8-24-2020

Alternative Economic Indicators

C. James Hueng, Editor
Western Michigan University

Citation

Hueng, C. James, ed. 2020. Alternative Economic Indicators. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

Alternative Economic Indicators

C. James Hueng, editor

Alternative Economic Indicators

Alternative Economic Indicators

C. James Hueng
Editor

2020

W.E. Upjohn Institute for Employment Research
Kalamazoo, Michigan

Copyright Page

ISBN 9780880996778 (ebook)

W.E. Upjohn Institute for Employment Research
300 S. Westnedge Avenue
Kalamazoo, Michigan 49007-4686

Cover design by Carol A.S. Derks.

Contents

Acknowledgments

I would like to acknowledge the supports from the W.E. Upjohn Institute for Employment Research and the Department of Economics and College of Arts and Sciences at Western Michigan University.

—C. James Hueng

1
Alternative Economic Indicators

C. James Hueng
Western Michigan University

Policymakers and business practitioners are eager to gain access to reliable information on the state of the economy for timely decision making. Traditional economic indicators have been criticized for delayed reporting, out-of-date methodology, and neglecting some aspects of the economy. Recent advances in economic theory, econometrics, and information technology have fueled research in building broader, more accurate, and higher-frequency economic indicators. The 2018–2019 Werner Sichel Lecture Series invited six prominent economists to speak on their current research on alternative economic indicators, including indicators in the financial market, indicators for business cycles, and indicators of economic uncertainty. Their lectures have been compiled in this volume.

In Chapter 2, William Barnett and Kun He argue that the growing complexity of financial instruments has made the traditional simple-sum monetary aggregates such as M1–M4 obsolete. The authors outline the evidence showing how the Fed's simple-sum monetary aggregates have provided misleading information about the economy and monetary policy. In contrast, they show how their Divisia monetary aggregates have been more in line with the true liquidity conditions in the economy. Unlike the simple-sum aggregates, which assume that all monetary components contribute equally to the aggregate, the Divisia monetary aggregates weight the growth of each component using a formula based upon its user cost to reflect its liquidity in making transactions. Barnett and He describe the latest efforts in constructing the Divisia monetary aggregates by incorporating credit card services and distinguishing between the demand-side and the supply-side money services.

In Chapter 3, Scott Brave introduces the National Financial Conditions Index (NFCI) and Adjusted National Financial Conditions Index (ANFCI) as measures of the overall financial market condition provided by the Federal Reserve Bank of Chicago. The latter is adjusted for the state of the business cycle and the level of inflation. Since the global financial crisis, economists at the Chicago Fed have constructed composite indices that aim to measure the overall tightness of the U.S. financial system. The NFCI is a weekly summary statistic estimated by a mixed-frequency dynamic factor model on a panel of 105 weekly, monthly, and quarterly financial time series. In the chapter, Brave shows that the index aligns closely with the historical episodes of financial stress and has been a useful tool in monitoring financial stability.

As to alternative business cycle indicators, numerous efforts have been devoted to replacing the traditional GDP measure. Many have adopted data-dimension reduction techniques such as principal components analysis and dynamic factor analysis to extract as much information from as many variables as efficiently as possible. Brave introduces the Chicago Fed National Activity Index (CFNAI), a monthly summary statistic for U.S. economic growth estimated by principal components analysis using 85 monthly indicators. The CFNAI has been shown to be roughly 95 percent accurate historically in identifying U.S. recessions (as defined by the National Bureau of Economic Research) since 1967.

Based on a dynamic factor model, the New York Fed Staff Nowcast is an early estimate of GDP growth for the current and subsequent quarters. In Chapter 4, Domenico Giannone and his coauthors, Patrick Adams, Eric Qian, Argia Sbordone, and Mihir Trivedi, present in detail this automated platform for monitoring U.S. macroeconomic conditions in real time. This nowcasting model synthesizes a large number of variables (macroeconomic big data) monitored by economists, incorporating new information within minutes of the data releases. It is entirely automated and mimics best practices without relying on any subjective judgment. This platform provides a model-based counterpart to the analysis traditionally based on expert knowledge. The authors show that the New York Fed Staff Nowcast provided accurate early estimates of the U.S. GDP during the Great Recession.

In Chapter 5, Alessandro Barbarino and Chiara Scotti provided an estimate of the probability of a recession occurring in 2019 by comparing various models and employing a mix of macroeconomic and financial indicators, including the Aruoba-Diebold-Scotti (ADS) real business condition index and Scotti's surprise and uncertainty indexes, explained below. The ADS index, maintained and updated by the Philadelphia Fed, is derived from a dynamic factor model as well. It is a daily index that measures the latent real business conditions in real time, emphasizing that a business cycle is about the dynamics and interactions of many economic indicators from various frequencies. The surprise and uncertainty indexes are daily measures of surprises and uncertainty about the U.S. real activity, as measured by the ADS index. Scotti and Barbarino conclude that real variables are more powerful in signaling recessions at shorter horizons, while financial variables are valuable leading indicators at longer horizons. Their model, using real variables, did not show a high recession probability in 2019 (as of mid-March of that year), contrary to what was suggested by the Congressional Budget Office and several published surveys.

In addition to Scotti's index of uncertainty about the real activity, the following chapter presents another index for uncertainty—the economic policy uncertainty index. In Chapter 6, Steven Davis details the construction of the index and highlights the effect of the shift in U.S. trade policy under the Trump administration on economic uncertainty. The index is constructed from newspaper coverage of policy-related economic uncertainty by using computer-automated newspaper searches. Davis shows that the U.S.-China conflict over trade and commercial policies has become a major source not only of economic policy uncertainty but also of increased equity market volatility. The trade conflict, however, has a limited impact on U.S. domestic investment. Conversely, the Chinese economy is more vulnerable to trade policy shocks and uncertainty.

In the past decade, thanks to revolutions in computer science, engineering, and geography, data compiled by sensors on satellites have become publicly accessible for researchers. The satellite night lights data have been increasingly used by social scientists as an alternative measure of economic activity. In the book's final chapter, Adam Storeygard highlights six key advantages of using satellite data for economic research and policymaking. These include 1) providing data for data-poor contexts, 2) high spatial resolution, 3) low-cost repeat measurements, 4) data available for the whole world, 5) consistency across borders of different systems, and 6) avoiding possible data manipulations by traditional data collectors. Storeygard provides examples of research on deforestation, pollution, urban growth, transportation, and political economy. Although not a complete substitute for traditional administrative or survey data, the satellite night lights technology holds great promise as the cost of obtaining these data goes down and the algorithms for analyzing them keep advancing.

All the economic indicators presented in the lecture series are publicly available. They are so well accepted that researchers around the world have been adopting the presented ideas and methodologies to build comparable indicators for many countries. Looking ahead, we expect emerging technologies in big data platforms and artificial intelligence to further advance the research in how data are collected and analyzed, which should lead to more innovative and informative economic indicators, resulting in better policymaking and business decisions.

2
Getting It Wrong

How Faulty Monetary Statistics Undermine the Fed, the Financial System, and the Economy

William A. Barnett
University of Kansas and Center for Financial Stability

Kun He
University of Kansas

Barnett (2012) documented the degree to which faulty monetary statistics have tended to undermine the Federal Reserve System (the Fed), the financial system, and the economy. That MIT Press book, which brings together nearly a half century of research on that subject, won a Professional and Scholarly Excellence (PROSE) Award for the best book published in economics during 2012, presented by the Association of American Publishers. The research in the book is primarily based on the use of the Divisia monetary aggregates, originated by Barnett (1980) and made available to the public by the Center for Financial Stability (CFS) in New York City. But newer, more sophisticated monetary aggregates are now available from the CFS. The new aggregates incorporate credit card services into the Divisia monetary aggregates, and they distinguish between the demand-side total monetary services consumed and the supply-side inside-money services associated with value added in financial intermediation. This chapter begins the process of updating Barnett (2012) to use the newer data, but with the need for more sophisticated econometric tests in the frequency domain.

Supply-side inside-money aggregates and demand-side total monetary aggregates are not equal, since total demand-side monetary aggregates include outside money not produced as outputs of private financial intermediaries. As economic indicators, they may perform differently in the short run and in the long run. Divisia monetary aggregates, on the demand side or supply side, can be expected to perform even better when credit-card transaction services are taken into account. In this chapter, we empirically compare credit card–augmented inside-money supply-side Divisia aggregates and total-money demand-side Divisia aggregates. In particular, we compare their correlations with major economic policy targets in the short term and long term. To acquire dynamic performances for time-series data at different frequencies, we transform their time series into the frequency domain using spectral analysis methods. Spectral coherence between the Divisia indexes and major final targets of policy at different frequencies can provide evidence of the role of inside-money supply-side Divisia and total-money demand-side Divisia in the short run and long run.

The original Divisia monetary aggregates measure demand-side monetary services using the economic aggregation and index number theory developed by Barnett (1980). The data are available from the Center for Financial Stability (CFS) in New York City. On the demand side, there is no reason to differentiate among inside money, outside money, regulated services, or shadow banking services. Demanders consume liquidity services supplied by all relevant sources. On the supply side, the manner in which the monetary services are produced is highly relevant to the transmission mechanism of monetary policy and to the indicator value of the resulting service flows.

On the supply side, traditionally, outside money has been measured as the monetary base, supplied by the Federal Reserve as the sum of currency and bank reserves. Inside money has been calculated as the difference between the total-money supply, measured as a simple sum, and outside money. In recent years, that measure of inside money has become conspicuously defective, with M1 inside money often being negative, despite the fact that most of the monetary services in the economy are now produced by private banks as value added in banking and, hence, properly representing inside money.

In recent decades, transaction and liquidity services have been augmented dramatically by the growth of privately supplied unregulated monetary services from bank-supplied credit cards and from the services provided by unregulated shadow banking. We consider inside money using aggregation and index number theory, not simple-sum accounting, and we augment our aggregates with credit-card service flows. We believe that the relationship between inside-money services on the supply side, total monetary services on the demand side, and final targets of policy can differ at different frequencies, since the transmission mechanism behaves differently in the short run from the long run.

Exploring those extensions of Barnett (2012) would best be done using harmonic analysis. As a first step in that direction, we investigate the properties of the data in the frequency domain using spectral analysis. But that approach, while being the appropriate first step in the intended direction, requires stationary data, which lose relevant information about the dynamics of the economy and of the monetary transmission mechanism. In addition, that approach is heavily sample-size dependent. In subsequent research, we plan to extend this approach to the time-frequency domain using wavelets, in accordance with the approach in Barnett, Ftiti, and Jawadi (2019). This chapter contains our first steps in that direction.

CREDIT CARD–AUGMENTED DIVISIA

Using accounting conventions, credit cards cannot be aggregated with monetary assets, since monetary assets are assets and credit-card balances are liabilities. Accounting conventions do not permit adding liabilities to assets. But aggregation and index number theory aggregate over service flows, regardless of whether produced by assets or liabilities. As shown by Barnett and Su (2020), services of credit cards and of monetary assets can be aggregated using aggregation and index number theory.

These are the definitions of variables used in Barnett and Su's (2020) model:

R_t = expected yield on the benchmark asset, representing the rate of return on pure capital;

$\boldsymbol{\mu}_t$ = vector of real balances, μ_{it}, of monetary asset deposit-account type i during period t;

$\boldsymbol{\tau}_t$ = vector of real expenditure volumes, τ_{jt}, with credit-card type j for transactions during period t;

$\mathbf{e}_t =$ vector of expected interest rates, e_{it}, on $\boldsymbol{\tau}_\mathbf{t}$;

$\varsigma_t =$ vector of rotating real balances, ς_{jt}, in credit-card type j during period t from transactions in previous periods;

$\bar{\mathbf{e}}_\mathbf{t} =$ vector of expected interest rates on $\boldsymbol{\varsigma}_\mathbf{t}$;

$c_t =$ real balances of excess reserves held by the intermediary during period t;

$\mathbf{L}_\mathbf{t} =$ vector of labor quantities receiving expected wage rates, ω_t, during period t;

$z_t =$ quantities of other factors of production;

$c_t =$ price of the factor z_t;

$\mathbf{k}_t =$ reserve requirements, where k_{it} is the reserve requirement applicable to μ_{it}, and $0 \leq k_{it} \leq 1$ for all I;

$R_t^d =$ Federal Reserve expected discount rate;

$\overline{R}_t = \min\{R_t, R_t^d\}$;

$\gamma_t =$ vector of expected yields paid by the firm on i_t ; and

$p_t^* =$ true cost of living index, used to deflate nominal balances to real balances.

The vector $\boldsymbol{\gamma}_t$ is defined so that the nominal user-cost price for produced monetary asset i_t is

$$\gamma_{it} = p_t^* \frac{(1-k_{it})R_t - \rho_{it}}{1+R_t} .$$

The vector $\tilde{\pi}$ is defined so that the nominal expected user-cost price for produced credit card services, τ_{jt} , is

$$\tilde{\pi}_{jt} = p_t^* \frac{e_{jt} - R_t}{1+R_t} .$$

The vector $\boldsymbol{\sigma}_t$ is defined so that the nominal expected user-cost price for carried-forward rotating credit card debt, ς_{jt}, is

$$\sigma_{jt} = p_t^* \frac{e_{jt} - R_t}{1 + R_t} \ .$$

The nominal expected user-cost price of excess reserves, c_t, is

$$\gamma_{0t} = p_t^* \frac{R_t}{1 + R_t} \ .$$

The corresponding expected real user costs are

$$\frac{\boldsymbol{\gamma}_t}{p_t^*}, \ \frac{\tilde{\boldsymbol{\pi}}_t}{p_t^*}, \ \frac{\boldsymbol{\sigma}_t}{p_t^*}, \text{ and } \frac{\gamma_{0t}}{p_t^*} .$$

Based on the aggregator function existence assumption of technology weakly separable in produced monetary asset service, the Divisia money index for produced inside-money services is acquired by solving the financial intermediary's decision problem. The result is

$$\log M_t^* - \log M_{t-1}^* = \sum_i \overline{s}_{it} (\log \mu_{it}^* - \log \mu_{i,t-1}^*) + \sum_i \overline{\mu}_{jt} (\log \tau_{jt}^* - \log \tau_{j,t-1}^*) ,$$

where M_t^*= the economic output quantity aggregate for financial firms.

Here, $\overline{s}_{it} = \frac{1}{2}\left(s_{it} + s_{i,t-1}\right)$, with s_{it} and u_{jt} computed from

$s_{it} = \mu_{it}^* \gamma_{it} / \left(\boldsymbol{\mu}_t^* \boldsymbol{\gamma}_t + \boldsymbol{\gamma}_t^{*'} \tilde{\boldsymbol{\pi}}_t\right)$, and u_{jt} is the solution to the constrained revenue maximizing problem:

$$\text{Max } \boldsymbol{\mu}_t^{*'} \boldsymbol{\gamma}_t + \boldsymbol{\gamma}_t^{*'} \tilde{\boldsymbol{\pi}}_t \text{ subject to } f\left(\boldsymbol{\mu}_t, \boldsymbol{\tau}_t, \boldsymbol{k}_t\right) = M_t^* .$$

Unlike conventional accounting inside money, the CFS credit card–augmented Divisia inside-money aggregates correlate very well with nominal GDP and can serve the central purposes of inside money, long contemplated in the literature on monetary economics. Further knowledge of its properties remains to be discovered and explored in the frequency domain.

The primary differences between the supply-side measure and the CFS demand-side Divisia monetary aggregates is the supply side's inclusion of credit card services and exclusion of currency and Treasury bills.

Another difference between the demand-side and the supply-side user-cost formulas for monetary asset services results from the existence of reserve requirements, producing an implicit tax on banks. But in recent years, that tax has been nearly zero because of sweeps, low interest rates, and Federal Reserve payment of interest on reserves.

In this chapter, we begin our empirical exploration of the inside-money and total credit card–augmented Divisia for broad M4 and narrow M1, beginning from July 2007. Moving from DM1 (Divisia Ma) to the higher levels of aggregation incorporates increasing amounts of shadow banking and negotiable money-market security liquidity services, properly weighted.

SPECTRAL ANALYSIS THEORY

For a finite series $u(j)$ of length $T = N\Delta t$, here with N referring to the sample size and Δt referring to the sampling periodicity, the discrete Fourier transform (DFT) $U(k)$ of $u(j)$ and its inverse (IDFT) for finite series (see, e.g., Iacobucci [2005]) are

$$U(k) = \frac{1}{N}\sum_{j=0}^{N-1} u(j)\, e^{-i2\pi jk/N}$$

and

$$u(j) = \sum_{k=-[N/2]}^{[(N-1)/2]} U(k)\, e^{i2\pi jk/N},$$

where $u_k = \dfrac{k}{N\Delta t}$ is the frequency and $T = N\Delta t$ is the time. In our power spectrum for real data in later parts of the paper, the label for frequency domain is v_k, and the period should be $1/v_k$. The power spectrum is given by

$$P_u(k) = |U(k)|^2 .$$

An estimator for the power spectrum is given by Schuster's Periodogram:

$$P_u(k) = \Delta t \sum_{J=-(N-1)}^{N-1} \gamma_{uu}(J) \cos\frac{2\pi Jk}{N},$$

where $\gamma_{uu}(J) = \gamma_{uu}(-J) = N^{-1} \sum_{j=-(N-J)}^{N-J} \left(u(j) - \bar{u}\right)\left(u(j+J) - \bar{u}\right)$ is the standard sample estimation at lag J of the autocovariance function.

To build a spectral estimator, which is more stable—i.e., it has a smaller variance than P_u(k)—we turn to the technique of windowing. This technique is employed both in time and in frequency domain to smooth all abrupt variations and to minimize the spurious fluctuations generated in time as a series is truncated. The smoothed spectrum is given by

$$\hat{S}_u(k) = \Delta t \sum_{J=-(N-1)}^{N-1} \omega_M(J) \gamma_{uu}(J) \cos\frac{2\pi Jk}{N},$$

where the autocovariance function is weighted by the lag window $\omega(j)$ of width M. It can be shown that this is equivalent to splitting the series in N/M subseries of length M, computing their spectra, and taking their mean with the spectral window $W_M(k)$ of width $M = M - 1$.

For two time series, $u_1(j)$ and $u_2(j)$, with cross covariance $\gamma_{12}(J) = \gamma_{12}(-J)$, the cross spectrum is

$$\hat{S}_{12}(k) = \Delta t \sum_{J=-(N-1)}^{N-1} \omega(J) \gamma_{12}(J) e^{-i2\pi Jk/N} = \hat{C}_{12}(k) - i\hat{Q}_{12}(k).$$

Here, the real part $\hat{C}_{12}(j)$ is the cospectrum, and the imaginary part is the $i\hat{Q}_{12}(j)$ quadrature spectrum. The coherency spectrum (correlation coefficient) is

$$\hat{K}_{12}(k) = \frac{\left|\hat{S}_{12}(k)\right|}{\sqrt{\hat{S}_1(k)\hat{S}_1(k)}} = \frac{\sqrt{\hat{C}_{12}(k)^2 + \hat{Q}_{12}(k)^2}}{\sqrt{\hat{S}_1(k)\hat{S}_1(k)}}.$$

The phase spectrum (time lag) is

$$\hat{\Phi}_{12}(k) = \arctan(-\frac{\hat{Q}_{12}(k)}{\hat{C}_{12}(k)}),$$

which measures the phase difference between the frequency components of the two series: 1) the number of leads (> 0) or 2) the number of lags (< 0) of $u_1(j)$ on $u_2(j)$.

DATA

Regarding the data sources, see Barnett and Su (2020). The credit card transaction services can be measured by the transaction volumes summed over four sources: Visa, Mastercard, American Express, and Discover. The credit-card-augmented Divisia aggregate does not apply to debit cards, nor to store cards, nor to charge cards not providing a line of credit. The model identifies credit card services as sources of value added in banking and therefore outputs of financial intermediation, since those credit card accounts provide deferred payment services. Cash and checking accounts do not provide that service. Debit cards do not, either. The services of debit cards are similar to the services of checking accounts, which are already included as services of demand deposit accounts but are not the source of value added we identify as credit card services.

Store cards are not outputs of financial intermediation, since they are maintained by the stores that supply the purchased products. In addition, the connection between store cards and those products sold by the stores is inconsistent with the assumption of blockwise weak separability of financial services and consumer goods on the demand side, since these cards can be used only to purchase the goods sold by the store. Charge cards that do not provide a line of credit are rarely provided by banks, and they are now largely limited to store cards.

SPECTRAL ANALYSIS RESULTS

The year-over-year growth rate for credit card–augmented inside-money and total-money Divisia are provided on the website of the Center of Financial Stability(CFS), dated from July 2007 to October 2018; the U.S. unemployment rate and CPI (consumer price index) are provided by the Bureau of Labor Statistics; and the U.S. inflation rate, along with inflation rates internationally, is provided by statbureau.org from July 2007 to April 2018. The total sample size is $N = 136$, but 131 for the inflation rate.

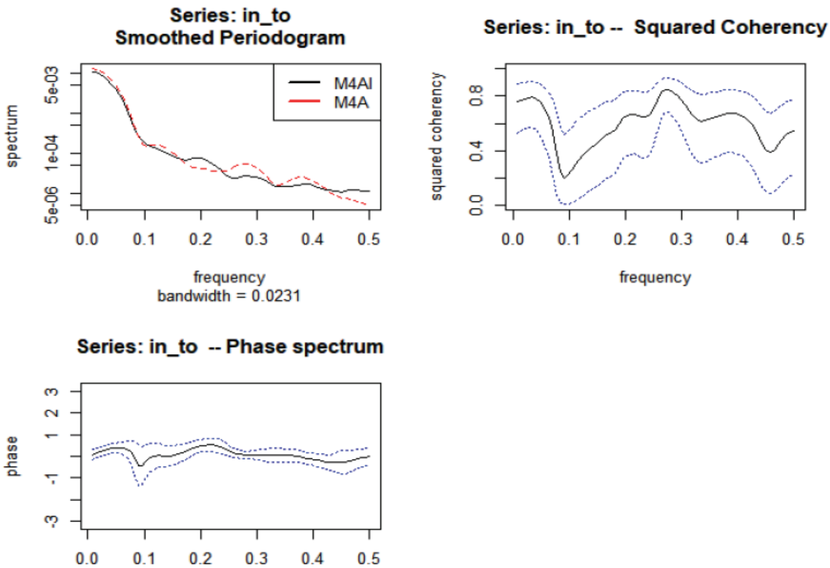
All data are monthly data, corresponding to periodicity of Δt = one month. All time series data were detrended when spectrum estimated. Here we chose modified Daniell smoothers as the smoothing function, with moving averages giving half weight to the end values. The smooth width $M = 8$ determines the trade-off between bias and variance for a fixed sample size. The larger the value of M, the smaller the variance of the estimated spectrum at a given frequency, but the larger the bias. To get a smoothed estimated spectrum without losing excessive information, we set $M = 8$.

Since the original value of the year-over-year growth rate of Divisia index is small, the power spectrum remains small after estimation. However, the periodic properties for coherence between inside-money Divisia and total-money Divisia, and with unemployment rate, inflation rate, and CPI index, are clear.

In the plots below, we have sample size $N = 136$ with an 11-year time range, from 2007 to 2018. Frequency domain results, with the frequency set at $v_k = 0.1, 0.2, 0.3, 0.4, 0.5$, correspond to periods of $1/v_k = 10.0, 5.0, 3.3, 2.5$, and 2.0 months, respectively. Although there is high correlation between inside-money and total-money Divisia, their behavior displayed differently at low frequencies over a long period. (See frequency of less than 0.1 with period exceeding 10 months.) Total demand-side money has a high coherency with the main economic indicators.

In the plots for the broad M4 level aggregates, M4AI denotes the inside-money Divisia M4 augmented with credit card services. M4A denotes the total Divisia M4 augmented by credit card services. Figure 2.1 displays the relationship between the total Divisia demand-side aggregate and the supply-side inside-money Divisia aggregate. The first plot displays their power spectrum. The second plot provides the squared coherency, measuring correlation between the two aggregates at different frequencies. The blue dashed lines above and below the coherency plot provide the 95 percent confidence band around the coherency plot. The third plot provides the phase spectrum and its confidence region.

Figure 2.1 Inside DM4 (supply-side) and Total DM4 (demand-side)



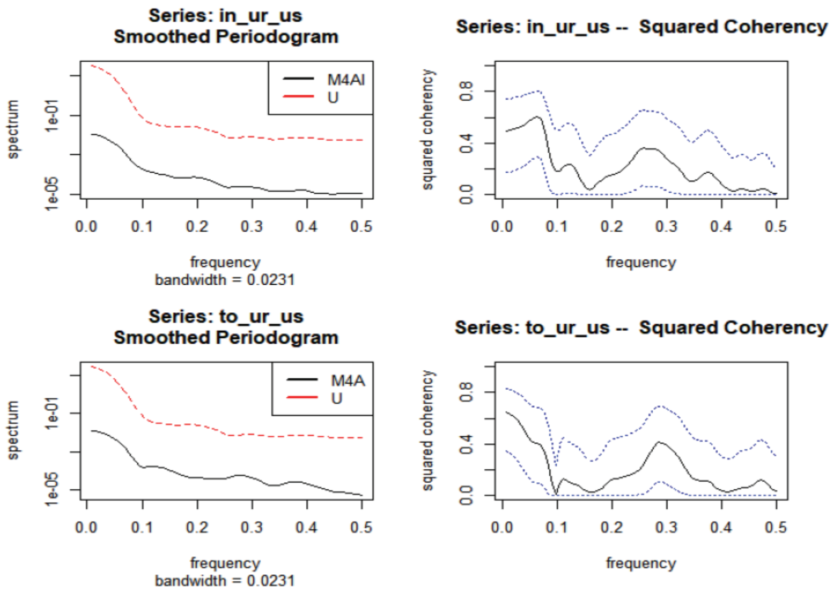
SOURCE: Center for Financial Stability.

Figure 2.2 provides the analogous results relating the broad monetary aggregates to the unemployment rate, while Figure 2.3 provides the results relating the monetary aggregates to the inflation rate and Figure 2.4 provides the results relating the aggregates to the CPI level at different frequencies.

In Figures 2.5–2.8, we similarly consider the Divisia index for the narrow M1 aggregate. Moving from M4 to the lower levels of aggregation incorporates decreasing amounts of shadow banking and negotiable money-market security liquidity services. The periodic behavior differences become less significant.

In Figure 2.9, we explore the relationship between unemployment and inflation and thereby the frequency properties of the Phillips curve. The cross correlation, ACF (auto-correlation function) in Figure 2.9, between the unemployment rate and inflation rate is displayed under different numbers of lag. Since the sampling periodicity is monthly, the correlation will be significantly positive only when the lag or lead

Figure 2.2 DM4 and Unemployment Rate



SOURCE: Bureau of Labor Statistics and Center for Financial Stability.

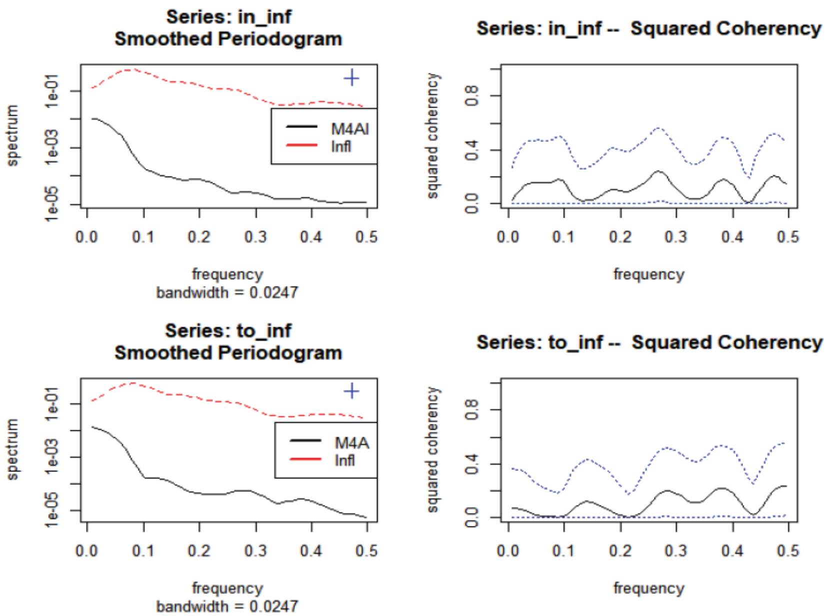
between the two indicators is more than 12 months. Also, there are phase differences under different frequencies or periods. As a result, it is not surprising that Divisia growth rates have different coherences with these two indicators.

PREVIOUS RESULTS

As this research advances, it will be relevant to compare with prior results that appeared in Barnett (2012) and Barnett and Chauvet (2011), but with the newer augmented aggregates now available from the CFS. Examples from the earlier research include the following figures.

Figure 2.10 displays the broadest Divisia monetary aggregate available from the St. Louis Federal Reserve Bank over a period of 40 years. The figure clearly displays the aggregate's correlation with the business

Figure 2.3 DM4 and Inflation Rate

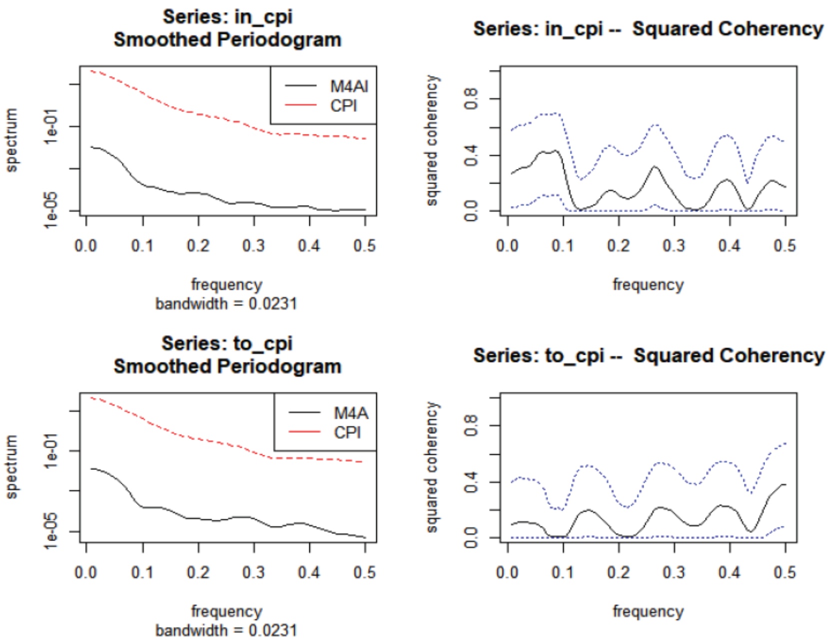


SOURCE: Bureau of Labor Statistics and Center for Financial Stability.

cycle and its predictive ability relative to the Great Recession, which began immediately after the end of that figure's time period.

Figure 2.11 displays M1 inside money computed in the conventional manner as total-money supply minus outside money. The M1 aggregate used in that computation is the Federal Reserve Board's measure using simple sum aggregation without sweep adjustment. The only available measure of outside money provided by the Federal Reserve is the monetary base. In that figure, the monetary base was acquired from the St. Louis Federal Reserve Bank's FRED database. Observe that inside money, by that measure, became negative during a period of time when most monetary services in the economy were provided as inside money by privately owned banks and other privately owned financial intermediaries. The error has two sources: 1) simple-sum M1 is biased downward by the Federal Reserve's failure to sweep adjust its component data, and 2) the Federal Reserve's measure of the mon-

Figure 2.4 DM4 and CPI



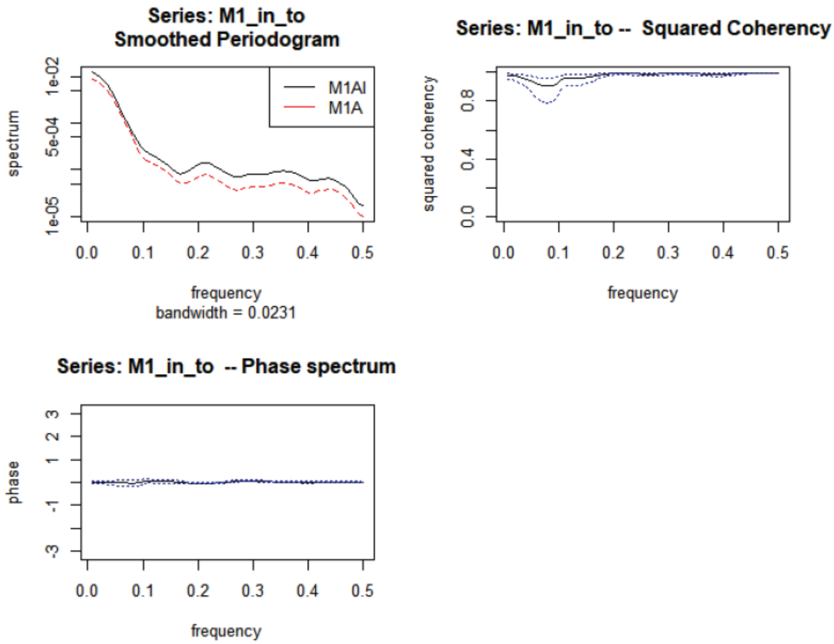
SOURCE: Bureau of Labor Statistics and Center for Financial Stability.

etary base has become an upwardly biased measure of outside money in recent years as a result of the Federal Reserve's nonstandard policies.

Figure 2.12 displays nonborrowed reserves as reported by the Federal Reserve Board. Nonborrowed reserves were the instrument of monetary policy adopted by Fed chairman Paul Volcker during the three-year period of the "monetarist experiment." Observe the period during which nonborrowed reserves became negative. That result is a contradiction in terms, since borrowed reserves, by definition, cannot exceed total reserves. The Federal Reserve's accounting error, producing that impossible result, occurs because it is including within borrowed reserves some bank borrowing not held as reserves.

Figure 2.13 displays the results of the Taylor rule, as provided by the St. Louis Fed's FRED database. The figure shows the target range for the Taylor rule and the actual path of the federal funds rate. Clearly

Figure 2.5 Inside DM1 (supply-side) and Total DM1 (demand-side)



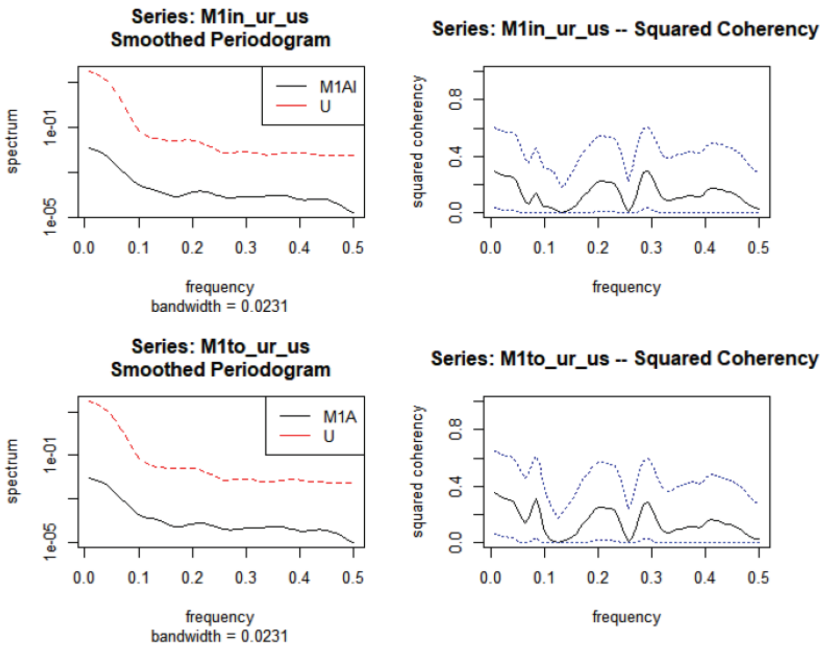
SOURCE: Center for Financial Stability.

the federal funds rate was below the target range for three successive years, casting doubt on the policy relevance of the interest rate target.

CONCLUSION

In this chapter, we begin our research on updating the results in Barnett (2012) to use the more sophisticated Divisia monetary aggregates recently available from the Center for Financial Stability. Those aggregates are extended to include credit card services and to distinguish demand-side total consumed monetary services from supply-side inside monetary services associated with value added in banking. Since

Figure 2.6 DM1 and Unemployment Rate

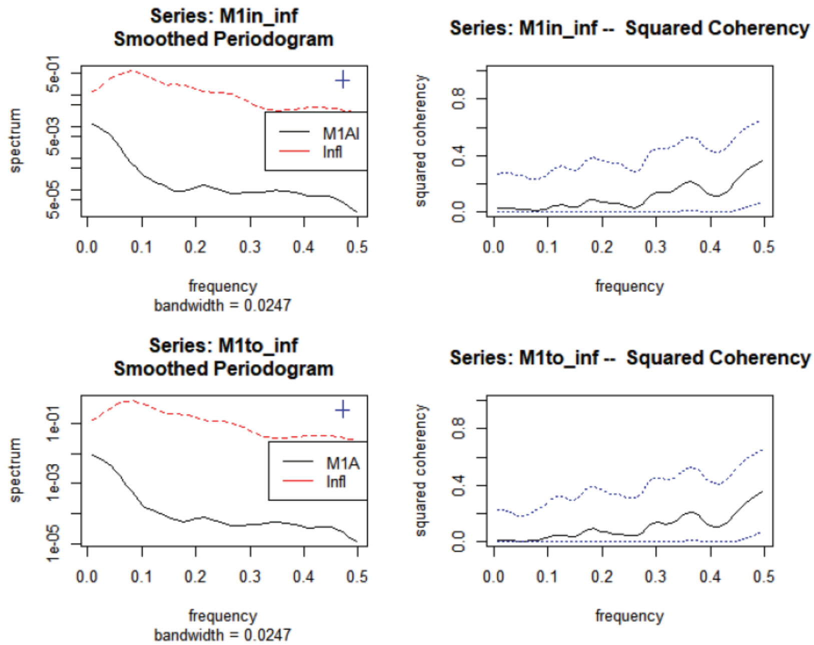


SOURCE: Bureau of Labor Statistics and Center for Financial Stability.

the transmission mechanism has lags resulting in different correlations with final targets in the long run versus the short run, we introduce into this literature tests in the frequency domain.

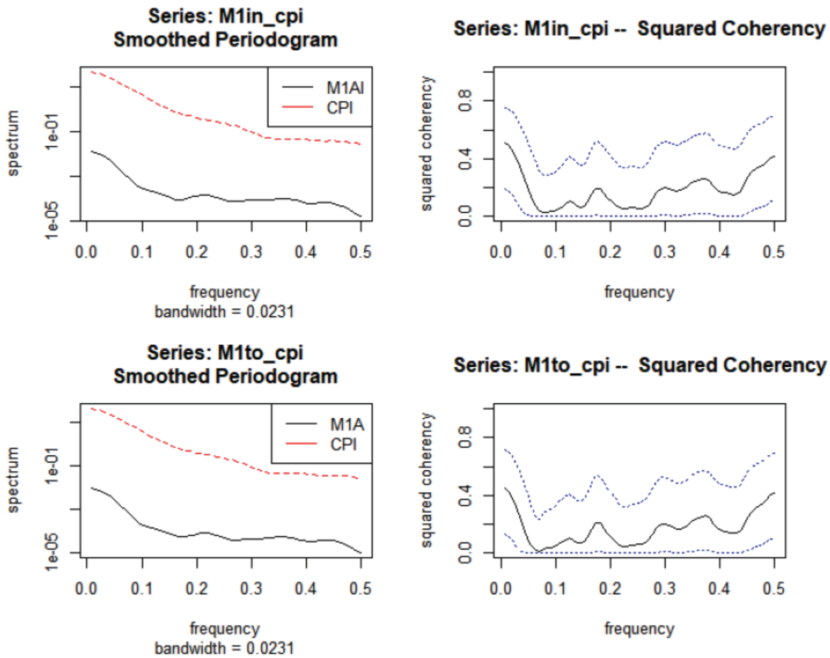
To acquire dynamic performances for time series data at different frequencies, we transform the time series into the frequency domain using spectral analysis methods. As the sample size becomes larger, more significant results will become available from data covering a complete business cycle. Although this approach is an appropriate first step in this direction, conversion to the frequency domain requires stationarity. However, such stationary data lose much relevant information about the economy. In subsequent research, we shall investigate nonstationary data with wavelet methodology in the time-frequency domain, following the approach of Barnett, Ftiti, and Jawadi (2019).

Figure 2.7 DM1 and Inflation Rate



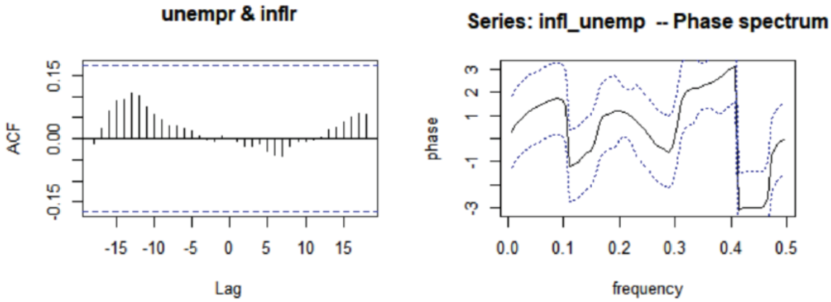
SOURCE: Bureau of Labor Statistics and Center for Financial Stability.

Figure 2.8 DM1 and CPI



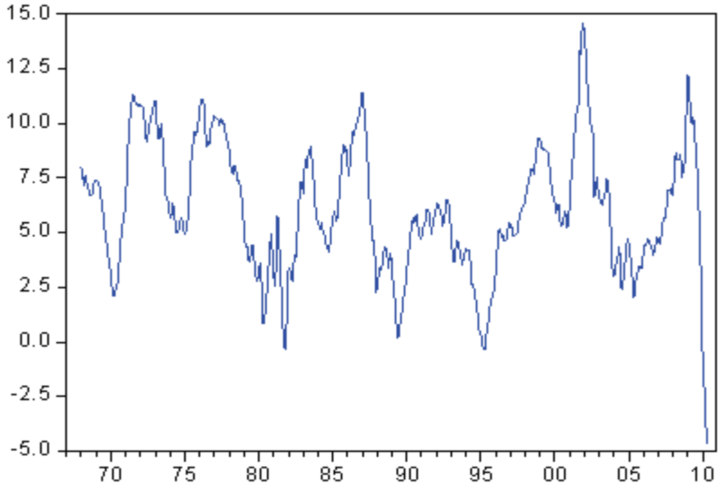
SOURCE: Bureau of Labor Statistics and Center for Financial Stability.

Figure 2.9 Possible Explanation of Phillips Curve



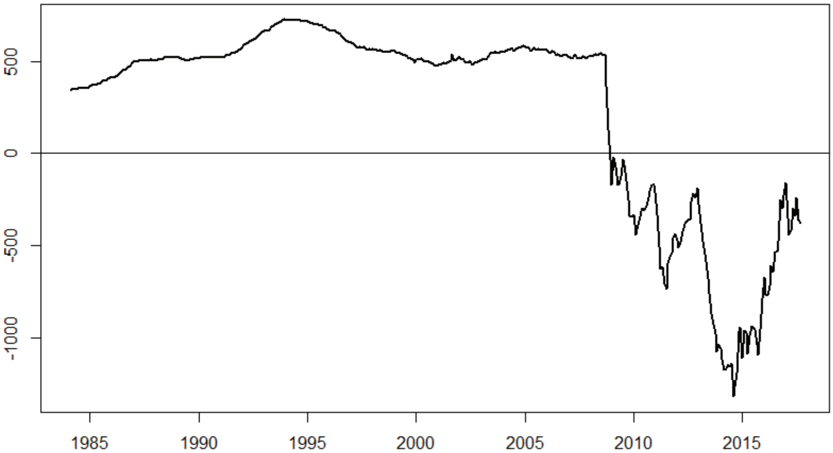
SOURCE: Bureau of Labor Statistics and Center for Financial Stability.

Figure 2.10 Year-over-Year Growth Rates of the Broadest Available Divisia Monetary Aggregate during 40 Years



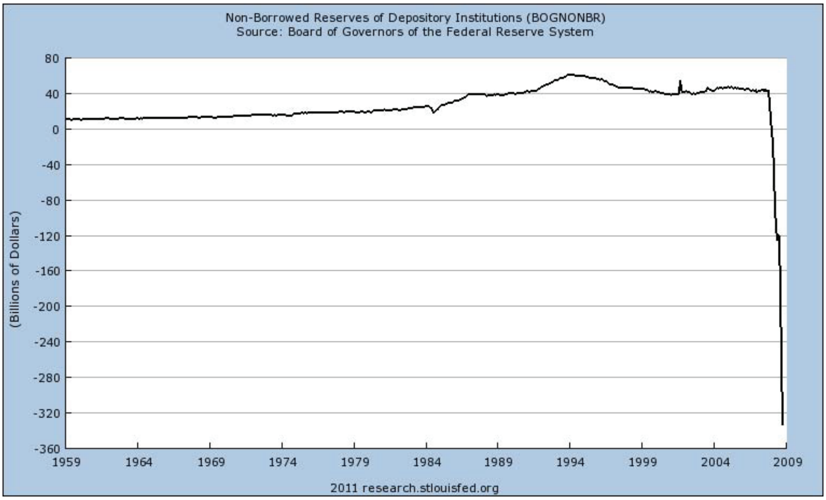
SOURCE: Barnett (2012).

Figure 2.11 M1 Inside Money, Computed as Federal Reserve Simple Sum M1 Minus St. Louis Federal Reserve Bank Monetary Base (outside money)



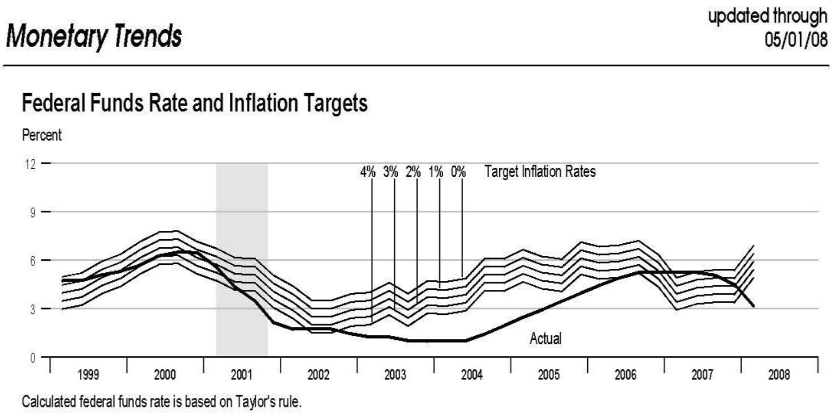
SOURCE: St. Louis Federal Reserve Bank's FRED Database.

Figure 2.12 Nonborrowed Reserves



SOURCE: Board of Governors of the Federal Reserve System.

Figure 2.13 Taylor Rule Federal Funds Rate



SOURCE: St. Louis Federal Reserve Bank's FRED Database. Reprinted from Barnett and Chauvet (2011).

References

Barnett, William A. 1980. "Economic Monetary Aggregates: An Application of Aggregation and Index Number and Aggregation Theory." *Journal of Econometrics* 14(1): 11–48.

———. 2012. *Getting It Wrong: How Faulty Monetary Statistics Undermine the Fed, the Financial System, and the Economy*. Cambridge, MA: MIT Press.

Barnett, William A., and Marcelle Chauvet. 2011. "Financial Aggregation and Index Number Theory." In *Surveys on Theories in Economics and Business Administration*, Vol. 2, Kar-yiu Wong, ed. Singapore: World Scientific, pp. 1–51.

Barnett, William A., Zied Ftiti, and Fredj Jawadi. 2019. "The Causal Relationships between Inflation and Inflation Uncertainty." University of Kansas Department of Economics working paper. Lawrence: University of Kansas.

Barnett, William A., and Liting Su. 2020. "Financial Firm Production of Inside Monetary and Credit Card Services: An Aggregation Theoretic Approach." *Macroeconomic Dynamics* 24(1): 130–160.

Iacobucci, Alessandra. 2005. "Spectral Analysis for Economic Time Series." In *Lecture Notes in Economics and Mathematical Systems*, Vol. 551: *New Tools of Economic Dynamics*, Jacek Leśkow, Martin Puchet, and Lionello F. Punzo, eds. Berlin: Springer, pp. 203–219.

3
A Closer Look at the Chicago Fed's Activity Indexes

Scott A. Brave
Federal Reserve Bank of Chicago

How does one go about summarizing the state of the U.S. economy? In the age of "big data," this may seem like a strange question to ask, but it is no less relevant today than it was when the National Income and Product Accounts were first developed in the early twentieth century. If anything, it may arguably be a more difficult question now than it was then, given the multitude of economic statistics produced by both government statistical agencies and private firms. While it remains common to characterize the health of the U.S. economy in terms of broad macroeconomic aggregates like gross domestic product (GDP), other measures are often used as well in order to capture the state of individual sectors of the economy or as potential indicators of the future direction of growth in GDP.

With so many indicators available to economic and financial analysts, using them effectively becomes a question of how best to make use of their common strengths while minimizing their individual weaknesses. Activity indexes are designed for just such a purpose. As an example of what is referred to as a *dense* modeling approach in statistics (Giannone, Lenza, and Primiceri 2018), these indexes aim to extract as much information on the overall state of the U.S. economy as they can, and to do it as efficiently as possible, while using all of the available data. In principle, this approach acknowledges that all of the available indicators might be important for measuring the health of the U.S. economy, despite their own individual influence potentially being small.

At the Federal Reserve Bank of Chicago (the Chicago Fed), we produce two types of activity indexes: 1) economic and 2) financial activity indexes. The former characterize business conditions in the U.S. economy at various levels of geographic detail, while the latter capture credit conditions in the financial sector, broadly considered. Both types of indexes are predicated on a common statistical framework—namely the Chamberlain and Rothschild (1983) approximate factor model, as discussed in the next section. While their estimation methods vary, both types of indexes rely on popular data-dimension reduction techniques such as principal components and dynamic factor analysis (Stock and Watson 2011).

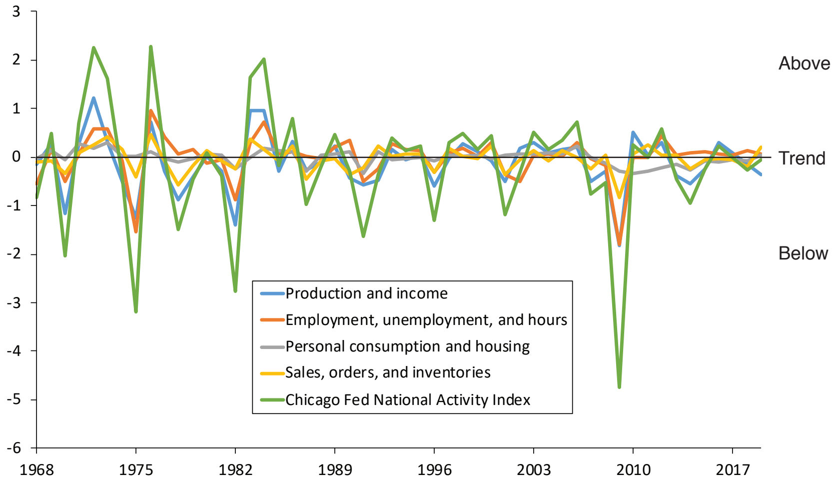
As an example of each type of index, Figures 3.1 and 3.2 plot the recent histories for the Chicago Fed National Activity Index (CFNAI) and the National Financial Conditions Index (NFCI), respectively. Based on the work of Stock and Watson (1999), the CFNAI was originally developed to help forecast inflation, but over time it has come to be viewed as a measure of the U.S. business cycle (Evans, Liu, and Pham-Kanter 2002). Positive values of the index are interpreted as representing above-trend economic growth; negative values as representing below-trend growth. The index is shown in standard deviation units based on a history extending back to early 1967. The section titled "The CFNAI" chronicles the nearly 20-year history of the production of this index as well as its offshoots and recent extensions.

The NFCI, in contrast, was developed more recently from research conducted during the global financial crisis. It aims at measuring the overall tightness of the U.S. financial system (Brave and Butters 2011). An increase in the NFCI implies an increase in *risk* or a decrease in *credit growth* or *leverage* in financial markets. Positive (negative) values denote tighter-than-average (looser-than-average) conditions in standard deviation units based on a history extending back to 1971. A separate index, the Adjusted National Financial Conditions Index (ANFCI), which rebenchmarks conditions relative to economic growth and inflation, is shown in Figure 3.2. The section titled "The NFCI" discusses some of the uses of the NFCI and ANFCI. Concluding remarks are offered in the final section.

THE APPROXIMATE FACTOR MODEL

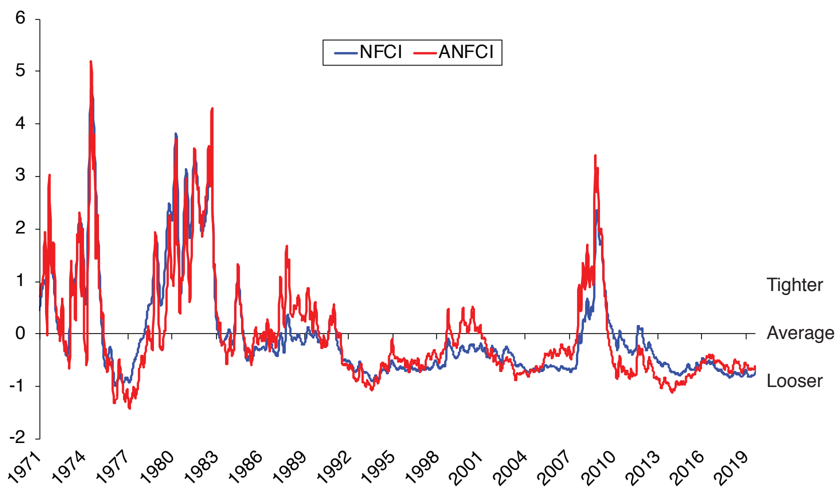
The Chamberlain and Rothschild (1983) approximate factor model has come to enjoy widespread use in economics and finance as a method

Figure 3.1 The Chicago Fed National Activity Index (CFNAI)



SOURCE: Federal Reserve Bank of Chicago (www.chicagofed.org/cfnai).

Figure 3.2 The National Financial Conditions Index (NFCI)



SOURCE: Federal Reserve Bank of Chicago (www.chicagofed.org/nfci).

to identify a small number of common components (i.e., *factors*, F_t) explaining the comovement of large panels of macroeconomic or financial time series, X_{it}. A commonly used parameterization of this model is shown in Equation 3.1,

$$(3.1) \quad \begin{aligned} X_{it} &= \Gamma_i F_t + \epsilon_{it} \\ \epsilon_{it} &\sim N(0,\sigma^2 I), \end{aligned}$$

where Γ_i are referred to as *factor loadings* for each time series *i*, and E_{it} represents the idiosyncratic variation in each time series that is uncorrelated with the factors. This framework can be used to capture, for example, sectors of the economy that vary together over the *business cycle* as well as financial markets that tend to tighten in concert over the *financial cycle*, with the single most important factor often serving as an *economic* or *financial activity index*, depending on the application.

The challenge faced by practitioners in applying this framework to construct activity indexes is that the econometrician does not typically observe the factors. Instead, latent variable estimators that can *extract* F_t up to a scale/sign rotation must be applied to X_{it}. In other words, one has to extract from the panel of time series the common *signal* in F_t from the *noise* of E_{it}. It is this feature of these estimators that was alluded to in the introduction as maximizing the common strengths of various economic and financial indicators while simultaneously minimizing their individual weaknesses in characterizing the state of the U.S. economy.

A technique commonly used for this purpose is *principal components analysis* (PCA). PCA can be viewed as a multidimensional restricted nonlinear least squares problem (Stock and Watson 2002a), e.g.,

$$(3.2) \quad \min_{\Gamma,F} \; V(\Gamma,F) = (NT)^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} (X_{it} - \Gamma_i F_t)^2 \quad s.t. \quad \frac{\Gamma'\Gamma}{N} = 1 \, .$$

Solving the error minimization problem above (in matrix form) for a single common factor produces an estimate of the activity index that is an intuitive *optimally weighted average* of large panels of time series. In other words, the index itself can be represented as a linear combination of the economic or financial time series that maximizes their *total variance explained*.

$$(3.3) \quad \textit{Activity Index:} \ \hat{F} = (\Gamma'\Gamma)^{-1}\Gamma' X * \frac{N}{N} = \frac{\hat{\Gamma}' X}{N} \ ;$$

$$\hat{\Gamma} \equiv \text{eigenvector associated with largest eigenvalue of } (X'X) \, .$$

For large panels of time series, PCA produces consistent estimates of the factors under general conditions (Bai and Ng 2002), and given its computational ease, it has become a standard for estimating the approximate factor model.

By extending the analysis of the approximate factor model along the time dimension, some of the restrictions implied by PCA can be relaxed using an alternative estimation technique called *dynamic factor analysis*. An example is given below:

$$(3.4) \quad \begin{aligned} X_{it} &= \Gamma_i F_t + \epsilon_{it}, \\ F_t &= \Phi F_{t-1} + \eta_t, \\ \epsilon_{it} &\sim N(0, \Sigma), \ \ Cov(\epsilon_{it}, \epsilon_{jt}) = 0 \ \ i = j, \\ \eta_t &\sim N(0, 1), \end{aligned}$$

where we now specify autoregressive dynamics (in companion form) for a single factor and allow for heteroskedasticity in the idiosyncratic errors of the panel.

Estimating the activity index in this case requires signal extraction methods for normal-linear state-space models that make use of the *Kalman filter* and routines for *maximum likelihood* estimation (Durbin and Koopman 2012). While we lose some of the simplicity of interpretation of the activity index by using this method versus PCA, we also gain the ability to directly forecast anything in the panel of time series. This feature has proven to be particularly attractive to researchers interested in forecasting the current state of the U.S. economy (Giannone, Reichlin, and Small 2008). Dynamic factor models can also be easily extended to handle common data irregularities, such as unbalanced panels and mixed frequencies of observation.[1]

The treatment of mixed-frequency data sets, in particular, is a strength of the state-space methods used to estimate dynamic factor models. For example, the practice of appending frequency-matching temporal aggregation constraints to the dynamic factor model (sometimes referred to as *accumulators*, as in Harvey [1989]) has been used to construct mixed-frequency indexes of both economic and financial activity for the United States (e.g., Aruoba, Diebold, and Scotti 2009; Brave and Butters 2012b; Mariano and Murasawa 2003). While these extensions are not commonly found in standard statistical software packages, their use is becoming more widespread. For further information, see Brave, Butters, and Kelley (2019), which describes the Matlab toolbox package MFSS.

Recent research has also developed computationally efficient methods that make the estimation of dynamic factor models feasible for large panels of time series. These include quasi maximum likelihood routines such as expectation-maximization (EM) algorithms (Bańbura and Modugno 2014; Doz, Giannone, and Reichlin 2012) as well as collapsing transformations that can simplify maximum likelihood estimation. An example of the latter can be found in Bräuning and Koopman (2014). Referred to as collapsed dynamic factor analysis, their application can be viewed as a hybrid case in which principal components are construed as observations of the latent factors up to the inclusion of classical measurement errors. The Chicago Fed's activity indexes make use of both PCA and dynamic factor estimation methods. In the sections that follow, we describe these indexes and summarize some of their applications.

THE CFNAI

The Chicago Fed National Activity Index (CFNAI) is a monthly summary statistic for U.S. economic growth. Estimated by PCA, it is the first principal component of 85 monthly indicators covering four broad categories of economic activity: 1) production and income; 2) employment, unemployment, and hours; 3) personal consumption and housing; and 4) sales, orders, and inventories. Many of the most commonly cited economic indicators for the United States fall within these categories, including industrial production, payroll employment, the unemployment rate, personal consumption expenditures, housing starts, and manufacturing and trade sales.

First introduced in Evans, Liu, and Pham-Kanter (2002), the CFNAI derived largely from earlier work examining the forecasting ability of real economic activity indicators for U.S. inflation (e.g., Fisher 2000; Stock and Watson 1999). Today, however, it is primarily seen as a coincident indicator of the U.S. business cycle, as this use of the index formed much of the motivation for its initial release at the onset of the 2001 recession, as well as much of the subsequent work with the index during and after the 2007–2009 recession (e.g., Brave 2009; Brave and Lichtenstein 2012).

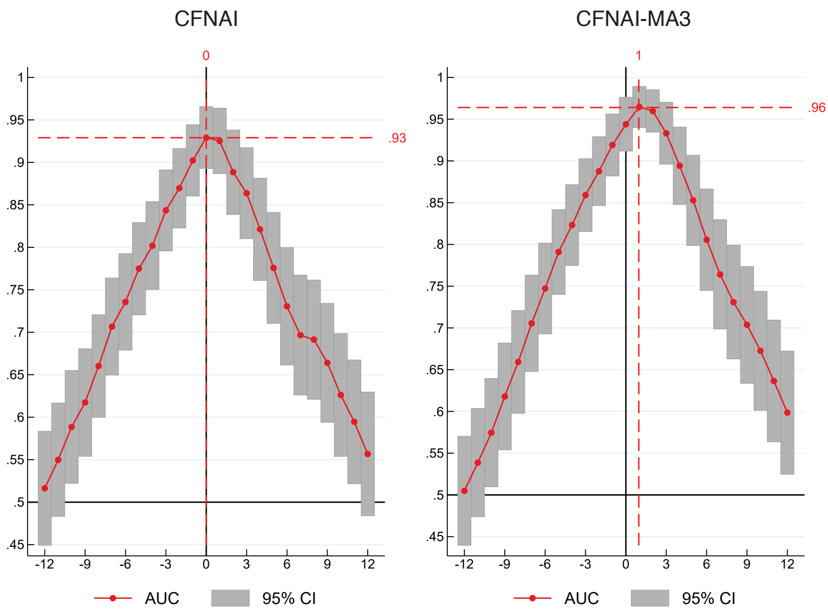
The CFNAI's performance in this regard has been quite good. For example, the index has been shown to be roughly 95 percent accurate historically in identifying U.S. recessions from expansions since 1967 based on a receiver operating characteristic (ROC) analysis of U.S. business cycles (Berge and Jordà 2011). This classification technique nonparametrically captures the trade-off between type I (false positive) and type II (false negative) errors based on the observed distribution of an indicator, C_t. To measure its accuracy for U.S. business cycles, the *ROC curve* is constructed over the range of realizations of C_t by applying the Cartesian convention $\{ROC(r), r\}$, in which $ROC(r) = TP(c)$ and $r = FP(c)$ and defining

$$(3.5) \quad \begin{aligned} TP(c) &= P[C_t \geq c | S_t = 1], \\ FP(c) &= P[C_t \geq c | S_t = 0], \end{aligned}$$

where S_t is a binary variable, with $S_t = 1$ representing a U.S. recession and $S_t = 0$ representing an expansion. The *area under the curve* (AUC) then represents C_t's accuracy in separating U.S. recessions from expansions.

Figure 3.3 depicts AUC values (red connected dots) for the CFNAI and its three-month moving average (CFNAI-MA3) at leads (negative x-axis values) and lags (positive x-axis values) in months over the U.S. business cycle. The dashed red horizontal lines in each panel correspond to the peak AUC value for each measure, while the gray bars are 95 percent confidence intervals. An AUC value statistically significant from 0.5 reflects an indicator that exhibits a significant ability to appro-

Figure 3.3 AUCs at Monthly Leads and Lags of the Business Cycle



SOURCE: Author's calculations based on data available at www.chicagofed.org/cfnai.

priately classify recessions and expansions as defined by the National Bureau of Economic Research (NBER).[2] The closer an AUC value is to 1, the more accurate the indicator. The closer an indicator's peak AUC value is to a zero monthly lag, the more coincident it is with the cycle, so that a peak value to the left of zero signifies a leading indicator and a peak value to the right of zero signifies a lagging indicator. It is clear from Figure 3.3 that both the CFNAI and CFNAI-MA3 are highly accurate coincident indicators of the U.S. business cycle, with peak AUCs of between 0.93 and 0.96 in the range of zero to one monthly lags. This result has also been borne out in practice, as the CFNAI led the NBER's dating of the 2001 and 2007–2009 recessions by 6–18 months on average in real time, according to the rules of thumb for the index used to judge the beginning and end of recessions (Brave and Butters 2010). Its success as a business-cycle measure has also led to its use in various forecasting applications for U.S. real GDP growth (Brave and Butters 2014), as well as the estimation of its trend (Brave and Butters 2013).

More recent work has expanded on the CFNAI by broadening the universe of data series and incorporating the latest advances in dynamic-factor analytic methods. For instance, Brave, Butters, and Kelley (2019) use mixed-frequency collapsed dynamic-factor analysis to summarize growth in 500 monthly real activity indicators and quarterly GDP growth to arrive at a measure of monthly GDP growth for the United States that can be decomposed into trend, cycle, and irregular components. The cycle component is then shown to be 99 percent accurate in capturing U.S. recessions and expansions and can be broken down further into leading and lagging elements that resemble the CFNAI and the Conference Board's Leading Economic Index, respectively.

The impressive résumé of the CFNAI has also spurred the development of other indexes used to measure growth in economic activity at a regional level (e.g., the Midwest Economy Index, or MEI; see Brave and Lu [2010]) and a local level (e.g., the Detroit Economic Activity Index, or DEAI; see Brave and Traub [2017]).[3] These indexes have been shown to be useful in filling gaps in our understanding of local economic conditions, given the longer publication delays and limited availability of data at state and local levels. For example, Brave and Wang (2011) used the MEI to predict gross state product growth in real time, and the DEAI was developed in order to measure the economic progress of Detroit after exiting bankruptcy. Summarizing annual, quarterly, and monthly data on city income, labor, real estate, and trade using mixed-frequency dynamic factor methods, the DEAI can also be used to estimate GDP for the city of Detroit as well as forecast its per capita income.

THE NFCI

The approximate factor model can also be applied to financial time series in order to capture periods of financial stress consistent with a *financial cycle*. Working with financial data, however, introduces additional complexities. For instance, financial time series are generally available at mixed and often higher frequencies of observation. Also, they tend to have richer correlation structures in which not all comovement can be captured in a single direction; i.e., there is generally a broader mix of procyclical and countercyclical indicators. Furthermore, if one is interested in isolating the state of financial markets from the state of the business cycle, adjustments must be made to either the data or the model to condition on this information.

All of these concerns are addressed in one form or another in the construction of the Chicago Fed's National Financial Conditions Index (NFCI). The NFCI is a weekly summary statistic for U.S. financial conditions and is estimated by mixed-frequency dynamic-factor analysis on a panel of 105 weekly, monthly, and quarterly financial time series. The index is representative of the entire U.S. financial system, containing broad coverage of money and debt and equity markets, as well as the traditional and "shadow" banking systems. It has been shown to be a useful tool in monitoring financial stability, aligning closely with historical episodes of financial stress (e.g., Brave and Butters 2011, 2012b).

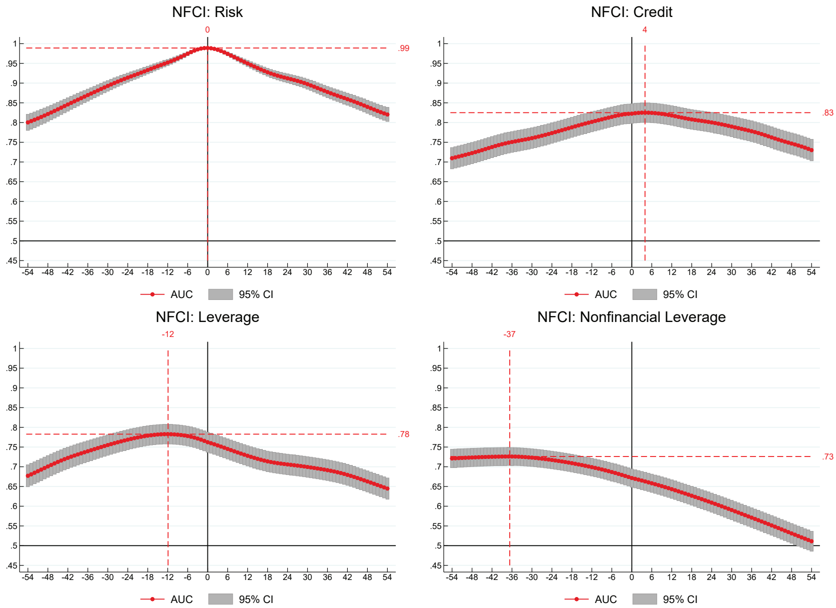
By conditioning the NFCI data on the state of the business cycle, a leading signal for financial stress can sometimes also be obtained (Brave and Butters 2011). This can be seen in Figure 3.2, which depicts the full-time series of the NFCI and its adjusted counterpart, the ANFCI. The ANFCI rebenchmarks U.S. financial conditions around a historical mean that is typical for a given level of economic growth and inflation (Brave and Kelley 2017). Positive (negative) values then denote tighter-than-average (looser-than-average) conditions on this basis. The ANFCI tends to display a slight lead on the NFCI in the run-up and aftermath of periods of financial stress. In addition, Brave and Genay (2011) find that it was also a useful predictor of Federal Reserve policy actions taken during the global financial crisis.

The indicators underlying the NFCI can be broadly classified into three types: 1) risk, 2) credit, and 3) leverage. These classifications are used in Brave and Butters (2012b) to construct subindexes of the NFCI (labeled *risk*, *credit*, and *leverage*) and highlight features of the financial cycle. Risk indicators capture volatility and funding risk in the financial sector and tend to be coincident indicators of financial stress. Credit indicators describe credit conditions in the nonfinancial sector and tend to be lagging indicators of financial stress. Finally, leverage indicators are measures of debt and equity in both sectors and tend to be leading indicators of financial stress.

In order to demonstrate these features, Figure 3.4 repeats the ROC analysis technique from the previous section on the three NFCI subindexes. For the subindexes, we classify the financial cycle based on a realization of the overall NFCI being positive or negative. Here, the x-axis values of the panels of the figure correspond to weekly leads or lags, while the y-axes continue to display AUC values. From the figure, it is clear that the risk subindex is a highly coincident indicator of financial stress (i.e., weeks where NFCI > 0), with a peak AUC value of 0.99 at a zero-week lag. On the other hand, the credit subindex tends to lag behind periods of stress by about a month, with a lower peak AUC value of 0.83, and the leverage subindex tends to lead periods of stress by about three months, with a lower peak AUC value of 0.78.

The leading signal for financial stress provided by leverage indicators can be further enhanced by isolating a subset of indicators for nonfinancial businesses and households. The resulting nonfinancial leverage

Figure 3.4 AUCs at Weekly Leads and Lags of the Financial Cycle



SOURCE: Author's calculations based on data available at www.chicagofed.org/nfci.

subindex tends to lead periods of stress by almost nine months, with a peak AUC value of 0.73, as seen in the bottom right panel of Figure 3.4. Brave and Butters (2012a) show that this particular subindex can be a useful early warning indicator, as it displays a significant lead with both the business and financial cycles and offers a superior view of potential financial imbalances in firm and household balance sheets in comparison with alternative measures like the private credit-to-GDP ratio. In addition, Brave and Lopez (2019) use this subindex to construct a probability of financial instability for the United States and then show how it can be used as a guidepost for macroprudential policymakers.

CONCLUSION

While it was not a point of focus in this chapter, it is worth mentioning that the activity index methodology can also be applied to qualitative, or survey-based, data just as easily as the quantitative government or market-based data focused on here. For example, Brave and Walstrum (2014) and Brave, Walstrum, and Berman (2015) develop an activity index methodology for quantifying survey responses collected for the Chicago Fed's *Beige Book* contribution. This work led to the introduction of the Chicago Fed Survey of Business Conditions (CFSBC).[4] Walstrum (2017) then showed how the CFSBC Activity Index could be used to forecast current-quarter GDP growth, much as with traditional activity indexes.

It is also worth noting that the development of activity indexes continues to be an active and expanding area of research, with the work of the Chicago Fed only a small part of that process. Within the Federal Reserve system, a number of indexes related to ours are also published, including various financial stress indexes and national and local business-conditions indexes.[5] Many foreign central banks and governmental agencies also produce similar indexes to ours in order to better understand fluctuations in their parts of the world. The success of these measures in capturing business and financial cycles and in aiding forecasting continues to demonstrate their value to policymakers and private-sector analysts.

Notes

Prepared for the 2018−2019 Werner Sichel Lecture Series at Western Michigan University. The views expressed herein are my own and do not necessarily represent those of the Federal Reserve Bank of Chicago or the Federal Reserve System. I would like to thank Dr. C. James Hueng and the WMU faculty and students for their hospitality; and the team of economists, research analysts, and assistants, as well as the editorial and public affairs staff, who support the production of these indexes at the Chicago Fed.

For more information on the Chicago Fed's economic and financial activity indexes, please go to chicagofed.org/research/data/index. To sign up for email notifications of any Chicago Fed index release, go to chicagofed.org/utilities/subscribe.

1. PCA can also be extended to handle some of these issues. See, for example, the alterations described in Stock and Watson (2002b).
2. As defined above, AUC values greater than 0.5 are consistent with a procyclical measure, and values less than 0.5 are consistent with a countercyclical measure. Wherever necessary, I have applied the convention of multiplying the indicator by 1 in order to assure that only AUC values greater than or equal to 0.5 are plotted. Without this sign convention, one would arrive at the overall accuracy of a countercyclical indicator by taking 1 minus its AUC value.
3. For more information on the MEI and DEAI, go to www.chicagofed.org/mei and www.chicagofed.org/deai.
4. For more information on the CFSBC, go to www.chicagofed.org/cfsbc.
5. See, for example, the metro business cycle indexes described in Arias, Gascon, and Rapach (2016) and maintained by the St. Louis Fed.

References

Arias, Maria A., Charles S. Gascon, and David E. Rapach. 2016. "Metro Business Cycles." *Journal of Urban Economics* 94(C): 90–108.

Aruoba, S. Borayğan, Francis X. Diebold, and Chiara Scotti. 2009. "Real-Time Measurement of Business Conditions." *Journal of Business and Economic Statistics* 27(4): 417–427.

Bai, Jushan, and Serena Ng. 2002. "Determining the Number of Factors in Approximate Factor Models." *Econometrica* 70(1): 191–221.

Bańbura, Marta, and Michele Modugno. 2014. "Maximum Likelihood Estimation of Factor Models on Datasets with Arbitrary Pattern of Missing Data." *Journal of Applied Econometrics* 29(1): 133–160.

Berge, Travis J., and Òscar Jordà. 2011. "Evaluating the Classification of Economic Activity into Recessions and Expansions." *American Economic Journal: Macroeconomics* 3(2): 246–277.

Bräuning, Falk, and Siem Jan Koopman. 2014. "Forecasting Macroeconomic Variables Using Collapsed Dynamic Factor Analysis." *International Journal of Forecasting* 30(3): 572–584.

Brave, Scott A. 2009. "The Chicago Fed National Activity Index and Business Cycles." Chicago Fed Letter No. 268. Chicago: Federal Reserve Bank of Chicago.

Brave, Scott A., and R. Andrew Butters. 2010. "Chicago Fed National Activity Index Turns Ten—Analyzing Its First Decade of Performance." Chicago Fed Letter No. 273. Chicago: Federal Reserve Bank of Chicago.

———. 2011. "Monitoring Financial Stability: A Financial Conditions Index Approach." *Economic Perspectives* 35(1): 22–43.

———. 2012a. "Detecting Early Signs of Financial Instability." Chicago Fed Letter No. 305. Chicago: Federal Reserve Bank of Chicago.

———. 2012b. "Diagnosing the Financial System: Financial Conditions and Financial Stress." *International Journal of Central Banking* 8(2): 191–239.

———. 2013. "Estimating the Trend Rate of Economic Growth Using the CFNAI." Chicago Fed Letter No. 311. Chicago: Federal Reserve Bank of Chicago.

———. 2014. "Nowcasting Using the Chicago Fed National Activity Index." *Economic Perspectives* 38(1): 19–37.

Brave, Scott A., R. Andrew Butters, and David Kelley. 2019. "A New 'Big Data' Index of U.S. Economic Activity." *Economic Perspectives* 43(1): 1–30.

Brave, Scott A., and Hesna Genay. 2011. "Federal Reserve Policies and Financial Market Conditions during the Crisis." FRB of Chicago Working Paper No. 2011-04. Chicago: Federal Reserve Bank of Chicago.

Brave, Scott A., and David Kelley. 2017. "Introducing the Chicago Fed's New Adjusted National Financial Conditions Index." Chicago Fed Letter No. 386. Chicago: Federal Reserve Bank of Chicago.

Brave, Scott A., and Max Lichtenstein. 2012. "A Different Way to Review the Chicago Fed National Activity Index." Chicago Fed Letter No. 298. Chicago: Federal Reserve Bank of Chicago.

Brave, Scott A., and Jose A. Lopez. 2019. "Calibrating Macroprudential Policy to Forecasts of Financial Stability." *International Journal of Central Banking* 15(1): 1–59.

Brave, Scott A., and Chenfei Lu. 2010. "A Snapshot of the Midwest Economy: Past and Present." Chicago Fed Letter No. 280. Chicago: Federal Reserve Bank of Chicago.

Brave, Scott A., and Paul Traub. 2017. "Tracking Detroit's Economic Recovery after Bankruptcy with a New Index." Chicago Fed Letter No. 376. Chicago: Federal Reserve Bank of Chicago.

Brave, Scott A., and Thomas Walstrum. 2014. "Using the Federal Reserve's *Beige Book* to Track Economic Activity." Chicago Fed Letter No. 328. Chicago: Federal Reserve Bank of Chicago.

Brave, Scott A., Thomas Walstrum, and Jacob Berman. 2015. "The Chicago Fed Survey of Business Conditions: Quantifying the Seventh District's *Beige Book* Report." *Economic Perspectives* 39(3): 77–89.

Brave, Scott A., and Norman Wang. 2011. "Predicting Gross State Product Growth with the Chicago Fed's Midwest Economy Index." Chicago Fed Letter No. 293. Chicago: Federal Reserve Bank of Chicago.

Chamberlain, Gary, and Michael Rothschild. 1983. "Arbitrage, Factor Structure, and Mean-Variance Analysis on Large Asset Markets." *Econometrica* 51(5): 1281–1304.

Doz, Catherine, Domenico Giannone, and Lucrezia Reichlin. 2012. "A Quasi-Maximum Likelihood Approach for Large, Approximate Dynamic Factor Models." *Review of Economics and Statistics* 94(4): 1014–1024.

Durbin, James, and Siem Jan Koopman. 2012. *Time Series Analysis by State Space Methods*. 2nd ed. Oxford Statistical Science Series, No. 38. Oxford: Oxford University Press.

Evans, Charles L., Chin Te Liu, and Genevieve Pham-Kanter. 2002. "The 2001 Recession and the Chicago Fed National Index: Identifying Business Cycle Turning Points." *Economic Perspectives* 26(3): 26–43.

Fisher, Jonas D. M. 2000. "Forecasting Inflation with a Lot of Data." Chicago Fed Letter No. 151. Chicago: Federal Reserve Bank of Chicago.

Giannone, Domenico, Michele Lenza, and Giorgio E. Primiceri. 2018. *Economic Predictions with Big Data: The Illusion of Sparsity*. FRB of New York Staff Report No. 847. New York: Federal Reserve Bank of New York.

Giannone, Domenico, Lucrezia Reichlin, and David Small. 2008. "Nowcasting: The Real-Time Informational Content of Macroeconomic Data." *Journal of Monetary Economics* 55(4): 665–676.

Harvey, Andrew C. 1989. *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge: Cambridge University Press.

Mariano, Roberto S., and Yasutomo Murasawa. 2003. "A New Coincident Index of Business Cycles Based on Monthly and Quarterly Series." *Journal of Applied Econometrics* 18(4): 427–443.

Stock, James H., and Mark W. Watson. 1999. "Forecasting Inflation." *Journal of Monetary Economics* 44(2): 293–335.

———. 2002a. "Forecasting Using Principal Components from a Large Number of Predictors." *Journal of the American Statistical Association* 97(460): 1167–1179.

———. 2002b. "Macroeconomic Forecasting Using Diffusion Indexes." *Journal of Business and Economic Statistics* 20(2): 147–162.

———. 2011. "Dynamic Factor Models." In *The Oxford Handbook of Economic Forecasting*, Michael P. Clements and David F. Hendry, eds. Oxford: Oxford University Press, pp. 35–59.

Walstrum, Thomas 2017. "Can the CFSBC Activity Index Nowcast U.S. Real GDP Growth?" Chicago Fed Letter No. 378. Chicago: Federal Reserve Bank of Chicago.

4
Nowcasting the Great Recession

Domenico Giannone
Eric Qian
Argia Sbordone
Mihir Trivedi
Federal Reserve Bank of New York

Patrick Adams
MIT Sloan School of Management

Economists at policy institutions, trading desks, and media outlets rely on economic data produced by various statistical agencies to understand the state of the economy and predict its future path. However, the highest-quality and most comprehensive economic data are published with long delays after the periods to which they refer. Most notably, gross domestic product (GDP), the most comprehensive measure of U.S. economic activity, is first published by the Bureau of Economic Analysis (BEA) one month after the end of each reference quarter, and these initial estimates are later revised.

Faced with the challenging task of monitoring macroeconomic conditions in real time, analysts track a wide variety of data releases, distilling signal from noise in incoming data and revising their beliefs about the state of the economy when these data diverge significantly from their expectations. The Nowcasting Report of the Federal Reserve Bank of New York (the New York Fed Staff Nowcast) formalizes and automates this process through an econometric model–based approach. The platform produces *nowcasts* of economic activity—predictions for the present, recent past, and near future—which are continually updated as new data become available. The platform's nowcasts of real GDP growth can be computed before the start of the reference quarter and updated each day to incorporate the most recent information, providing useful real-time readings on the state of the economy that can be used to guide key policy and private-sector decisions.

In this chapter, we provide a brief overview of the general challenge of monitoring macroeconomic conditions in real time and the methods underlying the New York Fed Staff Nowcast. We then present two case studies that assess the ability of the New York Fed Staff Nowcast to provide accurate early estimates of GDP during important real-world situations.

First, we study the day-by-day movements in the GDP growth nowcast during two critical quarters of the 2007–2009 recession. The model is able to predict major swings in economic activity (both upward and downward) long before the publication of the first official GDP estimates, providing confidence in its ability to track business cycle fluctuations.

Second, motivated by extensive data publication delays resulting from the 2018–2019 partial shutdown of the U.S. federal government, we conduct a counterfactual exercise to evaluate the performance of the nowcast during periods of severe yet realistic disruptions to the standard flow of macroeconomic data. Even during such periods, the nowcast can predict GDP growth with an accuracy comparable to the BEA's first official estimate; in particular, it can serve as a useful substitute for the official estimate if it is not published according to schedule (as was the case in early 2019).

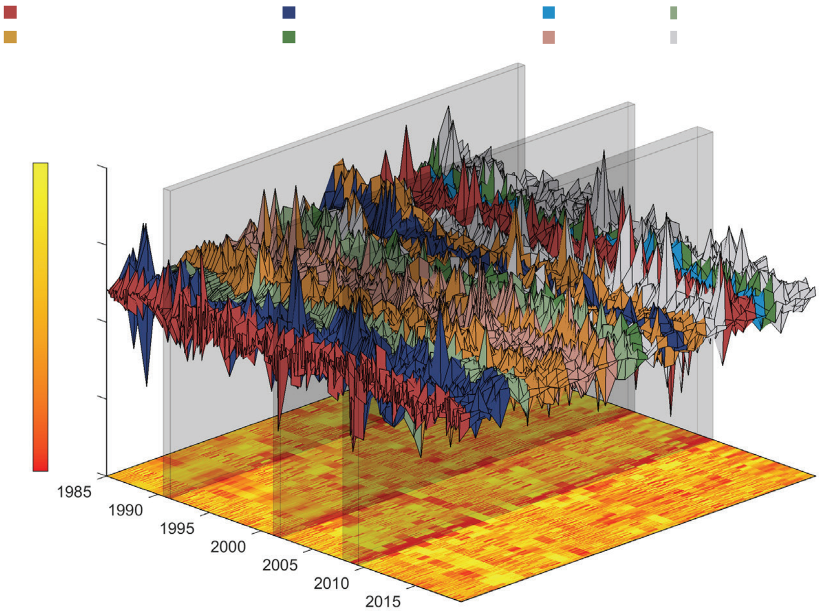
THE REAL-TIME DATA FLOW

As mentioned, in order to understand the state of the economy in real time, economists must extract signal from noise in a broad set of economic data. At any given point in time, economists face a trade-off between timeliness and quality when evaluating the most recent available data for each indicator. Business and consumer sentiment indicators—often referred to by market commentators as *soft data*—provide the first readings on economic activity during a particular reference period. Labor market indicators typically arrive next; most notably, the widely followed Employment Situation Report, released by the Bureau of Labor Statistics (BLS), provides estimates of the unemployment rate and payroll employment shortly after the end of the month to which the new data refer. Hard data on production, sales, and income begin to arrive several weeks later. Finally, the first estimate of gross domestic product—the total value of all goods and services produced in the United States over a given quarter—is published by the Bureau of Economic Analysis (BEA) roughly one month after the end of the reference quarter. Data of the highest quality and broadest coverage are thus only available well after the end of the period to which they refer.

The trade-off between timeliness and accuracy is present not only across different economic indicators, but also across different releases for the same indicator. In the case of GDP, the BEA produces benchmark estimates every five years that rely on data collected through a comprehensive economic census, covering around seven million businesses with paid employees and over 95 percent of the expenditures included in GDP. Between these benchmark estimates, annual and quarterly estimates are constructed using surveys conducted by the Census Bureau—with 150,000 and 35,000 reporting units, respectively—as well as administrative data (such as those from the Internal Revenue Service) and estimates from other sources (such as BLS employment data). In short, these benchmark revisions are the product of careful aggregation of detailed microeconomic information into national accounts. In contrast, the "advance" GDP release is the first official estimate available, with only a one-month delay. It is constructed using only half of the hard expenditure data ultimately used in the benchmark revisions and relies heavily on survey data gleaned by the Census Bureau. As a result of these unavoidable shortcuts required to produce timely estimates, these first estimates of GDP are subject to sizable revisions as higher-quality source data become available. What is gained through detailed microeconomic information is lost in timeliness.

With new data being released almost every day, and each release providing estimates for a large number of variables over a single reference period, economists face a big-data problem when attempting to monitor economic conditions. As an illustration, Figure 4.1 provides a useful visualization of the data at our disposal. The three-dimensional surface plot displays the path of 37 major economic indicators since 1985, with each data series colored according to its category.[1] The heat map on the horizontal plane presents a two-dimensional visualization of the same data: brighter yellow values indicate realizations above the sample mean for a given series, while darker red values indicate realizations well below the mean. The dark red areas are especially preva-

Figure 4.1 Big Data in Macroeconomics



NOTE: The three-dimensional surface plot displays the standardized time series for the major economic indicators since 1985, which are shaded by category as indicated in the legend. Recessions are marked by transparent gray surfaces. The heat map on the horizontal plane shows positive and negative readings of the data.
SOURCE: Authors' calculations, based on data accessed through Haver Analytics and the Federal Reserve Economic Database (FRED).

lent across many series during the recessions of the early 1990s, early 2000s, and (most notably) 2008 through 2009. In these periods, broad red strips across the heat map highlight the common downward movement across many series. However, despite the stark common movements across these series during both recessions and expansions, at any given point in time there are also individual series whose movements deviate from the others.

NOWCASTING

Nowcasting refers to the process of monitoring economic conditions by forming predictions for economic activity in the present, recent past, and near future. Nowcasting is a big-data problem, given the vast array of macroeconomic data at our disposal. The New York Fed Nowcast summarizes the rich and complex dataset depicted in Figure 4.1 using a parsimonious model motivated by the strong comovements evident among the series. The model formalizes the notion of a common business-cycle component present across all of these series and allows one to distill signal from noise by filtering out fluctuations specific to individual variables from incoming data.

The New York Fed Staff Nowcast is based upon a dynamic-factor model, which solves the "large *n*, small *T*" problem of relatively few time observations *T* compared to the large number of available data series *n*. It does this through dimension reduction: a small number of unobserved common factors are used to summarize the bulk of fluctuations in the observed variables. Forni et al. (2000) and Stock and Watson (2002a,b) presented the first applications of dynamic-factor models to large macroeconomic data sets, while Giannone, Reichlin, and Small (2008) demonstrate that these models can be used to reliably predict GDP growth in real time. Over the past decade, nowcasting models have been developed for many countries (see Bańbura et al. [2013] and Bok et al. [2018] for a survey).The dynamic-factor model can be easily cast in state space form, allowing for efficient estimation of both unknown parameters and unobserved common factors using the Kalman filter (Bańbura and Modugno 2014; Doz, Giannone, and Reichlin 2011). Moreover, the process by which the model's GDP growth forecasts are updated upon the release of new data can be interpreted in an intuitive manner that mimics market participants' processing of information. Before each data release, a new value is predicted for each variable based on previously available information. Once the new data are released, the model updates its forecast for GDP growth based on the discrepancy between predicted and realized values of all the variables; we refer to this discrepancy as *news*. If the model's predictions for each variable are exactly correct, its GDP growth forecast will remain unchanged, just as market participants would not revise their beliefs about the state of the economy in the absence of news. On the other hand, if the model's predictions are not exactly correct, its GDP growth forecast will be revised to account for the news, just as market participants who observe stronger- or weaker-than-expected data would revise their beliefs about the state of the economy accordingly.

More formally, revisions to the model's GDP growth forecasts are simply a weighted sum of news across all data releases. The sign of these weights encodes whether a higher-than-expected value for each release represents "good" or "bad" news (e.g., payroll employment versus the unemployment rate). The magnitude of the weights encodes the overall information content that the news provides on economic conditions in a particular period, taking into account factors like timeliness and the extent to which each variable is driven by common versus idiosyncratic fluctuations. The model is thus able to determine which data releases are most important for monitoring current economic conditions, just as market participants place greater emphasis on some data releases than others (as evidenced by sharp asset price movements typical upon the release of closely followed indicators like GDP).

The New York Fed Nowcast therefore provides a platform for interpreting the flow of data in real time. By determining each new observation's impact on predicted GDP growth, the model provides a "common unit" for comparing news across series like payroll employment and the unemployment rate. Additionally, the GDP growth forecasts are updated in a fully automated and judgment-free manner, allowing for continuous updates as soon as new information becomes available. And, these forecasts are available well before the publication of the first official estimate, which occurs one month after the end of each reference quarter. A detailed description of the model and the data is provided in Bok et al. (2018).

REAL-TIME ESTIMATES: TWO CASE STUDIES

We now present two case studies that illustrate the real-time performance of the New York Fed Staff Nowcast in two scenarios in which early and accurate GDP estimates serve a particularly important role in both policy and private-sector decision making. For both of these case studies, we make use of a daily real-time database that tracks the exact data available for each of the model's 37 input series on each date from October 1, 2001, to the present. By using real-time data to recursively estimate the model's parameters at the start of each quarter and update the GDP growth estimates using the new data available on each date, we are able to exactly reconstruct the estimates that forecasters would have obtained using our model in the past. Complete archives of both the reconstructed and real-time New York Fed Staff Nowcast estimates are available to the public and are described in Adams, Giannone, et al. (2019).[2]

Nowcasting the Great Recession

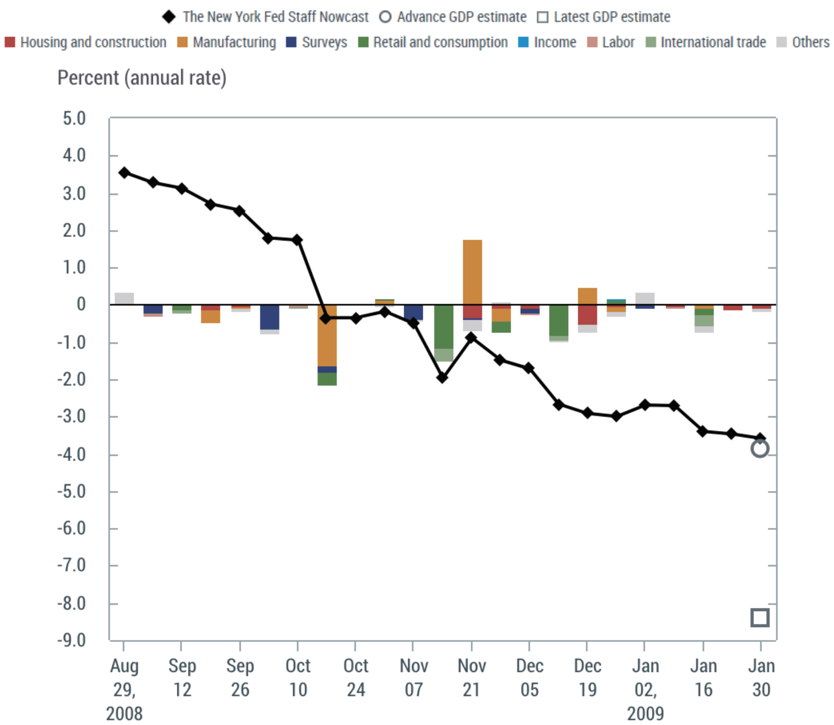
In our first case study, we track the day-by-day movements in the nowcast for GDP growth during two critical quarters of the 2007–2009 recession. The fourth quarter of 2008 saw the sharpest decline in real economic activity, while the third quarter of 2009 marked the end of the recession and the beginning of the recovery. For both quarters, we chart the progression of the GDP growth nowcast (represented by black diamonds), starting one month before the start of the quarter and ending one month after the end of the quarter after the BEA publishes the first official GDP estimate. Colored bars denote the overall contributions of data releases from different categories—surveys, retail and consumption, and more—to the weekly changes in the nowcast, based on the decompositions described in the second section. For comparison against official estimates, we also plot both the BEA's first and latest estimates for each quarter.[3]

The fourth quarter of 2008 was the worst of the recession, with real GDP contracting by 8.4 percent. On September 12, 2008 (right before the failure of Lehman Brothers was announced on September 15), our forecast for GDP growth for the fourth quarter of 2008 actually stood at a promising 3.1 percent. This estimate quickly changed as data on business sentiment, industrial production, and retail sales for the month of September became available, and our nowcast first dropped into negative territory roughly one month after Lehman Brothers' bankruptcy on October 17. The National Bureau of Economic Research (NBER) Recession Dating Committee officially announced on December 1, 2008, that the economy had been in recession for the past 12 months.

On the previous Friday, our nowcast for GDP growth in 2008:Q4 was −1.5 percent (Figure 4.2). Additional negative data releases over the next two months led to further declines in our nowcast, until our final prediction sank to −3.6 percent immediately before the advance GDP release in January 2009. Although we predicted the BEA's advance estimate almost exactly, this first estimate understated the severity of the contraction and was later revised downward substantially.

The third quarter of 2009 marked the end of the recession, as determined by the NBER Recession Dating Committee one year later. At the start of the quarter in July, our nowcast still predicted slightly negative GDP growth. However, over the next few months, a wide variety

Figure 4.2 Nowcasting the Great Recession in 2008:Q4

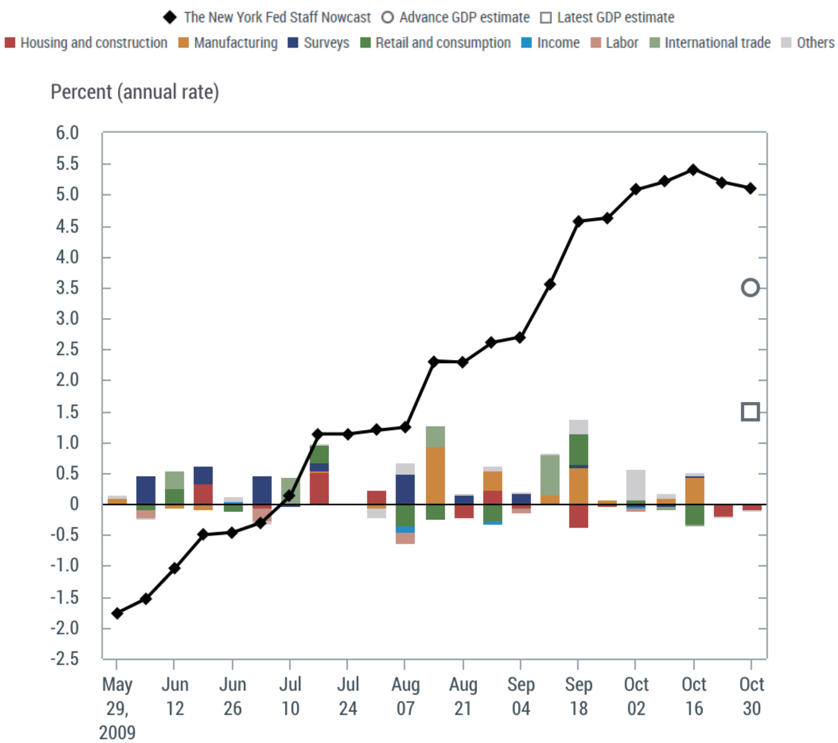


NOTE: Colored bars reflect the impact of each data release on the nowcast.
SOURCE: Authors' calculations, based on data accessed through Haver Analytics.

of better-than-expected data was released (especially for manufacturing, international trade, and business sentiment), and our nowcast for 2009:Q3 GDP growth (Figure 4.3) rose to over 5 percent by the end of the quarter. Our model successfully predicted the timing of the recovery but turned out to be overly optimistic in predicting its strength: the advance estimate of GDP growth in 2009:Q3 was 3.5 percent, but this estimate was later revised downward, as the latest available estimate reported growth of only 1.5 percent.

For both of these important quarters, the nowcast provided an early and reliable signal of the direction in which growth was headed, months before the publication of the first official estimate. These results provide

Figure 4.3 Nowcasting the Recovery in 2009:Q3



NOTE: Colored bars reflect the impact of each data release on the nowcast.
SOURCE: Authors' calculations, based on data accessed through Haver Analytics.

confidence that the New York Fed Staff Nowcast can provide useful early readings on upward and downward swings in activity by filtering through a variety of incoming data ahead of the publication of official GDP estimates. The large revisions from the first to the latest published estimates show that producing estimates with both minimal delay and high precision is a challenge even for the BEA. The contribution of the New York Fed Staff Nowcast is to extend the "accuracy timeliness" frontier in the period before official statistics are available.

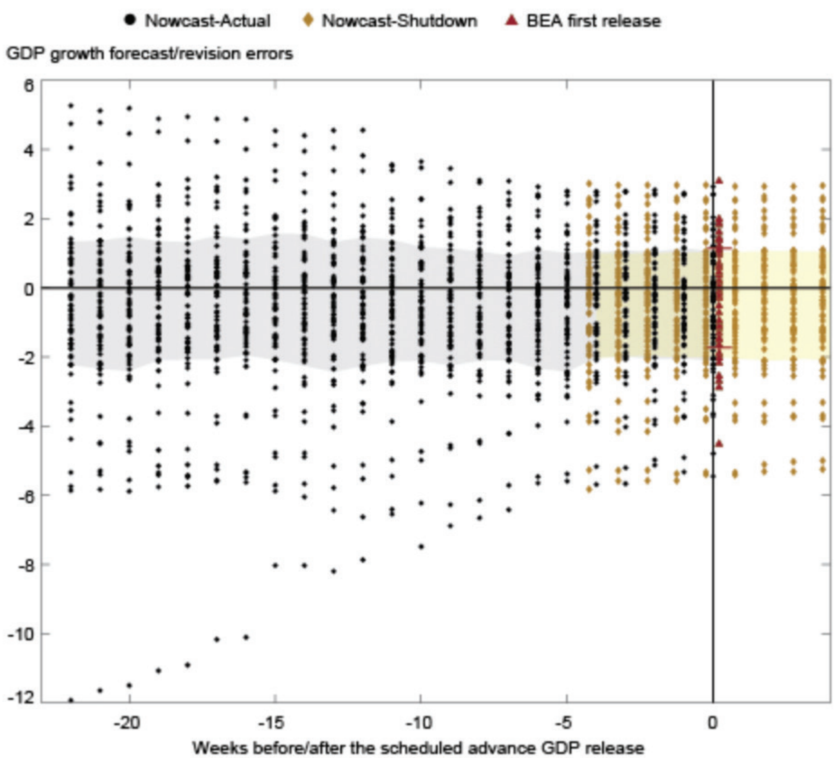
Nowcasting with Scarce Data

In our second case study, we evaluate the performance of the nowcast during periods of severe yet realistic disruptions to the standard flow of macroeconomic data. This exercise is motivated by the 2018−2019 U.S. federal government partial shutdown, during which the temporary closure of the Census Bureau and the Bureau of Economic Analysis delayed the publication of many scheduled data releases. While the most notable delayed release was the first estimate of 2018:Q4 GDP (which was postponed by one month), many other widely followed indicators of economic activity were released with substantial delays, including retail sales, home sales and construction, imports and exports, and durable goods orders.[4] However, a variety of other data sources—both hard data directly measuring activity and soft data measuring business sentiment—were published as previously scheduled.

Do such disruptions to the regular data publication schedule impair the ability of the New York Fed Staff Nowcast to accurately predict GDP growth? To answer this question, we conduct a counterfactual exercise in which we simulate similar dataflow disruptions for each quarter from 2002:Q1 to 2017:Q4, as if the Census Bureau and BEA had ceased publication of all new data releases starting one week before the end of the quarter. We assume that data previously published by these agencies remain available, and that new data published by other government agencies and private organizations become available as they are released in real time. We then evaluate the performance of our nowcasting model in this "scarce data" setting by studying the empirical distribution of its forecast errors for GDP growth, which provides a complete description of its historical forecasting performance.

Figure 4.4 plots GDP growth forecast errors for all quarters in our evaluation sample, based on the number of weeks remaining until the first GDP release at the date of each forecast (listed across the horizontal axis). The black dots represent the historical forecast errors for our nowcasting model (using the actual data available in real time),

Figure 4.4 Similar Data Delays Would Not Have Drastically Changed Past Predictions



NOTE: Points represent quarterly GDP growth forecast errors (for Nowcast-Actual and Shutdown) and revision errors (for the Bureau of Economic Analysis first release), computed with respect to the latest available estimates for the years 2002 through 2017. Shaded bands depict the 16th and 84th percentiles of the historical forecast errors, while the red lines at week 0 depict the same percentiles for revisions to the first release.

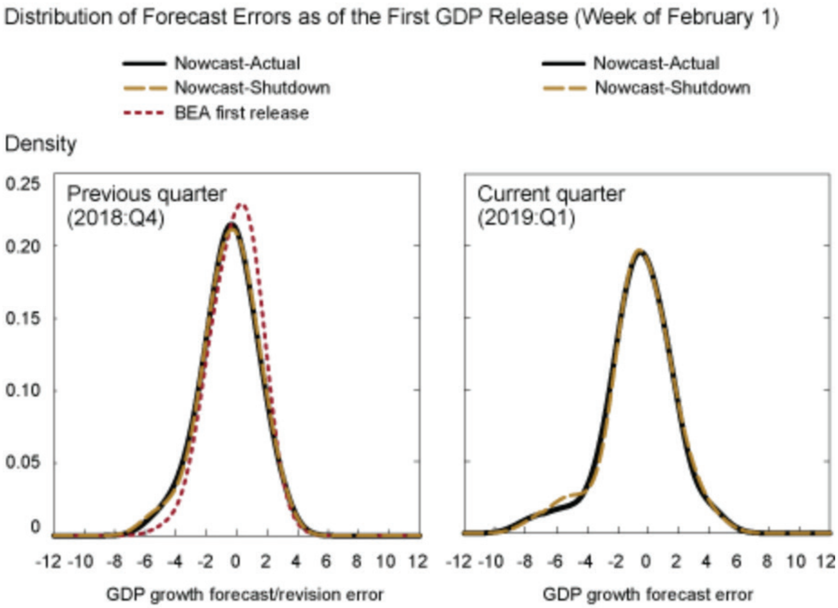
SOURCE: Authors' calculations, based on data accessed through Haver Analytics.

computed using the latest available GDP growth estimates.[5] The gold diamonds represent our model's forecast errors under the counterfactual scenario when key data releases are delayed, starting roughly four weeks before the first GDP release. The red triangles represent revision errors, from the first estimate published by the BEA for each quarter to the latest available estimates. Shaded bands depict the 16th to 84th percentile range of the errors for each forecast, while the red line on the week of the first GDP release depicts the same range for revision errors from the BEA's first to latest releases; these ranges provide an assessment of uncertainty for each of the three forecasts.

Overall, the model's forecasts for GDP growth remain accurate even when there are substantial disruptions to the usual pattern of data availability. For the weeks leading up to the first GDP release, the historical forecast error distributions represented by the black and gold markers are broadly similar, indicating that the accuracy of the nowcast is mostly unchanged when new Census Bureau and BEA data releases are not published. Moreover, under these conditions, the gold uncertainty bands reported for the nowcast are similar in width to the red bar at the week of the first GDP release, depicting uncertainty around this estimate. Therefore, the finding of Bok et al. (2018)—that uncertainty is similar around both the final nowcasts for a given quarter and the first GDP release—still holds, even when important data are not released according to their usual schedule, as was the case during the 2018–2019 federal government shutdown. Moreover, if the first release of GDP also happens to be delayed during these periods, the nowcast provides an alternative estimate of GDP growth of comparable accuracy to the first release.

Figure 4.5 presents an alternative visualization of the forecast error distributions for the nowcast (both the actual historical forecast errors and the errors under our counterfactual scenario based on the government shutdown) and first official GDP release. The left panel depicts the smoothed empirical distribution of the three sets of forecast errors plotted along the vertical line at the week of the first GDP release from the previous figure. The right panel depicts the distribution of forecast errors for current-quarter projections at the time of the previous quarter's first GDP release—e.g., the projections for 2019:Q1 GDP growth made near the end of January 2019.

Figure 4.5 Nowcast with Delayed Data Similarly Accurate to BEA First Release

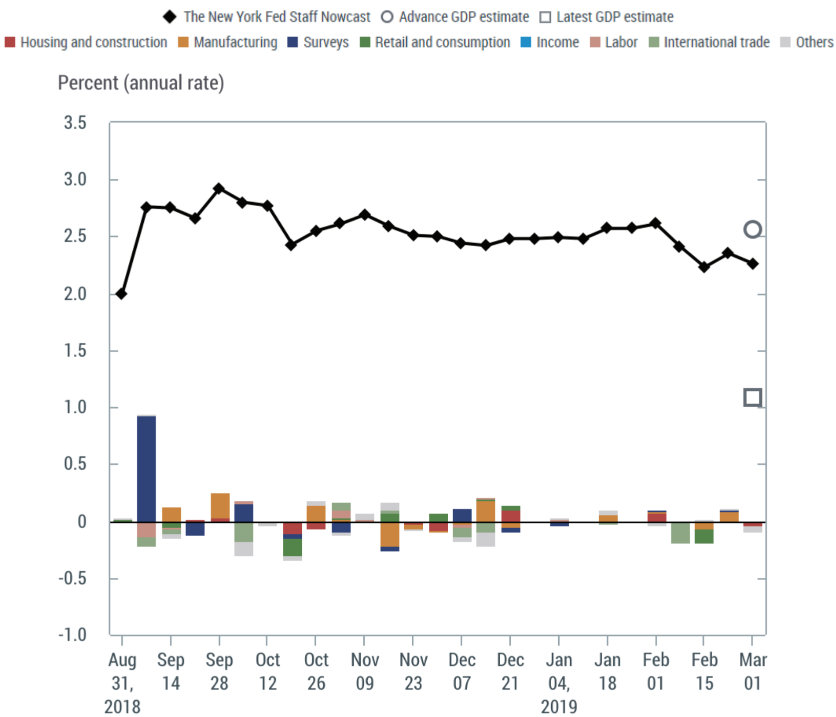


NOTE: The figures report kernel-smoother estimates of forecast error densities as of the scheduled first GDP release (week of February 1, 2019). The left panel gives error distributions for the previous quarter (2018:Q4). The right panel gives error distributions for 2019:Q1. Black lines refer to the actual nowcast errors, gold dashed lines refer to nowcast shutdown errors, and red dotted lines refer to the BEA revision errors.
SOURCE: Authors' calculations, based on data accessed through Haver Analytics.

As noted in the discussion of the previous figure, the three forecast error distributions are broadly similar, indicating that the nowcast provides estimates of GDP growth with comparable accuracy to the first official release, even under conditions of data scarcity similar to those resulting from the 2018–2019 U.S. federal government shutdown. The main differences between these distributions arise from large negative forecast errors, which are more likely to occur for the nowcast than for the BEA estimates. Moving to the right panel, both distributions for the model's longer-horizon forecast errors display greater dispersion than their shorter-horizon counterparts in the left panel, reflecting greater uncertainty when making predictions for the current (as opposed to the previous) quarter. The striking similarity of the model's actual and counterfactual forecast errors reported in the right panel illustrates that data scarcity, similar to the scarcity of data resulting from the 2018–2019 U.S. federal government shutdown, does not blunt the accuracy of early projections for the current quarter.

Finally, Figure 4.6 displays the real-time progression of the GDP growth nowcast for 2018:Q4. The effects of the shutdown-related data publication delays can be seen through the paucity of colorful bars from late December through early January. While our model's prediction was quite close to the first official estimate released in February 2019, the

Figure 4.6 Nowcasting during the Government Shutdown



NOTE: Colored bars reflect the impact of each data release on the nowcast.
SOURCE: Authors' calculations, based on data accessed through Haver Analytics.

latest available estimate is notably lower than both of these early estimates, highlighting the uncertainty about economic activity that prevails even after official statistics are initially published.

CONCLUSION

The New York Fed Staff Nowcast is able to produce accurate and early estimates of real GDP growth well before the publication of the BEA's first official estimates. We presented two case studies that evaluate the model's performance during the Great Recession and during the U.S. federal government shutdown at the beginning of 2019. We encourage interested readers to further study our model by exploring our online interactive archives, which collect both real-time forecasts for the period from 2016:Q1 to the present and reconstructed historical forecasts extending back to 2002:Q1.[6]

Notes

The views expressed in this chapter are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of New York, the Federal Reserve System, or the Upjohn Institute.

1. The full list of all 37 series is presented in Bok et al. (2018). Each series is appropriately transformed in order to induce stationarity, then standardized so that all variables have the same mean and variance over the sample period.
2. These archives can be explored in interactive form at the following link: https://www.newyorkfed.org/research/policy/nowcast.
3. The latest available estimates are based on data published by the BEA in July 2019.
4. For a full list of data releases delayed because of the shutdown, see Adams, Qian, et al. (2019).
5. We use the latest available estimates, since these reflect both 1) new source data that become available after the publication of the first estimate and 2) methodological changes intended to improve the quality of the estimates.
6. These archives (along with structured data files containing historical forecasts) can be found at the following link: https://www.newyorkfed.org/research/policy/nowcast.

References

Adams, Patrick, Domenico Giannone, Eric Qian, and Argia Sbordone. 2019. "Just Released: Historical Reconstruction of the New York Fed Staff Nowcast, 2002–15." *Liberty Street Economics* (blog), Federal Reserve Bank of New York, July 12. New York: Federal Reserve Bank of New York.

Adams, Patrick, Eric Qian, Argia Sbordone, and Domenico Giannone. 2019. "Monitoring Economic Conditions during a Government Shutdown." *Liberty Street Economics* (blog), Federal Reserve Bank of New York, February 5. New York: Federal Reserve Bank of New York.

Bańbura, Marta, Domenico Giannone, Michele Modugno, and Lucrezia Reichlin. 2013. "Now-Casting and the Real-Time Data Flow." In *Handbook of Economic Forecasting l*, Graham Elliott, Clive W. J. Granger, and Allan Timmermann, eds. Vol. 2. Amsterdam: Elsevier North-Holland, pp. 195–237.

Bańbura, Marta, and Michele Modugno. 2014. "Maximum Likelihood Estimation of Factor Model on Datasets with Arbitrary Pattern of Missing Data." *Journal of Applied Econometrics* 29(1): 133–160.

Bok, Brandyn, Daniele Caratelli, Domenico Giannone, Argia M. Sbordone, and Andrea Tambalotti. 2018. "Macroeconomic Nowcasting and Forecasting with Big Data." *Annual Review of Economics* 10(1): 615–643.

Doz, Catherine, Domenico Giannone, and Lucrezia Reichlin. 2011. "A Two-Step Estimator for Large Approximate Dynamic Factor Models Based on Kalman Filtering." *Journal of Econometrics* 164(1): 188–205. https://www.sciencedirect.com/science/article/abs/pii/S030440761100039X?via%3Dihub (accessed February 10, 2020).

Forni, Mario, Marc Hallin, Marco Lippi, and Lucrezia Reichlin. 2000. "The Generalized Dynamic-Factor Model: Identification and Estimation." *Review of Economics and Statistics* 82(4): 540–554.

Giannone, Domenico, Lucrezia Reichlin, and David Small. 2008. "Nowcasting: The Real-Time Informational Content of Macroeconomic Data." *Journal of Monetary Economics* 55(4): 665–676.

Stock, James H., and Mark W. Watson. 2002a. "Forecasting Using Principal Components from a Large Number of Predictors." *Journal of the American Statistical Association* 97(460): 1167–1179.

———. 2002b. "Macroeconomic Forecasting Using Diffusion Indexes." *Journal of Business and Economic Statistics* 20(2): 147–162.

5
Real-Time Measurement of Business Conditions, Macroeconomic Surprises, and Uncertainty

Is a Recession Looming?

Alessandro Barbarino
Chiara Scotti
Federal Reserve Board

As pointed out by the Congressional Budget Office (CBO 2019), the current economic expansion has lasted more than nine years, becoming one of the longest expansions since 1945.[1] Because of such a long boom, some people think that a recession might be due soon. In January 2019, for example, the *Wall Street Journal* published a survey showing a 25 percent probability of recession in the next year, the highest level since October 2011 and twice the probability of one year ago. Similarly, the *Blue Chip Indicators* for January 2019 reported a consensus survey result for the probability of a recession in 2019 at 25 percent and the probability of a recession in 2020 at 37 percent. Reportedly, one of the reasons the Dow 30 and S&P 500 indexes both fell by more than 15 percent in December 2018 was from a concern that the economy would fall into a recession in late 2019 or 2020, prompting negative earnings growth.

This chapter focuses first on evaluating current business conditions in the United States, based on real-activity economic indicators, as well as on gauging market participants' optimism or pessimism about the economy and the uncertainty around this evaluation. Second, it evaluates some recession probability models that make use of a variety of data to pinpoint whether indeed a recession is looming.

To evaluate business conditions, we look at the Aruoba, Diebold, and Scotti (2009) business condition index (henceforth, “ADS index”) as well as the Scotti (2016) surprise and uncertainty indexes updated with the most current data. The ADS index turned negative in early 2019, suggesting worse-than-average conditions for the U.S. economy over the preceding months.[2] Economic surprises and uncertainty are evaluated using the Scotti (2016) indexes. The surprise index spiked early in 2019, as market participants were more pessimistic than warranted by economic releases, but then sharply collapsed following the release of the February employment report in early March. The uncertainty index steadily increased at the beginning of 2019, reaching levels last seen in late 2017.

Given this assessment, should we conclude that a recession is looming, as suggested by the CBO analysis? In order to tackle this question, we review the prediction of extant recession forecasting models by feeding them a variety of data, including the real-time indicators discussed in the first part of this chapter, a larger set of individual macro variables, and financial variables (like level, slope, curvature, corporate spreads, and so forth). When using individual data, to make sure we have entries for all the variables until the last data point, we truncate the sample in December 2018 (even if for some data series we have data until the day before we ran the estimation in March 2019) and find an increased probability of recession in mid-2019, possibly due to the big correction observed in financial markets in late 2018. When we reestimate the recession probabilities employing only the ADS index as a summary statistic of the real indicators, which is available in real time and allows us to take care of the ragged edges of the data, the estimated probabilities significantly decrease. Our analysis also shows that, consistent with Berge (2015), real variables appear to be more powerful in signaling recessions at shorter horizons, while the term spread and some additional financial variables are valuable leading indicators at longer horizons—that is, at horizons of 6–12 months ahead and beyond.

The remainder of the chapter discusses the real-time measurement of business conditions in the next section, the real-time evaluation of optimism/pessimism and uncertainty about the state of the economy in the section after that, and the evaluation of recession probability models with a variety of data in the fourth section. The final section offers our conclusions.

REAL-TIME MEASUREMENT OF ECONOMIC CONDITIONS

Aruoba, Diebold, and Scotti (2009) state the following:

> Aggregate business conditions are of central importance in the business, finance, and policy communities, worldwide, and huge resources are devoted to the assessment of the continuously evolving state of the real economy. Literally thousands of newspapers, newsletters, television shows, and blogs, not to mention armies of employees in manufacturing and service industries, including the financial services industries, central banks, government and non-government organizations, grapple constantly with the measurement and forecasting of evolving business conditions.

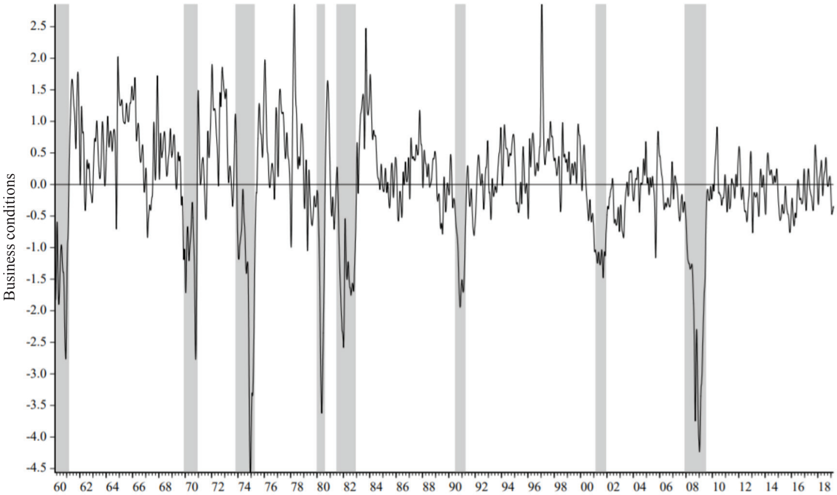
Complications to this assessment include the fact that business conditions are latent, data are released at different times and therefore not always all available at the time of the evaluation, and they are at different frequencies. The latency of the business cycle means that the business cycle is not directly observed, as it is not represented by any single variable, but rather, it is derived by information contained in a number of indicators like gross domestic product (GDP), industrial production (IP), employment, and so on. In fact, the National Bureau of Economic Research (NBER) does not define a recession in terms of only one indicator of activity, such as two consecutive quarters of decline in real GDP, but as a significant decline in economic activity spread across the economy, lasting more than a few months, normally visible in real GDP, real income, employment, industrial production, and wholesale-retail sales.[3]

Data are released at different times. For example, nonfarm payroll is announced the first Friday of the month, and initial jobless claims are released weekly on Thursday. Assume that it is Tuesday, February 26, 2019, and we are trying to assess the current state of the economy for the first quarter of 2019. We only have partial information available relative to that quarter. The January nonfarm payroll is available, but the February job market report will not be available until the coming Friday. Likewise, as of February 26, initial jobless claims releases are available only for the first seven weeks of the year. In addition, GDP data relative to the first quarter will be released only a quarter later. This ragged-edge structure of the data complicates the evaluation of real-time business conditions, as we need to juggle data series of different lengths.

In addition, data have different frequencies, covering various units of time. For example, GDP is quarterly, nonfarm payroll is monthly, and initial jobless claim is weekly. An assessment of business conditions needs to take this into account and be able to accommodate the different units and the aggregation that makes weekly series comparable to monthly or quarterly variables.

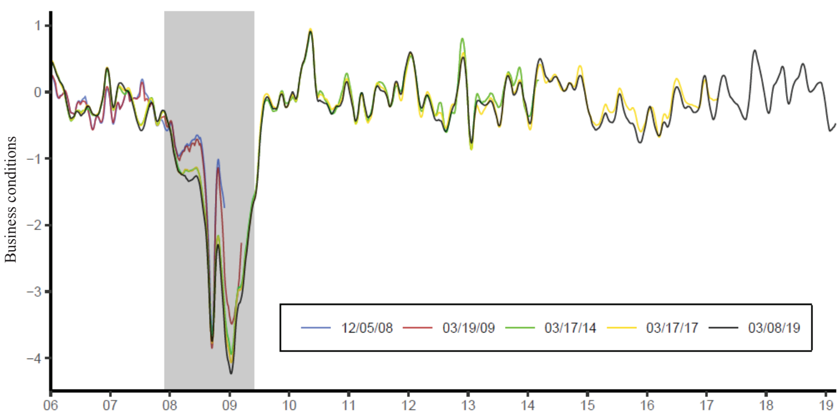
The empirical business cycle literature has dealt with these features through alternative approaches, including the dynamic factor framework, whether from the "small data" perspective, as in Aruoba, Diebold, and Scotti (2009); Chauvet (1998); Diebold and Rudebusch (1996); and Stock and Watson (1989), or the "big data" perspective, as in the seminal work of Bai and Ng (2006); Forni et al. (2000); and Stock and Watson (1991, 2002). Aruoba, Diebold, and Scotti (2009) propose a framework to measure economic activity in real time using a dynamic factor model that combines a small set of time series at different frequencies. In particular, the ADS index is designed to track real business conditions at a high frequency, combining information from (seasonally adjusted) economic indicators: weekly initial jobless claims, monthly payroll employment, monthly industrial production, monthly personal income less transfer payments, monthly manufacturing and trade sales, and quarterly real GDP. The Philadelphia Fed updates the index as soon as new data releases become available. Figure 5.1 displays the ADS index as of March 15, 2019. Of note, the average value of the index is zero, with progressively bigger positive values indicating progressively better-than-average conditions, and progressively more negative values indicating progressively worse-than-average conditions. The business condition index in Figure 5.1 is based on the information available as of March 2019. The index might look different, though, when computed on different data vintages. Figure 5.2, for example, shows the ADS index computed in real time in March 2019 and contrasts it against the index computed on different data vintages ranging from 11 to 2 years prior to March 2019. The ADS index computed on the December 2008 data vintage ends in December 2008 and, likewise, lines for the indexes computed on data vintages in 2009, 2014, 2017, and 2019 end in the respective years. Looking at the last recession, the real-time estimate of the index turned out to be overly optimistic, and it was subsequently revised downward, as shown by the wedge between the 2008 vintage (in

Figure 5.1 Aruoba-Diebold-Scotti (ADS) Business Condition Index, 3/1/1960–3/15/2019



NOTE: The ADS is constructed using the latest data available as of March 15, 2019. Grey shading indicates NBER-designated recessions. The limits used on the y-axis reflect the minimum and maximum values of the index over the entire history.
SOURCE: Federal Reserve Bank of Philadelphia, ADS Business Conditions Index.

Figure 5.2 ADS Index in Real Time



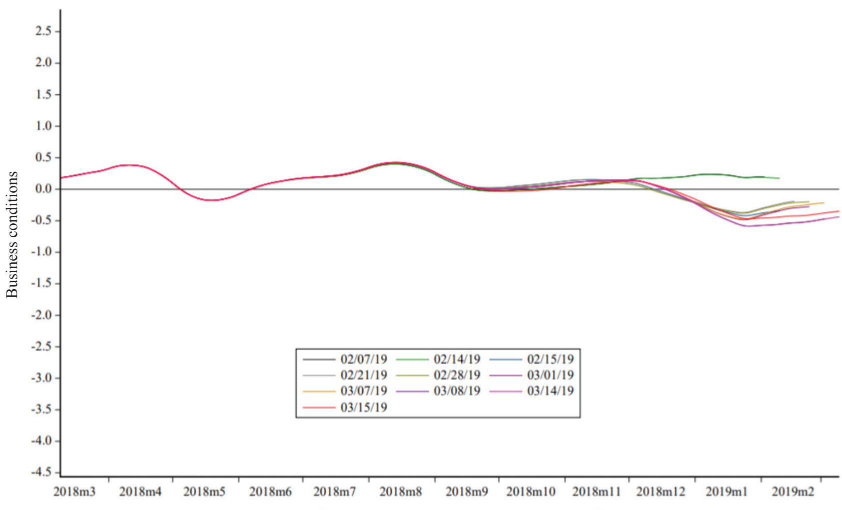
NOTE: The ADS indexes are constructed using data available up to the date indicated in the legend.
SOURCE: Federal Reserve Bank of Philadelphia, ADS Business Conditions Index.

blue) and the March 2019 vintage (the black line). Focusing on recent years, Figure 5.3 shows the tentacle plot of the 10 vintages of data prior to March 2019. After a couple of positive estimates at the beginning of the year, the ADS index was subsequently revised downward into negative territory relative to estimates computed on earlier data vintages. Both the new data releases and the revisions of previous data explain the downward revision.

REAL-TIME EVALUATION OF OPTIMISM/PESSIMISM AND UNCERTAINTY

While the ADS index measures the state of the economy and serves as a summary statistic of the information that market participants have received thus far about real activity, it is silent with respect to whether this information is in line with what market participants are expect-

Figure 5.3 ADS Index in Real Time, Recent Vintages

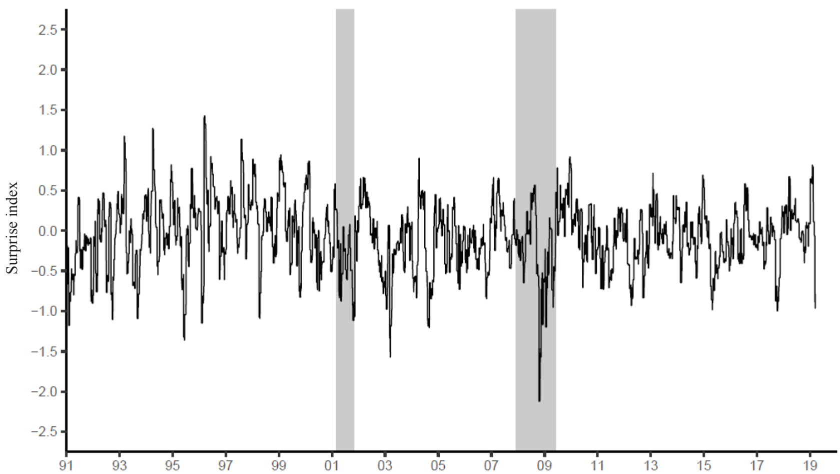


NOTE: The ADS indexes are constructed using data available up to the date indicated in the legend.
SOURCE: Federal Reserve Bank of Philadelphia, ADS Business Conditions Index.

ing and the uncertainty surrounding the data releases. The surprise and uncertainty indexes in Scotti (2016) speak to these issues. The surprise index summarizes recent economic data surprises and measures deviations from consensus expectations. A positive (negative) reading of the index indicates that agents were more pessimistic (optimistic), expecting economic data to be worse (better) than their actual realization. The uncertainty index measures the uncertainty related to the state of the economy. A greater (smaller) reading suggests that agents have, on balance, been more (less) uncertain about business conditions.

Figure 5.4 displays the surprise index computed as of March 2019. The surprise index reached its lowest value during the global financial crisis of 2008–2009, suggesting that as the crisis was unfolding, agents were less pessimistic about its possible outcome and its impact on the real economy. In contrast, the index turned positive during the beginning of 2019, indicating that agents were pessimistic about the state of the economy, harboring fairly low expectations relative to the actual

Figure 5.4 Scotti Surprise Index

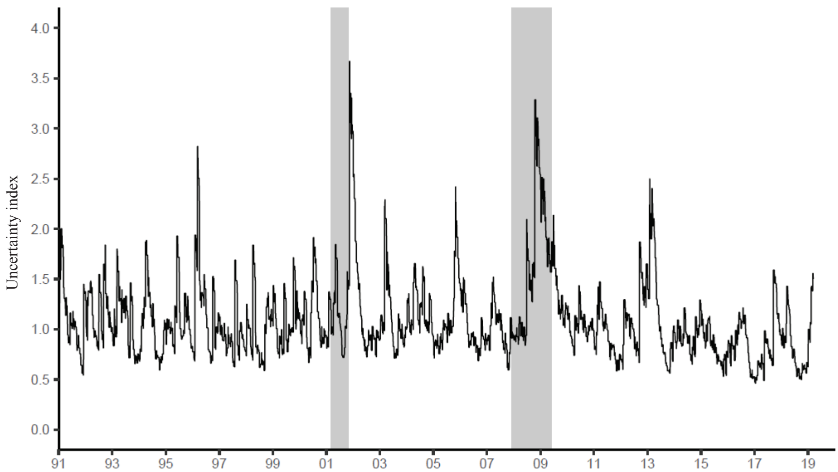


NOTE: The Scotti (2016) Surprise Index summarizes recent economic data surprises and measures deviations from consensus expectations. A positive (negative) reading of the surprise index indicates that agents were more pessimistic (optimistic), expecting economic data to be better (worse) than their actual realization.
SOURCE: Authors' calculations based on Bloomberg and NBER recession dates.

releases of GDP, IP, and nonfarm payroll. However, the index turned sharply negative in early March following the release of the February employment report.

Figure 5.5 portrays the uncertainty index, which tends to be elevated during recessions. Although there are some nonrecessionary periods in which the index spikes, it is interesting to note that the index increased in early 2019, reaching highs previously seen in 2017. This suggests that agents were less certain about the state of the economy.

Figure 5.5 Scotti Uncertainty Index



NOTE: The Scotti (2016) Uncertainty Index measures the uncertainty related to the state of the economy. A greater (smaller) reading suggests that agents have on balance been more (less) uncertain about business conditions.
SOURCE: Authors' calculations based on Bloomberg and NBER recession dates.

IS A RECESSION LOOMING?

Taken together, the several indexes presented so far suggest that business conditions in the United States have turned negative in early 2019, providing fertile ground for a downturn. Does this mean that a recession is looming?

In order to address such questions, we turn to the ADS index to directly inform the probability of a recession. First off, note that forecasting a recession is a hard task, given that the NBER tends to identify recessions only after a 12- to 18-month lag. As Hamilton (2011) puts it, “If people could predict recessions, they probably would not happen. Firms would not be stuck with inventories, labor, and capital they turn out not to need, and the Federal Reserve would probably ease its policy stance earlier.” Following a long-standing academic literature on estimating recession probabilities (see, for instance, Hamilton [2011] for an enlightening literature review), we use a probit model. We estimate the probability of being in a recession in the current quarter (that is, for a forecast horizon $h = 0$) only as a function of the ADS index instead of using a set of indicators. This stands as an alternative to customary estimates of recession probabilities, as can be seen from the review of relevant literature shown in Table 5.1, because we first embed in the ADS factor the information from the set of variables that constitute the ADS index, and then we feed the ADS index into the probit model. Our approach could be defined as “aggregate, then forecast,” paraphrasing the taxonomy laid out by Stock and Watson (2014).

Fossati (2016) estimates a small-scale dynamic factor model (DFM) in order to estimate a recession probability probit, but his DFM contains only monthly data (the same monthly data used in the ADS index), whereas the use of the ADS index allows us to automatically take care of mixed-frequency data and to include information on GDP.[4]

Figure 5.6 shows the recession probability based on the ADS index from 1960 to March 2019.[5] The recession probabilities computed before 2009 are based on the mid-March 2019 vintage of the ADS index, while the probabilities from 2009 to 2019 (to the right of the vertical red line) are computed in real time using the ADS vintage available at the time indicated on the x-axis. In other words, in the latter part of the sample, a probit model is recursively estimated with the new ADS index that summarizes available data up to a particular date. Generally, the model exhibits high spikes during NBER recession periods (the gray shaded areas). With an exception for the early-2000s recession, the estimated probability during all the recessions reached at least 75 percent. A word of caution on Figure 5.6: since the estimate of the probability is not in real time before 2009, looking at the estimate before the vertical red line could be deceiving if one intends to use the estimates to call recessions

Table 5.1 Summary of Literature on Predicting Recessions

Reference	Data scale	Data type	Data freq.	Model type	Forecast type	Real-time
Stock and Watson (1989)	Small	Real financial	Monthly	Factor from UC model	Backcasting Nowcasting	No
Stock and Watson (1991)	Medium	Real financial	Monthly	Logit + factor from UC model	Backcasting Nowcasting	No
Diebold and Rudebusch (1996)	Small	Financial	Quarterly	Markov switching DFM	None	No
Chauvet (1998)	Small	Real	Monthly	Markov switching DFM	Backcasting Nowcasting	No
Chauvet and Potter (2005)	Very small	Financial	Monthly	Dynamic probit	Backcasting Nowcasting	Yes
Harding and Pagan (2006)	Large	Real	Monthly	Dating algorithm	Backcasting	No
Wright (2006)	Very small	Real	Monthly	Probit	Forecasting	Yes
Chauvet and Piger (2008)	Small	Real	Monthly	Dating algorithm Markov switching DFM	Backcasting	Yes
Kauppi and Saikkonen (2008)	Very small	Financial	Monthly	Generalized dynamic probit	Backcasting Nowcasting	Yes
Berge (2015)	Medium	Real financial	Monthly	BMA logit	Backcasting Nowcasting	Yes
Fossati (2015)	Small/ large	Real	Monthly	Dynamic probit DFM Markov switching DFM	Forecasting	Yes
Giusto and Piger (2017)	Small	Real financial	Monthly	Vector quantization	Backcasting	Yes
This study	Small/ medium	Real financial/ other	Mixed: weekly/ monthly/ quarterly	BMA logit DFM	Nowcasting	Yes

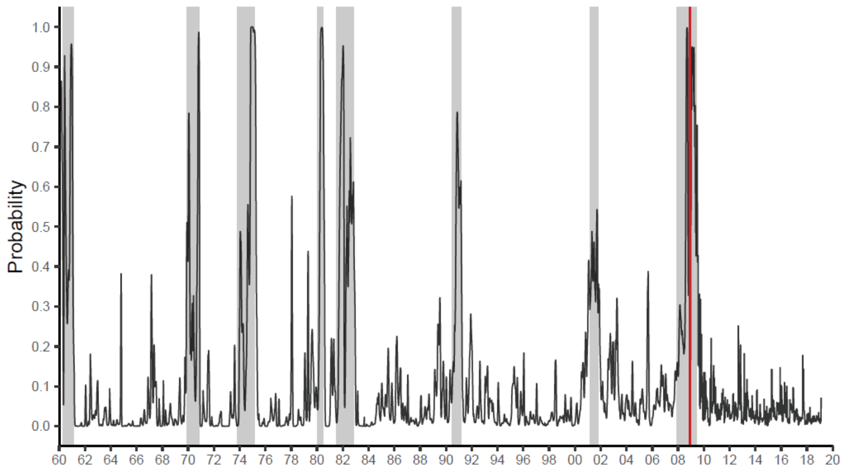
NOTE: The table presents a summary of the relevant literature on forecasting recessions at different horizons with macroeconomic and financial data. The model is formulated but not estimated. Financial data are not revised—therefore they are, by construction, real-time.

SOURCE: Authors' compilation.

in real time. That said, we verify that for the time window in which we have vintages of the ADS index, the real-time estimates are not drastically different from the estimate on the last vintage we used.

Figure 5.7 zooms in on the most recent 10 years and compares the recession probabilities estimated using the mid-March 2019 ADS vin-

Figure 5.6 Recession Probability Based on the ADS Index



NOTE: The recession probabilities computed based on the "Last Vintage ADS" use information as of mid-March 2019, whereas probabilities based on the "Real Time ADS" use the ADS vintages available at the time indicated on the x-axis. Grey shading indicates NBER-designated recessions.

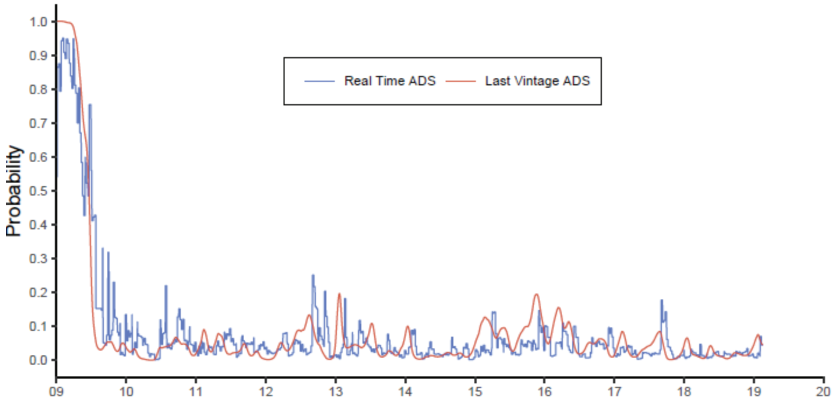
SOURCE: Authors' calculations based on Philadelphia Fed, ADS Business Conditions Index.

tage with the estimate using the real-time ADS index available at the time indicated by the x-axis. Although the real-time recession probability is more volatile when compared to the probability computed using the last vintage of data, the real-time estimate is substantially in agreement and provides the same signal as the estimate from the last vintage that we have used.

ADDITIONAL INDICATORS OF ECONOMIC ACTIVITY

The ADS index is a coincident indicator, and the probability computed above refers to the assessment that the economy is in recession in the corresponding month. This analysis could be expanded in two directions: 1) including alternative indicators as explanatory variables (individual macroeconomic series or financial series) and 2) forecasting

Figure 5.7 Recession Prob. Based on the ADS Index: Real-Time Versus Last-Vintage



NOTE: The recession probabilities computed before 2009 are based on the February 2019 vintage of the ADS index, while the probabilities from 2009 to 2019 are computed in real time using the ADS vintage available at the time indicated on the x-axis. Grey shading indicates NBER-designated recessions.
SOURCE: Authors' calculations based on Philadelphia Fed, ADS Business Conditions Index.

recession probabilities at different horizons. We investigate the second issue in the empirical analysis looking at a horizon of between 0 and 12 months. With respect to the first item, we discuss here some additional indicators that have been explored in the literature, starting from the seminal work by Stock and Watson (1989), and then we use them in conjunction with the ADS index to understand the quality of the incremental information that they provide in forecasting recessions. We find that financial indicators are very useful beyond the ADS index at longer forecast horizons.

The term spread. Stock and Watson (1989) introduce yield spreads—in particular the spread between 10-year and 1-year T-bonds—as useful indicators in forecasting economic activity and downturns. Estrella and Hardouvelis (1991) further explore the forecasting power of the *slope of the Treasury yield curve* as a leading indicator of downturns.[6] Corroborating the findings in these studies, academics and market participants point to the fact that a negative slope—a negative difference between a far-off maturity, typically 10 years, and a shorter maturity, typically between 2 years and 3 months—generally precedes economic recessions. Accordingly, recession probability models based uniquely on the term spread generally associate declining term spreads with an elevated probability of recession in the near/medium term.[7] For example, in early 2019, a simple probit model based only on the term spread would have predicted a much higher probability of recession in the next 12 months. In fact, the Federal Reserve Bank of Cleveland assessed the probability of being in a recession by January 2020 at about 30 percent using information from the yield curve as of mid-March 2019.[8] A variety of additional financial variables closely connected with the yield curve have been tested in forecasting turning points in the economy. For instance, Wright (2006) motivates the introduction of the average of the federal funds rate over a given quarter as it provides a measure of the impetus or restraint to the economy implied by the stance of monetary policy.[9] In addition, Wright (2006) also finds some evidence that a measure of expected excess returns on longer-maturity bonds, the *return forecasting factor* studied by Cochrane and Piazzesi (2005), is useful in predicting recessions. In our empirical evaluation in the section titled "Real-Time Measurement of Economic Conditions," we use both the level of the yield curve (as in Wright ([2006]) as well as slope and curvature.

Corporate bond spreads. Stock and Watson (1989) introduce yield spreads as useful indicators in forecasting economic activity and downturns, and they find that the spread between commercial paper and Treasury bills is a leading indicator of recessions. Several other authors have found that *corporate bond spreads*—also called *credit spreads*—are useful indicators in predicting recessions. Stock and Watson (1989) used the paper-bill spread, and Gertler and Lown (1999) studied the high-yield credit spread. Both of these spreads have predictive content on economic activity because they embed default risk, which incorporates investors' expectations of future corporate defaults. However, recent analysis in Gilchrist and Zakrajšek (2012) tries to distill the information on future economic activity in bond credit spreads beyond default risk and calls such a component the excess bond premium (EBP). Favara et al. (2016a) use the EBP in a probit model of recessions.[10] We include the most recent update of the EBP in the empirical explorations in the next section.

Other financial indicators. Beyond credit spreads, stock market returns and other financial information such as the implied stock market volatility—as measured by the Chicago Board Options Exchange volatility index (VIX)—have been used to predict economic activity. For instance, Danielsson, Valenzuela, and Zer (2018) use the VIX in trying to explain financial crises, and Engstrom (2014) uses option pricing in trying to predict stock market crashes. We include the VIX as a regressor in the models of the section "Real-Time Management of Economic Conditions."

USING THE ADS INDEX WITH ADDITIONAL INDICATORS IN FORECASTING RECESSIONS

Berge (2015) finds that the term spread and some additional financial variables are valuable leading indicators, but mostly at longer horizons—that is, at horizons of 6–12 months ahead, and beyond. Conversely, at shorter horizons, real variables appear to be more powerful in signaling recessions.

We follow Berge (2015) and use a series of logit models, one for each horizon from 0 to 12 months, and a variety of real and financial explanatory variables, in order to explore the usefulness of the ADS index. Through Bayesian Model Averaging (BMA), we select the models containing the most useful indicators at each forecast horizon, reducing the dimensionality of our big system of models (with N indicators, we have 2^N possible models).[11] The individual models are at the monthly frequency and contain a mixture of financial and real variables. Financial variables are as follows: the level, slope, and curvature of the yield curve; corporate bond spreads; the TED spread; the return on the S&P500; the trade-weighted dollar index; and the VIX. For macro indicators, we compare three different sets of variables:

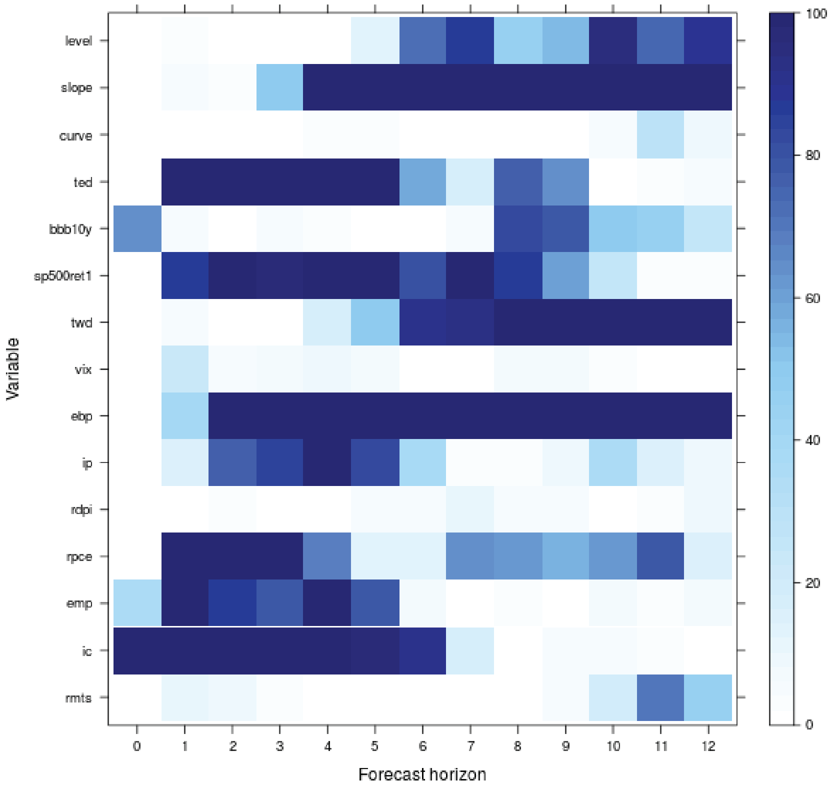
1) the ADS index;
2) the subset of real-activity variables used to compute the ADS index—specifically, nonfarm payroll, industrial production, retail sales, personal income, and a monthly average of initial jobless claims;[12]

3) a larger set of macro variables, including the variables that make up the ADS index, plus total light vehicle sales, the ISM purchasing managers' index, average weekly hours, housing permits, and the four-week moving average of unemployment claims.

Because the NBER announces turning points with a delay, and we do not want to incorrectly assume that a month in the recent past was or was not a recession (that is, assigning a 0 or 1 value to the logit dependent variable), we estimate the model up to December 2017 and use those parameter estimates in the evaluation of the recession probabilities. In this first step, our BMA approach selects the best combination of indicators at the different horizons, as shown by the heat map of the posterior inclusion probabilities in Figure 5.8. The darker color in the figure indicates a higher posterior probability that a particular variable (shown in the rows) is included in the model for that horizon (shown in the columns). We only show results for one of the models described above with financial and real variables.[13] It should be clear from the figure that there is a predominance of darker colors in the lower-left quadrant and in the top right, indicating that real variables have higher inclusion probabilities for shorter-horizon models, while financial variables have higher inclusion probabilities at longer horizons. An exercise in which we separate the estimation of a real-variable model and a financial-variable model highlights this finding even more, as already noted by Berge (2015). In fact, as shown in Figure 5.9, the Receiver Operating Characteristic (ROC) curves for the in-sample prediction from the 1- and 12-month-ahead models point to a superior performance of the real variables at the 1-month horizon, but a better performance of the financial indicators at the 12-month horizon.[14]

We then employ data up to December 2018 to forecast our indicators through December 2019, using a random walk, and compute the corresponding recession probabilities for each horizon based on the best model selected in the previous step. Figure 5.10 shows such probabilities for the three combinations of real variables outlined above. Interestingly, all three models point to an increased probability of recession in mid-2019, possibly due to the big correction observed in financial markets in late 2018. Of note, the model with the ADS index performs just as well as the model in which the underlying series enter one by one.

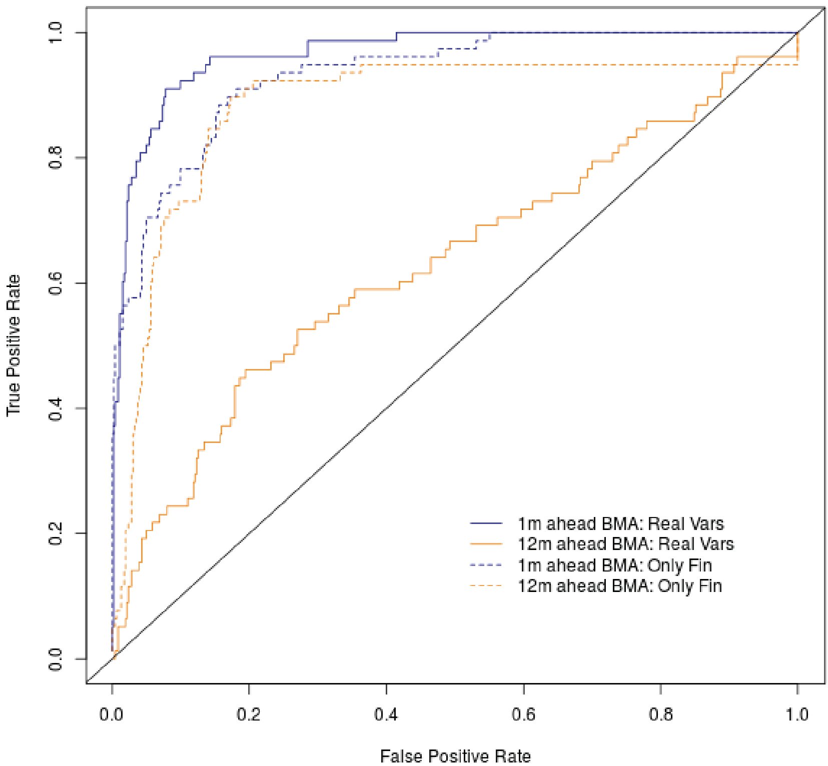
Figure 5.8 Heat Map of Posterior Inclusion Probabilities



NOTE: Posterior inclusion probabilities for each variable (rows) and each horizon (columns) related to the model with the ADS index and financial variables. Darker colors indicate a high inclusion probability.
SOURCE: Authors' calculations based on NBER recession dates, BEA real disposable income.

A drawback to this forecast exercise is that it does not allow for ragged edges in the data and mixed frequency, as it needs to stop at the last point in time for which all the series are available. To overcome this issue, we reestimate the recession probabilities employing only the ADS index as a summary statistic of the real indicators plus GDP, which is available in real time. The estimated probabilities based on the ADS index and financial variables as of mid-March are shown in Figure 5.11. Based on the additional information available between

Figure 5.9 Heat Map of Posterior Inclusion Probabilities, ROC Curves



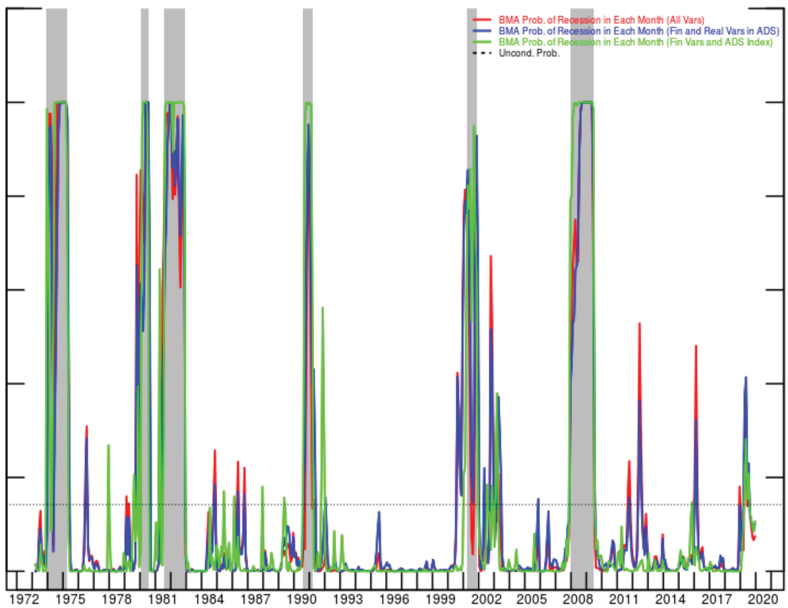
NOTE: ROC curves for 1- and 12-month-ahead models with real or financial indicators.
SOURCE: Authors' calculations based on NBER recession dates, BEA real disposable income.

the end of 2018 and mid-March, the probability of the NBER declaring a recession over the next year significantly decreased, in line with the ADS-only probability from Figure 5.7.

CONCLUSION

In this chapter, we update and evaluate a number of economic indicators as well as recession probability models. As pointed out by Berge

Figure 5.10 BMA Recession Probabilities, December 2018

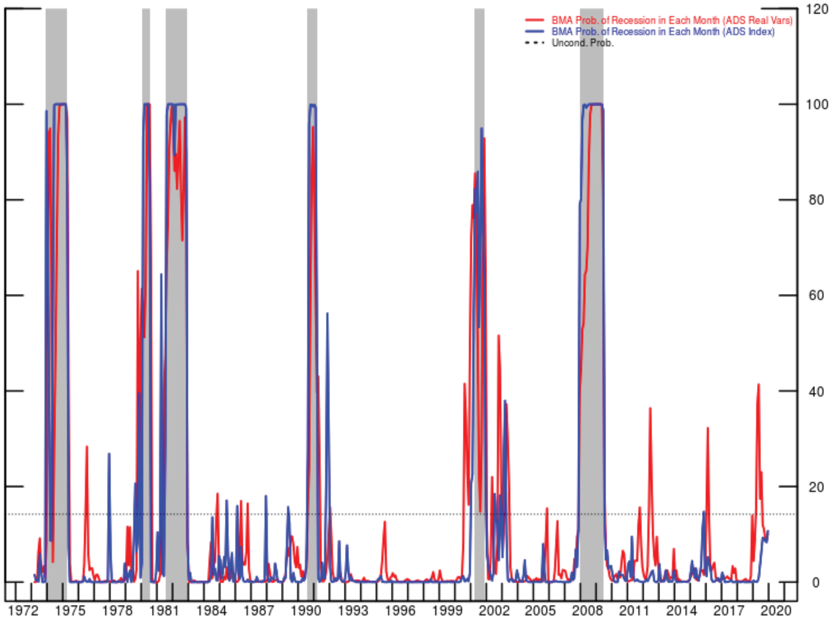


NOTE: Probability that the NBER will declare a recession in a particular month based on the BMA probit model. Forecast after December 2018 based on a random walk of the various indicators. Dotted line represents the unconditional probability.
SOURCE: Authors' calculations based on NBER recession dates, BEA real disposable income.

(2015), different mixtures of real and financial variables work best at different horizons, suggesting the need to maintain a set of models that work well at different forecasting horizons. Because of the inability of probit models to account for ragged edges, a real-time indicator of real activity like the ADS index might prove useful to have more up-to-date forecasts of recession probabilities.

The analysis of this topic, however, should not be limited to what is described above. For example, other indicators might be considered among the set of explanatory variables, along the lines of Engstrom and Sharpe (2018), who further qualify the most relevant term spread to forecast recessions. They argue that the *near-term forward spread*, computed as the difference between the implied forward rate on Trea-

Figure 5.11 BMA Recession Probabilities, Mid-March 2019



NOTE: Probability that the NBER will declare a recession in a particular month based on the BMA probit model. Forecast after March 2019 based on a random walk of the various indicators. Dotted line represents the unconditional probability.
SOURCE: Authors' calculations based on NBER recession dates, BEA real disposable income.

sury bills six quarters ahead and the corresponding yield on a three-month Treasury bill, is a better predictor compared to more traditional term spreads. The near-term forward spread can be interpreted as a measure of the market's expectations for the trajectory of conventional near-term monetary policy. When negative, it indicates that market participants expect monetary policy to ease on net over the next several quarters, presumably because they expect monetary policymakers to respond to the threat or onset of a recession.[15] The superiority of using forward rates comes from the fact that, because yields are averages of the forward rates spanning the period to maturity, they tend to be a noisier signal of the expected Fed actions. They corroborate this intuition by proving that their measure outperforms in sample the term spread of the

10-year Treasury constant maturity minus two-year Treasury constant maturity spread sometimes used in the classical probit model.

The national financial condition indicator (NFCI) maintained by the Federal Reserve Bank of Chicago could also potentially be used as a financial explanatory variable, just as the ADS index is used for real variables.[16] The NFCI index provides a comprehensive weekly update on U.S. financial conditions in money markets, debt and equity markets, and the traditional and "shadow" banking systems. Alternatively, because U.S. economic and financial conditions tend to be highly correlated, the adjusted NFCI (ANFCI)—an index that isolates a component of financial conditions uncorrelated with economic conditions to provide an update on financial conditions relative to current economic conditions—could be considered.

Recession probabilities based on macroeconomic and financial indicators could additionally be compared to news-count measures of recession probabilities, such as the LexisNexis index of Berge and Jordà (2011) or the Google trends recession index also reported in Berge and Jordà, and used in an original way to set priors of a Bayesian DFM in Monokroussos (2015). The horizon of these news-count probabilities is, however, not clear, as articles could talk about past, current, or future recessions. Therefore, a straight comparison with the measures described above might not be so straightforward.[17]

Notes

Alessandro Barbarino and Chiara Scotti are with the Federal Reserve Board, 20th and C Streets NW, Washington, DC 20551. The authors can be reached via email at alessandro.barbarino@frb.gov and chiara.scotti@frb.gov. We thank Charles Horgan for excellent research assistance. The opinions expressed here are our own and do not necessarily reflect the views of the Board of Governors or its staff. This chapter was originally prepared for the 55th annual Werner Sichel Lecture Series in 2018–2019 on "Alternative Economic Indicators" and delivered at Western Michigan University in March 2019.

1. See CBO (2019), https://www.cbo.gov/system/files?file=2019-01/54918-Outlook .pdf.
2. The ADS index is maintained and updated by the Federal Reserve Bank of Philadelphia at https://www.philadelphiafed.org/research-and-data/real-time-center/business-conditions-index.
3. See https://www.nber.org/cycles.html.

4. Fossati (2016) also estimates the factor from a larger scale DFM modeled after Stock and Watson (2002), and he finds in his out-of-sample exercise that the factor from the large-scale DFM performs better in forecasting recessions than the factor from the small-scale DFM. In this note, we did not compare the performance of the ADS index with the performance of the factor from a larger-scale DFM.
5. How negative should the ADS be in order to signal a recession? Berge and Jordà (2011) show that the level of the ADS index that maximizes the ROC when NBER recessions are the target is about −0.8017 (using the last vintage available to them and stopping the computation in December 2007).
6. See Estrella and Hardouvelis (1991) and Estrella and Mishkin (1998) for a thorough account of the earlier literature that links the term spread and real activity. More recently, papers that focus on the predictive power of the slope of the yield curve for recessions include Benzoni, Chyruk, and Kelley (2018); Chauvet and Potter (2005); Croushore and Marsten (2016); Kauppi and Saikkonen (2008); Rudebusch and Williams (2009); and Wright (2006).
7. Notice that the slope of the yield curve does not even need to be negative to obtain considerable spikes in the estimated probability.
8. See https://www.clevelandfed.org/our-research/indicators-and-data/yield-curve-and-gdp-growth.aspx.
9. Wright (2006) explores both the nominal federal funds rate and the real rate (for which inflation expectations are proxied by a four-quarter backward-looking moving average of the core personal consumption expenditures [PCE] price index).
10. Updated data on the EBP can be found in Favara et al. (2016b).
11. As in Berge (2015), we use the package BMA in R by Raftery et al. (2018) in order to reduce the dimensionality of the problem.
12. We exclude quarterly GDP, as the model is monthly.
13. Heat maps for the other models are available upon request from the authors.
14. An ROC curve illustrates the trade-off associated with achieving a particular true positive rate versus the corresponding false positive rate. The area under the ROC curve (AUC) is a summary statistic measuring the classification ability of an indicator/model. The higher the AUC, the better the classification ability.
15. The only noise in this measure would be term premiums or liquidity premiums embedded in shorter-term Treasury rates.
16. See https://www.chicagofed.org/publications/nfci/.
17. The progressive sophistication of available dictionaries for textual analysis will make it possible to sift through finer details in texts, which will circumvent these limitations.

References

Aruoba, S. Borayğan, Francis X. Diebold, and Chiara Scotti. 2009. "Real-Time Measurement of Business Conditions." *Journal of Business and Economic Statistics* 27(4): 417–427.

Bai, Jushan, and Serena Ng. 2006. "Confidence Intervals for Diffusion Index Forecasts and Inference for Factor-Augmented Regressions." *Econometrica* 74(4): 1133–1150.

Benzoni, Luca, Olena Chyruk, and David Kelley. 2018. "Why Does the Yield-Curve Slope Predict Recessions?" Chicago Fed Letter No. 404. Chicago: Federal Reserve Bank of Chicago.

Berge, Travis J. 2015. "Predicting Recessions with Leading Indicators: Model Averaging and Selection over the Business Cycle." *Journal of Forecasting* 34(6): 455–471.

Berge, Travis J., and Òscar Jordà. 2011. "Evaluating the Classification of Economic Activity into Recessions and Expansions." *American Economic Journal: Macroeconomics* 3(2): 246–277.

Chauvet, Marcelle. 1998. "An Econometric Characterization of Business Cycle Dynamics with Factor Structure and Regime Switching." *International Economic Review* 39(4): 969–996.

Chauvet, Marcelle, and Jeremy Piger. 2008. "A Comparison of the Real-Time Performance of Business Cycle Dating Methods." *Journal of Business and Economic Statistics* 26(1): 42–49.

Chauvet, Marcelle, and Simon Potter. 2005. "Forecasting Recessions Using the Yield Curve." *Journal of Forecasting* 24(2): 77–103.

Cochrane, John H., and Monika Piazzesi. 2005. "Bond Risk Premia." *American Economic Review* 95(1): 138–160.

Congressional Budget Office (CBO). 2019. *The Budget and Economic Outlook: 2019 to 2029*. Washington, DC: Congressional Budget Office.

Croushore, Dean, and Katherine Marsten. 2016. "Reassessing the Relative Power of the Yield Spread in Forecasting Recessions." *Journal of Applied Econometrics* 31(6): 1183–1191.

Danielsson, Jon, Marcela Valenzuela, and Ilknur Zer. 2018. "Learning from History: Volatility and Financial Crises." *Review of Financial Studies* 31(7): 2774–2805.

Diebold, F.rancis X., and Glenn D. Rudebusch. 1996. "Measuring Business Cycles: A Modern Perspective." *Review of Economics and Statistics* 78(1): 67–77.

Engstrom, Eric. 2014. "Forecasting Stock Market Crashes Is Hard—Especially Future Ones: Can Option Prices Help?" *FEDS Notes*, May 7.

Engstrom, Eric, and Steve Sharpe. 2018. "The Near-Term Forward Yield

Spread as a Leading Indicator: A Less Distorted Mirror." Finance and Economics Discussion Series 2018-055. Washington, DC: Board of Governors of the Federal Reserve System.

Estrella, Arturo, and Gikas A. Hardouvelis. 1991. "The Term Structure as a Predictor of Real Economic Activity." *Journal of Finance* 46(2): 555–576.

Estrella, Arturo, and Frederic S. Mishkin. 1998. "Predicting U.S. Recessions: Financial Variables as Leading Indicators." *Review of Economics and Statistics* 80(1): 45–61.

Favara, Giovanni, Simon Gilchrist, Kurt F. Lewis, and Egon Zakrajšek. 2016a. "Recession Risk and the Excess Bond Premium." *FEDS Notes*, April 8. Washington, DC: Board of Governors of the Federal Reserve System.

———. 2016b. "Updating the Recession Risk and the Excess Bond Premium." *FEDS Notes*, October 6. Washington, DC: Board of Governors of the Federal Reserve System.

Forni, Mario, Marc Hallin, Marco Lippi, and Lucrezia Reichlin. 2000. "The Generalized Dynamic-Factor Model: Identification and Estimation." *Review of Economics and Statistics* 82(4): 540–554.

Fossati, Sebastian. 2015. "Forecasting U.S. Recessions with Macro Factors." *Applied Economics* 47(53): 5726–5738.

———. 2016. "Dating U.S. Business Cycles with Macro Factors." *Studies in Nonlinear Dynamics and Econometrics* 20(5): 529–547.

Gertler, Mark, and Cara S. Lown. 1999. "The Information in the High-Yield Bond Spread for the Business Cycle: Evidence and Some Implications." *Oxford Review of Economic Policy* 15(3): 132–150.

Gilchrist, Simon, and Egon Zakrajšek. 2012. "Credit Spreads and Business Cycle Fluctuations." *American Economic Review* 102(4): 1692–1720.

Giusto, Andrea, and Jeremy Piger. 2017. "Identifying Business Cycle Turning Points in Real Time with Vector Quantization." *International Journal of Forecasting* 33(1): 174–184.

Hamilton, James D. 2011. "Calling Recessions in Real Time." *International Journal of Forecasting* 27(4):1006–1026.

Harding, Don, and Adrian Pagan. 2006. "Synchronization of Cycles." *Journal of Econometrics* 132(1): 59–79.

Kauppi, Heikki, and Pentti Saikkonen. 2008. "Predicting U.S. Recessions with Dynamic Binary Response Models." *Review of Economics and Statistics* 90(4): 777–791.

Monokroussos, George. 2015. "Nowcasting in Real Time Using Popularity Priors." MPRA Paper No. 68594. University Library of Munich, Germany.

Raftery, Adrian, Jennifer Hoeting, Chris Volinsky, Ian Painter, and Ka Yee Yeung. 2018. "Bayesian Model Averaging (BMA) R Package." Seattle: University of Washington.

Rudebusch, Glenn D., and John C. Williams. 2009. "Forecasting Recessions: The Puzzle of the Enduring Power of the Yield Curve." *Journal of Business and Economic Statistics* 27(4): 492–503.

Scotti, Chiara. 2016. "Surprise and Uncertainty Indexes: Real-Time Aggregation of Real-Activity Macro-Surprises." *Journal of Monetary Economics* 82(C): 1–19.

Stock, James H., and Mark W. Watson. 1989. "New Indexes of Coincident and Leading Economic Indicators." In *NBER Macroeconomics Annual*, Vol. 4, Olivier Jean Blanchard and Stanley Fischer, eds. Cambridge, MA: MIT Press, pp. 351–409.

———. 1991. "A Probability Model of the Coincident Economic Indicators." In *Leading Economic Indicators: New Approaches and Forecasting Records*, Kajal Lahiri and Geoffrey H. Moore, eds. Cambridge: Cambridge University Press, pp. 63–89.

———. 2002. "Macroeconomic Forecasting Using Diffusion Indexes." *Journal of Business and Economic Statistics* 20(2): 147–162.

———. 2014. "Estimating Turning Points Using Large Data Sets." *Journal of Econometrics* 178(P2): 368–381.

Wright, Jonathan H. 2006. "The Yield Curve and Predicting Recessions." Working paper, Finance and Economics Discussion Series. Washington, DC: Board of Governors of the Federal Reserve System.

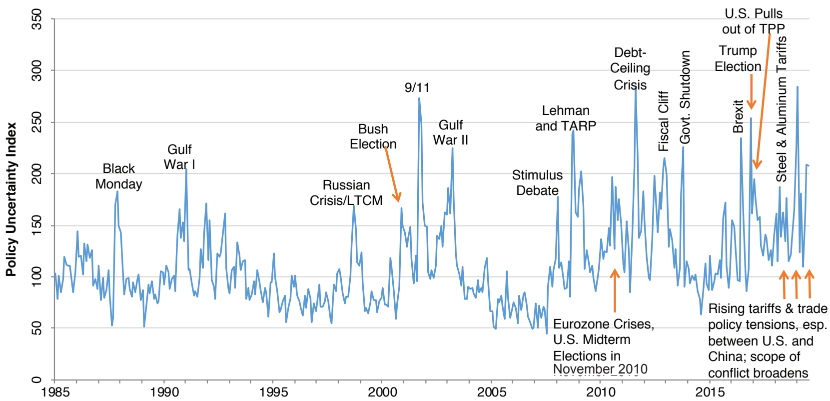
6
Rising Policy Uncertainty

Steven J. Davis
University of Chicago

Figure 6.1 displays a monthly index of economic policy uncertainty (EPU) for the United States that I developed with Scott Baker and Nick Bloom.[1] In constructing this index, we aim to capture uncertainty about *who* will make economic policy decisions, *what* economic policy actions will be undertaken and *when*, and the economic *effect*s of policy actions (or inaction)—including uncertainties related to the economic ramifications of "noneconomic" policy matters, e.g., military actions. To do so, we first count articles in 10 leading U.S. newspapers that contain the following triple of terms: 1) "economic" or "economy"; 2) "uncertain" or "uncertainty"; and 3) one or more of "Congress," "deficit," "Federal Reserve," "legislation," "regulation," or "White House." Next, we divide the raw EPU count by the number of all articles in the same paper and month, standardize the variability of the scaled EPU counts, and average over newspapers by month. Finally, we normalize the mean index value to 100 from 1985 to 2009. Thus, the index value of 284 in January 2019 is 2.84 times its 1985–2009 average.

Our U.S. EPU index spikes near tight presidential elections, Gulf Wars I and II, the 9/11 attacks, major fiscal policy battles from 2011 to 2013, and in reaction to the June 2016 Brexit referendum, Donald Trump's surprise election victory in November 2016, and escalating trade policy tensions in 2018 and 2019. The EPU index tends to rise in recessions, but many of the largest spikes and highest index values occur during the long, ongoing expansion that began in the second half of 2009. Measures of policy uncertainty derived from textual analysis of the Federal Reserve System's periodic *Beige Books* and from transcripts of quarterly earnings conference calls of publicly listed firms exhibit similar time-series patterns.[2] Baker et al. (2014) present and discuss evidence that policy-related economic uncertainty in the United

Figure 6.1 U.S. Economic Policy Uncertainty Index, 1985 to July 2019



NOTE: Monthly data normalized to 100 from 1985 to 2009.
SOURCE: Baker, Bloom, and Davis (2016), as updated at PolicyUncertainty.com.

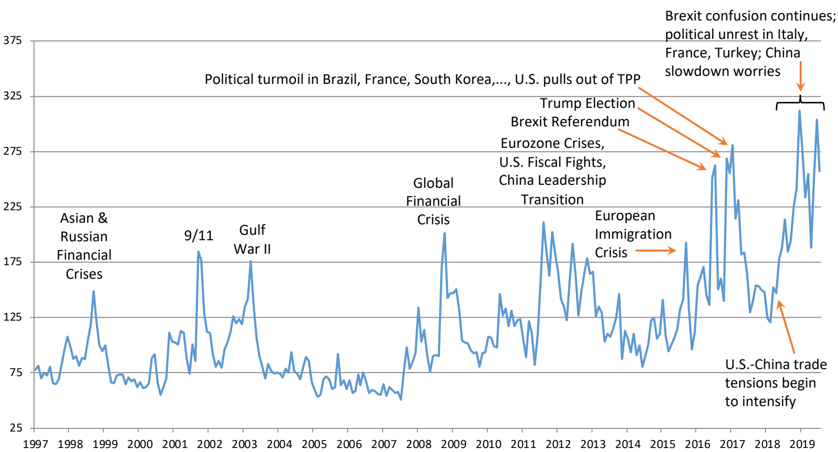
States followed an upward trajectory in the 1960s and 1970s, stabilized somewhat in the 1980s and 1990s, and rose again after the late 1990s.

Using the same approach, we and others construct monthly newspaper-based EPU indices for 20 additional countries: Australia, Brazil, Canada, Chile, China, Colombia, France, Germany, Greece, India, Ireland, Italy, Japan, Mexico, the Netherlands, Russia, South Korea, Spain, Sweden, and the United Kingdom.[3] We rely on own-country newspapers in constructing the national EPU indices and perform all searches in the language of the newspaper. To help develop suitable E, P, and U term sets, we consulted persons with native fluency and economics expertise in the relevant language and country. Our P term set differs across countries for reasons both obvious (e.g., using "BOJ" for Japan) and idiosyncratic (e.g., inclusion of "customs duties" for India). Monthly data for all 21 national EPU indices are available and regularly updated at www.PolicyUncertainty.com.

To construct an index of global economic policy uncertainty, I proceed as follows in Davis (2016): First, I renormalize each national EPU index to a mean of 100 from 1997 (or first year) to 2015. Second, I impute missing values for certain countries using a regression-based method.[4] This step yields a balanced panel of monthly EPU index values for 21 countries from January 1997 to the present. Third, I compute the global EPU index value for each month as the GDP-weighted average of the 21 national EPU index values, using GDP data from the IMF's World Economic Outlook Database. Figure 6.2 plots the resulting index.

The global EPU index rises sharply in reaction to the Asian and Russian financial crises, the 9/11 terrorist attacks, the U.S.-led invasion of Iraq in 2003, the global financial crisis in 2008–2009, the European immigration crisis in 2015, and several other developments.[5] It fluctuates around consistently high levels from mid-2011 to early 2013, a period characterized by recurring sovereign debt and banking crises in the Eurozone, intense battles over fiscal and health-care policies in the United States, and a generational leadership transition in China. Seven of nine members on the Standing Committee, China's most powerful decision-making body, were slated for retirement in 2012. Other senior leadership bodies in China experienced similarly high turnover rates because of retirement, leading Li (2011) to characterize 2012 as the fourth "generational transfer of power" in the history of Communist

Figure 6.2 Global Economic Policy Uncertainty Index, January 1997 to July 2019



NOTE: Using data for 21 countries that account for 80 percent of global GDP at current prices. Normalized to 100 from 1997 to 2015.
SOURCE: Baker, Bloom, and Davis (2016), as updated at PolicyUncertainty.com.

China. Two previous generational transitions coincided with tragedy and turmoil in the form of the Cultural Revolution and the 1989 Tiananmen Square protests and massacre.

Eurozone developments in the period from mid-2011 to early 2013 include a rescue package for Portugal in May 2011, a bailout package for Greece in July 2011 (amid widespread speculation that Greece would leave the Eurozone), large yield increases on Spanish and Italian government bonds in August 2011, April 2012, and June 2012, a May 2012 election in which most Greek voters rejected a proposed bailout agreement, and multiple extraordinary actions by the European Central Bank in response to these and other developments.

Across the Atlantic, bitterly partisan disputes over the direction of U.S. fiscal policy led to a "debt ceiling" fight in summer 2011 that threatened to curtail critical government functions and delay payments on U.S. Treasury securities, and an extraordinary "fiscal cliff" episode with last-minute resolutions of major uncertainties about tax and spending policies. Uncertainties surrounding U.S. health-care policy were also extraordinarily high in this period. For example, an appellate court struck down the Affordable Care Act ("Obamacare") in August 2011, concluding that Congress lacked the constitutional authority to require individuals to purchase health insurance or pay a penalty, and threatening the viability of the entire act (Cooper 2011). The issue remained unsettled until June 2012, when the Supreme Court reversed the appellate court in a surprise, closely divided decision (Bravin and Radnofsky 2012).

Russia's annexation of Crimea in 2014 and its military incursions in eastern Ukraine led to international sanctions and an uncertain environment that curtailed foreign investment in Russia and contributed to its weak economic performance (European Parliamentary Research Service 2016). The Russia-Ukraine conflict and its unsettled nature harmed the Ukrainian economy as well and deterred foreign investment there (Morelli 2016). China's aggressive pursuit of sovereignty claims in the South China Sea has raised concerns about threats to ship-borne trade in some of the world's busiest international waters (Schonhardt and Chaturvedi 2016). Recent geopolitical tensions in the Persian Gulf, U.S.-led economic sanctions on Iran, and the Iranian seizure of oil tankers in the Strait of Hormuz have renewed concerns about petroleum supplies (Rachman 2019).

Syria has been the epicenter of a many-sided military conflict and humanitarian catastrophe since 2011, with devastating consequences and highly uncertain long-term implications. The catastrophe produced a flood of migrants into neighboring countries and Europe in 2014 and 2015, stoking security fears, creating anxiety about social and economic consequences, and placing enormous strains on the Schengen Area arrangements for free mobility in a borderless Europe (Baker, Bloom, and Davis 2015; *BBC News* 2016; Dustmann et al. 2016; Halla, Wagner, and Zweimüller 2015).

Several major political and policy developments have rocked national economies and the global economic outlook since 2016. Leading examples include the June 2016 Brexit referendum, Donald Trump's upset electoral win in November of that year, and the strength of populist political movements in several European countries. These developments have injected new sources of political and economic uncertainties into the global economy.

There are many other recent examples of economic uncertainty emanating from political developments. In South Korea, political scandal led to the impeachment of President Park Geun-hye in December 2016 and her removal from office in March 2017. In Brazil, a long and severe recession, an extraordinary wave of corruption investigations, the criminal convictions of many leading political figures, and the impeachment and removal of President Dilma Rousseff in 2016 combined to upend the political landscape. Brazil's new president, Michel Temer, has promised to restore growth by reversing several major policies of his predecessor (*Economist* 2016). In Argentina, new fears that the Peronist party would regain political power in upcoming elections triggered a spectacular 15 percent depreciation of the Argentine peso on August 12, 2019, and a 38 percent plunge in the stock market the same day (Dube and Lewis 2019a,b; Mander 2019).

In Turkey, after squashing an attempted coup d'état in July 2016, the government set about arresting and firing more than 100,000 teachers, military officers, judges, mayors, civil servants, and others (Arango, Yeginsu, and Timur 2016; Yeginsu 2016). Ten weeks after the failed coup, Moody's Investor Service downgraded Turkey's sovereign credit rating, citing external funding risks, slowing growth, and "further concerns regarding the predictability and effectiveness of government policy and the rule of law" (Reuters 2016). The coup attempt and its aftermath have also strained international relations between Turkey and several of its allies (Birnbaum and DeYoung 2016). Intense pressures on the foreign exchange value of the Turkish lira have prompted dramatic policy moves by the Turkish Central Bank since 2018 and the dismissal of the bank's governor in July 2019 (Coskun 2019; Gauthier-Villars and Sindreu 2018; and Kantchev and Hannon 2019).

THE NEW PROMINENCE OF TRADE POLICY UNCERTAINTY

Trade policy has become both more uncertain and more protectionist under the Trump presidency. Particulars include the U.S. withdrawal from the Trans-Pacific Partnership (TPP) Agreement in January 2017, President Trump's early threats to jettison the North American Free Trade Agreement (NAFTA), doubts about U.S. congressional ratification of a NAFTA-replacement treaty, and a large number of tariff hikes, tariff threats, and tariff reversals. The average U.S. tariff rate rose from less than 2 percent in December 2017 to 4 percent in May 2019. It was slated to reach an estimated 5–8 percent by the end of 2019.[6] Threats of additional tariff hikes would have, if fully implemented, brought the average U.S. tariff rate to an estimated 9–11 percent by the end of 2019.[7] The trade-weighted average U.S. tariff on Chinese imports rose from 3.1 percent in 2017 to 12.4 percent in 2018 and 18.3 percent in May 2019. Current U.S. plans would take the average U.S. tariff rate on Chinese imports to an estimated 21.4 percent in December 2019 (Brown and Zhang 2019).

U.S. tariff hikes and President Trump's rhetorical attacks on trading partners invite retaliation. Indeed, tit-for-tat tariff hikes between the United States and China have been underway since April 2018. Canada, India, Mexico, Turkey, and the European Union have also imposed new tariffs on American imports in reaction to U.S. tariff hikes. In short, the shift to greater protectionism in U.S. trade policy has prompted other countries to respond in kind.

Trump administration officials often assert that the administration's aggressive trade policy stance will yield new trade deals that are more favorable to the United States. Developments to date offer scant support for this assertion. The U.S. and South Korea renegotiated and signed a revised Korea-U.S. Free Trade Agreement in September 2018, but the new agreement involves "only limited changes to the original pact" (Schott and Jung 2018). On November 30, 2018, Canadian Prime Minister Trudeau, Mexican President Peña Nieto, and U.S. President Trump signed the U.S.-Mexico-Canada Agreement (USMCA) to replace NAFTA. For months its ratification seemed unlikely; however, it was finally ratified by the U.S. Congress in January of this year.[8] The tone of recent statements from the U.S. and Chinese sides suggests dim prospects in the next few months for a significant resolution of outstanding trade policy conflicts and a reversal of recent tariff hikes. Nor is there any sign that the United States will soon resolve its trade policy conflicts with India, Turkey, or the European Union.

Trade policy under the Trump administration also has a capricious, back-and-forth character that amplifies uncertainty and undermines a rules-based trading order. Less than three months after withdrawing from the TPP, President Trump said he would consider rejoining for a substantially better deal, only to throw cold water on the idea a few days later (Trump 2018; Ungku and Greenfield 2018). Robert Lighthizer, the U.S. trade representative, justified steel tariffs on the laughable grounds that Canada, for example, presents a national security threat (Press 2018). But President Trump tweeted that tariffs on Canadian steel were really a response to Canadian tariffs on U.S. dairy products (Byrd 2018). In August 2018, the president, for reasons unclear, tweeted that he had "just authorized a doubling of Tariffs on Steel and Aluminum with respect to Turkey" (Ballhaus and Schlesinger 2018).

Under President Trump, tariffs are threatened, announced, delayed, reversed, announced again, imposed, and removed—often in quick succession. Some countries get tariff exemptions, some don't. Exemptions vary in duration, and they come and go in a head-spinning manner. The recent treatment of steel imports exemplifies this aspect of U.S. trade policy under President Trump. See Brown and Kolb (2019) for a detailed account.

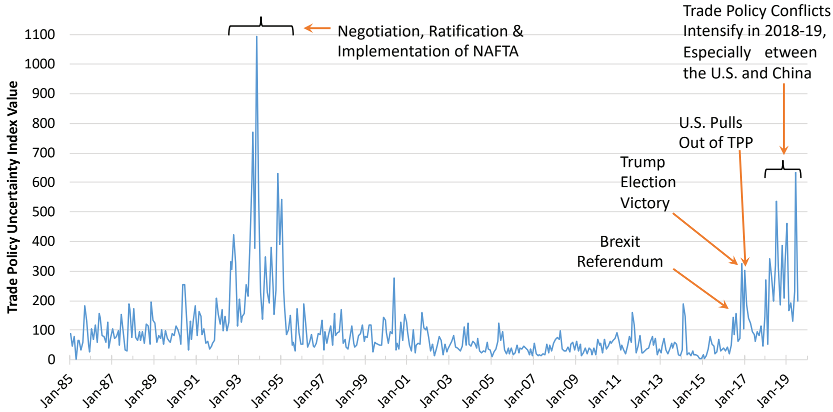
Another example involves the latest round of announcements about new tariffs on Chinese imports, which Kubota (2019) summarizes this way: "On Aug. 1, President Trump abruptly announced on Twitter that he would impose on Sept. 1 a 10 percent levy on roughly $300 billion in Chinese goods, an apparent response to what he described as China's failure to commit to promised U.S. agricultural purchases." Less than two weeks later, the plan was revised to "impose 10 percent tariffs on $112 billion of Chinese imports starting on September 1 . . . followed by a second round of duties on a different set of products, covering $160 billion of imports, on December 15" (Brown 2019, p. 1). China retaliated on August 23, announcing plans to levy new tariffs of 5 to 10 percent on $75 billion in U.S. imports. President Trump responded later the same day, announcing that he would raise existing and planned tariffs on $550 billon of Chinese imports by an additional 5 percentage points (Mauldin, Leary, and Deng 2019).

Trump administration trade policy also gives greater discretion over tariffs to bureaucrats, creating added complexity and uncertainty for individual businesses and compelling them, as a matter of business necessity, to become enmeshed in the tariff-exemption process. For example, the Department of Commerce rolled out a slow-working, burdensome process for requesting company-specific exemptions from steel and aluminum tariffs, as neatly recounted in the *Wall Street Journal (Wall Street Journal* Editorial Board 2018):

> Companies must submit a request attesting that their imports aren't made in the U.S. in "a satisfactory quality" or "sufficient and reasonably available amount." Companies must state the uses for their steel product, their average annual consumption of the product, as well as the number of days required to take delivery, manufacture and ship the product. They must also estimate the maximum and minimum composition of 24 chemical elements in their products including molybdenum, antimony and vanadium. There are dozens of other queries, but we'll spare you.
>
> Oh, and a separate request is required for each width, length, grade shape, and form of steel or aluminum product. A single company, Primrose Alloys, has submitted more than 1,200 steel product requests, according to Commerce's database. All 14 that have been reviewed so far were denied.
>
> Businesses may also submit statements to support their requests, which naturally turn political….

These various developments have led to a tremendous upsurge in anxiety and uncertainty about trade policy and its economic fallout. To attach some numbers to this point, Figure 6.3 displays a newspaper-based index of trade policy uncertainty (TPU) for the United States.

Figure 6.3 U.S. Trade Policy Uncertainty Index, January 1985 to July 2019



NOTE: Monthly data normalized to 100 from 1985 to 2009.
SOURCE: Baker, Bloom, and Davis (2016), as updated at PolicyUncertainty.com.

The TPU index reflects the frequency of articles in U.S. newspapers that discuss economic policy uncertainty *and* trade policy matters.

Two periods stand out. The first runs from August 1992 to March 1995 and reflects uncertainties around the negotiation, ratification, and implementation of the North American Free Trade Agreement (NAFTA). The second commences with Donald Trump's election victory in November 2016. The TPU index rose above 300 in reaction to the election outcome, the U.S. withdrawal from the TPP Agreement in January 2017, and U.S. tariffs on steel and aluminum imports imposed in March 2018. It rose even higher later in 2018 and in 2019, as U.S.-China trade policy conflicts intensified. The TPU index value averaged 301 from March 2018 to July 2019—7.7 times its average from 2013 to 2015 and 5.3 times its average from 1996 to 2015.

Table 6.1 presents evidence on the new prominence of trade policy uncertainty in China and Japan as well as the United States. Like Figure 6.3, Table 6.1 relies on frequency counts of own-country newspaper articles about trade policy uncertainty, but the scaling is different. In Figure 6.3, the raw TPU counts are scaled by the count of all articles in the same newspapers and normalized to 100 from 1985 to 2009. In con-

Table 6.1 Trade Policy Share of EPU Articles, Selected Time Periods for Three Major Economies

Time period	United States	Japan	China
1987–2015	4	8	16
2000–2015	2	7	20
NAFTA: Jan. 1992 to Dec. 2002	11	11	10
China WTO Accession: Jan. 2000 to Dec. 2002	3	5	36
November 2016 to December 2018	15	27	48
March–December 2018	15	29	48
January–July 2019	12	29	42

NOTE: Table entries report the percentage of articles about economic policy uncertainty that discuss trade policy matters in leading newspapers for the indicated countries.
SOURCE: Tabulated from data developed by Baker, Bloom, and Davis (2016) for the United States, Arbatli et al. (2019) for Japan, and Davis, Liu, and Sheng (2019) for China.

trast, Table 6.1 reports the percentage of EPU articles that discuss trade policy matters. All three countries show a dramatic rise in this percentage since November 2016, even more so since March 2018. Consider, for example, a comparison of the 2000–2015 period to the period from March to December in 2018: the trade policy share of EPU articles rose from 2 to 15 percent in the United States, from 7 to 27 percent in Japan, and from 20 to 48 percent in China. These comparisons support two conclusions: first, the rise in trade policy uncertainty under the Trump presidency has reverberated globally; second, the level of anxiety about trade policy is higher for major U.S. trading partners.

Trade policy concerns have also become a major source of stock market gyrations since 2018. For example, the S&P 500 index fell more than 2.5 percent on March 22, 2018, reacting to news about new U.S. tariffs on tens of billions of dollars of Chinese imports. Four days later, the index rose more than 2.7 percent on news that the U.S. and China had begun trade negotiations. Nevertheless, tariffs and tariff threats between the two countries have ratcheted upward in the ensuing 15 months.

In Baker, Bloom, Davis, and Sammon (2019), my coauthors and I examine the role of trade policy developments and 15 other news categories in large daily stock market moves. We first identified every daily move of more than 2.5 percent, up or down, in the U.S. stock market. By this criterion, there were 1,114 large daily moves from January 1900 to July 2019. For each large move, we read next-day news articles in the *Wall Street Journal* to classify perceptions of what moves the market.

Table 6.2 summarizes our evidence regarding the role of trade policy as a trigger for large daily moves in the U.S. stock market. The *Journal* attributed 7 of 1,103 large moves from 1900 to February 2018 mainly to news about trade policy, as compared to 4 of 11 large moves from March 2018 to July 2019.[9] By this metric, the prominent role of trade policy in recent U.S. stock market swings is historically unprecedented. In a complementary analysis, Huang et al. (2018) examine firm-level equity returns from March 21 to March 23. They find larger negative returns for U.S.-listed firms having greater exposure to trade with China over this period and larger negative returns for Chinese-listed firms with greater sales to the United States.

In Baker, Bloom, Davis, and Sammon (2019), my coauthors and I take a different approach to the analysis of newspaper content. We first use automated methods to identify articles about stock market volatility in 11 leading U.S. newspapers and to construct an "equity market volatility" (EMV) tracker. Our newspaper-based EMV tracker performs well in the sense that it moves closely with actual stock market volatility. Parsing the text in the EMV articles, we then quantify journalists' perceptions of what drives volatility in equity returns and classify the drivers into about 30 categories, many of which pertain to particular

Table 6.2 Trade Policy News Jolted the U.S. Stock Market in 2018 and 2019

	Number of daily stock market jumps greater than 2.5%	Number attributed to trade policy news	Percent (%)
January 1900 to February 2018	1,103	7	0.6
March 2018 to July 2019	11	4	35.7

NOTE: Table reports total number of jumps, up or down, in the indicated time periods and the number attributed primarily to news about trade policy, according to the human readings.

SOURCE: Tabulation of results from Baker, Bloom, Davis, and Sammon (2019).

types of policy. This approach lets us assess the importance of each category to the average level of stock market volatility and its movements over time.

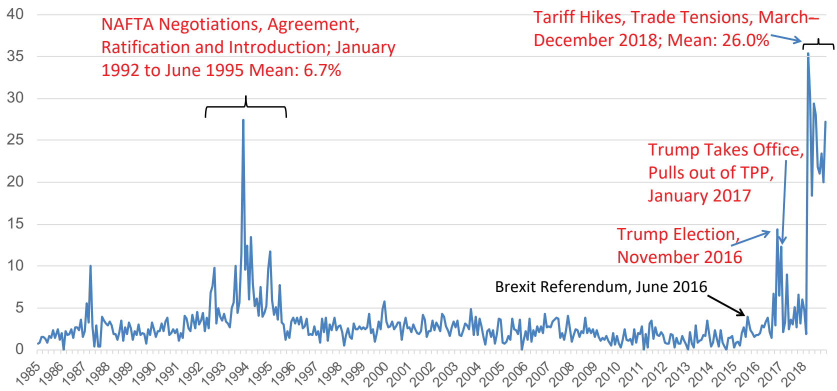
As seen in Figure 6.4, trade policy gets attention in 26 percent of articles about equity market volatility in leading U.S. newspapers from March to December 2018. In glaring contrast, trade policy matters receive attention in a mere 2.7 percent of articles about equity market volatility from 1985 to 2015.[10] In other words, trade policy went from a virtual nonfactor in U.S. equity market volatility in recent decades to one of its leading sources in 2018.

THE INTERPLAY BETWEEN POLICY UNCERTAINTY AND ECONOMIC PERFORMANCE

Politics and policy decision-making are often messy and fraught with uncertainty about political outcomes, policy decisions, near-term consequences, and long-term implications. The previous two sections offer a variety of examples, many of them recent, drawn from countries around the world. They include the U.S. debt-ceiling crisis in August 2011, the U.S. fiscal cliff and government shutdown episodes in 2012 and 2013, the Syrian catastrophe, multiple Eurozone crises since 2010, Russian military incursions in Ukraine, the European immigration crisis, the ongoing Brexit saga, a coup attempt and crackdown in Turkey, the removal of South Korea's president, corruption scandals and presidential removal in Brazil, a sharp escalation of U.S.-China trade policy conflicts in 2018 and 2019, and more. These examples illustrate the role of governments and political processes as sources of economic uncertainty. That uncertainty weighs negatively on economic performance. At least in a proximate sense, causality runs from political processes and policy uncertainty to aggregate economic performance in these examples.

Economic developments also give rise to uncertainty, directly and through their impact on policy making. As a leading example, the global financial crisis of 2007–2009 confronted policymakers with extraordinary and complex challenges, especially in the immediate wake of the financial panic in September 2008. There was great uncertainty about how policymakers should and would respond, and what would be the

Figure 6.4 Percent of Articles about Equity Market Volatility in Leading U.S. Newspapers That Discuss Trade Policy Matters, 1985 to 2018



NOTE: Computed from automated readings of newspaper articles about equity market volatility and (equity market volatility + trade policy) in 11 major newspapers.
SOURCE: Baker, Bloom, Davis, and Kost (2019).

economic consequences. In this episode, the financial crisis and its economic fallout drove a sharp rise in policy uncertainty. In turn, high policy uncertainty contributed to the severity of the crisis and the weakness of the ensuing recovery.

There is also evidence that major financial crises lead to higher levels of policy uncertainty for many years. Funke, Schularick, and Trebesch (2016) draw on data for many countries over 140 years to document a pattern of rising political polarization in the years following systemic financial crises, contributing to higher levels of policy uncertainty. Mian, Sufi, and Trebbi (2014) also find evidence that financial crises breed political polarization, which sometimes results in political gridlock and policy uncertainty.

A key point: the potential for negative shocks to drive policy uncertainty depends on the underlying environment, which is partly shaped by past policy decisions.[11] Consider again the global financial crisis. It was precipitated by a collapse in U.S. housing prices and mortgage-backed security values (Mian and Sufi 2014). The shock was large, and many banks were highly exposed to it. The shock led to a systemic financial crisis, because banks were poorly capitalized and heavily dependent on flight-prone forms of debt to fund their investments. If policymakers had required banks to rely more heavily on run-proof funding, the crisis would have been less severe—and perhaps would have been avoided altogether. In this and other respects, the precrisis regulatory regime set the stage for a major financial crisis (Admati and Hellwig 2013; Duffie 2019) and the ensuing policy uncertainty.

As another example, there is less need for discretionary fiscal stimulus in response to negative shocks when robust automatic fiscal stabilizers are in place. Automatic fiscal stabilizers lessen the political conflicts, decision delays, implementation lags, and policy uncertainty that come with efforts to deploy discretionary fiscal tools. Especially when monetary policy is hampered by an effective lower bound on policy rates, inadequate or poorly designed automatic fiscal stabilizers practically ensure that political leaders will turn to discretionary fiscal policy as a response to the next large economic downturn.

High policy uncertainty in the past decade has stimulated empirical research on its economic consequences. Durnev (2012), Giavazzi and McMahon (2012), Julio and Yook (2012, 2016), and Kelly, Pastor, and Veronesi (2016), among others, investigate the effects of election-related uncertainty on corporate investment, international capital flows, precautionary savings, and stock price volatility. By and large, this literature finds that election-related uncertainty reduces investment, discourages inward foreign direct investment (FDI), raises precautionary savings, and increases stock price volatility. Aaberge, Liu, and Zhu (2017) find that political uncertainty associated with the 1989 Tiananmen Square movement led to sharp savings increases by Chinese households. Wiemann and Lumsdaine (2019) find that increases in uncertainty about U.S. health-care policy lowers the consumption spending of married households, more so for those with worse health.

Handley and Limão (2015) develop evidence that lower uncertainty about trade policy stimulates investment in export capacity. Caldara et al. (2019) find evidence that higher trade policy uncertainty since 2017 has dampened U.S. business investment.[12] Gulen and Ion (2016) find negative effects of policy uncertainty on corporate investment using the U.S. EPU measure in Figure 6.1. Similarly, Baker, Bloom, and Davis (2016) find larger negative effects of EPU on investment rates and employment growth, and larger positive effects on stock price volatility, for firms with greater exposure to policy risks. Hassan et al. (2019)

use transcripts of earnings conference calls to construct time-varying measures of firm-level policy uncertainty. They also find that higher uncertainty discourages investment and employment. Using the EPU index for India, Anand and Tulin (2014) find negative effects of policy uncertainty on firm-level investment flows, with stronger effects on new projects.

A larger literature considers the effects of economic uncertainty in general, rather than policy uncertainty in particular. Important early analyses of how income uncertainty affects consumption behavior include Carroll (1997), Kimball (1990), and Zeldes (1989). Eberly (1994) finds that high uncertainty leads households to defer costly-to-reverse purchases of durable goods. Bloom (2009) finds that high uncertainty leads firms to cut or delay investment expenditures. These two studies and many others stress that heightened uncertainty provides an incentive to delay or forgo investments that are costly to reverse. Uncertainty can also depress investment by raising risk premiums, as stressed by Arellano, Bai, and Kehoe (2016); Christiano, Motto, and Rostagno (2014); Gilchrist, Sim, and Zakrajšek (2014); and Pastor and Veronesi (2013). Insofar as high uncertainty depresses investment and discourages the reallocation of capital and labor, it also slows the growth of productivity and output. See Bloom (2014) for a fuller discussion of how uncertainty affects economic activity.

Another branch of the literature investigates the dynamic relationship of policy uncertainty, or economic uncertainty more broadly, to macroeconomic performance. Examples include Arbatli et al. (2019); Baker, Bloom, and Davis (2016); Baker, Bloom, and Terry (2016); Colombo (2013); Ghirelli, Pérez, and Urtasun (2019); International Monetary Fund (2013); Jurado, Ludvigson, and Ng (2015); Leduc and Liu (2016); and Stock and Watson (2012). These studies typically find that higher (policy) uncertainty foreshadows a deterioration in macroeconomic performance. Romer (1990) marshals evidence that the 1929 stock market crash triggered a sharp rise in income uncertainty that led households to forgo purchases of consumer durables, accentuating the collapse of aggregate demand at the onset of the Great Depression. Evidence in Constantinescu, Mattoo, and Ruta (2017) suggests that high policy uncertainty depresses international trade in goods and services.

In summary, a variety of studies find evidence that high policy uncertainty undermines economic performance by leading firms to delay or forgo investments and hiring, by slowing productivity-enhancing factor reallocation, and by depressing consumption expenditures. This evidence points to a positive payoff in the form of stronger macroeconomic performance if policymakers can deliver greater predictability in the policy environment. A smaller literature finds that greater uncertainty causes households and firms to become less responsive on the margin to cuts in interest rates and taxes, in line with predictions of real options theory. See Aastveit, Natvik, and Sola (2013); Bertola, Guiso, and Pistaferri (2005); Bloom (2009); Bloom, Bond, and Van Reenen (2007); and Vavra (2014). These studies suggest that a stronger policy framework also increases the potency of countercyclical stabilization policies.

CONCLUSION

U.S. and global policy uncertainty have been highly elevated in recent years. According to Figure 6.1 and evidence in Baker et al. (2014), the past dozen years have seen the highest levels of U.S. economic policy uncertainty in the past 60 years. According to Figure 6.2, global EPU in 2017 and 2018 is running at even higher levels than during the global financial crisis. The huge rise in trade policy uncertainty since early 2018 is an extraordinary departure from recent history, as is the prominent role of trade policy in recent stock market volatility.

There is now a sizable body of empirical research that supports the proposition that high policy uncertainty harms macroeconomic performance. The evidence in this literature implies that greater predictability in the policy environment yields better macroeconomic performance. A smaller literature suggests that standard monetary and fiscal policy tools are also more effective in environments with greater policy predictability.

Notes

This chapter, prepared in connection with the Werner Sichel Lecture Series at Western Michigan University, draws on my research with Scott Baker, Nick Bloom, and others in Arbatli et al. (2019); Baker, Bloom, and Davis (2016); Baker, Bloom, Davis, and Kost (2019); Baker, Bloom, Davis, and Sammon (2019); Davis (2016); and Davis, Liu, and Sheng (2019). I gratefully acknowledge financial support from the U.S. National Science Foundation and the Booth School of Business at the University of Chicago.

1. See Baker, Bloom, and Davis (2016). Monthly updates are available at www.PolicyUncertainty.com.
2. See Baker, Bloom, and Davis (2016) on *Beige Books* and Hassan et al. (2019) on earnings calls.
3. See Cerda, Silva, and Valente (2016) on the EPU index for Chile; Baker, Bloom, Davis, and Wang (2013) for China; Gil and Silva (2018) for Colombia; Hardouvelis et al. (2018) for Greece; Zalla (2016) for Ireland; Arbatli et al. (2019) for Japan; Kroese, Kok, and Parlevliet (2015) for the Netherlands; Ghirelli, Pérez, and Urtasun (2019) for Spain; and Armelius, Hull, and Köhler (2017) for Sweden. EPU data for the other countries are updates to the indices developed in Baker, Bloom, and Davis (2016) and new indices that we developed using the same methods.
4. For example, I regress the EPU index for Australia from 1998 onward on contemporaneous EPU index values for all countries with no missing data. I then use predicted values from this regression to impute the missing Australian values for 1997.
5. Baker, Bloom, and Davis (2015) present and discuss a suite of newspaper-based indices of immigration-related fears and policy uncertainty for France, Germany, the United Kingdom, and the United States. Updates are available at http://www.policyuncertainty.com/immigration_fear.html.
6. I am unaware of authoritative, up-to-the-moment statistics on average U.S. tariff rates. Statistics cited in the text are a composite of estimates attributed to Deutsche Bank and UBS Group in Douglas (2019) and a chart attributed to Oxford Economics in Borodovsky (2019).
7. These figures are also estimates reported in Borodovsky (2019) and Douglas (2019). They do not incorporate President Trump's announcement on August 23, 2019, of additional tariff hikes on Chinese imports.
8. In August 2019, PredictIt assessed only a 30 percent probability that "both houses of the U.S. Congress shall ratify the United States-Mexico-Canada Agreement, by passing a bill to implement such agreement" by the end of 2019. See also Marcos (2019) and Werner, Lynch, and Rauhala (2019). However, the House ratified it that December, and the Senate followed suit on January 16 of this year.
9. The four dates and the corresponding value-weighted returns on the S&P 500 are March 22, 2018, −2.52 percent; March 26, 2018, 2.72 percent; December 4, 2018, −3.24 percent; and January 4, 2019, 3.43 percent.
10. To construct Figure 6.4, we first compute the ratio (count of EMV articles that

contain trade policy terms)/(count of all EMV articles) in each month from January 1985 to December 2018. The "count of all EMV articles" in the denominator is the number of articles in 11 leading U.S. newspapers that contain at least one term in each of the following three sets: (E)conomy: {economic, economy, financial}; Stock (M)arket: {stock market, equity, equities, Standard and Poors, Standard & Poors, Standard and Poor, Standard and Poor's, Standard & Poor's}; and (V)olatility: {uncertain, uncertainty, volatility, volatile, risk, risky}. The numerator is the count of the subset of EMV articles that also contain one or more terms in Trade Policy: {trade policy, tariff, import duty, import barrier, import restriction, trade quota, dumping, export tax, export duty, trade treaty, trade agreement, trade act, WTO, World Trade Organization, Doha round, Uruguay round, GATT, export restriction, investment restriction, NAFTA, North American Free Trade Agreement, Trans-Pacific Partnership, TransPacific Partnership, Federal Maritime Commission, International Trade Commission, Jones Act, trade adjustment assistance}.

11. The effects of policy uncertainty also depend on the environment. For example, Basu and Bundick (2017) and Nakata (2017) examine uncertainty shocks in New Keynesian models. Both papers conclude that higher uncertainty has a larger negative effect on output when the monetary authority's policy rate is closer to the zero bound. Caggiano, Castelnuovo, and Pellegrino (2017) find empirical support for this prediction.
12. Although they cannot cleanly disentangle uncertainty effects from (negative) anticipation effects, Altig et al. (2019) report survey evidence that trade policy developments in 2018 caused a small drop in U.S. business investment. Similarly, Bloom et al. (2019) find survey evidence that Brexit-related developments have caused a sizable drop in U.K. business investment over the past three years.

References

Aaberge, Rolf, Kai Liu, and Yu Zhu. 2017. "Political Uncertainty and Household Savings." *Journal of Comparative Economics* 45(1): 154–170.

Aastveit, Knut Are, Gisle James Natvik, and Sergio Sola. 2013. "Economic Uncertainty and the Influence of Monetary Policy." *Journal of International Money and Finance* 76(C): 50–67.

Admati, Anat, and Martin Hellwig. 2013. *The Bankers' New Clothes: What's Wrong with Banking and What to Do about It.* Princeton: Princeton University Press.

Altig, David, Nick Bloom, Steven J. Davis, Brent Meyer, and Nick Parker. 2019. "Tariff Worries and U.S. Business Investment, Take Two." *Macroblog* (blog), Federal Reserve Bank of Atlanta, February 25. https://macroblog.typepad.com/macroblog/2019/02/tariff-worries-and-us-business-investment-take-two.html (accessed January 13, 2020).

Anand, Rahul, and Volodymyr Tulin. 2014. "Disentangling India's Investment Slowdown." IMF Working Paper No. 14/47. Washington, DC: International Monetary Fund.

Arango, Tim, Ceylan Yeginsu, and Safak Timur. 2016. "Turks See Purge as Witch Hunt of 'Medieval' Darkness." *New York Times*, September 16. https://www.nytimes.com/2016/09/17/world/europe/turkey-erdogan-gulen-purge.html (accessed January 15, 2020).

Arbatli, Elif C., Steven J. Davis, Arata Ito, and Naoko Miake. 2019. "Policy Uncertainty in Japan." NBER Working Paper No. 23411. Cambridge, MA: National Bureau of Economic Research.

Arellano, Cristina, Yan Bai, and Patrick J. Kehoe. 2016. "Financial Frictions and Fluctuations in Volatility." *Journal of Political Economy* 127(5): 2049–2103.

Armelius, Hanna, Isaiah Hull, and Hanna Stenbacka Köhler. 2017. "The Timing of Uncertainty Shocks in a Small Open Economy." *Economics Letters* 155(C): 31–34.

Baker, Scott R., Nicholas Bloom, Brandice Canes-Wrone, Steven J. Davis, and Jonathan Rodden. 2014. "Why Has U.S. Policy Uncertainty Risen since 1960?" *American Economic Review* 104(5): 56–60.

Baker, Scott R., Nicholas Bloom, and Steven J. Davis. 2015. "Immigration Fears and Policy Uncertainty." *Vox,* a policy portal of the Center for Economic and Policy Research (CEPR), December 15. Washington, DC: Center for Economic and Policy Research. https://voxeu.org/article/immigration-fears-and-policy-uncertainty (accessed January 13, 2020).

———. 2016. "Measuring Economic Policy Uncertainty." *Quarterly Journal of Economics* 131(4): 1593–1636.

Baker, Scott R., Nicholas Bloom, Steven J. Davis, and Kyle J. Kost. 2019. "Policy News and Stock Market Volatility." NBER Working Paper No. 25720. Cambridge, MA: National Bureau of Economic Research.

Baker, Scott R., Nicholas Bloom, Steven J. Davis, and Marco Sammon. 2019. "What Drives Large Stock Market Movements?" Economic Policy Uncertainty website. www.policyuncertainty.com/research.html (accessed May 16, 2020).

Baker, Scott R., Nicholas Bloom, Steven J. Davis, and Xiaoxi Wang. 2013. "Economic Policy Uncertainty in China." Unpublished paper, University of Chicago, Chicago, IL.

Baker, Scott R., Nicholas Bloom, and Stephen Terry. 2016. "Does Uncertainty Reduce Growth? Using Disasters as Natural Experiments." Paper presented at the Melbourne Institute Macroeconomic Policy Meetings, held in Melbourne, Australia, October 5–6.

Ballhaus, Rebecca, and Jacob M. Schlesinger. 2018. "Trump Vows to Double

Metals Tariffs on Turkey as Dispute Escalates over Detained American." *Wall Street Journal*, August 11. https://www.wsj.com/articles/trump-vows-to-double-metals-tariffs-on-turkey-as-dispute-escalates-over-detained-american-1533906681 (accessed January 16, 2020).

Basu, Susanto, and Brent Bundick. 2017. "Uncertainty Shocks in a Model of Effective Demand." *Econometrica* 85(3): 937–958.

BBC News. 2016. "Schengen: Controversial EU Free Movement Deal Explained." *BBC News*, April 24. http://www.bbc.com/news/world-europe-13194723 (accessed January 15, 2020).

Bertola, Giuseppe, Luigi Guiso, and Luigi Pistaferri. 2005. "Uncertainty and Consumer Durables Adjustment." *Review of Economic Studies* 72(4): 973–1007.

Birnbaum, Michael, and Karen DeYoung. 2016. "The Aftermath of Turkey's Failed Coup Threatens Its Ties with Western Allies." *Washington Post*, September 21. https://www.washingtonpost.com/world/europe/the-aftermath-of-turkeys-failed-coup-threatens-its-ties-with-western-allies/2016/09/20/314bb754-6e0d-11e6-993f-73c693a89820_story.html (accessed January 15, 2020).

Bloom, Nicholas. 2009. "The Impact of Uncertainty Shocks." *Econometrica* 77(3): 623–685.

———. 2014. "Fluctuations in Uncertainty." *Journal of Economic Perspectives* 28(2): 153–176.

Bloom, Nicholas, Stephen Bond, and John Van Reenen. 2007. "Uncertainty and Investment Dynamics." *Review of Economic Studies* 74(2): 391–415.

Bloom, Nicholas, Philip Bunn, Scarlet Chen, Paul Mizen, Pawel Smietanka, and Gregory Thwaites. 2019. "The Impact of Brexit on U.K. Firms." SIEPR Working Paper No. 19-019. Stanford, CA: Stanford Institute for Economic Policy Research (SIEPR).

Borodovsky, Lev. 2019. "The Daily Shot: How High Will the Average U.S. Tariff Rate Get?" *Wall Street Journal*, June 7. https://blogs.wsj.com/dailyshot/2019/06/07/the-daily-shot-how-high-will-the-average-u-s-tariff-rate-get/ (accessed January 15, 2020).

Bravin, Jess, and Louise Radnofsky. 2012. "Court Backs Obama on Health Law." *Wall Street Journal*, June 29. https://www.wsj.com/articles/SB10001424052702304898704577480371370927862 (accessed January 15, 2020).

Brown, Chad P. 2019. "Trump's Fall 2019 China Tariff Plan: Five Things You Need to Know." News release, August 14. Washington, DC: Peterson Institute for International Economics. https://www.piie.com/blogs/trade-and-investment-policy-watch/trumps-fall-2019-china-tariff-plan-five-things-you-need-know (accessed January 14, 2020).

Brown, Chad P., and Melina Kolb. 2019. "Trump's Trade War Timeline: An

Up-to-Date Guide." News release, December 19. Washington, DC: Peterson Institute for International Economics. https://www.piie.com/sites/default/files/documents/trump-trade-war-timeline.pdf (accessed January 14, 2020).

Brown Chad P., and Eva Yiwen Zhang. 2019. "Trump's Latest Trade War Escalation Will Push Average Tariffs on China above 20 Percent." News release, August 6. Washington, DC: Peterson Institute for International Economics. https://www.piie.com/research/piie-charts/trumps-latest-trade-war-escalation-will-push-average-tariffs-china-above-20 (accessed January 14, 2020).

Byrd, Haley. 2018. "Trump Tweets That His 232 Tariffs Aren't *Really* about National Security." *Washington Examiner*, June 12. https://www.washingtonexaminer.com/weekly-standard/trump-tweets-that-his-232-tariffs-arent-really-about-national-security (accessed January 16, 2020).

Caggiano, Giovanni, Efrem Castelnuovo, and Giovanni Pellegrino. 2017. "Estimating the Real Effects of Uncertainty Shocks at the Zero Lower Bound." *European Economic Review* 100(C): 257–272.

Caldara, Dario, Matteo Iacoviello, Patrick Molligo, Andrea Prestipino, and Andrea Raffo. 2019. "The Economic Effects of Trade Policy Uncertainty." International Finance Discussion Paper No. 1256. Washington, DC: Board of Governors of the Federal Reserve System.

Carroll, Christopher D. 1997. "Buffer-Stock Saving and the Life Cycle/Permanent Income Hypothesis." *Quarterly Journal of Economics* 112(1): 1–55.

Cerda, Rodrigo, Álvaro Silva, and José Tomás Valente. 2016. "Economic Policy Uncertainty Indices for Chile." Working paper. Santiago: Instituto de Economía Facultad de Ciencia Económicas y Administrativas Pontificia Universidad Católica de Chile.

Christiano, Lawrence J., Roberto Motto, and Massimo Rostagno. 2014. "Risk Shocks." *American Economic Review* 104(1): 27–65.

Colombo, Valentina. 2013. "Economic Policy Uncertainty in the U.S.: Does It Matter for the Euro Area?" *Economics Letters* 121(1): 39–42.

Constantinescu, Cristina, Aaditya Mattoo, and Michele Ruta. 2017. *Trade Developments in 2016: Policy Uncertainty Weighs on World Trade*. Global Trade Watch. Washington, DC: World Bank.

Cooper, Michael. 2011. "Health Law Is Dealt Blow by a Court on Mandate." *New York Times*, August 12. https://www.nytimes.com/2011/08/13/us/13health.html (accessed January 15, 2020).

Coskun, Orhan. 2019. "Turkey's Central Bank Governor Was Sacked after Resisting 300 Point Rate Cut: Sources." Reuters, July 22. https://www.reuters.com/article/us-turkey-cenbank-governor/turkeys-central-bank-governor-was-sacked-after-resisting-300-point-rate-cut-sources-idUSKCN1UH142 (accessed January 15, 2020).

Davis, Steven J. 2016. “An Index of Global Economic Policy Uncertainty.” *Macroeconomic Review* 15(2): 91–97.

Davis, Steven J., Dingquian Liu, and Xuguang Simon Sheng. 2019. “Economic Policy Uncertainty in China since 1946: The View from Mainland Newspapers.” Working paper, PolicyUncertainty.com.

Douglas, Jason. 2019. “Costly Tariff Spat Masks Deeper Trade Problems.” *Wall Street Journal*, August 18. https://www.wsj.com/articles/costly-tariff-spat-masks-deeper-trade-problems-11566126002 (accessed January 15, 2020).

Dube, Ryan, and Jeffrey T. Lewis. 2019a. “Argentine Financial Markets Steady, but Political Instability Drives Uncertainty.” *Wall Street Journal*, August 13. https://www.wsj.com/articles/argentine-financial-markets-steady-but-political-instability-drives-uncertainty-11565727646 (accessed January 15, 2020).

———. 2019b. “Argentine Peso Dives after Populist Peronists Gain Edge in Vote.” *Wall Street Journal*, August 12. https://www.wsj.com/articles/argentine-peso-dives-after-president-defeated-11565621831 (accessed January 15, 2020).

Duffie, Darrell. 2019. “Prone to Fail: The Pre-Crisis Financial System.” *Journal of Economic Perspectives* 33(1): 81–106.

Durnev, Art. 2012. “The Real Effects of Political Uncertainty: Elections and Investment Sensitivity to Stock Prices.” Working paper. Montreal: McGill University.

Dustmann, Christian, Francesco Fasani, Tommaso Frattini, Luigi Minale, and Uta Schönberg. 2016. “On the Economics and Politics of Refugee Migration.” CEPR Discussion Paper No. DP11557. London: Centre for Economic Policy Research.

Eberly, Janice C. 1994. “Adjustment of Consumers’ Durables Stocks: Evidence from Automobile Purchases.” *Journal of Political Economy* 102(3): 403–436.

Economist. 2016. “Time for Temer: The New President Takes Over a Country in Crisis.” *Economist*, September 3.

European Parliamentary Research Service (EPRS). 2016. “Briefing—Sanctions over Ukraine: Impact on Russia.” Brussels: European Parliamentary Research Service. http://www.europarl.europa.eu/EPRS/EPRS-Briefing-579084-Sanctions-over-Ukraine-impact-Russia-FINAL.pdf (accessed January 14, 2020).

Funke, Manuel, Moritz Schularick, and Christoph Trebesch. 2016. “Going to Extremes: Politics after Financial Crises, 1870–2014.” *European Economic Review* 88(September): 227–260.

Gauthier-Villars, David, and Jon Sindreu. 2018. “Turkey Takes Action on

Strained Economy with Big Rate Rise." *Wall Street Journal*, September 13. https://www.wsj.com/articles/erdogan-moves-to-support-turkeys-embattled-lira-with-new-measures-1536826634 (accessed January 15, 2020).

Ghirelli, Corinna, Javier J. Pérez, and Alberto Urtasun. 2019. "A New Economic Policy Uncertainty Index for Spain." Banco de España Working Paper No. 1906. Madrid: Banco de España.

Giavazzi, Francesco, and Michael McMahon. 2012. "Policy Uncertainty and Household Savings." *Review of Economics and Statistics* 94(2): 517–531.

Gil, Mauricio, and Daniel Silva. 2018. "Economic Policy Uncertainty Indices for Colombia." Working paper, PolicyUncertainty.com. https://www.policyuncertainty.com/media/EPU_Colombia_Paper_GS.pdf (accessed January 14, 2020).

Gilchrist, Simon, Jae W. Sim, and Egon Zakrajšek. 2014. "Uncertainty, Financial Frictions, and Investment Dynamics." Finance and Economics Discussion (FEDS) Working Paper No. 2014-69. Washington, DC: Federal Reserve Board, Divisions of Research and Statistics and Monetary Affairs.

Gulen, Huseyin, and Mihai Ion. 2016. "Policy Uncertainty and Corporate Investment." *Review of Financial Studies* 29(3): 523–564.

Halla, Martin, Alexander Wagner, and Josef Zweimüller. 2015. "Immigration and Far-Right Voting: New Evidence." *Vox,* a policy portal of the Center for Economic and Policy Research (CEPR), November 29. Washington, DC: Center for Economic and Policy Research. https://voxeu.org/article/immigration-and-far-right-voting-new-evidence (accessed January 14, 2020).

Handley, Kyle, and Nuno Limão. 2015. "Trade and Investment under Policy Uncertainty: Theory and Firm Evidence." *American Economic Journal: Economic Policy* 7(4): 189–222.

Hardouvelis, Gikas A., Georgios Karalas, Dimitrios Karanastasis, and Panagiotis Samartzis. 2018. "Economic Policy Uncertainty, Political Uncertainty, and the Greek Economic Crisis." Working paper. Piraeus, Greece: University of Piraeus.

Hassan, Tarek A., Stephan Hollander, Laurence van Lent, and Ahmed Tahoun. 2019. "Firm-Level Political Risk: Measurement and Effects." *Quarterly Journal of Economics* 134(4): 2135–2202.

Huang, Yi, Chen Lin, Sibo Liu, and Heiwai Tang. 2018. "Trade Linkages and Firm Value: Evidence from the 2018 U.S.-China Trade War." IHEID Working Paper No. 11-2018. Geneva: Graduate Institute of International and Development Studies.

International Monetary Fund. 2013. "World Economic Outlook: Hopes, Realities, Risks." IMF Press, April 2013.

Julio, Brandon, and Youngsuk Yook. 2012. "Political Uncertainty and Corporate Investment Cycles." *Journal of Finance* 67(1): 45–83.

———. 2016. "Policy Uncertainty, Irreversibility, and Cross-Border Flows of Capital." *Journal of International Economics* 103(November): 13–26.

Jurado, Kyle, Sydney C. Ludvigson, and Serena Ng. 2015. "Measuring Uncertainty." *American Economic Review* 105(3): 1177–1216.

Kantchev, Georgi, and Paul Hannon. 2019. "Russia, Turkey Take Contrasting Approaches in Push for Looser Monetary Policy." *Wall Street Journal*, July 26. https://www.wsj.com/articles/russia-cuts-interest-rate-again-amid-global-push-for-looser-policy-11564139243 (accessed January 15, 2020).

Kelly, Bryan, Lubos Pastor, and Pietro Veronesi. 2016. "The Price of Political Uncertainty: Theory and Evidence from the Option Market." *Journal of Finance* 71(5): 2417–2480.

Kimball, Miles S. 1990. "Precautionary Saving in the Small and in the Large." *Econometrica* 58(1): 53–73.

Kroese, Lars, Suzanne Kok, and Jante Parlevliet. 2015. "Beleidsonzekerheid in Nederland." *Economisch Statistische Berichten* 100(4715): 464–467.

Kubota, Yoko. 2019. "U.S. Retreat on New Tariffs Gives Beijing Breathing Room on Trade Talks." *Wall Street Journal*, August 14.

Leduc, Sylvain, and Zheng Liu. 2016. "Uncertainty Shocks Are Aggregate Demand Shocks." *Journal of Monetary Economics* 82(C): 20–35.

Li, Cheng. 2011. "The 2012 Chinese Leadership Transition." Interview for "On the Record" series. Washington, DC: Brookings Institution. https://www.brookings.edu/on-the-record/the-2012-chinese-leadership-transition (accessed January 15, 2020).

Mander, Benedict. 2019. "Argentina Elections: Why Investors Believe Macri's Time Is Up." *Financial Times*, August 13.

Marcos, Cristina. 2019. "Hopes Dim for Passage of Trump Trade Deal." *The Hill*, July 5. https://thehill.com/homenews/house/451645-hopes-dim-for-passage-of-trump-trade-deal (accessed January 15, 2020).

Mauldin, William, Alex Leary, and Chao Deng. 2019. Trump Lashes Out at China, Sending Markets Falling." *Wall Street Journal*, August 23. https://www.wsj.com/articles/trump-orders-u-s-businesses-to-find-alternative-to-china-11566585967 (accessed January 16, 2020).

Mian, Atif, and Amir Sufi. 2014. *House of Debt: How They (and You) Caused the Great Recession, and How We Can Prevent It from Happening Again.* Chicago: University of Chicago Press.

Mian, Atif, Amir Sufi, and Francesco Trebbi. 2014. "Resolving Debt Overhang: Political Constraints in the Aftermath of Financial Crises." *American Economic Journal: Macroeconomics* 6(2): 1–28.

Nakata, Taisuke. 2017. "Uncertainty at the Zero Lower Bound." *American Economic Journal: Macroeconomics* 9(3): 186–221.

Pastor, Lubos, and Pietro Veronesi. 2013. "Political Uncertainty and Risk Premia." *Journal of Financial Economics* 110(3): 520–545.

Press, Jordan. 2018. "Lighthizer Calls Canada a National Security Threat in Defence of Steel Tariffs." *Globe and Mail*, July 26. https://www.theglobeandmail.com/business/article-lighthizer-calls-canada-a-national-security-threat-in-defence-of-steel/ (accessed January 15, 2020).

Rachman, Gideon. 2019. "What Happens When the World Cannot Rely on the U.S.?" *Financial Times*, August 9. https://www.ft.com/content/9e915a24-b9c5-11e9-96bd-8e884d3ea203 (accessed January 15, 2020).

Reuters. 2016. "Moody's Cuts Turkey's Credit Rating to 'Junk.'" Reuters, September 24. https://fortune.com/2016/09/24/moodys-turkey-credit-rating/ (accessed January 15, 2020).

Romer, Christina D. 1990. "The Great Crash and the Onset of the Great Depression." *Quarterly Journal of Economics* 105(3): 597–624.

Schonhardt, Sara, and Saurabh Chaturvedi. 2016. "South China Sea Ruling Increases Uncertainty for Shipping, Trade." *Wall Street Journal*, July 14. https://www.wsj.com/articles/south-china-sea-ruling-increases-uncertainty-for-shipping-trade-1468487095 (accessed January 15, 2020).

Schott, Jeffrey J., and Euijin Jung. 2018. "KORUS Amendments: Minor Adjustments Fixed What Trump Called 'Horrible Trade Deal.'" Policy Brief No. 18-22. Washington, DC: Peterson Institute for International Economics. https://www.piie.com/system/files/documents/pb18-22.pdf (accessed January 15, 2020).

Stock, James H., and Mark W. Watson. 2012. "Disentangling the Channels of the 2007–09 Recession." *Brookings Papers on Economic Activity* 2012(Spring): 81–156.

Trump, Donald J. 2018. "While Japan and South Korea would like us to go back into TPP, I don't like the deal for the United States. Too many contingencies and no way to get out if it doesn't work. Bilateral deals are far more efficient, profitable and better for OUR workers. Look how bad WTO is to U.S." Twitter @realDonaldTrump, April 17. https://twitter.com/realdonaldtrump/status/986436520444866560 (accessed May 16, 2020).

Ungku, Fathin, and Charlotte Greenfield. 2018. "Trump Says U.S. Could Rejoin TPP If Deal Improved. How Hard Would It Be?" Reuters, April 16.

Vavra, Joseph. 2014. "Inflation Dynamics and Time-Varying Volatility: New Evidence and an Ss Interpretation." *Quarterly Journal of Economics* 129(1): 215–258.

Wall Street Journal Editorial Board. 2018. "Trump's Political Tariff Bureaucracy." *Wall Street Journal*, August 6. https://www.wsj.com/articles/trumps-political-tariff-bureaucracy-1533597856 (accessed January 16, 2020).

Werner, Erica, David J. Lynch, and Emily Rauhala. 2019. "Trump's North American Trade Deal at Risk of Stalling in Congress." *Washington Post*, March 29. https://www.washingtonpost.com/business/economy/north-american-trade-deal-faces-early-headwinds-on-capitol-hill/2019/03/29/ef5b0d0c-5178-11e9-88a1-ed346f0ec94f_story.html (accessed January 15, 2020).

Wiemann, Thomas T., and Robin L. Lumsdaine. 2019. "The Effect of Health Care Policy Uncertainty on Households' Consumption and Portfolio Choice." Working paper. Washington, DC: American University.

Yeginsu, Ceylan. 2016. "Turkey Uses Post-Coup Emergency Decree to Purge Mayors and Teachers." *New York Times*, September 9. https://www.nytimes.com/2016/09/10/world/middleeast/turkey-uses-post-coup-emergency-decree-to-purge-mayors-and-teachers.html (accessed January 15, 2020).

Zalla, Ryan. 2016. "Economic Policy Uncertainty in Ireland." Working paper. Villanova, PA: Villanova University.

Zeldes, Stephen P. 1989. "Optimal Consumption with Stochastic Income: Deviations from Certainty Equivalence." *Quarterly Journal of Economics* 104(2): 275–298.

7
Measuring Economies from Space

Adam Storeygard
Tufts University

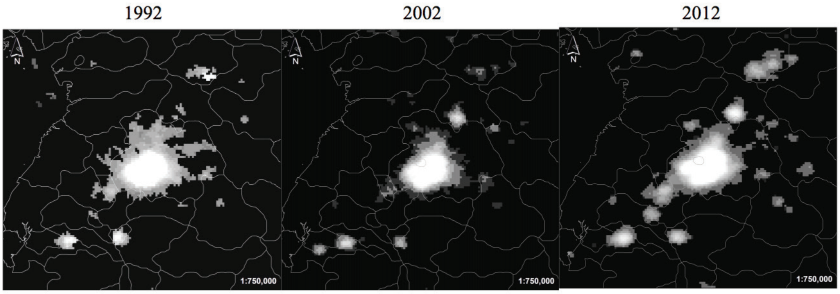
Why do we use satellite data for economic research and policymaking? Satellite data have several features that help us answer the kinds of causal questions about economic phenomena, and the effects of policies, that we care about as economists and policymakers. In particular, there are six advantages that I see as key. In this chapter, I address each one in turn, highlighting one or two pieces of economics research that use each and what we can learn from them. This chapter deals with data and methods, but throughout I'll show how they've been used to generate some very concrete lessons that are relevant to policy.[1]

SIX ADVANTAGES OF SATELLITE DATA

The first advantage of satellite data is that they exist where other data do not. Collecting data via household surveys and censuses is expensive and can be logistically difficult, especially in poor countries. The same is true of many kinds of administrative data that rich countries regularly collect.

To take an extreme example, Lee (2018) uses satellite data on lights at night from North Korea, a country that essentially does not publish credible economic data, to study the effect of sanctions there. Relying on the fact that changes in lighting are correlated with changes in economic activity, as demonstrated by Henderson, Storeygard, and Weil (2012), Lee uses the lights data to show that sanctions pushed economic activity toward Pyongyang (Figure 7.1) and to cities where trade with China was concentrated. China was not part of the sanctions regime. He argues that instead of primarily punishing ruling elites, as intended,

Figure 7.1 Lights Near Pyongyang, North Korea, in 1992, 2002, and 2012



SOURCE: Lee (2018).

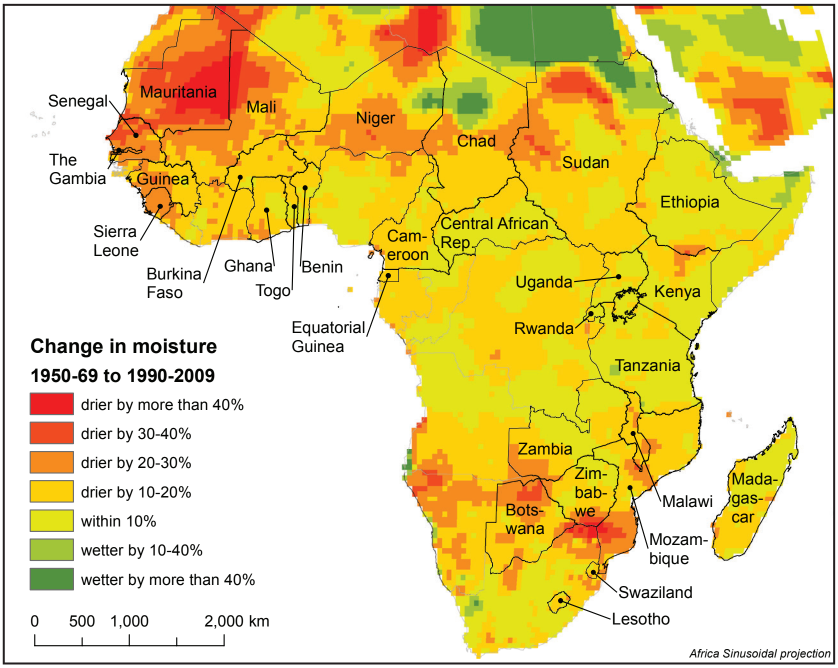
sanctions thus likely had their largest impacts on the "already marginalized hinterlands" (Lee 2018, p. 34).

A lack of data, of course, affects many places other than North Korea, in less extreme ways. In my work focused on how African cities grow, I have used the same night lights data as a measure of city-level economic activity, because other sources are rarely available, and almost never for every year. The paper by Henderson, Storeygard, and Deichmann (2017) considers how a drying climate throughout much of sub-Saharan Africa over a 50-year period (Figure 7.2) affected cities. The authors demonstrate that a drying climate appears to have pushed economic activity into some African cities but not others. Consistent with a simple theory, it is the cities most likely to have a preexisting export manufacturing base that attract new activity in times of drought, while cities that are more local in orientation are not affected. In the theory, the manufacturing-oriented cities are less affected by the drought because it affects neither their production technology nor the demand for their products. The more locally oriented cities, however, face a drop in demand from their customers: local farmers, whose production and therefore income fall.

These lights data have been used to address many other questions in data-poor environments, including how transport costs affect African cities (Storeygard 2016) and the effects of refugee camps on local economies in Kenya (Alix-Garcia et al. 2018).

A second advantage is that satellite data are often collected at extremely high spatial resolution, sometimes now less than one square

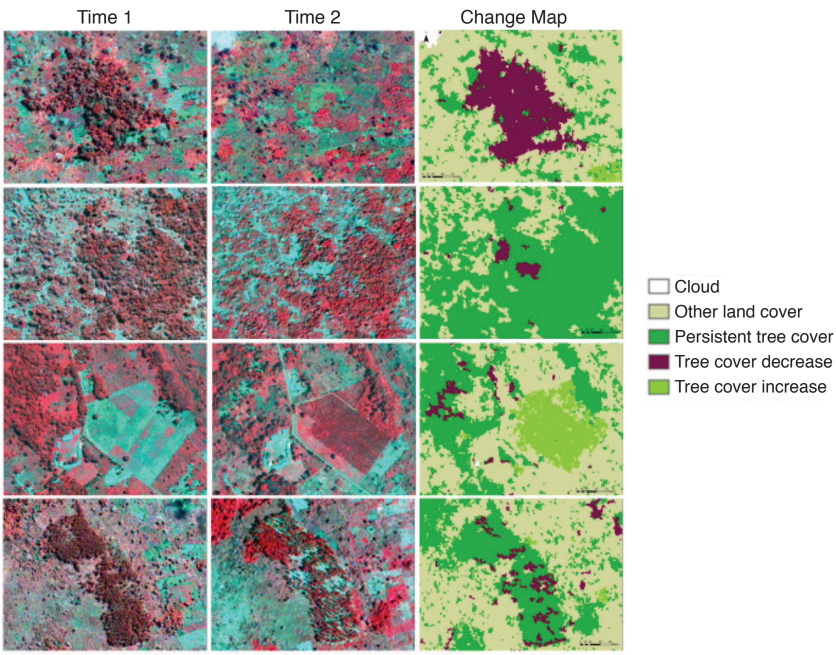
Figure 7.2 Moisture Change in Africa, from the Period of 1950–1969 to the Period of 1990–2009



SOURCE: Henderson, Storeygard, and Deichmann (2017).

meter. Jayachandran et al. (2017) have used this feature to evaluate a program of payments for ecosystem services in rural Uganda. The program paid private forest owners not to cut down trees for up to two years. The satellite data are fine enough that one can detect individual trees and assign them to the land of an individual household. Figure 7.3 shows what happened in four particular sample locations, each corresponding to a row in the figure. The first two columns of images show trees at the beginning and end of the two-year period, and the last column shows changes in tree cover. The authors link these data with a household survey about the program, using the location of each household's land. This link is of course critical—very little social science can be done with remote data alone. The authors found that people who were randomly assigned to the payments group cut down substantially

Figure 7.3 Changes in Forest Cover on Four Plots in Uganda



SOURCE: Jayachandran et al. (2017). Reprinted with permission from AAAS.

fewer trees than a control group that did not receive payments, and that the payments group did not displace tree harvesting onto other nearby lands. The difference was large enough that even if these effects were completely undone in four years, this delay in cutting down trees would be a cost-effective means of decreasing carbon emissions.

A third advantage is that repeat measurements with satellites are extremely cheap. Once analysts develop the methods for measuring something once, the satellite keeps orbiting and collecting data, so they, and often others, can apply the same algorithm to next month's or next year's data.

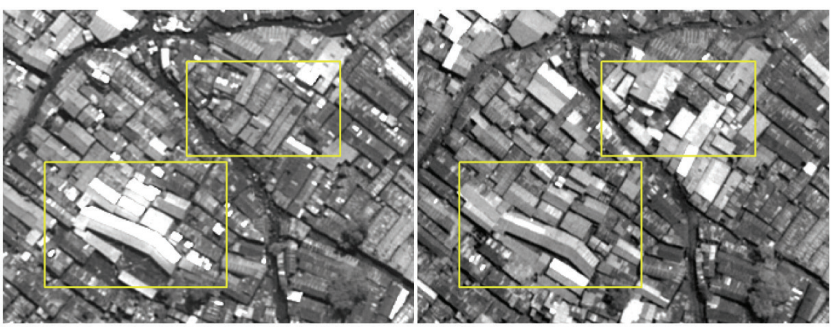
An excellent example of this is work done on ethnic patronage in the Kibera neighborhood of Nairobi, Kenya, by Marx, Stoker, and Suri (2019). These three authors are interested in the rents people pay relative to the quality of their house, but housing quality can vary over time.

To get a proxy for that, they measured the reflectance of the metal roofs (i.e., the amount of light reflected off the roofs). Roofs get less reflective as they get rustier (Figure 7.4, left and right lower yellow boxes), but when they get replaced, their shiny surfaces reflect more light (left and right upper yellow boxes). The authors were able to extract four measures of this reflectance for the whole neighborhood over a short period of time. Again, they could link the images of each house with a survey respondent.

The authors were able to document and quantify the fact that renters pay less rent, and get better housing quality, when they are of the same ethnicity as the local political boss. Conversely, they pay more rent and get worse housing quality when their landlord is of the same ethnicity as the local political boss.

In addition to reporting data for many points in time, most satellites report data for nearly the whole world on a regular basis. This was particularly useful for a study on the effect of subway systems on air pollution by Gendron-Carrier et al. (2018). The authors wanted to see whether air pollution fell in the weeks and months after subways

Figure 7.4 Changes in Roof Reflectivity from Old and New Roofs in Kibera, Nairobi, Kenya, July 2009 to August 2012



NOTE: "Both pictures are taken over the same area of the slum with the same resolution (0.5 meters panchromatic). The picture in the left panel was taken in July 2009 and that in the right panel in August 2012. The yellow rectangles highlight clusters of roofs that markedly evolved over the period. Roofs highlighted in the bottom rectangle degraded, while roofs within the top rectangle were upgraded in the same time frame" (Marx, Stoker, and Suri 2019, online appendix).
SOURCE: Marx, Stoker, and Suri (2019).

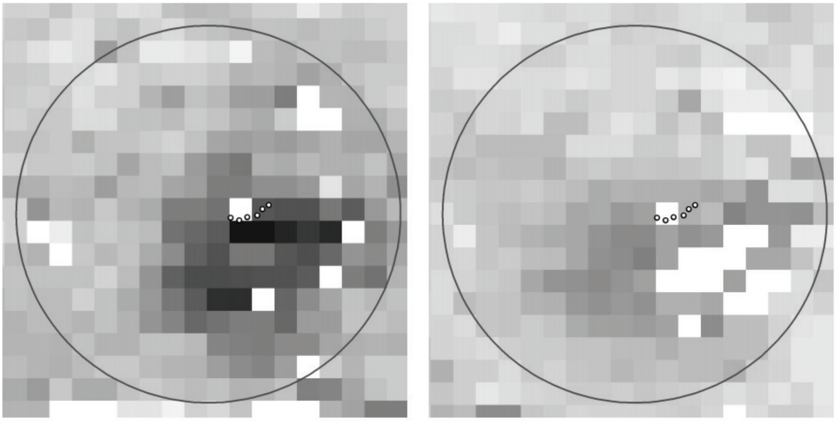
opened, as was the case for Bangalore in 2014 (Figure 7.5). But not many cities have seen subways open in the past 20 years.

Conveniently, satellite data on particulate matter are available for the whole world, so Gendron-Carrier et al. (2018) gathered administrative data on the opening of each subway stop in the 42 cities with new systems, which they were able to link to the pollution data. Without so many cities, they would not have had the statistical power for reliable inference.

Their results are quite striking. They find that subways substantially reduce particulates, and that the effect does not tend to decrease for as long as they can see in their sample, which is about eight years after the opening of the subway. This is somewhat surprising, as much other work, from Downs (1962) to Duranton and Turner (2011), predicts that new drivers will exploit any reduction in automobile traffic.

A feature related to the worldwide coverage is that satellites are measuring the same quantity everywhere. They don't turn off or change

Figure 7.5 Air Pollution before and after Inauguration of a New Subway Line in Bangalore



(a) Bangalore, June 2014 (b) Bangalore, December 2014

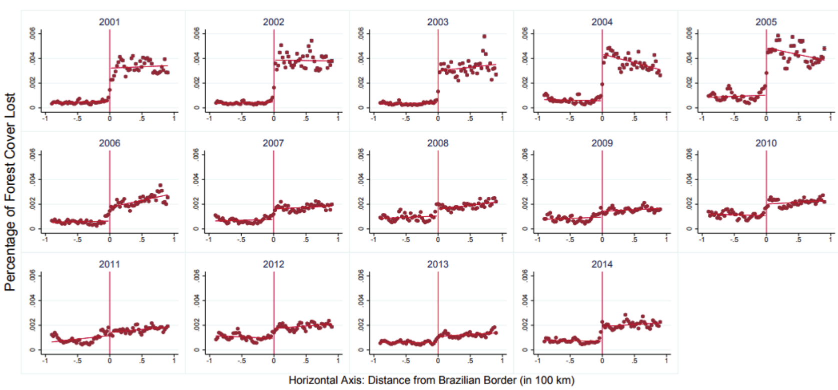
NOTE: Stations shown as small circles. Darker grid cells represent higher levels of particulates, and the large circle has a radius of 10 kilometers, centered on the central business district.

SOURCE: Gendron-Carrier et al. (2018).

methods when they cross a national border. This is different from even highly standardized surveys such as the Demographic and Health Surveys (DHS) carried out in many countries, because at a minimum, they must translate the same questions into different languages. Burgess, Costa, and Olken (2018) have exploited this idea of consistency across borders to consider the effect of a policy that Brazil introduced in 2006 to reduce deforestation. There are many reasons why deforestation rates change from year to year—including, for example, market prices of agricultural products farmers plan to grow on cleared land—so it's hard to distinguish the policy from other phenomena.

To address this, Burgess, Costa, and Olken (2018) use a spatial regression discontinuity design, as illustrated in Figure 7.6. Each point in each graph represents the share of forest in a 2.5-kilometer-wide swath of land parallel to Brazil's border cut down in a given year. This is equivalent to measuring deforestation along a transect crossing the border and then aggregating appropriately across all such transects. In the early years in the top row (2001–2005), when one moves from left to right into Brazil, rates of deforestation increase dramatically pre-

Figure 7.6 Deforestation Rates by Year and Distance to the Brazilian Border



NOTE: Average forest cover lost as a function of distance to the Brazilian border by year. Each point represents a 2.5-kilometer-wide band, indexed by distance into Brazil from the border, where negative values are in a neighboring country.
SOURCE: Burgess, Costa, and Olken (2018).

cisely at the border, represented by the vertical line in the middle. However, starting in 2006, that differential falls considerably, and by 2009 it is barely detectable. This is striking evidence that something important changed in Brazil relative to its neighbors in 2006, and the authors posit that this policy is the most likely candidate. Note that with coarser data, say at the district level, it would be more difficult to determine whether the jump happened precisely at the border.

The last advantage of satellites that I will highlight is their independence from typical reporting mechanisms. This is especially important when local officials might have incentives to underreport environmental damage, for example, but it has broader implications.

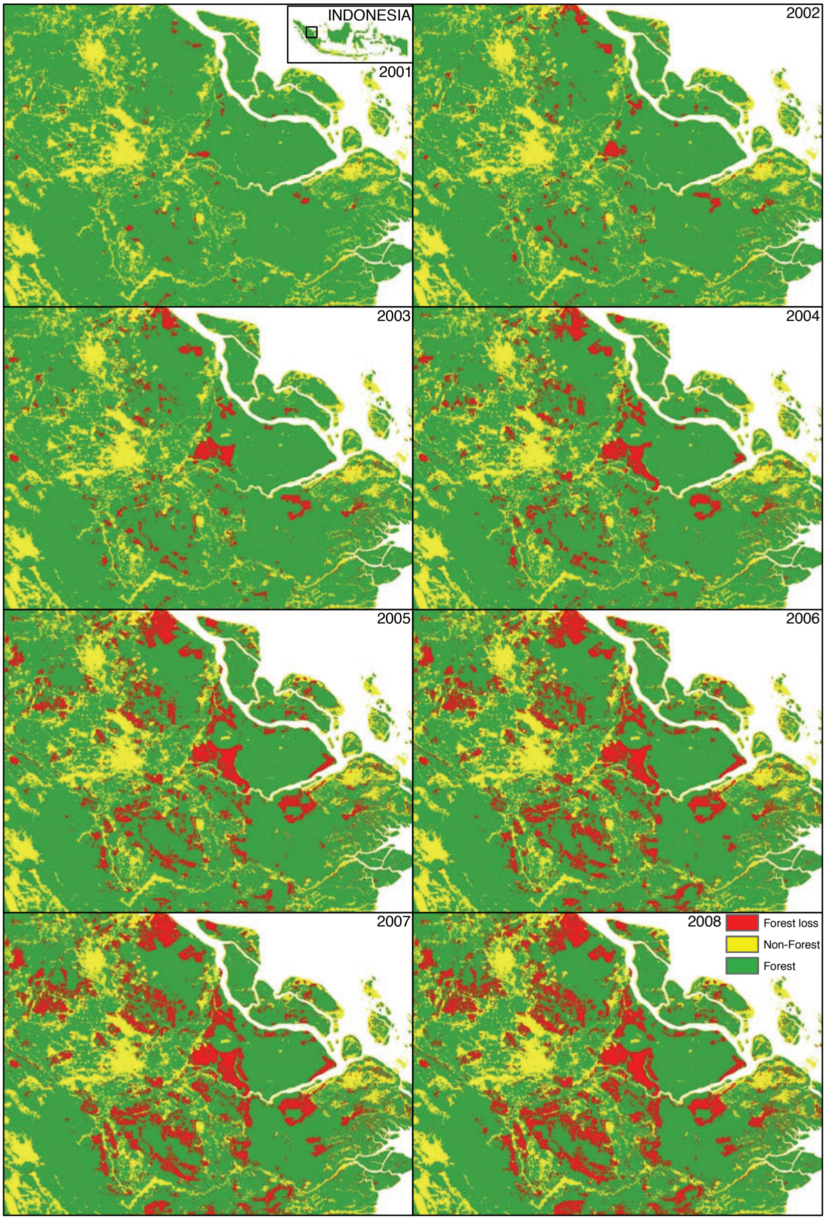
In an earlier paper, Burgess et al. (2012) look at how deforestation changed during rapid redistricting in Indonesia in the 2000s (Figure 7.7). By exploiting quirks in the timing of these changes, they show that redistricting led to more rapid deforestation. Their results are consistent with a model of Cournot competition in which the redistricting increased competition between districts for the revenue from legal and illegal logging.[2] A similar strategy has been used to investigate pollution in China, where the accuracy of official reporting has been called into question, with mixed results (Chen et al. 2012, 2013; Bombardini and Li 2019).

Even in the absence of outright falsification, the availability of a measure independent of traditional data sources is useful to reduce measurement error. Henderson, Storeygard, and Weil (2012) show this in combining data on lights growth with traditional GDP growth data. In essence, both lights and traditional GDP are subject to measurement error, but since the sources of their measurement error are very different, they are likely to be uncorrelated, and thus the two measures can be combined to form better estimates.[3] Pinkovskiy and Sala-i-Martin (2016) invoke a similar method in using lights to determine whether national accounts or household survey data provide more reliable estimates of national incomes.

CONCLUDING THOUGHTS

To highlight these six advantages of spatial data, I have described examples related to deforestation, pollution, urban growth, transporta-

Figure 7.7 Forest Cover in Riau Province, Indonesia, 2001–2008



SOURCE: Burgess et al. (2012), by permission of Oxford University Press.

tion, and political economy. There are many more in economics, on topics as varied as tourism (Faber and Gaubert 2019) and economic history (Pascali 2017), not to mention a much longer tradition in other fields, especially environmental science.

Another area that holds particular promise for the future is agriculture, especially in the developing world. It is not easy to learn a lot about crop choice or yield from a single satellite image. But once multiple images per growing season are available—or, even better, images every day—it starts becoming possible to learn an enormous amount about the agricultural economy. And with higher frequency high-resolution images, I expect that we will be able to learn about the choices of individual farmers at increasingly low cost.[4]

While I believe that this technology holds great promise, I do not want to give the impression that it can replace traditional data sources, or that it is without problems. The view from above is a powerful one, but it is not a complete one, and traditional administrative or survey data are critical to have in nearly all of the examples described above. Any given satellite image is a snapshot at one instant in time, not a summary of a day or a month that one could get, for example, from a pollution monitoring station. The most recent night-lights sensors,[5] for another example, provide data from two o'clock in the morning—as opposed to the early evening, as earlier satellites did—and so researchers will have to study whether that changes their relationship with economic activity. Satellites do not last forever, so repeat measurements over long periods require launching new satellites and adapting measurement techniques to them. It is still difficult to delineate objects like building footprints from a satellite image. Computers are getting better at that, but it's not yet routine, so it often requires lots of human labor. And while it is true that satellites generally operate the same way regardless of their location, context still matters. For example, an algorithm that is good at distinguishing a city from a surrounding forest does not always work as well in distinguishing a city from a surrounding desert. So, as in everything, it is important to know one's data well before attempting to interpret it.

To briefly summarize, satellites provide data for data-poor contexts, often at high resolution, with frequent repeat measurement, for the whole world, consistently across borders, in a way that is difficult to falsify because they are generally independent from traditional data providers. They are not magic, but as the price of data and processing power goes down and algorithms for analyzing them get better, they hold enormous potential for learning about economics and policy.

Notes

This chapter is based on the opening keynote address of the World Bank Land and Poverty Conference 2019, drawing on material from the Sichel Lecture delivered at Western Michigan University on October 10, 2018, and on Donaldson and Storeygard (2016). As Donaldson and Storeygard make clear, the economics literature using satellite data relies heavily on a much larger and older (but still rapidly developing) technical literature on the engineering and science of remote sensing.

1. For more details on satellite data and their use in economics, see Donaldson and Storeygard (2016), especially the references therein.
2. Note that redistricting could cause more mundane difficulties in reporting as well, as new district governments come into being, even in the absence of illegal motives.
3. This is slightly complicated by the fact that the units of the lights-based estimate are unknown. By analogy, if one weighs oneself with two different scales, both using kilograms as their unit of account, a simple mean of the two measurements is the optimal combination (unless one knows something about their relative precision). However, if one of the scales has an unknown unit of account, then its relationship to kilograms must be measured using the same data.
4. See Lobell (2013) and, for a recent developing-world example, Burke and Lobell (2017).
5. Visible Infrared Imaging Radiometer Suite (VIIRS).

References

Alix-Garcia, Jennifer, Sarah Walker, Anne Bartlett, Harun Onder, and Apurva Sanghi. 2018. “Do Refugee Camps Help or Hurt Hosts? The Case of Kakuma, Kenya.” *Journal of Development Economics* 130(C): 66–83. https://doi.org/10.1016/j.jdeveco.2017.09.005.

Bombardini, Matilde, and Bingjing Li. 2019. “Trade, Pollution, and Mortality in China.” Working paper. Vancouver, CA: University of British Columbia.

Burgess, Robin, Francisco Costa, and Ben Olken. 2018. “Wilderness Conservation and the Reach of the State: Evidence from National Borders in the Amazon.” NBER Working Paper No. 24861. Cambridge, MA: National Bureau of Economic Research.

Burgess, Robin, Matthew Hansen, Benjamin A. Olken, Peter Potapov, and Stefanie Sieber. 2012. "The Political Economy of Deforestation in the Tropics." *Quarterly Journal of Economics* 127(4): 1707–1754. https://doi.org/10.1093/qje/qjs034.

Burke, Marshall, and David B. Lobell. 2017. "Satellite-Based Assessment of African Yield Variation and Its Determinants in Smallholder African Systems." *Proceedings of the National Academy of Sciences* 114(9): 2189–2194. 10.1073/pnas.1616919114 (accessed March 13, 2020).

Chen, Yuyu, Ginger Zhe Jin, Naresh Kumar, and Guang Shi. 2012. "Gaming in Air Pollution Data? Lessons from China." *B.E. Journal of Economic Analysis and Policy* 13(3): 1–43.

———. 2013. "The Promise of Beijing: Evaluating the Impact of the 2008 Olympic Games on Air Quality." *Journal of Environmental Economics and Management* 66(2013): 424–443.

Donaldson, Dave, and Adam Storeygard. 2016. "The View from Above: Applications of Satellite Data in Economics." *Journal of Economic Perspectives* 30(4): 171–198. http://dx.doi.org/10.1257/jep.30.4.171.

Downs, Anthony. 1962. "The Law of Peak-Hour Expressway Congestion." *Traffic Quarterly* 16(3): 393–409.

Duranton, Gilles, and Matthew A. Turner. 2011. "The Fundamental Law of Road Congestion: Evidence from U.S. Cities." *American Economic Review* 101(6): 2616–2652.

Faber, Ben, and Cecile Gaubert. 2019. "Tourism and Economic Development: Evidence from Mexico's Coastline." *American Economic Review* 109(6): 2245–2293.

Gendron-Carrier, Nicolas, Marco Gonzalez-Navarro, Stefano Polloni, and Matthew Turner. 2018. "Subways and Urban Air Pollution." NBER Working Paper No. 24183. Cambridge, MA: National Bureau of Economic Research.

Henderson, Vernon, Adam Storeygard, and Uwe Deichmann. 2017. "Has Climate Change Driven Urbanization in Africa?" *Journal of Development Economics* 124(January): 60–82. https://doi.org/10.1016/j.jdeveco.2016.09.001.

Henderson, J. Vernon, Adam Storeygard, and David N. Weil. 2012. "Measuring Economic Growth from Outer Space." *American Economic Review* 102(2): 994–1028. http://dx.doi.org/10.1257/aer.102.2.994.

Jayachandran, Seema, Joost de Laat, Eric R. Lambin, Charlotte Y. Stanton, Robin Audy, and Nancy E. Thomas. 2017. "Cash for Carbon: A Randomized Trial of Payments for Ecosystem Services to Reduce Deforestation." *Science* 357(6348): 267–273. http://dx.doi.org/10.1126/science.aan0568 (accessed March 13, 2020).

Lee, Yong Suk. 2018. “International Isolation and Regional Inequality: Evidence from Sanctions on North Korea.” *Journal of Urban Economics* 103(C): 34–51. https://doi.org/10.1016/j.jue.2017.11.002.

Lobell, David B. 2013. “The Use of Satellite Data for Crop Yield Gap Analysis.” *Field Crops Research* 143(1): 56–64.

Marx, Benjamin, Thomas M. Stoker, and Tavneet Suri. 2019. “There Is No Free House: Ethnic Patronage in a Kenyan Slum.” *American Economic Journal: Applied Economics* 11(4): 36–70.

Pascali, Luigi. 2017. “The Wind of Change: Maritime Technology, Trade, and Economic Development.” *American Economic Review* 107(9): 2821–2854.

Pinkovskiy, Maxim, and Xavier Sala-i-Martin. 2016. “Lights, Camera,… Income! Illuminating the National Accounts–Household Surveys Debate.” *Quarterly Journal of Economics* 131(2): 579–631.

Storeygard, Adam. 2016. “Farther on Down the Road: Transport Costs, Trade, and Urban Growth.” *Review of Economic Studies* 83(3): 1263–1295.

Authors

Patrick Adams is a doctoral student in financial economics at the MIT Sloan School of Management.

Alessandro Barbarino is an economist in the Current Macroeconomic Conditions Section of the Federal Reserve Board.

William A. Barnett is the Oswald Distinguished Professor of Macroeconomics at the University of Kansas and director at the Center for Financial Stability.

Scott A. Brave is a senior policy economist in the Economic Research Department at the Federal Reserve Bank of Chicago.

Steven J. Davis is the William H. Abbott Distinguished Service Professor of International Business and Economics at the University of Chicago Booth School of Business.

Domenico Giannone is a senior principal economist at Amazon.com.

Kun He is a graduate assistant in the Department of Economics at the University of Kansas.

C. James Hueng is a professor in the Department of Economics at Western Michigan University.

Eric Qian is a senior research analyst in the Research and Statistics Division of the Federal Reserve Bank of New York.

Argia Sbordone is a vice president and function head in the Macroeconomic and Monetary Studies Function of the Federal Reserve Bank of New York.

Chiara Scotti is a special adviser to the Federal Reserve Board in the General Program Support Section.

Adam Storeygard is an associate professor of economics at Tufts University.

Mihir Trivedi is an advanced analytics associate at Komodo Health in Washington, D.C.

About the Institute

The W.E. Upjohn Institute for Employment Research is a nonprofit research organization devoted to finding and promoting solutions to employment-related problems at the national, state, and local levels. It is an activity of the W.E. Upjohn Unemployment Trustee Corporation, which was established in 1932 to administer a fund set aside by Dr. W.E. Upjohn, founder of The Upjohn Company, to seek ways to counteract the loss of employment income during economic downturns.

The Institute is funded largely by income from the W.E. Upjohn Unemployment Trust, supplemented by outside grants, contracts, and sales of publications. Activities of the Institute comprise the following elements: 1) a research program conducted by a resident staff of professional social scientists; 2) the Early Career Research Award program, which provides funding for emerging scholars to complete policy-relevant research on labor-market issues; 3) a publications program and online research repository, which provide vehicles for disseminating the research of staff and outside scholars; 4) a regional team that conducts analyses for local economic and workforce development; and 5) the Employment Management Services Division, which administers publicly funded employment and training services as Michigan Works! Southwest in the Institute's local four-county area.

The broad objectives of the Institute's activities are to 1) promote scholarship and evidence-based practices on issues of employment and unemployment policy, and 2) make knowledge and scholarship relevant and useful to policymakers in their pursuit of solutions related to employment and unemployment.

Current areas of concentration for these programs include the causes, consequences, and measures to alleviate unemployment; social insurance and income maintenance programs; compensation and benefits; workforce skills; nonstandard work arrangements; and place-based policy initiatives for strengthening regional economic development and local labor markets.