of time. Instead, they occur over many months, which allows time for capital markets and traditional social insurance programs to respond. However, to the extent we worry about financial amplification, there might well be a case for low-interest government loans to small firms. Clearly, the extraordinary support programs targeting the banking system during the Global Financial Crisis (GFC) forestalled a more severe recession in 2008 and 2009. As noted above, the macroprudential toolkit offers perhaps a better source of useful policies when the problem faced by firms is a dysfunctional supply of external finance.

Thus, we would not have advocated generous small-business support during the GFC. Unlike the GFC, which was largely caused by reckless corporate behavior, the COVID-19 pandemic was a natural disaster. Like U.S. households, U.S. firms had not self-insured against the risk of a deadly global pandemic, nor should we have expected this. Indeed, the pandemic triggered such a large and widespread economic shock that even prudently managed firms faced an elevated risk of failure. And, while it is certainly the case that some firms had unwisely taken on excessive leverage, the mere fact that firms had been returning profits to shareholders instead of building up vast cash buffers is a healthy feature of our economic system, not a reckless act to be punished. A central feature of the COVID-19 pandemic was that declines in a firm's revenue during the pandemic were not highly informative about the firm's post-pandemic prospects. Such a pattern might also characterize other natural disasters, major wars, or a large-scale cyberattack, in which case concerns that business support is a bailout, keeping alive "zombie" firms that should be liquidated, are relatively low.

Nonetheless, it is desirable to allocate losses to firms' equity holders, creditors, and other fixed claimants—both to protect taxpayers and for reasons of fairness. Our view is a support program should be designed to assign as much economic loss to these private sector actors as is practicable, while simultaneously reducing the scope for damaging deadweight losses and spillovers that would impede a broader economic recovery. Given the potentially catastrophic consequences of a wave of business bankruptcies for the broader economy, we believe that, when circumstances warrant such a program at all, it is prudent to err on the side of caution and make the program broadly available on terms that are not overly onerous.

4. Targeting and implementation principles

The social welfare framework described in the previous section motivates our view of which firms to target and how to implement a support program. Previously, we articulated with Jeremy Stein (2020) concrete implementation details for a business support program called Business Continuity Insurance to be rolled out in response to the COVID-19 pandemic. In this section, we describe a few of the high-level ideas from that piece and how they relate to the general principles above.[5]

4.a. Targeting

In the ideal world, assistance would be optimally targeted toward firms (1) with operations severely affected by the shock; (2) that are unable to smooth the shock on their own; or (3) for which bankruptcies would create substantial spillovers. In short, the program should target firms with the highest private and social insurance value relative to program cost. In practice, to minimize the program's administrative burden and maximize take-up, we believe any program should use relatively simplistic targeting that exploits information already available to the government.

We consider four dimensions of targeting, each of which were hotly debated during the pandemic: firm size, ownership, shock severity, and industry. First, consider firm size. Small- and medium-sized firms face more severe financial constraints in normal times, so it is natural to expect such firms to be less able to renegotiate obligations to lenders and landlords. There are also good reasons to believe that the costs of financial distress and bankruptcy are greater for smaller firms, which are more likely to be liquidated, than for larger firms. Financial constraints are difficult to measure directly; firm size—measured using past revenues or employees—is therefore a sensible proxy that is difficult to manipulate and relatively easy for policymakers to measure. Given that even larger midmarket firms with several hundred employees and between $10 million and $100 million in revenues might struggle to raise financing in private markets during a crisis, we believe firm size thresholds should include such firms in order to cover a substantial share of private sector employment.

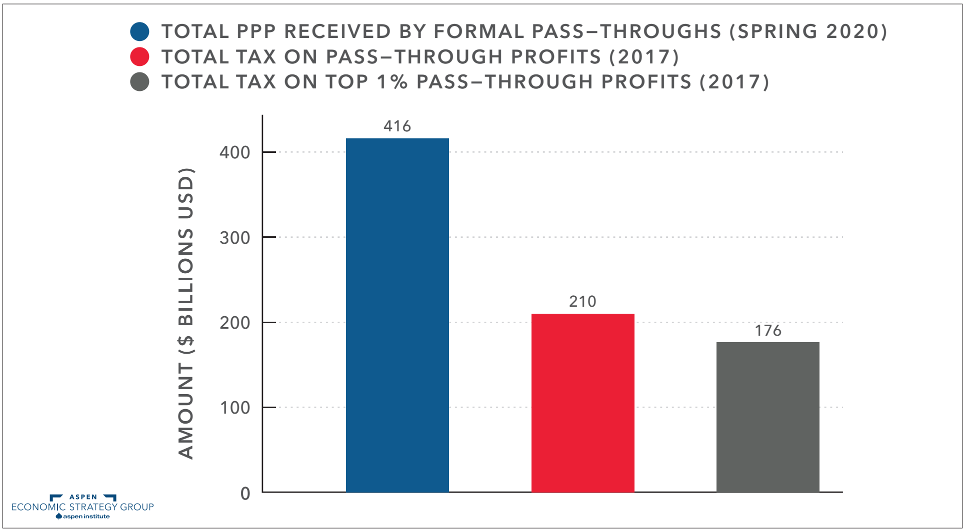
The second key dimension for targeting is firm ownership. Standalone, privately owned firms are more likely to face financial constraints and renegotiation frictions than firms with diversified ownership bases. The latter includes both publicly held firms and firms affiliated with financial investors, such as private equity or venture capital. Through their owners, these firms have access to large backstop balance sheets and the relationship capital that financial investors can deploy to renegotiate contracts when in distress. For these reasons, support programs should ideally exclude public and affiliated firms, for whom the welfare benefits of support are likely small.

5 See Hanson et al. (2020a) for our prior policy proposals and Hanson et al. (2020b) for two models that motivate business credit programs targeting larger firms.

Standalone firms are also more likely to feature insurance benefits for owners from supporting the firm in distress. The undiversified risk borne by standalone firm owners implies that sudden hardship could materially harm less-wealthy entrepreneurs, who often have limited savings outside the firm and use personal guarantees to secure financing.

On the other hand, business ownership is the typical path taken to reach the top of the income and wealth distributions. Supporting small and mid-market firms indiscriminately could disproportionately benefit the top 1% of the individual income and wealth distributions. Figure 5 helps put the magnitudes in perspective. The first two rounds of the PPP distributed $416 billion to formal pass-through firms, including S-corporations and partnerships. In 2017, such firms distributed $839 billion in ordinary business income to their owners, 70% of which typically goes to the top 1% of the income distribution. Assuming an average tax rate on this income of 30%, the tax payments on pass-through income made by these owners amount to $176 billion. Thus, if between half and all of the benefits of the PPP accrued primarily to owners, it is not unfair to characterize the program as a tax rebate to top-1% owners equivalent to between 118% and 236% of their typical business tax burden.[6]

Figure 5: The PPP transfer to business owners was potentially enormous



Source: The PPP data from the SBA. Tax data from SOI and author calculations based on Cooper et al. (2016).

6 This view of the PPP contrasts with public commentary because it distinguishes statutory incidence— which fell at least 60% on payroll—with economic incidence—the bulk of which may have fallen on owners who largely did not alter their business plans due to funding.

The third dimension of targeting that was rejected during the March and April 2020 rounds and included in the December 2020 round of the PPP is shock severity. The principle of targeting on shock severity seems obvious. Insurance benefits are typically paid in proportion to losses, for example, the revenue losses suffered during the pandemic or capital destroyed during a natural disaster. The rationale for this form of targeting is that consumption-smoothing can be achieved by replacing lost income. Analogously, the realized shortfall in a firm's revenue due to the crisis often directly determines a firm's ability to meet its recurring fixed obligations in the absence of government support. In the case of shock severity, and in contrast to firm size and ownership, measuring shock severity might be hard in real time and easier after the fact. This issue motivates a support program with backloaded payments, which can effectively implement "ex-post targeting" without delaying program rollout.[7]

Finally, in principle, policymakers could use information on a firm's industry or geography to approximate the expected revenue losses—*e.g.*, in the case of the pandemic, there were larger revenue shortfalls in the retail trade, restaurant, and hospitality industries than among firms that produce nondurable consumer goods. In practice, given uncertainty about the distribution of revenue shortfalls, the imperative to distribute funds in a timely manner, and the relative difficulty policymakers would have in verifying a firm's industry, we believe that any targeting along industry lines should be limited. Moreover, creating more generous support programs for certain firms based on characteristics that may only weakly correlate with shock severity also creates the opportunity for intense lobbying and distortion of the program by special interests. And if a program is adequately targeting based on firm size, ownership structure, and shock severity, using additional criteria to improve targeting is more likely to deter take-up and disadvantage less sophisticated firms than to promote program efficiency.

4.b. Implementation

First, we view the goal of a business support program as helping private firms cover the cost of their fixed and hard-to-renegotiate obligations, with the idea being that these costs would most threaten inefficient firm liquidations and spillover damage to the economy. We estimated firms' recurring fixed obligations using information available on corporate tax returns. The idea is simple. A firm's revenue must go toward: (1) variable costs of production as captured by the cost of goods sold; (2) compensating employees and managers; (3) depreciation; (4) equity holder profits;

7 A key lesson of mortgage modification programs from 2009 like the Home Affordable Modification Program is that too much emphasis on targeting and preventing moral hazard delivers programs that are slow and have limited impact.

or to (5) covering recurring fixed obligations. By subtracting off items 1–4 from revenue, we can reasonably approximate a firm's recurring fixed obligations. Across all corporate forms and excluding financial firms, we estimated these costs to be between $40 billion and $60 billion per week. Among S-corporations, which represent the typical private firms with employees, these costs are especially important in retail and wholesale trade and accommodations, industries with many small firms likely to be hardest hit by the pandemic.

Our approach is agnostic to the firm's choice of capital structure in terms of whether it chooses to own via mortgage borrowing or rent structures. Our approach treats leased equipment more generously than owned equipment because we exclude depreciation. We treat debt more generously than equity because we exclude profits. These exclusions align closely with the ease with which contracts can be flexibly renegotiated. Equity owners can forgo payments, while renegotiating debt is more difficult. Those who own capital outright save real depreciation expense when they are not actively using capital, implicitly reducing their user cost.

Second, given the scale of overhead costs and the heterogeneity across firms and industries, we argue that support in a noneconomic crisis should include repayment terms that are "soft," namely, such support should not take the form of traditional debt. For many firms, recurring fixed obligations are so large that it would be uneconomical for them to borrow to cover these costs. Even if firms were willing to borrow, providing support in the form of traditional loans is likely to impair firm balance sheets, creating "debt overhang" problems that could delay recovery. Well-designed repayment terms can help ensure that the only firms applying for assistance will be those that genuinely need help.[8]

For instance, if the program is implemented by the IRS, beneficiary firms would be required to gradually repay some or all of their benefits through a special corporate tax surcharge. Crucially, these surcharges should only begin once the emergency is over and the economic recovery is underway. Such a proposal would be straightforward for recipient firms and the IRS to administer: a firm's benefit account would be a simple tax account, just like depreciation or net operating loss carryforwards, that is debited each quarter to reflect the firm's tax surcharge payments.

From a corporate finance perspective, this tax-based repayment scheme is like having the government make preferred stock investments in firms affected by the crisis. Specifically, the owners of a beneficiary firm would retain the financial upside in their business, just as if the government had made a loan to the firm. However,

8 Our focus on breadth of eligibility and generosity aligns with the principles that Hubbard and Strain (2020) articulate. We contrast with their proposal in advocating that payroll be excluded from the targeted expense category, that the government is a better channel than banks for connecting firms to the program, and that forgiveness should be conditioned on some measure of hardship.

like preferred stock, this tax-based repayment scheme is “softer” than an ordinary debt contract, reducing the likelihood that beneficiary firms face insolvency following the emergency and suffer the kinds of debt overhang that limits the ability of near-bankrupt firms to hire, invest, and grow. In particular, like a preferred stock investment, this tax-based scheme ensures that repayment would automatically be extended to the extent that a firm has lower earnings in the wake of the crisis; and repayments would automatically be deferred each year if the firm operates at a loss.

Two programs that could be useful models for a tax-based scheme are (1) the net operating loss carryback refund program and (2) the First-Time Homebuyer Credit. In both cases, the tax forms for applications were short and clearly articulated the benefit formula and the conditions for eligibility.

The net operating loss carryback provision allows firms to apply for refunds for past tax payments when they incur a current loss. Historically, Form 1139 allowed firms to apply for a provisional refund with relatively few up-front requirements in terms of documentation. These refunds were applied for and approved quickly using existing IRS systems during the Great Recession. In implementing a business support scheme, the IRS could use a very similar form and process.[9]

The First-Time Homebuyer Credit was created in 2008 and allowed individuals to apply for a refundable tax credit when making an eligible home purchase via Form 5405. Individuals were required to certify that they were eligible, information that could have been used subsequently to prosecute fraud. In addition, individuals were allowed to apply the refund to past tax returns, which made them eligible to receive the credit shortly after application. Last, for those individuals who subsequently moved, there was a tax form that required them to calculate and repay a portion of the credit they received. In implementing a business support scheme, the IRS could use a similar form to determine any future repayment of grants received.[10]

4.c. Should the banks be involved?

We previously argued that small-business support should be deployed by the IRS. Our argument had three rationales. First, the IRS has direct access to the corporate

9 The case of net operating loss carrybacks highlights some practical pitfalls in designing business support programs. First, complexity in program design and application can deter take-up of such programs (Zwick 2021). Second, during the pandemic, the IRS was not prepared to accept electronic applications for net operating loss carrybacks, which delayed transmission of funds to eligible claimants.

10 It is unclear how much additional resourcing the IRS would require to implement a similar program for firms. The GAO conducted a brief review of the First-Time Homebuyer Credit along with compliance and implementation issues, including increased audits, fraud risk, and rollout of a new form, but does not provide a breakout of administration costs (GAO-10-166T, “First-Time Homebuyer Tax Credit: Taxpayers' Use of the Credit and Implementation and Compliance Challenges”). At the time of the GAO report, the IRS had processed 1.4 million claims totaling more than $10 billion.

tax returns needed to construct our measure of fixed obligations or alternative measures of benefits. As a result, it would be relatively easy for small- and mid-sized businesses to apply for and access the program. Under this implementation, firms would apply directly to the government for periodic assistance. Following an automated approval process (similar to how the IRS processes net operating loss refunds), the IRS would then send cash assistance to firm bank accounts using an electronic funds transfer. As noted above, some portion of this cash assistance could be treated as a grant and the rest would become a liability that the firm would repay over time. Specifically, beneficiary firms and the IRS would maintain a tax account tracking each firm's accumulated Business Continuity Insurance liabilities.

Second, in administering the program, the presumption should be toward disbursing funds quickly. However, there should be high penalties for fraud and abuse. This is an emergency relief program, so it should be made clear that regulatory, legal, and tax arbitrage will be dealt with more severely than in normal times. Firms who abuse the program should be held to account with high penalties, an immediate claw-back of all cash grants, and recourse to the personal wealth of managers and entrepreneurs (*i.e.*, limits on limited liability). The existing enforcement framework for tax evasion could be naturally extended to this program to lend credibility to the threat of prosecution.

Third, relative to delegating the loan application and underwriting process to banks, deploying funds through the IRS limits the extent to which agency frictions might deter private intermediaries from helping firms access socially valuable support. While preexisting relationships between banks and firms might speed access to funds, banks do not have the same incentives as the government to participate. They provide useful underwriting infrastructure, but at considerable cost that might be better internalized. In the deployment of the PPP, there was evidence suggesting that banks steered preferred clients in one direction and sped their access to loans in advance of new customers. The differential performance of some banks in supporting firm applications to the PPP materially affected the overall distribution of funds during the initial phase of the program, perversely leading regions in better economic shape to receive more funds. Furthermore, perhaps half of all small businesses do not have prior relationships beyond a checking account, and these firms tend to be those with less sophisticated and less wealthy owners who might benefit most from government support.

This principle implies that program simplicity and adequate resourcing are critical. Otherwise, administrative issues will severely limit program take-up—as illustrated by the performance of the MSLP during the pandemic. The design of the MSLP focused on avoiding losses on government loans, which resulted in a burdensome underwriting process for banks administering the loans and ultimately very low program take-up. In contrast, the simplicity of the PPP allowed it to disburse aid very quickly. Moreover, because the MSLP required risk retention by banks making loans, few banks were willing to participate despite significant underlying loan demand. Put bluntly, the rationales we highlighted above imply support programs should lend to firms at terms that are more lenient than those available from private lenders; risk retention essentially means that lending will take place at terms those lenders find attractive.

4.d. Relation to unemployment insurance

We view small-business support as a complement to, rather than a substitute for, traditional UI and expansions of UI during a crisis. We believe that the existing infrastructure for UI is a more effective vehicle for delivering aid to workers suffering hardship during a crisis. The record from the pandemic suggests that, despite initial rollout difficulties, the UI system was able to support more than 30 million newly separated or furloughed workers and disburse more than $700 billion in benefits since March 2020. Attempting to provide UI through a business support program is likely to subsidize inframarginal payroll expenses, which end up ultimately benefiting firm owners, as appears to have been the case for the PPP.

Nevertheless, the pandemic did highlight a need to invest in upgrading UI systems across states to ensure timely receipt of benefits for large numbers of workers, changes to benefit formulae, and the possible extension of benefits to the self-employed. To the extent we worry about congestion in labor markets spilling over to UI systems, it would be especially wise to invest in infrastructure to scale up or modify the program.

At the same time, the case of less-wealthy entrepreneurs demonstrates why a business support program is needed in addition to traditional UI. Traditionally, UI is not available to the self-employed, and for good reason, because moral hazard problems are too severe and because rich entrepreneurs can bear income shocks. In addition, any UI benefits provided are unlikely to scale beyond the entrepreneur's own income, thus failing to support a business struggling to cover fixed-cost payments.

Because it may be desirable to reallocate workers across firms during and after the crisis, a program that does not explicitly condition funding on retaining workers is ideal. It is also somewhat unnatural to expect firms to pay workers to idle, when accounting and payroll systems are based on hours worked and some workers are still expected to work. Tipped workers are also harder to support via a system in which the firm delivers benefits.

Of course, it would be possible to include payments covering employee wages in a support program. Programs like this were used in many other countries.[11] The total fiscal outlays of such programs are considerably larger than the type of support we propose, but would be similar to the combination of UI and business support. Given fiscal spillovers across programs, cost savings from a narrower program would only come through improved targeting or less generous benefit formulae. In this sense, a business support program that included payroll expenses could provide a backstop to UI.

Such programs have the additional possible benefit of preserving valuable firm-worker links, though at the expense of hindering reallocation. However, given large rates of recall to previous employers in pandemic recovery, fears that firm-worker links would be irreparably severed appear to have been oversold. An additional practical benefit of allowing payments to workers to be included in program support is that such payments are relatively easy for most employers to document, given existing payroll systems, compared to other expenses. This feature could help ease the process of applying for support and verifying eligible expenses for forgiveness.

Ultimately, aid to businesses and households should be paired to ensure that once the crisis ends: (1) household balance sheets are strong enough to drive a recovery in spending; and (2) business balance sheets are strong enough to drive a recovery in employment and investment. A strong recovery critically hinges on both conditions being met. A properly implemented and targeted business support program would undoubtedly help.

5. Policy tools to promote recovery

Having made the case for business support during the crisis and sketched elements of program design, a natural question is whether additional policy tools could promote recovery once the crisis has passed. Such tools might serve two welfare purposes. First, to the extent the crisis generates an aggregate demand shortfall, there is a case for traditional fiscal and monetary policy to close the output gap. Second, and perhaps more relevant in a disaster, there may be a case for promoting reallocation either by socializing startup costs or by taking other steps to facilitate firm entry and exit.

To us, the policy case for small-business support in the wake of the shock is considerably weaker than during the shock. The market failures targeted by business

11 See Hubbard and Strain and BPEA for surveys of many of these programs.

support programs are most severe during the crisis, when firms face recurring obligations, difficulty in renegotiation, and the absence of suitable liquidity support from private markets. For these reasons, we do not believe the government should be as involved in promoting recovery via direct small-business support once the crisis subsides. As bankruptcy court and labor market congestion risks diminish and output demand returns, the insurance and corrective value of support approaches zero. Small-business support therefore becomes a transfer to firm owners.

Ultimately, a large number of businesses will need to have their liabilities permanently restructured in bankruptcy, and many firms do need to be liquidated. For reasons of efficiency, fairness, and fiscal prudence, policymakers should not attempt to indefinitely delay business bankruptcies. The goal should be to ensure that this unavoidable bankruptcy wave plays out in a far more orderly fashion than it otherwise would have. Policymakers could continue the temporary extension that prevents debt forgiveness from being treated as taxable income.

One area to focus on is ensuring that forgiveness grants are easy to apply for, such that firms do not face surprise debts when the grace period for their PPP loans ends. Lenders have strong incentives to encourage borrowers to apply for forgiveness in order to avoid losses from loans that would otherwise default. Borrowers however may face significant challenges in submitting forgiveness applications, the forms for which are considerably longer and more complex than were the forms for initial loans. The risks of a nonstandard forgiveness process suggest that aligning the program with the standard tax-filing season would be an improvement.

A second challenge facing the recovery is how quickly the businesses that did close will be replaced. To promote and support entrepreneurship during the recovery, one possibility is to allow temporary continuation of UI to the self-employed when they start a new firm. The idea here is to subsidize start-up costs to correct for the loss in franchise capital due to inefficient liquidation. This program could be modeled after the French PARE program, which offers unemployment benefits to new entrepreneurs.[12] Of course, such a program would require careful monitoring to protect taxpayers from abuse. Another policy option is providing subsidized loans for new entrants to address financial frictions and unusually large demand. Usually, the SBA's 7(a) Loan Program requires attestations that a firm has sought and failed to receive financing in the private market. Such requirements could be temporarily waived.

To be sure, our position depends on the support offered during the crisis having been timely and adequate. Moreover, some of the insurance value described above

12 Hombert et al. (2020) find that the French program increased firm creation in the wake of a recession without decreasing the quality of new entrants.

remains relevant. For example, less-wealthy small-business owners who suffer considerable losses may benefit from prolonged access to social insurance programs.

We are less sanguine about the value of traditional fiscal policy levers targeting firms, such as investment or payroll tax credits. First, these policies often require firms to be operating with positive taxable income to benefit from them immediately. Such requirements result in less vulnerable firms receiving a disproportionate share of dollars spent. Second, the fundamental problem facing firms during the crisis is a fall in output demand, which makes supply-side policies less likely to be effective. Social insurance targeting firms in need is the better medicine.

We are also skeptical of proactive policy that uses a heavy hand to encourage reallocation, such as by selecting particular industries to subsidize. In general, it is unclear how much reallocation the economy needs and what form that reallocation should take. Commentators and policymakers have a tendency to underestimate mean reversion in economic systems, thus risk encouraging the wrong behavior while creating programs that may invite abuse.

References

Autor, David, David Cho, Leland Crane, Mita Goldar, Byron Lutz, Joshua Montes, William B. Peterman, David Ratner, Daniel Villar, and Ahu Yildirmaz. 2020. "An Evaluation of the Paycheck Protection Program Using Administrative Payroll Microdata." *Mimeo*.

Bartik, Alexander W, Zoe B Cullen, Edward L Glaeser, Michael Luca, Christopher T Stanton, and Adi Sunderam. 2020. "The Targeting and Impact of Paycheck Protection Program Loans to Small Businesses." National Bureau of Economic Research Working Paper #27623.

Chetty, Raj, John N Friedman, Nathaniel Hendren, Michael Stepner, and The Opportunity Insights Team. 2020. "How Did COVID-19 and Stabilization Policies Affect Spending and Employment? A New Real-Time Economic Tracker Based on Private Sector Data." National Bureau of Economic Research Working Paper #27431.

Chodorow-Reich, Gabriel, Olivier Darmouni, Stephan Luck, and Matthew C. Plosser. 2020. "Bank Liquidity Provision across the Firm Size Distribution."

Faulkender, Michael W, Robert Jackman, and Stephen Miran. 2020. "The Job Preservation Effects of Paycheck Protection Program Loans." *Available at SSRN 3767509*.

Granja, João, Christos Makridis, Constantine Yannelis, and Eric Zwick. 2020. "Did the Paycheck Protection Program Hit the Target?" National Bureau of Economic Research Working Paper #27095.

Hanson, Samuel, Jeremy Stein, Adi Sunderam, and Eric Zwick. 2020a. "Business Continuity Insurance and Loans: Keeping America's Lights on During the Pandemic." *Policy Brief*.

Hanson, Samuel, Jeremy Stein, Adi Sunderam, and Eric Zwick. 2020b. "Business Credit Programs in the Pandemic Era." *Brookings Papers of Economic Activity, forthcoming*.

Hombert, Johan, Antoinette Schoar, David Sraer, and David Thesmar. 2020. "Can unemployment insurance spur entrepreneurial activity? Evidence from France." *Journal of Finance,* 75 (3): 1247–1285.

Hubbard, Glenn R., and Michael Strain. 2020. "Has the Paycheck Protection Program Succeeded?" *Brookings Papers of Economic Activity, forthcoming*.

Humphries, John Eric, Christopher Neilson, and Gabriel Ulyssea. 2020. "Information Frictions and Access to the Paycheck Protection Program."

Pardue, Luke. 2021. "Small Business Experiences as the Paycheck Protection Program Ends: Evidence from Covered Period Expiration." *Journal of Public Economics, forthcoming*.

Smith, Matthew, Danny Yagan, Owen M Zidar, and Eric Zwick. 2020. "The Rise of Pass-Throughs and the Decline in the Labor Share." Working paper.

Data-Driven Opportunities to Scale Reemployment Opportunities and Social Insurance for Unemployed Workers During the Recovery

AUTHOR

Till von Wachter*

ABSTRACT

As unemployment insurance (UI) benefits and other economic stimulus measures are rolled back, less-educated, minority, younger, and female workers remain at risk of economic hardship and long-term consequences from ongoing job loss and unemployment during the recovery. This chapter discusses opportunities to quickly and effectively expand and scale economic supports and workforce services to assist low-income workers and those at risk of long-term unemployment. The initiatives I describe are designed to yield immediate benefits. They would leave the UI and workforce systems more resilient and more equitable going forward and they could readily be expanded into comprehensive reforms. The approaches recognize the new and uncertain economic environment by integrating ongoing learning and improved data collection. They also take into account shifting economic needs and political climate during the recovery with an employment-centered approach to cyclical labor market policy.

* University of California Los Angeles, California Policy Lab, National Bureau of Economic Research, Center for Economic and Policy Research. Email: tvwachter@econ.ucla.edu

Acknowledgements: I thank Alex Bell, TJ Hedin, Amy Ganz, Melissa Kearney, Roozbeh Moghadam, Geoffrey Schnorr, and Aparna Ramesh for helpful conversations.

1. Introduction

The US labor market is emerging from its deepest crisis since the Great Depression. The job losses and unemployment induced by the COVID-19 pandemic were of unprecedented scale and were concentrated among workers who were already vulnerable to adverse economic and social outcomes, such as less-educated workers, Blacks, Hispanics, younger individuals, and women. Despite the ongoing economic recovery, past experience indicates job loss and unemployment will continue for some time even after the recession is officially over and even absent economic effects from a resurgent pandemic. Moreover, a substantial number of individuals remain at risk of long-term scarring effects from the COVID-19 recession: job losers, the long-term unemployed, and young labor market entrants (*e.g.*, von Wachter 2020). At the same time, the economic recovery brings opportunities to deploy workforce services to reintegrate unemployed workers into employment and assist them in obtaining better jobs.

Job loss during recessions has been shown to lead to a range of long-lasting consequences for workers, including long-term losses in earnings, increases in mortality, and consequences for their families (*e.g.*, Davis and von Wachter 2011; Sullivan and von Wachter 2009). Long-term unemployment and unemployment insurance (UI) benefit exhaustion in particular have been associated with increases in poverty (*e.g.*, Rothstein and Valletta 2017; Ganong and Noel 2019). As in past major recessions, long-term unemployment is an important and potentially worrisome phenomenon during the recovery from the COVID-19 crisis. The fraction of the unemployed that have been out of work for 27 weeks or more stands at 40.9% (May 2021), close to the postwar peak during the Great Recession of 45.5% (April 2010) and is likely an underestimate of the true rate (BLS 2021). In California, 50% of all UI claimants and 20% of the pre-crisis labor force received more than six months of UI benefits, with higher incidence rates of long-term unemployment among more vulnerable workers and poorer and historically marginalized communities (Bell et al. 2021a).

In addition, the prolonged crisis has put financial strain on many of the often lower-income households most strongly affected by job loss and long-term unemployment. Furthermore, even though the United States has not fully recovered the employment lost at the start of the pandemic, vaccination rates are stalling, and COVID-19 remains a threat to public health, many states and the federal government have scaled back unemployment insurance benefits and other economic stimulus. While this is meant to help speed reemployment, it may put many low-income workers at risk of economic hardship since regular UI benefits are well below federal poverty thresholds (Bell et al. 2020a). Without further intervention, reentry into the workforce may also perpetuate existing inequalities if lower-income minority and female workers return to the same low-paying jobs they had before the crisis. Moreover, many workers whose jobs were permanently lost due to the pandemic may be better suited pursuing new training opportunities rather than immediate reemployment (Barrero et al. 2020).

A key question is what can be done to further reemployment during the recovery, and to assist those workers who are particularly at risk of economic hardship and of the long-term adverse effects of job loss and long-term unemployment. This chapter takes the view that existing programs and services can be effectively scaled to help avoid hardship while further speeding reemployment and assisting workers in obtaining better-paying jobs. The chapter discusses four broad opportunities to expand and scale economic supports and workforce programs effectively by harnessing existing programs and data infrastructures. The policies proposed are, in rough order of urgency:

1. Harness states' UI systems and similar large social programs to scale and target income support and workforce services to workers at risk of poverty or of adverse consequences from job loss and long-term unemployment.
2. Expand and subsidize Short-Time Compensation programs to speed rehiring, reduce churn, and allow and encourage job-related training during the recovery.
3. Institute a trigger-based policy grounded in economic theory that automatically adjusts benefits and eligibility for UI benefits to raise recipiency and equity.
4. Reform the UI data infrastructure to enable data-driven UI and workforce policy and support effective and equitable real-time decision making.

While all of these policies would provide substantial improvements to the US social insurance and workforce system in future recessions or another pandemic, a key aspect of these proposals is that they would have an immediate impact during the current economic recovery. Most, if not all, could be implemented by specific actors at the federal and state level without establishing new programs or creating new funding streams. Throughout, we refer to those workers that are either currently experiencing economic hardship or likely to experience adverse consequence from job loss or long-term unemployment as "at risk." During this crisis, many of these at-risk workers come from communities or have characteristics that had put them at a disadvantage in the labor market before the crisis, such as minorities, women, or lower-educated workers.

This set of proposals also recognizes and addresses several potential challenges to better insuring and reintegrating at-risk workers in the current environment. Given the scale of job loss, there are likely large numbers of long-term unemployed workers but limited funding for workforce development programs and other employment services. For example, in fiscal year 2018–2019, California served about 500,000 workers in federally funded workforce programs, such as the Workforce Innovation and Opportunity Act (WIOA) and the Wagner-Peyser Act programs. In contrast, nearly four million workers received more than six months of UI benefits in the year following the onset of the COVID-19 crisis. Hence, effectively targeting a potentially limited amount of services and resources may become crucial, as I discuss in Section 2.

Another concern is that the characteristics of long-term unemployed, discouraged, and other at-risk workers differ from those in past recessions due to the nature of the pandemic. Job losses in food services, retail, social, and personal services disproportionately affected lower-income, younger, and more vulnerable workers. These workers neither correspond to the profile of typically more mature "dislocated workers" who may have lost stable, higher-wage jobs due to economic restructuring, nor to the typical profile of hard-to-employ individuals who are the focus of WIOA Adult and Youth programs. Thus, labor market policy will be navigating uncharted waters during the recovery, and it will need to continuously adapt and improve. The proposals specify opportunities to structure outreach and services to allow for an ongoing learning process about the take up and effectiveness of programs among minorities, younger workers, women, and lower educated workers.

Finally, each proposal highlights specific actions that could be taken immediately by specific actors to scale workforce programs and other services. Where appropriate, the proposals also lay out medium- or longer-term actions to improve the nation's social insurance and workforce system. Although the proposals are not meant to offer a blueprint for wholesale reform, each proposal would constitute key components of reform and could be further scaled.

2. Connecting and targeting income support and workforce programs

2.a. Need

With a high rate of long-term unemployment and reduction in UI benefits in many states, a large number of often low-wage and vulnerable workers are at risk of adverse long-term consequences and economic hardship. At the same time, with the recovery gaining traction and available funding for workforce programs increasing, there are increasing opportunities to help these workers reenter the workforce and find better jobs. Job search assistance programs in particular have been shown to be impactful and cost effective, while job training programs can lead to long-term improvements in job outcomes (Card et al. 2018). Yet, these and other workforce services are often underutilized by the unemployed.

Policymakers need the ability to reach out to large groups of potentially at-risk workers with income support, workforce and other services quickly and effectively. Large programs such as UI or the Supplemental Nutrition Assistance Program (SNAP) serve millions of workers and collect information on earnings and family status that can in principle be used to assess need. Yet, these and other state and federal programs often operate independently with limited referral between them. While funding for income support and workforce services is often available, many unemployed or low-income workers are not aware of programs for which they might be eligible. Further, given the scale of the crisis, many workers may be in need of and eligible for government assistance for the first time, and hence not aware of available services.

2.b. Proposal

(1) Connect: Harness existing service relationships between large government programs, the workers they serve, and the data infrastructure used to provide services to quickly and effectively reach out to at-risk workers with information about additional income support and workforce services.

(2) Target: Systematically target workers most in need with information about income support and workforce services using administrative individual data that is already being used to assess eligibility and hence contains relevant information for predicting need and eligibility for other programs.

(3) Evaluate: Use large-scale, targeted outreach to build short-, medium-, and long-term evaluation mechanisms to improve effectiveness of services and refine targeting.

2.c. Details and discussion

2.c.1. Connect

Federal, state, and local government agencies already serve many low-income and other individuals and maintain databases of contact information, service records, and information on economic and family status used to assess eligibility for their clients. This network of existing service relationships can be used to reach out to individuals at high-risk of long-term unemployment or hardship with information about additional services for which they might be eligible. This outreach can occur via direct emails or text messaging, or through postings on online service accounts. By sharing relevant weblinks and information on how to access other programs, such messaging would point workers and other vulnerable individuals directly to available services. (In addition, agencies can use their standard communication channels for general messaging, such as their websites, press releases, or Twitter feeds.) Importantly, such outreach can take place based on data available within a given program and does not require potentially complex changes to data infrastructure, such as combining data of different agencies.

An example of a successful outreach campaign through the UI system occurred in California, where the UI agency (the Employment Development Department, EDD) sent messages to claimants about the availability of CalFresh benefits, California's SNAP program. Motivated by the potential expiration of UI benefits at the end of December 2020, staff from EDD and California's Labor and Workforce Development Agency (EDD's parent agency) coordinated with the CalFresh team at the California Department of Social Services (CDSS) to develop appropriate language to advertise CalFresh benefits to UI claimants at risk of benefit exhaustion.[1]

The messaging occurred through a post in the online accounts that claimants access for certifying UI benefits. EDD piloted the messaging for two days in December, when the extension of UI benefits had not yet been signed into law. Only claimants who were at risk of exhausting benefits at that point received the message. The outreach was highly effective in that it led to a rise in CalFresh applications of close to 40,000 in a single day. It was also efficient in that over 90% of applicants qualified for CalFresh benefits, allaying concerns that the messages would lead to a large number of ineligible claims. However, the large spike in applications led to bottlenecks at county welfare offices. In a second wave of outreach in June 2021, policymakers incorporated lessons learned from the first round and expanded messaging to include information about rental subsidies and health-care benefits.

This outreach was a success for several reasons. Staff at EDD and CDSS cooperated directly to coordinate the content and timing of messaging, including capacity constraints at county welfare offices throughout the state. An online screening tool for CalFresh allowed workers to assess eligibility in a few screens. EDD also included links to other services in its messaging throughout Spring 2021 without direct cooperation among agencies administering those programs, demonstrating that direct agency coordination was not always necessary.

1 The cooperation was important in particular since CalFresh eligibility rules had been modified throughout the crisis in response to federal legislation.

Social service providers that could be involved in such outreach are UI, SNAP, Medicaid, and Temporary Aid for Needy Families (TANF). All these programs are administered by state agencies with federal oversight. As a result, while state agencies take the lead on administering benefits, federal partner agencies such as the US Department of Labor (UI), the US Department of Agriculture (SNAP), or the US Department of Health and Human Services (Medicaid, TANF) can play an important role in developing and promoting blueprints for using the programs' data infrastructures for scaling outreach. In addition, the Social Security Administration and the Internal Revenue Service, among others, serve millions of potentially vulnerable individuals directly through Social Security Disability Insurance (SSDI), Supplemental Security Income (SSI), Social Security, or the Earned Income Tax Credit (EITC).

With the staggering increase in claimants during the crisis, the UI program is a promising candidate for connecting potentially vulnerable individuals to income support programs such as SNAP, or to workforce services programs such as job search assistance or job training. After many years of low uptake among the unemployed, state agencies administering the UI program now have expanded its reach to millions of workers at risk of long-term unemployment, and often administer federally funded job search assistance programs, such as Wagner-Peyser, and more intensive workforce services for dislocated or low-income workers funded by the WIOA. Since the unemployed must certify for benefits weekly or bi-weekly, depending on the state, UI agencies regularly communicate with their clients, and hence are able to inform them about the workforce services they manage, or other relevant services, such as SNAP, Medicaid, or state and local rental relief programs.

Messaging can go beyond sharing direct web links or contact information to other programs. A growing body of work in behavioral science examines the impact of many aspects of messages sent by government agencies, including the content, framing, and medium of the message (Thaler and Loevenstein 2008; Bhanot and Linos 2020). Such research is available to agencies to improve their outreach strategies. Furthermore, as we discuss in the next two sections, messaging can be greatly improved through targeting, and through data that is routinely collected in the process of messaging.

Outreach through existing service relationships, in particular the UI program, can be scaled quickly and extensively. However, not all individuals will be reached by such efforts. Many lower income and older individuals do not use cell phones, PDAs, PCs, or other devices to interact with government service providers. Some marginalized communities may not be receiving UI, SNAP, or other benefits in the first place, and hence will not benefit from this type of outreach.

The traditional approach to reach such workers is to harness the networks and relationships of local government agencies or community-based organizations. In addition, through the expanded reach of UI, SNAP, and other services during the crisis, it is worth exploring how clients' own social and work networks may be harnessed to further spread the word or be used as a referral mechanism. Such "respondent-driven" outreach is frequently used to survey hard-to-reach populations. A similar approach can be used to reach out to whole communities.

2.c.2 Target

Outreach should be targeted to those individuals most in need, most likely to benefit, and most likely to be eligible. Targeting increases the effectiveness of messaging and leads to an efficient use of client and staff resources spent on applying and screening for services. Most large social service providers have information about clients that allows for targeting, such as eligibility (*e.g.*, income or family status) or need (e.g., benefit exhaustion). If messaging is to be scaled quickly, limited targeting can still raise the effectiveness of messaging or help to control the potential flow of service applications. In the absence of any ability to target, simple and straightforward online screening tools can help effectively convey eligibility for a program (*e.g.*, the role of family status for eligibility for SNAP or the EITC).

With additional time and data, more sophisticated targeting is feasible. One approach is to predict the likelihood of an adverse event, often done with the goal of early intervention. For example, the UI system currently targets certain workforce services to workers based on their probability of exhausting UI benefits through the Worker Profiling and Reemployment Service (WPRS) and the Reemployment Services and Eligibility Assessment (RESEA) programs. The California Policy Lab (CPL) has helped Los Angeles County to target housing support based on the probability of experiencing a spell of homelessness (Bertrand et al. 2019). An alternative is to target a program to those workers for whom research suggests it will be most effective (Knaus, Lechner, and Strittmatter (2020) provide an example how this could work for the case of job search assistance in Switzerland). This kind of targeting can raise the impact of limited resources. Such an approach requires separate, credible estimates of treatment impacts for a meaningful number of client groups, which often are not available. Nevertheless, new statistical methods and IT capacity can harness increasing amounts of data to make such impact-based targeting potentially feasible for large social programs (von Wachter 2021).

2.c.3. Evaluate

Changing characteristics of workers at risk of long-term unemployment has pushed federal and state workforce systems into uncharted territory. Programs serving dislocated workers from traditionally shrinking sectors (such as WIOA's Dislocated Worker Program) or sectors affected by trade (such as Trade Adjustment Assistance, TAA) may not work in the same way for retail, restaurant, or other service workers seeing their lines of work diminished because of structural changes brought or accelerated by the pandemic. There is an urgent need to obtain additional evidence on the effectiveness of workforce services and other programs for such workers. Resulting impact estimates can be used to improve the services that are being provided. If such evidence is available for a sufficiently large number of client groups, they can also be used to improve targeting.

If outreach efforts are designed from scratch, they can be structured such that short-term impact estimates can be recovered in close-to-real time as the program is being rolled out. A well-designed outreach effort will collect basic statistics on how many clients accessed emails, text messages, and the web links they contained. In addition, the design of who is targeted or the design of the content of messages could be used to obtain impact estimates. Policymakers should consider designing targeting strategies to obtain program impact estimates, such as targeting workforce services based on the probability of exhaustion of UI benefits. Traditionally, UI claimants being considered for services are stratified into tiers by their probability of exhaustion. Outreach can begin with the top tier of claimants, while clients are assigned to workforce services, tier by tier, until capacity is reached. If the lowest tier cannot be served completely, participants within that tier can be randomly assigned.

Another design could instead select a random group of individuals within each tier who are assigned to more intensive workforce services and the control group is assigned to basic workforce services. Then, one would obtain valid impact estimates for each separate predicted exhaustion tier. The information can then be used to adjust the program and its targeting. Such a strategy of stratified randomization is particularly appropriate if the optimal approach to targeting is not known (for example, if targeting the highest tier is not necessarily optimal). In this example, those individuals at highest risk of benefit exhaustion may not necessarily be the ones who will benefit most from the services offered.[2]

2 If it is difficult to target based on individual characteristics, an alternative strategy is to randomly select groups of individuals, and then reach out to groups sequentially over time. Due to randomization, this can yield valid control groups unless economic conditions change very rapidly. Such an approach can be particularly useful if there is a concern with capacity in processing caseloads of new program applicants, such as in the SNAP-UI example from California.

Finally, there is a growing literature that tests the efficacy of the content, framing, and style of messages, especially when sent from a government provider (Linos et al. 2020). Testing the efficacy of messaging is valuable in its own right to continue improving the outreach strategy. In addition, different types of messages can be used to manipulate the number of individuals responding to the outreach. In that case, message types function as an experiment (with imperfect take up) and can be used to estimate program effectiveness.

3. Short-time compensation: employment-based labor market insurance

3.a. Need

Short-Time Compensation (STC, sometimes also called Work Sharing) provides workers with partial UI benefits while they remain employed at reduced hours and full benefits, and employers with the opportunity to reduce labor costs by reducing employee hours while avoiding layoffs. Currently, STC also allows firms to rehire previously laid off workers on a part-time basis. By temporarily subsidizing part-time work, STC provides flexibility to firms and helps to speed the rehiring process during the recovery. By limiting layoffs, it also helps to minimize the number of job seekers and hence crowding in the labor market. High expected and actual rates of recall among the unemployed, a substantial amount of partial UI receipt, and churn in and out of the UI system during the COVID-19 crisis suggests that attachment of workers and employers has remained high (Bell et al. 2021b). While STC has not seen broad uptake during this crisis, this could be remedied effectively in several ways, which are discussed below.

Importantly, even if firms permanently reduce employment as a result of the recession, as would be the case if the economy is undergoing reallocation between sectors, shifting such permanent layoffs into the future when the recovery has gained strength can reduce the long-term cost of layoffs for workers and society. More generally, STC insures workers against earnings losses over the business cycle by linking payments to employment rather than unemployment, helping to reduce some of the well-known drawbacks of UI. Low-income workers on STC are still eligible for the EITC, taking into account that today much of income support in the United States is now provided through the tax system.

3.b. Proposal

(1) Scale:

(1a) Develop targeted outreach to employers using approaches outlined in Section 2.

(1b) Allow payroll processors to file STC plans for their customers.

(1c) Require firms receiving government business loans to enroll in STC.

(2) Subsidize: Establish a direct subsidy for firms taking up the STC program instead of pursuing layoffs.

(3) Train: Allow workers and firms on STC to participate in subsidized training activities.

3.c. Details and discussion

3.c.1. Scale

To participate in the STC programs, firms have to first file an STC plan with the UI agency that specifies the number of workers involved, the number of hours reduced, and the number of layoffs avoided. Hours reductions usually cannot be more than 60% or less than 10%, and a minimum amount of a companies' employees have to participate. Once the plan is approved, the firm and participating employees jointly certify for UI benefits weekly, and workers receive prorated UI benefits based on the earnings loss. While UI claimants who work part time while receiving UI benefits can also receive prorated benefits by filing UI benefits on their own, these are lower than corresponding STC benefits, and workers are not guaranteed to keep receiving health and pension benefits.

A central challenge to the STC program, which is part of the UI program and available in more than 30 US states, is that it is not well known among employers. Evidence suggests that participating employers are satisfied by the program and that informing employers can raise awareness of the program (Houseman et al. 2017). The data that states employ to administer the UI program can be used to develop targeted outreach strategies in a similar fashion as discussed for workers in Section 2. For example, those firms could be initially targeted with information about STC that (1) have a lot of workers working part time while receiving partial UI benefits; (2) typically recall many workers or whose workers experience a lot of churn in and out of UI; and/or (3) have used STC in the past.[3]

3 To process incoming claims swiftly, it is helpful if the STC program is as fully automated and as accessible online as the rest of the UI program. This was a hold up for advertising and scaling the program in California before mid-2021. Earmarked federal funding for the improvement and establishment of STC programs can be used for this purpose.

STC has bipartisan support, in part because it promotes employment rather than subsidizing unemployment. Congress fully funded STC programs in 2012 to reduce crowding in the labor market after the Great Recession (Strain and Hassett 2014). Working with employer stakeholders, such as the Chamber of Commerce, or firms providing services to many employers, such as payroll processors, scheduling platforms, or human resource management platforms, would help with advertising the program. Similarly, as discussed in Section 2, harnessing insights from behavioral science for framing messaging and providing clearly accessible information will also help, as STC is a complex program involving the participation of both workers and firms. Several states have successfully scaled STC during the COVID-19 crisis (*e.g.*, Michigan or Washington) and their experience can be used as a model.

The US Congress and the US DOL should allow payroll processors to assist employers in filing STC plans. The administrative process of filing an STC plan can be burdensome for a single employer that does not know the program. Since payroll processors have to be notified of reductions in work hours, it makes sense to involve them in filing an STC plan. Given the large number of businesses they serve, payroll processors would quickly gain substantial expertise in filing such plans and would be able to effectively interface with the occasionally cumbersome web or paper forms provided by UI agencies.

Currently, businesses that operate in multiple states must comply with different STC program rules for each state. The complexity introduced by conflicting state rules can deter these employers from participation altogether. To avoid these complications, the US Congress should consider establishing a unified set of rules for states' STC programs and requiring the program in all states, which would also aid with scaling the STC program through the involvement payroll processors. Another alternative would be to institute a federally funded, national STC program, as outlined in von Wachter and Wandner (2020). To further aid in the take up and scaling of STC during recessions or national crises, a national program could simplify the STC benefit formula to allow payroll processors to directly compute and pay STC benefits to workers, provide information about adherence to program rules to the US DOL, and be reimbursed directly by the federal government.

US Congress should also require establishments to participate in an STC program if they receive emergency business loans (von Wachter 2020). While not all firms receiving loans will make employment adjustments, the fact that they applied for an emergency loan likely signifies the firm may need to do so during or after the period of the loan. STC provides these firms with a mechanism to reduce labor costs while avoiding layoffs as they are adjusting to changing economic conditions through the recession and recovery. Enrollment in an STC program also helps to guarantee that business loans stabilize jobs and ultimately benefit workers. Furthermore, by creating a direct link between business loans and worker-level data in the UI system, this step allows for measuring the impact of emergency business loans on employment outcomes.

3.c.2. Subsidize

Enrollment in the STC program should be subsidized because firms are unlikely to internalize the social value of reducing layoffs and crowding in the labor market. This is because the cost of job loss to workers occurs over the long run in the form of lower wages, especially (but not only) if job losses occur in recessions (Davis and von Wachter 2011; Lachowska, Johnston, and Mas 2020; Schmieder, von Wachter, and Heining 2020). In addition, enrolling in STC is more costly for firms than either full or partial UI because they must continue to pay for health care and pension benefits and incur administrative costs from joining the program. While in theory firms benefit from retaining skilled workers, the reality in a slack labor market is that firms are likely to be able to rehire laid-off workers.

In the Great Recession and the COVID-19 crisis, benefit payments were fully paid for by the federal government instead of by states' UI trust funds. However, not all states have passed the cost-savings from STC on to participating firms in the form of lower payroll tax rates. Firms that have laid off workers in the past will face higher payroll taxes, known as an experience rating. The federal government should automatically fund STC fully during recessions and exempt firms from increases in payroll taxes due to a rise in UI receipt by their workforce through participation in an STC program. In addition, participating firms should receive a payroll tax credit to offset some or all of their costs from paying for health care and other benefits. Since firms are usually aware of the tax penalty from a rise in UI participation due to experience rating, federal subsidies are likely to increase STC participation during downturns and reduce the cost from unemployment layoffs if these program terms are clearly communicated to firms.

3.c.3. Train

A structural challenge within the UI system is that dislocated unemployed workers are not able to apply benefits to worker training programs. Instead, unemployed workers who receive benefits must continuously search for full-time work.[4] The

4 An important exception to this rule is the TAA program, where earnings, subsidies, and training are typically combined. In addition, as of 2012, 16 states provided additional UI benefits to permanently laid off workers who require training to improve their skills (*e.g.*, National Employment Law Project 2012). For example, UI claimants who file for Extended Training Benefits by the 16th week of unemployment in California can obtain benefit extensions, but total benefits cannot exceed 52 weeks.

rationale is to avoid subsidizing investments in training that may take place in the absence of UI benefits and that would typically be funded by workers or their employers. Another concern is that training taking place during unemployment may be less effective than training occurring on the job. Training programs that tie workers to potential employers are often deemed more successful at improving labor market outcomes. However, in deeper downturns, an unintended consequence is that UI claimants are prevented from using a period of low job availability to invest in their skills.

The federal government should allow workers on STC to engage in training while their hours are reduced and while they are receiving partial unemployment benefits. Similarly, it should allow firms to establish training plans as part of STC that would aim to increase the skills of the workforce.[5] To minimize the risk of abuse, the implicit training subsidies provided by STC could be limited to downturns. Moreover, one could limit training to those employers expecting work sharing to last a certain minimum number of weeks. Nevertheless, the risk of abuse is relatively small compared to the potential benefits, since STC programs are currently small, and potentially large benefits would result from allowing workers to better use periods of slack work while also preventing layoffs. As STC programs grow, it will be important for the US DOL to pursue a formal, randomized evaluation of the STC as it has done with other programs.

4. Adjusting the UI system over the business cycle via automatic triggers

4.a. Need

It is widely recognized that the UI program should automatically adjust to labor market conditions, rather than relying on ad hoc action by US Congress and/or state legislation. Despite the presence of a trigger-based, state-level Extended Benefit program, removing discretionary action has proven difficult, partly because of a lack of agreement about appropriate automatic triggers. The COVID-19 crisis showed that automatic triggers are needed beyond increasing benefit durations in recessions. It is also necessary to adjust eligibility requirements and benefit levels over the business cycle as well. Improvements to and extensions of the current trigger-based system can be achieved by harnessing data generated from the UI system itself.

Current Extended Benefits programs provide additional weeks of UI benefits when the state's insured unemployment rate (the fraction of employees receiving UI benefits) or the state's unemployment rate as measured by the Bureau of Labor Statistics

---

5 A small number of states help defray employers' training costs as part of Back-to-Work programs (Kugler 2015).

(BLS) is above a certain threshold. While the trigger based on the unemployment rate usually provides extended benefits more often, not all states have adopted it. Moreover, the more common trigger for the Extended Benefit program based on the insured unemployment rate is flawed in that it only counts workers on regular UI benefits, not those receiving extended benefits. Extended Benefit programs have played a minor role in recent recessions, with the majority of benefit extensions enacted and paid for by US Congress.[6]

4.b. Proposal

(1) **Extend durations:** Use a measure of UI benefit exhaustion to design triggers for benefit extensions grounded in economic rationale.

(2) **Increase benefits:** To raise UI uptake and prevent hardship, UI benefits should be automatically increased during recessions.

(3) **Broaden access:** Eligibility criteria for UI plays an important role in determining UI access and should be relaxed during recessions to raise UI coverage and better assist claimants as they adjust to changing labor market conditions.

4.c. Details and discussion

4.c.1. Extend durations

Workers exhausting UI benefits have been found to be at heightened risk of poverty and long-term unemployment (Rothstein and Valletta 2017; Ganong and Noel 2019). Several reemployment programs of the UI system, such as the Worker Profiling and Reemployment Services (WPRS) and Reemployment Services and Eligibility Assessment (RESEA) programs, consider a UI claimant's risk of benefit exhaustion when determining eligibility for more intensive workforce services. The exhaustion rate among all UI claimants in a given state also reliably predicts the need for additional UI benefits in that state. In terms of economic theory, the incidence of benefit exhaustion is directly linked to a rise in the value of insurance provided by the UI system (Schmieder, von Wachter, and Bender 2012). Hence, estimates of the exhaustion rate should be used to trigger benefit extensions.

Current approaches rely on measures based on the insured unemployment rate or the overall unemployment rate. While these are common measures of labor market tightness and hence natural candidates, neither fully captures UI claimants' ability

6 In recessions, the federal government sometimes pays for half of state Extended Benefits or makes Emergency Unemployment Compensation available, also as a function of triggers based on the unemployment rate.

to find jobs. In particular, triggers based on the number of UI recipients have been flawed, since they do not count claimants receiving UI benefit through extension programs. Hence, as long-term unemployment increases in deep recessions and more individuals transition to extended UI benefits, the total number of regular UI recipients shrinks, which may trigger benefits to turn off prematurely (Bell et al. 2021c). This can be fixed by redefining the default triggers currently used by state extended benefit programs to better reflect the state of the labor market by also counting UI claimants receiving extended benefits.

However, the choice of the threshold for triggering changes in program benefits is arbitrary. Instead, devising triggers based on the exhaustion rate provides a more natural benchmark. For example, a sensible goal would be for the duration of UI benefits to be set such that the exhaustion rate during recessions is no higher than during expansions. Insofar as job losses are more costly in recessions, one could aim to keep the exhaustion rate during recessions lower than that in expansions.

4.c.2. Increase benefits

The primary focus of policy discussions about UI automatic triggers has been on the duration of benefits. In some cases, such as during the COVID-19 crisis, and to a lesser extent during the Great Recession, benefit levels have been increased as well. However, UI benefit increases during recessions should have a permanent place in the tool kit of state and federal policy for three reasons. The standard argument for raising UI benefits during recessions is that the insurance value of UI payments increases; this can arise because unemployment spells become longer on average, depleting workers' savings, or because the likelihood of job loss among members of the same household increases. In addition, the stimulus effects of UI payments for the economy are likely greater during recessions, since unemployed workers spend additional income. Since benefits apply to all UI claimants immediately, raising benefit levels can have a substantially higher stimulus effect than benefit extensions.

Finally, an important but typically overlooked argument is that increasing UI benefits raises the rate of UI receipt (Anderson and Meyer 2002). At the typical benefit rate of 50% of prior earnings, lower-income workers are likely unable to cover rent and necessities from UI benefits alone. At this level of income replacement, even middle-income workers may be compelled to skip UI for lower paying jobs since many may not have sufficient savings to supplement UI on their own. Raising the recipiency rate not only raises the insurance value of UI by reaching a larger number of workers, it's also likely to do so among more vulnerable workers. Higher and broader recipiency increases the value of the UI program as a platform through which workers can be connected to other economic support or workforce services as discussed in Section

2. Given this may speed reemployment, raise labor force attachment, or lead to better jobs, part of the cost of additional UI benefits may pay for itself.

To avoid ad hoc benefit changes that are subject to the political process, one approach is to tie benefit increases to the same triggers as those for benefit extensions. While the exhaustion rate is theoretically motivated and intuitive for triggering increases in benefit durations, no similar rule of thumb exists for triggering increases in benefit levels. Developing such triggers would be a valuable avenue for future research. Balancing the rate of UI receipt among earnings groups or linking benefit levels to the median duration of UI spells may be promising concepts to start with.

4.c.3. Broaden access

Automatically adjusting nonmonetary eligibility criteria during recessions should also be considered. Eligibility criteria specify the circumstances under which a worker can file for UI benefits, such as the reason for a job loss, the extent of job search required, and which type of jobs the worker must search for. For example, in many states previously full-time workers have to look for full-time work and are typically not able to engage in full- or part-time job training. Similarly, workers who quit their jobs because their spouse had to relocate or because they had to care for a family member are ineligible for benefits.

As layoffs increase and job finding rates decline during recessions, it is worth considering automatically relaxing certain benefit eligibility criteria. For example, it is likely that during recessions more workers need to relocate with their family or spouse for job opportunities, or that workers may need to take part-time jobs to make ends meet. Similarly, workers may be better served developing new skills than looking for work during deeper recessions. The American Recovery and Reinvestment Act (ARRA) provided financial incentives for states to adopt provisions relaxing eligibility along these margins (National Employment Law Project 2012).

Automatically adjusting UI eligibility criteria provides additional insurance coverage and thereby helps to raise the UI recipiency rate. As in the case of raising benefit levels or benefit durations, greater program uptake likely increases coverage among more marginalized workers. In turn, greater program uptake allows larger numbers of workers to be connected to additional services that may speed reemployment. A further advantage of visibly tying benefit criteria to the business cycle is that workers are aware of modified UI rules during recessions. This could be particularly valuable if workers can maintain UI benefits while pursuing re-training opportunities, which may help to prevent longer-term unemployment among dislocated workers.

5. Use UI data and research to enable data-driven policy

5.a. Need

The UI system has a wealth of untapped information that could be used to improve our understanding of the economy, the effectiveness of the UI program as a social insurance mechanism, and the administration of the UI program. This data can also be used to better target workforce services and identify opportunities to reduce program costs. One key advantage of UI data, relative to other economic data, is that it is available on a weekly or even daily basis. Broader and more informed use of UI data, along with improvements to data management systems would have an immediate benefit at the federal level and throughout the country.

5.b. Proposal

(1) Modernize: Modernize reporting requirements of states' UI systems to the US DOL to improve the ability to monitor the economy, to assess the functioning of the UI system, and to provide accurate information about the UI program to the public and policymakers, all in close to real time.

(2) Upgrade: Expand data collection during the administration of UI benefits to improve program administration and better target workforce services. Create a harmonized federal register of UI claims available for paying cross-states benefits and for evaluation and research purposes.

(3) Evaluate: Improve statistics generated and lessons learned about the UI system through evaluations and research by providing access to anonymized, individual-level UI claims and by fostering state and federal research partnerships with academic and other researchers.

5.c. Details and discussion

5.c.1. Modernize

Each state's UI program currently reports a set of statistics to the US DOL on a weekly basis, which in turn is made public. This includes the much-anticipated weekly release of data on initial and continuing UI claims, along with statistics such as the number of UI claimants first paid in a given week (so-called "first payments"). In addition, the US DOL publishes monthly statistics, such as the number of continuing claims by demographic characteristics or industry and provides information on state-specific UI rules and program administration, such as whether extended benefits are active.

First, the reporting system for UI data needs reform. The current system is cumbersome, even by the standard of other US government agencies, and leads to repeated misinterpretation of key statistics. For example, the total number of continuing claims is often reported without including workers who are receiving benefits through extended benefit programs, undercounting the number of individuals receiving UI. Many policymakers, researchers, and journalists are often not aware of which data is included in DOL reporting. For example, the amount of churn in the UI system can be approximated by the number of additional initial claims, but this data is seldom if at all used (Bell et al. 2020a). Moreover, some of the key data is not available in a machine-readable format, making reporting and analysis of the weekly news release difficult.

Second, additional statistics should be added to improve the value of UI data to policy makers and the public. Currently, relevant statistics are only partially provided, contain measurement errors, or are not provided at all. For example, one cannot currently calculate the number or fraction of new initial claims that were rejected. It is also impossible to calculate the number and fraction of UI claimants who exhaust their UI benefits. The number of initial claims often contains repeated claims, either because of duplicate claims, additional claims, or transitional claims (Cajner et al. 2020).

Another important statistic that can be misleading is the number of continuing claims. The number of continuing claims reported by the US DOL and most state UI offices corresponds to the number of weeks claimed by all UI recipients in a given calendar week. This coincides with the number of individuals receiving benefits for a week of unemployment during that same calendar week only if individuals certify for benefits on time (during the week of unemployment) and if they do not certify for multiple weeks. Backward certification, such as occurred frequently during the COVID-19 crisis, causes the standard continuing claims measure to be less informative. Using UI claims records, the CPL generated a measure of the number of individuals receiving benefits in a given calendar week that is robust to retroactive or delayed certifications (Bell et al. 2020b).

Part of these shortcomings can be addressed by expanding the statistics that states must provide to the US DOL. A partial list of such statistics currently not available that could be relatively easily calculated based on existing data is as follows:

- Report initial claims by type (new initial, additional, transitional), by program (regular, extended), and by demographic, industry, occupation, and county.
- Report the number of weekly unduplicated new initial and additional claims.

- Report the rate of benefit denial of new initial claims within a determined number of weeks (*e.g.*, within two, four, or six weeks).
- Report the number of continuing claims by the week of unemployment which the payment corresponds to instead of by the week of certification.
- Report the number of continuing claims by week of unemployment by demographics, by industry, by occupation, and by county.[7]
- Provide harmonized tabulations on race and ethnicity (see also "Upgrade" below).
- For each state, report the number of intra-state claims by state.

An expansion in the number of available statistics would be complementary and aided by improvements to the underlying individual data, discussed next.

5.c.2. *Upgrade*

Integrate. The data generated from administering the UI program is owned by the states and comprises three core data sets: (1) quarterly earnings records of total wages paid that a worker received from each employer (the so-called UI Base-Wage file); (2) quarterly employer records containing total earnings and total employment for each establishment, among others (the so-called Quarterly Census of Employment and Wages, QCEW); and (3) information on which individuals filed for and received UI benefit payments (in the initial claims and continuing claims files, respectively). While the states' UI Base-Wage files and the QCEW data have been integrated into common federal data registers, there is no single data register that contains all US UI claims information and that can be accessed for purposes of program administration, for program evaluation, or for statistical purposes. As a result, this data is rarely used for research on the UI system, and only occasionally accessible for evaluations of the UI program sponsored by the US DOL. The data is also rarely shared between states, limiting opportunities for improving program administration.

There is a history of collecting information generated at the state level in the US federal data system. Perhaps the most well-known example is state individual mortality records based on death certificates, which are sent to the Centers for Disease Control and Prevention (CDC). The CDC sets common standards, harmonizes the data, publishes aggregated statistics, and makes appropriately de-identified,

7 Since not all individuals certifying for a week of unemployment are actually paid (*e.g.*, some will not receive benefits because of excess earnings) these numbers should be further broken down into the number of individuals actually paid for a given week of unemployment.

individual-level mortality records available for research through a standard process. For the UI system, similar data collection, harmonization, and dissemination processes occur for the QCEW data (through the BLS) and the UI Base-Wage file (through the Census Bureau), though with some important differences from the mortality data as noted below. However, no such aggregated U.S.-wide data exists for the UI claims data.

There is a database that states can use for looking up individual claims and in some cases for exchanging data that could in principle be used to create a harmonized federal data register.[8] A federal government agency would have to be designated to regularly receive standardized extracts of the data from each state, to work with the states to agree on data extraction standards, and to produce a data set that is appropriate for research. The same agency would develop a data dissemination mechanism in cooperation with states. This agency could also be involved in generating statistical information from the program, but this could be taken on by another agency as well. For example, either the BLS or the Census Bureau could be the host agency, while the US DOL—with an appropriately staffed research department—could generate statistics.

A harmonized federal data system would yield immense benefits to state UI agencies, the US DOL, and federal and state policymakers. Among others, it would allow states to better assess reemployment outcomes from their programs, yield substantial improvements in terms of measuring key metrics of the UI system, provide better real-time information on the state of the labor market, and allow for improved and more comprehensive evaluations as a foundation for better and data-driven management of UI claims and implementation of evidence-based reforms.

There are two key steps in achieving such a harmonized data set. The first is a legal step, as currently each state would have to agree to share its data with the federal government. This has worked well for the QCEW, but the individual-level data is not available for research outside of the premises of the BLS. When the data is offered to the research community, the process has shown its shortcomings for the case of the UI Base-Wage file, where only a subset of states agreed to use its data for research, while a substantial number of states chose to review every research project. The data is also available with a three-year lag, limiting its usefulness for generating statistics about the economy. Hence, regulations by US Congress are likely needed to establish a functional, national integrated database of UI administrative records that can support program administration, statistics, and evaluation and research.

8 The database is called Unemployment Insurance Interstate Connection Network (ICON) and allows state UI agencies to request and receive data for use in the filing and processing of combined wage claims, military, and federal claims. The system provides for the exchange of data between state workforce agencies as well as federal partners.

A second, more technical step is to ensure the administrative data is processed appropriately to enable routine use. Since the UI data is based on spells of benefit receipt that evolve over time, and captures a large range of administrative actions, care has to be taken into how to structure an integrated database with the ability to support generating statistics or research and evaluation. A useful blueprint for such processing and dissemination is the UI data provided by the German government through the Institute for Employment Research, the research institute of the federal labor agency administering the UI program (Bender, Haas, and Klose 2000), which is now routinely used by US researchers to study unemployment and UI. Without such a defined structure, a simple collection of ad hoc extracts of UI claims data from different states is likely to be difficult to use and to miss important information.

Upgrade. The data currently available on UI claims is generated as part of the process of administering the program, and hence is not designed to be used for statistical analysis or program evaluation. As is typical with administrative program data, without further processing, individual-level data can be difficult to use and may not generate information that is useful in contexts other than correctly paying UI benefits. The underlying data is generated by individual claimants, by case workers, and by automated processes within a case management system.

This leads to two related issues. The first is that the data that is used for providing statistics or for evaluation has to be extracted from the case management system. Since there are a large number of potential administrative actions recorded in the system, typically only a subset of the data is extracted. Since the case management system is not designed to generate useful data for statistics or evaluation, the extracts may miss relevant information or may not be in a format that is conducive to learning about UI claims.

The case management data is typically extracted into two files. The initial claims file contains information on demographics and basic job background provided by the claimant when a new claim is filed, plus some information from the adjudication of the claim (*e.g.*, whether the claimant has sufficient earnings history to qualify for benefits). The continuing claims file contains information provided by the claimant during weekly (or bi-weekly) benefit certification, plus partial information from the adjudication of the claim (*e.g.*, whether a payment was denied due to excess earnings). The separate extraction can make it difficult to connect events in the initial claims file to actions recorded in the continuing claims file. That in turn makes it hard to reconstruct the various sequences of events that can occur while a claimant is receiving UI.[9]

9 For example, if after an initial flag that earnings levels are too low to qualify for benefits (shown in the initial claims file) an individual does not receive benefits (no record in the continuing claims file), it is not clear whether that individual appealed and got denied or whether that individual found a job. Alternatively, if an individual opens an additional claim after being laid off again, it can be difficult to associate the additional initial claim to the actual payment in the continuing claims file (or to the reason why payment was denied).

The second, related issue is that by its nature the case management system only records information that is needed for managing the UI claim and may not generate important or useful information needed for statistical or evaluation purposes. For example, the system generates an indicator for a "last payment" if a claimant has received the last payment for which she is eligible for a given program. However, these indicators by themselves cannot be used to construct an indicator for whether a claimant exhausted benefit eligibility across all programs, a crucial statistic for assessing the need for UI benefits (see Section 4). This is because after a last payment of regular benefits, during recessions the claimant might be eligible for a first extended benefit program; after which they might be eligible for additional extended benefits, and so on. This could in principle be resolved by combining the claimant's payment history with her prior earnings (an important factor in determining eligibility). However, because it is not needed for paying benefits, there is no single claim ID that can be used to string together payments received during the same claim.[10]

To be more useful for purposes of program administration and research, UI claims data should be updated to include variables that can be used to describe the evolution of UI claim from initial claiming to last payment. Such variables would include, among others:

- Add timing of different adjudication steps to the initial claims file.
- Generate a claim ID that is not affected by BYB changes (and hence can be used to connect continuous spells of benefit receipt that overlap the end of the BYB and receive a new BYB, but that really belong to the same period unemployment).
- Generate event IDs that link initial claims events (*e.g.*, an additional claim) and continuing claims actions (*e.g.*, a payment or denial).
- Add a system generated indicator measuring benefit exhaustion.

These additions would have to be harmonized across states' UI systems, so any modification of the processes of paying benefits would ideally occur following a coordination process led by a federal agency charged with improving, harmonizing, and disseminating UI data.

10 While this would be helpful for statistical purposes, it is not needed for paying benefits. The reason is that in general a claim is indexed to the first date benefits are paid, the so-called BYB (Benefit Year Beginning). However, due to the fact that the program requires a resetting of benefits after a year on the program, in longer recessions some claimants can experience changes in the BYB, making it difficult to construct full benefit histories for individuals.

5.c.3. Evaluate

Providing access to individual-level administrative data can be extremely valuable for research and evaluation purposes. In fact, many evaluations contracted by the US DOL, which are commonly viewed as the "gold standard" due to its randomized research design and data, are based on administrative UI claims data. However, this data is hard to access for researchers or for evaluation purposes outside of the context of these relatively infrequent and expensive federal evaluations. Hence, a mechanism for accessing processed and harmonized UI claims data at the state and federal level for research and evaluation purposes would make a tremendous difference for research on the UI program, the economy, and other federal programs.

However, frequent use of administrative data can also have direct returns for participating agencies, above and beyond specific evaluation and research findings. A more informed use of UI claims data can improve program administration or statistical analysis of local labor market conditions. However, as is typical for many government agencies, many UI agencies do not have the capacity to flexibly use UI data. In California, the CPL, a joint research institute at University of California–Berkeley and University of California–Los Angeles, has had an ongoing research partnership with EDD, the agency administering UI benefits. Such research partnerships can deliver value to participating agencies in numerous ways. For example, active use of the data helps to clarify potential measurement questions. Curated UI claims files can be easily used to generate dashboards or satisfy customers requests. Improved data can be used to better target workforce services.

For example, in California, CPL had the following research output, among others, during the COVID-19 crisis:

- Measurement of continuing claims, initial claims, and timing of churn in and out of the UI system.
- Provision of research files on initial claims by demographics, industry, and county used for a new dashboard.
- Measurement of the rate of UI exit, the frequency of reemployment, and the rate of benefit exhaustion.
- Analysis of UI recipiency rate and its differences and correlates across areas in the state.
- Analysis of the incidence of long-term unemployment and its differences and correlates across areas of the state.

In addition, better UI data can also be used to institute certain data-driven approaches to case management that can streamline operations and save costs. For example, a data-driven approach can be used to flag cases that could be fast tracked for approval, or support case workers' decision processes. Algorithms could also be used to flag which companies' data is updated frequently, and hence allow making certain wage records available earlier to obtain a more real time tracking of reemployment services.

The UI data is sufficiently complex that without a broader national user base and a user base in each state, it is unlikely that UI agencies will draw the full benefit for program administration. States will be in a better position to use the UI system to advance policy goals if insights are available from research based on UI data. This is why updating UI data systems is not just an academic exercise. It is a necessary input to being able to make policy improvements and will yield important insights that are valuable for all states. In addition, local research partnerships can help states improve their data infrastructure, support local uses of the data, and help build capacity.

6. Conclusion

While the US labor market is on track for recovery, important challenges and risks remain, not least from potential variants of the virus and lagging vaccination rates. The proposals outlined in this report not only have the potential to have immediate impact, but also to increase the resilience of the US social insurance and workforce system. This is because they not only seek to address flaws in the UI system—which typically receive the most attention—but strengthen and expand procedures and programs that are aimed at reintegrating workers in the labor market, but that currently do not sufficiently respond to changing economic conditions. Moreover, by focusing on reemployment, several of the proposals discussed here sidestep the passionate and recurring political debate as to the role of unemployment insurance benefits in prolonging the recession. In particular short-time compensation, but also the more effective and expanded use of workforce services and improvement of UI data infrastructure and statistics, are areas with potential for bipartisan support. The proposals also provide important concepts and infrastructure that can be applied to other programs discussed during the crisis, such as targeting and evaluating reemployment bonuses (O'Leary, Decker, and Wandner 2005).

References

Anderson, Patricia and Bruce Meyer. 1997. “Unemployment Insurance Takeup Rates and the After-Tax Value of Benefits.” *Quarterly Journal of Economics* 112 (3): 913¬–937

Barrero, Jose M., Nicholas Bloom, and Steven J. Davis. 2020. “COVID-19 is also a reallocation shock.” *Brookings Papers on Economic Activity,* Summer 2020 special edition.

Bell Alex, Thomas J. Hedin, Roozbeh Moghadam, Geoffrey Schnorr, and Till von Wachter. (2020a). “August 6th Analysis of Unemployment Insurance Claims in California During the COVID-19 Pandemic.” California Policy Lab Policy Brief.

¬——. 2020b. “July 2nd Analysis of Unemployment Insurance Claims in California During the COVID-19 Pandemic.” California Policy Lab Policy Brief.

——. 2021a. “March 18th Analysis of Unemployment Insurance Claims in California During the COVID-19 Pandemic.” California Policy Lab Policy Brief.

——. 2021b. “March 18th Analysis of Unemployment Insurance Claims in California During the COVID-19 Pandemic.” California Policy Lab Policy Brief.

Bell Alex, Thomas J. Hedin, Geoffrey Schnorr, and Till von Wachter. 2021c. “Why Extended UI Benefits were Turned Off Prematurely for Workers in 33 States.” California Policy Lab Policy Brief.

Bender, Stefan, Anette Haas, and Christoph Klose. 2000. “IAB Employment Subsample 1975-1995 Opportunities for Analysis Provided by the Anonymised Subsample.” IZA DP No. 117.

Bertrand, Marianne, Harold Pollack, Till Von Wachter, Janey Rountree, and Brian Blackwell. 2019. “Predicting and Preventing Homelessness in Los Angeles.” California Policy Lab Policy Brief.

Bhanot, Syon P. and Elizabeth Linos. 2020. “Behavioral Public Administration: Past, Present, and Future.” *Public Administration Review* 80 (1): 168–171.

BLS. 2021. “Of Total Unemployed, Percent Unemployed 27 Weeks & Over.” Retrieved from FRED, Federal Reserve Bank of St. Louis, July 4 2021. https://fred.stlouisfed.org/series/LNS13025703

Cajner, Tomaz, Andrew Figura, Brendan M. Price, David Ratner, and Alison Weingarden. 2020. “Reconciling Unemployment Claims with Job Losses in the First Months of the COVID-19 Crisis.” Finance and Economics Discussion Series 2020-055. Divisions of Research & Statistics and Monetary Affairs Federal Reserve Board, Washington, D.C.

Card, David, Jochen Kluve, and Andrea Weber. 2018. "What Works? A Meta Analysis of Recent Active Labor Market Program Evaluations." *Journal of the European Economic Association,* 16 (3): 894–931.

Davis, Steven J. and Till von Wachter. 2011. “Recessions and the costs of job loss.” *Brookings Papers on Economic Activity* (2).

Ganong, Peter and Pascal Noel (2019). “Consumer Spending during Unemployment: Positive and Normative Implications.” American Economic Review 109: 2383–2424.

Hassett, Kevin A. and Michael R. Strain. 2014. “Worksharing and Long-Term Unemployment.” Center for Budget and Policy Priorities.

Houseman, Susan, Frank Bennici, Susan Labin, Christopher O’Leary, Katharine Abraham, and Richard Sigman. 2017. “Modest outreach increases employer awareness and use of short-time compensation (STC).” Study Brief, Upjohn Institute.

Kugler, Adriana 2015. “Strengthening Reemployment in the Unemployment Insurance System.” Policy Proposal, Hamilton Project.

Knaus, Michael C., Michael Lechner, and Anthony Strittmatter. 2020. “Heterogeneous Employment Effects of Job Search Programmes: A Machine Learning Approach.” *Journal Human Resources,* published ahead of print March 26, 2020,

Linos, Elizabeth, Allen Prohofsky, Aparna Ramesh, Jesse Rothstein, and Matt Unrath. 2020. “Can Nudges Increase Take-up of the EITC? Evidence from Multiple Field Experiments.” California Policy Lab Working Paper No. 2020-1

O’Leary, Christopher J., Paul T. Decker, and Stephen A Wandner. 2005. “Cost-effectiveness of targeted reemployment bonuses.” *Journal of Human Resources* 40: 270–279.

National Employment Law Project 2012. “Modernizing Unemployment Insurance: Federal Incentives Pave the Way for State Reforms.” Briefing Paper May 2012.

Rothstein, Jesse and Robert G. Valletta. 2017. "Scraping by: Income and Program Participation After the Loss of Extended Unemployment Benefits," *Journal of Policy Analysis and Management* 36 (4): 880–908.

Schmieder, Johannes, Till von Wachter, and Joerg Heining. 2020. “The costs of job displacement over the business cycle and its sources: evidence from Germany.” Working Paper, UCLA.

Schmieder, Johannes, Till von Wachter, and Stefan Bender. 2012. “The Effects of Extended Unemployment Insurance Over the Business Cycle: Evidence from Regression Discontinuity Estimates Over 20 Years.” *Quarterly Journal of Economics* 127 (2): 701–752.

Sullivan, Daniel and Till von Wachter. 2009. “Job Displacement and Mortality: An Analysis Using Administrative Data.” *Quarterly Journal of Economics* 124 (3): 1265–1306.

Thaler, Richard H. and Cass R. Sunstein. 2008. *Nudge: Improving decisions about health, wealth, and happiness.* Yale University Press, New Haven, CT

von Wachter, Till. 2020. “Lost Generations: Long-Term Effects of the COVID-19 Crisis on Job Losers and Labour Market Entrants, and Policy Options.” *Fiscal Studies,* Special Issue.

——. 2021. “An Overview of Current Results and New Methods for Estimating Heterogeneous Program Impacts.” Mimeo UCLA, Prepared for the Social Security Administration.

von Wachter, Till and Steve A. Wandner. 2020. “Proposal for a U.S. national emergency short-time compensation programme during the COVID-19 crisis.” http://www.econ.ucla.edu/tvwachter/covid19/National%20Emergency%20STC%20and%20Covid-19.pdf.

Addressing Inequities in the US K-12 Education System

AUTHORS

Nora Gordon* and Sarah Reber**

ABSTRACT

American public schools do not successfully prepare all students for careers or college. Significant disparities in educational outcomes by race, ethnicity, economic disadvantage, and disability are persistent. At the same time, outcomes for schools serving similar student populations vary considerably. Making equitable progress will require shoring up fundamentals throughout the K-12 system: how staff are trained, recruited and retained, and supported in their work; curriculum; safe and healthy school buildings; and how these components are organized and used in schools. This would mark a shift away from an emphasis on "silver bullet" interventions to improve schools. Some schools and districts will require additional resources and supports. These efforts are complicated due to the federalist landscape of elementary and secondary education, where states, local school districts, and the federal government all play important roles. Quick fixes are few and far between, but improvements to school infrastructure stand out as low-hanging fruit. We should learn from past efforts to improve the impact of educational policy and philanthropy going forward, with careful attention to strengthening the research base.

* McCourt School of Public Policy, Georgetown University and National Bureau of Economic Research.
Email: nora.gordon@georgetown.edu
** UCLA Luskin School of Public Affairs, National Bureau of Economic Research, and Brookings Institution.
Email: sreber@ucla.edu

Acknowledgements: We are grateful to Eric Brunner, Carrie Conaway, Thomas Dee, Amy Ganz, Dan Goldhaber, Douglas Harris, Melissa Kearney, and Susanna Loeb for helpful comments, and to Emily Pelles for excellent research assistance. Any errors are our own.

1. Introduction

Despite decades of efforts by federal and state policymakers and major philanthropic and research investments, America's schools have so far fallen short of meeting the ideal of preparing all students to succeed in college or in a career. Racial disparities in outcomes are large and persistent. The ongoing public discussion about how to address these gaps, and how to improve schools overall, can feel like a rotating mash-up of buzzwords. Many of the buzzwords stand in for good ideas, but they are too often presented in isolation as silver bullets. The 2021 list would include things like socio-emotional learning, science-based reading, high-dosage tutoring, restorative justice, formative assessment, quality curriculum, and universal pre-K. Yet time and again silver bullets *du jour* come up short, with proponents blaming "implementation problems" for the failure to deliver. There are no easy explanations for why this pattern persists. To be sure, there are many under-used policies and programs that we should work harder to implement well. But we will also need to acknowledge the limitations of one-off, top-down prescriptions. American public schools operate within complex systems, and we will not be able to "fix" schools without ensuring that the fundamentals are sound. Further, a well-functioning system will necessarily leave some discretion in the hands of district, school, and classroom leaders, highlighting the importance of staffing and leadership at every level.

In this chapter, we argue that to reduce inequities in American education, we must improve systems in ways that benefit all students and schools. While schools involve many moving parts that could each be improved, reforms to these individual parts are too often evaluated and advocated for in isolation despite functioning in inherently interdependent ways. At the same time, attending to the fundamentals is a necessary, but not sufficient, condition to improve schools for students who have been poorly served—or discriminated against—both in and out of school, including students of color, students living in poverty, and students with special instructional needs. With improved fundamentals, seemingly logical but poorly implemented silver bullets from the past could prove more effective. Limited resources and political capital pit program against program in federal, state, and local budget cycles. But because each individual program will work better in the context of stronger systems with high-quality staff, these choices should not be viewed as zero-sum trade-offs. Ideally, we could improve outcomes for all students and reduce disparities at the same time. At times, however, changes that will help disadvantaged and lower-achieving groups the most may help other groups even more, increasing gaps while maximizing the absolute level of improvements for disadvantaged students.

We provide an overview of research on the fundamental inputs to schooling—staff, peer groups, curriculum, and physical infrastructure—as well as how these inputs are organized and combined in schools and classrooms. We also draw on research on attempts to "intervene" through add-on programs, which sometimes succeed according to rigorous evaluations but are less frequently successfully scaled, as well as "school turnaround" efforts targeting entire school buildings. Finally, we discuss reforms, such as test-based accountability, designed to improve student learning by changing governance and incentives at a higher level.

Even significantly improved schools will not, on their own, overcome the powerful influences of structural inequality and racism across American society, which contribute to persistent differences in educational outcomes across groups. A history of discrimination in housing, labor markets, policing, and other aspects of society means that low-income children and children of color often grow up in neighborhoods separate from their more-advantaged peers and are more likely to be exposed to a wide range of stressors that make succeeding in school difficult. Lead, air pollution, and other environmental hazards, violence in the community, police violence, economic despair and rising inequality, and adverse childhood events create obstacles to learning (Aizer et al. 2018; Currie and Schmieder 2008; Hardy, Logan, and Parman 2018; Nelson et al. 2020; Ang 2020; Kearney and Levine 2016; Sharkey et al. 2014). Policies that reduce child poverty and improve children's out-of-school environments may in some cases do more to improve learning and reduce inequality than school-based policies. While schools alone cannot solve these problems, they are a key policy lever for promoting economic mobility, equality, and well-being.

We first document patterns and trends in educational achievement and attainment, overall and by race and ethnicity. We then present a short primer on American public schools, which operate in complex institutional systems that vary considerably from state to state. We then turn to the policy levers available for improving educational equity; we see improvements to fundamental aspects of schooling, such as staffing and curriculum, as key. We conclude with a discussion of important lessons from research and suggestions for where policymakers and advocates should focus attention going forward.

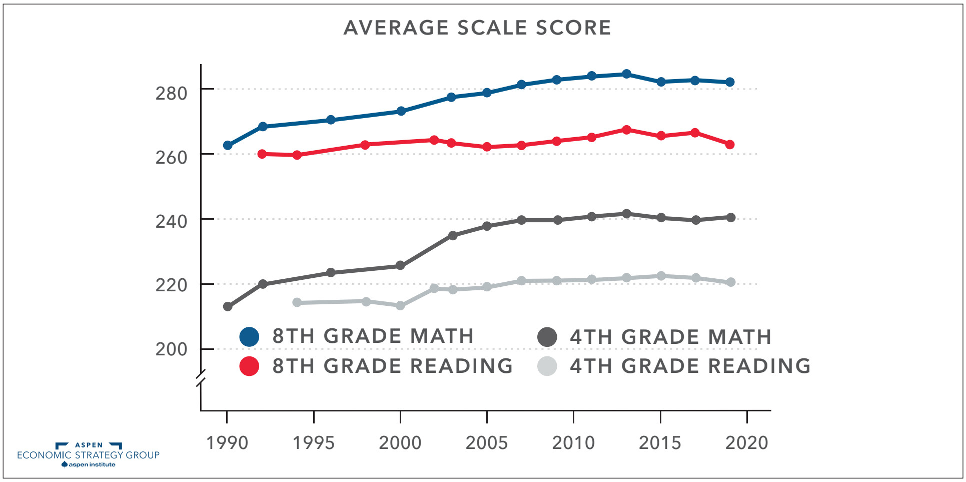
2. Patterns and trends in educational achievement and attainment

One thing is certain: America's schools do not consistently prepare all students for college and for careers. We present data on key outputs of the education system—academic achievement, as measured by test scores, and educational attainment—which we discuss with important caveats in mind. First, schools do many things, and these are not the only outcomes that matter. Second, although we focus here on differences by race, understanding how schools are working—and how outcomes vary—depending on other characteristics, such as English learner or disability status, is critically important. These categories are not consistently defined over time or across places, so data for these groups are often difficult to interpret or unavailable.

The No Child Left Behind Act required schools to report test scores by student "subgroups" including by race and ethnicity,[1] to shine a light on inequality and prevent systems from masking unacceptable outcomes for some groups behind school-level averages. It is in this spirit that we present data disaggregated by race and ethnicity in this section. We acknowledge that a focus on persistent disparities can foster a "deficit mentality," undermining the view that all students can succeed (Bertrand and Marsh 2021). To be clear, these patterns should not let schools off the hook. Rather, they point to the need for both better schools and robust social policies to support historically marginalized families and those struggling economically.

Figure 1 shows trends in math and reading scores on the National Assessment of Educational Progress (NAEP), known as "America's Report Card," for 4th and 8th graders. Math scores improved substantially between 1990 and 2005 or so, especially for 4th graders. But progress has since stalled, and while 4th graders' reading skills improved in the 2000s, scores have since plateaued, and 8th graders' reading skills have barely improved since the mid-1990s.

Figure 1: Progress in math has stalled for a decade and reading achievement is flat

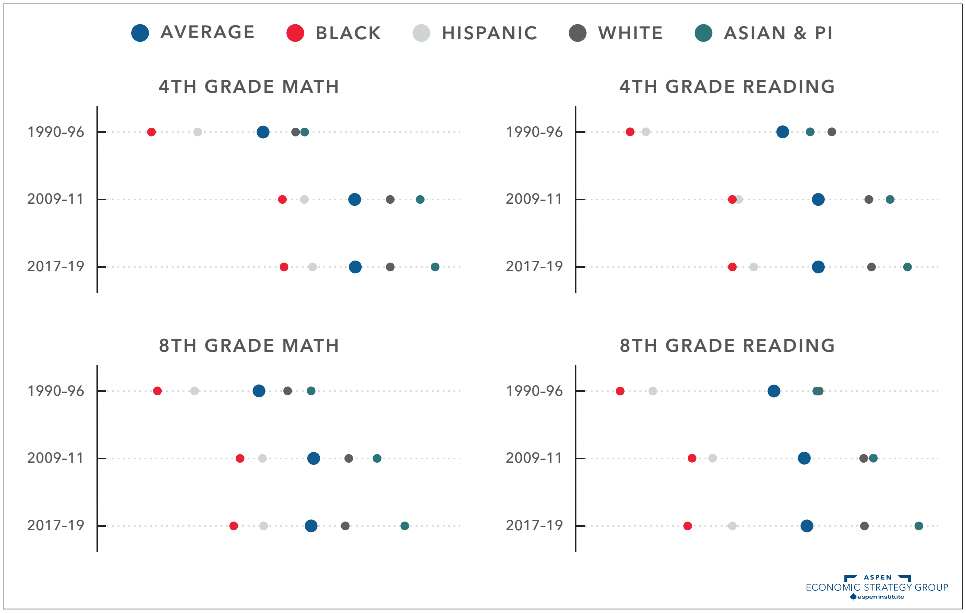


Source: National Assessment of Educational Progress (NAEP).

1 Office of Management and Budget (OMB) guidelines for classification of federal data by race and ethnicity specify that Hispanic or Latino origin (ethnicity) be collected in a separate question from race, and most of the data cited in this chapter follow that practice (Office of Management and Budget 1997). Unless otherwise noted, we categorize people of any race who identify as Hispanic or Latino as "Hispanic," and the other racial groups exclude people who identify as Hispanic or Latino; we exclude the "non-Hispanic" modifier for ease of exposition.

Differences in outcomes across racial groups are persistent (Figure 2). Scores for all groups improved between the 1990s and early 2010s, and gaps narrowed.[2] Unfortunately, little has changed since then, and Black 8th graders have even lost some ground. The differences across groups are large—for example, the difference in the average score between the highest- and lowest-scoring groups is equivalent to about three grade levels. Of course, scores vary considerably across students, and there are low- and high-scoring students in all groups. Research also shows that students scoring toward the bottom of the distribution are falling further behind, suggesting a growing share of students are not learning basic skills.

Figure 2: Test scores improved for all groups since the 1990s but progress has stalled and gaps are persistent

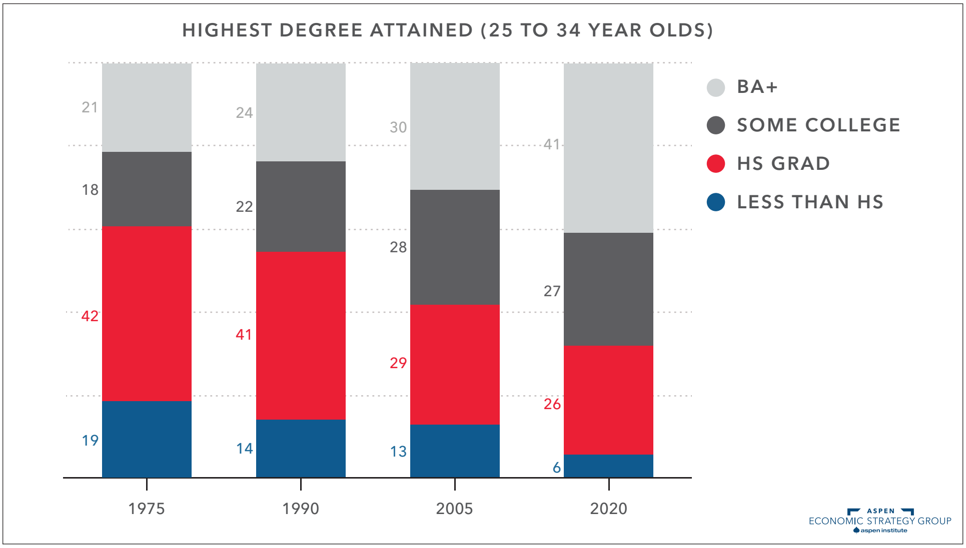


Source: NAEP.

Figure 3 shows trends in educational attainment of 25- to 34-year-olds, another important outcome of the education system. Since 1970, educational attainment has increased substantially: Today almost 40% of young people have a four-year college degree, and less than 10% are high school dropouts.

2 Since 2003, the NAEP reading and math assessments were given to 4th and 8th graders consistently every other year, so the 2009-11 data point is the average of 2009 and 2011, and the 2017-19 data point is the average of 2017 and 2019. The years used for the 1990-96 average vary based on data availability: the math assessment was given in 1990, 1992, and 1996 for both grades; the reading assessment was given in 1992 and 1994 for 8th grade and 1994 for 4th grade.

Figure 3: Educational attainment has increased over time

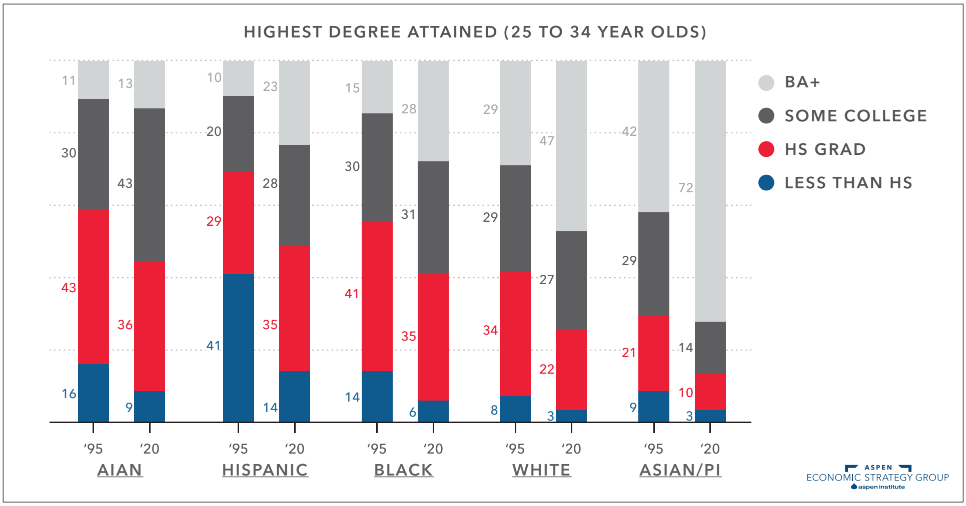


Source: Current Population Survey

While educational attainment has improved for all racial groups since 1995, there are substantial differences across groups (Figure 4). By most measures, educational attainment is lowest among American Indian/Alaska Native people, a group that has faced significant discrimination and neglect for centuries, though this group's disadvantage is frequently overlooked, in part because data are scarce. The COVID-19 pandemic has shone a light on these disparities, and early research suggests it has likely widened them further (Parolin and Lee 2021; Bacher-Hicks, Goodman, and Mulhern 2020) .

These documented gaps reflecting systematic inequalities are large, persistent, and deeply troubling. Less known is the substantial variance in school quality and outcomes even among schools serving students with similar backgrounds (Reardon 2019). That is, it is not just "troubled," high-poverty schools, or urban schools, or schools that serve mostly students of color, that are falling short; *many* schools serving students across the socioeconomic spectrum are not living up to their potential. Underperforming schools may be one explanation for the large increase in the use of private supplemental tutoring: Kim, Goodman, and West (2021) document a rough tripling of this sector from 1997 to 2016. These centers are concentrated in areas with higher parental income and education. We are unlikely to achieve an equitable distribution of opportunity or a more productive workforce without improving the quality of schools across the board.

Figure 4: Educational attainment has increased for all groups but significant gaps persist



Source: Current Population Survey

3. A primer on elementary and secondary education in the United States

Here we review the basics of the American elementary and secondary education system: Who does what and how do we pay for it? The answers to both questions vary considerably across states.[3]

3.a. Who does what?

Schools are the institution most visibly and directly responsible for educating students. But many other actors and institutions affect what goes on in schools. Three separate levels of government—local school districts, state governments, and the federal government—are involved in the provision of public education. In addition, non-governmental actors, including teachers' unions, parent groups, and philanthropists play important roles.

3.a.1. Schools

The vast majority of 5- to 17-year-old children attend public schools. (Expanding universal schooling to include up to two years of preschool is an active area of discussion which could have far-reaching implications, but we focus on grades

3 For state-specific information, consult state agency websites (e.g., Maryland State Department of Education) for more details. You can find data for all 50 states at the U.S. Department of Education's National Center for Education Statistics <http://nces.ed.gov>, and information on state-specific policies at the Education Commission of the States <ecs.org>.

K-12 here.) About 10% attend private schools; about a quarter of private school students are in non-sectarian schools, and the remaining three-quarters are about evenly split between Catholic and other religious schools. About 3% of students are homeschooled. Magnet schools are operated by local school districts but enroll students from across the district; magnet schools often have special curricula—for example a focus on science or arts—and were sometimes designed specifically to encourage racial integration. Charter schools are publicly funded and operate subject to state regulations; private school regulations and homeschooling requirements are governed by state law and vary across states. Nationally, 6.5% of public school students are enrolled in charter schools; the remainder attend "traditional public schools," where students are mostly assigned to schools based on their home address and the boundaries school districts draw. Washington, D.C. and Arizona have the highest rates of charter enrollment, with 45 and 18% of their public school students attending charter schools. Several states have little or no charter school enrollment. Prior to the COVID-19 pandemic, nearly all public schooling took place in person, with about 0.5% of students enrolled in virtual schools.

3.a.2. Local School Districts

Over 13,000 local education agencies (LEAs), also known as school districts, are responsible for running traditional public schools. The range of responses to the COVID-19 pandemic, with districts opting for fully in-person instruction, fully remote instruction, and everything in between, has highlighted the highly localized nature of school governance. The size and structure of local school districts, as well as the powers they have and how they operate, depend on the state. Some states have hundreds of districts, and others have dozens. District size is mostly historically determined rather than a reflection of current policy choices. But while districts can rarely "choose" to get smaller or larger, district size implicates important trade-offs (Andrews, Duncombe, and Yinger 2002; Gronberg et al. 2015). Having many school districts operating in a metropolitan area can enhance incentives for school and district administrators to run schools consistent with the preferences of residents, who can vote out leaders or vote with their feet by leaving the district. On the other hand, fragmentation can lead to more segregation by race and income and less equity in funding, though state laws governing how local districts raise revenue may address the funding issues. Larger districts can benefit from economies of scale as the fixed costs of operating a district are spread over more students and they are better able to operate special programs, but large districts can also be difficult to manage. And even though large districts have the potential to pool resources between more- and less-affluent areas, equity challenges persist as staffing patterns lead to different levels of spending at schools within the same district.

School boards can be elected or appointed, and they generally are responsible for hiring the chief school district administrator, the superintendent. In large districts, superintendent turnover is often cited as a barrier to sustained progress on long-term plans, though the causation may run in the other direction: Making progress is difficult, and frustration with reform efforts leads to frequent superintendent departures. School districts take in revenue from local, state, and federal sources, and allocate resources—primarily staff—to schools. The bureaucrats in district "central offices" oversee administrative functions including human resources, curriculum and instruction, and compliance with state and federal requirements. The extent to which districts devolve authority over instructional and organizational decisions to the school level varies both across and within states.

3.a.3. State Governments

The U.S. Constitution reserves power over education for the states. States have delegated authority to finance and run schools to local school districts but remain in charge when it comes to elementary and secondary education. State constitutions contain their own—again, varying—language about the right to education, which has given rise to litigation over the level and distribution of school funding in nearly all states over the past half century. States play a major role in school finance, both by sending aid to local school districts and by determining how local districts are allowed to tax and spend, as discussed further below.

State legislatures and state education agencies also influence education through mechanisms outside the school finance system. For example, states may set requirements for teacher certification and high school graduation, regulate or administer retirement systems, determine the ages of compulsory schooling, decide how charter schools will (or won't) be established and regulated, set home-schooling requirements, establish curricular standards or approve specific instructional materials, choose standardized tests and proficiency standards, set systems for school accountability (subject to federal law), and create (or not) education tax credits or vouchers to direct public funds to private schools. Whether and how states approach these issues—and which functions they delegate to local school districts—varies considerably.

3.a.4. Federal Government

The authority of the federal government to direct schools to take specific actions is weak. Federal laws protect access to education for specific groups of students, including students with disabilities and English language learners. Title IX prohibits sex discrimination in education, and the Civil Rights Act prohibits discrimination on the basis of race. The U.S. Department of Education issues regulations and guidance on K-12 laws and oversees grant distribution and compliance (Gordon and Pasachoff 2018). It also collects and shares data and funds research. The Bureau of Indian Education is housed in the Department of the Interior, not the Department of Education.

The federal government influences elementary and secondary education primarily by providing funding. Federal aid is typically allocated according to formulas targeting particular populations. The largest formula-aid federal programs are Title I of the Elementary and Secondary Education Act (ESEA), which provides districts funds to support educational opportunity, and the Individuals with Disabilities Education Act (IDEA), for special education. Both allocate funding in part based on child poverty rates. State and school district fiscal personnel ensure that districts comply with rules governing how federal funds can be spent and therefore have direct influence on school environments.

Since 1965, in addition to specifying how federal funds can be spent, Congress has required states and districts to adopt other policies as a condition of Title I receipt. The policies have changed over time, but most notably include requiring school districts to desegregate, requiring states to adopt test-based accountability systems, and requiring the use of "evidence-based" approaches. IDEA establishes protections for students with disabilities in addition to providing funding. The law guarantees their right to a free and appropriate public education in the least restrictive setting and sets out requirements for the use of Individualized Educational Programs. Because of these guarantees, IDEA allows students and families to pursue litigation.

Federal law prohibits conditioning funding on the use of any specific curriculum. The Obama Administration's Race to the Top program was also designed to promote specific policy changes—many related to teacher policy—but through a competitive model under which only select states or districts "won" the funds. For the major formula funds, like Title I and IDEA, the assumption (nearly always true) is that states and districts will adopt the policies required to receive federal aid and all will receive funds; in some cases, those policy changes may have more impact than the money itself (Cascio and Reber 2013).

The federal government has also allocated significant funding to support schools during the Great Recession and during the COVID-19 pandemic through specially created fiscal stabilization or relief funds; federal funding for schools during the COVID crisis is significantly larger than during the Great Recession.

The federal tax code, while perhaps more visible in its influence on higher education, also serves as a K-12 policy lever. The controversial state and local tax (SALT) deduction, now limited to $10,000, reduces federal tax collections and subsidizes progressive taxation for state and local spending, including for education. As of 2018, 529 plans, which historically allowed tax-preferred savings only for higher education expenses, can also be used for private K-12 expenses.

3.a.5. Non-Governmental Actors

Notable non-governmental actors in elementary and secondary education include teachers' unions and schools of education, along with philanthropists, vendors, and other advocates. The nation's 3.5 million public school teachers are a powerful political force, affecting more than just teachers' compensation. This has been highly visible during the pandemic, as local unions influenced district-level reopening decisions and the National Education Association sent suggestions that made their way into CDC guidance (Zilbermints 2021). Union strength varies considerably across U.S. states (Northern, Scull, and Shaw 2012).

Both states and institutions of higher education play important roles in determining who teaches and the preparation they receive. Policies related to teacher certification and preparation requirements, ranging from whether teachers are tested on academic content to which teachers are eligible to supervise student teachers, vary considerably across states.[4] Meanwhile, reviews of teacher training programs reveal many programs do not do a good job incorporating consensus views of research-based best practices in key areas (Pomerance and Walsh 2020). To date, schools of education have not been the focus of much policy discussion, but they would be critical partners in any changes to how teachers are trained.

Philanthropy has an important influence on education policy, locally and nationally. Not only do funders support individual schools in traditional ways, they are increasingly active in influencing federal and state laws (Reckhow, Tompkins-Stange, and Galey-Horn 2021). Part of these philanthropic efforts happen through advocacy groups, including civil rights groups, religious groups, and the hard-to-define "education reform" movement. Finally, the many vendors of curriculum, assessment, and "ed-tech" products and services bring their own lobbying power (Burch 2009).

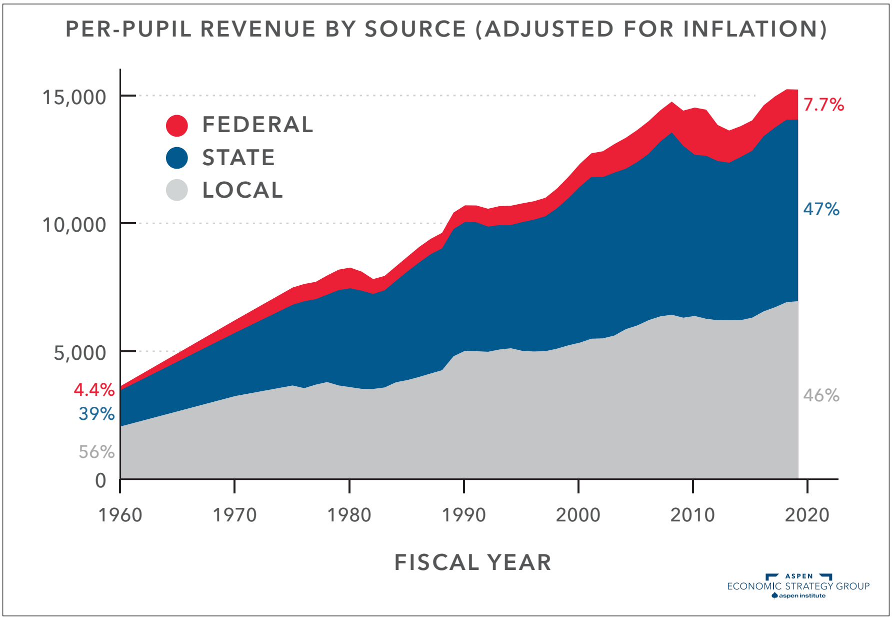
4 See the not-for-profit National Council on Teacher Quality <https://www.nctq.org/> for standards and reviews of teacher preparation programs, and descriptions of state teacher preparation policies.

3.b. Paying for school

Research on school finance might be better termed school district finance because districts are the jurisdictions generating and receiving revenue, and districts, not schools, are almost always responsible for spending decisions. School districts typically use staffing models to send resources to schools, specifying how many staff positions (full-time equivalents, or FTEs), rather than dollars, each school gets.

Inflation-adjusted, per-pupil revenue to school districts has increased steadily over time and averaged about $15,500 in the most recent year recorded (total expenditure, which includes both ongoing and capital expenditure, is similar but we focus on revenue because we are interested in the sources of revenue). Per-pupil revenue growth tends to stall or reverse in recessions, and has only recently recovered to levels seen prior to the Great Recession (Figure 5). On average, school districts generated 46% of their revenue locally, with about 80% of that from property taxes; about 47% of revenue came from state governments and about 8% from the federal government. The share of revenue raised locally has declined from about 57% in the early 1960s to 46% today, while the state and federal shares have grown.

Figure 5: Revenue per pupil has increased over time but declined during the Great Recession



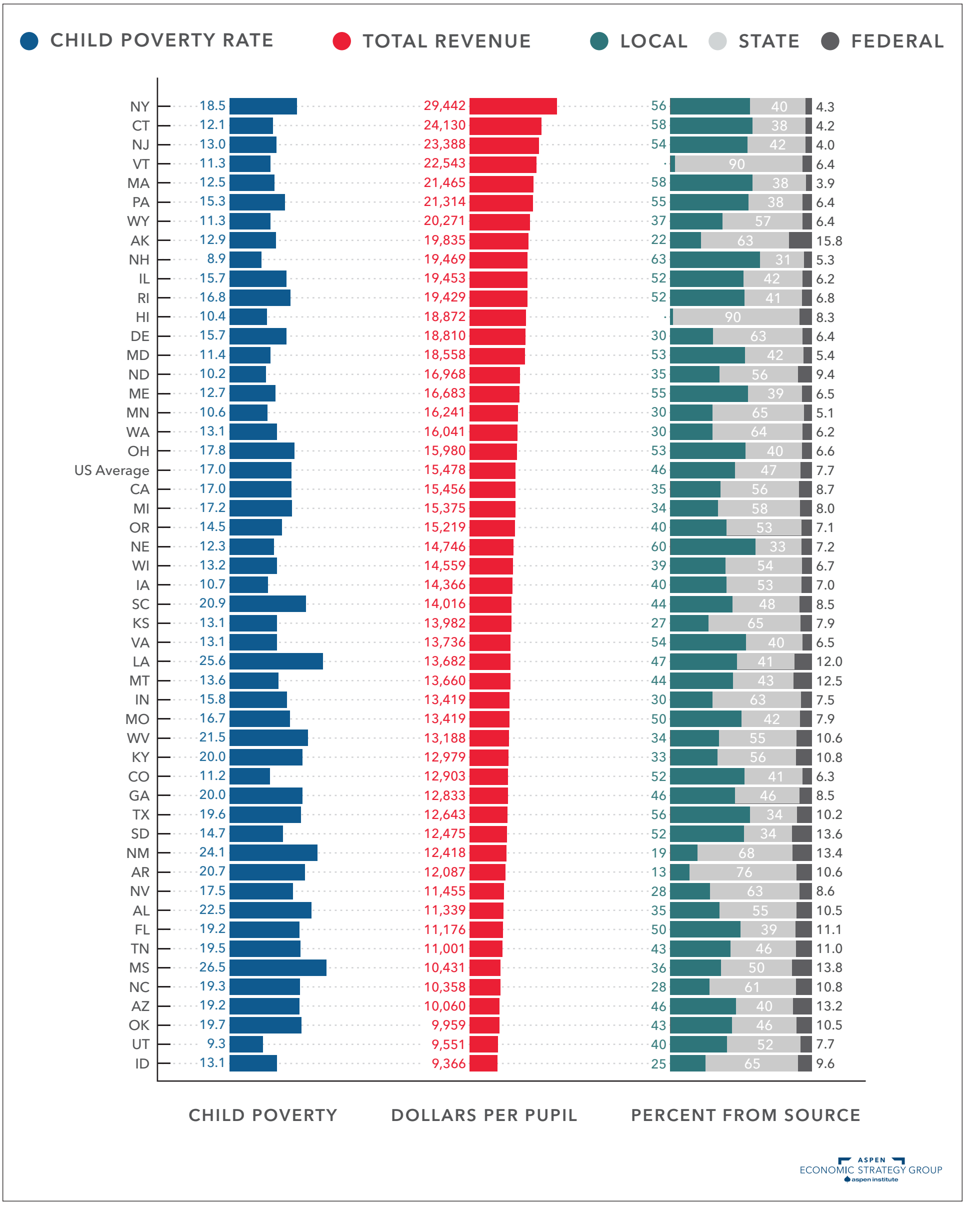
Source: Digest of Education Statistics

Notes: Per-pupil values have been adjusted to 2019 dollars using CPI-U.

Local revenue comes from taxes levied by local school districts, but local school districts often do not have complete control over the taxes they levy themselves, and they almost never determine exactly how much they spend because that depends on how much they receive in state and federal aid. State governments may require school districts to levy certain taxes, limit how much local districts are allowed to tax or spend, or they may implicitly or explicitly redistribute some portion of local tax revenue to other districts.

Both the level of spending and distribution of revenue by source vary substantially across states (Figure 6), with New York, the highest-spending state, spending almost $30,000 per pupil, while Idaho, Utah, and Oklahoma each spent under $10,000 per pupil. (Some, but far from all, of this difference is related to higher labor costs in New York.) Similarly, the local share of revenue varies from less than 5% in Hawaii and Vermont to about 60% in New Hampshire and Nebraska. On average, high-poverty states spend less, but there is also considerable variation in spending among states with similar child poverty rates.

Figure 6: Total revenue per pupil and revenue sources vary dramatically across states

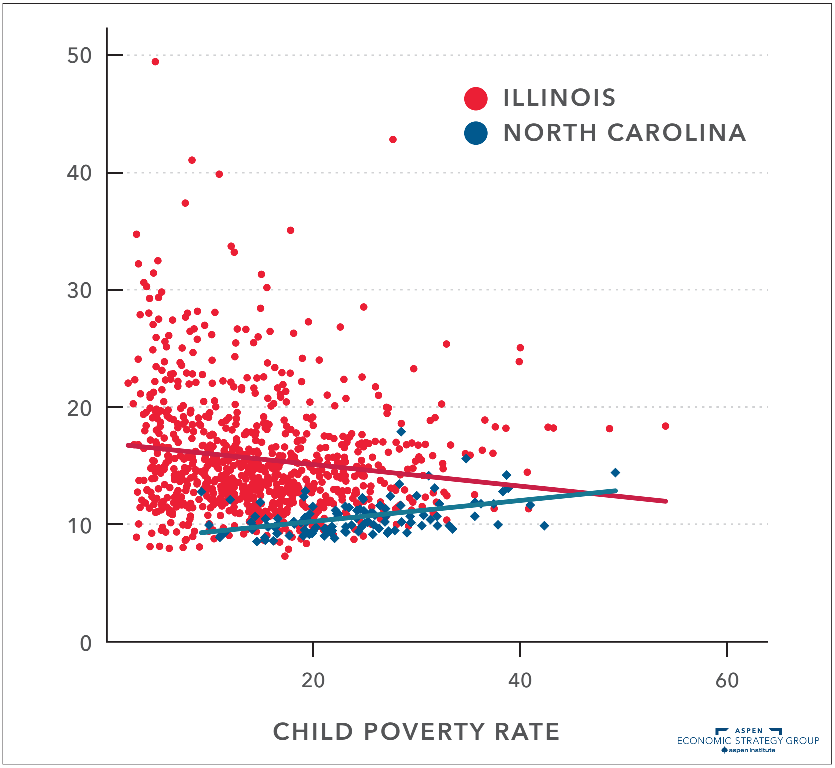


Source: Digest of Education Statistics

Notes: Per-pupil values are for the 2017-18 school year and have been adjusted to 2019 dollars using CPI-U.

Discussions of school funding equity—and considerable legal action—focus on inequality of funding across school districts *within the same state*. While people often assume districts serving disadvantaged students spend less per pupil than wealthier districts within a state, per-pupil spending and the child poverty rate are nearly always uncorrelated or *positively* correlated, with higher-poverty districts spending more on average. Typically, disadvantaged districts receive more state and federal funding, offsetting differences in funding from local sources. Meanwhile, considerable inequality exists between states, and poorer states spend less on average. Figure 7 illustrates this point, showing the relationship between district-level per-pupil spending and the child poverty rate in North Carolina (a relatively low-spending state with county- and city-based districts) and Illinois (a higher-spending state with many smaller districts). In North Carolina, higher poverty districts spend more on average; Illinois is one of only a few states in which this relationship is reversed. But this doesn't mean poor kids get fewer resources in Illinois than in North Carolina. Indeed, nearly *all* districts in Illinois spend more than most districts in North Carolina, regardless of poverty rate.

Figure 7: Per-pupil current expenditure in IL and NC

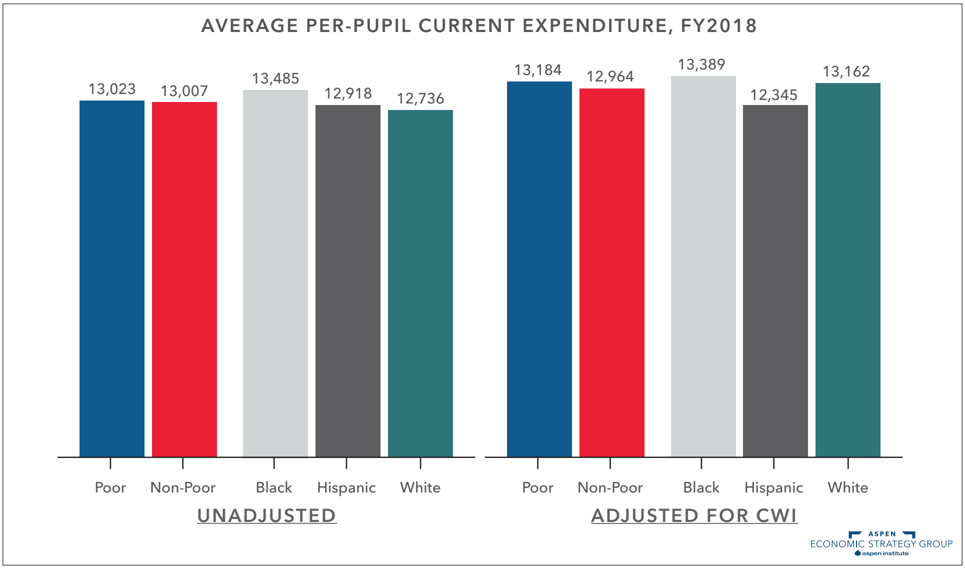


Source: Digest of Education Statistics

Notes: Per-pupil values are for the 2017-18 school year and have been adjusted to 2019 dollars using CPI-U.

Figure 7 gives a flavor of the wide variation in per-pupil school spending. Nationally, the district at the 10th percentile had per-pupil current expenditure of $8,800, compared to $18,600 at the 90th percentile (for these calculations we focus on current expenditure, which is less volatile year-to-year, rather than revenue). This variation is notably *not* systematically related to key demographics. For example, on average, poor students attend school in districts that spent $13,023 compared to $13,007 for non-poor students. The average Black student attends school in a district that spent $13,485 per student, compared to $12,918 for Hispanic students and $12,736 for White students.[5] School districts in high-wage areas need to spend more to hire the same staff, but adjusting spending to account for differences in prevailing wages of college graduates (the second set of bars) does not change the picture much.

Figure 8: On average, per-pupil current spending is similar across demographic groups



Source: Census of Governments/F-33, Common Core of Data, and Comparable Wage Index for Teachers.

Notes: Per-pupil values are for the 2017-18 school year and have been adjusted to 2019 dollars using CPI-U. Vermont is excluded due to data limitations.

5 These statistics may be particularly surprising to people given the widely publicized findings of the EdBuild organization that, "Nonwhite school districts get $23 billion less than white school districts." <https://edbuild.org/content/category/problems>. The EdBuild analysis estimates gaps between districts where at least 75% of students are non-White versus at least 75% of students are White. These two types of districts account for 53% of enrollment nationally. The $23 billion refers to state and local revenue (excluding federal revenue), whereas we focus on current expenditure (though patterns for total expenditure or total revenue are similar).

Does this mean the allocation of spending is fair? Not really. First, to make progress reducing the disparities in outcomes discussed above, schools serving more disadvantaged students will need to spend *more* on average. Second, these data are measured at the *school district* level, lumping all schools together. This potentially masks inequality across schools in the same district. The federal government now requires states to report some spending at the school level; states have only recently released these data and they have yet to be systematically analyzed, but past research on selected districts suggests that within-district spending differences are systematically correlated with student characteristics. How does this happen, when property taxes and other revenues for the entire district feed into the central budget and are not allocated based on neighborhood characteristics? Most of what school districts buy is staff, and compensation is largely based on credentials and experience. So schools with less-experienced teachers spend less per pupil than those with more experienced ones, even if they have identical teacher-to-student ratios. Research suggests schools enrolling more economically disadvantaged students, or more students of color, on average have worse working conditions for teachers and experience more teacher turnover. Together, this means that school districts using the same staffing rules for each school—or even allocating more staff to schools serving more economically disadvantaged students—would have different patterns in spending per pupil than *staff* per pupil.

4. Policy levers for improving schools

To improve the quality of education American schools offer, something—realistically, many things—will have to change, but what? There is not, unfortunately, a proven and specific prescription for running schools well, nor is it likely that any one prescription is best across the wide range of local contexts in which American schools operate.

We begin this section with a discussion of key inputs: school staff, peers, curriculum and materials, and infrastructure, and the ways in which these inputs are combined in different arrangements through programming and scheduling. These inputs most directly influence instruction. They also cost money, so we next turn to how money matters, and how schools can get more of it. Finally, we turn to approaches that attempt to improve schools by changing incentives at a high level, including desegregation, school choice, standards and accountability, and the use of evidence in policy and practice.

Throughout, we emphasize what credible research shows and where research is inconclusive. Some educational approaches are more difficult than others to evaluate rigorously, so it is important to keep in mind that a lack of evidence does not necessarily imply no impact. Much of the research in this area focuses on the average impact of these different approaches; some examine how different strategies affect different groups of students. Further, these strategies may have different effects (for some or all students) when implemented as part of more holistic efforts than in isolation, which is how they are typically evaluated.

4.a. Key inputs

4.a.1. School staff

Some of the most compelling findings in recent decades of education research show the importance of teachers and principals and how much effectiveness varies even in the same school district or school. A number of studies use "value-added" approaches to estimate teacher effectiveness, measuring how much each teacher's students improve on tests, controlling for student characteristics including prior test scores. Researchers have raised a number of objections to this approach; for example, they argue that controls for student background may not be sufficient, in which case a teacher who is "ineffective" according to value-added measures may just be one who teaches students who face more barriers to learning. Others point out that good teaching is about more than raising test scores. Despite these concerns, the weight of the evidence suggests that these measures are meaningful, if imperfect, and that teacher effectiveness varies substantially across teachers (Chetty, Friedman, and Rockoff 2014).

Such research points to potentially large benefits of improving average teacher quality, either by changing how teachers are hired and who is retained or by helping existing teachers improve. Each of these approaches is challenging. Although recent research has identified some more effective approaches to screening teachers at the time of hire, standard hiring and compensation practices mostly reward degrees and certifications that do not consistently predict teacher effectiveness (Staiger and Rockoff 2010; Jacob et al. 2018; Goldhaber, Grout, and Huntington-Klein 2017). And better screening at the time of hire without also improving pre-service and in-service training or attracting more people to the teaching profession can only go so far, as teachers screened out by one school or district may simply be hired in another. Most in-service professional development programs for teachers are not effective, though some programs evaluated with randomized controlled trials have shown effects on teaching practice and student achievement (Institute et al., n.d.). And setting a higher bar for tenure (currently nearly all teachers receive tenure after completing the requisite number of years) faces major resistance from teachers' unions. It would also mean more teachers would be inexperienced, and could influence entry into the profession.

Reducing barriers to entering the teaching profession is critical for several reasons. First, policies that would screen out more teachers at the time of hire or after a few years on the job would require a larger inflow to maintain staffing ratios. Further, would-be teachers cannot be expected to make major investments in certification if they face a significant chance of being asked to leave the profession after a few years. In addition, reducing barriers to entry could help improve diversity of the profession. While almost half of students are non-White, 80% of young teachers are White, and a growing body of research shows that students benefit from same-race teachers (Figlio 2017; Gershenson, Hansen, and Lindsay 2021). For example, one recent study showed that being assigned to at least one Black teacher in the early grades increased the likelihood that Black students graduate from high school by 9 percentage points (Gershenson et al. 2018). The under-representation of Black and Hispanic people among college graduates is a barrier to diversifying the teacher workforce, but there is scope to learn from states and school districts that have hired more diverse workforces and experiment with new approaches (Figlio 2017).

If teachers are so important, should we simply pay them more, or forgive teachers' student loans? Teacher pay has indeed been stagnant in recent years, though teachers have better-than-average benefits, the cost of which is increasing. Blanket loan forgiveness for teachers is not well-targeted for improving the recruitment and retention of highly effective teachers, though existing programs like the income-driven repayment and public service loan forgiveness programs need reform to make them more transparent and easier to use. Across-the-board pay increases are warranted in some places and could make teaching more attractive over time, but are unlikely to yield major improvements in student outcomes overall—and will do little to address disparities without other changes. Improving working conditions for teachers could also help; many of the "working conditions" teachers prefer—adequate planning time, appropriate training, reasonable student-teacher ratios, some autonomy—are key to effective instruction.

Although looming "teacher shortages" have been overstated, some types of teachers have been chronically under-supplied; for example, positions in math and science, special education, and teaching English learners are consistently difficult to staff (Dee and Goldhaber 2017). Higher salaries for these specialties are likely necessary to make sure those positions are consistently filled with effective teachers. Rural districts face additional staffing challenges and higher vacancy rates (Stoddard and Toma 2021).

More effective approaches would change the structure of teacher pay and create new career pathways for effective teachers. Most teacher salary schedules award higher salaries to teachers with degrees and certifications that are not consistently related to effectiveness, and more experienced teachers also earn higher salaries.

Together with the relatively generous pension benefits, this means that pay is back-loaded. Reforms that raise starting pay, put more emphasis on paying for credentials and training that are proven to improve teaching, or pay more for difficult-to-find expertise or to work in hard-to-staff schools are more promising.

A number of states and school districts implemented teacher evaluation reforms in the 2010s, sometimes paired with more fundamental changes to the structure of pay and dismissal practices. The approaches varied, but common components included new methods for evaluating teachers (usually based on student test scores or observations), one-time bonuses or permanent pay increases based on those evaluations, bonuses for teaching in particular schools, and new career pathways where teachers serve as mentor teachers or observe and evaluate other teachers. A common goal of this type of reform was to end the common practice of *pro forma* teacher evaluation, where vanishingly few teachers receive unsatisfactory ratings. By and large, these reforms were not well-implemented and did not lead to more differentiation in teachers' performance ratings, though in some cases they brought renewed attention to instructional quality (Kraft and Gilmour 2017). Even if few teachers are dismissed, more extensive (or different) teacher evaluations can prompt teachers to pay more attention to their practice, and some systems involved explicit feedback from a mentor teacher. One study found that teachers who participated in a more intensive evaluation program—where they were observed and rated several times during the year—were better teachers for several years following the evaluation (Taylor and Tyler 2012).

The District of Columbia school system implemented one of the more comprehensive reforms, IMPACT. The program was controversial and underwent two major revisions, landing on a system that gave strong incentives for low-performing teachers to improve or be dismissed and for highly effective teachers to work in the most disadvantaged schools (among other provisions), but did not rely too heavily on value-added measures of teacher performance or school-wide rewards. An evaluation of the IMPACT program showed it produced a broad and sustained improvement in teacher effectiveness (Dee, James, and Wyckoff 2021).

Aside from increasing pay or improving working conditions to attract and retain better teachers or formal training, other factors can influence teacher effectiveness. For example, pairing student teachers with more instructionally effective cooperating teachers—rather than just more experienced ones—improves their subsequent performance as new teachers (Ronfeldt, Brockman, and Campbell 2018). Allowing teachers who have been certified in one state to more easily teach in another state could alleviate shortages in some cases (Goldhaber, Grout, and Holden 2017).

Principals and school counselors have also been shown to influence student outcomes. Though individual counselors matter (Mulhern 2020), there is little research on improving counselor quality, but the Wallace Foundation has developed a promising approach to developing "principal pipelines" to identify aspiring principals from existing staff and prepare and support them as new principals. This relatively inexpensive program had positive impacts on student achievement and principal retention (Gates et al. 2019).

4.a.2. Peers

Students in a school influence each other through a number of channels. They may learn from each other, influence what type of curriculum is offered, influence the culture of the school, and use more or less of the teacher's time. A disruptive student can reduce the time students are actively learning. Parents care a lot about who their children will attend school with, in part for nonacademic reasons. Studying peer effects is difficult because students with similar characteristics tend to enroll in the same schools—"birds of a feather flock together"—so it is not surprising to find that, for example, students with high-scoring peers themselves have high test scores. And how peers affect each other can be complex and context-dependent. Estimates of peer effects cover a wide range, and some methodological questions remain. Several studies find that students benefit from having more girls and more high achieving peers, and that high achieving students benefit most from other high achieving peers (Sacerdote 2011). A recent study finds that U.S.-born students benefit from having more immigrants in their schools, and this effect is concentrated among lower-achieving students (Figlio et al. 2021). Some (but not all) studies suggest that students do better when variation in achievement is lower (Sacerdote 2011), possibly because it is easier to target instruction when students are working at similar levels (Duflo, Dupas, and Kremer 2011).

Students with behavioral problems appear to negatively impact their peers' learning, and this is a possible reason that having more girls in a class is beneficial (Pope and Zuo 2020; Carrell and Hoekstra 2010). One recent study finds that sharing a classroom with a student who is exposed to domestic violence affects educational outcomes years in the future and even reduces wages (Carrell, Hoekstra, and Kuka 2018). This points to the potentially enormous benefit of addressing the underlying problems of students who disrupt classrooms.

4.a.3. Curriculum and educational products

While curriculum is often operationalized in part by instructional materials, such as textbooks or software, educators conceptualize it more broadly. Grover "Russ"

Whitehurst, the founding director of the Institute of Education Sciences, defines curriculum as “the content and sequence of the experiences that are intended to be delivered to students in formal coursework.” (Whitehurst 2009) This definition encompasses much of the work of schools, so it’s hard to imagine efforts to improve schools that do not address curriculum. However, it is far from straightforward to improve schools by changing curriculum or instructional materials.

Not only must leaders choose the right curriculum for their contexts, they must ensure it is implemented well (Polikoff, Wang, and Kaufman 2021). These two issues are closely linked: A curriculum that improves student outcomes in one place will not work in another context if staff lack the resources or will to implement it well. Educators also regularly choose their own supplementary materials. For example, over half of teachers surveyed in the 2017 American Teacher Panel reported using the website TeachersPayTeachers.com at least once a week to obtain supplementary materials (Kaufman et al. 2017). Educators report wanting more information to help them choose materials that are cost effective and aligned with curricular standards (Polikoff and Campbell 2018). In practice, resources exist to help identify “evidence-based” materials (*e.g.*, the What Works Clearinghouse), or to identify materials that are aligned with standards (*e.g.*, EdReports), but not both in a one-stop shop. Further, information on the full set of costs associated with implementation is hard to come by.

Proponents of “culturally relevant pedagogy”—which can include a range of practices designed to nurture students’ ethnic and social identities, appreciating their own culture while developing fluency in at least one other culture—argue it can improve student learning, in part by strengthening the student-teacher relationship. These approaches have strong theoretical underpinnings but have rarely been evaluated systematically. Some programs supported by the Obama Foundation’s My Brother’s Keeper initiative promote culturally relevant pedagogy. And California has developed model ethnic studies courses for use in its schools. An evaluation of Oakland’s My Brother’s Keeper program, a high-school level course staffed by Black men, found it reduced dropout rates (Dee and Penner 2019); an ethnic studies program in San Francisco improved attendance, GPA, and credits earned substantially (Dee and Penner 2017). Further research should assess whether these promising programs are effective when adopted at scale.

In most cases, “evidence-based” educational materials are stand-alone curricular components, not a full curriculum.[6] The success of one component often depends on what else is happening in the school; for example, a social studies program centered around primary source materials may work well only if students have had reading

6 See, for example, the What Works Clearinghouse <https://ies.ed.gov/ncee/wwc/>.

instruction sufficient to understand the primary sources. This makes evaluations of such individual components hard to interpret. Another challenge comes from the nature of evaluation itself: Programs subject to evaluation must be defined clearly—and in some cases rigidly—if the results of the evaluation are to generalize to other settings. In other words, the programs must be designed to minimize educator discretion, which rules out many less rigid approaches from being evaluated in ways that would deem them evidence-based (Gordon 2018). These evaluation challenges are greater still when it comes to special education, which centers around providing and delivering an Individualized Education Program to each student.

The marketplace for educational products, such as textbooks and software, is segmented into "core" and "supplemental" products. The core corresponds to what you may think of as standard school: the instruction that goes on in general (not special) education classrooms. Supplemental products are designed for when students do not achieve proficiency, as determined by grade-level benchmarks. Supplemental products are widely used in part because the structure of state and federal funding mean that districts may find it easier to spend on supplemental services than on shoring up the core. School districts feel pressure to show that categorical funding streams, like Title I federal funds, state funds for English learners, or special education dollars, are cleanly allocated to expenditures benefiting *only* those students—even if spending on the core, which would benefit all students, might help the target population more than add-on interventions (Gordon and Reber 2015; Setren forthcoming).

4.a.4. School infrastructure

Research shows that spending on capital improvements or to build new schools improves test scores and other outcomes (Jackson and Mackevicius 2021). Though the CDC notes there is no safe level of lead exposure for children, 37% of schools that test for lead reported elevated lead levels; fewer than half of districts even tested (GAO 2018b). Evidence that poor indoor air quality and exposure to lead and other toxins impedes learning and can have long-term effects is now conclusive (Aizer et al. 2018; Stafford 2015). Approximately one-third of schools require HVAC updates (GAO 2020). Studies also show that heat impacts learning adversely, especially when schools do not have air-conditioning (Park et al. 2020; Park, Behrer, and Goodman 2021). Schools serving low-income students and students of color are more likely to lack air conditioning conditional on other factors, and Park et al. (2020) estimate that heat accounts for 1 to 13% of U.S. racial achievement gaps. Installing air conditioning could plausibly help shrink achievement gaps.

4.a.5. Organizing these inputs

The effectiveness of the individual inputs we have described above depends on how the inputs are combined and used in schools. Districts and schools have considerable discretion over scheduling, and over how they group students, teachers, and other staff in schools and classrooms. These organizational choices include determining school and class size, how students and teachers are assigned to each room, how students are grouped inside classes or "pulled out" to work with a paraprofessional or specialist, and how to handle student behavioral problems.

Class size is a key concern for many teachers and parents. After declining steadily for decades, student-teacher ratios[7] increased during the Great Recession and have remained somewhat elevated since, and are about average in the OECD. Student-teacher ratios vary substantially across states, ranging from about 23 in Arizona, California, and Utah, to about 12 in several states. Studying the effects of class size is difficult because a number of factors can influence which students and teachers are in smaller classes. In the 1980s, the state of Tennessee conducted a randomized experiment—Project STAR—to test the effects of small classes. Research on Project STAR found substantial benefits of smaller classes (Krueger 1999) and has contributed to a widespread view that class size reduction "works," though class size reduction is expensive, and Chingos (2013) concludes the STAR intervention was not cost-effective.

Large-scale class-size reductions are often not able to replicate the idealized conditions of the STAR experiment. For example, California's class-size reduction program, which focused on the same grades as STAR, produced a reshuffling of staff that worsened teacher experience inequities across schools (Schrag 2006). Compelling (but nonexperimental) work has failed to identify similarly large effects of class-size reduction in more recent data (Schwartz, Zabel, and Leardo 2017). Still, class size—or staff-to-student ratios more generally—may matter in ways that are not easily captured in this research. For example, class size is an important working condition and may affect which teachers can be attracted and retained, and some instructional approaches may only work in smaller classes.

Experimental assignment to small or large *schools* poses an even greater challenge, and the research base on this question is small. Most compelling studies suggest small schools have modest benefits (Gershenson and Langbein 2015; Barrow,

7 This is not quite the same as "class size" because it includes special education and other teachers who may not be "regular classroom teachers."

Schanzenbach, and Claessens 2015). Meanwhile, schools could reap large benefits—at little cost—by starting the school day later for middle and high schools (allowing adolescents to get enough sleep) and creating K-8 schools to avoid transitioning adolescents to new schools at the developmentally challenging time of 6th or 7th grade (Jacob and Rockoff 2011).

School leaders make many decisions about how students are grouped within buildings and rooms. For example, how many different "levels" of algebra II are offered in a high school? Are gifted and talented students pulled out of their regular classrooms one period per week for supplemental instruction, do they attend different schools, or are they granted access to additional materials while remaining in their regular classroom settings? Such choices could have important implications for student achievement and equity. Unfortunately, documenting the causal impacts of these grouping choices is challenging. Not only are data sparse, but local needs and capacity drive choices to use these practices and, independently, student outcomes (Nomi 2009). Further, the effectiveness of different grouping strategies is surely linked to the curriculum and pedagogical practices they are used alongside.

Despite the importance of and considerable variation in teacher effectiveness, the allocation of teachers to schools is generally not made purposefully. In most districts, new teaching vacancies are first offered to teachers within the district, and more experienced teachers tend to move to placements in more advantaged schools over their careers. This means that inexperienced teachers are disproportionately hired in schools serving low-income students and students of color. This turnover harms student achievement (Ronfeldt, Loeb, and Wyckoff 2013; Carver-Thomas and Darling-Hammond 2017). Lower-achieving students and Black students appear to be more affected by teacher quality, so improving the level and distribution of quality teachers could reduce disparities (Aaronson, Barrow, and Sander 2007). Collective bargaining agreements typically prohibit "forced placements" of teachers into specific schools; even without such agreements, teachers might leave a district rather than take an assignment they find undesirable. Research suggests it is far costlier to use financial incentives to induce teachers to switch schools than to retain them in schools where they already teach (Dee and Goldhaber 2017).

Schools must decide not only how to group students for instruction relating to the core curriculum, mapping to grade-level standards, but also for additional supports. "High-dosage" tutoring, in which a trained tutor works with no more than a few students at a time, either at least three times a week or in intensive, week-long programs (Robinson et al. 2021), is a promising approach for students performing below grade level; both the grouping of students and staff and what happens during the sessions are important for a successful program.

For some students, specialized instruction or services are required by law (*e.g.,* students with disabilities, special education, English learners, gifted and talented). Even for these students, schools have some discretion in how they group them; for example, students can be "pulled out" of general education classrooms for additional services or have additional resources "pushed in" to that classroom, or they can use separate classrooms or schools for some students. Teachers need more preparation to serve English learners and students with disabilities more effectively, and more research on how to improve outcomes for these students more broadly is sorely needed (Mavrogordato et al. 2021).

There is, however, a strong research consensus that the processes *identifying* students eligible for a range of specialized services are not equitable. Grissom and Redding (2016) find that Black students are half as likely to be referred for gifted programs compared with White peers, even after controlling for test scores, with a notable exception: when Black students have a Black teacher. Universal screening processes that test all students for gifted eligibility rather than those whose parents opt in are one strategy to support more inclusive identification; another is to rely on a range of assessments rather than just one. Some large districts are just beginning to eliminate the use of standardized tests for admission to competitive magnet schools.

Racial disparities in special education are known as "disproportionality" and related discussions often focus on over-identification—students being incorrectly classified as having a disability—perhaps due to concerns that the classification serves to remove students from general education classrooms rather than to provide useful services (Gordon 2018). However, under-identification—students who do have disabilities not receiving what would be an accurate classification—is also a serious problem, as it may prevent students from receiving important services to which they are entitled. Some states are turning to universal screening as a tool for more accurate identification of some high-incidence learning disabilities such as dyslexia. For those policies to be effective, positive screens need to result in timely full evaluations, and, perhaps most challenging, schools must be equipped to serve students identified with disabilities.

Students are sometimes removed from the general education setting for noninstructional reasons as well. The use of exclusionary discipline, including suspension and expulsion, disproportionately affects students of color, even controlling for the incident leading to disciplinary action (Barrett et al. 2019). Black boys have the highest rates of out-of-school suspension of any group, with 18% suspended at least once in a given school year, over three times the rate for White boys (GAO 2018a). Principals play a significant role in the prevalence of suspension use; for example, Bacher-Hicks et al. (2019) find that attending schools run by principals who encourage suspension increases the likelihood that students, especially males and students of color, are arrested and incarcerated later in life. Lindsay and Hart (2017) find that Black students are less likely to be suspended when their teacher is also Black.

Two popular alternatives (or supplements) to exclusionary discipline are restorative justice and positive behavioral interventions and supports. Both methods require staff time and investment in training (Gray et al. 2017). A random assignment evaluation found that restorative justice practices implemented in Pittsburgh public schools reduced suspensions but negatively impacted math scores for Black students in elementary and middle school (Augustine et al. 2018).

4.b. Budget

To what extent do better schools simply spend more? The idea that increasing funding would improve schools is intuitive, yet many early studies found a weak, or even negative, relationship between school spending and outcomes, leading some to shift focus away from how much schools spend toward how they spend it (Hanushek 1997). These older studies suffered from some methodological limitations, and a meta-analysis of more recent studies shows that additional funding does typically improve test scores and educational attainment (Jackson and Mackevicius 2021). The size of the effect varies considerably across studies, suggesting again that context matters. The analysis points to larger benefits of additional spending in schools that serve low-income students and finds little evidence of diminishing returns to additional spending so far. Effects of spending on educational attainment are typically larger than effects on test scores, suggesting that not all improvements in school quality are captured by test scores.

Unfortunately, data limitations mean most studies on spending do not provide much insight into how money is spent, although they do sometimes distinguish capital expenditures from ongoing spending and find that both types of spending benefit students. Meanwhile, research on how specific initiatives affect student outcomes historically has not reported costs, though this appears to be changing. In 2020, the Institute of Education Sciences began to require cost analysis in funded research projects.

Increasing school spending—and targeting aid to where it is needed most—is more difficult than it seems because lower levels of government can adjust their spending in response to aid from higher levels of government. For example, a school district that receives more funding from its state might cut taxes instead of increasing spending.

In the last half century, states have frequently changed the amount and method of distributing aid to local districts, often in response to litigation related to education provisions in state constitutions. Usually, the goal of these changes is to reduce inequality in spending across districts or increase spending in particular types of districts (defined by low spending, low income, or low tax bases, depending on the context). These reforms have reduced, but far from eliminated, spending disparities across districts within states and reduced inequality in outcomes (Card and Payne 2002; Lafortune, Rothstein, and Schanzenbach 2018). The average effects of school finance reforms on spending mask considerable heterogeneity across states; in some states, they have little impact on spending (Shores, Candelaria, and Kabourek 2020).

Unfortunately, key questions about state school finance programs remain unanswered. Successful legal challenges to state systems prompt judicial mandates that typically tell the state what distribution of spending is acceptable, but not how to achieve it. This may well be because there is not a consensus among researchers or advocates on the best way to achieve more equal or equitable funding. Recent research provides suggestive evidence that equalization plans and categorical aid correlate with more progressive post-reform outcomes (Shores, Candelaria, and Kabourek 2020). Court orders prescribing more and more equitable funding appear to be more likely to translate to action where unions are strong; in states with weaker unions, increased state aid more often leads to local property tax reductions instead of spending increases (Brunner, Hyman, and Ju 2020). In many states, litigation has not had its intended effects, while a number of states without court orders have increased the level of spending and adopted similar formulas as those with court orders.

Local school districts and state governments might respond in ways that “undo” federal aid, especially since federal funding is a small share of the total. Local school districts could reduce their own taxes (or fail to raise them as much as they would have) when they receive more federal funding, depending on state rules governing local taxation; evidence on this question is mixed (Gordon 2004; Cascio, Gordon, and Reber 2013). State governments could offset federal aid by reducing their own effort or giving less state aid to districts that receive more federal aid. This question is important but difficult to study because responses might play out over time and we don’t know what states would have done in the absence of aid. This is also why attempts to prevent this type of offsetting behavior, such as maintenance of effort requirements, are not effective.

The discussion above speaks to challenges in increasing budgets at the school *district* level, but resource differences across schools within districts are also important. Policymakers and advocates have only recently focused on this issue. One approach aiming to address these gaps is weighted student funding (WSF). In theory, WSF directs dollars rather than positions to schools, and advocates typically propose this in conjunction with devolving more decision-making to the school level. In practice, this model is challenging to implement and sustain. A recent review of formulas used in WSF districts found that on average, districts ran less than half of funds through these formulas (Roza, Hagan, and Anderson 2021). Moving from a system where districts fund actual salaries of the staff employed in a school to providing schools a pot of funds that depend on student characteristics would leave schools with more experienced teachers unable to maintain their current workforce, redirecting those funds to schools with less experienced teachers. Most districts have labor agreements that prohibit forced transfers of teachers from one school to another, and federal policymakers have maintained support for "no forced transfers" in negotiations over how federal funds are used to influence the allocation of resources across schools.

4.c. Changing systems: governance and incentives

So far we have discussed changing budgets and a range of specific educational inputs. Other approaches attempt to change systems in ways that advocates hope will improve incentives, governance, or accountability and encourage better outcomes.

4.c.1. School desegregation

The 1954 Supreme Court ruling in *Brown v. Board of Education* did not immediately dismantle segregation, but between 1965 and 1970, most Southern school districts did desegregate, sometimes, though not always, under the supervision of a federal court (Cascio et al. 2008). Larger city districts and non-Southern districts continued to desegregate through the 1970s (Reber 2005). Desegregation directly changes students' peers, but that is not the only or even the most important effect of these programs. Advocates hoped that the quality of schools attended by Black students would improve if White students attended those same schools, presumably due to some change in the political economy governing the allocation of resources. Research on desegregation in the 1960s and 1970s largely supports that notion: segregation declined, despite some White flight (Reber 2005), and Black students benefited from attending desegregated schools (Reber 2010; Johnson and Nazaryan 2019). This appears to be mostly because desegregation improved the quality of schools Black students attended, by giving them access to better-resourced, formerly White schools and increasing spending (Reber 2011; Tuttle 2019). School desegregation significantly reduced the number of Black teachers, which likely reduced the benefit to Black students overall (Thompson 2020). Fewer studies are able to examine the effects of desegregation on educational outcomes for White students, but those that do find little effect (Tuttle 2019).

In the 1970s, school districts were often required by a court to make race-conscious school assignments to counteract residential segregation, but the current legal environment limits districts' ability to use race-conscious school assignment policies (including some that were originally ordered by a court). Still, how district boundaries and school attendance zones are drawn influences segregation. Historically, schools were more segregated than neighborhoods; Linda Brown did not want to go to a *White* school, she simply didn't want to travel to the Black school. In most school districts today, schools are about as segregated as expected based on residential patterns, and a small number of districts have attendance zones that are less segregated than neighborhoods (Monarrez 2021). Districts could take some actions to promote more integrated schools, but addressing residential segregation is critical to making real progress on school segregation: Parents prefer their children attend school close to home (Phi Delta Kappan 2017), and the current legal environment is unfriendly to ambitious desegregation approaches.

4.c.2. Choice programs

A number of approaches to reform that have been proposed or tried involve introducing additional schooling options that break the link between where a student lives and the school they attend: vouchers, charters, open enrollment, magnet schools, education tax credits, virtual academies, and homeschooling options. School choice programs could improve schools through several channels. First, they create options that may be better than the alternative for the particular students who choose to attend. By definition, parents who take advantage of these programs prefer the "choice" school to their next best alternative. However, this may simply reflect the low quality of the available traditional public schools, and sometimes choice schools produce measurably worse outcomes. Second, choice schools can serve a research and development purpose, acting as laboratories to try out new educational approaches. This was initially a key argument for charter schools, but while research on charter schools has identified some promising practices, efforts to bring those practices to traditional public schools have been limited, and the "laboratory" role of charters has been deemphasized over time.

Finally, choice programs could have a broader impact by creating competitive pressure on all schools, including traditional public schools, to improve. Even in the traditional system, where students are assigned to schools based on where they live, parents have choice and schools face competitive pressure since families can choose where to live or can send their children to private schools, subject to their ability to pay. While choice—among traditional school districts or through open enrollment, charters, or other mechanisms—can provide incentives for quality, it can also increase sorting of students by family income or other characteristics or make it difficult to redistribute resources. Choice programs may encourage schools to compete by enrolling students who enter with higher achievement levels and require fewer supports— "cream skimming"—rather than by improving value-added. The details of how choice programs are designed and financed matter critically for how much students who enroll in them benefit and whether they have positive or negative spillovers to the rest of the system.

Among choice programs in the United States, charter schools have the largest reach and are the most researched. The charter authorization process and funding approaches vary considerably across states, and individual charter schools vary in their instructional approaches, so generalizing about charter schools is difficult. Large statewide or national studies of charter school effectiveness typically find that charter schools have a small positive or no effect on student achievement on average. A number of studies estimate the effectiveness of oversubscribed charter schools by comparing students who gained admission through a lottery to those who applied for the lottery but did not gain admission and tend to find larger benefits of attending a charter school, compared to the broader studies of all charters (Cohodes and Parham 2021). This makes sense because the schools that have lotteries are by definition oversubscribed so more likely to be better than the available alternatives, which are often low-performing, urban schools. Virtual charter schools are harder to study because students often choose them because they are already struggling, but research suggests they produce worse academic outcomes than in-person schools (Fitzpatrick et al. 2020; Ahn and McEachin 2017).

With respect to cream skimming and student sorting, the story is mixed. Overall, the presence of charter schools appears to increase racial segregation modestly, though this varies considerably across states (Monarrez 2019). Charter schools are not allowed to have selective admissions, but they could influence who enrolls with policies such as onerous enrollment procedures and parental involvement requirements, discriminatory recruitment strategies, or with discipline policies that formally or informally push some students out. On the whole, there is little evidence that charters cream skim high-achieving students or push out low-achieving students; in fact, they often target disadvantaged, low-achieving students (Kho 2021). However, evidence suggests that some charters try to avoid enrolling students with disabilities (Bergman and McFarlin 2020) or to push out students with behavioral problems (Kho 2021). Setren (forthcoming) found that students with disabilities in Boston Public Schools were more likely to lose their classification if they won a charter lottery than if they lost; their achievement also improved.

Whether competition from school choice improves traditional schools is the big question. The evidence on charter schools is mixed but on balance suggests moderate improvements in nearby schools due to charter competition (Cohodes and Parham 2021).

Evidence on other choice programs is more limited, in part because experience with those programs in the United States is limited. Traditional voucher programs give families money that can be used to attend private school, with some restrictions. More recently, some states have tax credit scholarship programs, which allow corporations and other taxpayers to subtract the value of contributions made to qualifying non-profit scholarship programs from their taxes, effectively making the donations (including to religious schools) free to the taxpayer. Because charter schools are public, they are subject to state testing requirements that do not always apply to voucher or tax credit programs, making those programs difficult to evaluate. In Florida, which operates the largest such program, participating in the program was found to increase both enrollment and graduation from college (Chingos, Monarrez, and Kuehn 2019).

Research on small voucher programs implemented in the 1990s and 2000s showed small to substantial benefits for participating students, depending on the study; they often estimate larger benefits for Black students (Epple, Romano, and Urquiola 2017). More recent voucher programs showed smaller benefits, and one study found substantial negative effects of a private school voucher program in Louisiana (Abdulkadiroğlu, Pathak, and Walters 2018).

Open enrollment programs—where students can enroll in schools other than the neighborhood school to which they are assigned—or ranked choice programs—where families list their school choices and are assigned a school based on their lottery number combined with an assignment algorithm—have grown in popularity, especially in large urban districts. Research on these programs mostly focuses on how families choose and what those choices imply about what they value in schools and less about the effects of choice on student outcomes.

Overall, choice programs can provide families schooling options they prefer to the traditional public school to which they are assigned. Rigorous evaluations of such programs tend to find participating students benefit, though the magnitude of these effects varies considerably and is often small and sometimes negative (Epple, Romano, and Urquiola 2017). Competition from schools of choice has so far not been a game-changer, and the benefits of choice programs are ultimately limited by the availability of high-quality schools near where disadvantaged students live.

4.c.3. Standards and accountability

Rather than dictating how local districts should operate their schools, the idea behind accountability is to define desired outcomes, and hold districts accountable for reaching them. This makes sense if the best approach varies by local context and local actors have the best information about what works. Accountability will work better when the important outcomes are well-defined and well-measured and when the actors being held accountable have the capacity to change those outcomes.

Since the No Child Left Behind (NCLB) Act was signed into law in 2002, eligibility for most federal funds requires each state to submit an accountability plan. They must set curricular standards, test students on material corresponding to those standards, and make the testing data public.[8] NCLB required every school to be on a path to have 100% of students meeting "proficiency" standards in a short time. As this proved wildly unrealistic, even considering the low proficiency standards set in many states, the U.S. Department of Education granted waivers to states exempting them from this requirement. When Congress reauthorized the law as the Every Student Succeeds Act (ESSA) of 2015, they removed the unenforceable consequences, but kept the standards, testing, and reporting requirements. Under both versions of the law, states develop their own accountability systems, subject to federal requirements such as reporting outcomes separately for student subgroups including by race and ethnicity, economic disadvantage, and disability status.

Research shows that these accountability regimes on average had modest positive impacts on test scores and induced some predictable perverse responses, such as teaching to the test, focusing instruction on students near the proficiency thresholds, and reduced emphasis on instruction in untested subjects and grades (Dee, Jacob, and Schwartz 2013). States crafted and implemented accountability in a range of ways; success was also variable (Dee and Dizon-Ross 2019; Bonilla and Dee 2020).

The 2015 reauthorization also changed the requirements for schools identified as needing improvement. NCLB offered four specific "turnaround" options, while ESSA is more flexible. A meta-analysis of NCLB school turnaround efforts finds they yielded only modest positive impacts on math, but points to promising practices: they found stronger impacts when efforts involved extending learning time or replacing a significant share of a school's teaching staff (Henry et al. 2020).

8 The Common Core State Standards, which have been adopted by 41 states and the District of Columbia, were a collaborative cross-state effort and not a federal effort. These are curricular standards, specifying what students in different grades should learn in different subjects, rather than curricular materials that explain how students should be taught.

4.c.4. Using evidence

Federal policy encourages, and sometimes requires, states and districts to take "evidence-based" approaches to educating students. Louisiana has taken the further step of offering financial incentives for districts choosing evidence-based curricula from a state-provided list.

The idea that decisions should be evidence-based makes sense in the abstract but presents many practical challenges. Many widely used products and, more broadly, teaching strategies, have not been subject to rigorous study. The research base for many important practices is therefore thin and typically does not incorporate information on cost effectiveness. Research typically does not delineate the conditions under which the intervention was successful. At times, evidence-based approaches are incongruent with educators' and parents' strongly held beliefs or values; see, for example, recent debates around the "science of reading" (Hanford 2018).

There is a central tension in crafting policy about evidence use. Offering more flexibility, as with the federal approach, is important given the need to consider whether the research suggests an evidence-based approach makes sense in a specific local context. At the same time, education leaders have little time or research training, enhancing the appeal of simple lists. But not everything that works can be easily evaluated. The materials that can be most "cleanly" evaluated and marketed as evidence-based are those that limit teacher discretion, preventing potential gains from customization, and potentially making teaching less attractive as a profession for some.

5. Conclusion

We argue above that there is no silver bullet that will transform elementary and secondary education. Instead, we need to ensure students receive quality instruction by supporting the fundamentals, especially staffing. We also must identify particular schools and students in need of more targeted support and devote resources to them specifically. Advocates for equity should embrace the need to shore up schooling for everyone as essential to their cause: Without attention to staffing, core curriculum, and infrastructure, piling on more interventions is unlikely to help. This attention to fundamentals realistically constitutes a challenging and long-term agenda. In this section we summarize key principles to guide future efforts.

5.a. Recognize the key role of states

Throughout this chapter, we have emphasized variation across states and the power that state governments wield when it comes to education policy. The federal government can play a key research and development function and provide funding, which is especially important in recessions (Henry et al. 2020). And attaching strings to federal aid has been a powerful tool for inducing particular policy changes in states and local districts, but it is a blunt instrument; federal policymakers can reasonably choose only a few strings at a time. School districts tax and spend, and schools are where the rubber hits the road, but states make most of the important rules.

5.b. Pay attention to fundamentals

While well-implemented, locally supported, evidence-based programs and interventions can be cost-effective tools, technocratic fixes cannot substitute for the fundamental work of "core instruction." A focus on the basics is warranted; there is no substitute for effective teachers, supported by good principals and staff, working with a reasonable number of students, using a strong core curriculum, working in a well-maintained building with access to necessary technologies and supplies—including sufficient planning time. Students with disabilities, those learning English, and students who are not working at grade level for whatever reason should receive effective and appropriate intervention, but a quality core is critical and could reduce the need for intervention. States can support local districts as they work on these fundamentals; schools of education and unions should be critical partners in this work. Increased flexibility over the use of funds and, critically, making sure districts understand they can use categorical aid on core instruction could help.

- Every school needs a deep bench, with diverse, well-trained, and supported staff. Policies like creating career pathways for teachers, removing barriers to entering the teaching profession that are not associated with better teaching, ensuring student teachers are matched to effective supervisors, incentivizing teachers to stay in schools that have traditionally experienced high turnover, and building principal pipelines can help.
- Federal financing of critical improvements to school infrastructure, such as removing lead and updating HVAC systems, could make sense in light of low interest rates.

5.c. Increase emphasis on vulnerable students

Data and research related to students with disabilities, English learners, and American Indian and Alaska Native students are scarce, but the data we do have suggests these populations are often not being served well by our schools. Students who fall into more than one of these categories—for example English learners who also have learning disabilities—are particularly likely to have their needs misidentified or overlooked. Sometimes, best practices for instruction or intervention are well understood in theory but have not been delivered effectively in practice. In any case, a new focus on these groups—including collecting better data, conducting more research, and better training teachers—is warranted.

5.d. Adopt proven policies and practices, mind the details

We should continue to encourage the thoughtful adoption of strategies that have been shown to work or might be expected to work based on what we know about learning. However, these efforts require greater attention to engaging with educators and communities to ensure the strategies can be implemented well and make sense in the local context.

We also need better and different research to realize the full potential of "evidence-based" practice. Most education research evaluates whether a particular approach was better than some unspecified, business-as-usual approach. Willingham and Daniel (2021) propose instead using research to identify "gold standard" options for different questions of practice based on cost-effectiveness so that new research can compare alternatives to best practice, rather than whatever happens to be in place.

5.e. Continue to focus on what happens out of school

In this chapter, we focus on schools. Public education has the potential to promote economic growth and equality, but policies addressing out of school factors may be equally or more important. Providing income support and access to health care, reducing exposure to lead, and reducing violence, including police violence against Black communities, among other things, could do more to improve learning than many of the education policies described above. And quality schools will be complementary with these other policies if they help students arrive at school more ready to learn.

References

Abdulkadiroğlu, Atila, Parag A. Pathak, and Christopher R. Walters. 2018. "Free to Choose: Can School Choice Reduce Student Achievement?" *American Economic Journal: Applied Economics* 10 (1): 175–206. https://doi.org/10.1257/app.20160634.

Ahn, June, and Andrew McEachin. 2017. "Student Enrollment Patterns and Achievement in Ohio's Online Charter Schools." *Educational Researcher* 46 (1): 44–57. https://doi.org/10.3102/0013189X17692999.

Aizer, Anna, Janet Currie, Peter Simon, and Patrick Vivier. 2018. "Do Low Levels of Blood Lead Reduce Children's Future Test Scores?" *American Economic Journal: Applied Economics* 10 (1): 307–41. https://doi.org/10.1257/app.20160404.

Andrews, Matthew, William Duncombe, and John Yinger. 2002. "Revisiting Economies of Size in American Education: Are We Any Closer to a Consensus?" *Economics of Education Review* 21 (3): 245–62. https://doi.org/10.1016/S0272-7757(01)00006-1.

Ang, Desmond. 2020. "The Effects of Police Violence on Inner-City Students." *The Quarterly Journal of Economics* 136 (1): 115–68. https://doi.org/10.1093/qje/qjaa027.

Augustine, Catherine H., John Engberg, Geoffrey E. Grimm, Emma Lee, Elaine Lin Wang, Karen Christianson, and Andrea A. Joseph. 2018. "Can Restorative Practices Improve School Climate and Curb Suspensions?: An Evaluation of the Impact of Restorative Practices in a Mid-Sized Urban School District," December. https://www.rand.org/pubs/research_reports/RR2840.html.

Bacher-Hicks, Andrew, Stephen B. Billings, and David J. Deming. 2019. "The School to Prison Pipeline: Long-Run Impacts of School Suspensions on Adult Crime." National Bureau of Economic Research working paper #26257. https://doi.org/10.3386/w26257.

Bacher-Hicks, Andrew, Joshua Goodman, and Christine Mulhern. 2020. "Inequality in Household Adaptation to Schooling Shocks: Covid-Induced Online Learning Engagement in Real Time." National Bureau of Economic Research working paper #27555. https://doi.org/10.3386/w27555.

Barrett, Nathan, Andrew McEachin, Jonathan N. Mills, and Jon Valant. 2019. "Disparities and Discrimination in Student Discipline by Race and Family Income." *Journal of Human Resources*, September, 0118-9267R2. https://doi.org/10.3368/jhr.56.3.0118-9267R2.

Barrow, Lisa, Diane Whitmore Schanzenbach, and Amy Claessens. 2015. "The Impact of Chicago's Small High School Initiative." *Journal of Urban Economics* 87 (May): 100–113. https://doi.org/10.1016/j.jue.2015.02.002.

Bergman, Peter, and Isaac McFarlin. 2020. "Education for All? A Nationwide Audit Study of School Choice," IZA DP #13007, 49.

Bertrand, Melanie, and Julie Marsh. 2021. "How Data-Driven Reform Can Drive Deficit Thinking." *Phi Delta Kappan* 102 (8): 35–39. https://doi.org/10.1177/00317217211013936.

Bonilla, Sade, and Thomas S. Dee. 2020. "The Effects of School Reform under NCLB Waivers: Evidence from Focus Schools in Kentucky," Education Finance and Policy, 15 (1): 75–103. https://direct.mit.edu/edfp/article/15/1/75/10315/The-Effects-of-School-Reform-under-NCLB-Waivers.

Brunner, Eric, Joshua Hyman, and Andrew Ju. 2020. "School Finance Reforms, Teachers' Unions, and the Allocation of School Resources." *The Review of Economics and Statistics* 102 (3): 473–89. https://doi.org/10.1162/rest_a_00828.

Burch, Patricia. 2009. *Hidden Markets: The New Education Privatization.* New York and London: Routledge.

Card, David, and Abigail Payne. 2002. "School Finance Reform, the Distribution of School Spending, and the Distribution of Student Test Scores." Journal of Public Economics 83: 49–82.

Carrell, Scott E., and Mark L. Hoekstra. 2010. "Externalities in the Classroom: How Children Exposed to Domestic Violence Affect Everyone's Kids." *American Economic Journal: Applied Economics* 2 (1): 211–28. https://doi.org/10.1257/app.2.1.211.

Carrell, Scott E., Mark Hoekstra, and Elira Kuka. 2018. "The Long-Run Effects of Disruptive Peers." *American Economic Review* 108 (11): 3377–3415. https://doi.org/10.1257/aer.20160763.

Carver-Thomas, Desiree, and Linda Darling-Hammond. 2017."Teacher Turnover: Why It Matters and What We Can Do About It," Learning Policy Institute, 60.

Cascio, Elizabeth, Nora Gordon, Ethan Lewis, and Sarah Reber. 2008. "From Brown to Busing." *Journal of Urban Economics* 64 (2): 296–325. https://doi.org/10.1016/j.jue.2008.02.001.

Cascio, Elizabeth, Nora Gordon, and Sarah Reber. 2013. "Local Responses to Federal Grants: Evidence from the Introduction of Title I in the South." *American Economic Journal: Economic Policy* 5 (3): 126–59.

Cascio, Elizabeth, and Sarah Reber. 2013. "The K-12 Education Battle." In *Legacies of the War on Poverty*, edited by Martha J. Bailey and Sheldon Danziger. Russell Sage Foundation.

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff. 2014. "Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood." *American Economic Review* 104 (9): 2633–79. https://doi.org/10.1257/aer.104.9.2633.

Chingos, Matthew M. 2013. "Class Size and Student Outcomes: Research and Policy Implications." *Journal of Policy Analysis and Management* 32 (2): 411–38. https://doi.org/10.1002/pam.21677.

Chingos, Matthew, Tomás Monarrez, and Daniel Kuehn. 2019. "The Effects of the Florida Tax Credit Scholarship Program on College Enrollment and Graduation." Urban Institute. February 1, 2019. https://www.urban.org/research/publication/effects-florida-tax-credit-scholarship-program-college-enrollment-and-graduation.

Cohodes, Sarah R., and Katharine S. Parham. 2021. "Charter Schools' Effectiveness, Mechanisms, and Competitive Influence." National Bureau of Economic Research working paper #28477. https://doi.org/10.3386/w28477

Currie, Janet, and Johannes Schmieder. 2008. "Fetal Exposure to Toxic Releases and Infant Health." *American Economic Review,* 99 (2): 177–83. https://doi.org/10.1257/aer.99.2.177.

Darling-Hammond, Linda, Maria E. Hyler, and Madelyn Gardner. 2017. "Effective Teacher Professional Development," Learning Policy Institute, 76.

Dee, Thomas S., and Elise Dizon-Ross. 2019. "School Performance, Accountability, and Waiver Reforms: Evidence From Louisiana." *Educational Evaluation and Policy Analysis* 41 (3): 316–49. https://doi.org/10.3102/0162373719849944.

Dee, Thomas S, and Dan Goldhaber. 2017. "Understanding and Addressing Teacher Shortages in the United States," Brookings Institute, The Hamilton Project, 28.

Dee, Thomas S., Brian Jacob, and Nathaniel L. Schwartz. 2013. "The Effects of NCLB on School Resources and Practices," *Educational Evaluation and Policy Analysis*, 35 (2): 252–279. https://journals.sagepub.com/doi/abs/10.3102/0162373712467080.

Dee, Thomas S., Jessalynn James, and Jim Wyckoff. 2021. "Is Effective Teacher Evaluation Sustainable? Evidence from District of Columbia Public Schools." *Education Finance and Policy* 16 (2): 313–46. https://doi.org/10.1162/edfp_a_00303.

Dee, Thomas S., and Emily K. Penner. 2017. "The Causal Effects of Cultural Relevance: Evidence From an Ethnic Studies Curriculum." *American Educational Research Journal* 54 (1): 127–66. https://doi.org/10.3102/0002831216677002.

———. 2019. "My Brother's Keeper? The Impact of Targeted Educational Supports." w26386. National Bureau of Economic Research. https://doi.org/10.3386/w26386.

Duflo, Esther, Pascaline Dupas, and Michael Kremer. 2011. "Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya." *American Economic Review* 101 (5): 1739–74. https://doi.org/10.1257/aer.101.5.1739.

Epple, Dennis, Richard E. Romano, and Miguel Urquiola. 2017. "School Vouchers: A Survey of the Economics Literature." *Journal of Economic Literature* 55 (2): 441–92. https://doi.org/10.1257/jel.20150679.

Figlio, David. 2017. "The Importance of a Diverse Teaching Force." *Brookings* (blog). November 16, 2017. https://www.brookings.edu/research/the-importance-of-a-diverse-teaching-force/.

Figlio, David N., Paola Giuliano, Riccardo Marchingiglio, Umut Özek, and Paola Sapienza. 2021. "Diversity in Schools: Immigrants and the Educational Performance of U.S. Born Students." National Bureau of Economic Research working paper #28596. https://doi.org/10.3386/w28596.

Fitzpatrick, Brian R., Mark Berends, Joseph J. Ferrare, and R. Joseph Waddington. 2020. "Virtual Illusion: Comparing Student Achievement and Teacher and Classroom Characteristics in Online and Brick-and-Mortar Charter Schools." *Educational Researcher* 49 (3): 161–75. https://doi.org/10.3102/0013189X20909814.

Frankenberg, Erica, Genevieve Siegel-Hawley, Jia Wang, and Gary Orfield. 2012. "Choice Without Equity: Charter School Segregation and the Need for Civil Rights Standards," June. https://escholarship.org/uc/item/4r07q8kg.

GAO. 2018a. "Discipline Disparities for Black Students, Boys, and Students with Disabilitites." https://www.gao.gov/assets/gao-18-258.pdf.

———. 2018b. "Lead Testing of School Drinking Water Would Benefit from Improved Federal Guidance." https://www.gao.gov/assets/gao-18-382.pdf.

———. 2020. "School Districts Frequently Identified Multiple Building Systems Needing Updates or Replacement." https://www.gao.gov/assets/gao-20-494.pdf.

Gates, Susan, Matthew Baird, Benjamin Master, and Emilio Chavez-Herrerias. 2019. *Principal Pipelines: A Feasible, Affordable, and Effective Way for Districts to Improve Schools*. RAND Corporation. https://doi.org/10.7249/RR2666.

Gershenson, Seth, Michael Hansen, and Constance A. Lindsay. 2021. Teacher Diversity and Student Success: Why Racial Representation Matters in the Classroom.

Gershenson, Seth, Cassandra M. D. Hart, Joshua Hyman, Constance Lindsay, and Nicholas W. Papageorge. 2018. "The Long-Run Impacts of Same-Race Teachers." National Bureau of Economic Research working paper #25254. https://doi.org/10.3386/w25254.

Gershenson, Seth, and Laura Langbein. 2015. "The Effect of Primary School Size on Academic Achievement." *Educational Evaluation and Policy Analysis* 37 (1_suppl): 135S-155S. https://doi.org/10.3102/0162373715576075.

Goldhaber, Dan, Cyrus Grout, and Kristian L. Holden. 2017. "Why Make It Hard for Teachers to Cross State Borders?" Phi Delta Kappan 98 (5): 55–60. https://doi.org/10.1177/0031721717690367.

Goldhaber, Dan, Cyrus Grout, and Nick Huntington-Klein. 2017. "Screen Twice, Cut Once: Assessing the Predictive Validity of Applicant Selection Tools." Education Finance and Policy 12 (2): 197–223. https://doi.org/10.1162/EDFP_a_00200.

Gordon, Nora. 2004. "Do Federal Grants Boost School Spending? Evidence from Title I." Journal of Public Economics 88 (9): 1771–92. https://doi.org/10.1016/j.jpubeco.2003.09.002.

———. 2017. "Education Funding: Tax Credits Cost the Federal Government Money | Econofact." March 21, 2017. https://econofact.org/education-funding-tax-credits-cost-the-federal-government-money.

———. 2018. "What Works vs. What We Can Evaluate." Brookings (blog). June 14, 2018. https://www.brookings.edu/research/what-works-vs-what-we-can-evaluate/.

Gordon, Nora, and Eloise Pasachoff. 2018. "Fiscal Compliance Rules for Federal Funding of Elementary and Secondary Education: Transparency, Reason-Giving, and Agency Accountability." The Oxford Handbook of U.S. Education Law. December 11, 2018. https://doi.org/10.1093/oxfordhb/9780190697402.013.9.

Gordon, Nora, and Sarah Reber. 2015. "The Quest for a Targeted and Effective Title I ESEA: Challenges in Designing and Implementing Fiscal Compliance Rules." RSF: The Russell Sage Foundation Journal of the Social Sciences 1 (3): 129–47. https://doi.org/10.7758/RSF.2015.1.3.07.

———. 2020. "Supporting Students and Promoting Economic Recovery in the Time of COVID-19." Brookings (blog). May 7, 2020. https://www.brookings.edu/blog/brown-center-chalkboard/2020/05/07/supporting-students-and-promoting-economic-recovery-in-the-time-of-covid-19/.

Gray, Abigail M, Philip M Sirinides, Ryan Fink, and Adrianne Flack. 2017. "Discipline in Context: Suspension, Climate, and PBIS in the School District of Philadelphia," *CPRE Research Reports*, retrieved from https://repository.upenn.edu/cpre_researchreports/106.

Grissom, Jason A., and Christopher Redding. 2016. "Discretion and Disproportionality: Explaining the Underrepresentation of High-Achieving Students of Color in Gifted Programs." *AERA Open* 2 (1): 2332858415622175. https://doi.org/10.1177/2332858415622175.

Gronberg, Timothy J., Dennis W. Jansen, Mustafa U. Karakaplan, and Lori L. Taylor. 2015. “School District Consolidation: Market Concentration and the Scale-Efficiency Tradeoff.” *Southern Economic Journal* 82 (2): 580–97.

Hanford, Emily. 2018. “Why Aren’t Kids Being Taught to Read?” September 10, 2018. https://www.apmreports.org/episode/2018/09/10/hard-words-why-american-kids-arent-being-taught-to-read.

Hanushek, Eric A. 1997. “Assessing the Effects of School Resources on Student Performance: An Update.” *Educational Evaluation and Policy Analysis* 19 (2): 141–64. https://doi.org/10.3102/01623737019002141.

Hardy, Bradley L, Trevon D Logan, and John Parman. 2018. “The Historical Role of Race and Policy for Regional Inequality,” Brookings Institute, The Hamilton Project, 23.

Henry, Gary T., Lam D. Pham, Adam Kho, and Ron Zimmer. 2020. “Peeking Into the Black Box of School Turnaround: A Formal Test of Mediators and Suppressors.” *Educational Evaluation and Policy Analysis* 42 (2): 232–56. https://doi.org/10.3102/0162373720908600.

Jackson, C. Kirabo, and Claire Mackevicius. 2021. “The Distribution of School Spending Impacts.” National Bureau of Economic Research working paper #28517. https://doi.org/10.3386/w28517.

Jacob, Brian A, and Jonah E Rockoff. 2011. “Organizing Schools to Improve Student Achievement: Start Times, Grade Configurations, and Teacher Assignments,” Brookings Institute, The Hamilton Project, 28.

Jacob, Brian A., Jonah E. Rockoff, Eric S. Taylor, Benjamin Lindy, and Rachel Rosen. 2018. “Teacher Applicant Hiring and Teacher Performance: Evidence from DC Public Schools.” *Journal of Public Economics* 166 (October): 81–97. https://doi.org/10.1016/j.jpubeco.2018.08.011.

Johnson, Rucker C., and Alexander Nazaryan. 2019. *Children of the Dream: Why School Integration Works*. Basic Books.

Kaufman, Julia, John Davis, Elaine Wang, Lindsey Thompson, Joseph Pane, Katherine Pfrommer, and Mark Harris. 2017. *Use of Open Educational Resources in an Era of Common Standards: A Case Study on the Use of EngageNY*. RAND Corporation. https://doi.org/10.7249/RR1773.

Kearney, Melissa, and Phillip Levine. 2016. “Income Inequality, Social Mobility, and the Decision to Drop Out of High School.” *Brookings Papers on Economic Activity* 2016 (March): 333–96. https://doi.org/10.1353/eca.2016.0017.

Kho, Adam. 2021. “Evaluating Claims That Charter Schools Practice ‘Cream Skimming’ and ‘Pushout.’” *Education Next* (blog). February 24, 2021. https://www.educationnext.org/do-charter-schools-cream-skim-high-achievers-push-out-low-performing-students/.

Kraft, Matthew A., and Allison F. Gilmour. 2017. “Revisiting The Widget Effect: Teacher Evaluation Reforms and the Distribution of Teacher Effectiveness.” *Educational Researcher* 46 (5): 234–49. https://doi.org/10.3102/0013189X17718797.

Krueger, Alan B. 1999. “Experimental Estimates of Education Production Functions.” *The Quarterly Journal of Economics* 114 (2): 497–532.

Lafortune, Julien, Jesse Rothstein, and Diane Whitmore Schanzenbach. 2018. "School Finance Reform and the Distribution of Student Achievement." *American Economic Journal: Applied Economics* 10 (2): 1–26. https://doi.org/10.1257/app.20160567.

Lindsay, Constance A., and Cassandra M. D. Hart. 2017. "Exposure to Same-Race Teachers and Student Disciplinary Outcomes for Black Students in North Carolina." *Educational Evaluation and Policy Analysis* 39 (3): 485–510. https://doi.org/10.3102/0162373717693109.

Magnuson, K., and J. Waldfogel. 2008. *Steady Gains and Stalled Progress: Inequality and the Black-White Test Score Gap*. Book Collections on Project MUSE. Russell Sage Foundation. https://books.google.com/books?id=lNAEyVdBEQUC.

Mavrogordato, Madeline, Rebecca Callahan, David DeMatthews, and Elena Izquierdo. 2021. "SUPPORTS FOR STUDENTS WHO ARE ENGLISH LEARNERS," EdResearch for Recovery, Brief No. 15, 6.

Monarrez, Tomás. 2019. "Do Charter Schools Increase Segregation?" *Education Next* (blog). July 24, 2019. https://www.educationnext.org/do-charter-schools-increase-segregation-first-national-analysis-reveals-modest-impact/.

———. 2021. "School Attendance Boundaries and the Segregation of Public Schools in the US," 66.

Mulhern, Christine. 2020. "Beyond Teachers: Estimating Individual Guidance Counselors' Effects on Educational Attainment," 76.

Nelson, Charles A., Zulfiqar A. Bhutta, Nadine Burke Harris, Andrea Danese, and Muthanna Samara. 2020. "Adversity in Childhood Is Linked to Mental and Physical Health throughout Life." *BMJ* 371 (October): m3048. https://doi.org/10.1136/bmj.m3048.

Nomi, Takako. 2009. "The Effects of Within-Class Ability Grouping on Academic Achievement in Early Elementary Years." *Journal of Research on Educational Effectiveness* 3 (1): 56–92. https://doi.org/10.1080/19345740903277601.

Northern, Amber M., Janie Scull, and Dara Zeehandelaar Shaw. 2012. "How Strong Are U.S. Teacher Unions? A State-By-State Comparison." The Thomas B. Fordham Institute. Accessed May 29, 2021. https://fordhaminstitute.org/national/research/how-strong-are-us-teacher-unions-state-state-comparison.

Office of Management and Budget. 1997. "Revisions to the Standards for the Classification of Federal Data on Race and Ethnicity." Accessed May 30, 2021. https://obamawhitehouse.archives.gov/node/15626.

Park, R. Jisung, A. Patrick Behrer, and Joshua Goodman. 2021. "Learning Is Inhibited by Heat Exposure, Both Internationally and within the United States." *Nature Human Behaviour* 5 (1): 19–27. https://doi.org/10.1038/s41562-020-00959-9.

Park, R. Jisung, Joshua Goodman, Michael Hurwitz, and Jonathan Smith. 2020. "Heat and Learning." *American Economic Journal: Economic Policy* 12 (2): 306–39. https://doi.org/10.1257/pol.20180612.

Parolin, Zachary, and Emma K. Lee. 2021. "Large Socio-Economic, Geographic and Demographic Disparities Exist in Exposure to School Closures." *Nature Human Behaviour* 5 (4): 522–28. https://doi.org/10.1038/s41562-021-01087-8.

Polikoff, Morgan, Elaine Lin Wang, and Julia H. Kaufman. 2021. "The Overlooked Support Teachers Are Missing: A Coherent Curriculum." March 3, 2021. https://www.rand.org/blog/2021/03/the-overlooked-support-teachers-are-missing-a-coherent.html.

Pomerance, L, and K Walsh. 2020. "Clinical Practice and Classroom Management." https://www.nctq.org/review/docs/NCTQ%202020%20Teacher%20Prep%20Review_Clinical%20Practice%20and%20Classroom%20Management_Final_10.19.pdf.

Pope, Nolan G, and George W Zuo. 2020. "Suspending Suspensions: The Education Production Consequences of School Suspension Policies," 49.

Reardon, Sean F. 2019. "Educational Opportunity in Early and Middle Childhood: Using Full Population Administrative Data to Study Variation by Place and Age." *RSF: The Russell Sage Foundation Journal of the Social Sciences* 5 (2): 40–68. https://doi.org/10.7758/RSF.2019.5.2.03.

Reber, Sarah J. 2005. "Court-Ordered Desegregation Successes and Failures Integrating American Schools since Brown versus Board of Education." *Journal of Human Resources* XL (3): 559–90. https://doi.org/10.3368/jhr.XL.3.559.

———. 2010. "School Desegregation and Educational Attainment for Blacks." *Journal of Human Resources* 45 (4): 893–914. https://doi.org/10.3368/jhr.45.4.893.

———. 2011. "From Separate and Unequal to Integrated and Equal? School Desegregation and School Finance in Louisiana." *The Review of Economics and Statistics* 93 (2): 404–15. https://doi.org/10.1162/REST_a_00090.

Reckhow, Sarah, Megan Tompkins-Stange, and Sarah Galey-Horn. 2021. "How the Political Economy of Knowledge Production Shapes Education Policy: The Case of Teacher Evaluation in Federal Policy Discourse." *Educational Evaluation and Policy Analysis*, April, 01623737211003906. https://doi.org/10.3102/01623737211003906.

Ritter, Gary. 2010. "A Closer Look at Charter Schools and Segregation." *Education Next* (blog). April 27, 2010. https://www.educationnext.org/a-closer-look-at-charter-schools-and-segregation/.

Ronfeldt, Matthew, Stacey L. Brockman, and Shanyce L. Campbell. 2018. "Does Cooperating Teachers' Instructional Effectiveness Improve Preservice Teachers' Future Performance?" *Educational Researcher* 47 (7): 405–18. https://doi.org/10.3102/0013189X18782906.

Ronfeldt, Matthew, Susanna Loeb, and James Wyckoff. 2013. "How Teacher Turnover Harms Student Achievement." *American Educational Research Journal* 50 (1): 4–36. https://doi.org/10.3102/0002831212463813.

Roza, Marguerite, Katherine Hagan, and Laura Anderson. 2021. "Variation Is the Norm: A Landscape Analysis of Weighted Student Funding Implementation." *Public Budgeting & Finance* 41 (1): 3–25. https://doi.org/10.1111/pbaf.12276.

Sacerdote, Bruce. 2011. "Chapter 4 - Peer Effects in Education: How Might They Work, How Big Are They and How Much Do We Know Thus Far?" In *Handbook of the Economics of Education*, edited by Eric A. Hanushek, Stephen Machin, and Ludger Woessmann, 3:249–77. Elsevier. https://doi.org/10.1016/B978-0-444-53429-3.00004-1.

Schrag, Peter. 2006. "Policy from the Hip: Class-Size Reduction in California." *Brookings Papers on Education Policy*, no. 9: 229–43.

Schwartz, Amy Ellen, Jeffrey Zabel, and Michele Leardo. 2017. "Class Size and Resource Allocation: ESE Policy Brief," Massachusetts Department of Elementary and Secondary Education.

Setren, Elizabeth. Forthcoming. "Targeted vs. General Education Investments: Evidence from Special Education and English Language Learners in Boston's Charter Schools." *Journal of Human Resources,* forthcoming.

Shores, Kenneth A., Christopher A. Candelaria, and Sarah E. Kabourek. 2020. "Spending More on the Poor? A Comprehensive Summary of State-Specific Responses to School Finance Reforms from 1990–2014." *EdWorkingPapers.Com*. Annenberg Institute at Brown University. https://www.edworkingpapers.com/ai19-52.

Staiger, Douglas O, and Jonah E Rockoff. 2010. "Searching for Effective Teachers with Imperfect Information." *Journal of Economic Perspectives* 24 (3): 97–118. https://doi.org/10.1257/jep.24.3.97.

Stafford, Tess M. 2015. "Indoor air quality and academic performance," *Journal of Environmental Economics and Management* 70 (March): 34–50.

Stoddard, Christiana, and Eugenia F. Toma. 2021. "Introduction to Special Topic: Rural Education Finance and Policy." *AERA Open* 7 (January): 23328584211011610. https://doi.org/10.1177/23328584211011607.

Taylor, Eric S, and John H Tyler. 2012. "The Effect of Evaluation on Teacher Performance." *American Economic Review* 102 (7): 3628–51. https://doi.org/10.1257/aer.102.7.3628.

Thompson, Owen. 2020. "School Desegregation and Black Teacher Employment." *The Review of Economics and Statistics*, October, 1–45. https://doi.org/10.1162/rest_a_00984.

Tuttle, Cody. 2019. "The Long-Run Economic Effects of School Desegregation." SSRN Scholarly Paper ID 3460993. Rochester, NY: Social Science Research Network. https://papers.ssrn.com/abstract=3460993.

Whitehurst, Grover J. "Russ." 2009. "Don't Forget Curriculum." *Brookings* (blog). October 14, 2009. https://www.brookings.edu/research/dont-forget-curriculum/.

Zilbermints, Regina. 2021. "White House Defends CDC Outreach to Teachers Union." TheHill. May 4, 2021. https://thehill.com/homenews/administration/551743-white-house-defends-cdc-outreach-to-teachers-union-in-developing.

America and International Trade Cooperation

AUTHORS

Chad P. Bown*

ABSTRACT

The Biden administration inherits a U.S. trade policy in transition. This chapter provides a factual and contextual assessment of that transition as well as a normative set of U.S. trade policy recommendations. It starts by allocating recent changes in U.S. trade policy into one of two categories. Some are "noncooperative" trade policy actions, driven by the perception that other countries were not following agreed-upon rules, so no longer should the United States. Others are tweaks to "cooperative" policy and indicate continued U.S. adherence to existing trade rules, but with policy modifications demanded by changes in underlying domestic economic, social, and national security preferences. China was the driver behind many U.S. policy changes of both type; however, only some arose from the perception that Chinese policies were noncooperative. Other U.S. changes appear motivated by the combination of these new U.S. preferences with American dissatisfaction over how supply chains relocated globally, partly as a result of China's integration into the global economy. A final set are not focused on China at all, but rather on trade policy changes due to cooperation on corporate tax reform, climate mitigation policy, and COVID-19 vaccines. After providing a descriptive analysis of these changes, the chapter then proposes a normative set of recommendations for how the Biden administration can implement its stated approach of a "worker-centered" trade policy with a commitment to work with allies to resolve bilateral issues, to work with allies on common concerns involving China, and to work with allies and China on global challenges.

* Peterson Institute for International Economics & Center for Economic Policy Research; 1750 Massachusetts Avenue, NW, Washington, DC 20036. Tel: +1.202.454.1306. Email: cbown@piie.com, web: www.chadpbown.com.

Acknowledgements: For helpful comments and conversations, thanks to Kathleen Claussen, Amy Ganz, Melissa Kearney, Soumaya Keynes, Kadee Russ, Matthew Slaughter, Robert Staiger, Robert Zoellick, and participants at the Aspen Economic Strategy Group 2021 Annual Meeting. All views and any remaining errors are my own.

1. Introduction

As of late 2021, the path of U.S. trade policy remained uncertain. Early signs from the Biden administration suggest the United States is in a transition period. The trade war that the Trump administration initiated with China in 2018 remained mostly unresolved, with bilateral ties having further soured, rendering future trade cooperation even less appealing politically. But there are other issues at play. First, the pandemic has presented pressing issues for international cooperation, including how to accelerate manufacturing and trade in COVID-19 vaccines to prioritize global public health. Second, after four years of U.S. neglect, tackling climate change is likely to be reasserted as a priority during Biden's presidency, also with potential trade implications. And third, domestic politics after a tight election remain paramount, with the new administration signaling a commitment to a "worker-centered" trade policy.

Despite the flux, and even after candidate Joe Biden campaigned against the Trump administration's trade strategy, one result was clear—there was unlikely to be a simple reversal of the Trump administration's trade policies. Many are likely to linger, as the new administration focuses its international engagement deliberately and modestly.

The Biden administration's non-reversal is consequential because the Trump administration made significant changes both to U.S. trade policy and to the multilateral, rules-based trading system. First, the Trump administration began to impose tariffs on China in July 2018, leading to retaliation and a trade war, resulting in a temporary and uneasy truce—even the terms of the Phase One agreement in effect since February 2020 mean that United States and China imposed permanently higher discriminatory tariffs on one another. In that sense, both were flouting the most basic rule of most-favored nation (MFN) treatment, a pillar of the multilateral trading system's World Trade Organization (WTO). Second, beginning in March 2018 and under the guise of protecting America's national security, the administration imposed other duties on tens of billions of dollars of steel and aluminum imports from countries aside from China, including mostly military and economic allies, that also remain in effect. Third, the U.S. administration ended the WTO's 25-year-old system of dispute resolution, meaning there is no longer a universally agreed upon way to resolve the inevitable trade frictions that arise between any of the WTO's 164 members, let alone something involving the United States. Finally, the Trump administration unilaterally implemented a series of export controls on the semiconductor supply chain—seeking to cut off Chinese access for national security reasons—but also with considerable commercial implications for firms operating in key economic and military allies.

To make sense of contemporary U.S. trade policy—and what the Biden administration is inheriting and likely continuing to some degree—this chapter establishes a simple analytical framework in order to characterize these and other recent actions. While U.S. policy is in flux, one question is whether each of these and other potentially forthcoming policy changes are "cooperative" and can take place with minor tweaks—and thus *within* the existing international trade rules—or whether the U.S. changes are so dramatic that they are "noncooperative" and require abandoning the system that has been in place since 1947 and negotiating (and adhering to) completely new rules.

The answers to that question are not uniform with regard to various elements of current U.S. trade policy. Some U.S. policy changes are tweaks and covered by WTO exceptions, others are inconsistent with the existing rulebook. However, while the need for "new rules" is a persistent slogan for trade negotiators, this time is different in one way: U.S. administrations appear no longer willing to constrain themselves by old rules until the new ones are in place. That being said, new rules cannot be determined by the United States alone. They are ultimately negotiated, with the outcome of those negotiations depending on China, as well as other countries. By deciding to no longer follow certain rules, the United States is finding that others will follow suit.

The second part of this chapter provides a more normative set of proposals for U.S. policy. These are themselves informed by the first part—*e.g.*, conditional on U.S. policymakers having decided to impose a noncooperative policy in a given context or to tweak its cooperative policy, what should that policy look like to achieve its (potentially noneconomic) objective at the least economic cost? Importantly, these proposals reflect a realism in the shifting policy environment that has emerged, albeit rather suddenly in the United States. Relative to their most recent predecessors, today's trade policymakers may be prioritizing political, social, environmental, and even national security objectives over economic efficiency. While economics remains critical in helping policymakers understand the trade-offs associated with different policy choices, the discourse is no longer between free trade and protectionism.

2. Framing America's trade policy re-evaluation

Consider the workhorse economic model of international trade agreements. Trade agreements are valuable because they solve what is known, in game theoretic terms, as the prisoner's dilemma.[1] (See Figure 1.)

In such a game, each player has two choices—"cooperate" or "do not cooperate." (The values in each box are the payoffs to each player if that is where they jointly end up.) To start, suppose there is no coordination between the players, so that each

1 These theoretical models were developed in a series of research beginning with Bagwell and Staiger (1999, 2002); see the survey of Bagwell, Bown, and Staiger (2016) which also reviews empirical evidence.

chooses its best response. The equilibrium outcome will be that each chooses "do not cooperate," and the payoff to each is 1. But the problem with this outcome is obvious. Even though neither of them has a unilateral incentive to change its behavior, if they both agreed to, each can be made better off and receive a payoff of 3.

Figure 1: The prisoner's dilemma of trade policy

● U.S. ● TRADING PARTNER

	COOPERATE	DO NOT COOPERATE
COOPERATE	3 , 3	0 , 5
DO NOT COOPERATE	5 , 0	1 , 1

Moving closer to the real world, imagine the two players are now countries, and the game they are playing is whether or not to cooperate over trade policy. Each country is a 'large' consumer of imports from the other, with market power. In the absence of cooperation, each would impose beggar-thy-neighbor policies like tariffs. For a large country, a small tariff can make itself slightly better off but only by making the other country much worse off. In a nutshell, trade agreements involve two countries jointly agreeing to cooperate by tying their hands and refraining from imposing tariff policies that are unilaterally optimal, but jointly suboptimal. Trade agreements are mutually beneficial because they stop the international cost-shifting externalities that arise when countries do not consider the negative impacts that their policies impose on those outside their borders.

Broadly speaking, these prisoner's dilemma models can be used to characterize the WTO and its core rules. Furthermore, more complex versions of these sorts of political-economic models of trade agreements do not necessarily result in "free" trade (zero tariffs) as being needed to sustain an equilibrium of cooperation. Equilibrium tariffs may arise when governments have political preferences, for example, that put higher weight on the wellbeing of one set of societal groups relative to others.[2] Alternatively, domestic constraints may leave a government without access to more efficient policy instruments—such as subsidies, to deal with market failures or local externalities—implying that tariffs may be a second best policy.

2 A straightforward representation is to take a two-sector model and a government objective function—which includes consumer surplus, producer surplus, and government revenue—and allow the government to have political preferences that weight the producer surplus of the import-competing sector with $\gamma>1$, while all other elements of the objective function receive a weight of 1.

A second basic feature of WTO rules is that they do not prevent countries from adjusting their cooperative policies—including tariffs—when certain shocks occur or if a government's preferences change. Suppose, for example, there is a sudden need to offer higher protection to some sector for redistributive reasons or to implement a policy to address a newly identified externality or market failure. WTO rules allow governments this flexibility, subject to two main caveats. The first is that such a policy change should be applied on a relatively nondiscriminatory basis. If it is a trade policy, it should follow the MFN principle and apply to all WTO members equally. If it is a domestic regulatory policy, it should follow another WTO principle—referred to as national treatment—and apply equally to local and foreign firms. The second caveat is that, if the policy change reduces previously agreed levels of market access, the government has to be prepared to "pay" for it. That is, adversely affected trading partners have a right to seek compensation so as to rebalance the benefits of the overall agreement—often that may mean an authorized, but limited, form of tariff retaliation.

Through the lens of this standard political-economic model of trade agreements, U.S. trade policy, as well as that of its key trading partners, was viewed for decades under the WTO as being "cooperative."[3] The starting point question for today is whether a variety of U.S. trade policy actions—some already taken, others under consideration—mark the United States as making a more dramatic shift toward reimplementing *noncooperative* policy. There are not many clear cut examples of the latter; one is the 1930 Smoot-Hawley Tariff Act, before the advent of the WTO's predecessor, the General Agreement on Tariffs and Trade (GATT). Another might be in the 1980s when U.S. "aggressive unilateralism" on trade policy toward a variety of trading partners resulted from its dissatisfaction with the GATT (Bhagwati and Hudec 1990). (Some credit the 1980s U.S. outbursts as helping convince trading partners to commit to deeper forms of international cooperation, resulting in the formation of the WTO in 1995.)

Today, and at one extreme, there is the possibility that policymakers beginning with the Trump administration perceive China as having implemented noncooperative policy. In this case, they may be shifting U.S. policy toward China away from the standard cooperative policy applied since granting Chinese exporters its low, MFN tariffs on a provisional basis starting in 1980 and that was locked in with China's WTO accession in 2001.[4] A noncooperative shift could imply that getting back to

3 For empirical evidence, see Broda, Limão, and Weinstein (2008), Ludema and Mayda (2013), and Bown and Crowley (2013).

4 For a review of U.S. trade policy toward China over 1980–2018, see Bown (2019a). For China and the WTO, see Wu (2016, 2020).

cooperation requires a major change in behavior by trading partners (China in our example), possibly accompanied by negotiating new trade agreement rules.

At the other extreme sits the possibility that the United States still seeks to implement cooperative policies—but it needs to update them in light of changes in the underlying (domestic) social, political, and national security environment. Such a situation would not necessarily involve a major rewrite of WTO rules, and there is no allegation that trading partners—including China—have implemented a noncooperative economic policy. Yet, even tweaks that the United States would like to make to its cooperative trade policy may come at a price, requiring some negotiations and potentially compensation to trading partners adversely affected by the changes.

Which U.S. policy changes are noncooperative versus tweaks to cooperative policy? The next subsections describe some of the major U.S. trade policy developments and attempts to allocate each into the appropriate category.[5]

China is the motivation for multiple changes in U.S. policy that have been enacted since 2018. The desire to treat China differently—relative to other trading partners and relative to how the U.S. treated China in the past—currently enjoys bipartisan support in the United States. Yet, how the United States will apply its trade policy toward China is still to be fully determined.

Before proceeding, it is also worth clarifying that much of the current policy reaction to China does not appear to be an attempt to reverse the so-called "China shock" to the U.S. labor market. The lack of American labor market and community-level adjustment resulting from the integration of the massive Chinese economy into the global trading system in the first decade of the 2000s has been the source of much debate over the last 10 years. The evidence suggests that *was* a real economic shock, and its labor market and local community implications were arguably mismanaged by U.S. policymakers who failed to deploy the complementary package of *domestic* policy initiatives to ensure mobility of workers and adjustment of communities that bore the brunt of the shock at the time.[6] However, reapplication of tariffs against China is highly unlikely to do much to remedy the suffering that continues to impact those workers and communities. (Some of the Biden administration's domestic policy agenda may address ongoing challenges introduced by that shock.) Thus, this

5 There is third category, described and analyzed by Mattoo and Staiger (2020) and referred to as "bargaining tariffs," that motivates the United States as imposing potentially noncooperative policy even though China may not have imposed its own noncooperative policy. *I.e.*, the Trump administration motivated some of its tariff increases as a temporary strategy to induce countries (with higher tariffs) to reduce their tariffs toward the United States (Ross 2017).

6 See Autor, Dorn and Hanson (2013) and the subsequent research literature, including surveys and updates in Autor, Dorn and Hanson (2016, 2021).

chapter evaluates the contemporary U.S. policy response to China as driven by other current and forward-looking reasons.

2.a. China and noncooperative U.S. policy

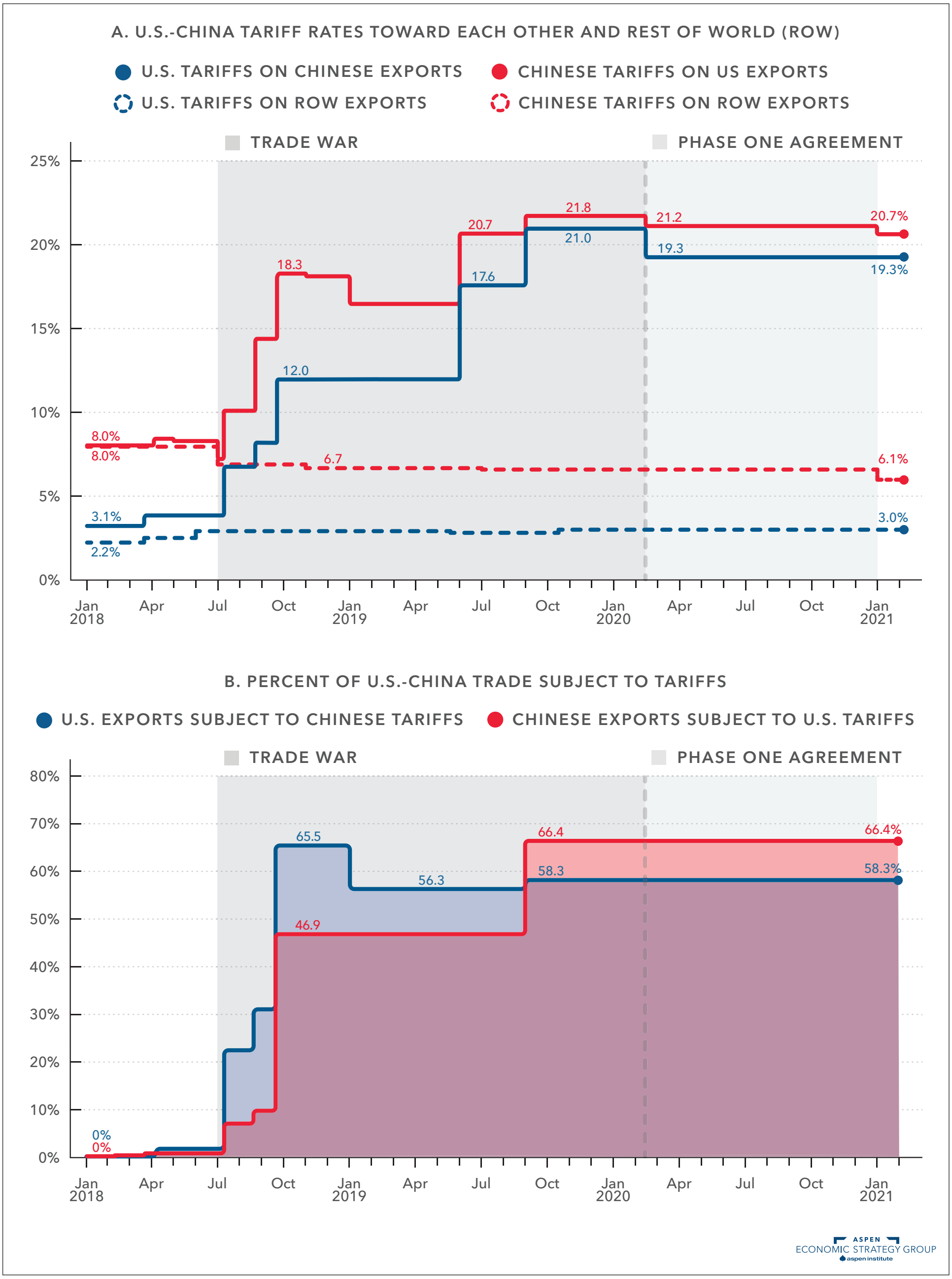
This section explores two questions. Prior to the trade war, was China playing its policy noncooperatively? Has the United States decided it must also shift its policy to do the same?

Historically, the United States has pushed for relatively low tariffs, applied on a nondiscriminatory basis to all members of the WTO.[7] One interpretation of the Trump administration's tariff war is the following. Even nearly two decades after its 2001 WTO accession, China had refused to engage in additional tariff liberalization. It was deploying other policies in ways symptomatic of noncooperative play, imposing costly externalities on trading partners. Thus, the United States imposed trade war tariffs as its best response; as a result, each country is now imposing its noncooperative policy on the other. (Both are economically worse off than if they agreed to cooperate—see again Figure 1—but the United States may now be better than off than it was when it was cooperating but China was not.)

Specifically, in 2018–2019, the United States increased tariffs considerably toward imports from China. U.S. average tariffs toward China increased from roughly 3% in early 2018, to over 19% by the end of the trade war. China responded by raising tariffs on U.S. exporters (from 8% to 20%), as well as *lowering* its applied MFN tariffs on imports from the rest of the world (from 8% to just above 6%). Most of the higher tariffs remain in place, despite the U.S.-China Phase One agreement implemented in February 2020. Another implication is that the United States imposes much higher tariffs on imports from China, on average, than it applies to imports from countries in the rest of the world (Figure 2). Through this lens, China and the United States are not cooperating with one another, but each is still implementing cooperative policy toward the rest of the world.

7 It has also negotiated some free trade agreements offering preferentially lower tariffs to a handful of countries—the most important being Canada and Mexico (NAFTA and now USMCA), and South Korea—but most U.S. partners and U.S. trade remains conducted under non-FTA (and thus WTO) rules.

Figure 2: U.S.-China trade war tariffs



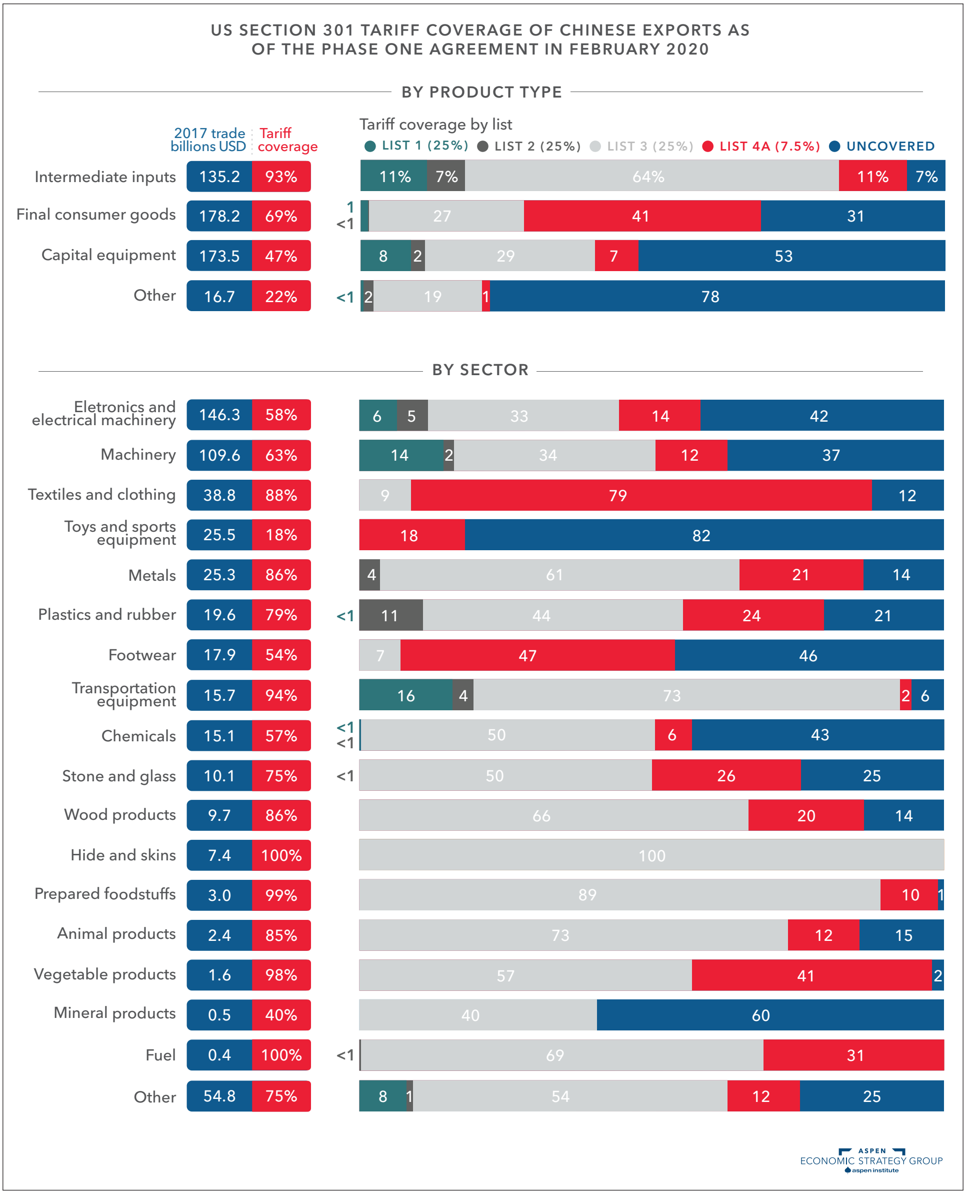
Source: Bown (2021a), trade-weighted by trading partner exports to the world at the six-digit HS level.

Consider first whether the United States is now imposing noncooperative tariffs toward China. While the average level of duties has certainly increased, there is little empirical evidence to date that the particular tariff profile—*i.e.,* products selected, rates chosen—that the United States implemented were in any sense "optimal," let alone better than the tariffs applied prior to the trade war. For example, despite President Trump's repeated assertion that China was paying for the tariffs, there is not yet evidence that the United States effectively exploited its market power by driving down the price received by Chinese exports (Amiti, Redding, and Weinstein 2019, Fajgelbaum et al. 2019, Cavallo et al. 2021).[8] Furthermore, while research generally concludes the tariffs increased U.S. product prices, there is not yet evidence that the tariffs led to employment increases even in import-protected sectors (Flaaen and Pierce, 2019).

It is worth noting at least two other economic implications of the particular U.S. tariffs chosen. First, the U.S. tariffs applied mostly to imports of intermediate inputs, as opposed to final consumer goods (Figure 3). The resulting higher input costs imply American downstream firms are at a disadvantage relative to their competitors in foreign markets—both for selling to Americans (Flaaen and Pierce, 2019) and globally as exporters (Handley, Kamal, and Monarch 2020). Second, the tariffs create an incentive for American businesses to source those inputs from countries other than China. This is consistent with desires to diversify certain economic activity out of China for either economic or noneconomic reasons. (More on this below.)

8 However, in their general equilibrium, quantitative model, Fajgelbaum et al. (2019) do find that there are terms-of-trade trade changes impacting the U.S. economy operating through the channel of relative wage or other factor-price changes.

Figure 3: U.S. tariffs on China disproportionately targeted imported inputs used by American businesses and workers



Source: Bown (2021a)

Consider next the question of whether China was imposing its policies noncooperatively in the first place. Here it is worth considering four different policy instruments.

Start with import tariffs. The fact that China applied higher tariffs than the United States prior to the trade war and that it had not significantly reduced its import tariffs further since its 2001 WTO accession is not sufficient evidence of noncooperative policy behavior. For example, evidence from the new tariffs resulting from China's WTO accession (Bagwell and Staiger 2011; Bown 2019b) was consistent with the negotiations having expunged China's ability at the time to shift the costs of its tariffs onto foreign exporters by forcing them to lower their prices. Admittedly, such analyses cannot speak to whether the situation had changed fifteen years later; perhaps China had increased its market power through industrialization and economic growth since 2001. However, even if China's tariffs had come to exert market power by 2016—implying the WTO could play an efficiency-enhancing role by facilitating China's engagement in reciprocal tariff liberalization negotiations with someone—the failure of such negotiations to materialize is also not necessarily evidence of China engaging in noncooperative behavior. The United States, for example, may have been unlikely to take on its historical role in leading such negotiations due to what is referred to as the "latecomers problem." By 2016, U.S. tariffs were already so low—resulting from multiple rounds of successful negotiations with other countries taking place since the 1940s, subsequently extended to China through its WTO membership and application of the MFN principle—that there was not much the United States could offer to engage China in further tariff liberalization.[9] Overall, whether China's currently applied MFN tariffs continue to exert market power—especially given the significant tariff reductions applied in 2018 and 2019 during the trade war—remains an open research question.

However, China had a number of other policies in place prior to the trade war that may have imposed other international externalities, and thus been closer to noncooperative behavior, aside from tariffs. One allegation involves shifting of rents from intellectual property rights—*e.g.*, the "forced technology transfer" arguments in the Section 301 reports (USTR 2018a, 2018b) that were the legal justification for the Trump administration's trade war tariffs.[10] The U.S. claim was that the Chinese government created an economic system that resulted in foreign companies having to share their technology with local firms involuntarily and at less-than-market rates, that this constituted "unfair trade," and thus the United States could respond

9 For discussions, see Bagwell and Staiger (2014) and Staiger (Forthcoming).

10 For a discussion, see Mavroidis and Sapir (2021).

unilaterally with tariffs.[11] Beijing's high tariffs for certain products meant that, in order to sell to Chinese consumers, foreign firms would have to produce locally. However, in order to produce locally, the Chinese government demanded foreign firms form joint ventures with local and often state-owned firms. Such relationships would require the sharing of technology at less than commercial terms. Finally, the Chinese government had not only failed to protect the intellectual property of foreign firms, but Beijing often set up institutional, regulatory and state-sponsored arrangements seemingly to facilitate its theft or expropriation.

A third example of potential noncooperative Chinese policy could be its complex system of subsidies. This includes not only the increased role played by China's state-owned enterprises (SOEs), but the economic subsidies that result through its deployment of other policies. For example, Chinese export restrictions on upstream (primary aluminum) products subsidize downstream (manufactured aluminum) products at the expense of downstream competitors in the rest of the world (OECD 2019a). Another was China's use of below market debt and equity to subsidize its domestic semiconductor industry (OECD 2019b). Even though China may have never promised to become a market-oriented economy (Wu, 2016), its more dramatic shift toward state-orientation under President Xi triggered alarms internationally.[12] While, under certain conditions, both the United States and China could be made better off by cooperating on government policy and agreeing to restrain subsidies, the lack of coordination—*e.g.,* China's refusal to stop subsidizing since its WTO accession, its increased support to state-owned enterprises after 2013, its 'Made in China 2025' industrial policy, etc.—may result in the United States also eventually abandoning its initial position of implementing a cooperative (non-subsidy) policy so that it also starts subsidizing domestic production.[13]

China's refusal to halt its subsidies could partially explain the sudden shift in U.S. policy in the semiconductor sector, where the United States has pivoted from concerns over other countries' subsidies to embracing legislation that may result in tens of billions of dollars of subsidies for the semiconductor supply chain.[14] Commercial aircraft may be another example of the United States changing its policy position on subsidies. The negotiated settlement to the recent E.U.-U.S. disputes at

---

11 The Trump administration argued that because such actions were not covered by WTO rules, filing a WTO dispute against China over the issue would be fruitless. Others have disagreed, arguing that the United States could have filed what is known as a "nonviolation complaint" at the WTO, claiming that China's policies still harmed American firms even without breaking any specific WTO rules. For discussions, see Hillman (2018) and Staiger (Forthcoming). For nonviolation complaints more generally, see Staiger and Sykes (2013).

12 For an assessment of the role of the changing role of the state in the Chinese economy, see Lardy (2014, 2019).

13 See, for example, Brander and Spencer (1985) for an early illustration of international subsidies in imperfectly competitive markets. In cooperative markets with political economy forces, see Bagwell and Staiger (2001).

14 See Bown (2020a), Varas et al. (2021), OECD (2019b), and Ip (2021).

the WTO regarding subsidies to Airbus and Boeing may have sought to resolve their differences so that they can both subsidize to compete with China's state-backed emerging industry.

A final example of Chinese noncooperative policy includes its use of export restrictions to take advantage of market power it possesses on the supply side. Historical examples include a series of export restrictions on rare earth minerals and raw materials that resulted in formal WTO disputes showing the inconsistency of Beijing's policies with its multilateral legal commitments. Indeed, foreign concerns over Chinese subsidies are often motivated by fears that China could become such a large player in a given sector that it would then be able to use export restrictions to exploit its market power at the expense of foreign consumers or consuming industries.

2.b. China and cooperative U.S. policy

In addition to the noncooperative scenarios described above, it is possible the United States could seek to adjust other parts of its trade policy toward China while maintaining a cooperative approach. This could be caused by changes to U.S. domestic preferences, emergence of some externality or market failure, or some other shock, not as a response to China's (perceived or real) noncooperative decisions. In such instances, the United States would like to *tweak* its policy, but cooperatively and in a manner that is broadly consistent with WTO rules. Put differently, some U.S. policy changes are not necessarily because of any perception that China is implementing a noncooperative *economic* policy.

One motivation for trade policy tweaks in a cooperative scenario could be American dissatisfaction with how economic activity has been reallocated globally over the 30 years since China's entry into the trading system. The sources of this change may be innocuous—*i.e.,* purely economic and technological, and not the result of any noncooperative Chinese policy. For example, the removal of trade barriers allowed comparative advantage to flourish. Inventions such as containerization, automation, and the information and communications technology revolution resulted in the fragmentation of production, emergence of cross-border supply chains, and certain economic activity concentrating geographically in Asia. The problem is that sourcing certain specific products primarily *from China* has resulted in an outcome that may no longer be economically, socially, or geopolitically desirable.

Take, for example, public health. China emerged as the concentrated source of residual foreign supply of certain medical gear during the pandemic. Global shortages of personal protective equipment (PPE) created a political firestorm in early 2020, including in the United States. The Trump and then Biden administrations ultimately responded by first imposing export restrictions and then providing over $1 billion of subsidies and industrial policy targeting the domestic supply chain for PPE—both outputs (*e.g.,* N95 respirators, hospital gowns, rubber gloves) and inputs (*e.g.,* melt blown fiber, filters, rubber)—to expand U.S. production capacity.[15] Some sort of quasi-permanent policy intervention may be required if the United States seeks to maintain preparedness and surge capacity once market conditions normalize—*i.e.*, the pandemic is resolved—or if it seeks to diversify foreign sourcing away from China.

National security is another important noneconomic example, especially given heightened awareness that China under President Xi Jinping seems to pose a more serious geopolitical threat. In this case, there are certain technologies that generate negative externalities to the United States merely by being exported to China, and thus export controls are arguably first-best policies. Without commenting on the national security threat posed by any particular technology, examples of such recent U.S. export controls include semiconductors and equipment meant to address the national security threat posed by companies such as Semiconductor Manufacturing International Corporation (SMIC) and Huawei (Bown 2020a,b) that are alleged to have ties to the Chinese military.

Other noneconomic examples include human rights and democracy. It may be that China is acting against American and western "values"—*e.g.,* any consumption of certain goods produced in China generates negative externalities to Americans. China's mistreatment of Uyghurs in Xinjiang, including allegations of forced labor, has resulted in the United States imposing "withhold release orders" (WROs, import bans) on certain products (Hendrix and Noland 2021). Beijing's suppression of democracy in Hong Kong has led the United States to reclassify the city as being equivalent to China for customs purposes. Although Hong Kong used to be a separate customs territory under U.S. trade law, imports from Hong Kong now face the trade war tariffs and other special U.S. duties imposed on imports from China.

Each of these are plausibly interpreted as the United States exercising "cooperative" trade policy, acting within the exceptions permitted by WTO rules.

A final motivation could be the United States adapting and learning from the Chinese model to potentially improve U.S. policy. For example, closer ties between China and its businesses during the recent pandemic may have been a contributing reason why China was able to more quickly scale up its "surge capacity" for PPE, relative to the United States, whose response to the shortage was less nimble despite having more advanced warning (Bown, Forthcoming).

15 For a discussion of the U.S. policy response to PPE shortages in 2020, see Bown (Forthcoming).

While the United States is unlikely to shift its market capitalism model to one that is more accommodative of state-owned enterprises, it may seek ways of developing closer ties with firms in industries that are critical for national security or public health—*e.g.,* through subsidies to maintain surge capacity or some basic market participation.[16] Especially during the pandemic, the United States has also shown an increased willingness to deploy the Defense Production Act to have firms reallocate resources toward government orders and priorities relative to the private sector that may reflect socially beneficial outcomes (positive externalities) and not simply market incentives (Bown and Bollyky, Forthcoming).[17] It is unclear whether this is a purely emergency phenomenon or a harbinger of a longer-term trend, but the evidence of increased use of U.S. export controls in 2019 in other sectors suggests it was not unique to the pandemic.

2.c. Foreign, non-China sources of changing U.S. preferences with regard to trade policy

There are other changes afoot in American domestic preferences for trade policy, many of which have less to do with China, but nevertheless have implications for international cooperation on trade.

2.c.1. Climate

The Biden administration has rejoined the Paris Accord and indicated combatting climate change is a policy priority (Tai 2021). This could have implications for U.S. trade policy in several ways.

Although proposals for a domestic carbon tax remain politically unpopular, the Biden administration could attempt to mimic one through a combination of other regulatory and subsidy policies. This combination of policies would then raise concerns about "carbon leakage," that is, subjecting only domestic industries to the tax, which would create an incentive for carbon-intensive activity to relocate to countries without one. The resulting imports would be unsustainable for political and economic reasons, and would undermine the policy goal of mitigating climate change.

A carbon border tax, also known as a border carbon adjustment mechanism (CBAM), could address this concern by applying the tax based on the carbon content of the import and whether it had been taxed abroad. In this sense, a CBAM has the appearance of an import tariff. While feasible in theory, estimating the carbon

16 The United States imposed tariffs on imported steel and aluminum beginning in 2018 out of the threat that imports imperiled national security. While the overall argument has been discredited, for specialized niche products—*e.g.,* aluminum needed for military needs—subsidies would be more efficient than tariffs.

17 On the other hand, the Emergent BioSolutions story shows the difficulty of maintain idle surge capacity (Stolberg, LaFraniere, and Hamby 2021).

content of imports is a practical challenge. Given the complexity, there is the additional concern that the policy could be abused (subject to regulatory capture) by special interests. Finally, if it were applied unilaterally, it could become subject to foreign retaliation, which could also reduce its sustainability as a viable policy.

A CBAM is on the agenda of the Green Deal of the von der Leyen Commission in the European Union.[18] In July, the European Commission issued a CBAM proposal that would include an import tax on carbon-intensive industries like steel, aluminum, cement and fertilizers. It is thus important for the United States to familiarize itself with this policy instrument as it may emerge elsewhere first, with the potential for other countries' CBAM to hit U.S. exporters if America's policymakers fail to price carbon emissions and get U.S. industries to internalize its societal costs.

The United States is also contemplating legislation that could include major investments in domestic infrastructure, including hundreds of billions of dollars in spending on transit (roads and bridges), rail, and electrification of vehicles (White House 2021). Some of this public spending may tilt away from subsidizing old industries (fossil fuels) and toward new industries (electric batteries, clean energy) to shift incentives both to new priorities and tackle market failures and externalities.

At this stage it is also unclear if adopting such subsidies would fall within the confines of existing trade rules, or if the United States would need to negotiate new rules to accommodate such subsidies and encourage other governments to do the same. The failure to agree internationally means that the current rules may permit foreign retaliation as compensation if U.S. subsidies impose adverse effects on trading partner industries.

2.c.2. Tariffs on steel and aluminum

The United States has been imposing higher tariffs on imported steel and aluminum products since March 2018. Other, previously imposed U.S. tariffs had mostly halted imports of such products from China before the 2018 actions, leading the United States to import from other sources (trade diversion), due to the relative homogeneity of each metal. Most of the new trade restrictions in 2018 thus hit imports from economic and military allies such as the European Union and Japan, even though they did not do anything "wrong"—the underlying policy concern was China's alleged subsidization of its industry.[19]

---

18 See European Commission (2021), Keynes (2021) and Bown and Keynes (2021).

19 Canada and Mexico were important early targets hit as well, but they negotiated voluntary export restraints in May 2019 in exchange for removal of their tariff retaliation as part of the deal to get the USMCA—the renegotiated NAFTA—to pass Congress.

This caused problems for U.S. relations with military allies, with many imposing retaliatory tariffs that hurt U.S. exporters in other sectors. The U.S. tariffs also make it harder for other American businesses to compete with firms in Japan, Europe, or elsewhere that do not need to pay higher input costs. Because the tariffs were imposed under the guise of protecting U.S. national security, and they have been disputed at the WTO, they also place the multilateral institution in the untenable position of having to rule upon whether a country's policy is in response to a legitimate national security threat.

Finally, the steel and aluminum tariffs have done little to address the underlying economic concern. There has been no international engagement by the United States or other countries with China on the underlying issue of its subsidies to the steel and aluminum industries.

2.c.3. Taxation of multinational corporations, especially digital companies

Taxation of multinational companies is a major area of political concern in the United States and elsewhere, and it has turned into one of international concern as it threatens to imperil trade cooperation (Treasury 2021). The failure of multilateral progress at the OECD had led a number of major economies, beginning with France, to impose Digital Services Taxes (DSTs) designed to have an equivalent economic effect to imposing a tariff on American high-tech companies, including Google, Apple, Facebook and Amazon (Hufbauer and Lu 2018). The DSTs led the United States to conduct investigations under Section 301 of the Trade Act of 1974 that could result in U.S. tariff retaliation against European countries.

In June, the G7 economies announced a framework agreement that when agreed by the larger group of G20 economies in July. Given this progress, the United States has suspended its retaliatory tariffs against the United Kingdom, Italy, Spain, Turkey, India, and Austria, after doing similarly with retaliatory tariffs against France in January.

2.c.4. COVID-19 and global public health

The COVID-19 pandemic and global public health is another area which requires global cooperation and highlights the importance of trade. The proliferation of the disease globally and the eruption of additional variants implies no one is safe until everyone is safe. But given the complexity of inventing, developing, and manufacturing vaccines, most countries will not be able to produce them locally, leaving international trade as critical to addressing the public health crisis. The failure to develop and deploy an explicit framework for international trade and equitable sharing in vaccines led to accusations of hoarding of vaccine-making equipment and raw materials, and to demands to waive patent protection for vaccines.

In June, the G7 announced a plan to "vaccinate the world." However, while the early commitments involved donating hundreds of millions of doses of vaccines, it did not yet articulate a long-run strategy to manufacture and ship enough vaccine to fully—let alone quickly—inoculate the global population.

2.d. Domestic sources of changing U.S. preferences toward U.S. trade policy

The Biden administration has indicated it seeks to develop a "worker centered" trade policy. Thus far, the practical implications of that emphasis have been threefold.

First, given the divisive nature of trade in the public debate, the administration has decided against immediate pursuit of any new trade-liberalizing agreements. The Biden administration has even put on hold a handful of limited negotiations it inherited from the prior administration, such as potential free trade agreements with the United Kingdom and Kenya.

Second, it has prioritized enforcing worker-centered provisions in existing trade agreements. That has involved the administration initiating investigations into potential labor violations taking place at plants in Mexico under the new USMCA, the renegotiated NAFTA.

Third, in the ongoing multilateral negotiations over new rules for fisheries subsidies, the administration has tabled a new proposal seeking rules protecting against the use of forced labor on fishing vessels (USTR 2021). In a related action, its first WRO involved all tuna, swordfish, and other seafood sourced from fishing vessels owned or operated by Dalian Ocean Fishing, a Chinese company, for allegedly relying on forced labor (CBP 2021).

3. Policy recommendations

The Biden administration did not define trade as an early policy priority. It had to address the public health needs created by the pandemic, develop emergency fiscal policy, and shore up the U.S. economic recovery battered by recession. It also prioritized policy concerns such as racial injustice, climate change mitigation, immigration reform, tax reform, and infrastructure investments. As the Biden administration emphasizes restoring the American domestic economy through its "building back better" agenda, the trade community should view support of this domestic policy agenda as the first step necessary for rebuilding American political support for a future U.S. policy of openness to the global economy.

Nevertheless, the Biden administration also needs a new strategy for international engagement. The Trump administration approach was to engage countries bilaterally, if at all. It viewed trade as zero sum rather than mutually benefitting (on net) both countries. Any country, especially one with a bilateral trade surplus with the United States, was viewed as an adversary. Without yet providing detail, the Biden administration has indicated a different approach. There has been no mention of continuing the Trump administration's focus on remediating bilateral trade deficits. Instead, it plans to "work with allies."

Implementation of a new international trade engagement should then have at least three components: (1) resolving old grievances and establishing a framework to work out new grievances that will inevitably emerge *with allies*; (2) establishing a framework to work with allies in areas of common concern involving China; and (3) establishing a framework to work with allies and China in areas of global concern.

3.a. A worker-centered trade policy begins at home

Improving the American workforce's competitiveness and adaptation to a changing and dynamic global economy should be the administration's top priority and would be foundational to any trade agenda.

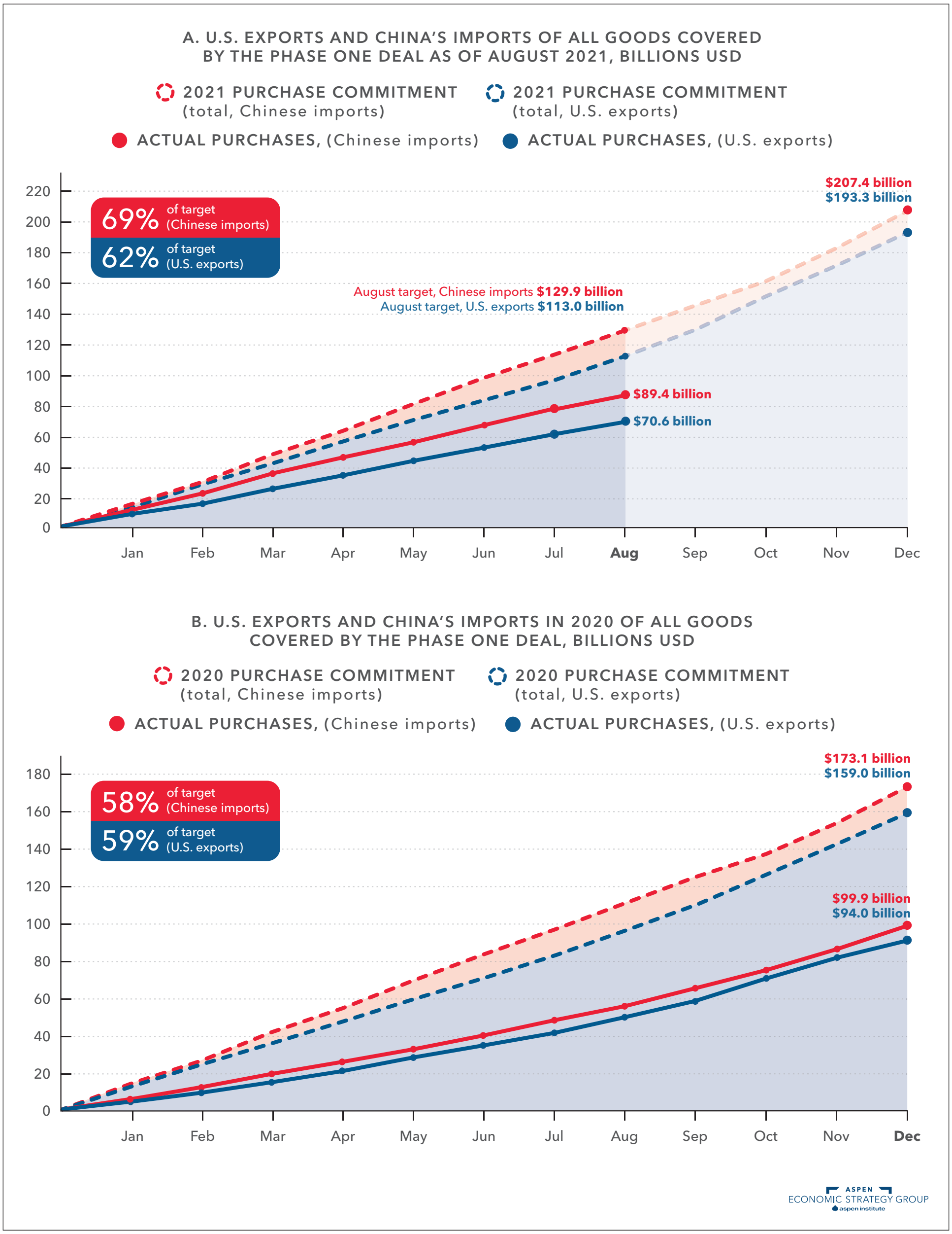
While trade has contributed to worker displacement, strongly impacting certain sectors and communities (Autor, Dorn, and Hanson 2013), there are multiple forces of disruption at work, including technological change and shifting consumer demands. U.S. policy needs to support any part of the American workforce facing disruption, regardless of the underlying cause of that disruption. Policies that focus on individual workers—rather than a particular set of jobs—such as promoting education, retraining, health care, childcare, and portability of benefits, will form the core of a worker-centered trade policy.

Enforcing the labor standards commitments that other trading partners have taken on is important, as is pushing for countries to take on stronger commitments to provide for their workers. But the economic impact on American workers of such actions is likely small relative to how much those workers would benefit from more supportive domestic labor market policies.

3.b. Review and adjust the U.S. tariffs unilaterally imposed on China

The Biden administration is conducting a review of the U.S.-China Phase One agreement that was signed in January 2020. However, U.S. tariffs remain on more than two-thirds of imports from China, and China has not lived up to its commitment to purchase an additional $200 billion of U.S. imports in 2020 (falling over 40% short) and 2021 (still well short), as indicated by Figure 4.

Figure 4: China's purchases of U.S. goods under the Phase One agreement



Source: Bown (2021b)

There is no reason the current and additional U.S. tariffs must remain as they are. Even if the United States is committed to maintaining a "noncooperative" policy toward imports from China, the optimal version of that policy is unlikely to be a blanket 25% tariff on nearly all imports of intermediate inputs from China that American businesses and workers rely on to remain competitive in the global economy. Most of those additional tariffs—or the ones imposed at lower (7.5%) rates—were chosen to avoid consumer products such as electronics, toys, clothing, and footwear (see again Figure 3). If U.S. tariffs on China are to remain a permanent part of U.S. trade policy, the products subject to the tariffs as well as the tariff rates should be reviewed and ultimately changed to reflect a policy that is in the best interest of the U.S. economy and its workers, and not just to avoid a negative reaction by American consumers facing higher prices.

3.c. Working with allies on non-China

The list of U.S. trade policy issues that do not involve China is always long. And it has been made more complex by the Trump administration having taken one important U.S. trade policy tool used to resolve frictions—WTO dispute settlement—off the table.[20] Nevertheless, there is evidence the Biden administration is seeking to resolve trade disputes with key allies, some of which involve complex negotiations and compromise.

For decades, the United States and European Union have each complained about the other's subsidies to commercial aircraft makers Airbus and Boeing. Under the Trump administration, the dispute had proceeded to the point of retaliatory tariffs. By removing those tariffs on imports from the European Union and United Kingdom, the Biden administration signaled a willingness to prioritize a quick and sustainable solution for subsidy disciplines that both sides would apply in this commercial space. This is also relevant because China has also been subsidizing commercial aircraft production, and so the two sides may need to jointly tackle subsidy disciplines with China.

A second important effort involves taxation of multinational corporations. A key element of the Biden plan agreed at the G7 in June suggests a willingness to accommodate the ability of other countries to also increase tax revenue collections from major multinationals. In effect, the United States would "share" some of its corporate tax base so that all governments might jointly collect more revenue. If it

20 WTO disputes take years to litigate. Even if the Biden administration were to work quickly reform the dispute settlement system and begin initiating new cases, it is unlikely the administration would see the political payoff of those disputes before the end of its first term. Thus, while in the long-term interests of the United States, fixing WTO dispute settlement is unlikely to deliver any political victories for the U.S. administration.

is agreed multilaterally at the OECD, and if it passes Congress and becomes U.S. law, the result would be cooperation with allies on trade, since it would head off the DSTs that allies deployed to unilaterally target U.S. tech companies.

Two other Trump-era grievances must be addressed with allies: U.S. tariffs on steel and aluminum, and a framework for future dispute resolution. These will be harder to tackle, because they are not purely bilateral issues and are caught up in the related question of how to address areas of common concern with China.

With steel and aluminum, the policy challenge is to find a solution that is rules-based and thus acceptable to allies, and ultimately addresses the underlying problem that is generating the demand for U.S. import protection in the first place: global overcapacity triggered by China and its system of subsidies. Before the Trump administration, that problem could have been addressed through a combination of additional transparency into the Chinese system and peer pressure arising through the OECD Steel Forum negotiations, potentially coupled with a multi-country trade dispute brought against China. (The Obama administration initiated a WTO dispute against China's *aluminum* subsidies that the Trump administration decided against pursuing.) For the moment, WTO dispute settlement is off the table because the entire system is dysfunctional. A solution will need to be found elsewhere, perhaps with a rejuvenated set of negotiations at the OECD. If packaged properly, China may be more willing to engage since its metal exports are facing increased protection outside of the United States, including and especially in the European Union.

The United States has offered protections to the steel industry in one form or another—voluntary export restraints, trigger price mechanism, antidumping and countervailing duties, safeguard tariffs, and most recently the national security tariffs—off and on since the 1960s. One potential resolution for the United States to renegotiate its tariff bindings to offer the U.S. steel industry permanently higher levels of import protection. There are WTO-consistent options available for doing so; the United States would simply need to negotiate compensation with trading partners. But trading partners are already extracting compensation, through retaliation against politically sensitive U.S. export sectors like Harley-Davidson motorcycles and Kentucky bourbon, thus the main change would be to convert this into a WTO-consistent form of protection. While costly to the U.S. economy, including to downstream using industries, this is a way to achieve such an end without threatening the integrity of the rules-based agreement. Thus, it could have benefits for cooperation elsewhere.

The second issue of ongoing concern to many economic allies is the absence of a functioning dispute resolution mechanism that resulted when the Trump administration ended the WTO's Appellate Body. This creates uncertainty for countries to take on new trade agreement commitments and risks escalating inevitable trade frictions into a trade war. There is bipartisan concern in the United States for how the WTO's original system performed, especially how its legal decisions constrained the United States' use of policies such as antidumping, countervailing duties, and safeguards (Bown and Keynes 2020). While most of those legal cases were brought by partners like the European Union, Japan, Canada, and South Korea, the major concern today is how the result of those rulings—and future rulings—would curtail the United States' ability to use those trade remedies to address imports from China.[21]

The dispute settlement problem requires fixing even if the United States is unwilling to do so with its bilateral relationship with China. The United States needs a mutually acceptable way of efficiently resolving trade frictions, especially with allies. Perhaps equally important, a viable system is needed so that the rest of the world can resolve its trade disputes, so that the entire rules-based trading system does not break down.

3.d. Working with allies on issues involving China

One key limitation of the Trump administration's approach toward China was that it was almost entirely bilateral. The trade war imposed considerable costs on the U.S. economy and the administration's signature Phase One agreement yielded minimal improvements in the areas of core concern. The European Union also negotiated with China bilaterally, albeit without suffering through a trade war. Nevertheless, the E.U.-China Comprehensive Agreement on Investment (CAI) made minimal progress with China on systemic issues. Each of these results were unsurprising, given the changes that the United States (and the European Union) would like China to make are systemic, generating benefits to lots of other countries as well. Because no single country would appropriate all of the gains from negotiating with China, no single country would be willing to offer up enough in negotiations to obtain them.

The U.S. bilateral strategy toward China may now be changing, as the Biden administration has indicated it will seek to work with allies. The European Commission in December 2020 issued a blueprint for how it might work with the new U.S. administration in this area (European Commission 2020). President Biden's June summit with the European Union provided some initial detail, including agreement to establish a new Transatlantic Trade and Technology Council.

---

21 This admittedly creates a bizarre parallel universe, because the main issue is that the United States and China are more importantly violating the basic rules of MFN by imposing their trade war tariffs on each other. Disagreeing over dispute settlement, as well as prior WTO legal decisions over "zeroing" or "public body," are less than second order, they are a nonbinding constraint in the U.S.-China relationship. Nevertheless, some of the issues are relevant for U.S. trade relations with other countries and could become relevant if the United States and China restored "cooperation" and once again treated one another like other WTO members.

An approach that collectively negotiates solutions with China will have pluses and minuses. On the positive side, convincing other partners to credibly threaten their own noncooperative policies toward China is more likely to convince Beijing of the benefits of adopting a cooperative policy. On the negative side, the United States often does not have the same offensive or defensive interests in negotiations as its allies. The first challenge will involve internal agreement and maintaining a common approach toward China. Working collectively will limit each country's ability to engage China unilaterally. It will prove difficult to remain united if and when China retaliates selectively and strategically, or offers something bilaterally, in an attempt to play allies one off another. This is the analog to China's strategic retaliation against U.S. farm interests—but not other sectors—when it attempted to pit one U.S. industry against another during the trade war.

Two important areas of concern to the United States and its allies are China's industrial subsidies and its system of forcibly transferring foreign technology. Admittedly, these are the areas on which the Trump administration had been working with the European Union and Japan. A "Trilateral" initiative began at the WTO Ministerial Conference in December 2017 to develop new international disciplines to address China's system of state capitalism and state-owned enterprises. Indeed, the three parties had made it sufficiently far to issue a joint statement on industrial subsidies in January 2020, proposing new types of unconditionally prohibited subsidies, which reversed the burden of proof in disputes.[22] Instead of the United States, for example, showing that China's subsidy has caused harm, China would now need to demonstrate that its subsidy has not harmed others.[23]

The Trilateral's progress on subsidies was halted with the COVID-19 pandemic, and there was little publicly acknowledged progress on the issue of the forceable transfer of technology. Though such a process could and should be re-engaged, an emerging question is whether any of the Trilateral partners would now stake out substantially revised positions on subsidies given the pandemic experience or other factors described earlier. However, if the three can agree between themselves, the next step would involve bringing their proposals to other like-minded countries before eventually approaching China to negotiate an agreement and return (for the United States and China) to implementing policies that are more cooperative.

A second and very new area involves potential allied coordination of export controls. With the implementation of U.S. export controls on semiconductors

22 For some of the challenges in addressing China's system of subsidies within a WTO framework, see Bown and Hillman (2019).

23 Joint Statement of the Trilateral Meeting of the Trade Ministers of Japan, the United States and the European Union Washington, DC, January 14, 2020.

and manufacturing equipment—which are also applied to semiconductor manufacturers in other allied trading partners, especially Taiwan, South Korea, Japan, and the European Union—these economies will increasingly seek to influence which technologies are being "controlled," since the policies affect their exports, too. A common solution will need to emerge, as the failure to control such technologies from all sources means the national security threat will not be addressed, despite the cost that the U.S. export controls impose on commercial interests of U.S. industry. However, the coordination of export controls falls outside of WTO rules.[24]

As a first step, the U.S.-E.U. decision in June to pursue a Trade and Technology Council looks to provide improved transatlantic coordination of export controls, as well as related national security concerns arising over inbound foreign direct investment. To be effective, establishing such an institutional framework would likely require expansion to a core group of other key economies—including the United Kingdom, Japan, South Korea, Canada, and Australia (plus or minus others, depending on the technology and involvement of the industrial supply chain). Policymakers need to better tailor export controls to limit their application to only where they are essential to protect national security threats. Overreliance on them will undermine cooperation and reduce their effectiveness.

A final and related area involves coordinating policies *against* forced labor and *in favor of* human rights and democracy. Like other examples, the failure to coordinate policy weakens their impact. It is already difficult for trade policy to impose costs on China without unintended consequences due to the global nature of supply chains.

All of this, of course, is designed to establish a clearer framework that will hopefully lead to a jointly preferable outcome, whereby the United States, its allies, China, and all WTO members return to a set of cooperative policies and participate in a mutually agreeable, rules-based trading system. That result would someday have the United States eliminating its Section 301 tariffs, China eliminating its retaliatory tariffs, and China and other countries taking on other commitments to stop implementing policies that impose externality costs on trading partners.

In the immediate term, it is unclear where these countries will end up. Will more of China's partners develop credible threats and ultimately deploy noncooperative policies of their own? Will China respond to them in kind? Or will the process of engagement more explicitly lead to cooperation? While the United States can

24 Export controls on dual use technologies are currently managed through the Wassenaar Arrangement, an agreement ill-suited to modern issues. The Wassenaar Arrangement replaced COCOM, which navigated the export control issue during the Cold War. It was designed in the early 1990s to control the flow of weapons of mass destruction from getting to rogue states. Its membership includes Russia, a country to which the United States and its allies might like to control the flow of certain technologies. For more see Bown (2020b).

establish, clarify, and incentivize the framework, the sovereign decisions of its allies and China will jointly determine the collective policy choices and overall outcomes.

3.e. Working with allies and China

Finally, as described earlier, there are at least two areas in which the United States and China, as well as other countries, must work collectively: climate and global public health.

As major emitters of carbon, both the United States and China must take on more stringent commitments to reduce emissions in a timely manner and on a larger scale. They are not alone, and the best approach would be not only to adhere to the commitments in the Paris Climate Accord, but also to cooperate and commit to the adoption of domestic policies that would allow them to go further. The Biden administration has been keen on prioritizing climate mitigation and cooperation, and the first visit by a Biden administration official to China was its climate envoy, John Kerry (Myers and Crowley 2021). While a carbon border tax is potentially in the offing if countries refuse to cooperate, if there is cooperation on reducing emissions (raising the explicit and implicit price of carbon) then CBAM would be a threat only and not required to be used in practice.

As discussed above, the pandemic has created a new global public health demand for cooperation on vaccine manufacturing, distribution, and thus trade. The emergence of viral variants has highlighted the risk that the COVID-19 pandemic will not really be over anywhere until it is under control everywhere, and that is likely to require vaccinating most of the world. Given the complexity of vaccine production, manufacturing in the near term is limited to only a handful of countries—including the United States and China—and thus trade will be essential to get vaccines distributed worldwide. Yet, the possibility of even manufacturing vaccines in sufficient quantity is complicated by the nature of global supply chains.

Effectively and quickly scaling up production globally requires additional collaboration and cooperation between the major economies. A handful of countries have proposed a Trade and Health Initiative at the WTO, but more is needed, including financing mechanisms, coordination of subsidies for inputs and outputs across countries, and agreements to not limit exports and keep trade lanes open. Bollyky and Bown (2020, 2021) have outlined one proposal for an explicit COVID-19 Vaccine Investment and Trade Agreement to help facilitate the transfer of technology as well as sufficient scaling up of production of inputs and raw materials in high demand by vaccine manufacturers.

The Biden administration has also committed to good faith negotiations over a waiver to the WTO rules for vaccine patents. Such a waiver may, over the long term, contribute to the transfer of technology for vaccine production globally, decreasing the current concentration in relatively rich countries, India, and China. But the rather long period of time before such benefits materialize means countries cannot waste existing opportunities to scale up current manufacturing capacity more quickly to produce more vaccines to inoculate the world. More trade in vaccines will be the quickest way to save lives.

4. Conclusion

Even with the entry of the Biden administration, U.S. trade policy was in for a long period of transition. Policy changes set in motion by the prior administration were only a start. While they might be modified, they are highly unlikely to be reversed. On almost all fronts, the future of U.S. trade policy looks to be very different from the past.

References

Amiti, Mary, Stephen Redding, and David Weinstein. 2019. "The Impact of the 2018 Trade War on US Prices and Welfare." *Journal of Economic Perspectives* 33 (4): 187–210.

Autor, David H., David Dorn, and Gordon H. Hanson. 2013. "The China Syndrome: Local Labor Market Effects of Import Competition in the United States." *American Economic Review* 103 (6): 2121–68.

———. 2016. "The China Shock: Learning from Labor Market Adjustment to Large Changes in Trade." *Annual Review of Economics* 8: 205–240.

———. 2021. "On the Persistence of the China Shock." Brookings Papers on Economic Activity, BPEA Conference Drafts, September 9.

Bagwell, Kyle and Robert W. Staiger. 1999. "An Economic Theory of GATT." *American Economic Review* 89 (1): 215–48.

———. 2001. "Strategic Trade, Competitive Industries and Agricultural Trade Disputes." *Economics and Politics* 13 (2): 113–28.

———. 2002. *The Economics of the World Trading System*. Cambridge, MA: The MIT Press.

———. 2011. "What do trade negotiators negotiate about? Empirical evidence from the World Trade Organization." *American Economic Review* 101 (4): 1238–73.

———. 2014. "Can the Doha Round be a development round? Setting a place at the table', in Robert C. Feenstra and Alan M. Taylor (eds.) *Globalization in an Age of Crisis: Multilateral Economic Cooperation in the Twenty-First Century*, Chicago, IL: University of Chicago Press, 91-124.

Bagwell, Kyle, Chad P. Bown and Robert W. Staiger. 2016. "Is the WTO Passé?" *Journal of Economic Literature* 54 (4): 1125–1231.

Bhagwati, Jagdish and Hugh T. Patrick (Eds.). 1990. *Aggressive Unilateralism: America's 301 Trade Policy and the World Trading System*. Ann Arbor, MI: University of Michigan Press.

Bollyky, Thomas J. and Chad P. Bown. 2020. "The Tragedy of Vaccine Nationalism: Only Cooperation Can End the Pandemic." *Foreign Affairs* 99 (5): 96–109.

Bown, Chad P. 2019a. "The 2018 US-China Trade Conflict After 40 Years of Special Protection." *China Economic Journal* 12 (2): 109–136.

———. 2019b. "What's Left for the WTO?" in Ben Zissimos (ed.) *The WTO and Economic Development,* Cambridge, MA: The MIT Press, 93-144.

———. 2020a. "How the United States Marched the Semiconductor Industry Into Its Trade War with China." *East Asian Economic Review* 24 (4): 349-388.

———. 2020b. "Export Controls: America's Other National Security Threat." *Duke Journal of Comparative and International Law* 30 (2): 283–308.

———. 2021a. "The US-China Trade War and Phase One Agreement." *Journal of Policy Modeling,* forthcoming.

———. 2021b. US-China Phase One Tracker: China's Purchases of US goods as of August 2021. PIIE Chart, Peterson Institute for International Economics, August 30. https://www.piie.com/research/piie-charts/us-china-phase-one-tracker-chinas-purchases-us-goods

———. Forthcoming. "How COVID-19 Medical Supply Shortages Led to Extraordinary Trade and Industrial Policy." *Asian Economic Policy Review.*

Bown, Chad P. and Thomas J. Bollyky. 2021. "Here's how to get billions of COVID-19 vaccine doses to the world." *PIIE Trade and Investment Policy Watch*, March 18.

———. Forthcoming. "How COVID-19 Vaccine Supply Chains Emerged in the Midst of a Pandemic." *The World Economy.*

Bown, Chad P., and Meredith A. Crowley. 2013. "Self-Enforcing Trade Agreements: Evidence from Time-Varying Trade Policy." *American Economic Review* 103 (2): 1071–90.

Bown, Chad P., and Jennifer A. Hillman. 2019 "WTO'ing a Resolution to the China Subsidy Problem," *Journal of International Economic Law* 22 (4): 557-578.

Bown, Chad P. and Soumaya Keynes. 2020. "Why Trump Shot the Sheriffs: The End of WTO Dispute Settlement 1.0." *Journal of Policy Modeling* 42 (4): 799–819.

Bown, Chad P. and Soumaya Keynes. 2021. "Tackling climate change with a carbon border adjustment mechanism." *Trade Talks,* Podcast episode 156, July 25.

Brander, James A., and Barbara J. Spencer. 1985. "Export Subsidies and International Market Share Rivalry." *Journal of International Economics* 18 (1–2): 83–100.

Broda, Christian, Nuno Limão, and David E. Weinstein. 2008. "Optimal Tariffs and Market Power: The Evidence." *American Economic Review* 98 (5): 2032–65.

Cavallo, Alberto Gita Gopinath, Brent Nieman and Jenny Tang. 2021. "Tariff Passthrough at the Border and at the Store: Evidence from US Trade Policy." *American Economic Review: Insights* 3(1): 19-34.

CBP (U.S. Customs and Border Protection). 2021. "CBP issues Withhold Release Order on Chinese fishing fleet." CBP Public Affairs press release.

European Commission. 2020. Joint Communication to the European Parliament, the European Council and The Council: A new EU-US agenda for global change. High Representative of the Union for Foreign Affairs and Security Policy, December 2.

European Commission. 2021. Proposal for a Regulation of the European Parliament and of the Council Establishing a Carbon Border Adjustment Mechanism. Brussels, July 14.

Fajgelbaum, Pablo D., Pinelopi K. Goldberg, Patrick J. Kennedy, and Amit K. Khandelwal. 2019. "The Return to Protectionism." *Quarterly Journal of Economics* 135 (1): 1–55.

Flaaen, Aaron and Justin Pierce. 2019. "Disentangling the Effects of the 2018-2019 Tariffs on a Globally Connected U.S. Manufacturing Sector." Finance and Economics Discussion Series 2019-086. Washington: Board of Governors of the Federal Reserve System, https://doi.org/10.17016/FEDS.2019.086.

Handley, Kyle, Fariha Kamal, and Ryan Monarch. 2020. "Rising Import Tariffs, Falling Export Growth: When Modern Supply Chains Meet Old-Style Protectionism." National Bureau of Economic Research Working Paper #26611. Cambridge, MA.

Hendrix, Cullen S. and Marcus Noland. 2021. "Assessing potential economic policy responses to genocide in Xinjiang." Peterson Institute for International Economics Policy Brief 21-14, June.

Hillman, Jennifer. 2018. "The Best Way to Address China's Unfair Policies and Practices Is Through a Big, Bold Multilateral Case at the WTO." Testimony Before the U.S.-China Economic and Security Review Commission Hearing on U.S. Tools to Address Chinese Market Distortions, June 8. https://www.uscc.gov/sites/default/files/Hillman%20Testimony%20US%20China%20Comm%20w%20Appendix%20A.pdf

Hufbauer, Gary Clyde and Zhiyao (Lucy) Lu. 2018. "The European Union's Proposed Digital Services Tax: A De Facto Tariff." Peterson Institute for International Economics Policy Brief 18-15, June.

Ip, Greg. 2021. "'Industrial Policy' Is Back: The West Dusts Off Old Idea to Counter China." *The Wall Street Journal*, July 29. https://www.wsj.com/articles/subsidies-chips-china-state-aid-biden-11627565906

Keynes, Soumaya. 2021. The EU proposes a carbon tariff on some imports. *The Economist*, July 17. https://www.economist.com/finance-and-economics/2021/07/15/the-eu-proposes-a-carbon-tariff-on-some-imports

Lardy, Nicholas R. 2014. *Markets over Mao: The Rise of Private Business in China*. Washington: Peterson Institute for International Economics.

———. 2019. *The State Strikes Back: The End of Economic Reform in China?* Washington: Peterson Institute for International Economics

Ludema, Rodney D., and Anna Maria Mayda. 2013. "Do Terms-of-Trade Effects Matter for Trade Agreements? Theory and Evidence from WTO Countries." *Quarterly Journal of Economics* 128 (4): 1837–93.

Mattoo, Aaditya and Robert W. Staiger. 2020. "Trade wars: What do they mean? Why are they happening now? What are the costs?" *Economy Policy* 35 (103): 561–584.

Mavroidis, Petros C. and Andre Sapir. 2021. *China and the WTO: Why Multilateralism Still Matters.* Princeton, NJ: Princeton University Press.

Myers, Steven Lee, and Michael Crowley. 2021. "John Kerry Heads To China for First High-Level Talks Under Biden." *The New York Times*, April 13, 2021. https://www.nytimes.com/2021/04/13/us/politics/john-kerry-china-climate-biden.html

OECD (Organization for Economic Cooperation and Development). 2019a. Measuring distortions in international markets: The aluminium value chain. OECD Trade Policy Papers, no. 218. Paris: OECD Publishing.

OECD (Organization for Economic Cooperation and Development). 2019b. Measuring distortions in international markets: The semiconductor value chain. OECD Trade Policy Papers, no. 234. Paris: OECD Publishing.

Ross, Wilbur. 2017. "Most favored nation rule hurts importers, limits U.S. trade," *The Wall Street Journal,* Opinion, May 25. https://www.wsj.com/articles/most-favored-nation-rule-hurts-importers-limits-u-s-trade-1495733394

Staiger, Robert W. Forthcoming. *A World Trading System for the Twenty-First Century*. Cambridge, MA: MIT Press.

Staiger, Robert W. and Alan O. Sykes. 2013. "Non-Violations." *Journal of International Economic Law* 16 (4): 741-775.

Stolberg, Sheryl Gay, Sharon LaFraniere, and Chris Hamby. 2021. "Top Official Warned That Covid Vaccine Plant Had to Be 'Monitored Closely.'" *The New York Times*, April 7. https://www.nytimes.com/2021/04/07/us/emergent-biosolutions-coronavirus-vaccine.html

Tai, Katherine. 2021. "Remarks from Ambassador Katherine Tai on Trade Policy, the Environment and Climate Change." Office of the United States Trade Representative. https://ustr.gov/about-us/policy-offices/press-office/press-releases/2021/april/remarks-ambassador-katherine-tai-trade-policy-environment-and-climate-change

USTR (US Trade Representative). 2018a. Findings of the Investigation into China's Acts, Policies, and Practices Related to Technology Transfer, Intellectual Property, and Innovation under Section 301 of the Trade Act pf 1974. March 22. Washington.

———. 2018b. Update Concerning China's Acts, Policies and Practices Related to Technology Transfer, Intellectual Property, and Innovation. November 20. Washington.

———. 2021. United States Urges WTO Members to Address Forced Labor on Fishing Vessels in Ongoing Fisheries Subsidies Negotiations. May 26. Washington.

US Department of Treasury. 2021. The Made in America Tax Plan. April.

Varas, Antonio, Raj Varadarajan, Jimmy Goodrich and Falan Yinug. 2021. "Strengthening The Global Semiconductor Supply Chain In An Uncertain Era." Semiconductor Industry Association and Boston Consulting Group, March.

———. 2020. "How Restrictions to Trade with China Could End US Leadership in Semiconductors." Boston Consulting Group and Semiconductor Industry Association.

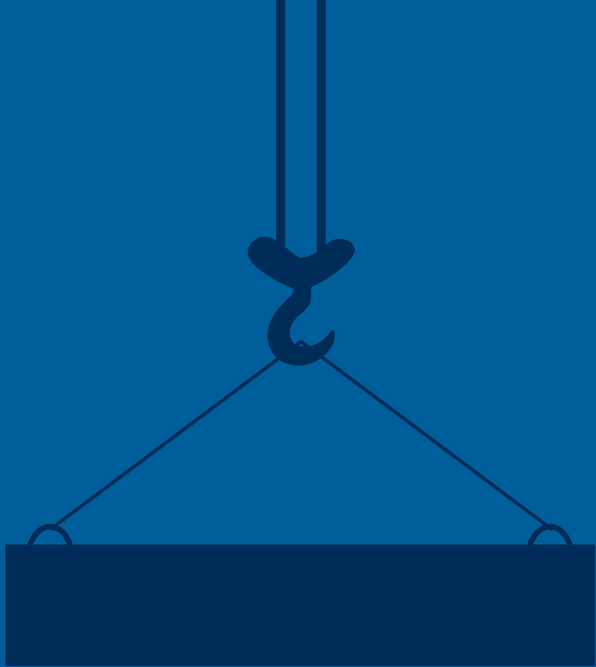
White House. 2021. Fact Sheet: The Bipartisan Infrastructure Investment and Jobs Act Advances President Biden's Climate Agenda. August 5.

Wu, Mark. 2016. "The 'China, Inc.' Challenge to Global Trade Governance." *Harvard International Law Journal* 57 (2): 261–224.

———. 2020. Managing the China Trade Challenge: Confronting the Limits of the WTO. Penn Project on the Future of U.S.-China Relations, Working Paper, October.

PART II

THE US INFRASTRUCTURE AGENDA



Economic Perspectives on Infrastructure Investment
Edward Glaeser and James Poterba

Challenges of a Clean Energy Transition and Implications for Energy Infrastructure Policy
Severin Borenstein and Ryan Kellogg

Science and Innovation: The Under-Fueled Engine of Prosperity
Benjamin F. Jones

Economic Perspectives on Infrastructure Investment

AUTHORS

Edward Glaeser* and James Poterba**

ABSTRACT

To determine the appropriate level of infrastructure spending, there is no alternative to aggregating the results of project-by-project cost-benefit analysis. With widespread variation in both the benefits and costs of projects within broad infrastructure asset classes, it is important to recognize that the returns to some additional highway lanes are much higher than others, and that the costs of extending wire-line broadband coverage in some locations may exceed the benefits relative to the next-best alternative technology. Because comprehensive project evaluation is enormously information-intensive and can be gamed, many of the widely discussed estimates of the infrastructure gap in the United States are based on alternative methodologies, such as benchmarking infrastructure spending levels against international or historical averages. Such exercises may not recognize that infrastructure projects in the United States often cost more than arguably comparable projects in other nations. As infrastructure spending ramps up, heightened attention to procurement

* Harvard University; National Bureau of Economic Research. Email: eglaeser@harvard.edu.

** MIT; National Bureau of Economic Research. Email: poterba@mit.edu.

Acknowledgements: This chapter was prepared for the July 2021 meeting of the Aspen Economic Strategy Group. We are grateful to Amy Ganz and Melissa Kearney for valuable comments, to Treb Allen, Ricardo Sanchez Gomez, Shane Greenstein, Dejan Makovšek, Lawrence Summers, Clifford Winston, and Faith Witryol for input and discussions, and to Hilary Gelfond for research assistance. Sections of this chapter draw on the infrastructure economics project that is summarized in Glaeser and Poterba (2021).

Disclosure: The authors, along with Stephen Redding of Princeton University, are co-investigators on a research grant focused on transportation economics that is funded by the US Department of Transportation.

practices and to project management could yield high returns in avoiding unnecessary spending. Cost-benefit calculations must also consider maintenance spending as an important infrastructure outlay, since the bias of the political system toward ribbon-cuttings for new projects can often short-change high-return upgrade and maintenance work. Financing infrastructure is a perennial challenge. User fees, while politically difficult to adopt, can be an important way of ensuring that infrastructure is used efficiently and of aligning funding with those who reap the benefits of new projects. In many cases, infrastructure use is progressive. In cases in which user fees may be regressive, such as fares on public transit buses, it may be possible to design compensatory policies, such as transit system vouchers for low-income households, to offset distributional concerns while preserving the efficiency benefits of use-related charges. Public-private partnerships require careful analysis on a case-by-case basis. While they can deliver operational and procurement benefits relative to similar projects managed exclusively by the public sector, they can also impose unpriced risks on taxpayers, and in some cases can impose a long-term increase in the cost of infrastructure use in return for a short-term relaxation of public sector liquidity constraints.

1. Introduction

Both sides of the political spectrum routinely call for increased public spending on infrastructure projects, although their justifications often vary. Federal support for infrastructure now seems likely to ramp up in the near future, although the spending plan's size, timing, financing and scope are not finalized. The prospect of a major infrastructure initiative makes this an appropriate time to review a number of economic insights related to such spending.

All infrastructure projects are not created equal, and the benefits per dollar spent can vary widely as a function of the nature of the project and the management of its construction and subsequent use. Different projects will benefit different constituencies, both geographically and across the economic spectrum, so infrastructure projects must be viewed as one part of the federal government's activities that impact the distribution of economic resources.

This chapter highlights policy relevant lessons from the voluminous research literature on the economics of infrastructure projects. We provide a selective introduction to the many studies on the topic and draw a number of conclusions that can inform the design of an infrastructure-spending program, regardless of its size or scope. Four conclusions deserve particular note.

First, because infrastructure projects differ widely in cost, complexity, and benefits, systematic cost-benefit analysis is a critical tool for identifying the highest-return opportunities. Studies of the return to expanding the interstate highway system, for example, point to very different benefits in different locations. The highest value derives from expansions in densely populated areas with congested roadways—but those are also often the most expensive places to build new highways or expand existing ones. While it is difficult to draw firm conclusions about the optimal size of a federal infrastructure initiative from existing research, the rigorous application of cost-benefit analysis can help direct spending to the highest return-for-cost projects. A number of studies suggest high returns to maintaining existing infrastructure. Comparing the returns to new projects with those on maintenance initiatives is a key component of any high-return infrastructure program.

Second, managing costs is essential for any infrastructure project. Building highways and subways is, by some indicators, significantly more expensive in the United States than in other high-income countries. Several potential factors may explain these high costs: infrastructure design, the extent to which the project must remediate potential adverse effects on communities and the environment, and construction management, starting with procurement and including the way delays and cost over-runs are handled. The best time to address these issues is before, rather than after, launching a major infrastructure spending initiative.

Third, user fees can play an important role in financing both new infrastructure projects and in maintaining existing ones. User fees are often ruled out in the policy process because they are claimed to be regressive. More honestly, they are politically difficult. Yet fees for vehicle miles traveled that vary by time of day, for parking in dense urban areas, for the use of airports and ports, and many other user charges can reduce the demands that an infrastructure program places on general revenues. If set to reflect the marginal cost of using infrastructure, they also represent an important step toward its efficient utilization. While there are many justifications for investing in infrastructure, there are few compelling reasons for making such infrastructure free to users, especially since that will lead to utilization above and beyond the efficient, cost-reflective level. Overuse of transportation infrastructure is not just economically inefficient, but can have other adverse effects, such as the generation of excessive carbon dioxide emissions and other forms of pollution.

Finally, while public-private partnerships and the privatization of infrastructure assets can sometimes enhance operational efficiency and improve both procurement and management, private provision also creates risks. At times, private providers have negotiated, or in some cases renegotiated, highly advantaged terms, or failed to serve key constituencies. Privatizing infrastructure assets as a means of raising capital warrants particularly scrutiny. It is only attractive when the private sector can secure funds on more favorable terms than the public sector, but the U.S. Treasury borrows at a particularly low rate and most state governments also have excellent bond ratings.

The remainder of this chapter describes the evidence that leads us to these four conclusions and elaborates on them. We begin by identifying key features of traditional, or physical, infrastructure assets. Then we describe various approaches to assessing the appropriate level of infrastructure investment and to undertaking cost-benefit analyses. Next, we discuss the cost and financing of infrastructure projects. Finally, after describing some of the political economy challenges that arise in infrastructure projects, we address the potential role of the federal government in a more localized process of project selection and implementation.

2. Defining features of physical infrastructure projects

The term "infrastructure" is a relatively recent addition to our national vocabulary, and its meaning has evolved over time. Carse (2016) explains that it was originally used by engineering writers to describe railroad tracks, which were a "piling up" (*structura* in Latin) below (*infra*) steam trains. The first appearance of "infrastructure" in an English language economics journal appears to be Wellisz' (1960) article on Dutch, French, and Italian economic development. He put the term in quotations and defined it to be synonymous with social overhead capital, investments that lowered costs for their users, while also delivering static externalities as well as dynamic externalities by encouraging private investment. When Joy (1967) used the term, it remained sufficiently esoteric to require definition, in this case as a "synonym for 'track' in its broader sense of earthworks, bridges, tunnels, permanent way and signage." In the last half century, "infrastructure" has expanded well beyond the railroad sector, but in most contexts it still refers to various types of fixed capital investments. In an influential study, the Congressional Budget Office (1988) referred to "public works infrastructure" and identified six sub-categories: highways, aviation, mass transit, wastewater treatment, water transportation, and groundwater and surface water resources.

The American Jobs Plan, when proposed in March 2021, included spending on public works, traditional infrastructure, as well as new outlays to retrofit homes and private commercial buildings, provide job training, increase R&D spending, and improve access to child and elder care. Spending programs in these categories are often labeled "social infrastructure."

This chapter focuses primarily on traditional infrastructure projects that involve fixed capital investments associated with the movement of goods—including electricity and digital content—or people. This definition encompasses all of the CBO's public works infrastructure, as well as broadband, fiber optic cables, and the electricity grid. Our focus on these topics should not be interpreted as a dismissal of the importance of social infrastructure spending. Such spending can be enormously valuable. However, the issues in the design and analysis of such spending programs are different from those associated with traditional infrastructure programs. [1]

To introduce our focus on physical infrastructure projects, we highlight four important features that are found in nearly all of them and that are central to their analysis:

1. ***Project valuation depends on future use.*** Predicting future use is essential to evaluating a potential infrastructure project. Use is determined by demand from potential users and by the supply of complementary inputs. For instance, rails have little value without trains, and highways are less valuable when the cost of vehicles or fuel are high, or when there is little parking available at potential destinations.

2. ***Projects generate location-specific benefits.*** Infrastructure projects typically generate benefits to users in a particular place. A fixed geographic location makes infrastructure, such as a rail line, riskier than other investments, like buses, that can be moved to adapt to changing circumstances. The place-based nature of infrastructure also implies that its beneficiaries are geographically concentrated, which means that it will have particular appeal to place-based politicians, and as a potential source of aid for disadvantaged places. Infrastructure investment may have direct benefits for a location, and it may also spur complementary private sector investment. Infrastructure investments can also generate negative externalities, such as road noise, that harm particular neighborhoods.

3. ***The marginal cost of facility use is often below the average cost of service delivery.*** Traditional infrastructure is a fixed investment, and the marginal cost of using it may be less than the average cost of building and maintaining it. That gap is a potential justification for government subsidies. Measuring the marginal cost of use is not always straightforward—especially when there are significant costs of congestion, accidents, pollution or when the depreciation rate of the physical capital depends on its use. These components can be much more difficult to assess then simple operating costs.

1 In some cases, the line between traditional and social infrastructure blurs. For example, while we do not discuss investments in hospitals or schools, neither of which are involved in moving goods or people, either could be included in traditional infrastructure, while also playing a key role in the provision of social infrastructure.

4. ***Projects are long-lived.*** Nearly all infrastructure projects involve both immediate costs and future benefits This raises issues about how to appropriately discount future benefits and costs, and makes the choice of a discount rate a key policy lever. The decline in real interest rates in the last three decades should be reflected in this process. In addition, inherent uncertainty about the future makes it difficult to accurately value the costs and benefits associated with any particular project. Uncertainly also means that flexibility is a desirable feature of long-lived infrastructure projects. For instance, some roads can take a variety of vehicles, while railroad tracks have limited applicability. More flexible infrastructure projects should command a lower risk premium than inflexible projects, since they can adapt more easily to technological or economic change.

Most infrastructure projects involve a period of investment, followed by a much longer period of use. This means that the timing of the investment period can matter. Advocates of fiscal stimulus see infrastructure as a natural tool for employing underutilized labor and capital during a downturn. Skeptics retort that infrastructure takes so long to plan and implement that most recessions will be over before meaningful work gets done. Recognizing and addressing these lags is essential if infrastructure spending is to be used as a tool of macroeconomic stabilization.

3. What determines the optimal level of infrastructure investment?

The initial proposal for the American Jobs Plan called for $932 billion of spending on traditional forms of infrastructure, including transportation ($621 billion), water ($111 billion), broadband ($100 billion) and the electric grid ($100 billion). Some call for even larger outlays: the American Society of Civil Engineers (ASCE) (2021) claims the United States needs nearly $2 trillion in spending to close its 10-year infrastructure-investment gap. What determines the optimal level of infrastructure investment, and the optimal size of the infrastructure capital stock? These are difficult questions to answer, and it is easier to describe an approach to answering them than to provide a specific answer.

The guiding economic principle is clear: the optimal level of infrastructure should be determined by comparing the costs of acquiring infrastructure capital with the benefits of using it. Benefits can be difficult to measure, however, and projected and completed project costs often diverge. This section describes several approaches to assessing the optimal level of infrastructure capital and the returns to infrastructure investment. It contrasts the "engineering" approach, which defines infrastructure need without emphasizing the trade-offs between marginal costs and marginal benefits, and the economic approach, which embraces them.

3.a. A collision of paradigms: engineering vs. economics

One of the most widely cited studies of the state of the U.S. infrastructure capital stock, which is commonly invoked in support of higher spending levels, is the ASCE's *Report Card for America's Infrastructure*. It assigns the United States a grade of C- for 2021. Another study by the McKinsey Global Institute (MGI), summarized by Woetzel et al. (2016), finds that "the world needs to invest an average of $3.3 trillion annually just to support currently expected rates of growth." Both studies determine infrastructure need by reference to standards, in the ASCE case engineering standards, and in the MGI case historical spending levels, that do not consider the cost of infrastructure investment. The implicit premise of the ASCE study is that "need" equals that cost of bringing all infrastructure capital up to best-practice engineering standards. Particularly in the case of upgrading currently safe and functional infrastructure that does not meet current design standards, some comparison of costs and benefits seems more appropriate.

Unlike engineers, who are often asked what it will cost to build a bridge but not asked to measure its benefits, economists are rarely asked to determine the cost of a bridge, but they are often asked whether the benefits of building it compare favorably with other potential uses of the public funds that building the bridge will require. Lionel Robbins famously defined economics as "the science which studies human behavior as a relationship between ends and scarce means which have alternative uses." It is difficult for an economist to consider infrastructure spending as a fixed requirement that must be satisfied before allocating funds to health care or education or national defense.

The economic approach to assessing the optimal level of infrastructure capital is project-driven. It begins by estimating returns on investing in a particular project and comparing them with its cost. Provided the costs include the distortions associated with tax finance or other funding mechanisms, and that there are no constraints on raising additional revenue, the decision rule "if benefits exceed costs, accept the project" will generate the set of projects that warrant public investment. If there is a fixed budget available for infrastructure investment, it may not be possible undertake all projects for which benefits exceed costs; in that case investment should flow to projects in the order of their benefit-cost ratio.

This approach, which endeavors to include benefits to users and to society as a whole, typically yields a list of high- and low-return activities. Gramlich (1994), an example of the application of the economic approach, presents a ranking of potential projects. In this framework, the optimal level of infrastructure spending equals the sum of the cost of all the projects for which benefits exceed costs. Importantly, the estimate of need would depend in part on economic parameters, such as the costs of inputs like steel and concrete, the construction wage rate, and the level of interest rates. If costs of construction rose, the optimal number of infrastructure projects to undertake would decline. If interest rates and discount rates fell, holding all else equal, the warranted level of infrastructure spending would rise, because the future benefits of infrastructure projects would be valued more highly relative to their current costs.

3.b. How are widely-cited estimates of infrastructure "need" constructed?

To frame the discussion of the optimal infrastructure capital stock, it is helpful to understand how two of the most widely cited studies of infrastructure need develop their estimates. The ASCE analyzes the current infrastructure capital stocks in a variety of different asset classes, and compares them with measures of need and engineering best practice. The ASCE *Report Card,* which estimates the spending needed to raise the nation's grade to an "A," are often thought to measure the level of spending required to preserve the safety and soundness of transportation and water infrastructure. That is not the case; the infrastructure grades target a different benchmark.

When ASCE refers to an infrastructure asset as structurally deficient, that does not mean that it is unsafe. In fact, "structural deficiency" is a more technical term. An asset can be classified as structurally deficient because it does not meet all of the current standards for constructing a new asset of the same type. In the case of a bridge, it may receive a structural deficiency label because of substantial water traffic delays at high tide. The CBO (2016) explains that "bridges with structural deficiencies have significant parts in a deteriorated condition and reduced load-carrying capacity. Bridges that are functionally obsolete do not meet current design standards.… Neither type of deficiency necessarily indicates that a bridge is unsafe." In addition, the grades assigned to various infrastructure classes depend on a number of subjective elements, such as the degree of innovation in infrastructure planning and construction, and the robustness of the funding plan by the government entity that is responsible for maintaining the asset. An infrastructure class may lose marks because of organizational weaknesses in the entity that oversees assets in that class, not because of limitations in the physical condition of the underlying assets. Such considerations may be relevant for discussions of infrastructure financing and governance, but the low grades do not necessarily reflect the quality of existing physical infrastructure.

The ASCE also assumes, implicitly, that the only way to remedy an infrastructure deficiency is by building new capacity (*e.g.*, reducing daily congestion on a particular roadway by adding more lane-miles). Taken on its own terms, the highway congestion example illustrates the shortcomings of defining infrastructure need by setting a fixed target, such as the absence of congestion. Calculating the reduction in congestion-hours on a particular road segment that will flow from a given road-building program is difficult, in part because highway lane supply creates its own demand. Duranton and Turner (2011) find that the amount of driving increases dramatically with the number of road miles built. Even assuming that it was feasible to expand the highway network enough to sharply lower traffic delays, the underlying goal of traffic-free roads is not the same as determining the optimal stock of highway capital. Why do we think that spending enough to get traffic-free roads is the best use of government funds relative to other uses of public funds, such as investing in early childhood education?

In contrast, a key element of the economic approach is recognizing that capital spending is only one way of addressing a given objective. There may be others. The same outcome that could be achieved by building additional highway lanes could also be achieved by adopting sophisticated time-of-day congestion pricing on the most-demanded routes, as some cities, such as Singapore, have done with some success. The optimal size of the infrastructure capital stock is likely to be smaller if utilizing the capital is priced rather than free. Engineering estimates of infrastructure need are likely to be overstated because of the failure to consider more efficient use of existing infrastructure assets. Cost-benefit analysis should be used to choose among the different approaches to reducing congestion.

Moreover, transportation innovations, like ride-sharing services, automated vehicles, and GPS-based road pricing, provide additional alternatives to public infrastructure. For example, instead of building train links to underserved populations in low-density locales, the poorer residents of those places could be provided with vouchers for ride-sharing services. An experimental program that allocates such vouchers, or combines mobility vouchers with Section 8 housing vouchers, could facilitate measuring the impact of such services. GPS-based road pricing can reduce congestion even more effectively than traditional tolls. Autonomous buses on dedicated lanes can move swiftly between cities and offer a plausible alternative to rail that is far less expensive.

While some of the infrastructure spending that is identified in the ASCE *Report Card* can perhaps be viewed as a "need," much is discretionary, and should be subject to standard cost-benefit analysis. Even for some decisions that may appear to be binary, such as whether or not to repave a road, there is often a continuous decision component, such as when to repave. Cost scales with the frequency of repaving. The timing of spending on new roads, bridges, and tunnels is flexible, yet this is not reflected as a consideration in the ASCE analysis.

Because the engineering analysis of infrastructure is a technical task, it can be difficult for nonexperts to find independent metrics to evaluate the ASCE grades and to thereby assess the spending recommendations. One example of where such a comparison is possible is road quality. The Department of Transportation (DOT) collects International Roughness Index (IRI) for U.S. roads. Relatively comprehensive data on U.S. roads are available since 1993. The ASCE awarded U.S. roads a "D" grade in both 2017 and 2021, up from D- grades in 1998 and 2009 but down from a C+ grade in 1988 and a D+ grade in 2001. Yet Duranton, Nagpal, and Turner (2020) observe that for at least one important class of transportation arteries—interstate highways—road roughness has improved over time.

Table 1 shows that for urban interstates, the percentage of miles of highway (distinct from percentage of miles driven on the highway) that is in the smoothest category has risen from 3.5% in 1993 to 40% in 2019. The share of smooth rural interstate road-miles increased from 8.3% to 53.4%. Interstate highways account for about 2.5% of U.S. roadway lane-miles but nearly one-quarter of all miles driven. Yet despite these significant improvements, the ASCE grade for highways fell from a C+ in 1988 to a D- in 1998. Road smoothness increased dramatically from 2009 to 2019, and yet the ASCE raised its assessment by only one-third of a grade.

Table 1: Roughness of U.S. highways, 1999–2019

ROAD TYPE AND ROUGHNESS MEASURE	1993	1999	2009	2019
Rural, IRI > 170	12.5%	12.6%	10.5%	12.5%
Rural, IRI < 60	6.2	9.5	12.0	16.3
Urban, IRI > 170	18.5	28.0	29.9	33.0
Urban, IRI < 60	4.0	9.0	9.1	8.0
Rural Interstates, IRI > 170	7.0	2.3	1.7	2.0
Rural Interstates, IRI < 60	8.3	21.5	34.0	53.4
Urban Interstates, IRI > 170	13.2	7.3	5.1	5.0
Urban Interstates, IRI < 60	3.5	12.0	20.9	40.1

Source: Authors' calculations using DOT Condition of U.S. Roadways by Functional System data as reported at https://www.bts.gov/content/condition-us-roadways-functional-system, accessed 5/31/2021. Rural and urban roads are divided into various categories (interstates, other principal arterials, minor arterials, and major collectors (rural) and collectors (urban)). The entries in rows 1-4 weight the IRI results for each category based on the number of road-miles in that category.

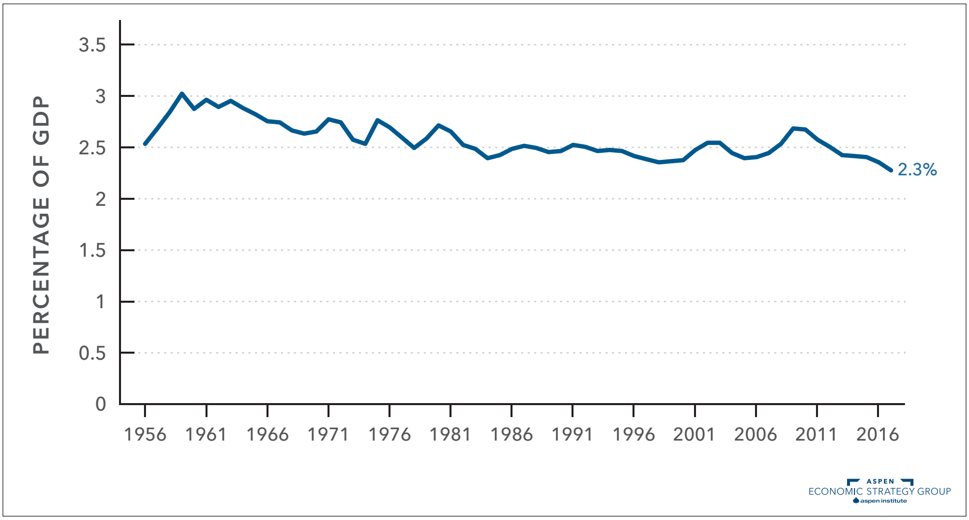
Bridges have also seen significant improvements in the last two decades, but only a slight increase in their ASCE grade. In 1998, the Federal Highway Administration reported that 6.9% of the bridges that were classified as part of the National Highway

System (NHS), and 16.5% of all bridges, were structurally deficient. ASCE assigned bridges a C- grade that year. By 2017, the share of bridges classified as structurally deficient had been cut in half, to 3.4% of the NHS bridges and 8.9% of all bridges. The ASCE grade rose only modestly, however, to a C+.[2]

While ASCE is probably the most widely discussed study of infrastructure need, another frequently cited source for infrastructure need is the McKinsey Global Institute (MGI). MGI adopts three approaches in assessing infrastructure need; none compares infrastructure benefits with infrastructure costs. The first uses historic spending patterns, as a percentage of GDP, for each country. Woetzel et al. (2016), who summarize the MGI findings, report that "global investment on roads, rail, ports, airports, power, water, and telecommunications infrastructure has averaged about 3.8% of global GDP." If this 3.8% is multiplied by global GDP projections through 2030 from IHS Global Insight, which assume a 3.3% annual growth rate, total global infrastructure spending "need" equals $62 trillion from 2013 through 2030.

Figure 1 shows the time series pattern of U.S. spending on transport and water infrastructure as a share of GDP. From a high of nearly 3% of GDP in the late 1950s, when the interstate highway system was being built, the level of spending has trended down, reaching about 2.3% of GDP in recent years.

Figure 1: Public spending on transport and water infrastructure, 1956–2017



Source: CBO (2018).

2 FHWA data on bridges may be found at https://www.fhwa.dot.gov/bridge/nbi/no10/defbr17.cfm (accessed July 01, 2021). Historical values of the ASCE Infrastructure Report Card are available at https://infrastructurereportcard.org/making-the-grade/report-card-history/.

Cross-country comparisons indicate that the United States is currently a low-spending nation with regard to transportation investment. Table 2 shows that in 2019, transportation infrastructure investment as a share of GDP in France, Germany, and the U.K. was higher than that for the United States, at 0.89%, 0.89%, 0.71%, respectively, as compared to 0.55% in the United States. China spent 5.64% of GDP on transportation investment, but the comparison is difficult because China is building infrastructure from a much lower base than other countries. Cross-country comparisons are very difficult to evaluate because of this initial condition issue, and because the costs of investing in infrastructure may different across countries.

Table 2: Transportation infrastructure investment, share of GDP, 2019

Country	Inland Transportation Investment/GDP
China	5.64%
France	0.89
Sweden	0.89
Germany	0.71
U.K.	0.91
U.S.	0.55

Source: OECDiLibrary, Total Inland Transportation Infrastructure Investment, 2019, per GDP, stats.oecd.org, accessed 5/31/2021.

The MGI study also employs a second approach to estimating the optimal stock of infrastructure capital: assuming a desirable ratio of the value of infrastructure capital to GDP, based on historical patterns. While this approach is based on the stock of capital rather than the flow of new investment, it suffers from the same limitations as the historical, investment-as-a-share of GDP analysis. It does not consider either the current costs, or current benefits, of modifying the infrastructure capital stock. MGI reports that the infrastructure stock for most economies averages about 70% of GDP. Under the assumption that this reflects some long-run optimum, the study then calculates the amount of annual spending needed for infrastructure to reach and remain at that level. Globally, this calculation suggests that $67 trillion of infrastructure investment is needed between 2013 and 2030. This approach, which yields a similar answer to that from the spending-as-a-share-of-GDP analysis, also suffers from similar shortcomings. The first references the average historical flow of spending, and the second, the average historical stock of infrastructure capital. If the stock were constant as a share of GDP, however, the observed flow of spending would indicate the level of spending needed to maintain that stock. The stock approach and the flow approach could only diverge if the stock was rising or falling significantly during the historical period being studied, as it would be, for example, in China. For the United States, with a relatively stable infrastructure capital-to-GDP ratio, the two approaches unsurprisingly yield estimates of infrastructure need that are very similar.

The MGI approach provides an uncertain guide for the appropriate level of future spending, because there is little guarantee that spending in the past was at the right level. Indeed, much of the discussion behind an infrastructure agenda assumes that the United States has been spending too little. While applying spending ratios from other nations to the United States suggests substantial infrastructure need, it is not obvious that the ratio of infrastructure spending to GDP should be the same in the United States as it is other countries with lower per-capita income. Optimal infrastructure spending is likely to be higher when a highway system is being first laid down, as it was in the United States in the 1950s and 1960s, than when that highway system is mature, as it is today. A backward-looking or cross-country comparative approach also neglects potential differences, over time or across nations, in the cost of building infrastructure. The United States today faces higher costs of construction than other developed nations, which could translate into a smaller optimal infrastructure capital stock than elsewhere.

The third approach to estimating the optimal level of infrastructure in the MGI study relies on third-party estimates of future asset-class-specific infrastructure demand. Estimates are drawn from the Organization for Economic Cooperation and Development (OECD), the International Energy Agency (IEA), and Global Water Intelligence (GWI). The OECD's numbers are "central projections … derived from a Reference Scenario, based on a set of assumptions about government policies, macroeconomic conditions, population growth, energy prices and technology." These projections are not derived from any cost-benefit framework, and they embody important assumptions about future policies, in particular with regard to regulations related to climate change and the evolution of the energy economy. These figures are best understood as estimates of the amount of infrastructure needed to deliver the future quantities of electricity, water, and transportation services that their models predict will be needed.

Neither the ASCE estimate of the spending needed to raise infrastructure grades, nor the MGI estimates of infrastructure gap, indicate how many infrastructure projects have costs that fall below the best estimate of their benefits, and how much would it cost to invest in all such projects?

3.c. The cost of meeting infrastructure needs: lead pipes, safe bridges, and robust dams

For some categories of infrastructure, estimates of "need" are accompanied by the observation that those elements of infrastructure can fail catastrophically if they are not maintained. The most commonly emphasized threats are collapsing dams, falling bridges, and lead poisoning from aged pipes. While these risks are real, the amounts needed to reduce them are only a small part of the aggregate infrastructure needs that are currently reported.

ASCE suggests that over the next two decades, the United States needs to spend $109 billion ($2019) per year on water infrastructure to close the water infrastructure gap. Yet, as Tabuchi (2017) reports, the average one-time cost for replacing a pipe is approximately $5,000 and there are approximately 10 million lead pipes remaining in the United States. That number fits closely with the $45 billion budgeted for lead-pipe removal in the American Jobs Plan.[3] This represents a one-time outlay that is less than one-half of the annual spending the ASCE recommends. Much of the additional spending may apply to water infrastructure with more modest benefits than lead pipe replacement.

With regard to bridges, the DOT reports a disturbing rise in the number of bridges in poor condition. Deadly bridge collapses were more common in the 1980s than in recent years, but in 2007, 13 people died in the collapse of the I-35 bridge in Minneapolis.[4] The DOT (2019) estimates that an annual investment of $12.9 billion is necessary to maintain the current condition and performance of U.S. bridges, and suggests $22.7 billion as the proposed spending level to generate improvement, a value that ASCE cites uncritically.

The initial American Jobs Plan proposed $115 billion to upgrade the roads and bridges that are in most critical need of repair. The difference between the DOT's recommendation, $22.7 billion, and current spending is about $8.3 billion of spending per year, so a $115 billion budget for bridges could fund such an increase for nearly 14 years. Moreover, the DOT report does not indicate that this level of spending is needed to avoid catastrophic collapse, only that it would "improve conditions and performance."

Like bridges, dams present a risk of catastrophic failure. While the Johnstown Flood that followed that failure of the South Fork Dam killed thousands of downstream residents in 1889, recent dam failures have involved far fewer fatalities. Over the past 30 years, "dam failures" have typically meant water overtopping dams, such

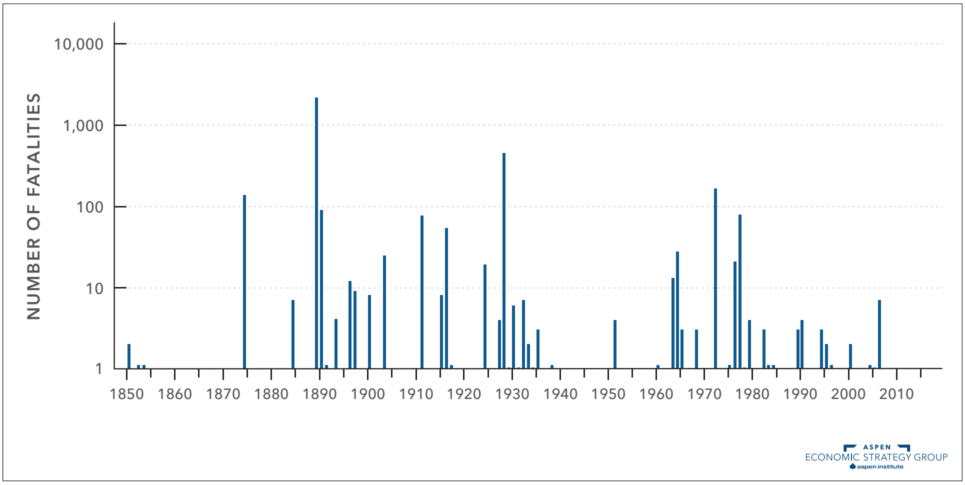
3 These funds are supposed to be given to the Environmental Protection Agency to add to its Drinking Water State Revolving Fund and to disburse the Water Infrastructure Improvement for the Nation grants. This fund and those grants, as currently specified, do many things beyond replacing lead pipes.

4 Penn (2018) reports that many of the deadliest bridge disasters involve the collapse of pedestrian bridges.

as Michigan's Edenville Dam (in 2020) and Iowa's Delhi Dam (in 2010) due to heavy rains, and they have rarely been deadly. The biggest near miss occurred in 2017 when 180,000 people were evacuated from areas downstream of the Oroville dam in California, but the dam held (KCRA 2017).

Figure 2 shows the number of dam fatalities by year in the United States in each year since 1850.

Figure 2: Timeline of fatal dam failures, United States



Source: Stanford University National Performance of Dams Program, accessed 5/31/2021 at http://npdp.stanford.edu/consequences_fatalities.

While the potential risks from dam-related catastrophes are real, the Association of State Dam Safety Officials (2017) estimates that rehabilitating all federally owned dams would only cost $4.2 billion, with $2.9 billion of that amount targeted to "high hazard" dams. Over two-thirds of all dams are owned by the private sector, and for them, the costs of rehabilitation are estimated to be much higher: $60.7 billion, with $18.7 billion devoted to "high hazard" dams. Presumably, the federal government's role should be to provide regulatory oversight and safety inspections, and then require the private owners to pay for maintenance. The relevant budgetary cost in this case, for inspections and enforcement, is likely to be a small fraction of the cost of dam repairs.

3.d. Universal wired broadband vs. alternatives?

The COVID-19 experience of remote schooling strengthened the case for investing in broadband in lower-density parts of the United States. Even before the pandemic, rural broadband access was on a sharply rising trajectory, due both to private market initiatives and public subsidies. The Federal Communications Commission (FCC 2021) reports that at the end of 2019, 94% of Americans lived in areas with access both to 25/3 Mbps fixed broadband service—what the FCC defines as high-speed broadband—and 10/3 Mbps mobile broadband service. Between 2016 and 2019, the number of rural residents lacking access to 25/3 service fell 46%. In 2019, 17% of rural residents did not have access to such service.

Public policy has been subsidizing rural broadband since the 1996 Telecom Act, which taxed telephone calls to finance rural broadband subsidies (Greenstein 2021). The Connect America Fund, established in 2011 by the FCC, subsidizes the development of rural high-speed broadband, Boik (2017) estimates that this costs $4.5 billion per year. There have also been sporadic bursts of investment in broadband for lower-income individuals or lower-density areas, such as the 2009 American Recovery and Reinvestment Act.

A key question is whether, for those who do not currently have access to 25/3 Mbps service, delivering such service requires fiber optic cable access, which can be expensive to provide to some remote areas, or whether other technologies, such as satellite broadband or 5G network access, can serve as a cost-effective alternative. The benefit-cost ratio for the three technologies is likely to vary by place, with fiber optic installation more expensive per household in very remote, low-density areas, and 5G only feasible in some areas with favorable local topography for line-of-site transmission. Boik (2017) examines a subsidy for broadband adoption in North Carolina and finds that many households seem to find satellite broadband an attractive alternative to high-speed wired broadband. He finds that "fewer than 43% of households adopt high-speed broadband in areas currently served by a single broadband provider," and relatively low willingness to pay for high-speed broadband among significant numbers of households that currently use slower options, DSL, or satellite. He concludes that "at most 64% of unserved census block regions in North Carolina warrant an entry subsidy to provide broadband quality comparable to urban areas," and that the cost of bringing high-speed broadband to households in the least dense 5% of the state would only be warranted if these households valued this service at more than $1500 per month.

Current satellite broadband speeds are now fast enough for most conferencing software applications, although since satellite data plans are typically less generous, full time students may run up against hard data walls. These data caps could cause hardship for families living in low-density areas if schools shift to online learning during another future emergency. Nonetheless, these financial shortfalls could be met with school-based subsidies, which might be much less expensive than a complete rural build-out of high-speed wired broadband. If high speeds are deemed to be essential, a central question is whether satellite broadband is fast enough. The satellite provider Viasat offers a 100 Mbps download speed plan. If that option meets other technical needs, then the high costs of providing hardwire broadband should be compared with the costs of subsidizing high-speed satellite service in remote areas. As with pricing infrastructure services as an alternative to building more infrastructure, multiple approaches to achieving the overall policy objective—in this case high-speed internet access—should be considered. Either alternative should be subject to standard cost-benefit analysis, the topic we now consider.

4. Cost-benefit analysis and infrastructure spending decisions

The application of project-specific cost-benefit analysis to individual infrastructure projects, and the aggregation of the results, is quite different from a budgeting process that seeks to come up with an aggregate spending number and then to enact legislation to spend that amount. Theoretically, it would be possible to do cost-benefit analyses on a vast number of projects, select only those projects with benefits greater than all-inclusive cost of raising the relevant funds, and to add the cost of those projects up to produce an optimal level of infrastructure spending. Yet that is not the way resource allocation decisions for infrastructure operate at present. In this section, we will briefly review the application of cost-benefit analysis to infrastructure projects, and then discuss ways in which that analysis might play a larger role in policy discussions going forward.

A key but not surprising insight of the project-based, bottom-up cost-benefit analysis approach, rather than the top-down, aggregate spending target approach, is that there is likely to be substantial heterogeneity in the returns to different projects within an infrastructure category. Just as subsidizing satellite broadband in some remote areas may offer a better cost-benefit trade-off than building fiber optic lines, some unsafe bridges in areas that attract relatively little traffic may be better closed and demolished than rebuilt. The cost of rebuilding may exceed all reasonable estimates of the return.

The starting point of cost-benefit analysis is calculating benefits by projecting the future uses of the infrastructure, evaluating the value of those uses, and then discounting that benefit flow using an interest rate, such as the government's cost of funds, adjusting as needed for risk. The benefits are then compared with the project's cost, which typically involves a large up-front cost as well as future periodic maintenance outlays, which must also be predicted and discounted.

While there is near universal endorsement of cost-benefit analysis among economists, it is not the cornerstone of infrastructure policy analysis. The time involved in estimating the returns to any one project are significant, and that makes the cost-benefit approach unattractive to anyone trying to craft legislation quickly. Moreover, cost-benefit analysis involves considerable uncertainty and inevitably, the evaluator has opportunities to exercise discretion. It is important to try, where possible, to develop institutions that can perform non-partisan, rigorous cost-benefit calculations. The public may, appropriately, be skeptical of estimates of the benefits of a bridge over a 30-year horizon. Cost-benefit analysis is often most useful when the gap between benefits and costs is large, which means that the project clearly should or should not be funded.

4.a. Cost-benefit fundamentals

There are a number of features of cost-benefit analysis that are similar across many forms of physical infrastructure. We identify six such components, and describe each of them with reference to transportation infrastructure projects.

4.a.1. Estimating future benefits.

The first ingredient of cost-benefit analysis involves estimating future usage of the infrastructure, and the benefits that flow to each user. In a standard private sector investment decision, payments to the investment's owner, such as profits on a business or royalties on a patent, are the benefits to the investor. Since users of publicly funded projects can often access them at a nominal cost, the benefits to users are usually assumed to be greater than the ticket price. Cost-benefit analyses in the transportation sector have been plagued by erroneous predictions for decades, with project boosters often overpredicting future expected demand (Kain 2007). Conversely, costs have frequently been underestimated, often by a wide margin. The rosy $35 billion projection of the cost of high-speed rail in California, made by the engineering firm Parsons Brinckerhoff in 2014, is a high-profile example (Parsons Brinckerhoff 2014). The next year, that firm won a $700 million contract to provide management services related to the high-speed rail system (Railway Technology 2015). By 2021, the estimated cost had reached $100 billion (Vartabedian 2021). While the mis-estimate may have been entirely innocent, engineering firms that stand to benefit from building infrastructure may have a conflict of interest when reporting on the benefits and costs of that infrastructure.

4.a.2. Measuring systemic impacts

A second ingredient of cost-benefit analysis for transportation projects involves measuring the impact of a new project on the usage of other routes and modes of transportation. A new rail line may alleviate the traffic on highways. Estimating the links between different modes, or even different routes within a given mode, is an even more difficult problem, one that can require sophisticated analytical tools. For example, Allen and Arkolakis (2019) develop a network model and estimate the system-wide benefits of building new highway capacity. Their results, shown in Table 3, find particularly high benefits to adding new highway lanes in some parts of the New York metropolitan area, which would seem to argue for more construction there, but they have limited information on the costs of adding new lanes in that area. More generally, their analysis highlights the range of benefits relative to costs for highway construction. There are of course many projects with much lower benefit-cost ratios than those in Table 3, including numerous projects with values less than one. The application of the network model illustrates the need, in some cases, for detailed analytical work that goes beyond the proposed project. Developing the capacity for that, perhaps with a federal infrastructure bank, could be an important part of an infrastructure initiative.

Table 3: Examples of high benefit-cost ratios for additions to interstate highway system

Project Location	Estimated Benefits ($M /year)	Estimated Costs ($M/year)	Benefit-to-Cost Ratio
White Plains, NY to Greenburgh, NY	$510.5	$3.8	135.8
North Hempstead, NY to Queens, NY	719.5	5.4	134.5
Islip, NY to Brookhaven, NY	257.5	1.9	135.5
Indianapolis, IN	206.9	2.2	100.4
Bayonne, NJ to Staten Island, NY	179.9	1.9	93.7

Source: Allen and Arkolakis (2019)

4.a.3. Assessing ancillary benefits and costs

A third step in cost-benefit analysis involves studying and evaluating the ancillary benefits of changing travel patterns. For example, new infrastructure spending might affect the total amount of carbon emitted in the United States, and the direction of the effect is likely to depend on the nature of the projects supported. The American Jobs Plan emphasizes investments that might reduce carbon emissions. Achieving that goal through new infrastructure requires a strong degree of substitution between new, low-carbon forms of transportation and older, more traditional modes. The degree of such substitution is an open question. A new rail line might reduce carbon emissions if it leads to reduced car traffic. Baum-Snow, Kahn, and Voith (2005), however, find that the train-for-car substitution is modest at best, which implies that the carbon emitted in building the rail line and in operating it every day may result, on net, in an *increase* in carbon emissions. To calculate the ancillary benefits of infrastructure construction that accrue through environmental channels, the estimated change in carbon emissions from new infrastructure must be multiplied by the welfare cost of carbon emissions. The precise magnitude of this cost is a subject of active debate (Stern and Stiglitz 2021).

4.a.4. Estimating macroeconomic effects

A fourth component involves measuring the macroeconomic effects of infrastructure projects, such as anti-recessionary stimulus and agglomeration economies. While infrastructure spending is often advanced as a job-creating program, there is uncertainty about the number of jobs that are created by each dollar spent on infrastructure, and about the social value that should be placed on such jobs. One estimate, using data from the American Recovery and Reinvestment Act (2009), is that $200,000 of infrastructure spending creates about one job for one year, while the spending is going on, although there is considerable uncertainty about that figure (Garin 2019). The jobs-per-infrastructure dollar ratio is likely to vary with the type of project and with the broader economic conditions that prevail when the project is undertaken.

The macroeconomic community is split about the value of infrastructure as a tool for fighting recessions. There are long-standing concerns about the capacity to time infrastructure spending to coincide with periods of economic slack. The prospect of long and variable lags in the implementation of fiscal policy, including infrastructure spending, was an important factor in the shift from fiscal to monetary policy as a primary tool for macroeconomic stabilization in the decades prior to the global financial crisis. The costs of labor can be lower in recessions, although often not by much, and that suggests that ordinary cost-benefit analysis should push infrastructure construction toward downturns. However, as long as the employment impacts of infrastructure spending remain uncertain, it will be difficult to resolve the differences between the advocates of counter-cyclical infrastructure spending, such as Summers (1988, 2017), and those who see minor macroeconomic effects at best, such as Ramey (2021).

It is also difficult to estimate the intrinsic benefits of job creation, which come from reduced spending on government-provided unemployment benefits as well as from personal benefits such as improved self-esteem. A reasonable approach to estimating these benefits involves multiplying three numbers: (1) the projected number of workers on the project; (2) the increase in total employment per worker hired, which captures the degree to which infrastructure employment crowds out other employment; and (3) the social benefits of switching workers from nonemployment to employment. The estimates of Garin (2019) and others indicate that even in a recession, public infrastructure employment significantly crowds out private employment, so the increase in employment per worker hired is likely to be far less than one.

4.a.5. Measuring impacts on GDP and productivity

A fifth element in cost-benefit calculations is the project's impact on economy-wide output and productivity. These effects are distinct from the jobs created in building the new project. New infrastructure may enhance the productivity of businesses, thereby raising total output. These benefits are linked primarily to use of the project, making it particularly important to accurately assess the prospective utilization. New infrastructure may also cause a relocation of economic activity that generates local externalities, such as agglomeration effects, which are benefits that accrue when firms and people locate near one another, thereby reducing transportation costs. Some evaluations of transportation infrastructure include agglomeration effects, which occur when an increase in the scale of a community increases the output of each of its members. London's Crossrail, for example, was supposed to create large-scale agglomeration benefits (Bhasin 2007). Yet if infrastructure projects just move people and activity from one area to another, then there will be offsetting agglomeration losses in the shrinking place, which must be weighed against the agglomeration benefits from the expanding area. Glaeser and Gottlieb (2009) argue that there is little economic certainty about the magnitude of these different effects.

4.a.6. Considering distributional impacts

A final component of cost-benefit analysis is the recognition of the distribution of project benefits and costs across the population. The standard approach is to treat benefits to one group as equivalent to benefits to another, and to sum the net benefits across groups. Alternatively, however, losses to vulnerable populations can be treated as far more serious than benefits to the prosperous. A dollar lost by the poor could be treated as the equivalent to two dollars gained by the rich. A job created in a low-income neighborhood could be valued more than a job created in a high-income location. Of course, such group-based weights must reflect moral and political values, not economic estimates, but cost-benefit analysis can always provide a range of estimates depending on the weights that are assigned to different populations. Similarly, it is possible to weigh losses more heavily than gains for all groups. The cost-benefit framework is flexible and can accommodate a wide range of social values.

As critical as cost-benefit analysis may be, there is little chance that a project-by-project analysis can be undertaken in the short time available when legislators begin debating a national infrastructure bill. This is surely one of the reasons a top-down, select-a-budget total approach is more common. One way to address this challenge is to maintain ongoing cost-benefit analysis in relevant federal agencies, such as DOT. Another, which can capture some of the benefits of cost-benefit analysis, is to apply this approach after the budget total has been determined. While the budget total may not be the same as the one that would arise from bottom-up cost benefit analysis, the allocation of the funds across projects will target those with the highest estimated benefit-to-cost ratio.

4.b. Expanding the role for cost-benefit analysis in the policy process

There are three ways in which cost-benefit analysis could be inserted more directly into the process of allocating U.S. infrastructure spending, even if bottom-up cost-benefit analysis is not possible. One is to focus on estimates of the benefits of infrastructure spending as a whole, rather than the benefits of particular projects. Increasing national spending on infrastructure makes sense when benefits per dollar spent are greater than the social cost of raising one dollar in taxes. The second is the creation of an infrastructure bank that would receive some fraction of federal infrastructure spending and deliberately allocate the funds to projects that appear to have particularly high benefits relative to costs. The third option would require states to make more use of cost-benefit analysis when they spend federal dollars, perhaps with input from a federal agency that develops and applies cost-benefit methods.

4.b.1. Applying cost-benefit analysis to overall spending levels

The first option is essentially "macro cost-benefit analysis." Instead of trying to figure out the impact of an individual bridge or highway, this begins by estimating the social benefit of spending an incremental amount, say $1 billion, on infrastructure overall or on a particular type of infrastructure such as highways. While this approach cannot ensure that all infrastructure projects deliver benefits greater than their costs, as long as future project choices resemble past ones, this approach provides a way for determining the return on new infrastructure spending.

There is a substantial literature on the aggregate output and productivity effects of infrastructure spending.[5] Estimates of the link between infrastructure spending

5 Pioneering studies by Aschauer (1989) and Munnell (1990) found a significant positive correlation between infrastructure capital and economic activity, while calling attention to the potential endogeneity of infrastructure spending. Shirley and Winston (2004), Gramlich (1994), and the CBO (1988) report that early post-war infrastructure investments had large returns, but that the returns on subsequent investments have been lower. Bom and Ligthart (2014) review the literature on infrastructure capital and aggregate output. Schanzenbach, Nunn, and Nantz (2017) summarize several recent empirical studies of how infrastructure affects productivity. Ramey (2021) includes infrastructure capital in a neo-Keynesian macro model, finding only modest productivity and output effects.

and productivity are not precise, however, and many such estimates are confounded by the potential endogeneity of infrastructure spending. Furthermore, determining the causal effect of infrastructure on economic growth is not easy to do. If states spend on infrastructure in anticipation of future growth, then it might look like infrastructure is causing growth, even though it is the anticipation of growth that is causing the spending. If other state attributes, such as lower density levels, that are associated with more spending exert an independent pull on economic activity, then empirical estimates will also be misleading.

There is another difficulty with this approach: Economic activity does not automatically represent social benefit. A dollar of GDP is not a dollar of extra welfare, since presumably there was some cost, such as the workers' time, of producing that GDP. Moreover, local GDP, which is often the outcome used in empirical analyses of infrastructure productivity, can increase because activity is displaced from one area to another. Estimates of infrastructure productivity based on local outcomes may tell us very little about aggregate economic activity.

4.b.2. Creating an infrastructure bank or adopting cost-benefit mandates

Even if cost-benefit analysis cannot provide a number for optimal overall infrastructure spending, the tools of cost-benefit analysis can be used to allocate appropriated funds across different projects. There are two natural ways to use these tools to improve the targeting of spending. The first possibility, which was originally proposed by Senators Chris Dodd (D-CT) and Chuck Hagel (R-NE) in 2007 and was much discussed during the late Obama administration, is for infrastructure spending to be allocated by a national infrastructure bank that would use cost-benefit analysis. The second possibility is to require cost-benefit analysis before states are granted federal funds for new infrastructure projects.

The basic idea of an infrastructure bank is to establish an independent entity with some form of appointed leadership, possibly subject to Senate confirmation, that would oversee a significant amount of infrastructure spending. A national infrastructure bank would have similarities to the World Bank or the Asian Development Bank. These institutions specialize in funding projects that are typically implemented by some other entity. In the U.S. context, states, localities, and public-private partnerships would ultimately be in charge of implementation. If sufficient resources were devoted to new infrastructure investments, it would be possible for the infrastructure bank to develop a robust cost-benefit analysis process, and to use the results of that process to determine funding. The bank could carry out cost-benefit analysis and determine which projects should be funded from a pool of resources provided by Congress. That would amount to ranking potential projects and funding the highest benefit-to-cost projects until funding is exhausted. Additionally, the infrastructure bank could provide guidance to legislators on the level of infrastructure spending that might be warranted by high benefit-to-cost projects, the result of ongoing analysis of potential projects. The entity would have some discretion for allocating spending, but its key objective would be to fund those projects with the highest level of net benefits. The track record of the international entities should motivate caution about the capacity of independent "banks" to always target the highest value-added projects. That objective could be written into law, but ultimately the entity's leadership would need to be selected so that they shared that objective. An infrastructure bank also might, more easily, time its spending to coincide with downturns (Haughwout 2019).

An alternative way to expand the use of cost-benefit analysis is to continue with the current procedure of providing funds to states and allowing them to make allocation decisions, but to subject them to cost-benefit related requirements. For example, new projects might have to meet a fixed internal rate-of-return threshold in order to go forward with federal support. Requiring cost-benefit analysis for the maintenance of existing infrastructure makes less sense; there is more consensus about the high rates of return for maintaining the existing infrastructure stock.

In this model of state autonomy checked by federal oversight, cost-benefit analysis must be done by an independent entity. If these regulations were to be imposed on states, the federal government would need to create and fund an agency capable of appraising state projects. The CBO provides one model of such an entity. Presumably, the evaluation organization would have close ties to the DOT, but it would ideally be sufficiently independent and apolitical so that its judgments would carry widespread respect. All cost-benefit analysis is subject to gaming, since assumptions about inputs are critical to the outcomes. When those who provide the estimates of costs and benefits are able to inflate the former and understate the latter, the results of the analysis may not result in an appropriate ranking of potential projects. Rather than selecting the most attractive projects, the use of cost-benefit analysis may only identify the projects with proponents with the greatest proclivity to overstate benefits relative to costs.

An infrastructure bank, or a federal requirement for state cost-benefit analysis, would run into some potential challenges. The infrastructure bank creates more executive branch discretion and therefore carries more risk of mismanagement, both in itself and because it needs to work through other entities, like state governments. If states are choosing and administering their own projects, the basic incentives are better aligned. When state governments are spending a fixed sum of money, they face stronger incentives to keep costs down than if they are spending the funds of an infrastructure bank. Of course, it may be possible for the infrastructure bank to design incentive contracts that restrict waste and abuse.

An infrastructure bank might have the additional effect of catalyzing public-private partnerships, a topic we will address in more detail below. Some states, like California and Texas, have been far more aggressive than others in supporting public-private partnerships, and a state-level approach might not make much headway in states that have been reluctant to adopt this approach. One important worry, which emerges from the work of Engel, Fischer, and Galetovic (2014), is that private companies have incentives to, and often succeed in subverting government agencies. The legislature would need to remain vigilant to ensure that the national infrastructure bank was not captured by related private companies.

With regard to cost-benefit analysis mandates, while there is an added cost of carrying out expanded cost-benefit analysis, and there could be delays in launching projects until such an analysis was complete, it seems difficult to object to requiring simple cost-benefit analyses for new federally funded transportation projects. The costs of these analyses are very low relative to the costs of infrastructure. Cost-benefit analysis might still permit some white elephant projects to go forward, but it is likely to be an improvement relative to the status quo. Imposing rate-of-return requirements may be difficult to do, however, in a way that passes constitutional muster. Another limitation is that such a process would not determine the allocation of funds across states, although it might be possible to use the results of start-level cost-benefit analysis to inform Congressional debates on allocation.

An infrastructure bank has more upside and downside risk than cost-benefit analysis mandates. In principle, it could be a nimble and intelligent agency that chooses really high return projects throughout the United States. It could also become a political tool that is largely beholden to pet ideas of both legislators and the administration.

5. The cost conundrum for new infrastructure projects

Much of the discussion around the need for additional infrastructure focuses on the benefits from additional investment. The optimal amount of infrastructure capital is also a function of its cost, and by international standards, infrastructure projects in the United States are extraordinarily expensive. The basic logic of cost-benefit analysis thus suggests that the United States should, all else equal, have less infrastructure that other comparable nations. If fixing potholes is more expensive in the United States than in other countries, one would expect to find more potholes here.[6] Given the high cost of U.S. infrastructure projects, even large increases in spending may have only modest effects on the quality of infrastructure services. Two key questions are therefore why infrastructure construction costs are so high in the United States, and whether it is possible to reduce them.

5.a. Why does it cost so much to build infrastructure in the U.S.?

Flyvbjerg, Bruzelius, and van Wee (2008) compare the capital costs for urban rail projects around the world. The costs for the six systems in the United States that were included in the analysis range from $88 million per kilometer (Atlanta) to $147.5 million per kilometer (Baltimore). Thirteen out of 17 of the European systems in the study, and five out of six in Asia or Latin America, had costs below $88 million per kilometer. Levy (2011) argues that these comparisons understate the cost disadvantage of U.S. projects, noting that “the American projects examined are quite old, from the 1980s, and many have large above-ground parts.” He further identifies three New York City projects with costs of $1.3 billion, $1.7 billion, and $4 billion per kilometer, as well as San Francisco’s Central Subway, which cost $500 million per kilometer even though, as a light rail tunnel, it was a less demanding project.

Levy’s updated Transit Cost database reports actual or projected cost-per-kilometer data, converted to $US using purchasing power parity exchange rates, on 540 different projects, including 256 that were completed by 2020. We inflation-adjust these cost estimates using the CPI and assume that the median dollar was spent in the year that was half-way between the start and end date of the project. We treat projects with average years beyond 2021 as having an average year of 2021.[7] Table 4 presents our findings.

For the 19 projects in the database that are in the United States, the average cost was $1,601 million per mile, compared with a non-U.S. global average of $478 million. The median U.S. project was $965 million per mile, compared with a non-U.S. median of $299 million. The database also contains information on the share of the rail system that is underground. When we restrict our analysis to the 255 projects that are 100% in tunnels, the median cost of the 11 U.S. projects is $1,379 million per mile, compared with a non-U.S. global median of $341 million. While precise comparisons are difficult, many projects in densely populated foreign cities have substantially lower per-mile costs than their U.S. counterparts.

---

6 However, this logic does not imply that the United States should *spend* less than other nations on infrastructure—spending is the product of the price of infrastructure and the quantity purchased.

7 This assumption seemed reasonable to us since estimates of future costs frequently fail to incorporate inflation. The Transit Cost database uses the middle of the state and end year for purchasing power adjustments.

Table 4. Cost-per-mile of large U.S. urban transit projects

City	Project Name	Start Date	End Date	Cost/Mile ($M)
Seattle	U-Link	2009	2016	637
Los Angeles	Purple Phase 3	2020	2027	1379
Los Angeles	Purple Phase 2	2018	2026	920
Los Angeles	Purple Phase 1	2014	2023	758
Los Angeles	Regional Connector	2014	2022	966
San Francisco	Central Subway	2010	2021	1115
Boston	Green Line Extension	2013	2021	523
San Francisco	BART to San Jose	2022	2030	1157
New York	7 extension	2007	2014	2921
New York	Second Avenue Phase 1	2007	2016	3156
New York	Second Avenue Phase 2	2019	2029	4271
New York	East Side Access	2007	2022	7081
New York	Gateway	2019	2026	2885
Honolulu	HART	2011	2026	528
Los Angeles	Crenshaw/LAX Line	2014	2021	266
Miami	Metrorail extension to MIA	2009	2012	253
Seattle	West Seattle and Ballard	2026	2036	1045
Washington	Silver Line Phase 1	2009	2014	304
Washington	Silver Line Phase 2	2013	2021	264

Note: The original source, https://transitcosts.com/data/, reports cost estimates in current dollars. All estimates have been converted to 2021 dollars.

There is no widely accepted source of global comparative data for highway costs, but one analysis by Brooks and Liscow (2021) finds that the United States has the highest highway construction costs in the world. U.S. highway construction costs have also risen over time. "Spending per mile on Interstate construction increased more than three-fold (in real terms) from the 1960s to the 1980s," a finding that is particularly remarkable because "neither changes in the observed geography of spending nor increases in material and labor prices explain these changes" (Brooks and Liscow 2019). The DOT National Highway Construction Cost Index increased by 32%, relative to the CPI, between 2003 and 2020.

The high costs of U.S. infrastructure can be analyzed at two levels: in an accounting sense, by asking which items add so much to the bill, and at a deeper level, by asking why the prices of some inputs are particularly expensive. These analytical approaches can be applied to consider the costs of the single most expensive project in the Transit Cost Database: New York City's East Side Access project. This completely underground project, at $4 billion per kilometer, is more than 20 times more expensive than the average all-tunnel project in other countries. Barone, Vitullo-Martin, and Pichardo (hereafter BVP) (2018) dissect the high cost of that project as well as the Second Avenue Subway and the #7 line extension, also in New York City. The two other projects are less expensive than East Side Access, but at $2 billion and $1.8 billion per kilometer, respectively, in inflation-adjusted terms, they are still ten times more expensive than the global median for urban rail projects.

East Side Access' $12 billion costs, as of 2016, included $9.7 billion of construction costs, expansively defined. These include what BVP (2018) categorize as construction ($7.3 billion), construction and production management ($890 million), design and engineering ($660 million) and vehicles and spare parts ($800 million). The other $2.3 billion reflects finance charges ($1.12 billion), unallocated contingency money ($720 million), administrative and regulatory costs ($259 million), and real estate and relocation ($192 million).

The cost breakdown highlights a number of important patterns. First, real estate costs are a tiny share of the project's total cost, despite New York City's sky-high property prices. Second, neither administrative and regulatory costs, nor the even smaller category of environmental mitigation (contained within construction and only $2.14 million) were significant causes of the high costs. Third, the two largest elements in construction costs were tunneling ($3.1 billion) and stations and intermodal facilities ($2.3 billion). The very expensive station construction is one reason why East Side Access was the most expensive project in the database, but the tunneling on its own is extraordinarily costly by global standards.

While direct environmental mitigation itself was a small component of the accounting costs, environmental factors play a much larger role in the overall cost of the project by changing the nature of construction itself. For example, BVP (2018) explain that "the Environmental Impact Statement (EIS) required construction activities in Manhattan to take place in the subterranean realm, with almost all equipment and spoils transitioning through the 63rd Street tunnels to staging sites in Sunnyside Yards," which typically meant "laborers filling burlap bags with spoils that were then loaded onto trains to Queens (or in some cases, the Bronx) and then unloaded and sorted by laborers." The Metropolitan Transit Authority (MTA) estimated that it could have saved $75 million in schedule-related costs alone by deploying a simpler system similar to that used in other projects.

Labor costs and procurement problems seem particularly critical in contributing to higher expenses associated with East Side Access.[8] New York City pays its infrastructure workers very high wages. BVP (2018) estimate the minimum labor costs for electricians at $127 per hour, for tunnel workers at $102 per hour, and for cement and concrete workers at $57 per hour. Labor costs would be higher on Sundays. The total labor costs on the project were between $2.9 and $4.3 billion. With regard to procurement, BVP (2018) report that "the contractual history of [East Side Access] is replete with examples of practices and decisions that led to unnecessary delays, defaults, and costs." They highlight in particular the decision by the MTA in 2012 that "all bids on Contract Modification 12 (CM12) were too high, upward of $950 million." This led to the cancellation of the bids. MTA then divided the work in CM12 into three sub-projects, which caused a delay that BVP (2018) estimate at three years, and a cost increase of at least $373 million overall.

Bosio et al. (2020) examine highway procurement globally, and find that in poorly governed countries, strict procurement rules lead to less corruption and better outcomes. They find the opposite in well-governed countries, where procurement rules limit the ability of project managers to avoid problematic companies who offer low bids. In New York City, the MTA strictly adheres to a low-bid rule. BVP (2018) note that "the adherence to accepting only the lowest qualified bid has led to less-experienced contractors defaulting on contracts."

The East Side Access project included $300 million in site preparation. The Second Avenue Subway required $335 million in site preparation, which came to 11% of its total construction budget. Site preparation is particularly difficult in New York City, because of the preponderance of electrical wires and pipes that are underground. This process involves bargaining between the MTA and New York City's utilities; it is not clear whether the infrastructure sponsors are striking their best bargains. Barro (2019) writes that, "if the city or the state brought more of its utility-oversight powers to bear to hold down costs for the MTA, we might be able to take a bite out of this particular cost problem."

Beyond accounting, there are three deeper explanations for why infrastructure costs are so high in the United States. First, it is possible that conditions are more demanding and that raw materials and labor are more expensive than elsewhere. This explanation suggests that higher costs are unavoidable but should be considered in discussions of optimal infrastructure policy, since higher costs are a good reason to build less infrastructure. Second, it is possible that the agencies charged with building infrastructure are poorly designed to manage costs. In this case, there may

8 Brooks and Liscow (2021) find that rising labor costs do not contribute substantially to the time series increase in construction costs between the 1960s and the 1980s. It may still be the case that labor costs in some large cities contribute to high infrastructure costs there.

be changes to infrastructure building practices that could lower costs and stretch infrastructure budgets. Third, it is possible that external factors, especially the threat of litigation or political backlash, lead to expensive forms of mitigation, which change the nature and cost of building projects (Altshuler and Luberoff 2005).

How can we assess these three competing explanations? Arguably, the conditions for tunneling in Manhattan are as difficult as anywhere in the world, although cost estimates for projects in London, which are all completely underground, are only one-third as high as those in New York City. Labor costs are higher in the United States than elsewhere, and especially so in New York City, but this reflects institutions as well as generally high labor costs. The Bureau of Labor Statistics (BLS) reports that in May 2020, the median hourly wage for electricians in the New York City metropolitan area was $36.13, and the mean was $40.48.[9] BVP (2018) report a minimum hourly wage for electricians of $65 on the East Side Access project and an added $62 dollars per hour in benefits, making the per-unit labor cost for the project a multiple of the prevailing wage.

Moving beyond labor costs, many of the capital goods that are used as inputs to highway construction, as well as most materials, possibly excepting some locally manufactured concrete, are bought and sold in regional or global markets. Even if the law of one price does not hold for these inputs, it is unlikely that deviations in prices across countries are large enough to be able to account for significant project cost differentials.

Procurement rules, which may achieve meritorious social goals, can also raise infrastructure costs. One study finds that allocations for minority contractors in California increase construction costs by 9% (Marion 2009). Such provisions also increase the number of Black-owned business, which highlights the cost-vs.-social goals trade-off (Chatterji, Chay, and Fairlie 2014). It is important to consider these trade-offs before embarking on a major infrastructure spending program.

While the extraordinary price tag of New York City projects reflects in part the challenges of building in an already hyper-dense locale, there is no equivalent explanation for the high costs of building highways in lower-density states that Brooks and Liscow (2021) report. Most of the United States is far less dense than most of Europe, and much of the country is reasonably flat. That turns the spotlight to labor costs. While labor costs may contribute to the high costs of highway work in the United States, the mean hourly wage in the industry labeled “Highway, Street, and Bridge Construction,” is under $30 per hour according to the BLS.[10] The labor

9 https://www.bls.gov/oes/current/oes_35620.htm

10 https://www.bls.gov/oes/current/naics4_237300.htm

share for highway work, as opposed to tunneling in New York City, is less than 30% (Garin 2019). Even 30% higher wages in the United States would only lead to an increase of 9% in total costs. This arithmetic suggests that there is still much to be done in accounting for the higher infrastructure costs in the Unites States.

5.b. Infrastructure costs have risen over time

Between the late 1950s and the early 1990s, Brooks and Liscow (2021) estimate, overall highway construction costs in the United States increased fourfold. They also systematically evaluate the role that input costs and geographic difficulty played in increase the costs of highway construction. They find that the real cost of materials and labor barely changed, so input costs cannot account for the overall increase. To control for changing difficulty in the geography of construction, they measure the average population density, hilliness (slope), and contact with water of new road segments in a state during a given year. They also control for state fixed effects to capture changes in the location of new highway segments; these factors only explain about 6% of the overall increase in the cost of building. They conclude that the increasing cost of accommodating citizen's complaints about the downsides of new highways has been a central source of cost increase.

Brinkman and Lin (2019) discuss "freeway revolts" in which neighborhoods fought to stop nearby road construction. This activity exploded in the 1970s, which also saw a dramatic increase in the number of newspaper articles about the environmental damage associated with interstate highways. Brooks and Liscow (2019) document a dramatic increase in the number of "wiggles" in new roads over time. These increase costs but may allow highways to bypass sensitive areas. The number of ramps and bridges has also increased: these also reduce the need to bulldoze existing structures and increase costs.

These facts are compatible with the narrative arc of Altshuler and Luberoff (2003), who focus on megaprojects, rather than highways. They split the post-war experience into three periods. During the first period, large urban construction projects occurred with little opposition. Robert Moses' New York projects, such as the Cross Bronx Expressway, perhaps epitomize this epoch. In the second period, neighborhood activists, such as Jane Jacobs, borrowed organizing techniques from the civil rights movement and learned how to block infrastructure projects. In the third period, which began in the 1970s and continues to this day, the public sector responded to a more empowered citizenry by avoiding relocation and offering expensive mitigation for the local consequences of new projects. Massachusetts' Big Dig epitomizes this era.[11] Its price tag ballooned from $2.5 billion in 1985 to $14.8 billion in 2008,

11 Fred Salvucci, the Massachusetts transportation secretary who shepherded the Big Dig project, had a grandmother who was relocated by an earlier megaproject. He was determined to complete the project without any relocations and with minimal resident discomfort.

reflecting a combination of delays, modifications to construction plans, and design changes that were adopted in response to various public interest groups. The actual costs may have been as high as $18 billion (Bearfield and Dubnick 2009).

The cost overruns in the Big Dig reflected in part the perpetual problem that optimistic figures are used to sell large projects to the public, but there were also genuine surprises that raised the project's cost. Boston is an old city with a great deal of underground infrastructure, and replacing and relocating unexpected pipes, electrical, and sewer lines added to costs. Moreover, to offset the air pollution that would be associated with the increased vehicle traffic after completion of the Big Dig, the project needed to include funding for pollution mitigation efforts such as the restoration of previously inactive commuter rail lines. We are not aware of any cost-benefit analysis in the selection of these mitigation projects.

With mega-projects, and even with ordinary highway construction, it is hard to distinguish between the cost impact of mitigation, which presumably delivers some value to impacted communities, and the cost impact of procurement and managerial problems. These two factors can compound. Managing a project that involves building a simple straight highway is vastly easier than managing a tunnel project, but tunnels are one way of reducing the impact on neighborhoods. Public bureaucracies that were up to the task of building relatively simple projects during the first era may find managing costs far harder during the current era of more complex construction projects.

5.c. *Making infrastructure more affordable*

Procurement practices, which are largely set at the state and local level, are potentially important determinants of the cost of infrastructure projects. The rules that govern state departments of transportation, labor negotiation, and environmental impact reviews are typically determined by state law and state and city politics. The federal government has access to only blunt tools for modifying the infrastructure procurement process and reigning in costs.

5.c.1. *Apply cost-benefit analysis*

If either a national infrastructure bank or a cost-benefit rule becomes the new norm, then higher project costs will make it harder, for a given level of public benefits, for a project to receive funding. The benefits of new projects would need to be higher in states where costs are particularly high, which could bring pressure to trim costs. It could also lead to more systematic over-statement of benefits in such states, or under-estimation of costs in the project planning stage; both would need to be monitored carefully.

5.c.2. *Purchase from low-cost suppliers*

Relaxing “Buy American” provisions is another way in which the federal government could reduce the cost of new infrastructure projects. Horrox and Casale (2019) claim that the average cost of electric transit buses is $750,000 in the United States, reportedly double that paid by the U.K., which has laxer rules about buying British products. The cost of an electric bus in Asia is lower still. Yet procurement in the bus industry is hampered by national and local regulations (Li, Kahn, and Nickelsburg 2015). Higher costs associated with domestic content rules will make the U.S. conversion to electric buses slower and far more expensive than that conversion elsewhere. Tariffs can have the same effect in raising project costs. The cost of the transition to solar energy will be much higher if tariffs raise the cost of solar panels and related products. There are many arguments for domestic content rules and tariff protection, and it is important to consider alternative policies that could address the underlying policy goals with fewer distortions. For example, when such policies are justified on the grounds of domestic income redistribution, it may be possible to achieve a similar degree of redistribution by other means such as transfer programs without incurring the cost of distorting infrastructure purchase decisions and confounding other price signals.

5.c.3. *Streamline environmental reviews*

If the federal government were committed to more rapid project completion and fewer cost overruns, it could model streamlined environmental reviews for infrastructure projects. This could be done in the context of directly funded projects, such as those under the TIGER/BUILD program described in Congressional Research Service (2019). In the context of this program, the federal government could directly assess whether mitigation expenses are excessive. The federal government also imposes its own environmental impact review process, which can go beyond the environmental impact reviews mandated by the states. The discretionary DOT grants associated with this program represent a small fraction of total infrastructure funding, but they offer an opportunity to make a statement about best practices.[12]

One option would be for DOT to make a public commitment to ensure that the social costs of environmental regulations do not exceed their benefits. It could also commit to increase the speed of these reviews to eliminate the costs of delay. DOT analysts could work with TIGER grant recipients to ensure that mitigation efforts satisfy cost-benefit analysis. Such actions would have a symbolic and informational

12 This analysis assumes that the federal government could move more quickly than the states with regard to project approval. Given state-level heterogeneity, it is possible that some states are already moving quickly relative to what a federal program could deliver. For other states, however, the cost saving could be substantial.

effect, suggesting to states that they should also be asking whether environmental impact reviews are too onerous or whether mitigation effects are excessive, and providing a model of how to do this.

States, which have primary control over transportation within the United States, are best positioned to reduce construction delays and mitigation-related costs. They control state-level environmental impact review processes, labor-related rules and the project choices that drive mitigation costs. Any reform must acknowledge that there often are environmental costs associated with infrastructure projects, and that cost-effective mitigation is appropriate. The key is to determine when benefits exceed costs, and to find ways to expedite project approval.

5.c.4. Harmonize implementation of prevailing wage requirements

The Davis-Bacon Act requires workers on federally funded projects to be paid the prevailing wage, and cost-cutting advocates have long urged its reconsideration. Yet prevailing wages are interpreted quite differently in different locations. In New York City, the Comptroller determines the prevailing wage. For electricians, in 2021 that was $58 per hour, with a supplemental benefit requirement of $58.46. After seven hours in a day, overtime kicks in, causing the wage to increase to $82 per hour and the benefit rate to rise to $62 per hour.[13] While a complete analysis would require more detailed information on the nature of the work being performed and the necessary skill sets for the workers, the Comptroller's prevailing wage is significantly higher than the average wage as reported by the BLS for the New York metropolitan area ($36 per hour). By comparison, in Houston, the prevailing wage is listed as $31 per hour and the benefit level is $9 per hour. The BLS reports that the average wage for an electrician in Houston is $25.47 per hour, about three-quarters of the New York City figure from the BLS. The non-overtime prevailing wage plus benefit for New York City is nearly three times that in Houston, but the effective cost difference may be even larger. Texas follows the general rule that overtime begins after a 40-hour week, while New York City has occupation-specific overtime rules that kick in sooner. New York also requires a higher minimum wage for nonstandard shifts, which Texas does not.

Even within the Davis-Bacon framework, the federal government could send much stronger signals to states and localities about using BLS data to establish prevailing wages. Similarly, simplified rules about benefits and overtime could be promulgated either in statute or through DOT. States and localities are free to require higher than prevailing wages on their projects, but at least it should be clear that this is the choice of local officials, not an adherence to federal law.

13 https://comptroller.nyc.gov/wp-content/uploads/documents/ConstructionWorkerSchedule-2020-2021.pdf

5.c.5. *Strengthen local procurement offices*

Procedures followed by local procurement agencies represent the largest direct contributor to total project cost, and the one that is probably most difficult to control. BVP (2018) emphasize choices about procurement made by the MTA that added significantly to the cost of key projects. Two examples illustrate this. First, the MTA voluntarily follows the procedures in New York's Wicks Law, which means that "systems for electrical, HVAC, and communications are individual bids separate from civil construction tasks such as tunneling and station construction." Instead of contracting with a single entity that does all these highly connected tasks, or that bids the total project and subcontracts them, MTA entertains separate bids for each. This process proliferates contract delays and is likely to increase costs. Second, the MTA must accept the lowest bid, even if that bidder seems unlikely to be able to complete the job. Bosio et al. (2020) point out that such rules, put in place to reduce bribery in a more corrupt era, still hamper procurement agencies today.

In many places, state and local government procurement rules are likely to raise the cost of infrastructure projects. Governments should quantify these costs, and assess whether the benefits of these rules justify their costs. Rigorous application of cost-benefit analysis would highlight the cost of these rules, because the higher prices for infrastructure projects associated with them might mean that the project is not funded.[14]

6. Beyond building: making better use of infrastructure, new and old

Most of the standard complaints about infrastructure in the United States refer to poor maintenance and congestion, rather a lack of roads or bridges. The CBO (1988) reported a 75% rate of return for urban road maintenance, and a 16% rate for rural road maintenance. Transportation economists generally assign a high value to road maintenance. Failure to maintain infrastructure can raise the cost of using that infrastructure, for example by imposing wear-and-tear on vehicles using roads with potholes. It can also raise the risk of more catastrophic losses. This section discusses using infrastructure better. It focuses on four issues: raising the priority on maintenance of existing infrastructure rather than new construction, the potential role of user fees in funding maintenance and reducing congestion, the possibilities and shortcomings of public-private partnerships, and non-infrastructure investments that complement infrastructure projects.

14 Makovšek and Bridge (2021) provide an overview of the practices for infrastructure procurement that are used in different nations and discuss their consequences for project costs and outcomes.

6.a. Prioritizing maintenance of existing infrastructure

Spending on highways is currently almost evenly divided between new construction and maintenance of existing roads. The CBO (2018) reports that for highways, operations and maintenance represented 47% of total government spending in 2017. Maintenance accounts for a larger share of other major infrastructure categories: 72% for water utilities and water resources, 66% for mass transit and rail, and 69% for aviation.

Gramlich (1994) suggested that at prevailing spending patterns, the return to maintaining existing roads is likely to exceed that of new construction. While there are counterexamples to any general rule of this form, the direction of this argument is that maintenance should receive greater priority than it currently does. The rule for efficient allocation is familiar: maintenance should be prioritized until the point where the rates of return are the same for maintenance and new construction. Adopting a rate of return threshold for new projects would do this explicitly, and if maintenance were included in the project set, projects that involve new construction would explicitly compete with maintenance in resource allocation. New construction projects would only get approved if their rate of return exceeds the rate of return to maintenance.

Another approach is to require that all funds dispersed by the National Highway Trust Fund (NHTF) be used for road maintenance, as proposed by Kahn and Levinson (2011). They also suggest that new roads would be supported by a National Highway Bank, which would lend but not grant funds to states for new construction. A less radical plan would require that a minimum percent of all NHTF payments be used for maintenance and embed the National Highway Bank in a larger national infrastructure bank that both lends and grants funds. Each of these rule-of-thumb options represents a step toward requiring new projects to meet a rate of return threshold that is calibrated to the return on maintenance spending.

The NHTF-for-maintenance proposal is the most straightforward of these proposals to implement, although it may be the most difficult politically. The fixed-share option is also straightforward. Implementing a rate-of-return threshold is somewhat more challenging because it requires estimating an average rate of return for road maintenance in the state and evaluating rates of return for all new projects. This is a benefit, not a cost, since forcing the public sector to estimate rates of return is an important step on the path toward better infrastructure policy.

While highway maintenance is the largest category of infrastructure maintenance, maintenance for two other types of physical infrastructure, bridges and dams, is important for avoiding potential catastrophic failures. Bridge safety is also funded from the NHTF. Under the National Bridge Inspection Program, state Departments of Transportation are required to inspect bridges longer than 20 feet at least every two years, and to report data to the DOT. The spending rule that prioritizes maintenance could be modified to require repairing structurally deficient bridges before spending on any other maintenance projects, perhaps with an opt-out mechanism allowing a state to petition DOT for a waiver if a structurally deficient bridge is not unsafe in any way.

Dam monitoring is currently more haphazard than bridge safety monitoring. Three separate federal agencies are involved, including the Federal Emergency Management Agency (FEMA), which provides grants and training, the Federal Energy Regulatory Commission (FERC), which inspects hydroelectric dams, and the U.S. Army Corps of Engineers, which maintains the National Inventory of Dams. An alternative to the status quo would be to create a single dam inspection agency, charged with regularly monitoring all significant dams in the United States. If a dam is deemed to enter into a danger zone, then its owners must remedy the issue within a fixed amount of time. The FERC inspection process already follows this structure.

6.b. Expanding the role of user charges and congestion fees

The Highway Trust Fund charges road users by levying a gasoline tax, and then it deploys those user fees to fund roads and road maintenance. Requiring users to pay for their infrastructure limits overuse and generates revenues. One can argue that the United States should build more infrastructure, and better maintain the infrastructure that it has, without believing that the federal government should pay for any of it. Levying a user charge on roads would not only help to fund these roads, but it would also offset the subsidy to carbon intensive driving that comes from federally funded roads.

The first case for user fee financing is that the size of the fee can be tied to the depreciation costs associated with infrastructure use. A U.S. Government Accountability Office study (1979) found that one five-axle tractor-trailer did as much road damage as 9,600 cars. Pais, Amorim, and Minhoto (2003) corroborate the estimates of the damage associated with heavy vehicles. A basic principle of public economics is that efficient outcomes occur when individuals pay for the social costs of their actions. Driving, and especially driving trucks, causes road damage; efficiency requires that drivers pay for those costs. The absence of such user charges implicitly encourages heavy trucks that create disproportional damage on roads. Winston (2013) reports that the absence of payment-damage charges for heavy trucks imposes an annual welfare loss of $15 billion ($2021).

Congestion presents a second argument for charging drivers more to use the roads. Starting with Singapore in 1975, congestion pricing has been adopted in a number of major cities, including London and Stockholm. At its best, congestion pricing can change by street and time of day and provide real-time incentives to reduce driving and make streets more fluid. At this point, congestion pricing can be implemented with sophisticated GPS monitoring systems that impose no time costs on drivers. Winston (2013) estimates that the absence of congestion pricing imposes a welfare cost of $62 billion per year ($2021) on households. This value is likely to be an underestimate of the total cost of congestion, since it excludes the cost of delays for shippers.

Politically, there is rarely support for raising gas taxes or imposing congestion fees. The debate over the American Jobs Plan illustrates this, since even though the federal gasoline excise tax has not been increased since 1993, there has been strong resistance to raising it. Part, but not all, of the objection stems from concerns about the distributional burden of such a tax increase. Voters appear to be more accepting of tolls on new roads than of new charges on existing roads that used to be free. One implication of that is that new infrastructure should be tolled immediately to eliminate a precedent for free use. One way to adopt congestion charges would be to apply them initially to autonomous vehicles; over time the charge could be extended to vehicles with drivers.

Opponents of user fees argue that they are regressive, but higher income households use many forms of infrastructure more than the poor and would consequently pay more of the user fees associated with them. A 2017 survey of air travel found that the average person living in a household with income of less than $25,000 made one airline trip per year, compared with 5.4 trips for the average person in a household with income of $150,000 or greater (Heimlich and Jackson 2018). A survey commissioned by the Connecticut Department of Transportation (2018) found that nearly 50% of the rail trips in the state were taken by individuals from households with income of more than $150,000, while less than 5% of rail travelers had household income of less than $50,000. In contrast, the Connecticut survey found just the opposite for local bus trips: Nearly 50% of riders had household income of less than $25,000. This suggests that there is a stronger case, on distributional grounds, for subsidizing bus service than train or airport use.

The distributional impact of highway charges and of taxes on gasoline and diesel fuel has been an important concern when higher taxes are proposed (Kile 2021). Gasoline is a significant budget item for many low-income households, and gasoline purchases as a share of household annual income are higher for low-income than higher-income households (Chernick and Reschovsky 1997), although Poterba (1991) points out that the ratio of all consumption to income is much higher at low than at high incomes, and that the share of consumption spending devoted to gasoline is lower at low incomes than middle incomes. Some estimates, such as Graham and Glaister (2004) suggest that the long-run elasticity of fuel demand with respect to income is over one, which means that fuel purchases relative to income rise with income, but other research, such as Small and Van Dender (2007), suggests that the vehicle miles driven rise less rapidly than income. There are also issues of geographic distribution when taxing gasoline, since average gasoline purchases vary substantially across location; they are higher in rural than urban areas. A key question is whether the distributional impact of a higher gasoline tax, or of a related user fee such as a vehicle miles traveled tax, could be offset by other policies, such as a targeted income tax provision or a SNAP-like program to reduce the cost of fuel purchases for low-income households.

For some types of infrastructure, the marginal cost of use is lower than the average cost of provision. Achieving economic efficiency in such cases requires charging the marginal cost of use and making up the difference with revenue from another source. Hong Kong Mass Transit Railway's (MTR) value capture program provides an excellent example of such cross-subsidization. MTR is both a transit company and a real estate company that erects tall buildings above or near their subway stops. The returns from the real estate development help to cover the revenue shortfall associated with the low fares.

Historically, a commitment to user fee financing has been helpful in securing low interest rate loans for entities like Robert Moses' Triborough Bridge Authority. Low interest rates today make debt-financed infrastructure investment more appealing because of the low cost of capital. If the debt was taken on by a special authority with the right to charge tolls, then the infrastructure could be financed with little impact on current budgets. That structure essentially replicates, inside governing, the financial model of public-private partnerships.

6.c. What role for public-private partnerships?

Discussions of new infrastructure programs often include the possibility of public-private partnerships as a means of providing financial support beyond that available from the public sector, or as a way of managing the projects to address issues such as service quality and maintenance. public-private partnerships can be a way of solving some problems that may confront the public sector in the construction or operation of infrastructure. They are not, however, well suited to all infrastructure projects, and they should only be used when it is clear what problem they are designed to solve, and how they will solve it. In some cases, when the answers to these questions are not clear and public-private partnerships are adopted to relax fiscal constraints on the public sector, they can actually raise the long-run cost of infrastructure provision.

Engel, Fischer, and Galetovic (2014) argue that the better alignment of operational incentives in private rather than public projects is a benefit, and possibly the most important benefit, of public-private partnerships. Transferring responsibility for maintenance out of the public sector and making the owner of the project dependent on user fee revenue can improve incentives for operations. When a private owner's return is a function of the future stream of user fees, there is an incentive for the infrastructure operator to keep quality high so that the user base is large and the revenue stream stays high. In contrast, the incentives to preserve the user base may be weaker for public sector entities that can rely on general tax dollars.

A recent illustration of the construction differences between public and private owners comes from India. Singh (2018) compares the roughness of public and private roads in India, exploiting the fact that on some highways, the road will alternative between publicly owned and privately owned segments. He measures road roughness using vertical acceleration measures and finds striking differences in road roughness by ownership structure: public roads are rougher. He argues that this difference reflects the fact that private providers anticipate having to pay for their own maintenance, which means that they ensure that initial road quality is high. Public roads are built by private contractors who have no stake in road maintenance or in road usage, and consequently they build roads that are not durable.

Private highways have a long history in the United States. The Philadelphia and Lancaster Turnpike, which opened in 1795, was the first long-distance, gravel road in America. It was also privately owned. Long-distance canals required so much investment that they were typically public in the United States. The Erie Canal, financed by the State of New York, was a tremendous success. The Potomac Company, a private enterprise led by George Washington before he became president, was not. Railroad companies were initially private, although they did receive subsidies, typically in the form of land grants. The current view that transportation infrastructure is naturally public reflects the particular experience of the 20th century, during which passenger rail moved from private to public hands, and highways were built by governments.

Toll roads are plausible candidates for privatization in many settings, although as Engel, Fischer, and Galetovic (2014) point out, the global track record with private roads includes many very poor outcomes. A report by the U.S. Department of Transportation – Federal Highway Administration (2016) provides a comprehensive review of the U.S. experience with public-private highway partnerships, and points out that, for a number of recent projects, early-year revenues have exceeded projections. The I-91 Express Lanes Project in California is an example of a successful private highway in the United States. The lanes that comprise this project run within the median of the Riverside Freeway which courses through Riverside and Orange Counties. The highway costs change with the time of day, and the peak price to travel the full 18 miles is over $20. Although this may well be the appropriate congestion charge and obviously plenty of customers are willing to pay that price, it is hard to imagine a government entity having the courage to charge drivers over one dollar per mile for access to a road.

Private roads are far less common in the United States than in the European Union; so are private airports. The Federal Aviation Administration began a pilot privatization program in 1997. While 12 airports applied, only one—Hendry Airport in the Everglades—is currently approved, and only one airport ever operated in the program —Stewart Airport in Newburgh, New York from 2000 to 2007.[15] The Branson Airport in Missouri was also built and operated by a private company with public support. In contrast, a large number of European and Asian airports, including Heathrow and Rome Airport, are operated by private companies, which are often partially owned by the government. Oum, Adler, and Yu (2009) look at privatization globally and find that "airports with government majority ownership and those owned by multi-levels of government are significantly less efficient than airports with a private majority ownership," corroborating the casual experience that many travelers have in public U.S. and private European airports. They also note that "airports with a private majority ownership derive a much higher proportion (56%) of their total revenue from non-aviation services." The fact that E.U. and U.K. airports can often feel like shopping malls illustrates the nature of those non-aviation services.

Even if privatized airports were no better at airport operations than their public sector counterparts, they do seem to be more entrepreneurial in the complementary task of selling goods and services to flyers. The revenue associated with such enterprises can help to fill the gap between the average cost of infrastructure use and the charges levied on users. The example of Hong Kong's MTR, which as noted above builds real estate that is connected with its rail service, suggests that this is a more general point. Private companies have the ability to more readily branch into different and relative businesses. Public entities that are focused on infrastructure just focus on infrastructure, perhaps because of the restrictions that the public sector places on the scope of departments and agencies.

---

15 https://www.faa.gov/airports/airport_compliance/privatization/

In some cases, privatization is primarily a financial transaction, designed to raise near-term revenue for a state or local government. The case of Chicago's sale of future revenues from parking meters to a private entity, which generated near-term revenue but reduced the city's long-term income, illustrates the challenges of privatization. Private firms will pay public entities up front for a stream of future revenues. In some cases, a lack of experience on the part of the public sector enables the private purchasers to underpay. In reviewing the sale of future parking meter revenue to a private firm, the Chicago Inspector General found that "the City was paid, conservatively, $974 million less for this 75-year lease than the City would have received from 75 years of parking-meter revenue had it retained the parking-meter system under the same terms that the City agreed to in the lease."[16]

The Inspector General's evaluation assumed a discount factor for the city of 5 to 5.5%, which is far lower than the private discount factor used to evaluate the stream of earnings (Hoffman 2009). Is this reasonable? If long-lived governments are more patient than private investors, there is no long-run benefit to the public sector from transferring a flow of future revenues to a private entity unless there is a major gain in efficiency. Similarly, large public entities are likely to be better able to bear risks than most private firms, which suggests that private entities may apply a greater risk premium in evaluating future revenue streams. This should reduce the value that a private sector bidder, relative to a public sector entity, would place on the revenue stream.

The financial case for public-private partnerships depends on their being able to borrow at better rates than a city government or to bear risk better. Yet even if governments, like the state of Illinois, are themselves seen as a credit risk, they can still set up special purpose, independent entities that will receive the dedicated flow of funds from the infrastructure project. Under Robert Moses, New York's Triborough Bridge Authority was regularly seen as a better risk by bond markets than New York City itself. The use of public-private partnerships to front-load revenue is often a sign of a failure in public sector decision-making.

Public-private partnerships make most sense when they can reduce costs, bring specialized expertise, or improve quality. We discussed previously the many constraints that bind the Metropolitan Transit Authority and increase the price of construction. While it might be better to reduce the constraints, a public-private partnership conceivably offers the possibility of bypassing them altogether. In some states, private providers cannot avoid the rules that bind public behavior. In Massachusetts, for example, the Pacheco laws require that privatization must not

16 https://igchicago.org/wp-content/uploads/2011/03/Parking-Meter-Report.pdf

only save money given current public practices but also that privatization would save money even if state employees work in the "most cost-efficient manner" and the private provider pays workers no less than their public sector equivalents.

One factor that may distinguish public and private entities is their capacity to levy fees and increase prices for infrastructure access. Once a private firm is charged with managing a project, it may face less political heat from proposed increases in user fees; this may make it easier for a privately managed infrastructure project to achieve efficient utilization.

6.d. Complementary investments and activities

There is a fundamental complementarity between transportation infrastructure and real estate development. Building infrastructure that increases access to workplaces and entertainment venues will cause the demand for space nearby to increase, generating a windfall to current property owners and an opportunity to deliver even more value by building at greater density levels. This fact lies behind the strategy pursued by Hong Kong's MTR, which offers an ingenious way of capturing rents created by infrastructure projects, but it is hard to imagine any large public transportation–related entity in the United States acting like a commercial developer.[17]

This motivates the movement toward tax increment financing and land value capture. The basic idea of both is to funnel some part of the increase in land values created by infrastructure projects back to pay for these projects. While attractive in principle, actually determining the impact of new roads on land values can be quite difficult in practice and would likely be subject to political gaming.

Tax increment financing is just about paying for infrastructure. Up-zoning areas near new infrastructure is a means of delivering greater overall social benefits from transportation infrastructure. If a new train station is built in an area, but density cannot be added to that area, then the benefits of the train will be minimal. A reasonable requirement is that areas that benefit from new infrastructure investment make it much easier to build homes nearby.

There are other services beyond real estate that can complement transportation, as the case of retail in commercial airports illustrates. For example, busses and trains

17 BVP (2018) point out that in the Second Avenue Subway project, "the MTA chose to forego development on six corner lots, building only vents and entrances instead of a larger building that could have combined residential and commercial uses with transit." Only three of those lots were good prospects for development, but estimates place their market value around $125 million in 2015. Unlike Hong Kong's MTR, the MTA sees itself as a transit authority, not as a real estate developer, and the MTA's staff are not particularly trained in real estate development.

have long featured advertisements within their vehicles. Trains also sell food. Such services can provide revenues and also improve the transportation experience for users. These are activities at which private entities are likely to have comparative advantage relative to public transit authorities.

In contrast, rural broadband is an area where the public sector is much better poised than the private sector to provide complementary services, such as education. Online education during the COVID-19 pandemic required broadband, a good student computer, and students and teachers who were comfortable with remote learning. Preparing students for continued remote learning in addition to traditional in-person learning will require public investments, especially in training, to ensure that broadband is fully utilized.

In principle, it is possible to have a private company providing a massively subsidized product, whether it is water for poor urbanites in 1880 or broadband for poor residents of rural states today, but private lobbyists are good at making the case for ever larger subsidies. If higher costs enable the company to make the case for larger subsidies, then there are few incentives for limiting expense. Voucher programs are a natural tool for subsidizing consumption, but they are a bad fit for situations where there is a single monopoly provider of either broadband or water.

7. The political challenges facing infrastructure investment

Building infrastructure is largely an engineering problem, but many of the factors that contribute to the poor performance of infrastructure in the United States do not involve engineering. New projects are built and maintenance is neglected. Small groups of empowered citizens can delay or block valuable projects. Public sector unions impose work rules that raise costs and delay schedules. These factors all stem from politics. Politicians like new projects because they get noticed. The media will applaud a new bridge or highway, and while an occasional media story may highlight potholes, repaving a road is unlikely to be a newsworthy event.

The approval process for infrastructure projects involves critical local review and inputs. Neighborhood activists exert sway because new infrastructure really does create tangible and significant costs to them. Consequently, they bother to fight, and they attract attention from politicians and administrators. At the same time, the dispersed thousands or millions who will benefit a small amount from the project pay far less attention. Glaeser and Ponzetto (2018) point out that the taxpayers who ultimately cover the added cost of abatement are even less attentive. Small interest groups are often more effective than dispersed alliances with weak incentives, as Olson (1965) argued decades ago. Some small citizen groups may extract costly project design modifications that raise the price tag for new infrastructure projects, even though the ultimate social benefits may be modest. At the same time, infrastructure projects that might benefit some less well-organized groups may never be built, because of the lack of experience in making the political case for such projects.

The problem of getting to yes becomes even harder when infrastructure spans multiple jurisdictions. Straightening out passenger rail between Boston and New York is practically unthinkable as long as that rail goes through Connecticut, which has little interest in speeding the journey. Anything that spans multiple states typically needs federal engagement.

There are several federal actions that might address some of these issues, such as reserving some portion of the Highway Trust Fund for road maintenance. That structure essentially fights against the political urge to favor new projects. However, many of these problems cannot ultimately be solved by the federal government alone. It does not have the power to change local politics, pass state laws relating the environment, or change the procurement process for state agencies. State governments do have that power. Consequently, reducing costs requires the federal government to interact with the states and to make reform a precondition for funding.

It is hard to imagine how that can happen for any standard pass-through program, like the trust fund. For the federal government to meaningfully impact state behavior, it will need to bargain with the states on a project-by-project basis. That cannot be done by the national legislature and it cannot really be done by any entity that is directly political. An effective organization would need independent authority to craft deals that would lower costs and increase value. It would need to be well staffed and well-funded. A national infrastructure bank might play such a role, but it is an untried idea and unlikely to remedy all of the difficult issues. We recognize the potential downsides of such a bank, but given the challenges of reducing infrastructure costs and improving infrastructure administration otherwise, the concept deserves serious consideration.

8. Conclusion

Our assessment of the role of economic analysis in infrastructure investment suggests several broad conclusions.

First, it is difficult to place high confidence in widely discussed measures of infrastructure "need." The most reliable way to develop such estimates would be by applying cost-benefit analysis on a project-by-project basis and aggregating the results. But that approach is expensive, given the vast array of potential infrastructure projects, and it is subject to gaming by overstating future benefits and low-balling costs. Estimates of the returns to maintaining existing infrastructure are often higher than estimates of the returns to undertaking new projects, which suggests the importance of guarding against "ribbon-cutting bias" toward new initiatives on the part of both elected leaders and the heads of government agencies. Any major infrastructure initiative should emphasize careful ex ante analysis of project costs and benefits, with oversight where feasible of padding by advocates of the assumptions regarding future costs and benefits.

Second, infrastructure projects in the United States are expensive relative to those in other nations. The precise reasons are difficult to identify, but they include project designs that incorporate many features that remediate adverse project effects, such as highway noise and the inconvenience of disruption while building, required wages for workers that may exceed area norms, project delay through regulatory processes, and weak procurement and project management by the relevant government agencies.

Third, user fees warrant greater consideration as a source of infrastructure project financing. Such fees, along with congestion charges, can improve the efficiency of infrastructure use. While there are concerns about the distributional effects of user fees and burdens on low-income groups in particular, the pattern of infrastructure use across income groups suggest that some user fees are progressive—higher income households use airports, for example, more than their lower-income counterparts. Public transit, particularly buses, is a notable exception. Rather than carry out income redistribution by exempting infrastructure use from charges, policymakers could consider targeted redistribution programs, such as transit vouchers for low-income households or infrastructure-use rebates mediated through the tax system. Some states currently provide income tax relief for renters or for commuters who can document their travel costs.

Finally, public-private partnerships can provide a means to increase operational efficiency, but arguments that they allow project sponsors to access low-cost capital should be viewed with caution. In some cases, the cost of capital for private entities may exceed that for public sector borrowers and relying on private finance rather than public funding may ultimately increase the cost of the project. These partnerships may nevertheless offer operational efficiencies or ways to circumvent political constraints that bind on public entities but not private sector actors. Some state and local governments may be attracted to these partnerships because they relieve current cash flow constraints, but they may come at a price in terms of the long-term cost of infrastructure services.

References

Allen, Treb and Costas Arkolakis. 2019. “The Welfare Effects of Transportation Infrastructure Improvements.” National Bureau of Economic Research Working Paper #25487, original version posted January 22, 2019.

Altshuler, Alan and David Luberoff. 2003. *Mega-Projects: The Changing Politics of Urban Public Investment*. Washington, DC: Brookings Institution.

American Society of Civil Engineers. 2021. *Report Card for America’s Infrastructure*. Reston, Virginia: ASCE. https://infrastructurereportcard.org/ (accessed June 10 2021).

Aschauer, David. 1989. “Is public expenditure productive?” *Journal of Monetary Economics* 23: 177–200.

Association of State Dam Safety Officials. 2017. *The Cost of Rehabilitating Our Nation’s Dams: A Methodology, Estimate, and Proposed Funding Mechanisms*. Lexington, KY: ASDSO.

Barone, Richard, Julia Vitullo-Martin and Alyssa Pichardo. 2018. *Building Rail Projects Better for Less: A Report on the Costs of Delivering MTA Megaprojects*. Regional Plan Association Report. https://s3.us-east-1.amazonaws.com/rpa-org/pdfs/RPA-Building-Rail-Transit-Projects-Better-for-Less.pdf (accessed on June 9, 2021).

Barro, Josh. 2019. “Why New York Can’t Have Nice Things,” *New York*. May 30, 2019. https://nymag.com/intelligencer/2019/05/new-york-infrastructure-costs.html

Baum-Snow, Nathaniel, Matthew E. Kahn, and Richard Voith. 2005. "Effects of urban rail transit expansions: Evidence from sixteen cities, 1970-2000." *Brookings-Wharton papers on urban affairs* (2005): 147–206.

Bearfield, Domonic A., and Melvin J. Dubnick. 2009. "All Mega-projects are Local? Citizen Participation Lessons from the Big Dig." *Journal of Public Budgeting, Accounting & Financial Management* 21 (3): 393–426.

Bhasin, A. 2007. “The Economic Benefits of Crossrail.” Colin Buchanan and Volterra Consulting. https://volterra.co.uk/wp-content/uploads/2013/02/Economic-Benefits-of-Crossrail.pdf

Boik, Andre. 2017. "The economics of universal service: An analysis of entry subsidies for high speed broadband." *Information Economics and Policy* 40: 13–20.

Bosio, Erica, Simeon Djankov, Edward Glaeser, and Andrei Shleifer. 2020. “Public Procurement in Law and Practice.” National Bureau of Economic Research Working Paper #27188.

Bom, Pedro R.D. and Jenny Ligthart. 2014. ”What Have we Learned from Three Decades of Research on the Productivity of Public Capital?” *Journal of Economic Surveys* 28: 889–916.

Brinkman, Jeffrey and Jeffrey Lin. 2019. “Freeway Revolts!” Federal Reserve Bank of Philadelphia Working Paper #19-29.

Brooks, Leah and Zachary Liscow. 2019. *Infrastructure Costs*. Washington: Brookings Institution, Hutchins Center Working Paper #54. https://www.brookings.edu/wp-content/uploads/2019/07/2019-07-12_infrastructure_costs_v2.pdf

———. 2021. “Can America Reduce Highway Spending? Evidence from the States.” In Edward Glaeser and James Poterba, *Economics of Infrastructure Investment* (Chicago: University of Chicago Press).

Carse, Ashley. 2016. "Keyword: Infrastructure: How a humble French engineering term shaped the modern world." In *Infrastructures and Social Complexity*., edited by Penelope Harvey, Casper Jensen, and Atsuro Morita, 45–57. London: Routledge.

Chatterji, Aaron K., Kenneth Y. Chay, and Robert W. Fairlie. 2014. “The Impact of City Contracting Set-Asides on Black Self-Employment and Employment.” *Journal of Labor Economics* 32: 507–561.

Chernick, Howard and Andrew Reschovsky. 1997. “Who Pays the Gasoline Tax?” *National Tax Journal* 50 (2): 233–259.

Congressional Budget Office. 1988. *New Directions for the Nation’s Public Works*. Washington: CBO.

———. 2016. *Approaches to Making Federal Highway Spending More Productive*. Washington: CBO.

———. 2018. *Federal Support for Financing State and Local Transportation and Water Infrastructure*. Washington: CBO.

Congressional Research Service. 2019. *The TIGER/BUILD Program at 10 Years: An Overview*. Washington: CRS.

Connecticut Department of Transportation. 2018. *Service and Fare Equity Analysis: Disparate Impact and Disproportionate Burden Analysis of Proposed Fare and Service Changes to Bus and Rail Services in Connecticut*. Hartford: Connecticut DOT. https://portal.ct.gov/-/media/DOT/CTDOTServiceandFareEquityAnalysisApril2018pdf.pdf?la=en

Duranton, Gilles, Geetika Nagpal, and Matthew A. Turner. 2021. “Transportation Infrastructure in the US.” In *Economics of Infrastructure Investment,* edited by Edward Glaeser and James Poterba. Chicago: University of Chicago Press.

Duranton, Giles, and Matthew A. Turner. 2011. “The Fundamental Law of Road Congestion: Evidence from the US.” *American Economic Review* 101: 2616–2652.

DOT. 2019. “Capital Investment Scenarios,” In *Status of the Nation’s Highways, Bridges, and Transit*. Federal Highway Administration. https://www.fhwa.dot.gov/policy/23cpr/chap7.cfm

Engel, Eduardo, Robert Fischer, and Alexander Galetovic. 2014. *The Economics of Public-Private Partnerships*. Cambridge: Cambridge University Press.

Federal Communications Commission. 2021. *Fourteenth Broadband Deployment Report*. Accessed June 22 2021 at https://docs.fcc.gov/public/attachments/FCC-21-18A1.pdf

Flyvbjerg, Bent, Nils Bruzelius, and Bert van Wee. 2008. “Comparison of Capital Costs per Route-Kilometre in Urban Rail.” *European Journal of Transport and Infrastructure Research* 8 (1): 17–30.

Garin, Andrew. 2019. "Putting America to work, where? Evidence on the effectiveness of infrastructure construction as a locally targeted employment policy." *Journal of Urban Economics* 111: 108–131.

Glaeser, Edward, and Joshua Gottlieb. 2009. “The Wealth of Cities: Agglomeration Economies and Spatial Equilibrium in the United States.” *Journal of Economic Literature* 47: 983–1028.

Glaeser Edward and Giacomo Ponzetto. 2018. "The political economy of transportation investment." *Economics of Transportation* 13: 4–26.

Glaeser, Edward and James Poterba. 2021. "Economic Analysis and Infrastructure Investment," in *Economic Analysis and Infrastructure Investment*, edited by Edward Glaeser and James Poterba. Chicago: University of Chicago Press.

Graham, Daniel J., and Stephen Glaister. 2004. "Road traffic demand elasticity estimates: a review." *Transport Reviews* 24: 261–274.

Gramlich, Edward M. 1994. "Infrastructure investment: A review essay." *Journal of Economic Literature* 32: 1176–1196.

Greenstein, Shane. 2021. "Digital Infrastructure" in *Economic Analysis and Infrastructure Investment,* edited by Edward Glaeser and James Poterba. Chicago: University of Chicago Press.

Hall, Jonathan D. and Joshua Madsen. 2020. "Can Behavioral Interventions be Too Salient? Evidence from Traffic Safety Messages." Working paper, University of Toronto Transportation Research Institute.

Haughwout, Andrew. 2019. "Infrastructure Investment as an Automatic Stabilizer." In *Recession Ready: Fiscal Policies to Stabilize the American Economy,* edited by Heather Boushey, Ryan Nunn, and Jay Schambaugh, 129–152. Washington: Brookings Institution.

Heimlich, John P. and Chris Jackson. 2018. *Air Travelers in America: Findings of a Survey Conducted by IPSOS.* Airlines for America. https://airlines.org/wp-content/uploads/2018/02/A4A-AirTravelSurvey-20Feb2018-FINAL.pdf

Hoffman, David H. 2009. "An analysis of the lease of the city's parking meters." Report of Inspector General's Findings and Recommendations. https://igchicago.org/wp-content/uploads/2011/03/Parking-Meter-Report.pdf

Horrox, James and Matthew Casale. 2019. *Electric Buses in America: Lessons from Cities Pioneering Clean Transportation.* US PIRG Education Fund. Accessed June 6 2021 at https://uspirg.org/sites/pirg/files/reports/ElectricBusesInAmerica/US_Electric_bus_scrn.pdf

Joy, Stewart. 1967. "The "Standby" Concept on Railways." *Journal of Transport Economics and Policy* 1: 243–48.

Kain, John F. 2007. "Deception in Dallas: Strategic Misrepresentation in Rail Transit Promotion and Evaluation." *Journal of the American Planning Association* 56: 184–196.

Kahn, Matthew E., and David Levinson. 2011. *Fix it first, expand it second, reward it third: A new strategy for America's highways.* No. 2011-03. Washington: Brookings Institution.

Kile, Joseph. 2021. *Testimony on Options for Funding and Financing Highway Spending.* U.S. Senate Committee on Finance, May 18. https://www.cbo.gov/publication/57222

KCRA. 2017. "Evacuation orders remain for more than 180K as officials monitor Lake Oroville." KCRA, February 13, 2017. https://www.kcra.com/article/evacuation-orders-issued-for-low-levels-of-oroville/8735215#

Levy, Alon. 2011. "US Rail Cost Comparisons." Pedestrian Observations Blog. https://pedestrianobservations.com/2011/05/16/us-rail-construction-costs/ (retrieved on June 9, 2021).

Li, Shanjun, Matthew E. Kahn, and Jerry Nickelsburg. 2015. "Public transit bus procurement: The role of energy prices, regulation and federal subsidies." *Journal of Urban Economics* 87: 57–71.

Makovšek, Dejan, and Adrian Bridge. 2021. "Procurement Choices and Infrastructure Costs." In Edward Glaeser and James Poterba, eds., *Economics of Infrastructure Investment*. Chicago: University of Chicago Press.

Marion, Justin. 2009. "How costly is affirmative action? Government contracting and California's Proposition 209." *Review of Economics and Statistics* 91: 503–522.

Munnell, Alicia, with Leah Cook. 1990. "How Does Public Infrastructure Affect Regional Economic Performance?" *New England Economic Review* (September-October): 11–32.

Olson, Mancur. 1965. *The Logic of Collective Action*. Cambridge: Harvard University Press.

Oum, Tae H., Nicole Adler, and Chunyan Yu. 2006. "Privatization, corporatization, ownership forms and their effects on the performance of the world's major airports." *Journal of Air Transport Management* 12: 109–121.

Pais, Jorge C., Sara IR Amorim, and Manuel JC Minhoto. 2013. "Impact of traffic overload on road pavement performance." *Journal of Transportation Engineering* 139: 873–879.

Parsons Brinckerhoff. 2014. "2014 California High-Speed Rail Benefit-Cost Analysis." 2014 Business Plan Technical Supporting Document. https://hsr.ca.gov/wp-content/uploads/docs/about/business_plans/BPlan_2014_Sec_7_CaHSR_Benefit_Cost_Analysis.pdf

Penn, Alexandrea. 2018. "The deadliest bridge collapses in the US in the last 50 years." CNN, March 15, 2018. https://www.cnn.com/2018/03/15/us/bridge-collapse-history-trnd

Poterba, James. 1991. "Is the Gasoline Tax Regressive?" In *Tax Policy and the Economy: volume* 5, edited by David Bradford. Cambridge, MIT Press.

Railway Technology. 2015. "Parsons Brinkerhoff-Network Rail consortium wins $700m California high-speed rail programme." Railway Technology, June 14, 2015. https://www.railway-technology.com/news/newsparsons-brinckerhoff-network-rail-consortium-wins-700m-california-high-speed-rail-contract-4600924/

Ramey, Valerie. 2021. "The Macroeconomic Consequences of Infrastructure Investment," National Bureau of Economic Research Working Paper #27625, in *Economics of Infrastructure Investment*, edited by Edward Glaeser and James Poterba. Chicago: University of Chicago Press.

Schanzenbach, Diane Whitmore, Ryan Nunn, and Greg Nantz. 2017. *If You Build It: A Guide to the Economics of Infrastructure Investment*. Washington: Brookings Institution, Hamilton Project.

Shirley, Chad, and Clifford Winston. 2004. "Firm inventory behavior and the returns from highway infrastructure investments." *Journal of Urban Economics* 55: 398–415.

Singh, Ram. 2018. "Public–private partnerships vs. traditional contracts for highways." *Indian Economic Review* 53: 29–63.

Small, Kenneth A., and Kurt Van Dender. 2007. "Fuel efficiency and motor vehicle travel: the declining rebound effect." *The Energy Journal* 28 (1).

Stern, Nicholas and Joseph E. Stiglitz. 2021. “The Social Cost of Carbon, Risk, Distribution, and Market Failures: An Alternative Approach,” National Bureau of Economic Research Working Paper #28472.

Summers, Lawrence H., 1988. "Economic Priorities for the Next President." *Challenge* 31 (5): 21–26.

———. 2017. “The Case for a Proper Programme of Infrastructure Spending.” *Financial Times.* January 16. https://www.ft.com/content/cf700fad-6e87-3b32-867c-8686853c76f1 accessed June 6 2021.

Tabuchi, Hiroko. 2017. “$300 Billion War Beneath the Street: Fighting to Replace America’s Water Pipes.” *New York Times*, November 10, 2017. https://www.nytimes.com/2017/11/10/climate/water-pipes-plastic-lead.html

U.S. General Accounting Office. 1979. *Excessive Truck Weight: An Expensive Burden we can Longer Support.* Washington: GAO. http://archive.gao.gov/f0302/109884.pdf

Vartabendian, Ralph. 2021. “High-speed rail to run on a single track in Central Valley as overall cost rises.” *Los Angeles Times,* February 10, 2021. https://www.latimes.com/california/story/2021-02-10/high-speed-rail-central-valley-single-track-rising-cost

Wellisz, Stanislaw. 1960. "Economic Planning in the Netherlands, France, and Italy." *Journal of Political Economy* 68: 252–83.

Winston, Clifford. 2013. “On the Performance of the US Transportation System: Caution Ahead.” *Journal of Economic Literature* 51: 773–824.

Woetzel, Jonathan, Nicklas Garemo, Jan Mischke, Martin Hjerpe, and Robert Palter. 2016. *Bridging Global Infrastructure Gaps.* June 14. https://www.mckinsey.com/business-functions/operations/our-insights/bridging-global-infrastructure-gaps

Challenges of a Clean Energy Transition and Implications for Energy Infrastructure Policy

AUTHOR

Severin Borenstein* and Ryan Kellogg**

ABSTRACT

The United States faces the challenge of dramatically reducing carbon emissions while simultaneously ensuring the reliable supply of on-demand energy services that its residents have come to expect. Federal policy will be instrumental in driving investments in energy infrastructure that will be required to transition the U.S. energy supply to zero-emission sources. This chapter discusses the major barriers that policy will need to overcome in order to successfully execute this transition at a reasonable cost. A core problem is that wind and solar generation are intermittent. Provision of reliable zero-emission supply therefore requires combining wind and solar resources with investments in dispatchable zero-emission sources (such as nuclear, hydroelectric, geothermal, and fossil-fueled power plants with carbon capture and sequestration), long-distance transmission, demand flexibility, and storage technologies. But given uncertainties about technological progress, it is difficult to know which combination of investments will be most cost-effective. We argue that broad incentives – such as carbon pricing, clean energy standards, or clean energy subsidies – that do not discriminate across zero-emission resources will be essential for directing capital toward cost-effective investments in clean energy infrastructure.

* Haas School of Business, University of California, Berkeley; Energy Institute at Haas; and National Bureau of Economic Research. Email: severinborenstein@berkeley.edu.

** Kellogg: Harris Public Policy, University of Chicago; and National Bureau of Economic Research. Email: kelloggr@uchicago.edu.

Acknowledgements: We thank Sara Johns and Jonathan Rockower for excellent research assistance.

We also argue, however, that such incentives on their own will be insufficient to meet the overall challenge. Policy must also address a suite of additional problems in energy markets that clean energy pricing incentives alone will not address. These problems include motivating global emissions reductions, overcoming regulatory barriers to long-distance transmission construction, addressing deficiencies in wholesale energy markets, reducing utilities' inclusion of non-marginal costs in volumetric retail rates, eliminating inequities in the distribution of clean energy's benefits and costs, and funding infrastructure decommissioning at the end of its useful life.

1. Introduction

The provision of reliable, on-demand energy services to households and businesses is a central feature of the modern U.S. economy. Americans generally take it for granted that their local gasoline stations will have fuel on-hand, that a steady supply of natural gas is available to run their furnaces in the winter, and that electricity will be there to power their lights, refrigeration, or air conditioning.

Reliable, on-demand energy supply has required substantial investments in both physical infrastructure – including wells, mines, refineries, long-distance transmission pipelines and wires, electric generators, and distribution networks – and the human capital to use it effectively. Historically, U.S. energy policy has emphasized investing in and operating this infrastructure at lowest cost, while maintaining reliability. Energy policy in the late 20th century focused on reducing regulation of energy markets in order to lower end-user costs. These market reforms – some of which were federal changes, but many of which were enacted unevenly across the states – eliminated gasoline price regulation, liberalized natural gas extraction, and broke up integrated utility monopolies, thereby permitting independent power producers to enter into electric generation markets. The core idea motivating these reforms was to lower costs by using competitive forces to drive efficiency improvements relative to the well-known inefficiencies of regulated monopolies.

The core challenge for energy policy in the 21st century is how to supply low-cost energy that is not only reliable but also does not harm the environment. This chapter focuses on the problem of global climate change and the need to substantially reduce the emissions of carbon dioxide (CO2) and other greenhouse gases that are associated with the widespread use of fossil fuels as primary energy sources. There is also, however, an increasing imperative to address long-standing local environmental damages and inequities associated with energy production and use. So climate change solutions must be paired with improvements in environmental justice.

The problem of climate change is vexing, and not one that private markets will solve on their own. The core challenge arises from the fact that fossil fuels became dominant energy sources in the first place because they can be reliably supplied at low cost. Oil, coal, and (to a lesser extent) natural gas are energy-dense and storable. Oil and gas can easily be transported in bulk over long distances by pipeline, and oil can be transported across the oceans via tanker at low cost. These resources are therefore amenable to energy infrastructure that allows end-users to consume energy when they want it and where they want it.

Unfortunately, current zero-emission technologies mostly lack fossil fuels' desirable characteristics. Compared to fossil fuels, electricity is extremely costly to store with current technology. Generation from wind and solar are also not *dispatchable*. They produce only when the wind is blowing or the sun is shining, which are not necessarily the times when end-users demand energy. Moreover, even with recent technological improvements in solar and wind generation, in many locations they remain more costly than fossil fuel–fired generation due to weak availability of the solar or wind resource. An alternative approach, fossil fuel generation with carbon capture and storage (CCS), also remains costly given current technology, as do more storable, low-carbon fuels like hydrogen.

These challenges imply that massively reducing CO_2 emissions, while simultaneously maintaining reliable supplies of energy for on-demand consumption, will not be easy. Achieving these twin goals will likely require large, policy-driven investments in new energy infrastructure for generation, transmission, storage, and distribution of zero-emission power. The policy question is then how to catalyze the necessary investments without unduly increasing costs to end-users.

One potential approach to energy infrastructure policy would follow models of public investments in transportation infrastructure, in which agencies identify particularly promising investment opportunities (say, a specific transmission line or renewable power generation facility) and then direct public funds to those investments. Alternatively, policy could take a less targeted approach of providing blanket incentives for zero-emission energy production (either by taxing emissions or by subsidizing clean production), continuing to provide strong incentives for grid reliability, and otherwise getting out of the way to let "the market" figure out the lowest-cost solution.

Direct public investment into specific projects or technologies is fraught; when technology is evolving rapidly, governments picking winners often does not end well. The history of U.S. biofuel policy provides a cautionary tale. In the early 2000s, there was widespread hope for production of liquid fuels from plants, such as switchgrass, or from agricultural waste products. Federal legislators were sufficiently optimistic about these technologies that the Energy Policy Act of 2005 and the Energy Independence and Security Act of 2007 specified minimum blending requirements for advanced biofuels into automotive gasoline. Yet these technologies never came to fruition, and U.S. biofuels today consist almost entirely of corn-based ethanol. Despite questions over corn-based ethanol's CO2 emissions relative to gasoline, the large implicit subsidies that sustain this industry have proven politically difficult to remove.[1]

We view a pricing policy that supports zero-emission energy production, coupled with robust reliability incentives, as a strong foundation for U.S. energy infrastructure policy. Still, we argue that very large carbon taxes or renewable subsidies on their own would be an insufficient response to the global task at hand. First, direct costs are not the only barrier to investments in zero-emission energy infrastructure. Most importantly, construction of long-distance, high-capacity transmission lines – which will be essential to address renewables' non-dispatchability limitation – is beset by multi-jurisdiction regulatory and hold-up problems. Regulatory reform that invests federal authorities with the power to certify interstate transmission lines will be essential to enabling needed transmission investments.

Reforms to wholesale and retail electricity markets also have a role to play in promoting efficient investment and dispatch of zero-emission generation resources. Increased use of organized, transparent wholesale power auctions can increase the dispatch of low-cost generators while sending clearer price signals for resource investment. In retail electricity markets, the standard pricing paradigm used nearly ubiquitously in the U.S. excessively marks up end-users' electricity prices by folding recovery of fixed distribution costs into retail rates. Retail pricing reforms that reduce or eliminate these mark-ups can improve end-users' incentives to switch from natural gas to electricity for space heating, and to switch from gasoline-powered to electric vehicles. Such reforms can also improve the progressivity with which the fixed costs of electricity distribution are shared across consumers.

The transition to a net-zero emission energy system will also lead to stranded fossil fuel assets. These assets will include both physical infrastructure and human capital. Policy will need to ensure the repurposing or proper decommissioning of fossil fuel infrastructure while also helping identify – and in some cases directly provide – opportunities for displaced workers to apply their skills. The impacts of a clean energy transition on the oil, gas, and coal workforce is one of several ways in which energy market outcomes can be inequitable. Policies on stranded assets,

1 See Mullins, Griffin, and Matthews (2011), Hoekman and Broch (2018), and Scully et al. (2021) for discussions of uncertainty in estimates of corn-based ethanol's life cycle emissions and reviews of recent literature.

recovery of fixed distribution costs, infrastructure siting, and addressing legacy infrastructure's impacts on disadvantaged communities all have a role to play in ensuring a just transition that equitably shares benefits and costs.

One area where we will argue for increased direct public funding is research, development, and deployment of early-stage clean energy technologies. The climate challenge is a global challenge, and emissions reductions from the U.S. alone will not allow the planet to avoid the worst effects of climate change. An important pathway for driving global emissions reductions is technological improvement. While U.S. intellectual property policy is typically concerned with preventing free access to IP by other countries, the climate problem is a situation where the optimal policy is likely to be investing in technology and then licensing it for free, at least to low-income countries. Private innovators do not have an incentive to behave in this manner.[2] Thus, there is an important role for public funds to drive the development and export of low-cost, low-carbon technologies.

The remainder of the chapter is divided into three primary sections. Section 2 presents background on the energy industry in the U.S., divided into the oil and gas industry and the electricity industry. Section 3 discusses seven areas of challenges the country faces in transitioning to a low-carbon energy system while maintaining affordability, equity, and reliability. Section 4 then discusses potential solutions and the role that infrastructure investment can play in addressing these challenges. We conclude in Section 5.

2. Background

The energy industry has more than the usual level of institutional and operational complexities that are important for discussing the relevant policy issues. In this section, we briefly outline those factors in order to establish a baseline understanding of the industry before examining the challenges it presents and possible solutions.

2.a. Oil and gas extraction, transportation, and distribution

We begin by discussing the current state of markets and infrastructure for the dominant source of energy in the U.S.: fossil fuels. Our discussion focuses on oil and natural gas, highlighting their energy density, storability, and transportability. These private advantages, combined with the unpriced pollution externalities of these fuels, will pose significant challenges for climate policy.

2 Jones (2021) discusses at greater length why science and innovation are public goods that merit substantial publicly funded investment, not just in the energy sector but in a variety of settings.

Our discussion, for the most part, omits coal. While coal remains a significant share of U.S. primary energy consumption, its importance has steadily declined over the past decade, in large part due to the shale boom that dramatically increased U.S. natural gas supplies. Even in the absence of aggressive climate policy, the trend of disinvestment from coal is likely to continue unabated.[3] We partition our discussion of the oil and gas industry into its three main activities: extraction, long-distance transportation, and local distribution.

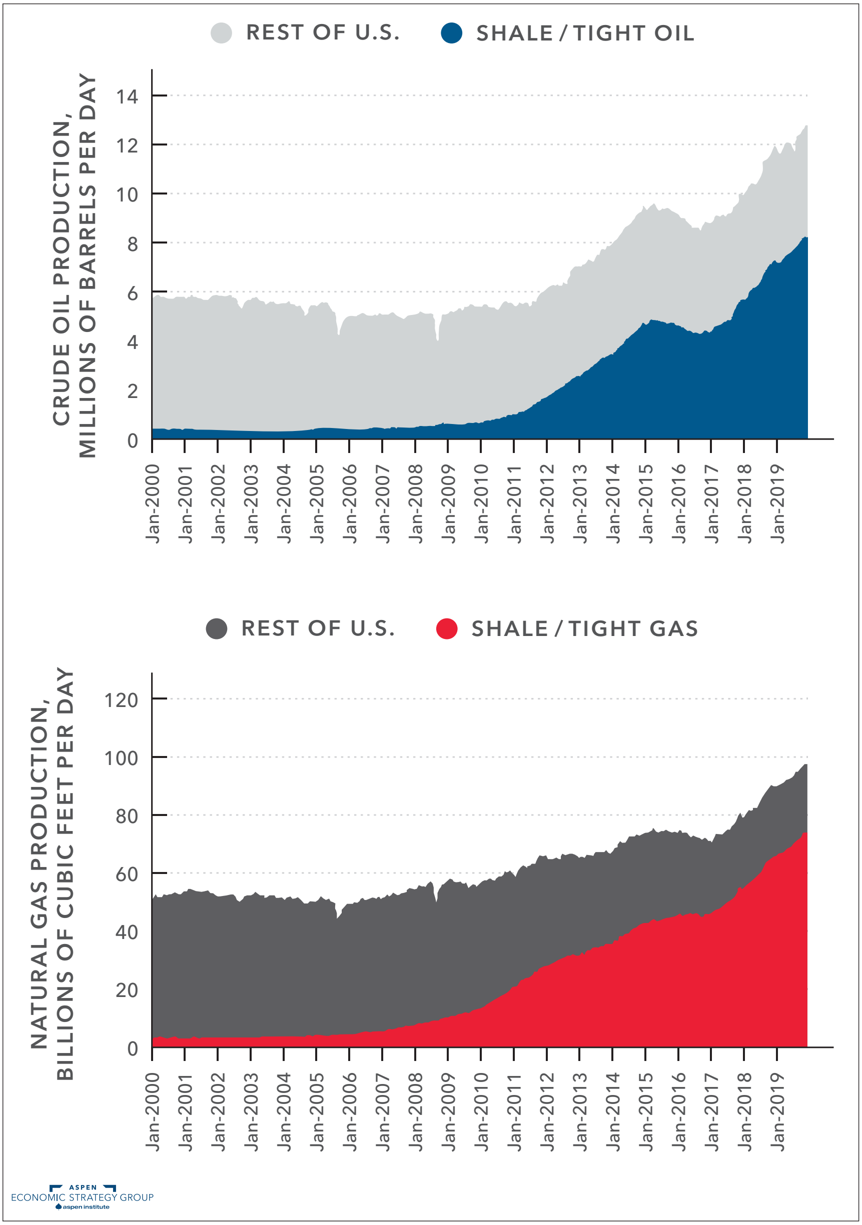
2.a.1. Upstream extraction

Oil and gas extraction has not faced substantive economic regulation since the 1980s, when the last federal wellhead price control regulations were lifted. Despite the presence of the international OPEC cartel on the world oil market, oil and gas extraction in the U.S. can be described as competitive. There are dozens of large integrated producers and large independents, as well as thousands of small independent producers.

The most important decisions that U.S. oil and gas producers regularly make are when and where to invest in drilling new wells. These decisions weigh wells' expected production and local output prices against their drilling and completion cost. Once drilled, a well's production follows a decline curve. Except in extreme situations, producers do not have an incentive to change wells' production rate in response to output prices (Anderson, Kellogg, and Salant 2018).

3 Globally, however, coal remains a substantial threat to progress on reducing greenhouse gas (GHG) emissions, as highlighted in the International Energy Agency's Global Energy Review 2021. https://www.iea.org/reports/global-energy-review-2021.

Figure 1: U.S. crude oil and dry natural gas production, 2000–2019



Source: Energy Information Administration (EIA) data on shale oil production, available at https://www.eia.gov/energyexplained/oil-and-petroleum-products/data/US-tight-oil-production.xlsx; data on total oil production, available at https://www.eia.gov/dnav/pet/hist/LeafHandler.ashx?n=PET&s=MCRFPUS2&f=M; data on shale dry gas production available at https://www.eia.gov/naturalgas/weekly/img/shale_gas_202104.xlsx; data on total dry gas production, available at https://www.eia.gov/dnav/ng/hist/n9070us2m.htm.

Beginning in the mid-2000s, U.S. production of oil and gas increased substantially due to the exploitation of shale oil and gas reserves. Profitable extraction from these deposits had previously been considered impossible, but combining horizontal drilling with high volume hydraulic fracturing (injecting large volumes of water, sand, and chemicals to fracture the shale formation itself) was a technology breakthrough that transformed the U.S. oil and gas sector in just a few years. As shown in Figure 1, shale oil has led U.S. oil production to more than double since 2010, and shale gas has led U.S. natural gas production to nearly double.

2.a.2. Long-distance oil and gas transmission

Oil and gas deposits are often located far from demand centers and therefore require long-distance, overland transportation. Pipelines have long been the dominant (and for natural gas, the only) technology for doing so, though significant volumes of crude oil also travel via railroad.

Because long-distance transmission pipelines are characterized by economies of scale, service between any two distant locations is often provided by just one pipeline firm, or at most a handful of firms. To prevent abuse of market power, the transmission rates that pipelines may charge are therefore regulated by the Federal Energy Regulatory Commission (FERC).

Per the Natural Gas Act of 1938, FERC not only regulates service rates but also authorizes interstate pipeline investments by issuing certificates of public convenience and necessity. These certificates can allow pipeline firms to use eminent domain to acquire rights-of-way, if necessary. To obtain a certificate, pipeline companies are typically required to demonstrate demand by signing firm shipping agreements with prospective customers.

Despite FERC's authority to regulate oil pipelines' service rates, it does not have certification authority over new oil pipeline construction. Oil pipelines must instead be approved individually by each state that the line passes through. Finally, both oil and gas pipelines must comply with environmental regulations and procedures, and in particular with those prescribed by the National Environmental Policy Act.

2.a.3. Refining, local distribution, and storage

Crude oil is converted to end-use products like gasoline and diesel fuel at refineries that are dispersed broadly throughout the U.S. Due to the ease with which oil and refined products can be transported and stored, their distribution and storage are decentralized. Gasoline and diesel are typically distributed to retail fueling stations via truck, and storage is so cheap that individual consumers store days of fuel on board their own vehicles. These distribution markets generally operate with little or no sector-specific economic regulation.

Natural gas is inherently less energy-dense than refined oil products, and it is therefore more costly to distribute and store on an energy-unit basis. Pipelines are the only cost-effective method for moving natural gas, so its distribution requires a pipeline connection to each customer. The large up-front capital expense of building a gas distribution network, along with ongoing maintenance costs, make gas distribution a classic example of a natural monopoly. Throughout the U.S., local natural gas distribution companies are therefore monopolies licensed by the state or local government. They are usually owned by private investors but sometimes owned and operated by local governments, with rates regulated under cost-of-service principles.

It is more economical to store natural gas at scale, and substantial storage is held by local distribution companies, large industrial users, and third-party storage firms. Nearly all of this storage makes use of underground geologic formations rather than above-ground vessels. This gas storage plays an essential role in managing both seasonal and short-term fluctuations in natural gas demand. At the start of the winter heating season, the quantity of natural gas in storage is typically sufficient to satisfy more than one month of average winter gas consumption.[4]

2.b. Electricity markets

For most of the last century, electricity has been the primary energy source for lighting and appliance services, including air conditioning. But it has played a smaller role in space and water heating, industrial energy, and transportation. Increasingly, however, electricity is being viewed as the most probable path to decarbonizing not just the activities that have historically used it, but also many of the activities for which energy has been supplied by direct combustion of coal, natural gas, and refined oil products. For electrification to be a credible pathway to decarbonization, however, electricity generation must become almost entirely carbon free, while still remaining extremely reliable and cost competitive.

2.b.1. Structure of the electricity industry

The electricity industry comprises four major functions: generation, transmission, distribution, and procurement/retailing.[5] In a few parts of the U.S., all four functions are provided by a single entity (as was the case in most areas thirty years ago), but in most of the country generation is now competitively supplied. In 2020, about 40% of U.S. electricity came from burning natural gas, 19% from coal, 20% from

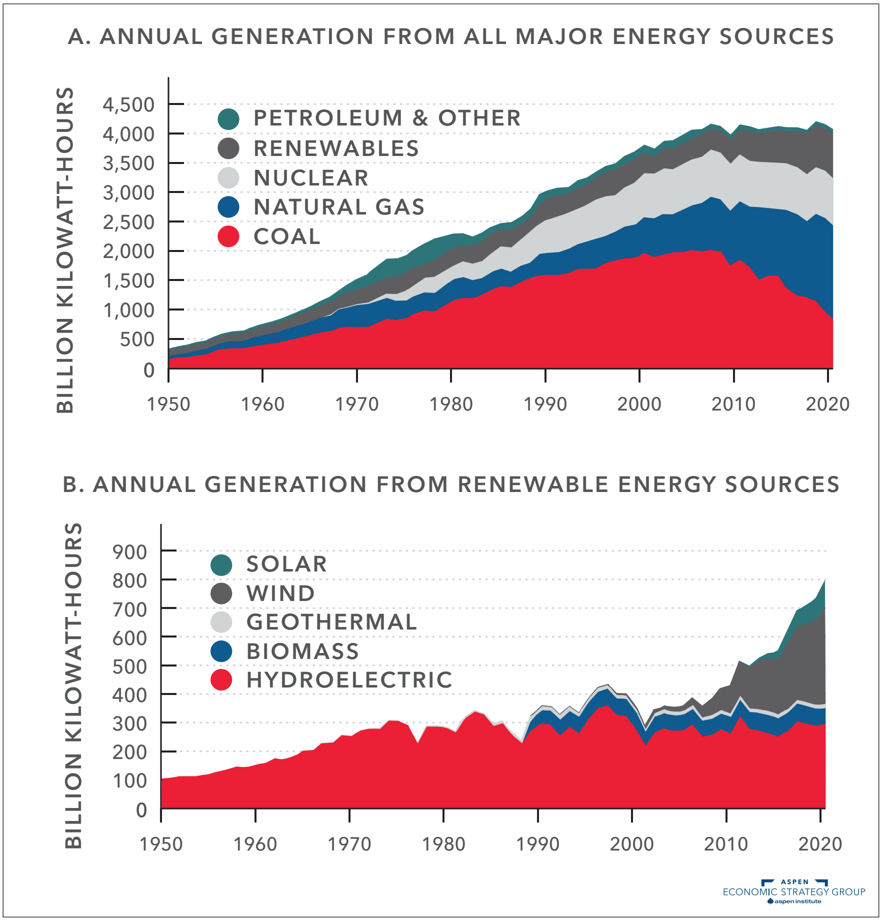
4 U.S. storage volumes typically peak at around 4 trillion cubic feet just before winter, and January gas consumption has been roughly 3 trillion cubic feet in recent years. See EIA data on storage and consumption at https://www.eia.gov/dnav/ng/hist/nw2_epg0_swo_r48_bcfw.htm and https://www.eia.gov/dnav/ng/hist/n9140us2m.htm.

5 Transmission refers to long-distance transport at very high voltages while distribution is done at lower voltage for local delivery to customer locations.

nuclear power, and about 20% from renewables (wind 8.4%, hydro 7.3%, solar 2.3%, biomass 1.4%, and geothermal 0.4%).[6] As shown in Figure 2, this fuel mix is a drastic change from a decade ago, when coal provided a much larger share of generation fuel. Figure 3 shows the share of electricity generation from non-utility producers.

Transmission is generally viewed as a natural monopoly in the sense that the lowest-cost capacity along a given corridor is a single transmission facility. Interstate transmission lines, 80% of which are owned by investor-owned utilities (IOUs), are

Figure 2: Fuel sources for U.S. electricity generation, 1950–2020



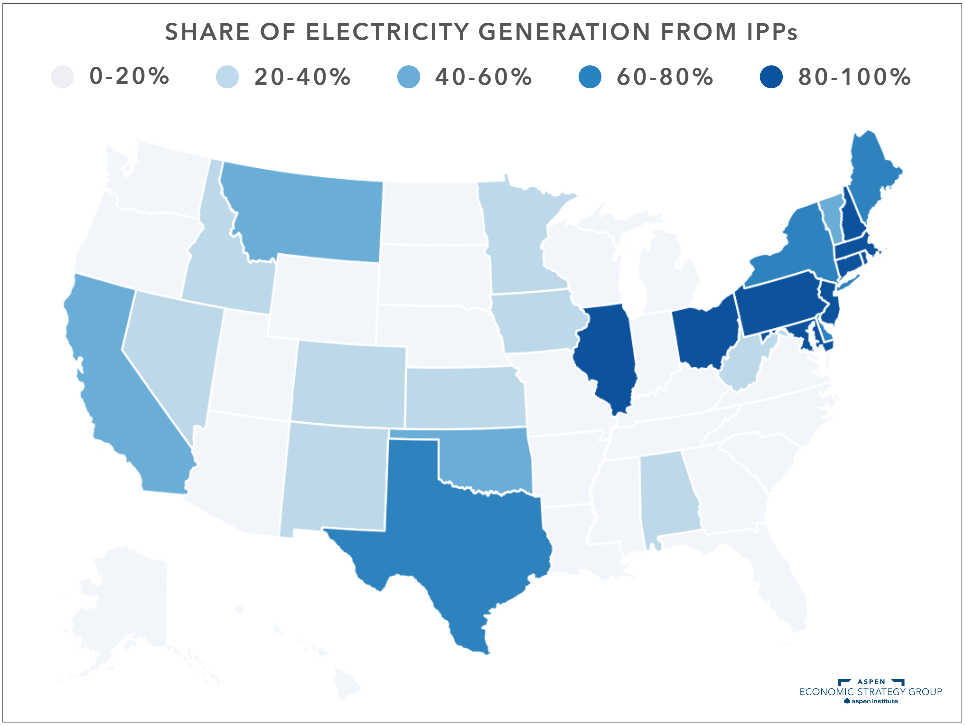
Note: Electricity generation from utility-scale facilities.

Source: U.S. Energy Information Administration, Monthly Energy Review, Table 7.2a, January 2021 and Electric Power Monthly, February 2021, preliminary data for 2020.

6 https://www.eia.gov/energyexplained/electricity/electricity-in-the-us.php.

authorized and regulated by FERC. Nonetheless, siting new transmission lines requires approval from multiple state authorities as well as FERC. Local distribution is indisputably a natural monopoly, with about 70% of electricity distributed by state-regulated IOUs and the remainder by nonprofit municipal utilities and co-ops.[7]

Figure 3: Share of electricity generated by independent power producers (IPPs) in 2019, by state



Source: EIA data on "Net Generation by State by Type of Producer by Energy Source," available at https://www.eia.gov/electricity/data/state/

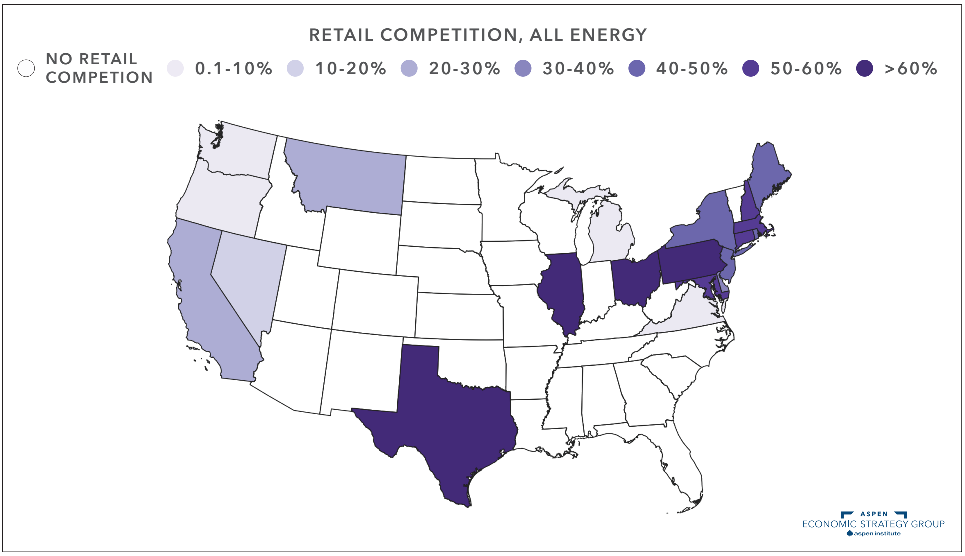
Procurement/retailing is a financial broker function carried out by entities that contract with (or may own) generators for electricity supply and with customers for electricity demand. In most of the country, the same regulated or municipal/co-op utility that provides distribution services also does the procurement on behalf of customers. In a few areas, "retail choice" is available to some or all customer classes, in some cases through for-profit companies (most notably in Texas) and in other cases through nonprofit local governmental organizations – known as "community

7 Unlike in cable television, there is essentially no "overbuild," where two distribution companies serve overlapping areas.

choice aggregation" in California or "municipal electricity aggregation" in Illinois, two states with large numbers of customers served by such entities.[8] Figure 4 shows the share of all electricity purchased that is sold by these non-utility retailers.

Over the last twenty-five years, many electricity markets have significantly restructured, changing from mostly vertically integrated organizations that generated high proportions of the power they distributed to customers, to be much more a supply chain of separate companies that perform the four major functions and interact with one another through market transactions. There is strong evidence that the widespread move to wholesale market competition and trading has increased efficiency and reduced electricity generation costs (Fabrizio, Rose, and Wolfram 2007; Davis and Wolfram 2012; Mansur and White 2012; Cicala 2015; Cicala forthcoming). The more limited evidence on retail choice suggests it has produced fewer net benefits (Wilson and Waddams 2010; Hortaçsu, Madanizadeh, and Puller 2017; Byrne, Martin, and Nah 2019).

Figure 4: Share of electricity purchased under retail choice in 2019



Source: Authors' calculations based on EIA Form EIA-861.

8 More information on these government retailers is at https://www.leanenergyus.org/.

2.b.2. What makes electricity different from other markets?

A combination of physical properties differentiates electricity markets from any other good:

1. Electricity itself is non-storable, and storage of potential energy for rapid conversion to electricity – such as through hydroelectric dams, chemical batteries, or compressed air in underground caverns – requires large capital investments and/or is subject to significant energy loss.

2. Electricity is transported through a network of common carriers – transmission wires – that are somewhat akin to pipes carrying water or gas, except the flow of electricity across connected wires is determined by laws of physics ("loop flow") and very costly to control through an equivalent of valves that can direct it along specific lines.

3. The aggregate supply of electricity injected into this grid of transmission wires must nearly exactly equal the aggregate demand for extracting power from a grid on a second-by-second basis. The same is true for supply and demand at each location if transmission capacity or loop flow constrains the flow of electricity to where it is demanded on the grid. This nearly exact balance between demand and supply must hold despite the fact that there can be substantial demand variation hour-to-hour, some of which is predictable, but some of which is not.

4. Imbalances that are more than very slight and very brief will disrupt the power quality (*e.g.,* electrical frequency) and potentially damage both generating equipment and consuming equipment. As a result, generating equipment is designed to detect frequency deviations and disconnect from the grid if the magnitude of such deviations is too large. Thus, even a small and transitory supply shortage (or surplus) can turn into cascading outages if not managed extremely rapidly. In February 2021, frequency fluctuations brought the Texas grid to within minutes of a cascading outage situation, causing the grid operator to order disconnection of large swaths of customers in order to maintain supply/demand balance. In August 2003, an imbalance in Ohio that was not managed by the local grid operator cascaded across the upper Midwest and Northeast, as well as parts of Ontario, Canada, causing more than 500 generating units to go off-line and customers to lose power across these areas for up to four days.

5. Due to factors 1-4, in every grid there is a central balancing authority that determines at least marginal supply and demand adjustments in order to maintain balance overall and at each node of the grid. Due to the complexity of power flow on the grid and other information constraints, these balancing authorities cannot match supply that has been contracted for specific buyers

with those buyers in real time. So they cannot cut off supply to specific buyers if their counter-parties fail to produce and deliver electricity to the contracted locations, resulting in insufficient supply. Thus, in real time, supply shortages are socialized across all buyers in the market.[9] After the fact, balancing authorities can impose fines on retailers whose customers withdrew more electricity than their suppliers injected into the grid for a specific time interval and location, but these fines are limited by bankruptcy constraints.

As with electricity, most service industries sell products that are non-storable and for which demand is volatile. With other non-storable services, however, a shortage from one supplier is not socialized among all suppliers in the market. Instead, one seller's supply shortage simply leads to queueing or stockouts. For the reasons discussed above, neither is an option with electricity.

These physical attributes of electricity combine with unusual financial, regulatory, and administrative attributes:

1. Due to large economies of scale in construction, it typically makes sense to build transmission and distribution system capacities beyond near-term demand, so the marginal cost of usage is well below the average total cost (inclusive of the fixed costs of construction and upkeep). In such a natural monopoly situation, recovering the full cost from volumetric rates pushes retail prices up relative to marginal cost.

2. Electricity generation from fossil fuels creates significant negative externalities, including climate change externalities, which are for the most part not priced. These unpriced externalities push the retail price down relative to the full social marginal cost of providing electricity.

3. In many areas, electricity rates are used to pay for public policy priorities that are not a marginal cost of providing electricity, such as assistance to low-income customers, subsidies for energy efficiency programs and rooftop solar, procurement of power from immature technologies in order to support their development (climate mitigation), and aggressive vegetation management in increasingly fire-prone areas near electrical wires (climate adaptation).

4. For nearly all customers, electricity is sold under a "requirements contract," meaning that the retailer agrees to supply whatever quantity the customer demands at a preset price. Meeting this obligation while keeping the grid in balance requires investment in substantial generation and transmission capacity that is seldom used.

9 This approach is particularly problematic when there are many small retail providers buying in the wholesale market, each internalizing little of the grid risk of being short, as was the case in Texas in February.

5. Dynamic pricing – under which retail prices adjust at roughly the same time scale as wholesale costs – is extremely uncommon in electricity. Historically, and even today, customers have very little information or technology to respond to changes in the supply/demand balance or wholesale market prices. [10]

3. Energy challenges

The U.S. is at a critical point in the arc of the energy industry and the fight against climate change. The country must rapidly reduce greenhouse gas (GHG) emissions while maintaining energy affordability, reliability, and resilience, and address the historical socioeconomic and racial disparities in both access to energy consumption and harm from energy production. These imperatives present a series of interwoven challenges to the industry and to plans for infrastructure investment.

3.a. Low-carbon technologies

If the primary pathway for a clean energy transition is to be electrification, then the electricity industry must develop systems for delivering near zero-carbon electricity while controlling costs and maintaining grid reliability. Breakthroughs in the costs of nuclear power or carbon capture and sequestration could greatly ease the path to meeting this challenge, but neither seems imminent. The more likely strategy over at least the next decade will be widespread use of intermittent renewables: wind and solar power.

Wind and solar have become cost competitive with fossil-fired generation in many locations when deployed in large-scale developments. However, continued improvements in renewables' supply cost, relative to oil and gas, will require continued innovation for at least two reasons. First, the shale boom has already substantially reduced the cost of production of U.S. oil and gas, and further cost decreases cannot be ruled out. Second, as renewables displace oil and gas, the high-cost oil and gas will be displaced first, leaving lower-cost supply in the market, and lower prices. Even with current technologies, there are large quantities of oil and gas that can be produced at costs well below the current market prices (Asker, Collard-Wexler, and De Loecker 2019). Cutting even 20% or 30% from the demand for these fuels will almost certainly result in declines in their prices. Thus, absent a lucky run of technology breakthroughs in renewables – and none in oil and gas – large-scale substitution away from oil and gas toward renewables will require either substantial mandates and incentives, or direct government expenditures on renewable generation infrastructure.

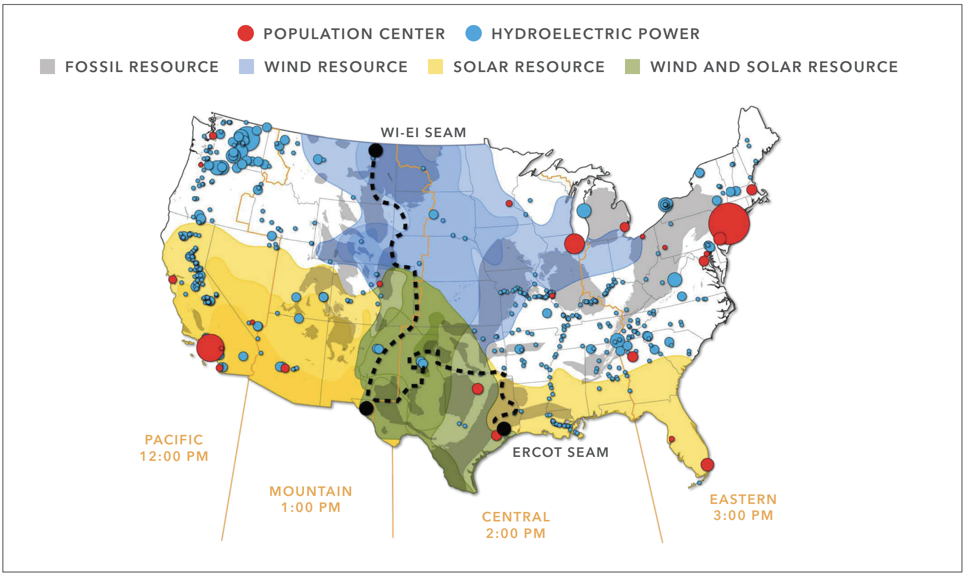
10 In the last decade, regulators have moved toward retail prices that more accurately reflect seasonal and time-of-day average differences in costs, but these average differences capture very little of the actual variation in wholesale costs.

Wind and solar, however, will not be able to support a stable electricity system on their own because their production pattern is determined by nature. Other resources are needed to make sure that supply and demand balance every minute. A combination of four different approaches could maintain grid stability in a system with very high levels of intermittent generation: long-distance power transmission, storage, demand flexibility, and dispatchable generation. The last of these four approaches could come from burning a storable renewable fuel such as hydrogen or biomass, fossil fuel generation (possibly with CCS), or nuclear, hydroelectric, or geothermal power. Each of these four approaches has its own challenges, however.

Transmission

Transmission is critical to taking advantage of resource-rich locations for sun and wind, many of which are far from population centers, as shown in Figure 5. Transmission can also help address the problem that wind and solar are not time-shiftable energy sources by enabling a diversified portfolio of renewable generation. Such a portfolio increases supply stability by decreasing the likelihood that all of a demand center's renewable power sources are idle at the same time. At the same time, transmission also provides more market options for these intermittent resources. More market options mean less need to curtail renewable production if the output of a given facility exceeds the needs of nearby customers, which then implies higher capacity utilization and lower cost per kilowatt-hour.

Figure 5: Map of U.S. energy resources and population centers



Source: National Renewable Energy Laboratory (NREL) "Interconnection Seams Study," available at https://www.nrel.gov/analysis/seams.html

By connecting markets, transmission also enhances competition among generators. Because electricity demand is relatively inelastic, when an electricity market is strained, that is a particularly attractive moment for a generator to reduce output in order to drive up the wholesale price (Borenstein 2002; Borenstein, Bushnell and Wolak 2002) Transmission undermines that incentive by creating more potential competitors in every market (Borenstein, Bushnell, and Stoft 2000).

Though the direct cost of transmission is significant, by most accounts the larger barrier is siting authority, and particularly the web of authorities – including FERC and various agencies in every state through which a line passes – each of which must approve any interstate transmission proposal. These multiple veto points currently raise an extraordinarily high hurdle to interstate transmission investment, as documented in Gold's (2019) account of the failed attempts to connect wind resources in Oklahoma to population centers in the Southeast. Overall, progress on transmission is likely more dependent on reducing the regulatory barriers than on providing government funding.

Storage

Just as a battery in your basement can substitute to some extent for a reliable grid connection, large-scale battery storage can partially substitute for long-distance transmission connections by filling in when intermittent generation fades or demand surges. Battery storage has come down in cost dramatically over the last decade, and it likely has a significant role to play in balancing short-term fluctuations in supply and demand. Current battery technologies, however, are not cost competitive for long-duration storage and not likely to become so in the next decade. Challenges in long-duration storage are particularly relevant if the U.S. is to electrify space heating and depend on it for getting through cold winter periods when renewable generation is often limited.

Demand flexibility

Electricity demand is becoming less predictable and more extreme due to climate change, which has made extreme weather more common over the last 50 years.[11] This increased demand volatility amplifies the challenge of meeting demand with intermittent generation sources. Overall demand for energy services is also growing, though energy efficiency improvements – from light bulbs to refrigerators, air conditioning, and electrical heating – have reduced the growth of demand for electricity itself.

11 https://www.epa.gov/climate-indicators/weather-climate.

One tool for managing electricity systems with demand volatility and large shares of intermittent generation is dynamic demand adjustment. The 2009 American Recovery and Reinvestment Act funded the rollout of electric meters that can measure consumption on a sub-hourly basis and communicate data directly to the utility, but since then the pace of introducing dynamic rates, which incentivize customers to change consumption as supply availability changes, has been slow.

Dispatchable generation

Storable (and transportable) renewable fuels, such as hydrogen, are in some ways technologically ideal, but at this point they are not cost competitive at scale. Credible pathways exist for them to potentially become a significant part of the solution, but thus far they are aspirational (Meyer and Thomas 2021). The same is true for fossil fuel generation with CCS.

Existing nuclear and hydroelectric (and to a lesser extent, geothermal) generation are the front runners for low-carbon, long-duration grid balancing in the near-term as wind and solar ramp up. But the finances of these resources have been weakened by low wholesale prices, a topic we return to below, which threatens the viability of some existing plants and discourages new construction.

Uncertain technological progress

Overall, energy technologies are evolving more rapidly now than at any time in the last century. Despite the confident claims of entrepreneurs and advocates of specific technologies, there is tremendous uncertainty about the pace and direction of future technological change. For every exciting success story, like solar photovoltaics (PV) or lithium ion batteries over the last decade, there are many disappointments, like advanced biofuels or tidal power. Investments in the industry must balance the urgency of reducing carbon emissions and demonstrating new low-carbon technologies with the option value of delaying major financial commitments to obtain more information about which technologies are likely to be most successful. The challenge is not only to balance these factors, but also to design market and policy processes that can adapt rapidly and allow the technologies that deliver the lowest costs and greatest societal benefits to succeed.

3.b. Wholesale market design

While the technological barriers to a stable, near-zero carbon grid are widely discussed, the design challenges for market mechanisms and compensation receive less attention. The resource transition imperative brought about by the climate crisis unfortunately comes at a time when there is also wide disagreement about the best way to operate electricity markets. This debate has many facets, but probably the most important among them is how to reward supply resources whose value does not fit the standard per-kilowatt-hour compensation paradigm. These resources include those that are particularly valuable for their ability to change output rapidly on demand – such as battery storage and hydroelectric generation – as well as capacity that can stand by for very long periods – months or even years – and then reliably operate when needed in a crucial situation.[12]

Market designs created over the last two decades have not settled on how these non-standard electricity market players should be paid, but their importance is growing as weather becomes more extreme and wind and solar generation become a larger share of the portfolio. Texas pioneered one market model that is premised on extremely high electricity prices during times of grid stress. These high prices are intended to incentivize generators to be available and incentivize buyers to sign long-term contracts for supply, helping to finance the construction and continued operation of those generators. Unfortunately, the Texas grid operating under this model fell far short of delivering adequate power supply during the February 2021 Texas energy crisis. On the other hand, alternative models that require energy retailers to contract for standby capacity have also had disappointing outcomes at some peak demand times. In both cases, the risk of extreme financial losses from failures to perform do not seem to have provided sufficient incentives for ensuring supply performance during critical events.

Electricity markets continue to grapple with striking the right balance between market incentives and regulatory guardrails that can lower the cost of providing power without increasing threats to reliability.[13] This problem is not caused by intermittent renewables – as was demonstrated during the February energy crisis in Texas. But it is exacerbated by large-scale deployment of intermittent generation, which adds greater uncertainty on the supply side to the growing uncertainty on the demand side – as was demonstrated by the (much smaller) power shortage California experienced in August 2020.

Wholesale electricity markets have also been strained by renewable electricity policies. While economists generally argue for pricing GHG emissions (as well as

12 A centralized grid operator could, in theory, optimize use of each type of resource in order to minimize costs, though the information and computational requirements of doing so would be enormous. In a market setting, with dispersed ownership and control of supply resources, the technological complexities described in the previous section make it difficult to create compensation mechanisms that incentivize suppliers to behave as efficiently as this theoretical optimum.

13 See, for instance, the discussion of the 2021 Texas outages at https://www.wsj.com/articles/a-failure-of-texas-size-proportionsstate-struggles-to-overhaul-its-power-market-11618565415.

emissions of other pollutants), policymakers have instead preferred to mandate and subsidize renewable power. Rather than increasing wholesale prices, as a carbon price would do, these approaches have reduced wholesale prices.[14] This price reduction has squeezed the profits of nuclear, hydro, and other generation sources that are near carbon free but have not been included in states' subsidies and mandates. This problem has been especially severe for existing nuclear power plants, many of which are not earning revenue sufficient to cover their ongoing operating costs. Because wholesale prices are depressed by renewable generation policies, the standard market test doesn't convey appropriate information about whether these plants should continue to operate.

3.c. Sustainable retail pricing and distributed energy resources

Most models of a clean energy transition suggest that it can be done at a fairly modest cost to the overall economy (National Academies of Science, Engineering and Medicine 2021; Larson et al. 2020; Phadke et al. 2020), but the way that some of the leading states and countries in the low-carbon movement are financing their transitions raises concerns that it could undermine the goals of both large-scale electrification and equity. In California, New York, Germany, and other market-based economies that are aggressively pursuing decarbonization, much of the cost has been covered through increased volumetric electricity rates. As a result, rates in these areas, particularly for residential customers, are now many times higher than the social marginal cost (*i.e.,* the marginal cost inclusive of externalities) of providing incremental electricity.

High volumetric rates are used to cover many costs that do not vary with the volume of electricity delivered, such as most costs of electric infrastructure, climate mitigation (*e.g.*, contracting for new low-carbon forms of generation or storage technologies when they are still quite expensive), climate adaptation (*e.g.,* increased vegetation management), subsidies for rooftop solar and low-income customers, and energy efficiency programs. Many of these programs are worthwhile expenditures, but paying for them through high volumetric rates discourages switching from direct combustion of fossil fuels to electricity. Borenstein and Bushnell (2019) shows that in California, New England, and a few other areas of the U.S. that have made significant investments in climate change mitigation and other social programs, retail prices are two to three times higher than the social marginal cost. These high

14 Subsidizing renewables differs from pricing GHG emissions in other ways. Both approaches encourage renewable generation, but subsidies for renewables don't distinguish between renewable resources that crowd out coal-fired generation versus those that displace natural gas-fired generation (which emits half as much GHGs as coal), or even displace other zero-carbon generation such as hydro, nuclear, or incumbent renewables. Pricing the "bad" directly avoids this problem by appropriately changing the costs of each type of generation (Borenstein 2012).

retail volumetric rates pose a problem if electrification is the preferred pathway to reducing GHG emissions.

High volumetric electricity rates are also a challenge to equity. Borenstein, Fowlie, and Sallee (2021) suggests that high volumetric rates disproportionately burden low-income households. It also shows that paying for climate mitigation and adaptation by raising retail electricity rates is a more regressive approach to funding these programs than a sales tax or gas tax. Historically, the alternative to high volumetric rates has been to impose uniform monthly connection charges for residential customers – which are even more regressive then volumetric rates – or forms of charging commercial and industrial customers based on their own peak usage (known as "demand charges"), which create another set of perverse incentives (Borenstein 2016).

The states that have taken the early steps on climate mitigation and adaptation – and have generally funded their environmental and equity programs from electricity rates – have also been leaders in promoting behind-the-meter (BTM) resources, including rooftop solar installations and batteries. These technologies hold promise for improving system efficiency, but they are also incentivized by retail pricing that exceeds the actual value of additional supply or reduced demand to the grid. Residential rooftop solar, for instance, is thriving mostly in California and other areas of the country with very high retail rates, where more than half of those rates are covering fixed costs of the system or paying for state policies that do not affect the marginal cost of consumption.

BTM resources deployed efficiently can be a valuable part of a decarbonization strategy, but if adopted in response to retail prices that far exceed social marginal cost, they will raise the total cost of achieving carbon reduction goals. In addition, when utilities rely on high volumetric rates to recover their fixed costs, BTM resources reduce utilities' revenue, thereby forcing them to raise their volumetric rates to maintain cost recovery. And because early BTM solar and battery adoption tilts strongly toward wealthier households, the resulting cost shift to other customers is highly regressive (Borenstein and Davis 2016).

Finding funding sources for the non-marginal costs of electricity, along with the related programs that are now paid for through electricity rates, is an under-appreciated challenge of the energy transition. Though this retail pricing distortion is most evident in the states that are leading on climate action, it is likely to become a concern in other areas should federal policies press them to meet a clean energy standard or other carbon reduction mandate.

3.d. Energy infrastructure and equity

Retail pricing problems are far from the only way in which energy systems can impose disproportionate costs on disadvantaged communities. Energy infrastructure has been at the heart of U.S. economic growth for more than a century, but that same infrastructure has produced negative local environmental impacts, for which there is mounting evidence of greater health effects than were recognized even a decade ago.[15] These local pollutants have disproportionately harmed disadvantaged communities. For years, economic analyses have often attributed that disproportionate impact to a differential willingness to pay for environmental amenities as a function of wealth. Recent work, however, shows that the cause goes beyond wealth differentials, and includes lack of political power and racial bias (Hamilton 1995; Mohai, Pellow, and Roberts 2009; Hausman and Stolper 2020).

There are a variety of ways that the transition to zero-emission energy supply might be managed inequitably or prolong the impact of past inequities. One potential mechanism for disproportionate impact is the selection of which fossil fuel–powered installations close early versus later during the transition to zero-emission sources. Broad incentive programs for zero-carbon power, like carbon taxes or clean energy standards, may not discriminate between facilities that emit high versus low levels of local pollutants, or facilities that are near to or far from vulnerable populations. Thus, there is a risk of disproportionate impact if broad, zero-carbon incentive policies fail to reduce emissions from sources of local pollution that are concentrated in “hotspots” near disadvantaged communities. A related concern is that if CCS is a significant mechanism for achieving zero emissions, then plants that continue to burn fossil fuels might still emit local pollutants, even if the carbon is captured.

Evidence from research on emissions markets to date suggests that these markets have not increased disadvantaged communities’ relative exposure to local pollution. Hernandez-Cortes and Meng (2021) studies California’s CO_2 cap-and-trade program using detailed facility-level emissions data from 2008 to 2017, spanning the start of the program in 2013. The paper finds that while local pollution disparities were increasing prior to 2013, they have been decreasing since that time, in a way that appears causally attributable to the CO_2 cap-and-trade program. In related work, Shapiro and Walker (2021) studies markets for local air pollution, using data in California and Texas, and finds that these markets neither increase nor decrease disparities in pollution exposure. While recent experience with pollution markets is therefore reassuring, it remains true that the U.S. does not have experience with

15 See the U.S. Environmental Protection Agency’s Integrated Science Assessment for Particulate Matter (Final Report, Dec 2019), https://cfpub.epa.gov/ncea/isa/recordisplay.cfm?deid=347534.

CO_2 emissions markets at a scale sufficient to reach (or at least draw near to) zero-carbon emissions. The possibility that such a transition may lead to unjust exposure to local pollutants therefore remains a concern.

Finally, inequities can also arise through the siting of new energy infrastructure. Even renewable generation can have adverse impacts on nearby population and property values (Jarvis 2021). History suggests that disadvantaged communities and communities of color are particularly vulnerable because they lack the resources and political influence to fight for fair compensation for these impacts. Energy infrastructure policy should not repeat the mistakes of 20th century highway policies, which displaced and divided low-income urban communities (Rose and Mohl 2012).

3.e. Stranded infrastructure, decommissioning, and human capital

Displacement of oil and gas capital with zero-emission energy infrastructure will also lead to stranded assets. Abandoned wells are already a problem in the upstream extraction industry even in the absence of large-scale climate policies. The U.S. Environmental Protection Agency (EPA) estimates that there are at least 3.6 million abandoned wells in the U.S., two-thirds of which are unplugged and potentially leaking methane (a potent GHG) and other chemicals that endanger the local environment and human health (EPA 2021; Raimi et al. 2021). Aggressive climate policies are likely to exacerbate these problems and spread them to the transportation, refining, and distribution sectors. Furthermore, just as with the operational infrastructure discussed in the previous subsection, this stranded infrastructure is located disproportionately near disadvantaged communities.

These stranded assets present multiple challenges. First, if the last known owner of an abandoned well is insolvent or unknown, the well is then considered to be "orphaned." In this case, the cost of decommissioning the well must be borne by the public. Recent counts have documented 56,600 orphaned wells in the U.S., but the number of undocumented orphaned wells is likely an order of magnitude larger (IOGCC 2019; Raimi, Nerurkar, and Bordoff 2020).

Second, stranded assets include not just physical capital but also human capital. Much of the oil and gas industry workforce is heavily invested in specialized skills. Abrupt economic transitions are known to impose severe burdens on affected workers (Walker 2013), and there is no reason to believe the transition to a green economy will be different. A plan for green infrastructure investment will therefore need to include provisions for maintaining the employment of these workers' skills, both for their own sake and to blunt political opposition to zero-emission policies.

3.f. Grid and supply chain security

We recognize here two other areas of concern, though we don't delve further into them. The first, grid security, would require a paper unto itself best written by a security specialist. The grid is vulnerable to both physical and cyber attacks, much like the vulnerability of hydrocarbon transportation systems demonstrated by the May 2021 Colonial Pipeline system ransomware attack. As with grid reliability more generally, security threats to the grid pose a greater risk as society becomes more dependent on electricity and less diversified in energy delivery modes. While there are compelling arguments for "electrify everything," it is worth recognizing that such a strategy comes with a need for greater grid security investment. Grid security risks also potentially increase the value of distributed generation paired with storage.

The second area is developing and securing the materials supply chain for grid-related hardware. Currently, this issue is discussed most in relation to critical metals and minerals for battery production, but concerns are also raised in relation to solar PV and the electronic components of wind turbines, transformers, and other grid hardware. Many of the necessary materials have only been produced in limited quantities, and there is incomplete knowledge of the available reserves for massively scaling up supply. For some materials, the known supply is mainly in countries that are insecure – such as the Democratic Republic of the Congo, where most of the world's known cobalt reserves are located – or in countries that can be described as having an adversarial relationship with the U.S. in at least some respects – such as the many rare earth minerals for which China is the primary supplier.

3.g. Decarbonizing the world

In 2018, the U.S. was responsible for 15% of worldwide CO2 emissions, down from 24% at the turn of the 21st century.[16] The U.S. share of emissions is likely to continue falling, due in part to U.S. emission reductions but primarily to increases in emissions from other countries whose incomes are increasing rapidly. In the next few decades, much of the world's population will be coming out of extreme poverty and increasing their energy consumption for industrial production, commercial activities, homes, and transportation (Wolfram, Shelef, and Gertler 2012). If they do not have a low-carbon pathway for growth, they will almost certainly pursue high-carbon pathways that are available, just as was done by what are now the richest countries in the world. Climate change is a global challenge, so every domestic policy to address it must also be evaluated for its global impact.

16 https://ourworldindata.org/co2-emissions

4. Implications for infrastructure policy

The challenges facing the energy industry make clear that there will be a need for substantial infrastructure investment this decade and beyond. Broad clean energy incentive policies that effectively put a price on GHGs would boost the energy sector's pivot to a more sustainable path, but as we discuss below there are multiple reasons that other tools will be needed as well in order to meet the challenges we have outlined.

4.a. Low-carbon technologies

The advantages that fossil fuels possess in their low private costs, energy density, transmissibility, and storability make it difficult for zero-emission energy sources to compete with these fuels in energy markets. Absent strong policies, energy markets on their own do not penalize fossil fuels for their emissions. Nor, conversely, do they reward clean energy sources.

A cornerstone of energy infrastructure policy then needs to include significant, broad incentives for zero-emission infrastructure. These incentives can come from carbon pricing, clean energy standards (with tradable credits), clean energy subsidies, or some combination of these policies. To be maximally cost-effective, these policies should broadly include all zero-emission sources, including generation from sources like nuclear, hydro, and fossil-fueled power with CCS that have often been excluded from state-level renewable portfolio standards. In addition, the tradeable credits associated with a clean energy standard should be freely tradeable across state lines, in order to best direct clean energy investment to regions where renewable resources are abundant.

Yet even if the federal government were to adopt a robust clean energy standard or a carbon price equal to the full social cost of GHG emissions, standard economics, equity considerations, and institutional constraints all still suggest that other government action will be important in the pursuit of a low-carbon economy. These actions will be even more needed and valuable if a significant carbon price, clean energy standard, or low-carbon subsidy is not established.

Invest in research and development

Probably the most important reason for additional government action is that rapid innovation across many technologies will be necessary to avoid the worst impacts of climate change, and markets for innovation do not function well. This failure in innovation or knowledge markets occurs because knowledge producers are able to capture only a tiny fraction of the value they create. Thus, without government policies, the incentive to create new knowledge is suboptimal.

The inability of innovators to fully capture the value of their innovations is the fundamental justification for intellectual property protection, such as patents and copyrights. However, such mechanisms create monopolies and legal processes that inhibit use of new knowledge. In other sectors, such as medicine and basic science, knowledge creation is supported directly by the government to a much greater extent than in energy, and the norm is to make some of that knowledge freely available through academic journals.[17]

Climate change, however, creates even more of an imperative for uninhibited knowledge sharing than in other sectors, particularly sharing with the developing world. This argument follows in part from the ethical consideration that most of the GHGs in the atmosphere today were put there by what are now wealthier nations. But wealthy countries also have a self-serving incentive to share low-carbon technologies in order to avoid having nations now coming out of poverty ramp up their growth through intensified use of fossil fuels, because emissions from that pathway will create spillover damages to the wealthier countries. Sadly, the current pandemic is a close analog, creating large private benefits to wealthier nations from sharing vaccine knowledge with poorer nations in return for little or no compensation. This knowledge market failure exists quite apart from appropriate GHG pricing in the U.S. or in all advanced economies. Direct investment in research, development, and deployment of new technologies can create appropriate incentives to innovate in GHG-reducing technologies without creating barriers to worldwide diffusion of those technologies.[18]

Promote charging infrastructure

With transportation electrification playing a major role in virtually every strategy for deep decarbonization, there will be a need for widespread investment in electric vehicle (EV) charging infrastructure, including the electrical distribution system upgrades needed to support the service. Overcoming the network economics challenge of rapidly growing both EVs and EV charging infrastructure is important, but there is still great uncertainty about the best technologies, locations, and business models for EV charging. This uncertainty suggests the need for a flexible

17 See Table 2 in https://ncses.nsf.gov/pubs/nsf21315#data-tables.

18 We include deployment in this discussion of knowledge market spillovers because competitors gain immense information value from the success or failure of another firm's attempt at innovation. In the recent past of the energy industry, examples of valuable information from failures include attempts to build new conventional nuclear power plants in Georgia and South Carolina that resulted in cancellation or massive cost overruns and a number of carbon capture and sequestration facilities that have been canceled before completion. On the positive side, solar panel manufacturers have greatly benefited from learning about process improvements of their competitors, and battery manufacturers are able to extract valuable information from the chemistry of competitors' products without violating intellectual property laws.

approach to government support – such as incorporating energy use for personal transportation into broad clean energy standards or carbon pricing – while leaving EV charger siting decisions and business model experimentation to private firms.

One important area for policy to play a hands-on role in facilitating EV adoption is in harmonizing and enforcing standards for high-speed ("level 3") charging stations. At present, there are three standards for level 3 chargers, and vehicles built to one standard cannot be charged at a station built for a different standard. Li (2019) shows that, since drivers' willingness to adopt EVs depends on the availability of charging, harmonizing standards (or at least requiring charging stations to have adapters) can substantially increase take up of EVs.

Support technology for dynamic demand response

As the country expands the range of energy services powered by electricity, it also has an opportunity to expand the array of devices that can time-shift electricity demand and help to integrate intermittent renewables on the grid. Electric hot water heaters and vehicles likely offer the greatest residential and commercial opportunities for automated demand shifting, which can be enabled with low-cost communicating technologies. Similar technologies have been shown to be effective in shifting electricity demand for space heating/cooling and refrigeration as well (Callaway 2009). Industrial demand also offers demand response potential, which tends to be more specific to the industrial process, but could be equally or more valuable. Efficient dynamic price signals will provide incentives for all customers to adjust their demand in ways that help balance a system with large quantities of intermittent supply.

Simplify regulatory pathways for long-distance interstate transmission

Finally, a massive expansion of long-distance electricity transmission is likely to be a key factor in attaining deep decarbonization at reasonable cost. Government incentives to support transmission that enhances grid access for zero-carbon generation would obviously help drive this expansion, but more important will be policies that simplify – and give investors greater clarity and certainty about – the processes for building these facilities.

Currently, investors seeking to build interstate transmission lines must obtain permission from each state through which the line passes, providing multiple opportunities for incumbent interests to extract rents or exercise veto power. This tortuous regulatory pathway contrasts sharply with that for natural gas pipelines, where FERC is the central certification authority for rights-of-way and use of eminent domain. As discussed at greater length in Cicala (2021), a statute that gives FERC this same authority for interstate transmission (along with institutional capacity with which to execute that authority) would provide a much-needed clear regulatory pathway for transmission investors. In addition, conflicts with local governments and landowners can be minimized by developing a regulatory process that endeavors to use existing public infrastructure rights-of-way – such as railroads, highways, and pipelines – to the maximum extent possible.

4.b. Wholesale market design

The challenges facing wholesale electricity markets reflect the complexity of operating an electricity grid and the uncertainty about how to meet the goals of clean, safe, reliable, and affordable operations as intermittent renewable generation becomes a larger share of the total energy market. It is not possible to predict how specific generation and storage technologies will progress, so it is important to design markets and policies to reward capabilities and performance that help meet these goals.

One implication of these challenges is the need for greater awareness of the impact of subsidizing wind and solar on wholesale market outcomes, an impact that will grow as wind and solar become a larger share of the market. More broadly, these market design challenges argue for supporting research not just on the science and engineering of new electricity resources, but also on the business, economics, and engineering of designing markets that allow the resources to work together efficiently, as well as on how firms will respond, for good or bad, to the incentives created by various market designs.

For instance, Mansur and White (2012) and Cicala (forthcoming) study reforms to wholesale market design in some parts of the U.S. that transformed systems of decentralized, bilateral power trades to highly organized and centralized wholesale auctions. Both papers demonstrate that these reforms substantially increased inter-regional trade and reduced generation costs by promoting the efficient dispatch of relatively low-cost generation units. This work suggests that geographic expansion of such reforms can help promote the dispatch of renewable generation (which has nearly zero marginal generation cost once installed) while increasing the transparency of price incentives that guide investment in new generation and transmission infrastructure.

The February 2021 electricity crisis in Texas also highlights the need for wholesale market designs to address the challenges that extreme weather events pose for grid reliability. These challenges will only become more acute over time as climate change increases the frequency, duration, and severity of extreme temperatures and storms. The Texas disaster demonstrated that even very large financial incentives for energy production were insufficient to compel the electricity supply chain to deliver power during critical conditions.

The challenges that extreme weather events pose to grid reliability are in some ways similar to the challenges that systemic risk poses to financial markets. During these events – and as occurred in Texas – suppliers that fail to perform can avoid the full costs of non-performance through insolvency. This judgment-proof problem dampens incentives to take steps in advance, such as weatherization, that help ensure performance during critical events. Non-performance can cascade through the system physically as discussed in Section 2. In addition, non-performance and insolvency can cascade financially via the contracts that link utilities, retailers, generators, and fuel (especially natural gas) providers. For instance, non-performance by a natural gas supplier can then lead to non-performance by generators dependent on that gas. Those generators then must, in order to fulfill their contracts with utilities and retailers, purchase large volumes of extraordinarily expensive electricity on the spot market, leading to their insolvency (and possibly the insolvency of the non-performing gas supplier as well).

The problem of ensuring reliability in the face of these challenges is daunting. And much like the problem of ensuring financial stability, the ideal solution is far from clear. One route is to increase the use of direct regulations such as weatherization standards. Such approaches can be valuable, though they run the risk of "responding to the last crisis" rather than being forward-looking, and they require dedicated enforcement by government agencies. Another approach would be to require financial assurance for performance during critical events. Such assurance could be provided by diversified financial entities, who would have the ability to demand preventative steps (such as weatherization) as a condition of underwriting performance contracts.

4.c. Sustainable retail pricing and distributed energy resources

As discussed in Section 3, states (and many countries) that have pursued the most aggressive climate policies have financed them in large part by raising retail electricity prices such that those prices are now greater than social marginal cost. These high prices discourage electrification, create perverse incentives for BTM generation and storage, and disproportionately burden low-income customers. These problems have a number of implications for infrastructure policy.

First, policy evaluations and support for investment in BTM resources should be based on the true avoided social marginal cost of that BTM generation in the long-run decarbonization plan, not on retail price nor on counterfactuals premised on historical fossil fuel generation. Going forward, BTM generation and storage will be substituting primarily for grid-scale low- and zero-carbon resources, and for storage technologies that benefit from scale economies. BTM generation and storage has characteristics that still give it value – such as enhanced customer resilience, avoided land use disputes, and potential for reduced demands on transmission and distribution networks. Those factors must be weighed against the higher cost per kilowatt-hour of energy delivered. Retail pricing that reflects social marginal cost is a powerful tool with which to incorporate these tradeoffs into households' and firms' decision-making.

Second, any electricity revenues collected above social marginal cost should be seen as a distortionary tax, just as taxes on labor, capital, and consumer goods are distortionary. In the case of the states leading on climate change, the electricity tax raises the prices to two or more times the social marginal cost, a far larger gap than is seen in most other sectors of the economy. One justification for this pricing has been that electricity demand is highly inelastic, so the deadweight loss from this tax has still been small. However, as customers are now more able to substitute among fuels for many of their energy services – electricity versus gasoline for vehicles and electricity versus natural gas for space and hot water heating – and as they now have more options for BTM generation and storage, the inefficiency from these distorted prices is growing.

Third, this electricity tax should also be recognized to be highly regressive. In California, the location of about half of the country's residential BTM solar capacity, *net* electricity consumption (after adjusting for BTM generation) of residential customers in the wealthiest quintile of census block groups is now only slightly higher than for poorer areas. Nearly any other funding source, including sales taxes and gasoline taxes, would be more progressive than this electricity tax. Fixed monthly charges that are indexed to income may be part of the solution, but there isn't an equivalent charge for commercial and industrial customers, which consume 62% of all electricity.[19]

Many costs similar to those covered by the electricity tax in the energy sector are paid from state or federal budgets in other sectors. Food and healthcare subsidies for low-income families are typically government budget items, but electricity subsidies for the poor come from electricity prices. Climate adaptation, such as sea walls, road and bridge reinforcement, and forest management, are paid primarily from local, state, and federal budgets. Yet adapting transmission and distribution

---

19 https://www.eia.gov/electricity/monthly/epm_table_grapher.php?t=table_5_01

infrastructure, as well as greatly expanding utility vegetation management, is currently financed from electricity prices.

Covering the costs of infrastructure, climate mitigation, and climate adaptation policies through funding mechanisms that do not rely on retail electricity pricing should be a priority. Direct government support for policies that will otherwise raise volumetric electricity rates – including energy efficiency programs, early investment in new low-carbon energy technologies, and climate mitigation and adaptation programs, among others – can mitigate excessive electricity rates in areas that have invested in clean energy and encourage such investment in areas that have not.

4.d. Energy infrastructure and equity

Policies that incentivize zero-emission energy infrastructure must also be accompanied by policies that guard against the possibility of local pollution "hotspots" during the transition. One model for doing so might be California, which passed legislation (Assembly Bill 617) in 2017 that augmented its CO2 market with policies to address the needs of communities affected by local pollution. Fowlie, Walker, and Wooley (2020) discusses how this legislation responded to communities' frustrations with California's carbon policies, largely by empowering community organizations with enhanced monitoring capabilities and authority to address local pollution problems. Fowlie, Walker, and Wooley (2020) concludes that, while it is still too early to definitively evaluate this program, California's community driven approach to local pollution could be a useful model for a complement to federal climate policies.

Transportation electrification, particularly if it can substitute for diesel engines in heavy trucks and rail locomotives, can also contribute to reducing local pollution in populated areas. Public funding of research and development can help lower technological hurdles to electrification of heavy transportation.

Distributive and procedural justice must also be a consideration for the siting of new, zero-emission energy infrastructure. The federal transmission siting authority that we argue for in Section 4.a. above cannot be a blank check for use of eminent domain to condemn the land of disadvantaged households. Instead, and as argued by Cicala (2021), new transmission infrastructure should primarily leverage existing rights-of-way, with judicious use of eminent domain as a last resort option.

4.e. Stranded infrastructure, decommissioning, and human capital

When energy infrastructure reaches the end of its useful life, it can present environmental hazards if not decommissioned properly. Orphaned oil and gas wells are perhaps the most prominent example of improperly decommissioned infrastructure that then becomes a burden to the public. But zero-emission infrastructure can present environmental hazards as well. Batteries, for instance, contain toxic compounds. Investment in energy infrastructure must therefore be accompanied by policies that provide assurance that decommissioning will occur.

Liability rules alone are insufficient to assure decommissioning due to what is known as the "judgment-proof" problem (Shavell 1986, 2002). If the owner of an asset at the end of its life is highly leveraged, it will be able to avoid decommissioning liabilities via bankruptcy.

One way to address the judgment-proof problem is to require firms to post a decommissioning bond at the time of investment. Bonding policies only work, however, to the extent that the bond amounts are commensurate with decommissioning costs. For oil and gas wells, state and federal bond requirements are often well below estimated decommissioning costs, leading directly to the widespread existence of orphaned wells (Boomhower 2019; GAO 2019).

Evidence from oil and gas bonding reforms in Texas shows that increased bond amounts can reduce orphaned wells and environmental incidents, primarily by encouraging small, poorly capitalized firms to divest their assets to better-capitalized entities (Boomhower 2019). Such reforms must become more widespread in order to end the proliferation of orphaned wells in the U.S. Moreover, adequate bonding requirements need to be imposed on new energy infrastructure development in order to save public funds from being saddled with decommissioning expenses for future abandoned assets.

Improvements in bonding policy cannot, however, remedy the sins of past policy inadequacy. The environmental hazards presented by the current stock of orphaned wells can only be addressed by direct public investment in their decommissioning. The decommissioning of all of these wells – including restoration of their surroundings – is likely to cost more than $10 billion (Raimi et al. 2020; Raimi, Krupnick, Shah, and Thompson 2021).

A silver lining to this necessary expenditure is that it provides an opportunity to employ oil and gas workers who would otherwise be displaced by the transition to zero-emission energy infrastructure. Raimi et al. (2020) estimates that decommissioning the full inventory of orphaned wells would require 100,000 person-years of labor. The natural source of such labor would be displaced oil and gas workers, whose skills would naturally transfer to well decommissioning work.

Even putting aside the inherent value of eliminating the health and environmental threats posed by orphaned wells, providing opportunities to displaced oil and gas workers conveys benefits of its own. Well decommissioning would employ a valuable stock of human expertise that would otherwise depreciate away were these workers to move to another sector or drop out of the labor force. These job opportunities may also soften political opposition to the zero-emission energy transition, and from a distributive justice perspective help reduce inequitable impacts of green infrastructure policies.

Well decommissioning on its own is not going to be a long-run solution to displacement of oil and gas workers, whose workforce has numbered around 400,000 people over the past several years (Raimi et al. 2020). Nonetheless, it could at least serve as a temporary bridge to future employment opportunities, including opportunities in hydrogen, geothermal energy, offshore wind, and carbon sequestration, that could play a role in a zero-carbon economy.

4.f. Grid and supply chain security

We are not in a position to offer specific guidance on investments to address grid security or supply chain security. We will note, however, that both are areas with immense spillovers beyond the organizations that are directly affected by security failures. The Texas crisis of February 2021 demonstrated the cascading economic and health impacts when the grid fails to deliver extremely high reliability. If decarbonization makes electricity more dominant in energy supply, any vulnerabilities become more of a threat to society. Supply chain security presents less risk of a short-term acute crisis than does grid security. But it conveys serious risk of stalling long-term progress on energy goals, with implications beyond the firms directly involved in the supply chain.

4.g. Decarbonizing the world

Recognizing that the only GHG reduction goal that really matters is the global one has implications that are often lost in the U.S. energy and climate debate. Every proposed investment should be judged on how it could scale to reduce global emissions and on the knowledge that can be gained and applied globally. One immediate implication is that independent ex post evaluation is central to the value of U.S. energy infrastructure policies and investments that address climate change. To make such evaluation possible, programs should be designed at the front end in ways that maximize the reliability of evaluation at the back end.

Unfortunately, it now also seems possible that the world will not meet the challenge of reducing greenhouse gas emissions as quickly or effectively as necessary to avert extreme temperature changes and other catastrophic consequences. Though carbon dioxide removal and solar radiation modification in some ways lie outside the area of energy infrastructure, it is necessary to also note the high potential returns to investments in RD&D in these technologies (Keith and Deutch 2020).

5. Conclusion

The defining energy challenge of the 21st century is to transition the provision of energy services to zero-emission sources, while simultaneously controlling costs and ensuring the reliability of energy supply. This transition will require historic investments in zero-emission energy generation, transmission, storage, and distribution infrastructure. Federal policy choices will play a leading role in determining whether, where, and when these investments will occur, how costly they will be, and who will bear those costs.

The zero-emission technologies that have thus far become cost competitive at scale are wind and solar PV. However, the intermittent nature of generation from these resources will present ever greater challenges to grid reliability as their share of generation capacity increases. Investment in technologies such as long-distance transmission, battery storage, hydrogen, nuclear, geothermal, or carbon capture and sequestration will be needed to ensure reliable, on-demand energy services.

Because the rate of future technological advances is uncertain, it is difficult to know in advance which combination of technologies will provide the lowest-cost solution. One way to avoid over-investment of public funds into technologies that ultimately fail is to employ policies – such as carbon prices, clean energy standards, or clean energy subsidies – that provide broad incentives for zero-emission energy supply without discriminating across different technologies.

Broad incentive policies should therefore be a core component of energy infrastructure policy, but on their own they will leave unsolved a variety of challenges and market failures. Some of these challenges concern investment barriers that cannot be surmounted by investment incentives alone, and others concern ensuring that the benefits and costs of the clean energy transition are shared equitably. We therefore advocate that energy infrastructure policy include a number of features in addition to broad clean energy production incentives:

- ***Invest in research, development, and early-stage deployment of novel technologies.*** The climate challenge is a global challenge, and an essential way to encourage other nations to reduce their own emissions will be to develop low-cost zero-emission technologies and then export those technologies around the globe.

- ***Improve the design and price transparency of wholesale power markets.*** Improvements in wholesale market design can help accommodate increased use of intermittent generation resources and ensure reliability in the face of increased climate-driven demand and supply uncertainty. Possibilities include increased use of wholesale power auctions to improve price transparency and facilitate inter-regional trade, development of mechanisms to compensate resources that can change output rapidly on demand, and implementation of more targeted reliability regulations to guard against the risk of system-wide failures.

- ***Enhance federal authority over long-distance transmission siting.*** Construction of long-distance transmission is currently hobbled by multi-jurisdictional control over approvals. Significant new investments will require centralized federal authority, like that which currently exists for natural gas pipeline construction.

- ***Reform retail electricity rates to more accurately reflect society's full marginal cost.*** In many parts of the country, the current practice of relying on volumetric charges to recover the costs of climate response, fixed infrastructure, and other public purpose programs discourages electrification, distorts end-user investment incentives, and disproportionately burdens lower-income households. Covering many of these costs through state and local budgets would be more efficient and equitable. Retail rates in these locations could also make greater use of fixed connection charges, particularly by making them income-based.

- ***Address local pollution and involve local communities.*** During the transition, policies on local pollutants must ensure that disadvantaged communities are not disproportionately affected by pollution from fossil fuel power plants, industrial facilities, and homes that are slow to switch to alternative fuels or shut down. Local communities should be empowered to play a role in local pollution monitoring and enforcement.

- ***Ensure funding for infrastructure decommissioning.*** Infrastructure inevitably depreciates and must be decommissioned. Authorizing agencies should require that new investments be accompanied by bonds that ensure that future decommissioning costs are covered. And to address the orphaned wells problem – a consequence of insufficient bonding requirements in the past – public funds can cover the required decommissioning while simultaneously providing employment opportunities to oil and gas workers who are displaced by the clean energy transition.

References

Anderson, Soren T., Ryan Kellogg, and Stephen W. Salant (2018), "Hotelling Under Pressure," *Journal of Political Economy,* 126 (3), 984–¬1026.

Asker, John, Allan Collard-Wexler, and Jan De Loecker (2019), "(Mis)Allocation, Market Power, and Global Oil Extraction," *American Economic Review,* 109 (4), 1568–1615.

Boomhower, Judson (2019), "Drilling Like There's No Tomorrow: Bankruptcy, Insurance, and Environmental Risk," *American Economic Review,* 109 (2), 391–426.

Borenstein, Severin (2002), "The trouble with electricity markets: understanding California's restructuring disaster," *Journal of Economic Perspectives,* 16 (1), 191–211.

Borenstein, Severin (2016), "The Economics of Fixed Cost Recovery by Utilities," *The Electricity Journal,* 29, 5–12.

Borenstein, Severin and James Bushnell (2019), "Do Two Electricity Pricing Wrongs Make a Right? Cost Recovery, Externalities, and Efficiency," Energy Institute at Haas working paper #294, Revised July 2019.

Borenstein, Severin, James B. Bushnell, and Steven Stoft (2000), "The competitive effects of transmission capacity in a deregulated electricity industry," *RAND Journal of Economics,* 31(Summer).

Borenstein, Severin, James B. Bushnell, and Frank A. Wolak (2002), "Measuring market inefficiencies in California's restructured wholesale electricity market," American Economic Review, 92 (5), 1376–1405.

Borenstein, Severin, and Lucas W. Davis (2016), "The distributional effects of US clean energy tax credits," Tax Policy and the Economy, 30 (1), 191–234.

Borenstein, Severin, Meredith Fowlie, and James Sallee (2021), "Designing Electricity Rates for an Equitable Energy Transition," Energy Institute at Haas working paper #314.

Byrne, David P., Leslie A. Martin, and Jia Sheen Nah (2019), "Price discrimination, search, and negotiation in an oligopoly: A field experiment in retail electricity." Working paper. Available at SSRN: https://ssrn.com/abstract=3352209

Callaway Duncan S. (2009), "Tapping the energy storage potential in electric loads to deliver load following and regulation, with application to wind energy," Energy Conversion and Management, 50, 1389–1400.

Cicala, Steve (2015). "When does regulation distort costs? Lessons from fuel procurement in US electricity generation," *American Economic Review,* 105 (1), 411–444.

Cicala, Steve (2021), "Decarbonizing the U.S. Economy with a National Grid," U.S. Energy and Climate Roadmap, Energy Policy Institute at the University of Chicago.

Cicala, Steve (forthcoming), "Imperfect Markets versus Imperfect Regulation in U.S. Electricity Generation," *American Economic Review.*

Davis, Lucas W., and Catherine Wolfram (2012). "Deregulation, consolidation, and efficiency: Evidence from US nuclear power," American Economic Journal: Applied Economics, 4 (4), 194–225.

Environmental Protection Agency (2021), "Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990-2019," Washington, D.C. https://www.epa.gov/sites/production/files/2021-02/documents/us-ghg-inventory-2021-main-text.pdf.

Fabrizio, Kira R., Nancy L. Rose, and Catherine D. Wolfram (2007). "Do markets reduce costs? Assessing the impact of regulatory restructuring on US electric generation efficiency," American Economic Review, 97 (4), 1250–1277.

Fowlie, Meredith, Reed Walker, and David Wooley (2020), "Climate Policy, Environmental Justice, and Local Air Pollution," Brookings Economic Studies Program.

Gold, Russell (2019), Superpower: One Man's Quest to Transform American Energy, New York: Simon & Schuster.

Government Accountability Office (2019), "Bureau of Land Management Should Address Risks from Insufficient Bonds to Reclaim Wells," report GAO-19-615.

Hamilton, James T. (1995), "Testing for Environmental Racism: Prejudice, Profits, Political Power?" *Journal of Policy Analysis and Management,* 14 (1), 107–132

Hausman, Catherine and Samuel Stolper (2020), "Inequality, Information Failures, and Air Pollution," National Bureau of Economic Research working paper #26682.

Hernandez-Cortes, Danae and Kyle C. Meng (2021), "Do Environmental Markets Cause Environmental Injustice? Evidence from California's Carbon Market," National Bureau of Economic Research working paper #27205.

Hoekman, S. Kent and Amber Broch (2018), "Environmental Implications of Higher Ethanol Production and Use in the U.S.: A Literature Review. Part II - Biodiversity, Land Use Change, GHG Emissions, and Sustainability," *Renewable and Sustainable Energy Reviews*, 81 (2), 3159–3177.

Hortaçsu, Ali, Seyed Ali Madanizadeh, and Steven L. Puller (2017), "Power to choose? An analysis of consumer inertia in the residential electricity market," American Economic Journal: Economic Policy, 9 (4), 192–226.

Interstate Oil and Gas Compact Commission (2019), "Idle and Orphan Oil and Gas Wells: State and Provincial Regulatory Strategies," https://iogcc.ok.gov/articles/iogcc-issues-report-idle-and-orphan-oil-and-gas-wells.

Jarvis, Stephen (2021), "The Economic Costs of NIMBYism: Evidence from Renewable Energy Projects," Energy Institute at Haas working paper #311, revised January 2021.

Jones, Benjamin F. (2021), "Science and Innovation: The Under-Fueled Engine of Prosperity," in Rebuilding the Post-Pandemic Economy, Aspen Institute Economic Strategy Group.

Keith, David W. and John M. Deutch (2020), "Climate Policy Enters Four Dimensions,,, in Securing Our Economic Future, Aspen Institute Economic Strategy Group.

Larson, E., C. Greig, J. Jenkins, E. Mayfield, A. Pascale, C. Zhang, J. Drossman, R. Williams, S. Pacala, R. Socolow, EJ Baik, R. Birdsey, R. Duke, R. Jones, B. Haley, E. Leslie, K. Paustian, and A. Swan (2020), "Net-Zero America: Potential Pathways, Infrastructure, and Impacts," interim report, Princeton University, Princeton, NJ, December 15, 2020.

Li, Jing (2019), "Compatibility and Investment in the U.S. Electric Vehicle Market," working paper.

Mansur, Erin T. and Matthew W. White (2012), "Market Organization and Efficiency in Electricity Markets," Dartmouth Tuck Working Paper.

Meyer, Gregory and Nathalie Thomas (2021), "Hydrogen: the future of electricity storage?" *Financial Times,* April 4, 2021.

Mohai, Paul, David Pellow, and J. Timmons Roberts (2009), "Environmental Justice," *Annual Review of Environment and Resources,* 34, 405–430.

Mullins, Kimberley A., W. Michael Griffin, and H. Scott Matthews (2011), "Policy Implications of Uncertainty in Modeled Life-Cycle Greenhouse Gas Emissions of Biofuels," Environmental Science and Technology, 43 (1), 132–138.

National Academies of Sciences, Engineering, and Medicine (2021), The Future of Electric Power in the United States, Washington, DC: The National Academies Press. https://doi.org/10.17226/25968.

Phadke, Amol, Umed Paliwal, Nikit Abhyankar, Taylor McNair, Ben Paulos, David Wooley, Ric O'Connell (2020), "Plummeting Solar, Wind, and Battery Costs Can Accelerate Our Clean Electricity Future," Goldman School of Public Policy report, UC Berkeley.

Raimi, Daniel, Alan Krupnick, Jhih-Shyang Shah, and Alexandra Thompson (2021), "Decommissioning Orphaned and Abandoned Oil and Gas Wells: New Estimates and Cost Drivers," preprint.

Raimi, Daniel, Neelesh Nerurkar, and Jason Bordoff (2020), "Green Stimulus for Oil and Gas Workers: Considering a Major Federal Effort to Plug Orphaned and Abandoned Wells," Joint Report by the Columbia SIPA Center on Global Energy Policy and Resources for the Future.

Rose, Mark H. and Raymond A. Mohl (2012), Interstate: Highway Politics and Policy Since 1939, 3rd edition. University of Tennessee Press.

Scully, Melissa J., Gregory A. Norris, Tania M. Alarcon Falconi, and David L. MacIntosh (2021), "Carbon Intensity of Corn Ethanol in the United States: State of the Science," *Environmental Research Letters,* 16 (4), 043001.

Shapiro, Joseph S. and Reed Walker (2021), "Where is Pollution Moving? Environmental Markets and Environmental Justice," National Bureau of Economic Research working paper #28389.

Shavell, Steven (1986), "The Judgment Proof Problem," *International Review of Law and Economics*, 6 (1), 45–48.

Shavell, Steven (2002), "Minimum Asset Requirements," National Bureau of Economic Research working paper #9335.

Walker, W. Reed (2013), "The Transitional Costs of Sectoral Reallocation: Evidence from the Clean Air Act and the Workforce," *Quarterly Journal of Economics,* 128(4), 1787–1835.

Wilson, Chris M., and Catherine Waddams Price (2010), "Do consumers switch to the best supplier?" *Oxford Economic Papers,* 62 (4), 647–668.

Wolfram, Catherine, Orie Shelef, and Paul Gertler (2012), "How will energy demand develop in the developing world?" *Journal of Economic Perspectives,* 26 (1), 119-38.

Science and Innovation: The Under-Fueled Engine of Prosperity

AUTHORS

Benjamin F. Jones*

ABSTRACT

Science and innovation are central to human progress and national economic success. Currently, the United States invests 2.8% of GDP in research and development, which is supported by a range of public policies. This chapter asks whether the United States invests enough. To answer that question, the conceptual case for government intervention and skepticism about that case are reviewed. The chapter then turns to systematic evidence, including the very latest evidence, regarding the operation of the science and innovation system and its social returns. This evidence suggests a clear answer: We massively underinvest in science and innovation, with implications for our standards of living, health, national competitiveness, and capacity to respond to crisis.

* Kellogg School of Management and National Bureau of Economic Research. Email: bjones@kellogg.northwestern.edu

1. Introduction

Scientific and technological advances have long been recognized as engines of economic growth and rising prosperity. The fruits of these advances—instantaneous global communications, vaccines, airplanes, heart surgery, computers, skyscrapers, industrial robots, on-demand entertainment, to name a few—might seem almost magical to our ancestors from not-too-many generations ago. The power of this progress has been broadly evident since the Industrial Revolution and was recognized at the time, including by political leaders. As the British Prime Minister Benjamin Disraeli noted in 1873, "How much has happened in these fifty years … I am thinking of those revolutions of science which … have changed the position and prospects of mankind more than all the conquests and all the codes and all the legislators that ever lived."

Disraeli was talking of things like the steam engine, the telegraph, and textile manufacturing. In the century and a half since Disraeli's observation, standards of living have advanced remarkably amidst the continued progress of science and technology. Real income per-capita in the United States is 18 times larger today than it was in 1870 (Jones 2016). These gains follow from massive increases in productivity. For example, U.S. corn farmers produce 12 times the farm output per hour since just 1950 (Fuglie et al. 2007; USDA 2020). Better biology (seeds, genetic engineering), chemistry (fertilizers, pesticides), and machinery (tractors, combine harvesters) have revolutionized agricultural productivity (Alston and Pardey 2021), to the point that in 2018 a single combine harvester, operating on a farm in Illinois, harvested 3.5 million pounds of corn in just 12 hours (CLASS, n.d.). In 1850, it took five months in a covered wagon to travel west from Missouri to Oregon and California, but today it can be done in five hours—traveling seven miles up in the sky. Today, people carry smartphones that are computationally more powerful than a 1980s-era Cray II supercomputer, allowing an array of previously hard-to-imagine things—such as conducting a video call with distant family members while riding in the back of a car that was hailed using GPS satellites overhead.

Improvements in health are also striking: Life expectancy has increased by 35 years since the late 19th century, when about one in five children born did not reach their first birthday (Murphy and Topel 2000). Back then, typhoid, cholera, and other diseases ran rampant, Louis Pasteur had just formulated the germ theory of disease, which struggled to gain acceptance, and antibiotics did not exist. In the 1880s, even for those who managed to reach age 10, U.S. life expectancy was just age 48 (Costa 2015). Overall, when examining health and longevity, real income, or the rising productivity in agriculture, transportation, manufacturing, and other sectors of the economy, the central roles of scientific and technological progress are readily apparent and repeatedly affirmed (Mokyr 1990; Solow 1956; Cutler et al. 2006; Alston and Pardey 2021; Waldfogel 2021).

But the stakes in science and innovation go beyond the longer-run rise of economic prosperity and health. Science and innovation are also central to confronting emergent threats. The COVID-19 pandemic, and the key role that novel vaccines have played in the U.S. recovery, demonstrate the importance of science and innovation to national resilience. Similarly, rapid scientific and technological advances were key to U.S. success in World War II (Snow 1959; Gross and Sampat 2020). Indeed, keeping ahead of one's adversaries through science and technology has long been recognized as central to national defense (Bush 1945; NAS 2007; Center and Bates 2019). Whether facing a pandemic, climate change, cybersecurity threats, outright conflict, or other challenges, a robust capacity to innovate—and to do so quickly—appears central to national security and national resilience.

Further, in a globalized world, workers compete in a global context. The productivity and comparative advantage of a nation's workforce depend on the advanced tools and skills that a nation can bring to its production. Innovation is thus key to creating high-paying jobs and maintaining national economic leadership, a benefit that the United States and its workforce have long enjoyed. And to the extent that systems of government depend on the visible success of the nation's economic models, U.S. scientific and technological progress have supported the attractiveness and durability of democracy and free market systems, playing a key role in resolving the Cold War in favor of liberty and facing new competition with authoritarian systems like China (Shirk et al. 2020). As with national security, the position of the U.S. economy in a global landscape hinges on keeping ahead—on continual progress in science and technology. Ultimately, scientific and technological advance not only drive improved standards of living and longer and healthier lives, but these advances underpin national economic success, national security, and the attractiveness of the national model.

These perspectives all point to the central role of science and innovation in the national interest. At the same time, there is a distinction between recognizing the deep contributions of science and innovation and saying that *government and public policy* have big roles to play in driving science and innovation. The purpose of this chapter is to focus precisely on this case. What is the case for a substantial government role in science and innovation? What is the evidence? How are we doing? What policy changes do we need? In answering these questions, the available evidence, including the very latest evidence, will suggest important answers. Namely, based on what we know now, the United States (and the world) appear to greatly underinvest in science and innovation. Investing in science and innovation is perhaps the world's greatest market failure and policy changes will be essential to doing more.

This chapter proceeds in three parts. First, it considers the role of science and innovation as "public goods," the conceptual basis for understanding why private markets underinvest in science and innovation and why there is an essential role for public action. At the same time, this chapter directly engages common forms of public skepticism about the value of scientific research, including the view that science may be isolated from the broader public interest and the issue that science and innovation investments often fail. These perspectives are laid out in Section 2 together with the usual kinds of anecdotal evidence that are used to illuminate them. Section 3 then turns to systematic evidence, drawing in the latest findings to see what is true in general, as opposed to in isolated anecdotes. This section argues that, when examined in light of systematic empirical evidence, there is a clear case for strengthening public support. Finally, Section 4 considers policy innovations that can bring the United States greater scientific and technological success—to create higher standards of living, longer and healthier lives, an increasingly competitive workforce, a more resilient nation, and a more effective model for the world. Section 5 concludes.

2. Science and innovation policy: public goods and public skeptics

Why should government and public policy have an important role to play in encouraging scientific and technological progress? The case hinges on the idea that the market, left to itself, provides insufficient incentive to invest in new ideas and thus underinvests in science and innovation. This perspective in turn motivates an array of approaches where public policy may work to increase science and innovation investments in line with their social value. This section will first lay out these conceptual arguments, with examples. Several forms of skepticism about science and innovation policy are then considered, also with examples. Systematic evidence will be considered in the following sections.

2.a. Science and innovation as a public good

Markets may function efficiently for the production of many ordinary goods and services. But the outputs of science and innovation are not ordinary goods. The outputs are, at root, new ideas: new knowledge and ways of doing things. And ideas have unusual properties. As Thomas Jefferson once observed, "He who receives an idea from me, receives instruction himself without lessening mine; as he who lights his taper at mine, receives light without darkening me." Jefferson is observing that an idea, once it is created, can bring benefit not just to the creator but to many additional parties—it can light other candles, creating benefits far beyond that first candle's light. When Isaac Newton discovered calculus, or Henry Ford introduced the assembly line, or Emmanuelle Charpentier and Jennifer Doudna developed the gene-editing tool CRISPR, they shined a new light on the world, a light that others could use.

This potential for broader use elevates the social value of ideas. Yet the value in this broader use may be difficult for the initial innovator to capture. Rather, the benefits from that spreading light may be largely captured by others. For example, others may use the exact same discovery, tool, or idea—say, calculus, the structure of DNA, or a machine-learning algorithm. Others may similarly use the original idea as inspiration for distinct variants—electric vehicles, mRNA-based vaccines, or cloud-based computing services. With the advance of ideas, some party engages in costly and risky work to discover or develop a new idea. Then, inspired by the original innovator, the social value spreads to many other parties. To the extent that the original creator does not capture this broader social value, the private value can fall short of the social value it creates. Then the private incentives to invest in creating the idea may be well below the social interest in making that investment. This is the basic market failure, and incentive problem, that surrounds the advance of ideas.

Beyond more immediate imitative spillovers, additional social value comes over time, where one advance unlocks doorways to further scientific or technological progress. These so-called “intertemporal spillovers” can be both valuable and unpredictable—and difficult for the initial innovator to capture. Examples in marketplace innovation include the personal computer, Internet, or smartphone, the creation of which opened the doorway to enormous arrays of novel software applications and business models. Intertemporal spillovers are also particularly germane in science. In science, an advance typically has no direct marketplace application but rather is a step forward in the deeper understanding of nature. Yet this deeper understanding of nature may prove essential to future marketplace innovation. Vannevar Bush, who led the U.S. science and technology efforts in World War II, evocatively described science along these lines as “the fund from which the practical applications of knowledge must be drawn” (Bush 1945).

To illuminate such spillovers concretely in today’s context, consider two examples. The first concerns the relationship between Uber and Albert Einstein. Uber is a novel business model that has disrupted the transportation sector, and to the user Uber might appear as a simple mobile app enabling a new business idea. But Uber relies on a string of prior scientific achievements. Among them is GPS technology, embedded in the smartphone and in satellites overhead, which allows the driver and rider to match and meet. The GPS system in turn works by comparing extremely accurate time signals from atomic clocks on the satellites. But because the satellites are moving at high velocity compared to app users and experience less gravity, time is ticking at a different speed on the satellites, according to Einstein’s mind-bending theories of special and general relativity. In practice, the atomic clocks are adjusted according to Einstein’s equations, before the satellite is launched, to account exactly for these relativistic effects. Without these corrections, the system would not work. There is thus a series of intertemporal spillovers from Einstein to the GPS system to the smartphone to Uber (not to mention all the other innovations, mobile applications, and new businesses that rely on GPS technology).

As another example, consider the modern biotechnology industry and its many applications—genetic testing, cancer diagnosis, gene-based drug development, paternity tests, criminal forensics, testing for COVID-19, etc.—that depend on the analysis of DNA. To study DNA, it must first be replicated into measurable quantities, and this replication process depends on many prior scientific advances. One critical if unexpected advance occurred in 1969, when two University of Indiana biologists, Thomas Brock and Hudson Freeze, were exploring hot springs in Yellowstone National Park. Brock and Freeze were asking a simple question: can life exist in such hot environments? They discovered a bacterium that not only survived but thrived—a so-called extremophile organism—which they named *Thermus aquaticus*. Like Einstein's work on relativity, this type of scientific inquiry was motivated by a desire for a deeper understanding of nature, and it had no obvious or immediate application. However, in the 1980s, Kary Mullis at the Cetus Corporation was searching for an enzyme that could efficiently replicate human DNA. Such replication faces a deep challenge: it needs to be conducted at high heat, where the DNA unwinds and can be copied, but at high heat replication enzymes do not hold together. Mullis, in a Eureka moment, recalled the story of *Thermus aquaticus*, knowing that this little bacterium must be able to replicate its DNA at high heat given its environment. And indeed, *Thermus aquaticus* turned out to provide what was needed. Its replication enzyme was declared by *Science Magazine* to be the "molecule of the year" in 1989. Mullis would be awarded a Nobel Prize soon after, and the biotechnology industry would boom, opening new chapters of human progress.

These examples highlight several features that we will return to later with systematic evidence. First, we see essential roles that science can play in enabling marketplace innovations. Second, we see that the spillovers from science can be highly unpredictable. Finally, we see a key limitation of market-based investment incentives in the context of new ideas. Namely, the market value of Einstein's insights or Brock and Freeze's discoveries are essentially zero—there is no marketable product or service that they directly provide, and markets not surprisingly provided no funding for their research. Yet their discoveries form foundations for entire industries. Even when there is a marketable product or service, such as Mullis's DNA replication approach, the imitative and intertemporal spillovers that follow suggest that the private returns captured by the initial innovator can be much lower than the social value created.[1]

1 Indeed, Kary Mullis and the Cetus Corporation would receive a tiny sliver of the social value enabled by their advance.

In a modern context, economists recognize Jefferson's candle, where the light of one candle becomes the light of many, as defining an aptly named "public good."[2] In general, public policy can play key roles in the provision of such goods. In the context of idea production, policy interventions take many forms—government-sponsored research funding, intellectual property systems, research and development (R&D) tax credits, prizes, public research contracts, demand-side "pull" mechanisms like advanced purchase commitments, and others. All of these approaches seek to encourage the advance of ideas, recognizing the high social returns that may greatly exceed the private returns. In each case, these policies attempt to repair relatively weak incentives in markets to produce new ideas, and bring greater resources to these efforts, in line with the social returns.

2.b. Science and innovation as a stumble in the dark

While the value of an effective new idea, once it is in hand, may be high, a different perspective emphasizes how hard it is to light the first candle. Discovering an important new insight about nature or creating a valuable new product or service is difficult, and investments in science and innovation by nature have unclear prospects. They are steps into the unknown, with results that are fundamentally uncertain (Arrow 1962). The image of light spreading from one candle to another happens later in the process. The actual, up-front activity of science and innovation is more like a stumble in the dark, searching for a light that may or may not be there.

This fundamental uncertainty means not only that the right direction for investment is not obvious, but also that failure is common. Well-intended investments fail to produce value, and experts often make incorrect bets. In science, many research projects are abandoned and those seen through have widely varying impact (de Solla Price 1965; Yin et al. 2019). Beyond individual projects, larger streams of research can fail. The same scientist can see great success in one agenda and little or no success with another (Liu et al. 2018). Even Nobel Prize–winning researchers regularly produce failed work streams.[3]

With intertemporal spillovers, judging success and failure is even more difficult. Even when the idea is in hand, there is enormous uncertainty about its future prospects, and eventual success is often preceded by apparent failure. For example, although the science has been advancing since the 1990s, mRNA-based medication had faced a litany of failures—for cancer treatment, heart disease, kidney disease, and other

2 These are goods with two features: first, many people may benefit from it without impinging each other's use; and second, excluding people from its benefit is either difficult or undesirable. National security, public parks, and clean air are other examples of public goods.

3 See https://www.nobelprize.org/failure/ for perspectives on failure from Nobel Prize winners themselves.

areas. At the start of 2020, no mRNA-based vaccine or drug treatment had ever been approved for use in humans. Yet mRNA vaccines proved extremely effective against COVID-19 and are now seen as a breakthrough in treating infectious disease, with renewed prospects for other diseases.[4] Scientists refer to specific ideas that are initially underappreciated as "sleeping beauties" (Ke et al. 2015), and sequential failures are often part of an iterative learning process that leads to eventual success (Yin et al. 2019).

Uncertainties and regular failure are not just common in basic research. They are common in marketplace innovation, too. In the pharmaceutical industry, a survey of the top 10 pharmaceutical firms found that only one in nine new compounds that reached human testing were ultimately approved for use (Kola and Landis 2004). In other words, leading pharmaceutical firms fail the vast majority of the time. Venture capitalists also fail. Consider Bessemer Venture Partners, a prominent and successful venture capital firm. In an exercise of public humility, Bessemer maintains an "anti-portfolio" on its website, noting all the new ventures that it reviewed and decided *not* to invest in. These missed early opportunities include Apple, Airbnb, Facebook, FedEx, Google, Intel, and Zoom, to name a few. In a study of another venture capital firm, researchers examined the return on each investment made to its prospects as initially judged by the venture firm's partners (Kerr et al. 2014). These are private sector investors, investing their own money and making their best bets. Yet the partners had essentially zero predictive success across the portfolio of their investments. Ultimately, it appears that in science and innovation, nobody has a crystal ball.

While the inherent uncertainty in science and innovation investments means that they inevitably produce many disappointments, the fact of regular failure can also breed doubts about the benefits of these investments more generally. If success is rare, and failure common, the social returns imagined from the "public goods" perspectives may be heavily reduced. One form of skepticism may then simply be that science and innovation success stories are relatively few and that science and innovation is a poor investment overall.

Other forms of skepticism focus on the allocation of research funding. Most science funding, especially in basic research, comes from the federal government (*i.e.*, from taxpayers), and observers have questioned the capacity of government officials to identify and invest in good opportunities. C.P. Snow famously suggested a cultural

4 As another high-profile example, artificial intelligence research also had a long history of failures before recent breakthroughs. Machine learning and neural networks methods, which developed in fits and starts over many decades, were for long periods seen as unpromising (Minsky and Papert 1969; Wooldridge 2021). But these methods are now driving innovation across the U.S. economy and the world and are the subject of increasingly intense international competition.

disconnect between scientists and policymakers that disrupts good decision-making with regard to science (Snow 1959). The resulting view, and concern, is that the public funding of science and technology is not allocated in line with the public interest. In the U.S. government, Senator Proxmire's Golden Fleece Awards regularly called out questionable lines of publicly supported research (Hatfield 2006). More recently, Solyndra has been held up as an example of poor public investment choices in the more applied, marketplace context. The Nobel Prize–winning economist Milton Friedman once argued that the government was likely to make poor R&D investment choices and suggested that perhaps R&D investments should be left to the private sector (Kealey 2013).

Skepticism can also focus on scientists and experts themselves. Scientists and researchers are often depicted as living in an "ivory tower" (especially in universities), disengaged from the real world and a practical understanding of the world's problems. Amidst rising skepticism about experts in general (*e.g.,* Nichols 2017), scientists and their priorities can be viewed with doubt, and the fact that their projects and ideas regularly fail can fuel the sense that their expertise is not especially useful. Meanwhile, the public readily sees examples of very young individuals—with little initial experience or advanced education—starting companies that bring transformative innovations to the economy. Examples include a young Steve Jobs, Bill Gates, and Mark Zuckerberg. Amidst regular examples of failure, the tension between the seemingly remote world of scientists and technology researchers and the readily apparent success of young innovators can breed skepticism about the value of deep expertise and scientists themselves.

Ultimately, the fundamental uncertainty in science and innovation and the related regularity of failure engenders several forms of skepticism: about the overall returns to science and innovation investments; about the capacity of the public sector to allocate research dollars; and about who actually drives breakthroughs and the value of experts themselves. In these more skeptical perspectives, the advance of science and technology might still be seen as a public good, but if public agencies, universities, or scientists themselves are poor at investment in practice, perhaps the social returns that public policy aims for are not actually realized. And while anecdotes can be marshaled on all sides of these debates, they cannot be settled with stories. Assessing these perspectives requires systematic evidence and data. What is actually happening on average? Are the social returns to science and innovation investment high or low in practice? Is public research funding, and its allocation across fields, aligned with the public interest? Who drives the progress of science and marketplace innovation, and where do the big breakthroughs come from? The following section addresses these questions.

3. Science and innovation in practice: what the evidence says

This section collects systematic evidence, including the most recent evidence, regarding the value and operation of the science and innovation system in practice. Much recent work has been enabled by the methodological advances, as well as the revolution of "big data," which produces comprehensive views. These studies strengthen the empirical foundations for assessing the science and innovation system, and while there are still many gaps in our understanding, a number of striking facts and important insights have emerged.

This section focuses on three specific questions. First, is the United States overinvested or underinvested in science and innovation? Second, are public science investments allocated in a way that is commensurate with the public interest? Third, who drives breakthroughs in science and innovation? Answering these questions is central to policy questions of whether, how, and how much the United States could successfully scale the science and innovation system.

3.a. The social returns to R&D

The question of whether to invest more in science and innovation is essentially a question of estimating the social returns to these investments. If the social returns are high, meaning that the benefits are large compared to the costs, then additional investment will be worthwhile. One way to measure this is a "social benefit–cost ratio," which calculates how many dollars of benefit society receives per dollar of investment cost. If the benefit-cost ratio exceeds 1, so that $1 of investment cost returns more than $1 of social benefit, then innovation investments are worthwhile and, from society's point of view, more than pay for themselves. An alternative calculation is a rate of return measure, in percentage terms per annum, which can then be compared to rates of return per annum on other investments (e.g., stock market returns or other benchmarks). Researchers have studied many industries, and used many methods, to ascertain the social return to science and innovation investments. The outcome measure is usually the increase in value-added output or productivity in an industry and the cost is usually the expenditure on R&D. In studying *social* returns, researchers are working to find not just the value of the R&D investment to the investing party, but also the additional benefits or costs to other parties. The headline of these studies is that, while estimates vary, the social returns to investment in R&D tend to be remarkably large, and much larger than the private returns to R&D and to ordinary private investment returns in other contexts (Griliches 1958; Mansfield 1977; Hall et al. 2010). See Table 1. For example, reviewing hundreds of studies on agricultural R&D, Alston et al. (2000) and Evenson (2001) find that median social rates of return estimates are over 40%, an investment return that is many multiples of stock market returns or the interest rate on government bonds. Similarly, a review by Hall et al. (2010) examines dozens of studies of manufacturing and other industries and finds similarly large median social rates of return.

Table 1: The social returns to R&D

Study	Industry / Context		Social Rate of Return	Social Benefit-Cost Ratio
Alston et al. (2000)	Agriculture (review of 292 studies)		44% (median)	--
Mansfield et al. (1977) and Tewksbury et al. (1980)	Industrial Innovations (37 case studies)		71% (median)	--
Bloom et al. 2013	Publicly-traded firms, All industries		55%	--
Azoulay et al. 2019	Biomedical research from the NIH		--	> 3
Jones and Summers (2020)	Overall U.S. Economy	Baseline estimate Conservative estimate	67% 20%	13.3 5

Notes: This table summarizes estimates of the social return to R&D investment. The social benefit-cost ratio conveys the number of dollars in benefit per dollar invested, where a ratio greater than 1 indicates that the investment pays back more than it costs. The social rate of return can be compared to standard private rates of return, as a percentage gain per year. See also Hall et al. (2010) for a review of methodologies and results. Overall, using many methods, industries, and research contexts, the social returns to R&D appear extremely high, pointing to enormous un-reaped rewards from further R&D investment.

Despite this tendency to find high social returns, some doubts have remained about these calculations, for three reasons. First, what is true for the industry, technology area, or time frame studied may not be generalizable. One may be concerned that studies are often "picking winners," focusing on technology areas that we know have advanced successfully and thus may not be representative of overall returns. Second, the causal linkage between R&D investment and the following output or productivity gains can be difficult to establish. Third, spillovers are messy. It is very difficult to trace the imitative or intertemporal benefits from a given advance, and some spillover effects may impose costs on other parties, not benefits.[5] Studies

5 For example, the spillover benefits of widely used advances like electricity, computers, the Internet, or the Human Genome Project are difficult to enumerate and assess. The spillovers from such "general purpose technologies" would appear to be extremely positive. On the other hand, private R&D returns in a business context can also exceed the social return through "business stealing," where private investors do well in part by reallocating business from other firms to themselves. For example, if Amazon earns income selling books online, it succeeds in part at the expense of existing bookstores, and here the private return to Amazon investors may (on this dimension) exceed the return to society as a whole.

might either over-attribute broader benefits to a given innovation or perhaps fail to account for the spillovers in a complete way. Recent methodological advances have led to new insights that confront these challenges explicitly, and here we consider three recent studies that make important headway—and once again find extremely high returns.

The first study, Bloom et al. (2013), examines industrial R&D. The analysis is particularly focused on isolating the causal impact of R&D and estimating its spillovers within related industries. Methodological advances in this study are both a causal research design, based on how businesses respond to changes in federal and state R&D tax incentives, and an analysis that distinguishes between potential positive and negative spillovers among industry participants.[6] Netting out the spillovers, the analysis finds that industrial R&D has a social rate of return of 55%, which is several times the private return experienced by the investing firm. The findings imply not only enormous social benefits to industrial R&D but also that private R&D investment is very low compared to its benefits for society.

The second study, Azoulay et al. (2019), examines scientific investments in biomedicine. The analysis is notable for its focus on the linkage between upstream basic research and downstream marketplace application, and in isolating causal impacts. The authors use features of the National Institutes of Health (NIH) funding system to isolate quasi-random changes in the funding for particular biomedical science areas. They then trace the effect of this marginal funding on new scientific research and on later inventions that build on this research. The central finding is that $10 million in additional NIH funding in a given area leads eventually to an additional three private sector patents, including novel drugs. Looking purely at the private returns to these patents suggests that private value of the patents greatly exceeds the expenditure by the NIH. This paper further demonstrates the unexpected ways that science propels technological progress, as the additional NIH funding, directed at a particular disease area, is often taken up in patents targeting other applications.

The third study, Jones and Summers (2020), examines the social returns across the U.S. economy. Whereas most studies examine the returns to R&D investment in particular sectors, this study takes a broader and longer-run view. In particular, this study calculates an overall return to science and innovation investment in the United States, both by examining total R&D spending on the cost side and total, valued-added output gains on the benefit side. By looking at all R&D spending

6 The potentially negative spillovers are from business stealing, where the innovating firm takes business from product market rivals, while the potentially positive spillovers are on businesses that may not compete with the focal business but build on related technologies.

the method accounts for the costs of both successes and failures. By looking at the overall growth path of the economy, the method can incorporate and net out spillovers with an unusually broad view.[7]

Although the Jones and Summers methodology is quite different from other approaches, it once again points to very high social returns. In the baseline estimates, the social rate of return to R&D expenditure in the U.S. economy appears to exceed 50%. Put in perhaps more salient terms, the analysis indicates that $1 invested in innovation produces, conservatively, at least $5 in social benefits on average—and quite possibly $10 or even $20 in social benefits per $1 spent.

In sum, a consistent picture of high returns emerges from these studies. This is true not only across numerous settings, but also across a wide range of methods, including new studies that use increasingly sophisticated and diverse estimation approaches. Notably, these social returns are not just good: They are enormous. Effectively, the science and innovation system is akin to having a machine where society can put in $1 and get back $5 or more. If any business or household had such a machine, they would use it all the time. But this machine is society's. The gains from investment largely accrue to others—not so much to the specific person who puts the dollar into the machine. This brings us back to Jefferson's candle and the public goods nature of innovation. The spreading light of new ideas brings large benefits and pays for its costs many times over, but these exceptional benefits are not captured by the private investor. Public policy thus has essential roles to play in elevating these investments and realizing the returns.

The pandemic provides an additional and salient example of the high social returns to science and innovation investments. Operation Warp Speed sought to accelerate the invention, manufacture, and delivery of novel vaccines, with the goal of overcoming the enormous public health and economic consequences that the pandemic has imposed. This public investment cost approximately $25 billion (Gross and Sampat 2021), and it is not difficult to see that this cost appears very small compared to the benefits vaccines have brought in helping solve the pandemic, whether the benefits are measured in lives saved or in the rekindling of economic activity.[8] Indeed, if all

---

7 For example, the method incorporates the impact of science as well as general-purpose technologies, from the Internet to smartphones, where the spillovers are difficult to catalogue and trace. More generally, by looking at net gains in value added, it encompasses imitative spillovers, business stealing spillovers, and other positive and negative impacts of the advance of new ideas.

8 As vaccines entered use in December 2020, COVID cases in the United States were rising past 200,000 per day and deaths were rising past 3,000 per day. Using "value of a statistical life" measures for the United States of approximately $7-13 million in current dollars (Bosworth et al. 2017), the loss of life in one day (and in just the U.S.) would be valued at or above the entire cost of Operation Warp Speed. Meanwhile, the United States was down 10 million jobs in December 2020 compared to February 2020, and GDP was at least 4% below trend, which equates to several billion dollars lost per day. And this daily GDP loss comes on top of several trillion dollars of government expenditure to stabilize the economy. Indeed, the expenditure on Operation Warp Speed was also tiny (less than 1%) compared to the $3 trillion the U.S. government has spent in pandemic relief through March 2021 (Gross and Sampat 2021).

Operation Warp Speed did was to bring the end of the pandemic one day forward in time, then it easily paid for itself (Azoulay and Jones 2020).

The enormous social return to R&D investments raises a simple question: Why don't we spend more? A striking feature of Operation Warp Speed, at less than 1% of U.S. government expenditure on the pandemic, or of overall U.S R&D expenditure, at 2.8% of GDP, is that we devote a very small share of our resources to these endeavors. Society has a machine that pays back far more than we put into it, yet we put few dollars into the machine. We will return to these issues when discussing policy opportunities in Section 4.

3.b. The public use and funding of science

Even after acknowledging the high returns to R&D as a whole, one may still be doubtful about the role of science investments in this system. These doubts are especially relevant from a policy perspective because the government is a lead funder of scientific research. On the one hand, the logic of public goods and conceptual case for public investment is especially powerful for science, where the immediate marketplace value of new understandings of nature, on their own, may be very low, and therefore are especially unlikely to be provided by the private sector. Yet, as discussed in Section 2, a skeptic may wonder whether scientific research in practice tends to be useful. Perhaps most of scientific research provides no spillovers to support valuable applications. Perhaps the government makes bad investment choices. Perhaps scientists themselves are isolated from practical problems, operating in communities that tend to serve their peculiar and remote curiosities. These issues would all undermine the public case for investing in science specifically, even where the average returns to R&D on the whole are high.

One answer is the study by Azoulay et al. (2019), discussed above, which finds high marketplace returns caused by additional research funding at the NIH. At the same time, that study has a narrow context, focusing on biomedicine and the NIH channel. To generalize, we consider here several "big data" analyses that study linkages between the entire corpus of scientific research, across all fields, and public use in multiple dimensions.

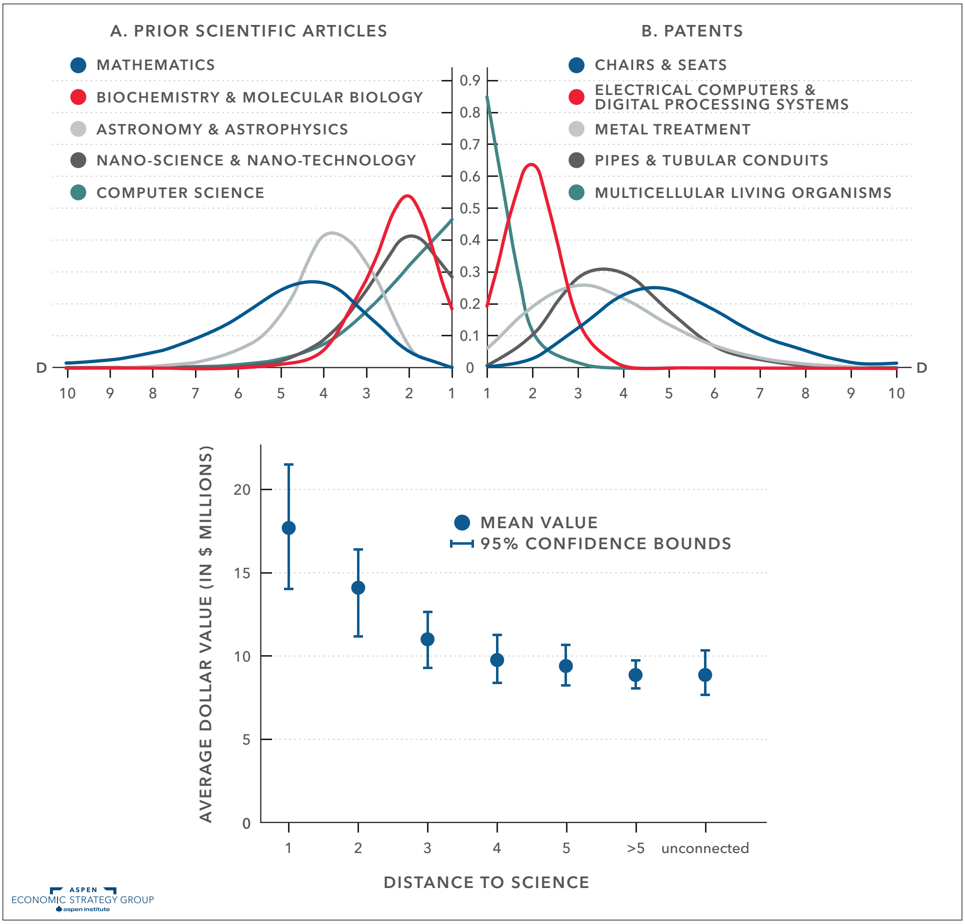
In a recent study, Ahmadpoor and Jones (2017) studied how U.S. patents build on prior scientific research, studying all U.S. patents since 1975 and tens of millions of scientific articles. The analysis investigates the connections between ideas, focusing on ideas that a given patent denotes as relevant prior art. This prior art can be prior patents but may also be scientific research articles.[9] Similarly, scientific research articles build on prior scientific articles. Using the references between documents, one can then trace knowledge flows within and between the domains of science and patenting, and study these flows across the entire landscape of research. Several facts and insights emerge. First, there is “majority connectivity” between the patenting and scientific domains. Conditional on a research article being cited at least once by other scientists, a large majority of scientific articles (79.7%) are part of a stream of knowledge that flows through to a specific future patent.[10] Second, the patents that draw directly on science are the most valuable patents.[11] In particular, these patents are the ones that are most heavily built upon by future inventions. Using similar big data, Watzinger and Schnitzer (2021) show that patents that directly draw on science have an average market value of $17.9 million, which is double the average market value of patents that are disconnected from science (see Figure 1). Finally, the data reveal the institutional sources of advances: In practice, universities and government laboratories produce the vast majority of the scientific articles that patents cite, and private sector businesses produce the vast majority of the patents that cite these articles. Overall, the flow of knowledge from publicly supported science into marketplace invention appears both highly valuable and remarkably widespread.

9 Studying prior art in patent documents has long been used in smaller samples to trace how one new idea builds on another within patenting and between science and patenting (*e.g.,* Carpenter and Narin 1983; Jaffe et al. 1993).

10 Patents directly cite science in research fields with applied orientations (*e.g.,* computer science, nanotechnology, and virology) but most of the connectivity is indirect, with these directly cited science advances building on other scientific advances, tracing back to increasingly basic science fields like mathematics and physics.

11 The patenting technology areas that are closest to science include areas such as biomedicine, artificial intelligence, and novel chemical compounds. Conversely, the patenting technology areas most distant from science (and with low market value) include inventions in things like cardboard boxes, ladders, envelopes, and chairs.

Figure 1: The use of scientific research in marketplace invention



Notes: Upper panel: Ahmadpoor and Jones (2017) consider the linkages between patents (right) and prior scientific articles (left), constructing a distant metric, D. The distance D=1 indicates a patent that directly references a scientific article. The distance D=2 is a second-degree citation (a patent that cites a patent that cites a scientific article on the right, or a scientific article cited by a scientific article that is cited by a patent on the left), and so on for higher measures of D. In some technology areas, like electrical computers, patents are close to science, while in others, like chairs and seats, patents are distant from science. Similarly, some fields of science, like nanotechnology, tend to be close to patenting while more basic research fields, like mathematics, tend to be more distant. Lower panel: Watzinger et al. (2021) consider the market value of patents based on how close they are to science. Patents that directly build on scientific articles have twice the market value as those most distant from science.

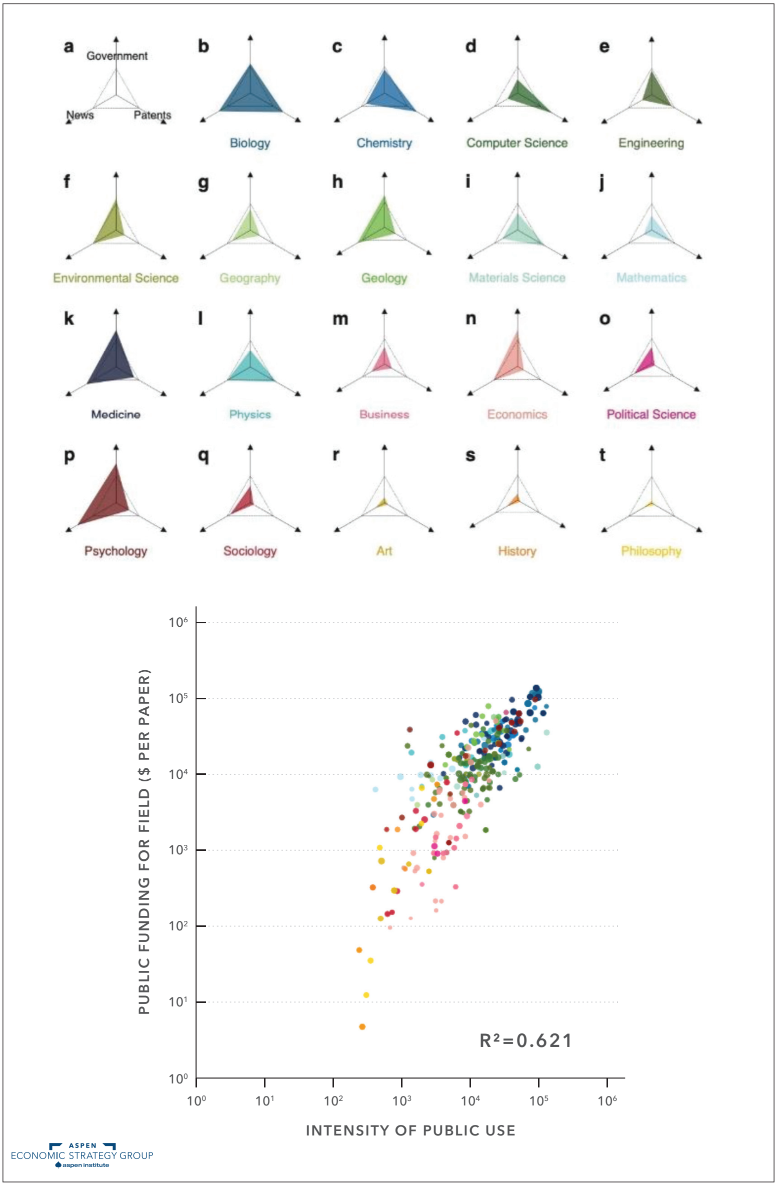
Further "big data" research has extended the study of science to other public uses. For example, in addition to supporting technological progress, research insights can support public policy and further engage basic human curiosity in the public at large. Specifically, Yin et al. (2021) further link the corpus of tens of millions of research articles not only to follow-on patenting but also to follow-on uses in U.S. federal government documents, across all federal agencies, and follow-on reporting in the general news media. What emerges is a diverse array of specialized use cases for different scientific and social scientific fields. For example, materials science research is used heavily in patents but is rarely referenced in government policy documents or in the news. Economics research, by contrast, is rarely referenced in patents but is regularly referenced in government and in the news. Government uses are very diverse and agency dependent,[12] while the news proves especially interested in human-centric subjects, such as psychology and medicine. Finally, Yin et al. further integrate funding information from major public sources.

This study allows insight on whether public funding, across hundreds of different research fields, is or is not allocated in line with public use of scientific research. What is especially striking is that a field's intensity of use in a given public domain—whether patents, policy, or news—strongly predicts public funding of that field. Pulling all three types of public use together, one can predict the public funding of different scientific fields with remarkable accuracy (see Figure 2).

The picture that emerges from these studies is not one of science and scientists being isolated from the public interest. Rather, science and social science have rich interfaces with public use, whether for marketplace invention, government policy, or general human interest. The science system appears metaphorically like a series of public parks. Many fields are like neighborhood parks—embedded in particular and often specialized communities of use. A few fields—like biomedicine—are more like a large national park, drawing in wide communities of public users and receiving proportionally more funding. Overall, these studies reject views that science is isolated from public use or funded in ways that don't track public interest. The widespread public use of science—and the value science brings—appear striking in the data.

12 For example, agencies like the Departments of Agriculture, Energy, and Transportation especially consume research in specific hard science and engineering areas related to their missions. The Department of Treasury especially consumes economics and business research, the Department of State draws heavily on political science research, and the Department of Defense is an unusual consumer of history.

Figure 2: The public use and public funding of scientific research



Notes: Upper Panel: When considering references to scientific articles from three public domains– patenting, government documents, and the news media–we see that different scientific fields are drawn upon in distinct and typically specialized ways. For example, computer science research is drawn upon directly and heavily in patenting, but less so in government policy documents or the news media. Psychology research, by contrast, is drawn upon especially by government agencies and in the news media, but much less so in patents.

Lower panel: Dividing the 19 top-level research fields into their 294 constituent subfields, we see that the intensity of public use is highly predictive of the public funding of the field. Source: Yin et al. (2021).

3.c. The people who drive breakthroughs

At the root of the science and innovation system are innovative people—the individuals who drive the advance of ideas. Understanding these "people inputs" is central to understanding the sources of advances and, consequently, to investing successfully in science and innovation. Who are these innovative people, and where do they come from?

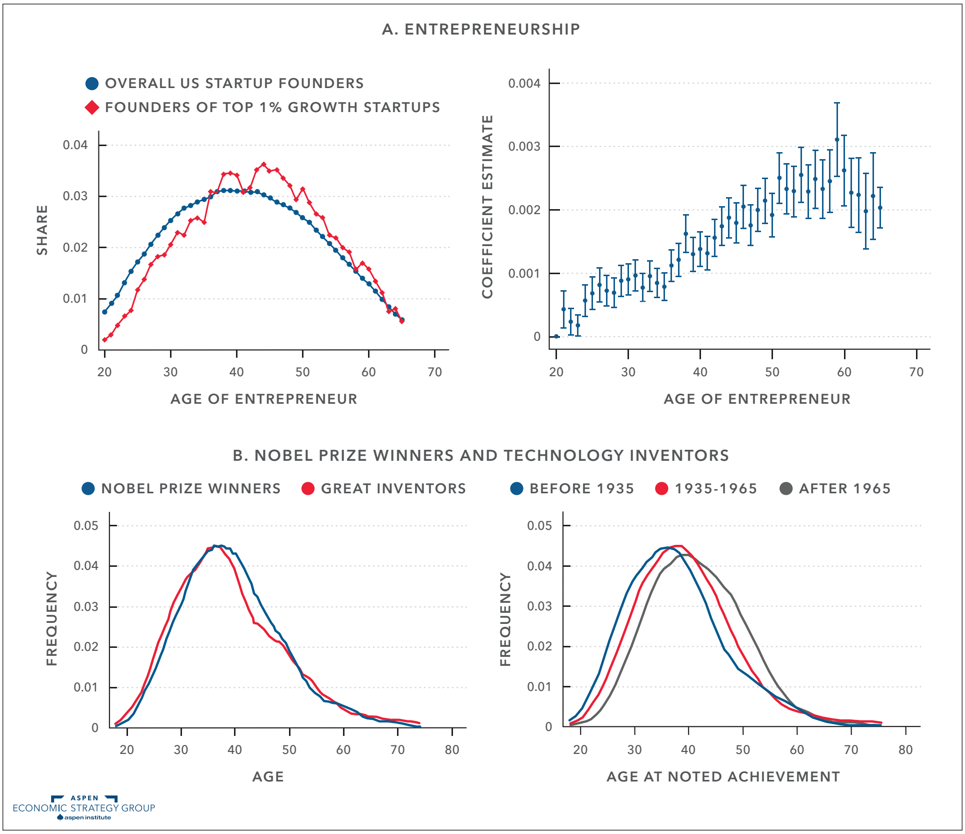
Public perceptions frequently suggest that very young people, often without substantial training, produce the big ideas. This view appears both in science and in marketplace innovation, and it is typically grounded in various viewpoints where younger people have greater levels of creativity, energy, and/or raw intelligence (Jones et al. 2014). The technologist and investor Paul Graham has said "the cutoff in investors' heads is 32 … after 32 they start to be skeptical" when discussing the right age for entrepreneurs, a view widely reflected in both the news media and venture capital behavior (Azoulay et al. 2019). And, in the sciences, people like Albert Einstein, Werner Heisenberg, and Paul Dirac made Nobel Prize winning contributions by the age of 25, suggesting the power of youth. Paul Dirac once opined, in a short poem, *"Age is of course a fever chill / that every physicist must fear / he's better dead than living still / when once he's past his 30th year"* (Jones 2010). These views have strong implications for the "people" part of the science and innovation system, including who should be hired and funded, and whether and how we can scale the relevant workforce.

Recent large-scale data studies have provided increasingly decisive insights on the demographic dimensions of scientists and innovators. First, consider new venture creation. Azoulay et al. (2019) used U.S. administrative data, including demographic information and tax records, to study every new business and every founder in the U.S. economy over the 2007–2014 period. They studied founder characteristics as well as the technology orientation of the business and its growth performance over ensuing years. Because this study considers millions of new businesses, it can focus not only on average outcomes but also on the very "upper tail" outcomes, including the 1 in 1,000 new businesses that saw the greatest sales or employment growth. The findings are striking: Rather than new venture success being the domain of founders in their 20s, or even their 30s, the upper tail successes came from individuals who start businesses at an average age of 45. Moreover, studying the employment histories of each founder, closer and longer work experience in the exact industry in which the new venture operates is extremely predictive of higher success rates. In other words, in contrast to the common ideas that (1) young people and (2) industry outsiders produce the exceptional successes, the reverse is true. Ultimately, age and relevant experience appear as signatures of success (see Figure 3A).

Turning to scientists and inventors, the major breakthroughs also tend to come in middle age. Studying all Nobel Prize winners and famous inventors over the 20th century, Jones (2010) finds not only that their signature breakthroughs tend to come in middle age, but also that they are coming at older ages with time (see Figure 3B). Today, one is more likely to produce a Nobel Prize–winning insight beyond age 55 than before age 30. Overall, in science, invention, and entrepreneurship, breakthroughs tend to come not from the young but from more seasoned individuals, deep in their domains.

In studying breakthroughs, one can also look more precisely at the role of expert knowledge. Here there is a key challenge that confronts science and innovation and is reshaping the "people" part of the science and innovation system. In particular, the very progress of science and technology means that there is more collective knowledge in each generation. This is one reason scientific advance is shifting away from breakthroughs by young people – who, in deepening areas, have more to learn before producing the next big steps (Jones 2010). But more generally this accumulation of knowledge across generations means that experts are increasingly specialized (Jones 2009). As Albert Einstein once said, "[K]nowledge has become vastly more profound in every department of science. But the assimilative power of the human intellect is and remains strictly limited. Hence it was inevitable that the activity of the individual investigator should be confined to a smaller and smaller section…" (Einstein 1949). Following Einstein's dictum, studies of the entire landscape of scientific research and patenting show exactly this: patterns of increasingly narrow expertise with time (Jones 2009; Jones 2011; Schweitzer and Brendel 2019; Hill et al. 2021).

Figure 3: Sources of scientific and innovative breakthroughs



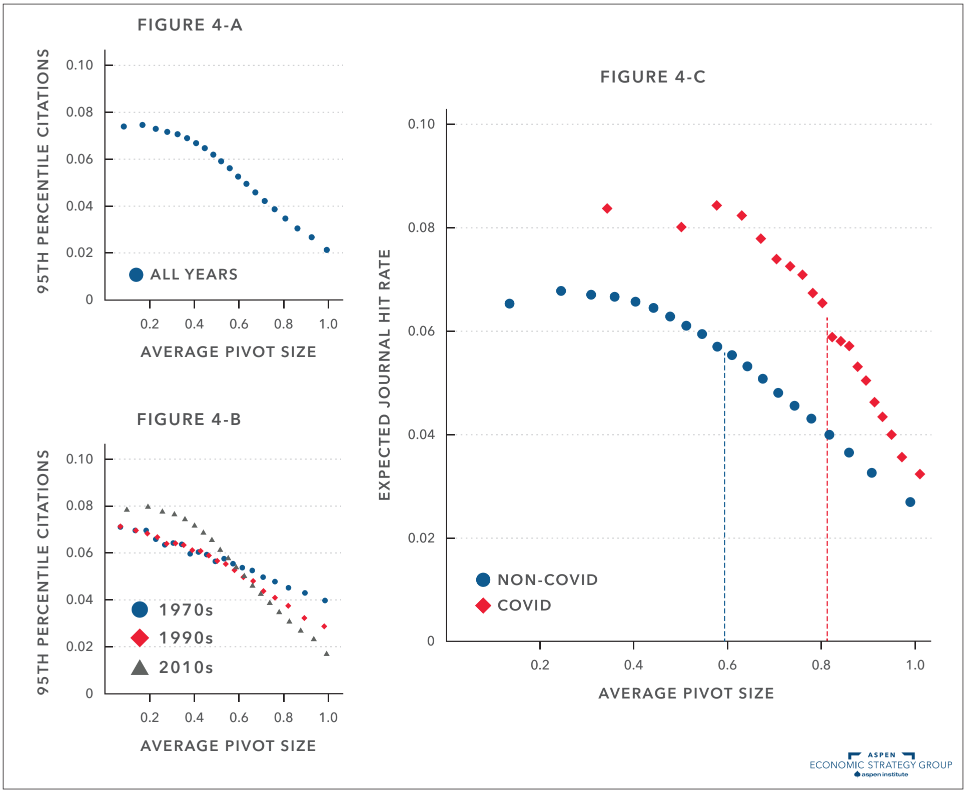
Notes: Contrary to common perceptions, science and innovation are not driven by young people with little domain expertise. **(3A)** The highest-growth, new ventures in the United States come from middle-aged founders (left). In fact, conditional on starting a firm, the probability that a founder has an upper tail success is increasing steadily and substantially with the founder's age (right). The data are all new ventures and founders in the United States from 2007–2014. Source: Azoulay et al. (2020). **(3B)** Similarly, the most notable science and technology breakthroughs come from individuals in middle age (left)–and from increasing ages with time (right). The data here are all Nobel Prize–winning contributions and the major inventions of the 20th century. Source: Jones (2010).

This narrowing of individual expertise has key implications for how we find breakthroughs. Namely, across all research areas, scientists and inventors increasingly work in teams, which act to aggregate expert knowledge and skills and allow researchers to attack problems more successfully (Mesmer-Magnus and DeChurch 2009; Lee et al. 2015). Critically, the highest impact science and the most valuable patents—whether from universities or private sector firms—increasingly come from larger teams (Wuchty et al. 2007). Today, the "people" part of scientific and technological progress has become not only a story of expertise, but a story of increasing specialization and collective expertise.

The narrowing of expertise also has implications for another key role of science and innovation: managing major crises. For example, Hill et al. (2021) study the response to COVID-19 across the entire landscape of scientific research. They show that an enormous range of scientists pivoted their research streams to engage the pandemic. In fact, nearly 6.3% of publishing scientists wrote a COVID-19 research article in 2020. However, the highest-impact COVID-19 research came, by far, from people who pivoted the least: Those who were already working on the very particular specialized topics that were closely positioned to engage COVID-19. This includes the University of Texas and NIH researchers who identified the COVID-19 spike protein as a key therapeutic target (Wrapp et al. 2020). Similarly, vaccines came not from outsiders but from specific scientists in private sector firms (such as Moderna and Pfizer-BioNTech) who were already specialized in the relevant mRNA platforms and could rapidly create solutions that targeted this spike protein. Stepping back from the pandemic and looking across all scientific research over the last five decades, researchers have become increasingly impactful when staying in their narrow domains and increasingly unable to make high-impact contributions outside their narrow domains (Hill et al. 2021). See Figure 4.

Ultimately, the picture of scientists, inventors, and entrepreneurs that emerges in these "big data" studies is one that emphasizes the importance of expertise—and the increasing importance of expertise. While science and innovation investments are probabilistic bets, and young and relatively inexperienced individuals can and do make large contributions, the weight of contributions increasingly come from older individuals with deep domain knowledge, and from specialists working in expert teams. A key implication is that critical resources of science and innovation depend on substantial human capital investments, which cannot be made overnight, but rather require effort and time to develop. The policy implications will be further considered in Section 4.

Figure 4: Expertise, specialization, and the pivot penalty



Notes: Scientists are increasingly specialized, and science increasingly relies on specialized domain experts to produce high-impact work. These figures examine this phenomenon by asking what happens when researchers work within or outside their domain expertise. In each panel, the pivot size measures how far a scientist is moving from their prior domain knowledge when writing a new article. The vertical axes are the probability of a high impact research article. Upper left: High impact work comes when a scientist stays close to their prior research expertise. Lower left: The penalty for moving away from one's specialized area is getting worse with time, indicative of an increasing advantage of domain expertise. Right: In 2020, enormous numbers of researchers pivoted to engage COVID-19, and this research experienced an impact premium given the critical demands of the pandemic. But even as scientists pivoted to help, the pivot penalty prevailed, with the highest impact work coming from individuals with relevant pre-existing domain expertise and who pivoted the least. The data are tens of millions of scientific articles across all of science, historically and today, and all COVID-19 research articles. Similar findings appear in patenting. Source: Hill et al. (2021).

4. Policy opportunities

We have now considered evidence, including recent studies and systematic evidence, to sharpen understanding of the science and innovation system in practice. In light of this evidence, we now ask how we can reshape public policy, emphasizing first-order policy dimensions to better engage the opportunities in science and innovation investment and meet the national interest.

4.a. The scale of investment

The United States appears to greatly under-invest in R&D. Studies, including the latest studies, find that the social benefits from these investments are extremely high. A central number from Jones and Summers (2020), which looks at the returns to science and innovation investment across the U.S. economy, suggests *conservatively* that $1 invested brings society $5 back on average. While one can debate specific numbers, the point that R&D investments bring extraordinary social returns appears highly robust.[13]

The question then for society is why we don't put more investment dollars into the science and innovation machine. R&D expenditure in the United States has averaged 2.8% of GDP over the past decade, representing a small share of economic activity. Even a 50% increase in total R&D expenditure, to 4.2% of GDP, would still call on a modest share of resources. Since some other countries already surpass such high R&D investment rates,[14] it seems practicable for a nation to invest substantially more in R&D. To understand why economies fall short, and leave such high-return investment opportunities untapped, we return to the public goods nature of innovation and the role public policy in putting additional dollars into the machine, whether to invest in basic research or to help encourage the private sector. As things currently stand, we appear to have a massive investment failure. Society has this incredible machine to raise standards of living, health, and worker productivity, yet we collectively fail to engage the machine to an extent commensurate with the benefit it appears to deliver.

That the United States doesn't invest more is even more striking in light of recent and evolving challenges. The United States has faced a slowdown in productivity growth and rising concern about the international competitiveness of the U.S. workforce over many years (*e.g.,* Gordon 2012; Autor et al. 2013), where real wages for the median household have struggled to rise and failed to keep pace with the gains in prior generations (*e.g.,* Council of Economic Advisers 2011; Autor et al. 2006). Yet, even as productivity has lagged, U.S. R&D intensity has slipped compared to other countries. In the mid-1990s, the United States was in the top five of countries globally in both total R&D expenditure as a share of GDP and public R&D expenditure as a share of GDP (Hourihan 2020). Today, the United States ranks 10th and 14th in these metrics, and U.S. public expenditure on R&D as a share of GDP is now at the lowest

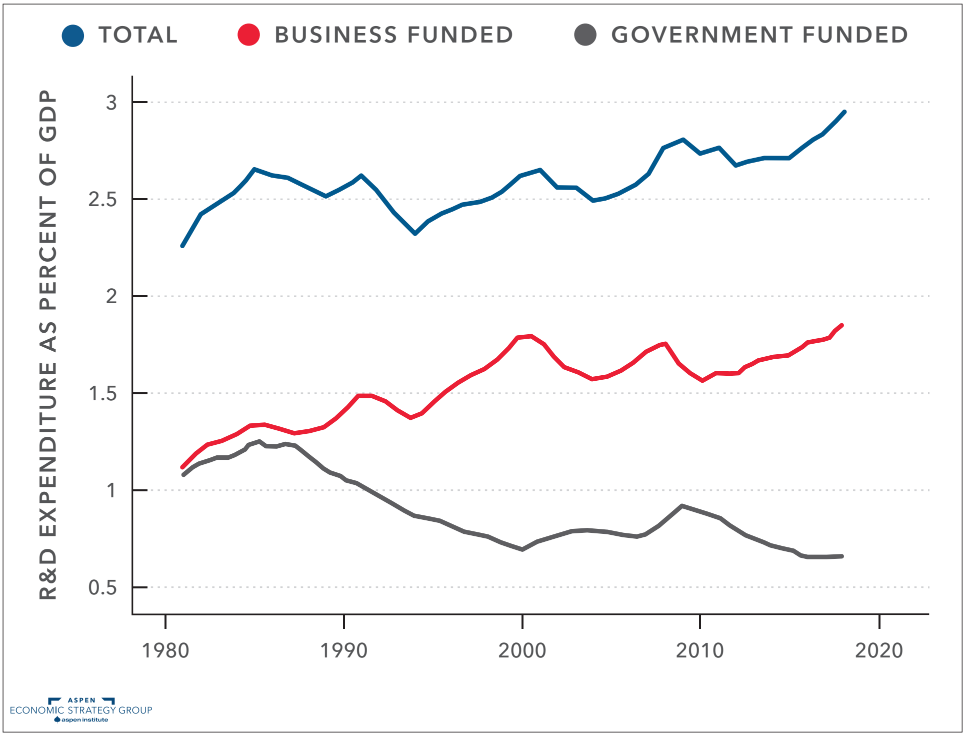
---

13 A related question concerns the optimal level of R&D investment. With more investment, we eventually hit "diminishing returns," where the value of additional R&D investment will decline. But we appear to be very far from that point now. For example, studies like Azoulay et al. (2019) and Bloom et al. (2013) show directly that additional investment in R&D produces enormous social returns on the margin. Jones and Williams (1998) suggest that optimal R&D investment levels in the United States should be, conservatively, two to four times higher than actual investment. For additional discussion, see Jones and Summers (2020).

14 For example, South Korea (4.6%) and Israel (4.9%) greatly exceed even an ambitious target like 4%, and many leading economies now substantially exceed U.S. R&D intensity, including Japan (3.2%) and Germany (3.2%).

level in nearly 70 years. See Figure 5 for U.S. trends. By contrast, China has massively increased its science and innovation investments in pursuit of leading the world economically and strengthening its hand in global affairs. China's R&D expenditure has grown 16% annually since the year 2000, compared to 3% annually in the United States. If China implements its current five-year plan, it will soon exceed the United States in total R&D expenditure.

Figure 5: U.S. R&D spending over time



Notes: Trends in U.S. R&D expenditure as a share of total U.S. GDP. Government R&D investment has been in a long decline. Federally funded R&D support as a share of GDP is now approximately half its level from the 1980s and more generally is at its lowest level in nearly 70 years. Source: OECD Main Science and Technology Indicators (2021).

Partly in response to these patterns, the U.S. Senate passed bipartisan legislation in June 2021 that authorizes expansions of public R&D investment. Within the provisions of the U.S. Innovation and Competition Act, there are authorizations to increase public R&D expenditure by approximately $90 billion, spent over five years, with the additional investment flowing primarily through the National Science Foundation, Department of Energy, and Defense Advanced Research Projects Agency (DARPA). This legislation is moving to seize the social returns to greater R&D

investment. It promises to increase the productivity and competitiveness of U.S. businesses and the U.S. workforce, and it is being promoted by policymakers with a competitiveness orientation.

An observation about U.S. science and innovation policy is that policymakers appear to go big (or, at least, go bigger) when perceiving specific threats. This was the case in World War II. It was the case with Sputnik and the Apollo program. It was the case with COVID-19 and Operation Warp Speed. And it is the case currently with China and developing legislation like the U.S. Innovation and Competition Act. The past efforts have tended to produce notable and world-leading advances—from radar and jet engines, to walking on the moon, to effective vaccines (Gross and Sampat 2021). What is an open question about our political economy is why we don't go even bigger, and do it all the time. For example, the U.S. Innovation and Competition Act is moving substantively in the right direction but still envisions only a modest increase in R&D intensity, raising the R&D share of GDP by about 0.1 percentage points. This is modest compared to what is already achieved in some other countries, and it is modest compared to the rising expenditure in China. Most importantly, it is modest compared to the gains that are on offer.

Looking purely at the social returns, the standard findings suggest that doubling the total investment in R&D would easily pay for itself (Jones and Williams 1998; Bloom et al. 2013; Jones and Summers 2020). That is, the additional expansion in standards of living in terms of GDP per person would be much larger in present value than the additional investment cost. How much potential is the United States leaving on the table? Using the general approach in Jones and Summers (2020), a sustained doubling of all forms of R&D expenditure in the U.S. economy could raise U.S. productivity and real per-capita income growth rates by an additional 0.5 percentage points per year over a long time horizon. This would lead to enormous increases in standards of living over time. It would greatly advance the competitiveness of U.S. businesses and workers and the overall position of the U.S. economy in the world. And this economic orientation leaves out the health gains of longer and healthier lives, which are among the most valuable deliverables from the science and innovation system (Cutler et al. 2006; Murphy and Topel 2007; Jones and Summers 2020).

4.b. The people pipeline

Successfully scaling up the science and innovation system, and achieving its many benefits, will rely on more than just increasing R&D expenditure. It also requires scaling the science and innovation workforce. These are the people who actually produce the breakthroughs, and systematic evidence about the people part of innovation (see Section 3.c) emphasizes that breakthroughs come from people with particular characteristics. While big ideas can come from many corners, they tend not to come from young people with little domain knowledge but rather from domain experts in middle age and beyond. The people who hit the "home runs" are typically individuals steeped in an industry when creating new ventures and typically specialized experts working in teams in both marketplace invention and in scientific research.

An immediate policy implication is that the people part of innovation cannot easily be extended overnight. Rather, expanding the science and innovation workforce requires investment to cultivate individuals with relevant training and talent. A number of recent studies, all utilizing high-scale data, further inform the sources and constraints in expanding this workforce. Here we consider the medium and longer-run opportunity through the U.S. educational system as well as the relatively rapid scaling opportunities through immigration policy.

4.b.1. Domestic investment in STEM workers

Recent studies have used comprehensive data to study the childhood backgrounds of inventors, including recent U.S. inventors (Bell et al. 2019), historic U.S. inventors (Akcigit et al. 2017), and inventors outside the United States (*e.g.,* Aghion et al. 2017). A striking finding is that inventors come from quite narrow parts of the overall population. Specifically, they tend to be male, they tend to be born in high-income households, and they tend to have been exposed to inventive careers as children. These studies and others further emphasize that there is enormous potential to expand entry into these career pathways—that is, the talent demonstrated at young ages is far wider than the set of people who enter these careers. In identifying career impediments, one also sees concrete opportunities to expand entry.

Consider for example mathematical ability demonstrated at young ages. Bell et al. (2019) study the 3rd grade test scores throughout New York City and observe the career pathways that eventually develop for these children. While very high math scores in 3rd grade are highly predictive of entry to invention later, this effect is much weaker among kids with equally high math scores if they come from lower-income households. Similarly, girls with extremely high math scores in 3rd grade are much less likely to enter invention later. At the same time, exposure to inventive career opportunities appears to be a powerful mechanism to encourage future entry. Studying the entire United States, children that grow up in neighborhoods with high invention rates are more likely to become inventors and will tend to patent in exactly the same technology area that they have been exposed to as children. Further, children who move to more inventive regions during childhood become far more likely to enter inventive careers. And girls who move to regions that are especially populated with female inventors become far more likely to become inventors themselves.

Altogether, these findings suggest two key things. First, there appear to be many "lost Einsteins" in the U.S. science and innovation landscape, where very talented kids miss out on these career opportunities. Talent does not appear to be a constraint on the U.S. capacity to scale science and innovation efforts. Second, among other potential educational interventions, exposure to innovation career pathways offers potentially low-cost, high-return policy approaches. Extending mentoring and social networks between the nation's existing inventive workforce and children from lower-income backgrounds, girls, other underrepresented groups, and those in neighborhoods with less inventive activity appear as large opportunities to expand pathways into the science and innovation system.

Stepping back, opening pathways into the STEM workforce would not only help propel standards of living, health improvements, and the U.S. position in the world, but it would also directly expand individual opportunity and reduce inequality. In particular, rising inequality in the United States over many decades is a story of increasing labor market and wage gains for highly educated workers, and a corresponding weakening job market for those with less education (*e.g.,* Goldin and Katz 2010). Sending more children into STEM careers will serve to reduce these wage gaps.[15] Cultivating untapped STEM talent among under-represented groups and in currently less-inventive areas, whether in cities or in rural areas, may have especially impactful job and wage effects. Thus, expanding the STEM workforce along these lines would appear as a win across many dimensions of society, not only accelerating standards of living gains and competitiveness but also help address inequality, including regional inequality, and structural labor market issues. From this perspective, a big push on developing the STEM workforce, could be a unifying, bipartisan policy step.

4.b.2. Immigration opportunities

The opportunities discussed above provide major pathways to expanding the people part of the innovation system. However, developing the talent pool for science and innovation through the education system, and especially early life-cycle efforts, will bear fruit relatively slowly. More rapid pathways are also available. In particular, a country can import talent (*i.e.,* through immigration). Recent, systematic studies of entrepreneurship and invention in the United States help inform this channel.

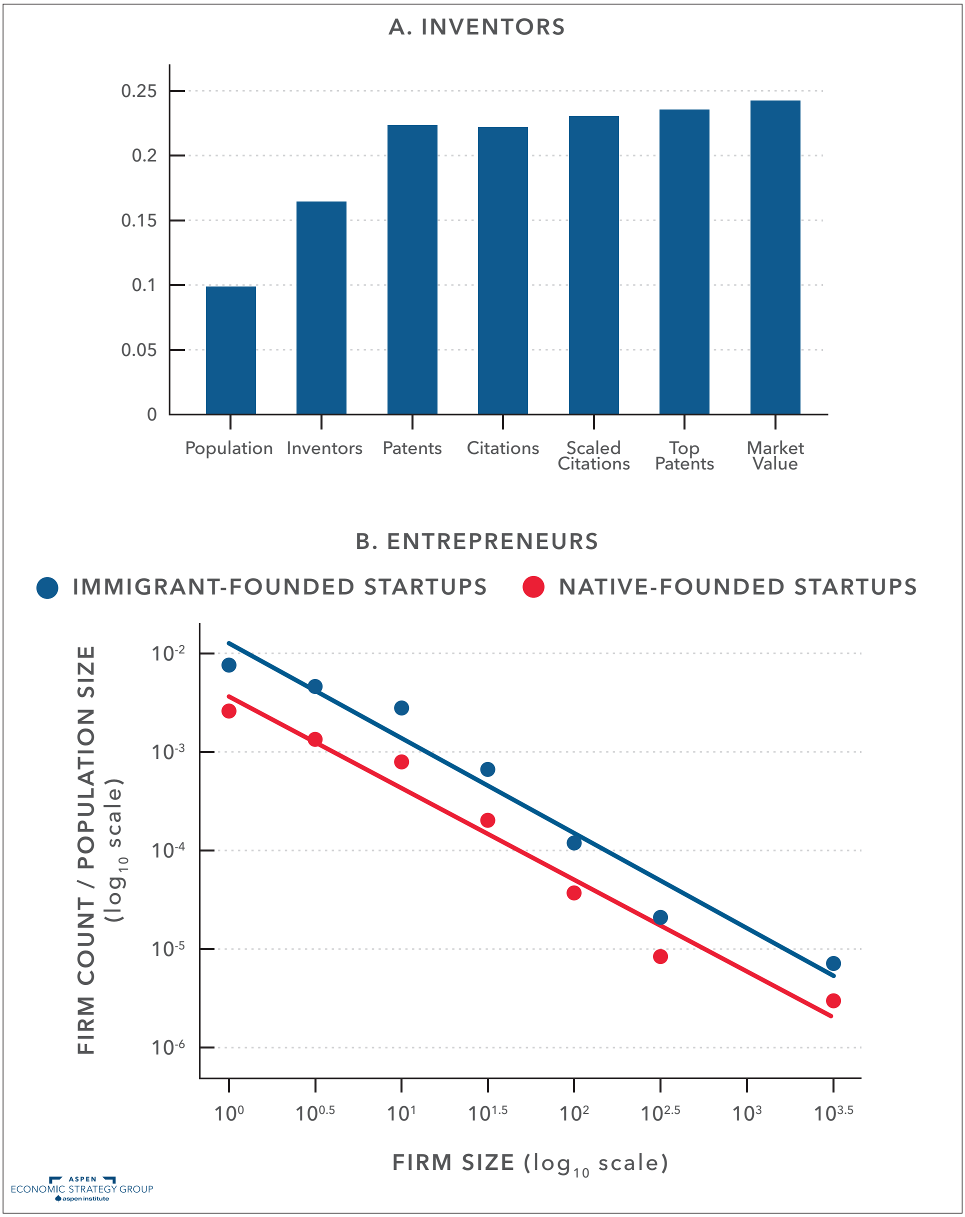
15 This is a point about supply and demand. Namely, labor force adjustment that makes highly trained STEM workers more abundant (and less-educated workers less abundant) will help those who remain less-educated: they will see more job opportunities per person and relative wage gains.

In a systematic study of inventors in the United States, Bernstein et al. (2019) examine the role of immigrants in U.S. invention. The central finding is that immigrants are especially productive in inventive activity. Not only do immigrants patent more often than U.S.-born individuals, but their patents are both more impactful for future invention and have greater market value. Overall, immigrants produce twice as many patents as one would expect from their population share. This is consistent more broadly with the STEM orientation of the immigrant workforce. While immigrants make up about 14% of the U.S. workforce, they account for 29% of the college-educated science and engineering workforce and 52% of science and engineering doctorates (Kerr and Kerr 2020). Overall, immigrants have accounted for about 30% of U.S. inventive activity since 1976 (Bernstein et al. 2019).

A similar picture emerges when examining entrepreneurship. Azoulay et al. (2021) study every new venture in the United States founded from 2007 through 2014 and examine whether the founders were born in the United States or abroad. They find that immigrants are 80% more likely to start a company than U.S.-born individuals. Moreover, immigrant founders are more likely to start companies of every size, including the highest-growth and most successful new businesses (see Figure 6).[16] Indeed, looking at Fortune 500 firms today and tracing them back to their founding roots, one similarly finds a disproportionate role of immigrant founders—from Alexander Graham Bell to Sergey Brin to Elon Musk. A remarkable finding here is that immigrant-founded firms employ more people in total than there are immigrants in the U.S. workforce.

16 Moreover, looking at the technology of these firms, and consistent with the patenting findings discussed above, immigrant-founded firms are also more likely to patent at all sizes.

Figure 6: Immigrants in the U.S. innovation system



Notes: Immigrants play especially productive roles in U.S. invention and entrepreneurship. **(6A)** While immigrants have represented 10% of the population since 1975, they have produced more than 20% of the patents, and an even greater share of the highest value patents. Source: Bernstein et al. (2019). **(6B)** Immigrants start new businesses at a rate 80% higher than U.S.-born individuals do. Moreover, immigrants start more firms of all sizes, including the most successful businesses that have the largest sales or employment. Source: Azoulay et al. (2020).

These recent, systematic studies show that immigrants are especially inventive and entrepreneurial. Moreover, the immigration channel may be a relatively rapid way to scale the people pipeline into the U.S. science and innovation system. Given that U.S. immigration policy currently constrains the entry of high-skill workers, there appears to be substantial further opportunity to rapidly expand the science and innovation workforce through immigration policy channels. Kerr and Kerr (2020) examine a range of policy options, including relatively small policy changes, that could make a difference to the innovation system along these lines.

4.c. The portfolio of investment

There are many specific directions of travel when thinking about the problems that we might scale R&D to solve—from Alzheimer's disease to violent crime to quantum computing to space travel. And there are many levers of public policy that can increase investment in science and innovation—from scaling basic research funding to expanding businesses' research and experimentation tax credit. When scaling the national investment in R&D, how should we think about the portfolio of investments? This final section considers these questions.

4.c.1. The importance of independent bets

In searching for as-yet undiscovered solutions, it is essential to remember a key feature of science and innovation investments: the regularity of failure (see Section 2.a). This inherent feature in creative search has important implications for the set of investments that are made as part of a successful R&D policy. First, we must embrace risk. That is, we must not only tolerate failure but embrace it in pursuit of opening new doors to progress. Second, we must engage a wide portfolio of bets. This approach can produce more efficient search, lower collective risk, and increase returns in the science and innovation system.

To illustrate this advantage, consider a search process to find a solution to a particular disease. Let's say that there are a number of pathways to try, but each has a low chance of success—say just 10%. Now let's say that we can make 10 investments in attempts to solve the disease. If all these investments try the same pathway, then the chance of producing a success is still only 10%. But if each investment tries a different pathway, each with an independent 10% chance of success, the collective probability of at least one success rises to 65%. By spreading out the bets, the chance of success is multiples higher, and for the same investment cost.

Public policy can play a key role in pushing for diverse pathways. And the U.S. government has taken this approach explicitly, particularly in crisis situations. For example, Operation Warp Speed explicitly chose to invest in four different vaccine platforms, with two vaccine candidates in each platform. The former director of Operation Warp Speed emphasized this diversification as the first principle of the policy design (Slaoui and Hepburn 2020). Similarly, in World War II, the U.S. Office of Scientific Research and Development (OSRD) was created to coordinate an enormous range of science and innovation investments that would help win the war. These efforts explicitly deployed a portfolio approach, engaging multiple pathways toward a given objective. The development of radar, sonar, high-scale antibiotics, early computing, and the atomic bomb were among the many rapid achievements of the OSRD's efforts (Gross and Sampat 2021).

The key lesson here is that science and innovation investments gain large advantages by spreading out along the frontier of opportunities. This can greatly accelerate progress, and not just in crisis but in ordinary times. But it's not clear that either the private sector or the public research institutions bet widely enough. Rather, we seem to crowd into particular areas. This can be true in the private sector, where businesses may duplicate others' R&D efforts as they compete for a market (Zeira 2011; Bryan and Lemus 2017). And it appears true in scientific research, too. For example, the Human Genome Project unveiled an enormous range of new pathways—*i.e.,* genes that encode proteins, the function of which we do not understand and may be key to advancing human health. Yet Edwards et al. (2011) describe "too many roads not taken," where 75% of protein research continues to focus on the 10% of genes that researchers already knew about prior to the Human Genome Project. In other words, scientists herd, too.[17] Research institutions thus need to focus on seeding and building new communities to further explore the roads not traveled amidst the vast unknown.

Ultimately the constraint on scientific progress is not the set of problems we would like to solve. Nor does it appear to be available pathways of discovery. Rather there is enormous opportunity to scale and diversify these efforts. In many ways, the vision of science and innovation needs to be the opposite of "picking winners." Rather, we need to "pick portfolios," with an emphasis on both increasing the scale of funding and human capital, and the diversity of approaches that are taken. The OSRD and Operation Warp Speed examples provide explicit institutional models, whereby public policy has appeared to play key roles in diversifying bets in effective manners.

17 Among the reasons that scientists herd into particular research areas is that they depend on having a relevant community of co-specialists around their work—scientists who collaborate, listen, evaluate, and collectively propel progress in very specific areas (Stoeger et al. 2018). As knowledge deepens, and scientists become more specialized (see Section 3.c), this need for community is likely only to intensify and becomes critical for advance. That is, science is a team sport, and scientist communities may have an effective minimum size. This will inhibit diversification of research pathways, and it links the capacity to diversify pathways of research to the overall scale of the scientific enterprise.

4.c.2. R&D policy levers and uncertainty

The range of mechanisms by which public policy works to expand R&D investment is large and complex. For basic and applied research, an array of federal government agencies solve the market failures by funding projects up-front. Lead investors include the Department of Defense, the NIH, Department of Energy, and the National Science Foundation. Their funding largely goes to a network of national laboratories and to research universities but can also work through private-sector research contracts.[18] Meanwhile, to increase market incentives for invention, the U.S. government supports the intellectual property system, including the U.S. Patent and Trademark Office. Other prominent, market-oriented policies include the R&E tax credit, which lowers innovation costs for private sector businesses, and the Small Business Innovation Research (SBIR) program, which helps fund R&D efforts by small technology businesses and new ventures. On education and workforce dimensions, policies that develop the STEM workforce, from early childhood through graduate school, as well as through immigration, further support the science and innovation system.

Examining this wide range of policy levers, one may ask which approaches are especially effective. And important progress has been made in evaluating specific policy approaches. For example, R&D tax credits for firms, the SBIR program, and NIH research funding appear quite effective at raising innovative investment and with high returns. Recent reviews of specific policy opportunities include Bloom et al. (2019) and Jones and Goolsbee (2021), which provide guides and assessments across wide arrays of policy areas.

At the same time, there remains much about R&D policy that we do not know. This is true especially in a comparative sense across different levers. For example, despite enormous progress in understanding science and innovation, we cannot yet credibly determine whether the investment returns are ultimately higher for basic research (say, in pure mathematics) compared to applied research (say, in nanotechnology materials) or how the social returns to upstream science investments compare to marketplace levers like the R&E tax credit. This puts policymakers in a seemingly uncertain position when assessing how to allocate budgetary resources across the science and innovation system. However, what we do know, and what this chapter has emphasized, is that the social returns to R&D investment overall are extremely high. And this point has strong implications for policy.

Consider again the social returns to R&D. Society has an R&D machine, where we put in $1 and receive at least $5 back. However, extending the picture, this machine

18 Federal agencies, and subcomponents of these agencies, use a wide array of funding models. See Azoulay and Li (2020) for an overview and discussion.

turns out to have a complex interface: There are many input slots in this machine, each able to take a dollar—one input slot for the NIH, one for DARPA, one for R&E tax credits, etc. Which slots should we put our dollars in? While there is substantial evidence that many of these input slots produce high returns, lingering uncertainty over which options are best may create pause, debate, and a failure to act. But paralysis would be a huge mistake. Yes, one might get even more dollars back if we knew better how to allocate investments across these slots. But the true failure is not to put more resources into the machine, because with what we already know—based on the allocation we already do—we are getting an enormous return.

Separately, policy uncertainties can be resolved with time, and explicit effort, through science itself. Much has been learned about how science and innovation operate, and where breakthroughs come from. Further advancing our understanding will depend on continued research effort, and the scientific toolkit is powerful here, from the expanding access to comprehensive data about scientists, inventors, entrepreneurs, and their funders, to the expanding set of empirical tools, which include experimental, network, and machine learning methods. Continuing research will sharpen our choices and promises to raise the social returns even further.

5. Conclusion

Science and innovation investments are central to the national interest. These investments can create higher standards of living, longer and healthier lives, and an increasingly competitive workforce. They can support national resilience in the face of crises, like the global pandemic, and they can sustain national leadership in the world, including on economic, political, and security dimensions. Given these potential benefits, this chapter has considered whether the United States invests enough in science and innovation, and specifically whether greater public support is warranted. We have asked several related questions: What are the arguments for or against a public role in the science and innovation system? What is the evidence? How are we doing? What policy changes do we need?

A primary case for public action sees new ideas – the fruits of science and innovation -- as "public goods" that the private sector will underprovide. Meanwhile, skeptical perspectives emphasize the regularity of failed R&D efforts, doubt the capacity for successful public investment, and question the role of science and domain experts in driving practical and important advances. After laying out these different perspectives and illuminating them with examples, the chapter turned to systematic evidence, including the very latest evidence. The conclusions from systematic evidence are clear. The social returns to R&D investments are enormous and greatly in excess of the private returns. Public investments in science appear closely aligned with public use, and domain experts are the primary drivers of breakthroughs in both science and marketplace innovation. In short, the U.S. science and innovation system as it stands delivers far more than its resource costs, and we underinvest in science and innovation to an enormous degree. For every $1 we invest, we conservatively receive $5 in benefit. Effectively, the public has at hand an extraordinary machine to benefit human progress and the national interest, yet we fail to use this machine anywhere close to its full capacity.

To meet the national interest, policy can adapt in first-order, high-return ways. This chapter has emphasized three poles of action to reap the rewards: (1) scaling funding resources; (2) scaling the people pipeline into science and innovation careers; and (3) making diverse investments across the landscape of opportunities. These investments promise to raise our standard of living, accelerate progress against disease, increase the competitiveness of the American workforce, solve for national and global crises, and secure the nation's leadership in the world.

References

Aghion, Philippe, Ufuk Akcigit, Ari Hyytinen, and Otto Toivanen. 2017. "The Social Origins of Inventors." National Bureau of Economic Research Working Paper #24110.

Ahmadpoor, M. and B.F. Jones. 2017. "The Dual Frontier: Patented Inventions and Prior Scientific Advance," *Science* 357, 583–587.

Akcigit, Ufuk, John Grigsby, and Tom Nicholas. 2017. "The Rise of American Ingenuity: Innovation and Inventors of the Golden Age." National Bureau of Economic Research Working Paper #23047. Cambridge, MA: National Bureau of Economic Research.

Alston, Julian M., Connie Chan-Kang, Michele C. Marra, Philip G. Pardey, and TJ Wyatt. (2000). *A Meta-Analysis of Rates of Return to Agricultural R&D, Ex Pede Herculem?* International Food Policy Research Institute.

Alston, Julian M. and Philip G. Pardey. 2021. "Innovation, Growth and Structural Change in American Agriculture," in *The Role of Innovation and Entrepreneurship in Economic Growth,* edited byAaron Chatterji, Josh Lerner, Scott Stern and Michael J. Andrews. University of Chicago Press.

Arrow, Kenneth. 1962. "Economic Welfare and the Allocation of Resources for Invention," in *The Rate and Direction of Inventive Activity: Economic and Social Factors*, 609-625. Princeton, NJ: Princeton University Press.

Autor, David H., Lawrence F. Katz and Melissa S. Kearney. 2006. "The Polarization Of The U.S. Labor Market." *American Economic Review* 96, 189–194.

Autor, David, David Dorn, and Gordon H. Hanson. 2013. "The China syndrome: Local labor market effects of import competition in the United States." *American Economic Review* 103, 2121–2168.

Azoulay, Pierre, Josh Graff Zivin, and Danielle Li. 2019. "Public R&D Investments and Private-sector Patenting: Evidence from NIH Funding Rules," *Review of Economic Studies* 86, 117–152.

Azoulay, Pierre, Benjamin F. Jones, J. Daniel Kim, and Javier Miranda. 2020. "Age and High-Growth Entrepreneurship," *American Economic Review: Insights* 2, 65–82.

____. 2021. "Immigration and Entrepreneurship in the United States," *American Economic Review: Insights*, forthcoming.

Azoulay, Pierre and Benjamin F. Jones. 2020. "Beat COVID-19 through Innovation," *Science* 368, 553.

Azoulay, Pierre and Danielle Li. 2020. "Scientific Grant Funding," in *Innovation and Public Policy,* edited by Austan Goolsbee and Benjamin F. Jones, University of Chicago Press.

Bell, Alexander, Raj Chetty, Xavier Jaravel, Neviana Petkova, and John Van Reenen. 2019a. "Who Becomes an Inventor in America? The Importance of Exposure to Innovation," *Quarterly Journal of Economics* 134 (2), 647–713.

Bernstein, Shai, Rebecca Diamond, Timothy McQuade, and Beatriz Pousada. 2019. "The Contribution of High-Skilled Immigrants to Innovation in the United States," Stanford University.

Bloom, Nicholas, Mark Schankerman, and John Van Reenen. 2013. "Identifying Technology Spillovers and Product Market Rivalry," *Econometrica* 81, 1347–1393.

Nicholas Bloom, John Van Reenen, and Heidi Williams. 2019. "A Toolkit of Policies to Promote Innovation," *Journal of Economic Perspectives* 33, 163–184.

Bosworth, Ryan C., Alecia Hunter, and Ahsan Kibria. 2017. "The Value of a Statistical Life: Economics and Politics." Strata. https://strata.org/pdf/2017/vsl-full-report.pdf.

Bryan, Kevin A. and Jorge Lemus. 2017. "The Direction of Innovation," *Journal of Economic Theory* 172, 247–272.

Bush, Vannevar. 1945. *Science, the Endless Frontier: A report to the President.* (Government Printing Office, 1945).

Center, Seth and Emma Bates. 2019. "Historical Perspectives on Innovation, Technology, and Strategic Competition," CSIS Brief.

CLAAS. n.d.. "Farmer sets two new grain maize harvest world records with CLAAS," CLAAS Group press release. https://www.claas-group.com/press-corporate-communications/press-releases/farmer-sets-two-new-grain-maize-harvest-world-records-with-claas-/1890846

Council of Economic Advisers. 2011. *The Economic Report of the President,* Chapter 3: The Foundations of Growth, 53–79.

Costa, Dora L. 2015. "Health and the Economy in the United States from 1750 to the Present." *Journal of Economic Literature* 53, 503–570.

Cutler, D., A. Deaton, and A. Lleras-Muney. 2006. "The Determinants of Mortality," *Journal of Economic Perspectives* 20, 97–120.

de Solla Price, Derek J. 1965. "Networks of scientific papers," *Science* 149, 510–515.

Edwards, Aled M., Ruth Isserlin, Gary D. Bader, Stephen V. Frye, Timothy M. Willson and Frank H. Yu. 2011. "Too Many Roads Not Taken," *Nature* 470, 163–165.

Einstein, Albert. 1949. *The World as I See It.* Citadel Press, Secaucus NJ.

Evenson, R. 2001. "Economic impacts of agricultural research and extension," in the *Handbook of Agricultural Economics,* 573¬–628.

Fuglie, Keith O., James M. MacDonald, and Eldon Ball. 2007. "Productivity Growth in U.S. Agriculture." USDA Economic Brief Number 9.

Goldin, Claudia and Lawrence F. Katz. 2010. *The Race between Education and Technology,* Harvard University Press.

Goolsbee, Austan and Benjamin F. Jones (editors). *Innovation and Public Policy,* University of Chicago Press, forthcoming.

Gordon, Robert J. 2012. "Is U.S. Economic Growth Over? Faltering Innovation Confronts the Six Headwinds," National Bureau of Economic Research Working Paper #18315.

Griliches, Zvi. 1958. "Research Cost and Social Returns: Hybrid Corn and Related Innovations," *Journal of Political Economy* 66, 419–431.

Gross, Daniel P. and Bhaven Sampat. 2020. "Organizing Crisis Innovation: Lessons from World War II," National Bureau of Economic Research Working Paper #27909.

______. 2021. "Crisis Innovation Policy from World War II to COVID-19," National Bureau of Economic Research Working Paper #28915.

Hall, Bronwyn, Mairesse, Jacques, and Pierre Mohnen. 2010. "Measuring the Returns to R&D," in *Handbook of the Economics of Innovation,* 1033–1082.

Hatfield, E. 2006. "Proxmire's Golden Fleece Award," *Relationship Research News (Newsletter of the International Association for Relationship Research)* 4, 5–9.

Hill, Ryan, Yian Yin, Carolyn Stein, Dashun Wang, and Benjamin F. Jones. 2021. "Adaptability and the Pivot Penalty in Science," Northwestern University.

Hourihan, Matt. 2020. "A Snapshot of U.S. R&D Competitiveness: 2020 Update," American Academy of Arts and Sciences, https://www.aaas.org/news/snapshot-us-rd-competitiveness-2020-update.

Jones, Benjamin F. 2009. "The burden of knowledge and the 'death of the renaissance man': Is innovation getting harder?" *Review of Economic Studies* 7, 283–317.

______. 2010. "Age and Great Invention," *Review of Economics and Statistics* 99.

______. 2011. "As Science Evolves, How Can Science Policy?" National Bureau of Economic Research *Innovation Policy and the Economy,* edited by Josh Lerner and Scott Stern, 11.

Jones, Benjamin F., E.J. Reedy, and Bruce Weinberg. 2014. "Age and Scientific Genius," in the Handbook of Genius (Dean Simonton, editor), Wiley.

Jones, Benjamin F. and Lawrence H. Summers. 2020. "A Calculation of the Social Returns to Innovation," in *Innovation and Public Policy*, edited by Austan Goolsbee and Benjamin F. Jones, University of Chicago Press.

Jones, Charles. 2016. "The Facts of Economic Growth." In Handbook of Macroeconomics, Editor(s): John B. Taylor, Harald Uhlig, Handbook of Macroeconomics, Elsevier, Volume 2, 3–69.

Jones, Charles I. and John C. Williams. 1998. "Measuring the Social Return to R&D," *Quarterly Journal of Economics*, 1119–1135.

Ke, Qing, Emilio Ferrara, Filippo Radicchi, and Alessandro Flammini. 2015. "Defining and Identifying Sleeping Beauties in Science," *Proceedings of the National Academy of Sciences* 112, 7426–7431.

Kealey, T. 2013. "The Case against Public Science," *Cato Unbound* 5.

Kerr, William, Ramana Nanda, and Matthew Rhodes-Kropf. 2014. "Entrepreneurship as Experimentation," *Journal of Economic Perspectives* 28, 25–48.

Kerr, Sari Pekkala and William R. Kerr. 2020. "Immigration Policy Levers for US Innovation and Startups," in *Innovation and Public Policy,* edited by Austan Goolsbee and Benjamin F. Jones, University of Chicago Press.

Kola, Ismail, and John Landis. 2004 "Can the Pharmaceutical Industry Reduce Attrition Rates?" *Nature Reviews Drug Discovery* 3, 711–715.

Lee, You-Na, John P. Walsh, and Jian Wang. 2015. "Creativity in Scientific Teams: Unpacking Novelty and Impact." *Research Policy* 44 (3): 684–97.

Lu Liu, Yang Wang, Roberta Sinatra, C. Lee Giles, Chaoming Song, and Dashun Wang. 2018. "Hot Streaks in Artistic, Cultural, and Scientific Careers," *Nature* 559, 396–399.

Mansfield, Edwin, John Rapoport, Anthony Romeo, Samuel Wagner, and George Beardsley. 1977. "Social and Private Rates of Return from Industrial Innovations," *Quarterly Journal of Economics* 77, 221–240.

Mesmer-Magnus, Jessica R., and Leslie A. DeChurch. 2009. "Information Sharing and Team Performance: A Meta-Analysis." *Journal of Applied Psychology* 94 (2): 535–46.

Minsky and Papert. 1969. Perceptrons: An Introduction to Computational Geometry.

Mokyr, Joel. 1990. *The Lever of Riches: Technological Creativity and Economic Progress.* Oxford University Press: New York.

Murphy, Kevin M. and Robert Topel. 2007. "The Value of Health and Longevity," Journal of Political Economy 114, 871–904.

National Academy of Sciences. 2007. *Rising Above the Gathering Storm: Energizing and Employing America for a Brighter Economic Future.* Washington, DC: The National Academies Press. https://doi.org/10.17226/11463.

Nichols, Tom. 2017. *The Death of Expertise: The Campaign against Established Knowledge and Why It Matters.* Oxford University Press.

Schweitzer, Sascha and Jan Brendel. 2019. (A Burden of Knowledge Creation in Academic Research: Evidence from Publication Data (November 9, 2019). Available at SSRN: https://ssrn.com/abstract=2895685 or http://dx.doi.org/10.2139/ssrn.2895685.

Shirk, Susan L. and Orville Schell. 2020. “Meeting the China Challenge: A New American Strategy for Technology Competition,” 21st Century China Center, UC San Diego School of Global Policy and Strategy.

Slaoui, Moncef, and Matthew Hepburn. 2020. “Developing Safe and Effective Covid Vaccines -- Operation Warp Speed’s Strategy and Approach,” *New England Journal of Medicine* 383, 1701–1703.

Snow, C. P. 1959. *The Rede Lecture*. Cambridge University Press.

Solow, Robert M. 1956. “A Contribution to the Theory of Economic Growth.” *Quarterly Journal of Economics* 70 (1): 65–94.

Stoeger, Thomas, Martin Gerlach, Richard I. Morimoto, and Luis A. Nunes Amaral. 2018. “Large-scale Investigation of the Reasons Why Potentially Important Genes are Ignored,” PLOS Biology, PLoS Biol 16(9): e2006643.

United States Department of Agriculture. 2020. “Agricultural Productivity in the U.S.: Summary of Recent Findings,” https://www.ers.usda.gov/data-products/agricultural-productivity-in-the-us/summary-of-recent-findings/.

Waldfogel, Joel. 2021. “Digitization and Its Consequences for Creative-Industry Product and Labor Markets,” in *The Role of Innovation and Entrepreneurship in Economic Growth*, edited by Aaron Chatterji, Josh Lerner, Scott Stern and Michael J. Andrews, University of Chicago Press.

Watzinger, Martin, Joshua Krieger, and Monika Schnitzer. 2021. “Standing on the Shoulders of Science,” Centre for Economic Policy Research Discussion Paper #13766.

Wooldridge, Michael. 2021. “Artificial Intelligence Is a House Divided: A decades-old rivalry has riven the field. It’s time to move on,” Chronicle of Higher Education.

Wrapp, Daniel, Nianshuang Wang, Kizzmekia S. Corbett, Jory A. Goldsmith, Ching-Lin Hsieh, Olubukola Abiona, Barney S. Graham, and Jason S. McLellan. 2020. “Cryo-EM Structure of the 2019-nCoV Spike in the Prefusion Conformation,” *Science* 367, 1260–1263.

Wuchty, Stefan, Benjamin F. Jones, and Brian Uzzi. 2007. “The Increasing Dominance of Teams in the Production of Knowledge.” *Science* 316 (5827): 1036–39.

Yin, Yian, Yang Wang, James Adams, and Dashun Wang. 2019. “Quantifying the Dynamics of Failure across Science, Startups and Security,” *Nature* 575.

Yin, Yian, Yuxiao Dong, Kuansan Wang, Dashun Wang, and Benjamin F. Jones. 2021. “Science as a Public Good: Public Use and Funding of Science,” National Bureau of Economic Research Working Paper #28748.

Zeira, Joseph. 2011. “Innovations, Patent Races and Endogenous Growth,” *Journal of Economic Growth* 16, 135–156.

Author Biographies

JOSE MARIA BARRERO

Assistant Professor of Finance, Instituto Tecnológico Autónomo de México (ITAM)

Jose Maria Barrero is an Assistant Professor of Finance at Instituto Tecnológico Autónomo de México (ITAM), where he conducts empirical and quantitative research in macroeconomics, finance, and labor economics, much of it focusing on firm behavior under uncertainty and more recently on the labor market impact of the COVID-19 pandemic. Professor Barrero holds a BA in Economics and Mathematics from the University of Pennsylvania and an MA and PhD in Economics from Stanford University.

NICHOLAS A. BLOOM

William D. Eberle Professor of Economics, Stanford University

Nick Bloom is a Professor in the department of economics and Professor, by courtesy, at the Graduate School of Business. He is also the Co-Director of the Productivity, Innovation and Entrepreneurship program at the National Bureau of Economic Research (NBER), and a fellow of the Centre for Economic Performance, and the Stanford Institute for Economic Policy Research. He is currently a Faculty Research Fellow in the NBER programs on public economics and corporate finance. Nick was an undergraduate in Cambridge, a masters student at Oxford, and a PhD student at University College London. While completing his PhD he worked part-time at the Institute for Fiscal Studies, a London based tax think-tank. After completing his PhD Nick worked as a business tax policy advisor to the UK Treasury, and then joined McKinsey & Company as a management consultant. In 2003 he moved to the London School of Economics to focus on research, before joining Stanford University in 2005. Professor Bloom's research focuses on measuring and explaining management practices. He has been working with McKinsey & Company as part of a long-run effort to collect management data from over 10,000 firms across industries and countries. The aim is to build an empirical basis for understanding what factors drive differences in management practices across regions, industries and countries, and how this determines firm and national performance. More recently he has also been working with Accenture on running management experiments. He also works on understanding the impacts of large uncertainty shocks—such as the credit crunch, the 9/11 terrorist attacks and the Cuban Missile crisis—on the US economy, for which he won the Frisch Medal in 2010. Nick lives on Stanford campus with his wife and three children. As a born and bred Londoner, married to a Scottish wife, with kids attending US schools, he lives in a multi-lingual English household.

SEVERIN BORENSTEIN

Professor of Economics, Haas School of Business, University of California Berkeley

Severin Borenstein is E.T. Grether Professor of Business Administration and Public Policy at the Haas School of Business and Faculty Director of the Energy Institute at Haas. He is also Director emeritus of the University of California Energy Institute (1994-2014). He received his A.B. from U.C. Berkeley and Ph.D. in Economics from M.I.T. His research focuses on business competition, strategy, and regulation. He has published extensively on the airline industry, the oil and gasoline industries, and electricity markets. His current research projects include the economics of renewable energy, economic policies for reducing greenhouse gases, alternative models of retail electricity pricing, and competitive dynamics in the airline industry. Borenstein is also a research associate of the National Bureau of Economic Research in Cambridge, MA. He served on the Board of Governors of the California Power Exchange from 1997 to 2003. During 1999-2000, he was a member of the California Attorney General's Gasoline Price Task Force. In 2010-11, Borenstein was a member of U.S. Secretary of Transportation Ray LaHood's Future of Aviation Advisory Committee. In 2012-13, he served on the Emissions Market Assessment Committee, which advised the California Air Resources Board on the operation of California's Cap and Trade market for greenhouse gases. In 2014, he was appointed to the California Energy Commission's Petroleum Market Advisory Committee, which he chaired from 2015 until the Committee was dissolved in 2017. From 2015-2020, he served on the Advisory Council of the Bay Area Air Quality Management District. In 2019, he was appointed to the Governing Board of the California Independent System Operator.

ERSKINE BOWLES

President Emeritus, The University of North Carolina;
Co-Chair, Aspen Economic Strategy Group

Erskine Bowles is co-chair of the Aspen Economic Strategy Group. He began his business career at Morgan Stanley and went on to found and lead Bowles, Hollowell, Conner, which became the preeminent investment bank focused on middle market mergers and acquisitions. He subsequently co-founded the middle market private equity firm Carousel Capital, served as a senior advisor for Credit Suisse First Boston, was a partner with the global buyout firm Forstmann Little & Co. and the chairman of the board of advisors of BDT Capital Partners. He previously served on the boards of Morgan Stanley (Lead Director), First Union Corporation, Merck, VF, Cousins Properties (Lead Director), Norfolk Southern Corporation, General Motors, Belk and Facebook. In his public sector career, Erskine served in the Clinton Administration as SBA Administrator, Deputy White House chief of staff, and chief of staff. During his time in the White House, he negotiated the first balanced budget in a generation. Erskine co-chaired with Senator Alan Simpson the National Commission on Fiscal Responsibility and Reform, and from 2005-2011, he served as the president of the University of North Carolina.

CHAD P. BOWN

Reginald Jones Senior Fellow, Peterson Institute for International Economics

Chad P. Bown is Reginald Jones Senior Fellow at the Peterson Institute for International Economics in Washington, DC. With Soumaya Keynes, he cohosts Trade Talks, a podcast on the economics of international trade policy. Bown previously served as senior economist for international trade and investment in the White House on the Council of Economic Advisers and most recently as a lead economist at the World Bank. He was a tenured professor of economics at Brandeis University, where he held a joint appointment in the Department of Economics and International Business School for 12 years. He has also spent a year in residence as a visiting scholar in economic research at the World Trade Organization (WTO) Secretariat in Geneva. Bown is also currently a research fellow at the Centre for Economic Policy Research (CEPR) in London and a member of the Council on Foreign Relations. His work has been published in Foreign Affairs, The Atlantic, Wall Street Journal, and Financial Times, as well as in academic journals such as American Economic Review, Journal of Economic Literature, Review of Economics and Statistics, Journal of International Economics, and Journal of Development Economics. Bown is author of the book Self-Enforcing Trade: Developing Countries and WTO Dispute Settlement (Brookings Institution Press, 2009), and coeditor, with

Joost Pauwelyn, of The Law, Economics, and Politics of Retaliation in WTO Dispute Settlement (Cambridge University Press, 2010). His volume on the global economic crisis, The Great Recession and Import Protection: The Role of Temporary Trade Barriers (CEPR and World Bank, 2011), was built from a trade policy transparency project that he initiated at the World Bank in 2004. Bown received a BA magna cum laude in economics and international relations from Bucknell University and a PhD in economics from the University of Wisconsin-Madison.

STEVEN J. DAVIS

William H. Abbott Distinguished Service Professor of International Business and Economics, The University of Chicago Booth School of Business

Steven J. Davis is distinguished service professor of business and economics at the University of Chicago Booth School of Business and senior fellow at the Hoover Institution. He is also a research associate with the National Bureau of Economic Research, advisor to the U.S. Congressional Budget Office, consultant to the Federal Reserve Bank of Atlanta, senior academic fellow with the Asian Bureau of Finance and Economics Research, senior adviser to the Brookings Papers on Economic Activity, past editor of the American Economic Journal: Macroeconomics, and elected fellow of the Society of Labor Economists. Davis is an applied economist who studies business dynamics, worker outcomes, economic uncertainty, public policy, and other topics. He is known for influential research using longitudinal data on firms and establishments to explore business outcomes, labor market consequences, and economic performance. He is a co-founder of the Economic Policy Uncertainty project, the Survey of Business Uncertainty, and the Stock Market Jumps project. He co-organizes the Asian Monetary Policy Forum, held annually in Singapore. In addition to his scholarly work, Davis has written for the Atlantic, Bloomberg View, Financial Times, Forbes, Los Angeles Times, Wall Street Journal and other popular media. He has appeared on BBC, Bloomberg TV, CGTN, Channel News Asia, CNBC, CNN, Fox News, NBC Network News, Sinclair Broadcast Group, and the U.S. Public Broadcasting System.

AMY GANZ

Deputy Director, Aspen Economic Strategy Group

Amy Ganz is Deputy Director of the Aspen Economic Strategy Group. She has worked in several policy research and analysis roles focused on tax policy, housing and urban policy, and state and local finance. She holds a BA in economics and political science from Wheaton College and an MPA in public policy analysis from NYU Wagner School of Public Service.

EDWARD GLAESER

Fred And Eleanor Glimp Professor of Economics, Harvard University

Edward Glaeser is the Fred and Eleanor Glimp Professor of Economics and Chair of the Economics Department in the Faculty of Arts and Sciences at Harvard University, where he has taught since 1992. He regularly teaches courses on microeconomics and the economics of cities. His books include Triumph of the City, Survival of the City (joint with David Cutler) and Fighting Poverty in the US and Europe: A World of Difference (joint with Alberto Alesina). He leads the NBER's Urban Economics Working Group and co-directs the Cities Programme of the International Growth Centre. He received his PhD from the University of Chicago in 1992.

NORA E. GORDON

Professor, Georgetown University

Nora Gordon is a Professor at the Georgetown University McCourt School of Public Policy. She is a Research Associate of the National Bureau of Economic Research, Non-Resident Fellow of the Urban Institute, and a member of the FutureEd Advisory Board. She is an Associate Editor of Education Finance and Policy and the Journal of Human Resources. Dr. Gordon is the co-author of Common-Sense Evidence: The Education Leader's Guide to Using Data and Research (Harvard Education Press, 2020). She studies American education policy, with an emphasis on equity, the federal role in elementary and secondary education, and the use of evidence in policy and practice. Dr. Gordon has testified before Senate and House committees on implementation of the Every Student Succeeds Act. She has served on the Institute of Education Sciences Expert Panel on the Study of the Title I Formula and DC's state Title I Committee of Practitioners, and currently serves on the Professional Advisory Board of the National Center for Learning Disabilities.

SAMUEL G. HANSON

Professor of Business Administration, Harvard Business School

Samuel G. Hanson is a Professor in the Finance Unit at Harvard Business School, and a Faculty Research Fellow at the National Bureau of Economic Research. He teaches the Investment Strategies course in the MBA elective curriculum and PhD courses in Corporate Finance and Empirical Methods. Professor Hanson holds a Ph.D. in Business Economics from Harvard University and a B.A. in Quantitative

Economics and Philosophy from Tufts University. Before beginning his doctoral studies, he worked as an investment banking analyst at Lehman Brothers and as an assistant economist at the Federal Reserve Bank of New York. During 2009 Hanson worked at the U.S. Treasury Department where he served as a Special Assistant and Liaison to the White House National Economic Council. Professor Hanson's research interests lie in corporate finance, behavioral finance, and asset pricing. His recent research has focused on corporate supply responses to fluctuations in investor demand for different types of securities and on optimal financial regulation. Hanson's research has appeared in the Quarterly Journal of Economics, the Journal of Finance, the Journal of Financial Economics, the Review of Financial Studies, and the Journal of Economic Perspectives.

BENJAMIN F. JONES

Gordon and Llura Gund Family Professor of Entrepreneurship;
Professor of Strategy, Kellogg School of Management Northwestern University

Benjamin F. Jones is the Gordon and Llura Gund Family Professor of Entrepreneurship, a Professor of Strategy, and the faculty director of the Kellogg Innovation and Entrepreneurship Initiative. An economist by training, Professor Jones studies the sources of economic growth in advanced economies, with an emphasis on innovation, entrepreneurship, and scientific progress. He also studies global economic development, including the roles of education, climate, and national leadership in explaining the wealth and poverty of nations. His research has appeared in journals such as Science, the Quarterly Journal of Economics and the American Economic Review, and has been profiled in media outlets such as the Wall Street Journal, the Economist, and The New Yorker. A former Rhodes Scholar, Professor Jones served in 2010-2011 as the senior economist for macroeconomics for the White House Council of Economic Advisers and earlier served in the U.S. Department of the Treasury. Professor Jones is a non-resident senior fellow of the Brookings Institution, a research associate of the National Bureau of Economic Research, and a member of the Council on Foreign Relations.

MELISSA S. KEARNEY

Director, Aspen Economic Strategy Group;
Neil Moskowitz Professor of Economics, University of Maryland

Melissa S. Kearney is the Neil Moskowitz Professor of Economics at the University of Maryland. She is also director of the Aspen Economic Strategy Group; a research associate at the National Bureau of Economic Research; and a nonresident senior fellow at the Brookings Institution. Kearney's scholarly research focuses on poverty, inequality, and social policy in the United States. Her work is published is leading academic journals and is frequently cited in the press. She is a senior editor of the Future of Children, an editorial board member of the American Economic Journal: Economic Policy, Journal of Economic Literature, and Demography, and she was previously co-editor of the Journal of Human Resources. She serves on advisory boards for the Notre Dame Wilson-Sheehan Lab for Economic Opportunities and the Smith Richardson Foundation. Kearney previously served as Director of the Hamilton Project at Brookings and as co-chair of the Massachusetts Institute of Technology J-PAL State and Local Innovation Initiative. Kearney teaches Public Economics at both the undergraduate and Ph.D. level at the University of Maryland. She holds a B.A. in Economics from Princeton University and a Ph.D. in Economics from MIT.

RYAN KELLOGG

Professor, The University of Chicago Harris School of Public Policy

Ryan Kellogg is a professor at the University of Chicago Harris School of Public Policy and Deputy Dean for Academic Programs. He is also a research associate at the National Bureau for Economic Research. Ryan's research bridges industrial organization, energy economics, and environmental policy, examining topics such as the response of investment to uncertainty, the economic consequences of the shale boom, the economics of oil and gas leasing, the effectiveness of policies to reduce emissions from the transportation sector, and the economics of carbon regulation when the economic environment is uncertain. Ryan earned a Ph.D. in Agricultural and Resource Economics from the University of California, Berkeley, in 2008. Prior to his graduate studies, he worked for BP in Houston, TX, and Anchorage, AK, for four years as an engineer and economic analyst. Ryan earned a B.S. in Chemical Engineering and a B.A. in Economics from Rice University in 1999. He grew up outside of Cleveland, OH.

HENRY M. PAULSON, JR.

Chairman, Paulson Institute; Co-Chair, Aspen Economic Strategy Group

Hank Paulson is executive chairman of TPG Rise Climate, the climate investing platform of the global private equity firm TPG. He is also the founder and chairman of the Paulson Institute, which aims to foster a more balanced and productive US-China relationship focusing at the intersection of green finance, conservation, and economics. Paulson is the co-chair of the Aspen Economic Strategy Group, and co-chair of the Bloomberg New Economy Forum Advisory Board. Paulson served as the 74th Secretary of the Treasury under President George W. Bush. Prior to that, he had a thirty-two year career at Goldman Sachs, ultimately serving as chairman and chief executive officer. A lifelong conservationist, Paulson was Chairman of The Nature Conservancy Board of Directors and, prior to that, founded and co-chaired the organization's Asia-Pacific Council. He also founded and co-chaired the Latin American Conservation Council. In addition, he co-chaired the Risky Business Project, a non-partisan initiative that quantified and publicized the economic risks of climate change in the United States. Paulson is the author of the best sellers On the Brink and Dealing with China. He is also the co-author of two books with Ben Bernanke and Tim Geithner, First Responders and Firefighting. Paulson graduated from Dartmouth College and received an M.B.A. from Harvard University.

JAMES POTERBA

President, National Bureau of Economic Research;
Mitsui Professor of Economics, Massachusetts Institute of Technology

James Poterba is the Mitsui Professor of Economics at the Massachusetts Institute of Technology and the president of the National Bureau of Economic Research, a nonprofit research organization with nearly 1600 affiliated economists. His recent research has analyzed the determinants of retirement saving, the draw-down of assets after households reach retirement, and the role of tax-deferred retirement saving programs such as 401(k) plans in contributing to retirement security. He has served as president of the Eastern Economic Association and the National Tax Association, and as vice president of the American Economic Association. He is a member of the National Academy of Sciences. He is a trustee of the College Retirement Equity Fund, the TIAA-CREF mutual funds, and of the Alfred P. Sloan Foundation. He holds an undergraduate degree from Harvard College and a D. Phil. in economics from Oxford University, where he was a Marshall scholar.

SARAH J. REBER

Associate Professor of Public Policy, University of California Los Angeles

Sarah Reber is an associate professor of public policy at the Luskin School of Public Affairs at UCLA, a research associate with the National Bureau of Economic Research, and a David M. Rubenstein Fellow in Economic Studies at the Brookings Institution. After receiving her PhD in economics from Harvard University, Reber was a Robert Wood Johnson Foundation Scholar in Health Policy Research at the University of California, Berkeley. She was also a Staff Economist at the Council of

Economic Advisers during the Clinton Administration. Professor Reber's research in education has focused on school desegregation, elementary and secondary school finance, and college access. Her research in health economics has examined racial and ethnic disparities in covid-19 mortality and the advantages and disadvantages of promoting competition in health insurance markets.

ADI SUNDERAM

Professor of Business Administration, Harvard Business School

Adi Sunderam is a professor of business administration in the Finance Unit and a Research Associate at the National Bureau of Economic Research. He teaches Investment Management in the MBA elective curriculum and Ph.D. courses in Corporate Finance and Empirical Methods. Professor Sunderam holds a Ph.D. in business economics and an A.B. in computer science and economics, both from Harvard University. In 2009 and 2010, he served in the U.S. Treasury Department as a special assistant and liaison to the White House National Economic Council. Professor Sunderam's research interests are in corporate finance, asset pricing, and financial intermediation. His work focuses on the organization of financial markets and its effect on asset prices and corporate investment. It has appeared in the Quarterly Journal of Economics, the American Economic Review, the Journal of Finance, the Journal of Financial Economics, and the Review of Financial Studies. He serves as an associate editor for the Journal of Finance.

TILL VON WACHTER

Professor of Economics, University of California Los Angeles

Till Von Wachter is Professor of Economics at the University of California Los Angeles, Faculty Director of the California Policy Lab, Director of the Federal Statistical Research Data Center, and Associate Dean for Research for the Social Science Division. Prof. von Wachter's research examines how labor market conditions and institutions affect the well-being of workers and their families. This includes the analysis of unemployment and job loss on workers' long-term earnings and health outcomes, as well as the role of unemployment insurance and disability insurance in buffering such shocks. Current research projects focus on job and earnings mobility of young workers over their careers, the effects of minimum wages, as well as several projects on homelessness using administrative data from Los Angeles. Professor von Wachter's research has been supported by the National Science Foundation, the National Institute on Aging, the Social Security

Administration, the Sloan Foundation, the Russell Sage Foundation, the Hilton Foundation, and Arnold Ventures. Prof. von Wachter has been an expert witness in numerous testimonies before committees of U.S. Congress, and has provided expert assistance to the City and County of Los Angeles, the U.S. Department of Labor, the Canadian Labor Ministry, the OECD, the United Nations, and the IMF.

ERIC ZWICK

Associate Professor of Finance and Fama Faculty Fellow, The University of Chicago Booth School of Business

Eric Zwick studies the interaction between public policy and corporate behavior, with a focus on fiscal stimulus, taxation and housing policy. His research draws insights from finance and behavioral economics while using a variety of methods: new data, natural experiments, theory and anecdotal exploration. Zwick is particularly interested in the problems that small and medium-sized private firms and new ventures face, from the perspective of owners, investors, managers and workers. A secondary area of interest concerns the role of bounded rationality and imperfect information in the design of policies that promote behavior change. This work focuses on determinants of habit formation in health and workforce productivity settings. Zwick earned a Ph.D. and M.A. in business economics from Harvard University and a B.A. in economics and mathematics with high honors from Swarthmore College. Prior to grad school, he worked as a research assistant at the National Bureau of Economic Research and as a web and software developer for several start-ups and non-profits.

ASPEN ECONOMIC STRATEGY GROUP
aspen institute

CO-CHAIRS

Henry M. Paulson, Jr.
74th Secretary of the US Treasury
Founder and Chairman
Paulson Institute

Erskine Bowles
President Emeritus
The University of North Carolina

ASPEN INSTITUTE CHAIRMAN

James S. Crown
Chairman and CEO
Henry Crown and Company

CHAIRMAN EMERITUS

Robert K. Steel
Partner and Chairman
Perella Weinberg Partners

ASPEN INSTITUTE PRESIDENT

Daniel Porterfield
President and CEO
The Aspen Institute

DIRECTOR

Melissa S. Kearney
Neil Moskowitz Professor of Economics
The University of Maryland

DEPUTY DIRECTOR

Amy Ganz
The Aspen Institute

MEMBERSHIP

Vickee Jordan Adams
Partner
Finsbury Glover Hering

Kelly Ayotte
Former United States Senator
New Hampshire

Mary T. Barra
Chairman and CEO
General Motors Company

Ben Bernanke
Distinguished Fellow in Residence
Brookings Institution

Joshua Bolten
President and CEO
Business Roundtable

Raphael Bostic
President and CEO
Federal Reserve Bank of Atlanta

Sylvia M. Burwell
President
American University

Kenneth I. Chenault
Chairman and Managing Director
General Catalyst

Susan Collins
Provost & Executive Vice President for Academic Affairs
University of Michigan

Bob Corker
Former United States Senator
Tennessee

David Cote
Executive Chairman
Vertiv Holdings

James S. Crown
Chairman and CEO
Henry Crown and Company
Chairman of the Board of Trustees
The Aspen Institute

Ray Dalio
Founder, Co-Chief Investment Officer, and Co-Chairman
Bridgewater Associates LP

Mitchell E. Daniels, Jr.
President
Purdue University

Susan Desmond-Hellmann
Senior Advisor
Bill and Melinda Gates Foundation and Google Ventures

Roger Ferguson
Steven A. Tananbaum Distinguished Fellow for International Economics
Council on Foreign Relations

Laurence D. Fink
Founder, Chairman and CEO
BlackRock

Kristin J. Forbes
Jerome and Dorothy Lemeson Professor of Management and Global Economics
MIT Sloan School of Management

Michael Froman
Vice Chairman & President,
Strategic Growth Mastercard

Jason Furman
Aetna Professor of the Practice of Economic Policy
Harvard University

Timothy F. Geithner
75th Secretary of the US Treasury
President
Warburg Pincus

Austan D. Goolsbee
Robert P. Gwinn Professor of Economics
The University of Chicago Booth School of Business

Danielle Gray
Executive Vice President & Global Chief Legal Officer
Walgreens Boots Alliance, Inc.

Jon Gray
President and Chief Operating Officer
Blackstone

Bill Haslam
49th Governor
Tennessee

Keith Hennessey
Lecturer in Economics
Stanford University Graduate School of Business

James Poterba
President
NBER
Mitsui Professor of Economics
Massachusetts Institute of Technology

Penny Pritzker
Founder and Chairman
PSP Partners

Robert E. Rubin
70th Secretary of the US Treasury
Counselor
Centerview Partners

Paul Ryan
Former US Speaker of the House
Founder
American Idea Foundation

Matthew Slaughter
Dean
Dartmouth Tuck School of Business

Margaret Spellings
President and CEO
Texas 2036

William Spriggs
Professor of Economics
Howard University

Robert K. Steel
Chairman and Partner
Perella Weinberg Partners

Michael R. Strain
Director of Economic Policy Studies; Arthur F. Burns Scholar in Political Economy
American Enterprise Institute

Lawrence H. Summers
71st Secretary of the US Treasury
Charles W. Eliot University Professor; President Emeritus
Harvard University

Kevin Warsh
Shepard Family Distinguished Visiting Fellow
Hoover Institution

Mark Weinberger
Director
Johnson & Johnson; MetLife; Saudi Aramco

Tom Wilson
Chair, President and Chief Executive Officer
The Allstate Corporation

Robert Zoellick
Senior Counselor
Brunswick Group

ASPEN ECONOMIC STRATEGY GROUP

aspen institute

ABOUT THE ECONOMIC STRATEGY GROUP

The Aspen Economic Strategy Group (AESG), a program of the Aspen Institute, is composed of a diverse, bipartisan group of distinguished leaders and thinkers with the goal of promoting evidence-based solutions to significant US economic challenges. Co-chaired by Henry M. Paulson, Jr. and Erskine Bowles, the AESG fosters the exchange of economic policy ideas and seeks to clarify the lines of debate on emerging economic issues while promoting bipartisan relationship-building among current and future generations of policy leaders in Washington.

CONTRIBUTORS

Jose Maria Barrero, Instituto Tecnológico Autónomo de México
Nicholas Bloom, Stanford University
Severin Borenstein, University of California, Berkeley
Chad P. Bown, Peterson Institute for International Economics
Steven J. Davis, The University of Chicago
Edward Glaeser, Harvard University
Nora Gordon, Georgetown University
Samuel Hanson, Harvard University
Benjamin F. Jones, Northwestern University
Ryan Kellogg, The University of Chicago
James Poterba, NBER and Massachusetts Institute of Technology
Sarah Reber, University of California, Los Angeles
Adi Sunderam, Harvard University
Till von Wachter, University of California, Los Angeles
Eric Zwick, The University of Chicago

21/002