Yale University
EliScholar – A Digital Platform for Scholarly Publishing at Yale

Yale Graduate School of Arts and Sciences Dissertations

Spring 2021

Essays in Environmental and Energy Economics

Wade Ivan Davis
Yale University Graduate School of Arts and Sciences, wade.davis@yale.edu

Follow this and additional works at: https://elischolar.library.yale.edu/gsas_dissertations

Recommended Citation
Davis, Wade Ivan, "Essays in Environmental and Energy Economics" (2021). *Yale Graduate School of Arts and Sciences Dissertations*. 35.
https://elischolar.library.yale.edu/gsas_dissertations/35

This Dissertation is brought to you for free and open access by EliScholar – A Digital Platform for Scholarly Publishing at Yale. It has been accepted for inclusion in Yale Graduate School of Arts and Sciences Dissertations by an authorized administrator of EliScholar – A Digital Platform for Scholarly Publishing at Yale. For more information, please contact elischolar@yale.edu.

Abstract

Essays in Environmental and Energy Economics

Wade Davis

2021

This dissertation contributes developments in modeling and policy analysis in environmental and energy economics. All three chapters are useful to ongoing debates in climate change policy and the regulation of greenhouse gas emissions. My first chapter develops a model of consumer decision-making in an analysis of the electricity retail choice market in Texas. This project explores (1) the limitations of consumer decision-making in a setting with large choice sets and (2) the relationship between competition and product variety after deregulation. I find strong evidence of inattention and search costs as explanations for consumers' widespread failure to choose cost-minimizing contracts. These findings suggest that policymakers could improve welfare with interventions that reduce search costs and inattention, such as removing the legal obstacles to concierge services or introducing a web-based tool to find consumers' cost-minimizing contract based on their consumption history. My findings also suggest that these interventions could lead to higher adoption of time-varying rates, which could lead to more efficient allocation of grid resources and lower emissions levels. My other main finding is that consumers are constrained in the monopoly setting from expressing their heterogeneous preferences for contract variety. This insight may guide regulators in monopoly settings to consider increasing variety. Of course, the possible benefits of increased variety face a trade-off with the costs of search and inattention.

My second chapter is co-authored with Robert Mendelsohn and Paula Pereda. We propose a model to estimate the economic damages from weather shocks and climate change. We contrast our model with the models used in previous literature, and we show that our model estimates substantially different effects than this earlier work, a finding the emphasizes the importance of model selection

and careful consideration of the implicit assumptions. We demonstrate our method in the contexts of both agricultural profits and GDP, but this model could be easily transported to a variety of other settings and sectors in the climate change damages literature.

My final chapter is co-authored with Kenneth Gillingham and James Stock. We compare several time series models to estimate the price elasticity of new vehicle sales, addressing the classic challenges of price and sales endogeneity and simultaneity in time series analysis with aggregate data. Correctly identifying the price elasticity of new vehicle sales is especially important for estimating the impacts of fuel economy standards because changing fuel economy stringency is assumed to cause a shock to new vehicle prices. The resulting effect on new vehicle sales has broad implications beyond the immediate impact on the vehicle industry. In particular, new vehicles generally have the best safety features and pollution controls, so reducing replacement of used vehicles has consequences for public safety and pollution levels. This project is also a novel application of a structural vector autoregression with instrumental variables (SVAR-IV), a relatively new methodology borrowed from the monetary policy literature. We compare the SVAR-IV with other time series approaches, some of which have been considered in policymaking for fuel economy standards.

Essays in Environmental and Energy Economics

A Dissertation
Presented to the Faculty of the Graduate School
Of
Yale University
In Candidacy for the Degree of
Doctor of Philosophy

By
Wade Davis

Dissertation Director: Kenneth Gillingham

June 2021

©2021 by Wade Davis

All rights reserved.

Acknowledgments

I am most thankful to my dissertation committee, Ken Gillingham, Robert Mendelsohn, and Matt Kotchen. Their comments and advice have greatly elevated the quality of these projects and contributed invaluably to my education in environmental and energy economics. As my principal advisor and the chair of my dissertation committee, Ken Gillingham has consistently challenged me to improve my research, expand my breadth, and learn new methods. I have been fortunate to work with Ken and benefit from his wide-ranging expertise across energy economics, economic tools, and research designs. As my mentor, Ken has been generous with his time as he has guided me in my search for impactful questions and the methods to answer them. Ken has encouraged me to become a more thoughtful researcher as well as a more precise and careful communicator. As both his student and Teaching Fellow for multiple semesters, I have also had the pleasure to learn from Ken in the classroom. His example will be a model for all of my future mentorship and teaching endeavors.

As a pioneer in the climate change literature, Robert Mendelsohn has been an ideal mentor for my foray into this subject and the research that has become the second chapter of this dissertation. Robert Mendelsohn has been a patient teacher throughout our work together, and the experience and knowledge he has shared have widened my thinking tremendously. As we have developed our project, I have come to more deeply appreciate the role of iteration and critical review in the research process. Next, Matt Kotchen's comments have been especially helpful in finalizing the structure and strategy of my first chapter. As his Teaching Fellow, I have been inspired by his compassionate approach to teaching the dismal science and his skill in presenting controversial conclusions to a sometimes-skeptical audience.

I would also like to thank my many other instructors, mentors, and fellow students whom I have encountered on the road towards a doctoral degree. I especially thank my fellow students and post-docs in the Yale Environmental and Natural Resource Economics (ENRE) Group, many of whom have been my late-night study companions and always indispensable sources of feedback. I thank my undergraduate instructors, Sarah Jacobson and Anand Swamy, who were among my earliest role models in economics. I also thank my Brattle Group mentors, Ahmad Faruqui, Ryan Hledik, and Neil Lessem, who all contributed so much to my professional development and igniting my passion for energy economics. Finally, I thank my partner, Sandra Lee, for her love and steadfast support, and I thank my parents, Wes and Elizabeth Davis, who have always been my champions.

Contents

Chapter 1: Extreme Contract Variety After Deregulation: Electricity Retail Choice in Texas

Abstract– This paper studies (1) the limitations of consumer decision-making in a setting with large choice sets and (2) the relationship between competition and product variety after deregulation.[1] In retail choice electricity markets, consumers choose their electricity contract from a competitive market of retailers, who are then responsible for procuring electricity from generators in the wholesale market. Retailers in this market compete on price, contract features, and additional services. I develop a model of consumer decision-making where consumers choose their multi-period sequence of cost-minimizing contracts subject to various constraints on their information sets and behavior. Using this model, I find that consumers usually fail to cost-minimize. Consumers spend a mean of $33 per month more on their actual contract choices than they would have on their ex post cost-minimizing contracts. I present evidence that it is unlikely that these missed savings are rationalized by discounting, risk aversion, or other behavioral explanations. Therefore, the value of missed savings captures the combined costs of search, inattention, and mis-selling in this market. This paper also presents the descriptive result that there is substantial heterogeneity in consumer preferences across contract features, which is not the case in regions where the monopolist utility has been allowed to remain in the retail market.

[1] **Acknowledgments and disclosures:** I did not receive any compensation for this work, and I do not have any conflicts of interest. I am grateful to ERCOT for their responses to my data requests and the anonymous electricity retailer who provided essential data for this project. I also thank my PhD advisor Kenneth Gillingham for his valuable advice and insights.

1.1 Introduction

Regulated monopolies and oligopolies have existed in many industries, including airlines, telecommunications, natural gas, and electricity, and regulators have had varying success in their efforts to open these markets to competition. Since the 1990s, there has been increasing deregulation in the electricity sector where markets have been largely controlled by monopoly utilities. Electricity deregulation has primarily focused on the wholesale market and enabling independent electricity generators to connect to the grid and compete on price. However, some jurisdictions have also introduced retail competition where consumers choose their electricity contract from a market of retailers, which are then responsible for procuring electricity from generators in the wholesale market. Under a system of retail choice, consumers continue to pay the monopoly utility for use of the transmission and distribution grid, but a consumer's bill for electricity consumption is determined by their chosen contract with their chosen retailer.[2] In marketing to consumers, retailers may compete on price, contract features, customer service, or various extras. Examples of these extra services include websites with information on electricity consumption, various smartphone apps, and compatibility with appliances, such as a programmable thermostat.

Proponents of retail choice argue that competition leads to lower prices, fewer price distortions, innovation, and variety. Researching the first five years of Texas' electricity retail choice market, one of the first in U.S., Puller and West (2013) estimate that retailers achieved modest price reductions relative to comparable regulated utilities over the same period. Indeed, Texas' electricity prices have remained among the lowest in the country and approximately 20 percent below the U.S. average.[3] However, previous research in retail choice markets has found strong evidence of consumer search frictions, including inattention, and inertia (e.g., Hortaçsu et al., 2017, and Giulietti et al., 2014). These frictions may enable retailers to exert market power and reduce consumers' ability to choose the contracts that would improve their welfare relative to the regulated setting. Therefore, policymakers implementing retail choice deregulation must trade off the goal of increasing competition

[2]Utilities that only operate transmission and distribution network are sometimes referred to as "wires companies."

[3]The Energy Information Administration estimates the 2019 average retail price in Texas was 8.6 cents per kWh as compared to 10.54 cents for the U.S. at large. U.S. EIA. State electricity profiles. November 2, 2020. https://www.eia.gov/electricity/state/

against concerns over consumer welfare and distributional effects.

This paper concentrates on the Texas residential electricity retail choice market. Policies that affect electricity prices are important because electricity is a ubiquitous good as both an essential household service and as an input to almost all other goods and services. Further, electricity contributes over 30 percent of greenhouse gas emissions.[4] Texas is a particularly important context as the largest producer and consumer of electricity and the largest emitter among all states.[5] I focus on the residential sector because commercial consumers tend to negotiate their contracts, so commercial offers are not typically advertised online, making it much more difficult to define a discrete choice set. Commercial consumers are also more likely to hire experts to optimize electricity contracts, which reduces concerns about inertia and inattention.

Electricity retail choice has been controversial with only 17 U.S. states having opened their electricity markets to some form of retail competition. Outside of the U.S., electricity retail choice is available to consumers in Alberta, Australia, Japan, and over half of Europe. However, many of these deregulated markets retain elements of their regulated predecessors. For example, some markets restrict retail choice to large consumers (i.e., commercial and industrial), and many have allowed the monopoly utility to remain in the retail market while continuing to operate the distribution and transmission network. Within the U.S., only Texas has required the monopolist utilities to exit the retail market. In all other retail choice states, the utilities retain over 50 percent market share and in most cases, over 80 percent.[6] That is, the utility offer is generally regarded as the default electricity contract, and there is relatively little switching to alternative retailers. In the states that have larger shares of consumers on alternative retailers (e.g., Ohio and Illinois with 46 and 35 percent, respectively), switching is largely driven by municipal aggregation where municipalities negotiate contracts with retailers on behalf of their residents.[7] In contrast, the participation in Texas' retail choice market is not driven by municipal aggregation but by the fact that there is no

[4]U.S. EIA. Energy-related carbon dioxide (CO2) emissions by source and sector for the United States, 2019. September 2020. https://www.eia.gov/tools/faqs/faq.php?id=75&t=11

[5]a) U.S. EIA. State total energy rankings, 2018. https://www.eia.gov/state/

b) U.S. EIA. State energy-related carbon dioxide emissions by year, unadjusted (1990-2017). May 20, 2020. https://www.eia.gov/environment/emissions/state/

[6]U.S. EIA. Electricity residential retail choice participation has declined since 2014 peak. January 17, 2019. https://www.eia.gov/todayinenergy/detail.php?id=37452#

[7]Municipal aggregation is sometimes referred to as community choice, as in California.

default offer from a monopolist utility. That is, Texas retail choice consumers not only have the power to choose, they are required to choose if they desire electricity.[8]

The most surprising development in Texas' residential retail choice market is the large number of retailers and the extreme variety of contract features. Texas now has over 40 retailers offering hundreds of contracts. The Texas market stands in stark contrast to the regulated monopolies in other states where consumers typically have one or two contracts to choose between. For example, Texas' numerous contracts offer a variety of nonlinearities and many different levels of fixed, variable and time-varying components. Contracts also vary in durations from one to 36 months, usually with steep fees for early termination. Some contracts are branded as green and promise to procure a certain percentage of a consumers' electricity consumption from renewable energy sources. Some retailers even include sign-up bonuses.

Texas retailers have had the ability to offer time-varying contracts since the utilities deployed smart meters, which they mostly completed by 2015. Smart meters can record 15-minute interval electricity consumption and transmit this information to consumers and retailers in real time. Since Texas wholesale prices are determined on 5-minute intervals, time-varying contracts send consumers the most accurate price signals, which can lead to more efficient allocations of generation as well as generator and transmission investments. More efficient allocation of electricity grid resources could lead to lower electricity costs and also emissions because the power plants that most increase their generation during peak times tend to have the highest emissions of greenhouse gases and air pollutants. Thus, a retail choice market may lead to environmental benefits to the extent that the market incentivizes consumers to switch to time-varying contracts and shift consumption from peak hours. For example, if time-varying contracts encourage consumers to switch consumption to times with more solar or wind generation, that may incentivize investments in these renewable technologies.

My first contribution is a model of exhaustive consumer search under various restrictions on con-

[8] A minority of Texas consumers, fewer than 20 percent, fall outside of the retail choice market either because they are served by utilities outside of the ERCOT territory or because they are still served by municipal or cooperative utilities that were not required to adopt retail choice. ERCOT is the nonprofit independent system operator for most Texas. Austin Energy and CPS Energy serving San Antonio are two of the largest municipal utilities that were not required to participate in retail choice.

sumers' information sets. That is, consumers seek their cost-minimizing multi-period sequence of contracts subject to their individual consumption characteristics and various degrees of uncertainty about future prices and contract offers. The most complex aspect of this model is accounting for the various contract durations. The opportunity cost of signing a long-duration contract is determined by the probability that a more favorable contract will become available. Conversely, the opportunity cost of waiting to sign a long-duration contract is determined by the probability of missing out on a more favorable offer.

Comparing my model estimates to consumers' observed contract choices, I find consumers spend a mean of at least $33 per month more on their actual contract choices than they would have on their ex post cost-minimizing contracts. I present a series of evidence to rule out possible rationalizations for consumers' behavior including (1) discounting; (2) uncertainty about future consumption, prices, and contract features; (3) risk aversion; and (4) green preferences. Thus, I conclude that consumers' deviations from the cost-minimizing contract choices are the result of some combination of true search costs, inattention (i.e., mistakes), and mis-selling. Mis-selling may occur when a sales representative guides a consumer onto a sub-optimal contract.

While the literature on consumer search and decision-making is vast, the most relevant studies are settings with choices over contracts. This literature includes Gruber and Verboven (2001) and Miravete (2003) on phone service providers and plans; Nevo et al. (2016) on broadband contracts; Abaluck and Gruber (2011) on Medicare Part D supplemental health insurance; Giulietti et al. (2005) on natural gas retail choice in Britain; and Giulietti et al. (2014) on electricity retail choice in Britain. The most closely related paper is Hortaçsu et al. (2017), which I refer to hereafter as HMP. HMP develop a two-stage model of residential consumer search for the Texas electricity market. The first stage is the decision whether or not engage in search (i.e., the decision to choose), and the second stage is the choice of retailer. HMP conclude that the market had substantial search frictions driven by inattention and brand loyalty.

While HMP study the first five years of the market (2002 to 2006), I study the market over a decade later (2015 to 2019). The first five years of the market were a transition period in which the monopolist utility remained in the retail market charging a regulated "price-to-beat." In fact, by the end of 2006, less than 40 percent of the residential market had switched from the monopolist utility to a new entrant retailer. Thus, a key feature of the HMP model is the brand advantage of the utility. Since 2007, the utility is no longer a choice in the market. Also, by studying the first years of the market, it is possible that HMP captured a snapshot of the consumer learning process, and indeed, HMP hypothesize that one source of the utility brand advantage may have stemmed from consumers misunderstanding the principle that "it is all the same power." Furthermore, the HMP data sample ends before the key developments of smart meter deployment, the emergence of time-varying contracts, the substantial entry of new retailers, extreme contract variety, and the rise of multi-contract retailers. In the HMP data, consumers choose between fewer than ten retailers, and most retailers only offered a single contract. In fact, in the few cases where retailers offered multiple contracts, HMP modeled them as single-contract firms. A principal feature of my more recent retail choice context are the tradeoffs between short- and long-term contracts, which are not accounted for in HMP.

The many changes in the retail choice market since the end of the HMP empirical setting have led to my proposal for a model where consumers engage in frictionless, mistake-free search for their cost-minimizing sequence of contracts.[9] One motivation for my model is the emergence of concierge services, which could presumably solve this problem on consumers' behalf. Alternatively, the functionality of the regulator-run marketplace (*powertochoose.com*, discussed below) could be improved or another tool developed to assist consumer search. Even if consumers choose to engage in search in every period in the first stage of the HMP model, they would still make mistakes with nonzero probability in the second stage due to the probabilistic nature of the logit specification. On the other hand, my model evaluates observed consumer decisions relative to my estimates of their hypothetical mistake-free decisions.

Another difference between my model and HMP is that HMP assume that consumers have perfectly inelastic electricity demand and that the level of consumers' electricity consumption does not affect their choice probabilities.[10] Lastly, my model estimates contract choices among the offers of a single

[9] I discuss the relationship between cost minimization and utility maximization in Section 1.3.

[10] HMP define their price variable to be the average bill for 1,000 kWh with each individual retailer in each time period. This assumption could be relaxed by modifying the definition of the price variable.

retailer, which allows me to avoid the problem of separately identifying brand advantage. This is a particularly nice feature of my model in the modern context where brand advantage could play a substantial role given the heterogeneity in customer service, websites, smartphone apps, and other retailer-level characteristics.

Wilson and Price (2010) (hereafter WP) also present a model that is closely related to mine in their study of the British electricity retail choice market. Using a survey of residential consumers in combination with retailer contract offers, WP compare consumers' chosen contracts to their cost-minimizing contracts. Similar to my results, WP found that only 8 to 20 percent of consumers chose the cost-minimizing contract, but the costs of failing to do so were much lower– only \$3 to \$7 dollars per month.[11] WP conclude that inattention, pure decision errors, and mis-selling are the most likely explanations for consumers' failures to select cost-minimizing contracts. As with HMP, my paper has several important differences from WP with respect to context, data quality, and modeling requirements. In particular, the monopolist utilities were allowed to remain in the market and functioned as the default retailer. Interestingly, WP found that 17 percent of consumers actually chose a higher-cost contract when switching away from the utility.

Also in contrast to my setting, WP note the lack of contract variety, which began to increase in the period after the end of their data. The market had approximately 15 retailers each offering essentially one contract.[12] As the WP setting featured a relatively small choice set where contracts did not feature different durations or time-varying components, my model accounts for a much more complex consumer decision-making process. Most importantly, while WP model a single contract choice, I model a a multi-period sequence of contract choices. Furthermore, WP do not observe actual consumer consumption data. Instead, the authors rely on consumers' consumption estimates provided in their survey responses. Thus, while WP rely largely on various survey questions to explore possible rationalizations for the observed decisions, I am able to exploit the richness of the contract variety and observed consumption data to dismiss such possibilities as discounting, risk aversion, and uncertainty. Lastly, in estimating my model over a single retailer's contracts, I am

[11] I converted WP's values from 2010 British pounds to 2018 USD for comparison with my estimates.

[12] In fact, retailers offered one contract for each of three payment methods– credit, debit, and pre-pay, but the prices and features were similar across payment methods.

able to eliminate the explanation of brand advantage.

My second contribution in this paper is the descriptive result that by revealed preference, consumers have a high degree of heterogeneity in their preferences across electricity contract features. A cursory internet search of Texas' retail choice market reveals the large number of retailers and wide variety of contract features, but confidential data is necessary to observe contract uptake. I obtained this data from an anonymous Texas retailer I refer to hereafter as Retailer A. To my knowledge, the only other data on consumer contract choices in the Texas market is the annual ERCOT Demand Response Survey, which I describe in more detail in Section 1.2. While the Demand Response Survey only describes uptake within a few broad categories of time-varying contracts, my data describes consumer uptake across a wide variety of contract features.

Beginning with Hotelling (1929), there is a substantial literature on the relationship between competition and product variety, but these models do not provide obvious predictions in the electricity context. In his model of firms competing on price and location, Hotelling finds a tendency towards imitation and that competition may actually reduce variety. However, subsequent literature, beginning with Spence (1976) and Lancaster (1979 and 1980), finds that the relationship depends on the tradeoff between firms' optimization with respect to consumer preferences versus reducing unit costs via scale economies. These papers model monopolistic competition between firms producing products with various degrees of substitutability.

While there are certainly administrative and marketing costs of additional electricity contract offers, the marginal costs of producing additional contracts is very low compared to other settings featured in the literature, such as producing variety in durable goods, television channels, and even breakfast cereals. Thus, as multi-product firms, the electricity retailer's optimal contract offerings are likely to depend on additional considerations, including (1) cannibalization, (2) defining contract features to achieve separating equilibriums, and (3) the heterogeneity in search frictions, such as inertia and inattention. Given the complexity of the retailer problem and the degree of consumer search frictions, I leave it to future work to develop a model to describe the welfare-maximizing level of contract variety as well as the relative welfare effects at different levels of contract variety. However, my evidence does suggest that consumers are constrained in the regulated monopoly setting with only one or two contract offers. That is, consumers have heterogeneous contract preferences that they cannot express in the monopoly setting.

My paper is organized as follows. Section 1.2 provides more details and history of Texas' retail choice market. Section 1.3 lays out my model of frictionless consumer search. Section 1.4 describes my proprietary data obtained from anonymous Retailer A. Section 1.5 presents the heterogeneity in the consumer preferences across electricity contract features. Section 1.6 compares my model estimates of consumers' cost-minimizing contracts to observed contract choices. Section 1.7 concludes.

1.2 Retail Choice in Texas

Beginning in 2002, Texas required its five primary utilities to allow retail competition for all customers in their service territories, encompassing for over 80 percent of Texas households.[13] There is no default retailer or contract, and consumers must make a choice when signing up for electricity service.[14] As mentioned above, the first five years of the market were designated as a transition period during which utilities still participated in the retail market offering a "price-to-beat" contract. Since 2007, utility contracts are no longer available, and consumers must make their choices among the other retailers. The regulated monopolies now act strictly as transmission and distribution utilities (TDUs), and their rates for these services remain regulated by the Public Utility Commission of Texas (PUCT). It is the responsibility of retailers to collect these TDU service charges from their customers. Thus, consumers receive only one monthly bill for their electricity consumption and related services.[15]

[13] a) Texas' five investor-owned utilities (IOUs) are Oncor (formerly TXU), CenterPoint (formerly Reliant), AEP Central (formerly Central Power and Light), AEP North (formerly West Texas Utilities), and Texas-New Mexico Power (TNMP). The size of these utilities ranges from TNMP with only 250,000 customers to Oncor, the sixth largest utility in the United States with 3.5 million customers.

b) The five service territories opened to retail competition comprise the majority of the Electricity Reliability Council of Texas (ERCOT), the independent system operator (ISO) for most of Texas. As mentioned above, utilities not required to participate in retail choice include cooperatives, municipal electricity utilities, such as Austin Energy, and utilities outside of ERCOT.

[14] The PUCT has designated certain retailers "providers of last resort" (POLR) with PUCT-accepted contracts in place. Consumers are only switched to a POLR contract if their chosen retailer fails or for some other reason exits the market.

[15] Bills also include administrative charges imposed by ERCOT and the PUCT.

Retailers are also responsible for procuring their customers' electricity on the wholesale market. Bilateral contracts account for over 80 percent of Texas' wholesale electricity sales, but retailers may also procure electricity on the public day-ahead and real-time auction markets.[16] If retailers fail to procure their customers' exact electricity consumption at each fifteen-minute interval through bilateral contracts and the day-ahead market, they must purchase the residual on the real-time market. Retailers can reduce costs by negotiating more favorable contracts with generators and improving their strategies to hedge their uncertainty about future prices and their customers' consumption profile.

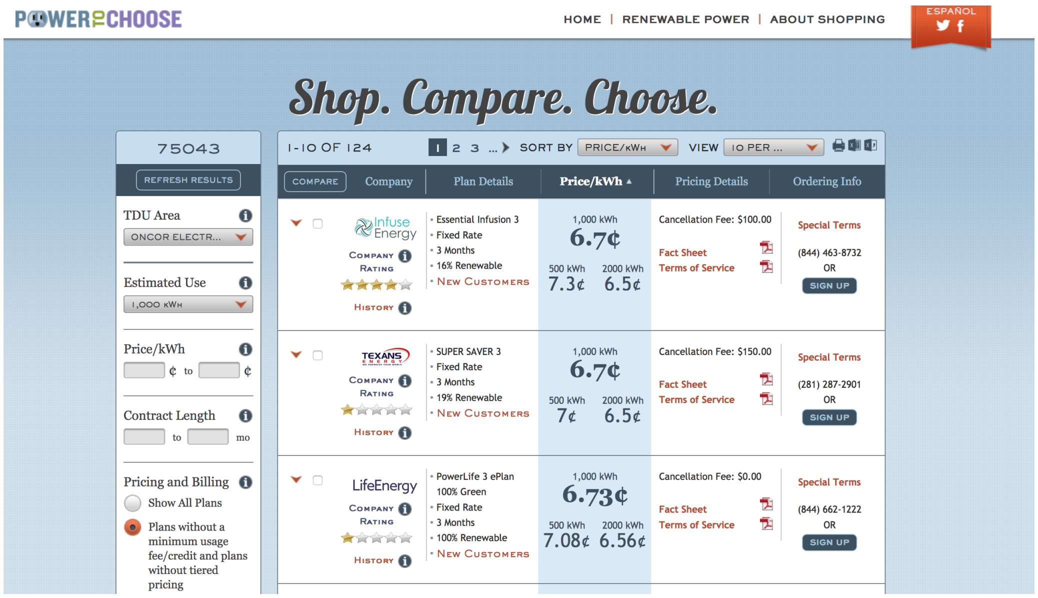
Consumer inertia and inattention may be mechanisms through which retailers continue to exercise market power. The retail divisions of the five TDUs were largely incorporated into two retailers, TXU and Reliant. These two retailers now operate across all five markets, and each have nearly 20 percent market share. Although over 40 subsequent retailers have entered the market, none have achieved more than five percent market share.

To facilitate consumer search, the PUCT introduced the *powertochoose.com* website, a screenshot of which is shown in Figure 1.1.[17] However, this website is not the only marketplace for consumers to learn about contract offers, and some retailers do not even advertise there. Other channels through which retailers advertise their offers include their own websites and also other aggregator websites, some of which charge retailers to list their contracts. As mentioned above, commercial consumers tend to negotiate their contracts, so specific commercial offers are rarely posted online.

[16]Day-ahead and real-time auctions are operated by ERCOT. Generators actually bid into 5-minute-interval wholesale market, but consumptions and prices are aggregated to 15-minute intervals for settlement with retailers.

[17]The PUCT periodically changes the rules for sorting and listing contracts on *powertochoose.com*. Currently, the default settings sort contracts on the average price per kWh for a consumer with an average load profile using 1,000 kWh per month. Consumers can also sort on a variety of other retailer and contract characteristics. Retailer ratings are determined by consumer complaints filed with the PUCT. Recent policy changes restrict the number of contracts each retailer can list on the website. Consumers must enter their zip code because offers vary by TDU service territory.

Figure 1.1: *powertochoose.com* Contract Offers



Notes: (1) powertochoose.com website results for zip code 75043 in Dallas. (2) These are the first three of over 100 offers. (3) The sidebar and dropdown menus show some of the options for filtering contract results.

Legally, retailers may select on consumer characteristics and engage in price discrimination through (1) targeted advertising, (2) defining contract features to achieve separating equilibrium, and (3) mis-selling.[18] Retailers may also freely switch customers onto any variable-rate month-to-month contract when their original contract expires, but they must send customers a reminder notice before doing so. Some disincentives for switching customers onto exorbitantly-priced contracts are (1) the possibility of receiving a poor rating from the PUCT to be displayed online, (2) inducing customer switching to a competitor, and (3) legal actions by customers.

Recently, concierge services have emerged offering to periodically search and switch consumers to lower cost contracts. For example, the company Energy Ogre charges customers $10 per month for this service.[19] However, large retailers have filed legal challenges, arguing that concierge services

[18] An example of potentially illegal price discrimination occurred in 2018 when retailers were found to be listing different contracts on the Spanish and English versions of the *powertochoose.com* website. See: Sixel, L.M.. Houston Chronicle. Regulators crack down on power companies that don't offer same deals to English, Spanish speakers. September 14, 2018. https://www.houstonchronicle.com/business/article/Texas-regulators-crack-down-on-electricity-13231266.php

[19] https://www.energyogre.com

do not have the authority to terminate and sign contracts on behalf of customers.[20] The findings of this paper suggest that such services could improve consumer welfare to the extent that that they reduce search costs and inattention. One concern with these new services is the potential for retailer capture where a retailer could pay a concierge service to switch consumers onto its own less favorable contracts.

Retailers have had the ability to offer time-varying contracts since the TDUs deployed smart meters, which they began around around 2009 and mostly completed by 2015.[21] Prior to smart meters, retailers only knew their customers' aggregate consumption based on monthly meter reads, and retailers were billed for their customers' estimated consumption on the hourly wholesale market based on the assumption that consumers had a market average hourly consumption profile. Thus, the introduction of smart meters has ensured that retailers are more accurately billed for their customers' energy consumption, and it has allowed retailers to offer time-varying contracts. The introduction of smart meters appears to have changed retailer entry incentives, and indeed, most of the increase in the number of retailers and contract offerings has occurred since 2010, prior to which there were only about 10 retailers in the market, primarily offering only a single contract each.

Interestingly, time-varying contract offers have mostly taken the form of free nights or free weekends. Regulated utilities in other states are increasingly offering time-of-use (TOU) rates, which typically feature weekend and off-peak rates that are several cents per kWh less expensive than the on-peak price.[22] In contrast to the Texas offers, however, these TOU contracts almost never provide free electricity. Smart meters have also allowed retailers to offer real-time pricing (RTP) where prices change throughout the day to reflect the wholesale market. For example, the retailer Griddy charges customers $9.99 per month and then the 15-minute wholesale price for all of their consumption.[23]

[20] Sixel, L.M.. Houston Chronicle. Power companies can still reject you if you paid to find low price. May 7, 2020. https://www.houstonchronicle.com/business/energy/article/Sites-that-find-electricity-deals-get-little-help-15254309.php

[21] a) The TDUs began smart meter deployment around 2009 with Oncor and CenterPoint achieving full deployment in 2012 and AEP and TNMP following slightly slower roll-out schedules.

b) The Edison Foundation Institute for Electricity Innovation. Utility-scale smart meter deployments, plans, & proposals. 2012. http://www.edisonfoundation.net/iei/publications/Pages/publications.aspx?category=Report

c) Cooper, Adam. The Edison Foundation Institute for Electricity Innovation. Electricity company smart meter deployments: Foundation for a smart grid. 2016. http://www.edisonfoundation.net/iei/publications/Documents/Final Electricity Company Smart Meter Deployments- Foundation for A Smart Energy Grid.pdf

[22] A typical TOU offer at a regulated utility features a higher rate for pre-defined peak hours, such as weekdays from 2PM to 10PM.

[23] a) Griddy also encourages customers to use smart appliances, such as programmable thermostats. Customers

Although RTP is the most economically efficient pricing mechanism (i.e., providing consumers the most accurate price signals, e.g., Borenstein and Holland 2005), regulated utilities in other states rarely offer it. As of 2019, approximately 7.7 percent of ERCOT residential consumers were on TOU rates, while 0.2 percent were on RTP.[24]

1.3 Model

In my model, consumers choose a sequence of contracts $(c_{nt})_{t=1}^{T}$ to minimize their expected sum of discounted bills from the initial period $t = 1$ to some future $t = T$. A contract c_{nt} is indexed by both type n and period t because two contracts of the same type may have a different prices or features depending on when a consumer signed up (i.e., $c_{nt} \neq c_{n,t-1}$). In the context of electricity contracts, it is helpful to think of a type n as a brand name or a dominant feature. For example, "Truly Free Weekends" is the brand name of one of the 12-month contracts offered by the retailer Reliant, and while the brand name and general tariff structure remain the same over time, prices and exact terms are routinely updated. In any given period, a consumer's contract choice set is the set of all contracts offered in that period $\{c_{nt} : t = t, n \in N_t\}$ where N_t is the set of contract types offered in that period. Thus, a consumer's cost-minimization problem is

$$\begin{aligned} &\min_{(c_{nt})_{t=1}^{T}} E_{t=1}\Big[\sum_{t=1}^{T} \beta^t c_{nt}(q_t)\Big] \\ &\text{such that} \\ &1) \quad c_{nt} = c_{n,t-1} \quad \text{if } c_{n,t-1} \text{ ends in period } t \text{ or later, and} \\ &2) \quad c_{nt} \in \{c_{nt} : t = t, n \in N_t\} \quad \text{otherwise,} \end{aligned} \tag{1.1}$$

can program these appliances to turn down or turn off in response to various price thresholds via Griddy's partnership with the company IFTTT (i.e., If This Then That). https://www.griddy.com/

b) ERCOT currently imposes a wholesale price cap of $9,000 per MWh (i.e., $9 per kWh).

[24] ERCOT's 2019 Demand Response Survey only includes consumers in ERCOT's retail choice jurisdictions and does not require responses from small retailers. The survey counts premises, not customers. I note the distinction because one customer, such as a landlord, could operate at multiple premises. The 7.7 percent of residential premises on TOU rates was up from 4.1 percent in 2014. Also, 6 percent of residential premises were on a peak-time rebate where customers are paid for estimated consumption reductions after receiving alerts on high-price days. Peak-time rebates and TOU rates are not mutually exclusive contract features.

where $c_{nt}(q_t)$ is the consumer bill in billing period t; q_t is the consumer's consumption; and β is a discount factor. Note that in the case of electricity, consumption q_t may be a complex object, such as a set of billing determinants, including consumption during particular hours of the day. The main constraint on consumer decisions is that they do not end contracts prematurely. This constraint is especially reasonable in the Texas retail choice setting where early termination fees are prohibitive (often over $100).

The time indices in this model are relative, and the process in dynamic. That is, consumers choose an initial sequence of contracts, but whenever the first contract in a sequence expires, they search again and re-solve Equation 1.1. Also, for a period in which a consumer makes a contract choice, the choice set is defined at the beginning of the period, while consumption and the subsequent bill are realized at the end of the period. This model requires consumers to be perfectly attentive to their contract end dates and engage in search rather than be switched onto a more expensive default contract. In the results of Section 1.6, I show that this type of inattention only partially explains consumers' deviations from their cost-minimizing contract sequences.

The main source of dynamics in this model is the tradeoff between long-term and short-term contracts. Consumers may pay a premium for a long-term contract if they expect higher prices or less favorable contract terms in the future. They may sign a short-term contract if they expect future offers to feature lower prices or more favorable terms. That is, consumers must account for the opportunity cost of signing a contract of a particular duration.

Consumers face two forms of uncertainty in their decision-making: (1) uncertainty about their future consumption and (2) uncertainty about future prices and contract offerings. As a set of starting assumptions, my baseline estimates assume that consumers have perfect information:

Assumption 1 (Perfectly predict consumption) $E_{t=0}q_t = q_t$ *for all* t*;* and

Assumption 2 (Perfectly predict prices and contract offers) $E_{t=0}\{c_{nt} : t = t, n \in N_t\} = \{c_{nt} : t = t, n \in N_t\}$ *for all* t.

The perfect information case of Assumptions 1 and 2 is equivalent to estimating a consumer's ex post cost-minimizing contracts. As I search for rationalizations of consumer choices in Section 1.6, I sequentially relax these assumptions about consumers' information sets. In particular, I explore whether consumer decisions could be explained by their uncertainty about future consumption, prices, or contract offers.

In this model, consumers are cost-minimizers, not utility-maximizers. Thus, it is possible that consumers' deviations from the model's estimates of the cost-minimizing contracts could be explained by parameters of consumers' utility functions, including risk aversion, preferences for particular contract features, and search costs. If this model were estimated over contracts from multiple retailers, the model would also need to account for firm-level differences as found in HMP. However, my results in this paper only consider consumer search over the contracts of a single retailer, which eliminates concerns about differences in retailer quality brand advantage. Even the contract offerings of a single retailer present a large choice set as described in Section 1.4. A consumer's degree of risk aversion may explain why consumer choices deviate from the model's cost-minimizing contracts because in the presence of uncertainty. For example, a consumer's degree of risk aversion will affect the level of premium they are willing to pay for a long-term contract to insure against future price shocks.

The model also assumes that consumers have no inherent taste for contract features– instead, a consumer's preferences across contracts are purely determined by how the contract features affect the consumer's bills in the cost-minimization problem of Equation 1.1. One counterexample could be green contracts where some consumers are willing to pay a premium to guarantee that a certain share of their electricity is procured from renewable generators. The value of such contracts could include a warm glow effect from the knowledge that the consumer may be reducing pollution. To the extent that search costs discourage search and constrain consumers' consideration of the full choice set, search costs could also rationalize deviations from my model predictions. In fact, consumers might even pay a premium for long-term contracts to increase the time until the next search and to reduce the number of opportunities to be inattentive and get switched onto an unfavorable default contract. In Section 1.6, I perform a series of tests to explore the extent to which these components of consumers' utility functions could rationalize their observed choices.

Three more qualifications should be considered when comparing my model estimates to actual consumer choices. First, the price elasticity of demand is another parameter of consumer utility functions that is not captured in my model. Within a conventional range of elasticities, I expect consumers to reduce consumption on more expensive contracts, increase consumption on less expensive contracts, or shift consumption on time-varying contracts. Thus, in not fully accounting for consumer elasticities, my model understates welfare gains from switching contracts. Second, while the model could accommodate contracts from multiple retailers, my estimates only include the offers of a single retailer. This also leads to underestimating the potential savings of consumers who failed to choose a cost-minimizing contract. Third, the model only estimates partial equilibrium results. If all consumers attempted to switch from their observed contract choices to their cost-minimizing contract as estimated by the model, retailers would surely respond with updated prices and contract offerings.

1.4 Data

My data consists of a random sample of 5,000 residential customers in Texas, who were customers of an anonymous Retailer A at any point between January 2017 and August 2019. The earliest customer joined in August 2015, and data extends through November 2019 for those customers who did not switch away to another retailer. The shortest duration customers only stayed with Retailer A a few days, while some customers stayed the entire sample period.[25] I observe customers' 15-minute smart meter interval data, contract choices, and monthly invoices. After accounting for various data discrepancies, 4,208 customers remain in the sample.

In many periods, Retailer A offers over 40 unique contracts across all of Texas' retail choice service territories, which provide a good representation of the contracts in the marketplace. These contracts include a variety of durations, time-varying features, and renewable energy content. Unlike some of

[25]Appendix Figure A.1 shows the histogram of contract durations in the data. Of course, this distribution is skewed towards zero, since the data includes many contracts that are still active.

its competitors, Retailer A is not specialized in a particular contract type. For example, the retailer Griddy only offers real-time price contracts, while the retailer Green Mountain Energy focuses its marketing on its renewable energy contracts. Table 1.1 compares the demographics of the customers in Retailer A's sample to Texas at large, suggesting that Retailer A customers live in zip codes that are wealthier, whiter, more educated, and more English-speaking. This comparison suggests that customers at other Texas retailers might be even less likely to select cost-minimizing contracts to the degree that lower education and English language ability exacerbate inattention and search costs. In particular, retailers do not always offer the same contracts in Spanish as they do in English, which could make Spanish-only speakers an especially vulnerable group in this market.[26]

Table 1.1: Sample Demographics of Retailer A Compared to Texas

		Texas	Retailer A sample	t-test p-value
[1]	Median household income (2018 dollars)	\$59,570	\$77,790 (453)	0.00***
[2]	Share of population with Bachelor's degree	19%	26% (0.16)	0.00***
[3]	Share of housing units that are owner occupied	62%	58% (0.31)	0.00***
[4]	Share of population that is white	42%	48% (0.30)	0.00***
[5]	Share of population that is Hispanic or Latino	39%	28% (0.28)	0.00***
[6]	Share of population that speaks English poorly	12%	8% (0.13)	0.00***

*Notes: (1) Standard errors in parentheses. (2) ***$p < 0.01$. (3) The Retailer A sample is described in Section 1.4. (4) Demographic data is from the American Community Survey 5-year estimates (2014-2018). (5) The sample demographics are a weighted average of the zip codes represented by Retailer A's customers.*

Summary statistics in Table 1.2 show that median customer in the Retailer A sample pays a lower price for electricity than both the Texas and U.S. average. Also, the median Retailer A customer uses less electricity than the Texas average but more than the U.S. average. Specifically, the median Retailer A customer pays \$0.102 per kWh, which compares to \$0.118 for Texas at large and \$0.130

[26] See Footnote [18].

for the entire country.[27] The median customer in the sample uses uses 918 kWh per month, which is below the Texas average of 1,140 kWh but above the U.S. average of 887 kWh.[28] In the Retailer A sample, the median customer bill is $88 per month, 56 percent of which is the retailer portion and 44 percent of which are charges passed through from the respective TDU in that service territory.

Table 1.2: Retailer A Sample Bills and Electricity Consumption

		All invoices
[1]	Median monthly electricity usage (kWh)	918
	Median monthly bill	
[2]	Total	$88
[3]	Retailer portion	$51
[4]	Transmission and distribution	$39
[5]	Median electricity price ($/kWh)	$0.102
[6]	Average electricity price ($/kWh)	$0.104
[7]	Number of customers	4,208
[8]	Number of invoices	41,356

Notes: (1) The Retailer A sample is described in Section 1.4. (2) Medians are taken across consumers, not invoices.

1.4.1 Defining Consumer Choice Sets

In addition to the customer consumption, invoices, and contract choices data, Retailer A also provided their contract database, which I used to define the consumer choice set in each period. This choice set is the object $\{c_{nt} : t = t, n \in N_t\}$ defined in section 1.3. Recall that a contract c_{nt} is indexed by both type n and period t and that N_t is the set of contract types offered in that period t. Retailers frequently offer a relatively stable set of contract types over time, often indicated by a brand name like "Truly Free Weekends," but they routinely update the precise prices and terms. I assume that a contract is in a consumer's choice set if another consumer selected the same contract within 30 days of the consumer's choice date. Indeed, my many visits to Retailer A's website confirmed the availability of most contract offerings over time. Most consumers had 40 to 50 Retailer A contracts in their choice set in each period. In Section 1.6, I explore relaxing the

[27] See Footnote [28]

[28] U.S. EIA. Residential average monthly bill by Census Division, and State. October 6, 2020. https://www.eia.gov/electricity/sales_revenue_price/

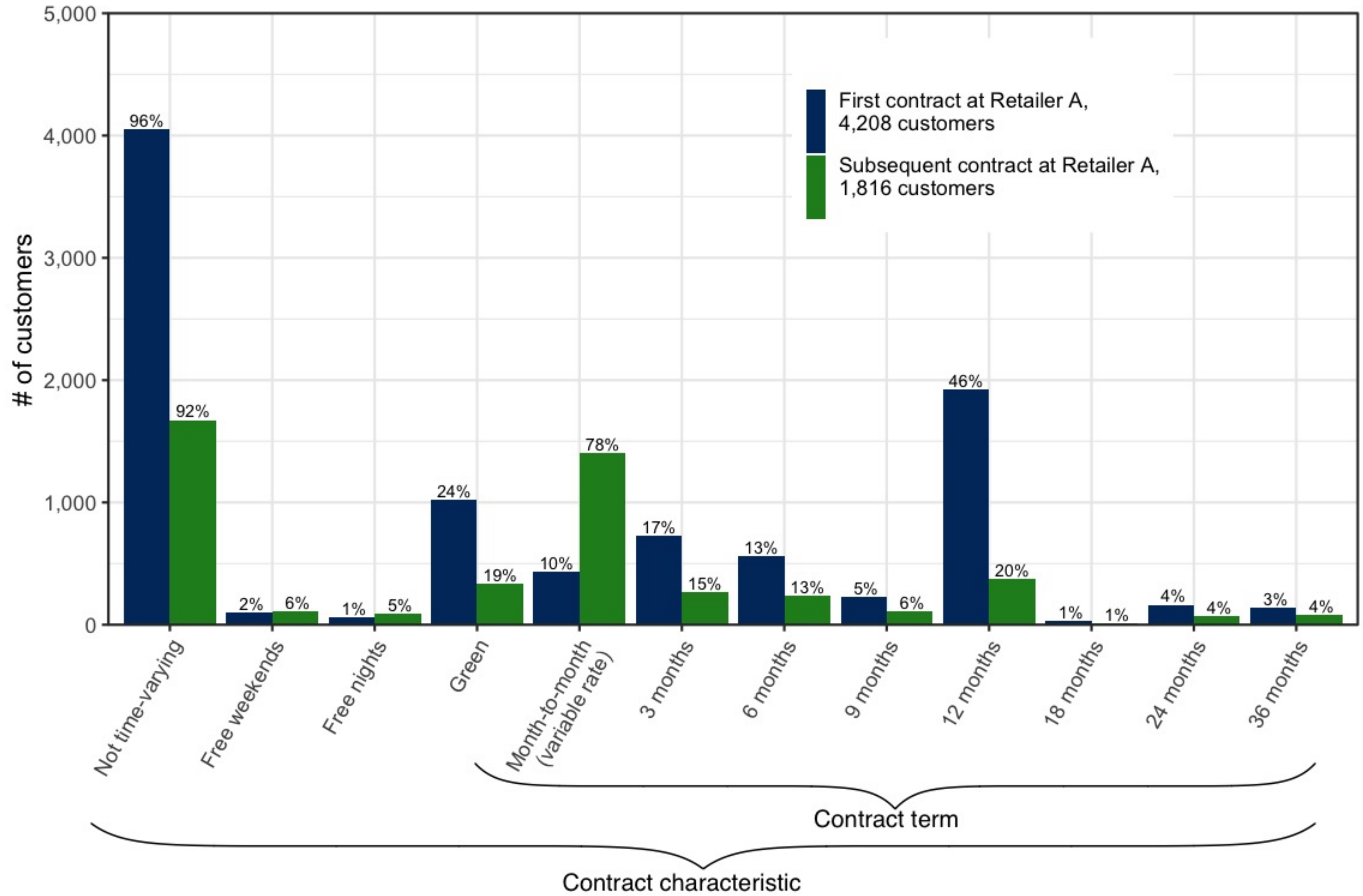
definition of the choice set to include all contracts in the Retailer A database on a given date.

1.5 Consumer Taste Heterogeneity and Switching Behavior

Figure 1.3 shows consumer choices across a variety of contract features, evidencing the heterogeneity in consumer preferences. While 12-month non-time-varying contracts are the most popular, no particular set of contract features is dominant. Interestingly, 24 percent of new Retailer A customers select a green contract, which Section 1.6 shows are about 3 percent more expensive per kWh than the Retailer A alternatives. Only three percent of consumers select time-varying contracts as their first contract with Retailer A, which is below the 7.7 percent Texas retailer average cited in Section 1.2. However, 11 percent of Retailer A consumers selected a time-varying contract as one of their subsequent contract choices.

The marked increase in month-to-month contracts as consumers' subsequent contract is the result of consumers failing to make another selection at then end of their contract term and being placed onto one of Retailer A's default contracts. In fact, I observe 29 percent of all consumers and 56 percent of consumers that stayed a full contract term being switched onto a default contract. Section 1.6 shows that a default contract is approximately forty percent more expensive per kWh than the Retailer A alternatives, which is an additional $36.72 for the median consumer. Among the consumers who were switched to a default contract, I observe 55 percent remaining on these contracts for more than two months, and 70 percent eventually leave Retailer A rather than choosing one of Retailer A's alternatives. Actually, 58 percent of all Retailer A customers leave Retailer A during the sample period, which suggests that consumers engage in relatively frequent switching between retailers as well as contracts. This rich switching behavior is explored more fully in Appendix Table A.1.

Figure 1.2: Most Popular Contract Features at Retailer A



Notes: (1) Blue bars indicate consumers' first contract choices at Retailer A, while green bars indicate their subsequent contract decisions. I observe 1,816 consumers make subsequent selections. (2) Bars are not mutually exclusive because contracts may have multiple features and consumers may have been with Retailer A long enough to select multiple successive contracts. For example, one 3-month contract could have both free nights and higher renewable energy content. (3) The high share of consumers switching to monthly contracts is because many consumers are defaulted onto these contracts when they fail to sign a new contract upon the expiration of their existing contract. Consumer experiences with the default contract are described in Appendix Table A.1. (4) The percentage labels on each bar indicate the share of consumers experiencing each contract characteristic. For the green bars illustrating subsequent contract choices, the percentages reflect the share of the 1,816 consumers.

1.6 Comparing Consumer Choices to Model Estimates

I now compare consumers' observed contract choices to my model estimates of their cost-minimizing contracts. Under a variety of assumptions and subsets of the data, I sequentially rule out various behavioral rationalizations for consumer choices and conclude that deviations from cost-minimizing contracts are best explained by a combination of true search costs, inattention, and mis-selling. Figure 1.3 compares the relative frequency of contract features between consumers' observed contract choices and my model's baseline estimates of their cost-minimizing contracts. As described in Section 1.3, the baseline estimates are based on the perfect information Assumptions 1 and 2 (i.e., ex post analysis) with no discounting.

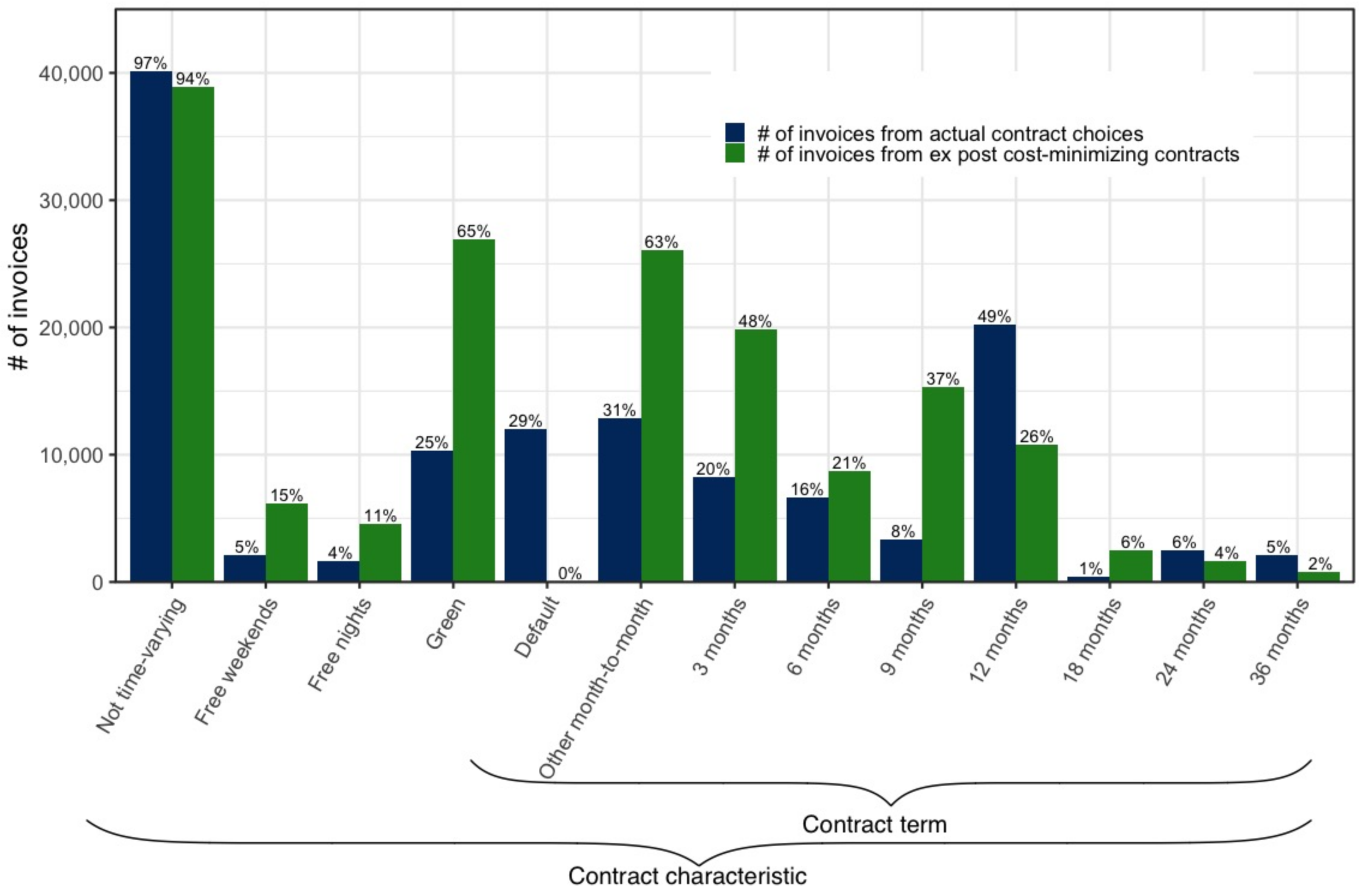
Figure 1.3 shows that consumers should have chosen short-term and time-varying contracts much more frequently to realize their cost-minimizing contracts. In particular, consumers should have chosen 12-month contracts much less frequently, non-time-varying contracts slightly less frequently, and the default contracts almost never. Instead, consumers should have much more frequently chosen month-to-month, 3-month, 9-month, and time-varying contracts. However, realizing the cost-savings of the month-to-month contracts would require consumers to have a high level of attention and low search costs because they would need to actively choose contracts almost every month to avoid being switched onto the default contract. Importantly, though, the model only prescribes this frequent-switching strategy for 63 percent of consumers, and even among these consumers, month-to-month contracts are only cost-minimizing in a subset of billing periods. For example, the model prescribes 48 percent of consumers at least one 3-month contract; 37 percent of consumers at least one 9-month contract; and 2 percent of consumers a 36-month contract.

Interestingly, the model also finds that 65 percent of consumers could have cost-minimized by choosing at least one green contract, as compared to the 25 percent of consumers who actually did so. This finding runs counter to the the intuition the green contracts should charge a premium, since procurement is constrained to the renewable energy portion of the wholesale market.[29] While the median consumer on a green contract pays approximately three percent more per kWh (see

[29] Green contracts guarantee a certain share of renewable energy procurement, which may be less than 100 percent.

discussion of Table 1.5), there are many periods where some green contracts are less expensive for a large subset of consumers.

Figure 1.3: Features of Ex Post Cost-Minimizing Contracts, $\beta = 1$



Notes: (1) Blue bars indicate consumers' actual contract choices, while green bars indicate my model estimates of their ex post cost-minimizing contract sequences. (2) These results assume consumers had perfect information, which is equivalent to the ex post analysis. I also assume no discounting of the future. These assumptions are relaxed in Appendix Tables A.3 to A.6. (3) Bars are not mutually exclusive because contracts may have multiple features and consumers may have been with Retailer A long enough to select multiple successive contracts. (4) The percentage labels on each bar indicate the share of consumers experiencing each contract characteristic.

More results of the baseline model are provided in column 1 of Table 1.3. This table provides my headline finding that the mean consumer could have saved \$33 per month on the cost-minimizing contract sequence, which is 38 percent of the median consumer's total bill and 65 percent of the retailer portion (compare to Table 1.2). While row 1 provides the mean consumer's mean monthly savings, row 2 provides the mean consumer's savings over the entire period at a monthly rate, which turns out to be an almost identical value. Row 3 shows that 75 percent of all consumers' observed invoice periods would have been strictly dominated by contracts on the cost-minimizing contract path. Row 4 shows that 32 percent of consumers have all of their invoices strictly dominated. These estimates are very conservative, since I only consider a contract strictly dominated in a particular month if the observed monthly bill is more than \$5 more expensive than the model's estimate of the cost-minimizing bill. I use this conservative definition to account for idiosyncrasies in rounding and pro-rating billing periods that lasted less than one full month.

Columns 2 and 3 of Table 1.3 confute the hypotheses that consumer behavior is rationalized by either discounting or uncertainty about future prices and contract offerings. In particular, column 2 estimates my model with a discount factor of $\beta = 0.95$ and yields almost identical results to the baseline case with no discounting. Column 3 provides evidence against the explanation of consumer uncertainty. This set of results relaxes Assumption 2 that consumers have perfect information about future prices and contract offerings. Instead, consumers choose their contract sequence as if the current prices and contract choice set will persist forever, and they only update their beliefs in the periods that they actually choose to engage in search. This belief-formation process aligns with the conclusions of Anderson et al. (2013), who find that Michigan consumers consistently forecast constant future gasoline prices. Although these estimates require no foresight about the future of the market, consumers perform almost exactly the same against this scenario as they did the perfect information case.

Table 1.3: Missed Savings on Cost-Minimizing Contracts

		(1) Ex post optimal, β=1	(2) Ex post optimal, β=0.95	(3) Imperfect information optimal, β=1
[1]	Mean monthly savings	$33*	$33*	$32*
		($16.47)	($16.47)	($15.98)
[2]	Mean discounted savings (at a monthly rate)	$32	$32	$31
[3]	Share of invoices that are strictly dominated	75%	75%	75%
[4]	Share of customers for whom all invoices are strictly dominated	32%	32%	34%

*Notes: (1) This table shows the potential savings if consumers had chosen their cost-minimizing contract sequences. (2) Column 1 assumes consumers had perfect information, which is equivalent to the ex post analysis. Column 1 also assumes no discounting of the future. (3) Column 2 introduces discounting. (4) Column 3 is an imperfect information case where consumers choose cost-minimizing contracts believing that their choice set will remain the same in the future. (5) Rows 3 and 4 consider a contract strictly dominated in a particular month if the observed monthly bill is more than $5 more expensive than the model's estimate of the cost-minimizing bill. I use this conservative definition to account for idiosyncrasies in rounding and pro-rating billing periods that lasted less than one full month. (6) Standard errors in parentheses. (7) *p* < 0.1*, **p* < 0.05*, ***p* < 0.01*. (8) Means are taken across consumers, not invoices.*

As further evidence against the explanations of discounting and uncertainty, columns 2 and 3 of Table 1.4 subset the data to only consumers' first and second invoices, respectively. These are the first and second invoices *after* an observed contract selection, so one consumer may have multiple first and second invoices. The similarity of the missed savings among first and second invoices to the entire sample shows that consumers are missing near-term cost-savings just as much as they are missing cost-savings in later periods. First, this provides additional evidence that consumer discounting is not rationalizing missed savings. Second, the failure to cost-minimize among first and second invoices suggests that consumers are not missing savings on account of uncertainty.

Even if consumers have long-term uncertainty, they should have fairly accurate expectations about both their near-term consumption and near-term prices and contract offerings. However, the fact that consumers are missing cost-savings as much in the near-term as in the long-term suggests that (1) uncertainty does not rationalize consumer decisions, and (2) consumers are not trading off near-term costs for future savings. For uncertainty about future consumption to rationalize consumers' failure to cost-minimize, consumers would need to consistently and systematically misestimate their future consumption. I cannot think of a mechanism for this type of bias in consumer expectations, especially given the ever-increasing availability of electricity consumption monitoring and control technologies in the market. Also, as discussed in the Introduction, WP used a different method to draw a similar conclusion in the U.K. electricity retail choice market.

Table 1.4: Missed Savings on Cost-Minimizing Contracts, Subsetting by Consumer Tenure and Invoice Timing

		(1) All invoices	(2) First invoice	(3) Second invoice	(4) Customers with 12 or more invoices	(5) Customers with 3 or more invoices
Panel A: Sample summmary (monthly)						
[1]	Median usage (kWh)	918	925	956	806	924
[2]	Median total bill	$88	$97	$84	$74	$87
[3]	Average price ($/kWh)	$0.104	$0.109	$0.096	$0.106	$0.104
[4]	Number of customers	4,208	4,201	3,901	1,367	3,958
[5]	Number of invoices	41,356	5,458	4,982	25,012	40,948
Panel B: Cost-minimizing contracts						
[6]	Mean monthly savings	$33*	$33**	$26***	$36***	$32**
		($16.47)	($14.76)	($10.37)	($10.36)	($14.97)
[7]	Mean discounted savings (at a monthly rate)	$32	$33	$26	$35	$31
[8]	Share of invoices that are strictly dominated	75%	74%	70%	76%	75%
[9]	Share of customers for whom all invoices are strictly dominated	32%	67%	63%	16%	30%

Notes: (1) This table shows the potential savings if consumers had chosen cost-minimizing contracts in each invoice period, examining various subsets of the data. (2) The table assumes consumers had perfect information, which is equivalent to the ex post analysis. The table also assumes no discounting of the future. (3) Column 1 is the same as Column 1 of Table 1.3. (4) Columns 2 and 3 are the first and second invoices after an observed contract selection, so one consumer may have multiple first and second invoices. (5) Rows 1 through 7 compare the sample characteristics of the subsets. (6) Standard errors in parentheses. (7) $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$. *(8) Also, see notes for Table 1.3.*

Another possible explanation for consumer deviations from the cost-minimizing contracts could be inherent preferences for particular contract features for which consumers are willing to pay a premium. I consider renewable energy content to be the most plausible feature to have this effect. Indeed, comparing columns 1 and 3 to column 2 of Table 1.5 shows that the median green contract consumer pays about three percent more per kWh than their non-green counterparts. However, the mean of $34 per month missed savings for green contract consumers is only slightly higher than the $33 missed savings of non-green consumers, so green preferences clearly do not rationalize the observed selections. Similarly, consumers on time-varying contracts miss out on approximately the same savings as non-time-varying consumers (compare columns 4 and 5), which also belies the hypothesis that consumers who select time-varying contracts are more savvy or attentive.

Column 6 of Table 1.5 shows that there is one group of consumers who perform far worse than the rest. This group is the consumers who got switched to a default contract when their previous contract expired. In fact, the median consumer on a default contract pays 40 percent more per kWh. Evidently, being switched to a default contract is very costly. However, column 7 shows that consumers on other month-to-month contracts miss out on less savings than consumers at large. This finding is consistent with Figure 1.3, which shows that choosing at least one month-to-month contract would have been the cost-minimizing strategy for 63 percent of consumers. However, as noted in the discussion of that figure, this strategy requires frequent search and a high degree of attention to avoid being switched onto a default contract at the end of the month.

Table 1.5: Missed Savings on Cost-Minimizing Contracts, Additional Subsets

		(1) All invoices	(2) Green	(3) Not green	(4) Time-varying	(5) Not time-varying	(6) Default	(7) Other month-to-month
Sample summmary (monthly)								
[1]	Median usage (kWh)	918	993	882	959	906	782	568
	Median monthly bill							
[2]	Median total bill	$88	$92	$86	$95	$87	$109	$64
[3]	Average price ($/kWh)	$0.104	$0.104	$0.104	$0.101	$0.104	$0.142	$0.110
[4]	Number of customers	4,208	1,052	3,308	328	4,070	1,230	491
[5]	Number of invoices	41,356	8,683	32,673	2,057	39,299	6,133	722
[6]	Mean monthly savings	$33*	$34*	$33**	$34***	$33*	$68***	$19
[7]		($16.47)	($19.77)	($15.31)	($10.18)	($16.77)	($24.65)	($10.22)
Cost-minimizing contracts								
[8]	Mean discounted savings (at a monthly rate)	$32	$33	$32	$34	$32	$69	$19
	Share of invoices that are strictly dominated	75%	67%	77%	91%	74%	95%	68%
[9]	Share of customers for whom all invoices are strictly dominated	32%	26%	35%	69%	31%	85%	65%

*Notes: (1) Standard errors in parentheses. (2) *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. (3) See notes for Table 1.4.*

Since my model treats consumers as cost-minimizers as opposed to utility-maximizers, a consumer's parameter of risk aversion is another omitted explanation that could rationalize consumer decisions. However, my data captures over four years of market evolution and various consumption and price shocks, including the wholesale price hitting the market cap of $9 per kWh in August 2018. Thus, to the extent that this data captures the true distribution of events consumers would might wish to insure against, failure to cost-minimize might be evidence of a misperception of risk, but it could not be rationalized by risk aversion itself. Over the true distribution of possible events, risk aversion can rationalize different contract preferences but not a failure to cost-minimize. Following this logic, columns 4 and 5 of Table 1.4 restrict the analysis to consumers with longer data series, and once again, I find similar results to consumers at large. This finding also suggests that the results are not overly influenced by idiosyncratic consumers who either entered the sample late or switched away from Retailer A very soon after signing up. In the next iteration of this paper, I will further explore the risk aversion hypothesis by restricting consumer choice sets to contracts of the same duration as their observed selections.

Although the alternative estimates of Tables 1.3 to 1.5 show generally similar results to the baseline model, it is possible that the distribution of cost-minimizing contract features shown in Figure 1.3 is more sensitive. However, I show that this is not the case in Appendix Tables A.3 to A.6. In particular, Appendix Table A.3 is the tabular version of Figure 1.3, and Appendix Tables A.4 to A.6 provide the contract feature comparison for various sensitivities.

1.7 Conclusion

I have presented evidence contrary to a series of explanations for consumers' failure to cost-minimize, including (1) discounting; (2) uncertainty about future consumption, prices, and contract features; (3) green preferences; and (4) risk aversion. Thus, I conclude that consumers' deviations from the cost-minimizing contract choices are most likely the result of a combination of search costs, inattention, and mis-selling. In this context, inattention includes being placed on a default contract after neglecting to select an alternative choice at the end of a contract term; failure to consider all contracts among the large choice set; and true mistakes. Recall, though, that there are three caveats, which are more fully discussed in Section 1.3. First, my model does not account for the price elasticity of demand, which may lead to estimates that understate the welfare gains from switching contracts. Second, in estimating my model over a single retailer's contracts, I am able to eliminate the explanation of brand advantage, but my results likely understate the potential savings of alternative contracts. Lastly, the model only estimates partial equilibrium results, as I do not consider the retailer response to consumer switching.

Consumers' widespread failure to choose cost-minimizing contracts suggests that policymakers could improve welfare with interventions that reduce search costs, inattention, and mis-selling. Some options include removing the legal obstacles to concierge services or introducing a web-based tool to find consumers' cost-minimizing contract based on their consumption history.[30] My findings also

[30]For example, the PUCT could add this functionality to their website, *powertochoose.com*.

suggest that these interventions could lead to higher adoption of time-varying rates, which could lead to more efficient allocation of grid resources and lower emissions levels. The severe consequences of inattention and search frictions in this market design appeared in sharp relief during the winter 2021 Texas blackouts in which residential consumers on real-time price contracts faced the market cap price of $9 per kWh for multiple days. In one widely reported incident, a consumer ended the month with a $16,752 bill.[31] To avoid financial disaster, a consumer on this contract would have had to either curtail their electricity consumption or quickly switch to an alternative retailer and contract.

My other main finding is that consumers are constrained in the monopoly setting from expressing their heterogeneous preferences for contract variety. This insight may guide regulators in monopoly settings to consider increasing variety. However, this policy would be best complemented by an intervention to reduce search costs because increased variety may actually be welfare-decreasing if the levels of search costs and inattention are too high. In future work, I plan to model the supply-side retailer decisions as well as estimate the equilibrium response to consumers switching to cost-minimizing contracts. Through this work, I hope to better explain the high level of retailer entry and extreme variety of contract offerings that have proliferated in the Texas market since 2010.

1.8 References

Abaluck, Jason and Jonathan Gruber (2011) "Choice inconsistencies among the elderly: Evidence from plan choice in the Medicare Part D program," *American Economic Review*, 101 (4), 1180–1210, 10.1257/aer.101.4.1180.

Anderson, Soren T., Ryan Kellogg, and James M. Sallee (2013) "What do consumers believe about future gasoline prices?," *Journal of Environmental Economics and Management*, 66 (3), 383–403, 10.1016/j.jeem.2013.07.002.

[31] McDonnell Nieto del Rio, Giulia, Nicholas Bogel-Burroughs, and Ivan Penn. "His Lights Stayed on During Texas' Storm. Now He Owes $16,752." The New York Times. February 20, 2021. https://www.nytimes.com/2021/02/20/us/texas-storm-electric-bills.html

Berry, Steven, James Levinsohn, and Ariel Pakes (1995) "Automobile prices in market equilibrium," *Econometrica*, 10.2307/2171802.

Borenstein, Severin and Stephen Holland (2005) "On the efficiency of competitive electricity markets with time-invariant retail prices," *The RAND Journal of Economics*, 36 (3), 469–493, Article.

Giulietti, Monica, Catherine Waddams Price, and Michael Waterson (2005) "Consumer choice and competition policy: A study of UK energy markets," *Economic Journal*, 115 (506), 949–968, 10.1111/j.1468-0297.2005.01026.x.

Giulietti, Monica, Michael Waterson, and Matthijs Wildenbeest (2014) "Estimation of search frictions in the british electricity market," *Journal of Industrial Economics*, 62 (4), 555–590, 10.1111/joie.12062.

Gruber, Harald and Frank Verboven (2001) "The evolution of markets under entry and standards regulation - The case of global mobile telecommunications," *International Journal of Industrial Organization*, 19 (7), 1189–1212, 10.1016/S0167-7187(01)00069-8.

Hortaçsu, Ali, Seyed Ali Madanizadeh, and Steven L. Puller (2017) "Power to choose? An analysis of consumer inertia in the residential electricity market," *American Economic Journal: Economic Policy*, 9 (4), 192–226, 10.1257/pol.20150235.

Hotelling, Harold (1929) "Stability in competition," *The Economic Journal*, 39 (153), 41–57, 10.2307/2224214.

Lancaster, Kelvin (1979) *Variety, equity, and efficiency*: Columbia University Press.

——— (1980) "Competition and product variety," *The Journal of Business*, 53 (3), 79–103, 10.1086/296102.

Miravete, Eugenio J. (2003) "Choosing the wrong calling plan? Ignorance and learning," *American Economic Review*, 93 (1), 297–310, 10.1257/000282803321455304.

Nevo, Aviv, John L. Turner, and Jonathan W. Williams (2016) "Usage-based pricing and demand for residential broadband," *Econometrica*, 84 (2), 411–443, 10.3982/ecta11927.

Puller, Steven L. and Jeremy West (2013) "Efficient retail pricing in electricity and natural gas markets," *American Economic Review: Papers and Proceedings*, 103 (3), 350–355, 10.1257/aer.103.3.350.

Spence, Michael (1976) "Product differentiation and welfare," *American Economic Review*, 66 (2), 407–414.

Wilson, Chris M. and Catherine Waddams Price (2010) "Do consumers switch to the best supplier?," *Oxford Economic Papers*, 62 (4), 647–668, 10.1093/oep/gpq006.

Chapter 2: Estimating a Nonlinear Model of Climate and Weather

with Robert Mendelsohn and Paula Pereda[1]

Abstract– The standard nonlinear model of climate and weather is misspecified. If weather and climate have nonlinear effects on economic outcomes, the functional form must include an interaction between weather and climate. With this interaction, the nonlinear effect of climate and weather can be identified. We reproduce the results from two prominent studies, Deschênes and Greenstone (2007, 2012) and Burke et al. (2015), and show that the nonlinear climate response is quite different from the nonlinear weather response in both papers. We also show that fixed effects must be applied to nonlinear variables before they are transformed, not afterwards as in the literature.

2.1 Introduction

The idea that both weather and climate have a nonlinear effect on economic outcomes is well known in the economics literature. There is ample evidence from weather panel studies that weather has a concave effect on agricultural outcomes (see review by Blanc and Schlenker 2017) and other economic outcomes (see review by Kolstad and Moore 2020). There is also evidence from cross-sectional research that climate has a concave effect on agricultural outcomes (see review by Mendelsohn and Massetti 2017) and also energy demand (Mansur and Morrison 2008). Finally, there is evidence from process-based and simulation models that weather and climate have nonlinear effects on out-

[1]**Acknowledgments and disclosures:** Authors in alphabetical order. Robert Mendelsohn is a Professor at the Yale School of Forestry and Environmental Studies, and Paula Pereda is a Professor at the Department of Economics, University of São Paulo. The authors declare that they have no conflict of interest. We are grateful to the São Paulo Research Foundation (FAPESP BPE grant 2018/02081-1) for funding Dr. Pereda's visit to Yale University.

comes in the agriculture, energy, and water sectors (Mendelsohn and Schlesinger 1999; Mendelsohn and Neumann 2004; Smith and Mendelsohn 2007). Although linear models can estimate marginal changes in climate or weather, forecasting the impacts of climate change requires a nonlinear model.

The other important insight from the literature is that the nonlinear effect of climate (the long-run expected weather) and the nonlinear effect of weather deviations (random surprises) on economic outcomes are not identical. Economic systems adapt to long-run climate, but each weather deviation remains a short-run surprise. For example, farmers grow very different crops in different climate zones, air conditioning is much more widespread in warmer climates, and water saving devices are more prevalent in arid regions. The nonlinear economic responses to climate and to weather are different. This is a common result in economic modeling where the short-run responses to income and prices are more inelastic than the long-run responses. Comparisons of results from long-run cross sectional and short-run weather panel studies suggest that the long-run response is less concave than the short-run response (Mendelsohn and Schlesinger 1999; Mendelsohn and Dinar 2009; Moore and Lobell 2014). This distinction between a forecasted outcome and a surprise is also exploited by several papers that try to use weather forecasts to distinguish between the dynamic response to weather shocks versus weather forecasts (Deryugina and Hsiang 2017; Severen et al. 2018; Lemoine 2020).

However, the standard nonlinear weather panel model in the literature starting with the pathbreaking article by Deschênes and Greenstone (2007), argued that economic outcomes depend simply on observed weather. Observed weather is the sum of the expected weather and random weather deviations. By assuming economic outcomes are a nonlinear function of observed weather, the model effectively assumes that weather deviations and expected weather are interchangeable and have identical effects. However, there is ample evidence that weather deviations and climate are not interchangeable.[2] The economic response to climate is different from the economic response to weather deviations (shocks). As demonstrated in this paper, the nonlinear functional form in the

[2]Technically, the climate is the probability distribution of weather outcomes. In this paper, we are only focusing on the first moment of this distribution by defining climate as the expected value of weather, which we measure as the long-run mean. The variance and other moments of the weather distribution are also technically the climate. However, there are fewer adjustments that firms and households can make to these higher order moments, so we ignore them in this paper and leave it for future research to explore their effect.

standard literature based on observed weather leads to biased estimates of the effect of weather. Although we do not examine a bin model in this paper, the specification issues raised in this paper apply to bin models as well.

This paper develops a theoretical model of the effect of climate deviations on economic outcomes. The model distinguishes the effect of climate from the nonlinear effect of weather deviations. Starting with a climate response function and a nonlinear weather deviation response, we deduce what the weather-climate model must look like. We build on the idea that the marginal effect of climate and weather should be the same when the weather happens to be equal to the climate (because the ideal level of long-run adjustment is present) (Hsiang 2016; Mendelsohn 2016; Mérel and Gammans 2021). However, the effect of weather and climate will not be the same when the weather deviates from the climate. In this case, the long-run adjustment is different from the response to the weather.

The theoretical model suggests that if the climate function is linear, the effect of weather depends just on weather deviations. The nonlinear model of observed weather in the literature is a function of the sum of climate and the weather deviation. As we show in the theory section, such a model is possible only if it is linear. The standard model is misspecified.

However, if the climate function is nonlinear, the weather function must include an interaction term between weather deviations and the marginal effect of climate. The effect of weather deviations will vary systematically across sites. Failure to capture this interaction effect with climate will lead to a biased estimate of the effect of local weather at each site.

With both the linear and nonlinear climate model, one can estimate both the marginal effect of climate and the marginal effect of weather. In the linear model, the marginal effect of climate is a constant estimated by the linear weather term. In the nonlinear climate model, the marginal effect of climate is determined by the linear weather term and the interaction term between weather and climate. The intertemporal within-group curvature of the weather versus the between-group curvature of the climate effectively identifies the difference between the effect of climate and the effect of weather. The two effects can be estimated and compared using a single weather panel regression.

The idea that weather and climate might interact has been explored in the literature but not in a consistent theoretical framework. Several weather panel studies have tried to estimate weather-climate interactions with a two stage procedure that first regresses intertemporal impacts on weather at each site. This linear regression leads to an average weather effect at each site. A second-stage regression is then estimated across sites that includes an interaction term between this average weather effect and climate (see reviews by Blanc and Schlenker 2017 and Kolstad and Moore 2020). The literature has often found a significant interaction effect. But the average weather effect at each site is not a good measure of the nonlinear effect of weather deviations.

In the process of studying the nonlinear effect of weather deviations, we have rediscovered a problem first noted by McIntosh and Schlenker (2006). The fact that it is difficult to use fixed effects with a nonlinear regression is well known (Neyman and Scott 1948, and Greene 2003).[3] However, McIntosh and Schlenker first described the problem of nonlinear variables with fixed effects in the nonlinear weather literature. Despite the warnings, however, the nonlinear weather literature has not addressed this estimation problem. The current estimation used in the nonlinear weather literature raises observed variables to higher orders and then conditions the transformed variables using fixed effects. This leads to biased coefficients. The correct approach is to condition the underlying variables first using fixed effects and then apply any nonlinear transformations.

A critical portion of our paper reproduces the proof of McIntosh and Schlenker (2006) showing the problematic constraints and implicit assumptions of a nonlinear weather model with fixed effects. However, McIntosh and Schlenker and their successors Mérel and Gammans (2021) propose an alternative DGP that maintains the same form of the problematic fixed effects specification with the addition of a so-called "climate penalty" term. We argue that this climate penalty term does not solve the problem that the authors have identified, leading us to propose an alternative specification with a weather-climate interaction.

We take two prominent examples from the literature to demonstrate how to correctly specify a non-

[3]Now known as the incidental parameters problem, the challenge of using fixed effects in models with nonlinear variables was first discussed in Neyman and Scott 1948. Discussion of the incidental parameters problem typically focuses on binary regression models, including logit and probit.

linear weather panel model. We reproduce the Deschênes and Greenstone (2007, 2012)[4] (hereafter D&G) analysis of farm net revenue in United States counties and the Burke et al. (2015) (hereafter BHM) analysis of GDP per capita in countries around the world. We first assume that climate has a linear effect and demonstrate that the reported weather coefficients in both papers are biased because of both a specification error and an estimation error. We then assume that climate has a quadratic effect and compare the resulting nonlinear weather effect with the nonlinear climate effect. We find that the weather effect and the climate effect are quite different in both papers. In particular, the climate effect is less concave than the weather effect. For example, in a poor hot country, a temporary warming of $3°C$ would reduce annual GDP per capita by 4 percent, while a permanent warming of $3°C$ would reduce annual GDP per capita by 3 percent.

2.2 Theory

This analysis addresses how to model the nonlinear effect of weather variations on economic outcomes across time given the effect of climate on expected economic outcomes. We adopt the following nomenclature throughout for each variable, V. The observed value of each variable, $V_{i,t}$, in location i at time t is the expected value of that variable, $\bar{V}_i$, plus an intertemporal deviation, $v_{i,t}$:

$$V_{i,t} \coloneqq \bar{V}_i + v_{i,t} \quad \text{where the expected value of variable } V \text{ at location } i \text{ is} \quad \bar{V}_i \coloneqq E_t[V_{i,t}|\gamma_i] \qquad (2.1)$$

where γ_i are the location-specific characteristics of location i. In the empirical applications of this paper, we generally define the expected value as the observed mean of the variable over time such that:

$$\bar{V}_i \coloneqq E_t[V_{i,t}|\gamma_i] \coloneqq \sum_{t=0}^{T} V_{i,t}/T. \qquad (2.1a)$$

In particular, we assume that observed weather, $W_{i,t}$, at each location i and time t is a random draw from the climate probability distribution at that location. Following equation (2.1), we define the climate, $\bar{W}_i$, as the expected value of weather, which we approximate as the long-run mean of

[4]The fixed effect nonlinear weather model was first developed in Deschênes and Greenstone (2007) but the final empirical results are in Deschênes and Greenstone (2012).

weather measured as $\bar{W}_i := \sum_{t=0}^{T} W_{i,t}/T$.[5] The observed weather, $W_{i,t}$, is the sum of climate, $\bar{W}_i$, plus a weather deviation, $w_{i,t}$:

$$W_{i,t} = \bar{W}_i + w_{i,t}.$$

Note that when the observed weather is equal to the climate, the weather deviation is zero.

Using past observed values of weather, economic decision makers have a reasonable measure of expected near-term climate.[6] Both firms and households can use that information to adjust capital to climate and improve their long-run performance. Firms will adjust capital to maximize expected profits, $\bar{\pi}_i$, and households will adjust capital to maximize their utility. For example, given the price of outputs, P_Q, inputs, P_X, and capital, P_K, a firm chooses a mean level of inputs, $\bar{X}_i$, and capital, $\bar{K}_i$, given the climate, $\bar{W}_i$, at their location i to produce output, $\bar{Q}_i$, that maximizes the expected profit at that location:

$$\max_{\bar{X}_i, \bar{K}_i} E[\bar{\pi}_{i,t}] = P_Q \bar{Q}_i(\bar{X}_i, \bar{K}_i, \bar{W}_i) - P_X \bar{X}_i - P_K \bar{K}_i. \tag{2.2}$$

The first order conditions for equation (2.2) are that marginal revenue is equated with marginal cost for inputs and capital:

$$P_Q \frac{d\bar{Q}_i(\bar{X}_i, \bar{K}_i, \bar{W}_i)}{d\bar{X}_i} = P_X \tag{2.2a}$$

$$P_Q \frac{d\bar{Q}_i(\bar{X}_i, \bar{K}_i, \bar{W}_i)}{d\bar{K}_i} = P_K. \tag{2.2b}$$

The choice of both inputs and capital are a function of the underlying climate (expected weather, $\bar{W}_i$) at each location. In contrast, in the short-run, as weather varies, firms cannot adjust capital and may even struggle to adjust inputs. The response to climate is a long-run adjustment where both inputs and capital can be adjusted whereas the response to weather is a short-run adjustment

[5] We are simplifying the analysis by using the observed mean over the sample period as a predictor of the expected value. Technically, we should use the observed mean over the immediate past. If the underlying climate was rapidly changing, this could be important.

[6] The private sector will engage in climate adaptations that have a positive net present value over the lifetime of the adjustment. We argue that climate change is slow enough that the climate from recent weather is a good forecast of the climate over the lifetime of an adaptation. A precise analysis should use past weather to estimate climate, not the mean observed weather during the panel. We are assuming that the decision maker can adjust to the expected value of weather. It may also be possible to adjust to the variance of weather, though we leave that possibility for future research.

where capital is fixed and even inputs may only partially adjust. Economic theory and empirical evidence suggests that short-run and long-run supply functions are quite different. There is every reason to believe that the observed response to climate and weather will be different. A parallel argument can be made on the consumer side about short-run and long-run demand.

The theory above suggests that long-run (mean) economic outcomes are a nonlinear concave function of expected climate (Mansur and Morrison 2008, and Mendelsohn and Massetti 2017):

$$\bar{Y}_i = f(\bar{W}_i). \tag{2.3}$$

Empirical evidence from laboratory studies and economic analyses all imply that long-run productivity is a concave nonlinear function of both temperature and precipitation (e.g. Ritchie and Nesmith 1991; Mendelsohn et al. 1994; and Mendelsohn and Schlesinger 1999). Casual observation also suggests that long-run economic productivity tends to be low at very cold temperatures and low precipitation (e.g. above 60 degrees latitude) and very hot places (e.g. Death Valley) and very wet places (moist tropical forests). Between these extremes, productivity is much higher.

It is also true that random fluctuations in weather have a nonlinear effect on short-run economic outcomes. This is what the authors intended to test in the nonlinear weather studies (D&G; BHM; and Schlenker and Roberts 2009). We adopt the general principle in these nonlinear weather panel studies that short-run variations in economic outcomes are a nonlinear function of weather deviations:

$$y_{i,t} = g(w_{i,t}). \tag{2.4}$$

The observed economic outcomes, $Y_{i,t}$, are the sum of the effect of climate, $\bar{W}_i$, and weather deviations, $w_{i,t}$:

$$Y_{i,t} = \bar{Y}_i + y_{i,t} = f(\bar{W}_i) + g(w_{i,t}). \tag{2.5}$$

One interesting feature of this model of long-run adaptation is that competitive firms will choose the profit-maximizing level of capital and inputs given the expected climate. Similarly, households will choose the utility-maximizing bundle of consumption given the expected climate. When the weather deviation is zero, the marginal effect of short-run weather and the marginal effect of expected climate are identical (Hsiang 2016; Mendelsohn 2016; Mérel and Gammans 2021). The marginal short-run and long-run effects are identical because the chosen capital and inputs (or goods) happen to be ideal when the observed weather are equal to the expected climate:

$$g'(0) = f'(\bar{W}_i). \tag{2.6}$$

We follow the weather literature and assume that $g'(w_{i,t})$ is linear in $w_{i,t}$, which implies that $g(w_{i,t})$ is quadratic in $w_{i,t}$. Without loss of generality, we substitute:

$$g'(w_{i,t}) = g'(0) + 2\beta_3 w_{i,t}. \tag{2.7}$$

Substituting equation (2.6) into (2.7) yields:

$$g'(w_{i,t}) = f'(\bar{W}_t) + 2\beta_3 w_{i,t}. \tag{2.8}$$

Integrating this expression in terms of $w_{i,t}$ yields:

$$g(w_{i,t}) = \beta_0 + f'(\bar{W}_t) w_{i,t} + \beta_3 w_{i,t}^2. \tag{2.9}$$

In general, the effect of weather depends on an interaction term between the weather deviation and the partial derivative of climate. It also includes the weather deviation squared. If climate is nonlinear, the effect of weather deviations will depend on the climate, not just the weather deviation. Substituting this back into equation (2.5) yields:

$$Y_{it} = \bar{Y}_i + y_{i,t} = \beta_0 + f(\bar{W}_i) + f'(\bar{W}_i) w_{i,t} + \beta_3 w_{i,t}^2. \tag{2.10}$$

In the spirit of fixed effects, we demean each variable of equation (2.10) by its location-specific mean, thus yielding the following within-group regression:[7]

$$y_{i,t} = f'(\bar{W}_i)w_{i,t} + \beta_3 w_{i,t}^2. \tag{2.11}$$

Equation (2.11) is the general formula that captures both the nonlinearity of climate and the nonlinearity of weather deviations. As we show below for both a linear and nonlinear climate function, equation (2.11) cannot become the standard nonlinear weather model in the literature.

2.2.1 Linear Climate Example

If the climate function, $f(\bar{W}_i)$, is linear then $\bar{Y}_i = \alpha_0 + \alpha_1 \bar{W}_i$. In this case, $f'(\bar{W}_i) = \alpha_1 = \beta_1$ because $\beta_1 = g'(0)$. The marginal effect of climate is a constant. Substituting into equation (2.11) yields:

$$y_{i,t} = \beta_1 w_{i,t} + \beta_3 w_{i,t}^2. \tag{2.12}$$

Economic outcomes depend strictly on weather deviations and they are independent of climate. So this model cannot be estimated as a function of observed weather since that embodies climate.

With a linear climate model, the marginal climate effect is a constant measured by $\hat{\beta}_1$. The linear effect of climate can be forecasted with the linear coefficient β_1. The nonlinear effect of weather deviations can be forecasted with the linear and squared terms in equation (2.12). The climate effect and the effect of weather deviations are different.

2.2.2 Quadratic Climate Example

Another useful special case to examine is when long-run economic outcomes are a quadratic function of climate:

$$\bar{Y}_i = \alpha_0 + \alpha_1 \bar{W}_i + \alpha_2 \bar{W}_i^2. \tag{2.13}$$

[7]Due to the nonlinear term $\beta_3 w_{i,t}^2$, we do not derive equation (2.11) by fixed effects estimation of equation (2.10). Instead, we remove the variable means *before* applying the nonlinear transformation. That is, $w_{i,t}^2 - \frac{\sum_{t=0}^T w_{i,t}^2}{T}$ may not equal $\left(w_{i,t} - \frac{\sum_{t=0}^T w_{i,t}}{T}\right)^2 = w_{i,t}^2$. We explain the importance of this distinction in Section 2.2.3.

The partial derivative of equation (2.13) is:

$$f'(\bar{W}_i) = \alpha_1 + 2\alpha_2 \bar{W}_i. \tag{2.14}$$

Substituting equation (2.14) into (2.11) yields:

$$y_{i,t} = \beta_1 w_{i,t} + \beta_2 \bar{W}_i w_{i,t} + \beta_3 w_{i,t}^2. \tag{2.15}$$

In this quadratic climate model, $\beta_1 = \alpha_1$, and $\beta_2 = 2\alpha_2$. The marginal effect of climate can be estimated from $\beta_1 + \beta_2 \bar{W}_i$. The marginal effect of weather is $\beta_1 + \beta_2 \bar{W}_i + 2\beta_3 w_{it}$. The effect of weather can be differentiated from the effect of climate. Note that the marginal effect of weather and climate are never $\beta_1 + 2\beta_3(\bar{W}_i + w_{it})$ as assumed in the standard nonlinear weather panel literature. β_2 can never be equal to $2\beta_3$, unless $\beta_2 = 2\beta_3 = 0$, the models are all linear. The estimated model reveals both the marginal effect of weather and the marginal effect of climate.

2.2.3 Empirical Methodology

Although the problem of using fixed effects with a nonlinear independent variable was first described in McIntosh and Schlenker, this problem persists in the nonlinear weather literature. Thus, we provide a simple solution to estimate fixed effects with nonlinear independent variables.

Let us use the nonlinear weather model, such as estimated in D&G and BHM to illustrate the estimation error:

$$Y_{i,t} = \beta_1 W_{i,t} + \beta_2 W_{i,t}^2 + \boldsymbol{X'_{i,t}\varphi} + \gamma_i + \varepsilon_{i,t}. \tag{2.16}$$

One interpretation of this literature is that they intended to assume climate has a linear effect on economic outcomes leading to a correctly specified nonlinear weather model that is a quadratic function of the weather deviations (equation 2.12). But the authors estimated equation (2.16), not equation (2.12).

We now reproduce the proof of McIntosh and Schlenker (2006). Rewriting the terms in equation

(2.16) as means and deviations yields:

$$Y_{i,t} = \bar{Y}_i + y_{i,t} = \beta_1(\bar{W}_i + w_{i,t}) + \beta_2(\bar{W}_i + w_{i,t})^2 + (\bar{\boldsymbol{X}}'_{i,t} + \boldsymbol{x}'_{i,t})\boldsymbol{\varphi} + \gamma_i + \varepsilon_{i,t}. \tag{2.16a}$$

Using a standard fixed effects approach, the above equation first transforms each nonlinear variable to a higher power and then conditions the transformed variables on the fixed effects leading to the following model:

$$y_{i,t} = \beta_1 w_{i,t} + \beta_2\big[(\bar{W}_i + w_{i,t})^2 - E_t[(\bar{W}_i + w_{i,t})^2|\gamma_i]\big] + \boldsymbol{x}'_{i,t}\boldsymbol{\varphi} + \varepsilon_{i,t}. \tag{2.16b}$$

Expanding the expression in brackets for the β_2 coefficient:

$$\begin{aligned}
&\big[(\bar{W}_i + w_{i,t})^2 - E_t[(\bar{W}_i + w_{i,t})^2|\gamma_i]\big] \\
&= \bar{W}_i^2 + 2\bar{W}_i w_{i,t} + w_{i,t}^2 - E_t[\bar{W}_i^2 + 2\bar{W}_i w_{i,t} + w_{i,t}^2|\gamma_i] \\
&= 2\bar{W}_i w_{i,t} + w_{i,t}^2 - E_t[w_{i,t}^2|\gamma_i].
\end{aligned} \tag{2.17}$$

By conditioning the transformed variables on the fixed effects, the literature has inadvertently turned the quadratic term in equation (2.17) into three terms: $2\bar{W}_i w_{i,t} + w_{i,t}^2 - E_t[w_{i,t}^2|\gamma_i]$ leading to:

$$y_{i,t} = \beta_1 w_{i,t} + \beta_2\big[2\bar{W}_i w_{i,t} + w_{i,t}^2 - E_t[w_{i,t}^2|\gamma_i]\big] + \boldsymbol{x}'_{i,t}\boldsymbol{\varphi} + \varepsilon_{i,t}. \tag{2.18}$$

The two additional terms in equation (2.18) and the restriction that the coefficient must be the same on all three terms lead to biased estimates of the weather parameters in equation (2.16). Note that one can reproduce the reported coefficient, $\hat{\beta}_2$, in D&G or BHM by estimating equation (2.16) using fixed effects or by estimating equation (2.18) with OLS.

McIntosh and Schlenker and their successors Mérel and Gammans (2021) propose that the solution to the restrictions of equation (2.18) is to introduce a climate penalty term in equation (2.16) of the form $\beta_3(W_{i,t} - \bar{W}_i)^2$. However, this approach is flawed because it still preserves the problematic second term of equation (2.18).

We claim that the correct approach to estimating nonlinear variables in (2.16) with fixed effects follows two steps. First, condition each of the original variables, $V_{i,t}$, on its fixed effects, $\phi_{i,t}$. Thus, each variable is turned into a deviation, $v_{i,t}$, from its expected value, $\bar{V_i} = E_t[V_{i,t}|\phi_{i,t}]$. Second, transform the conditioned variables, $v_{i,t}$ (e.g., raise to higher powers). The resulting equation is:

$$y_{i,t} = \beta_1 w_{i,t} + \beta_2 w_{i,t}^2 + \boldsymbol{x}'_{i,t}\boldsymbol{\varphi} + e_{i,t}. \tag{2.19}$$

Conditioning the variables on the fixed effects is relatively straightforward. The conditioned variable is simply: $v_{i,t} = V_{i,t} - \bar{V_i}$ where $\bar{V_i} = E_t[V_{i,t}|\phi_{i,t}]$ can be estimated by an OLS regression of the variable $V_{i,t}$ on the fixed effects $\phi_{i,t}$. A weather-climate interaction can then be explicitly added as in equation (2.15).

There is one limitation in our approach, which requires that panel observations not move between classifications over time (i.e., a country cannot have its status change from poor to rich, and a county cannot change from rainfed to irrigated). A similar restriction applies to time. One cannot use a fixed effect such as modern time that is defined to be different time periods across observations. Introducing endogenous fixed effects also biases the coefficients. See Appendix D for more detail.

2.2.4 Empirical Application

We start the empirical application by replicating two prominent studies of nonlinear weather effects: D&G and BHM. These papers estimate a nonlinear within-group intertemporal weather effect with no interaction terms with climate. The underlying models that are estimated in D&G and BHM resemble equation (2.16) with quadratic temperature and precipitation terms (that are a function of the sum of climate and weather deviations, $W_{i,t}$), other controls, and fixed effects. We replicate the results in the two papers with the same data[8], a spatial fixed effect, the same variables, the same weighting schemes (when applicable), and the same clustering methods as the authors.

D&G estimate agricultural profits per farmland acre, $Y_{i,t}$, in county i in year t:

$$Y_{i,t} = \beta_0 + \gamma_i + \delta_t + \lambda_i D_i + \boldsymbol{X}'_{i,t}\boldsymbol{\alpha} + \Sigma_{j=1}^2[\boldsymbol{\beta}_{1,j} D_i \boldsymbol{W}_{j,i,t} + \boldsymbol{\beta}_{2,j} D_i \boldsymbol{W}_{j,i,t}^2] + \varepsilon_{i,t} \tag{2.20}$$

[8]See Appendix B for more detail on the D&G and BHM data.

where β_0, λ, $\boldsymbol{\alpha}$, and $\boldsymbol{\beta}$ are estimated parameters; D_i is an indicator for whether a county is rainfed or irrigated; $\boldsymbol{X}_{i,t}$ is a vector of other determinants of agricultural production including soil quality; γ_i is a county fixed effect; δ_t is a year fixed effect; $\boldsymbol{W}_{i,t}$ includes the two weather variables, growing season degree days (GDD) and growing season precipitation, indexed by j; and $\varepsilon_{i,t}$ is an error term. An interaction between weather, $\boldsymbol{W}_{i,t}$, and irrigation status, D_i, allows D&G to estimate separate weather effects for rainfed and irrigated counties. Note that unlike equation (2.16), which includes only a location fixed effect, equation (2.20) includes two fixed effects— county, γ_i, and year, δ_t. We discuss the challenges of multiple fixed effects in Appendix C. We do not complicate the analysis by reproducing the sensitivity analysis in D&G around which fixed effects to include or how to cluster standard errors, as these variations do not change the underlying issues. Equation (2.20) makes one substantive change from D&G by changing the definition of irrigation to make it exogenous. D&G allowed the dummy for irrigation to change over time in some observations. We have changed this assumption so that irrigation status, D_i, is determined at the beginning of the sample (year 1987).[9] [10]

BHM estimate the change in GDP per capita, $Y_{i,t}$, in country i in year t according to:

$$Y_{i,t} = \beta_0 + \gamma_i + \delta_t + \Sigma_{j=1}^{2}[\boldsymbol{\beta}_{1,j} Poor_i \boldsymbol{W}_{j,i,t} + \boldsymbol{\beta}_{2,j} Poor_i \boldsymbol{W}_{j,i,t}^2] + \lambda_i t + \alpha_i t^2 + \varepsilon_{i,t} \tag{2.21}$$

where $\boldsymbol{\beta}$ are parameters of weather; γ_i is a country fixed effect; δ_t is a year fixed effect; $Poor_i$ is an indicator for whether or not a country is below the global median GDP per capita; $\boldsymbol{W}_{i,t}$ includes the two weather variables, annual country temperature and precipitation, indexed by j; $\lambda_{i,t}$ and σ_{it}^2 are country-specific trends (levels and quadratic); and $\varepsilon_{i,t}$ is an error term. An interaction between weather, $\boldsymbol{W}_{i,t}$, and income status, $Poor_i$, allows BHM to estimate separate effects for high-income and low-income countries. The regression is unweighted and the standard errors are clustered by country.

To investigate the estimation problem with nonlinear fixed effects, we begin by replicating D&G and BHM using standard fixed effects software. This transforms the variables first and then conditions

[9] See Appendix D for more detail.

[10] D&G consider a county irrigated if 10 percent or more of its farmland is irrigated.

the transformed variables with fixed effects. In Appendix E, we show that the resulting coefficients from the available software are identical to the coefficients from estimating equation (2.18) using the three terms explicitly. We then compare these results with a correctly specified and estimated model. We start with the implicit assumption in the literature that the climate effect is linear. We estimate equation (2.19). The difference between (2.18) and (2.19) reveals the size of the bias in the literature.

We then estimate a consistent nonlinear climate and weather model using the D&G and BHM data. We assume that the nonlinear climate and weather models are both quadratic leading to the functional form of equation (2.15). This involves including an explicit interaction term between weather deviations and climate. The result reveals the nonlinear effect of climate and the nonlinear effect of weather in a single panel regression. The results can be used to forecast the effect of both a temporary weather surprise or a long-run change in the underlying climate.

2.3 Results

Table 2.1 columns 1 and 2 present the original biased linear and quadratic weather coefficient estimates of equation (2.20) reported by D&G and of equation (2.21) reported in BHM using equations (2.16) or (2.18) to estimate the coefficients. Columns 3 and 4 show the corrected estimates for these two models by first conditioning each variable on fixed effects and then making the nonlinear transformations of equation (2.19). The temperature results are in the upper half of the table, and the precipitation results are below. We find that correcting the bias leads to different results in magnitude, direction (sign), and statistical significance. The coefficients of the quadratic terms change the most, but the coefficients of the linear terms are also affected by the misspecification.

With the BHM data, the linear weather coefficients, which measure marginal climate change, all change sign from positive to negative (except for rich countries). BHM find that the linear temperature coefficients is positive and significant for the entire sample as well as for rich countries. This implies that global warming increases GDP per capita. BHM also find that the quadratic term on temperature is negative and significant, though only at a 10 percent significance level for rich countries. However, the corrected parameters reveal that none of the temperature coefficients in the BHM model are significant. With respect to precipitation, BHM find that the linear precipitation coefficient is positive for poor countries and that the quadratic precipitation coefficient is negative for the entire sample and for poor countries. This implies that rainfall is beneficial to long-run GDP per capita but that rainfall deviations have a hill-shaped effect. When the BHM results are corrected, the linear precipitation term is no longer significant, and the quadratic precipitation coefficient is negative for the entire sample and for poor countries.

The D&G linear temperature coefficients change in magnitude when the coefficients are correctly estimated, but they do not change sign. D&G report that the linear temperature coefficients are negative (for the whole sample and rainfed counties), and all the quadratic temperature coefficients are insignificant. Correcting the estimation, the linear temperature coefficients are similar but with higher significance. The corrected quadratic coefficients are positive and significant for the entire sample and rainfed counties, which suggests that temperature deviations have a U-shaped effect on annual net farm profits. D&G report that the linear precipitation coefficient is negative and the quadratic coefficient is positive for rainfed counties. This implies more long-run rainfall is harmful and that annual rainfall has a U-shaped effect on farm profits. In the corrected results, the linear precipitation coefficient is still negative for rainfed counties but positive with only a 10 percent significance for irrigated counties. The corrected quadratic precipitation coefficient becomes negative for the entire sample and rainfed counties. This implies that rainfall has the expected hill-shaped effect on the annual profits of rainfed farms.

Table 2.1: Biased and Corrected Coefficients of Linear Climate with Quadratic Weather

		[1] Biased Linear Coefficient [Eq. 2.16 or 2.18]	[2] Biased Quadratic Coefficient [Eq. 2.16 or 2.18]	[3] Correct Linear Coefficient [Eq. 2.19]	[4] Correct Quadratic Coefficient [Eq. 2.19]
Temperature					
BHM[a] Percent GDP/capita (°C)	All	1.272***	-0.049***	-0.100	-0.211
		(0.379)	(0.012)	(0.206)	(0.131)
	Poor	2.543	-0.077**	-0.586	-0.401
		(1.765)	(0.037)	(0.486)	(0.273)
	Rich	0.890**	-0.032*	0.141	-0.117
		(0.441)	(0.018)	(0.225)	(0.161)
D&G[b] Farm profit/acre (100 growing degree days °F)	All	-1.453**	0.002	-1.279***	0.192***
		(0.594)	(0.008)	(0.301)	(0.047)
	Rainfed	-1.129**	-0.001	-1.089***	0.199***
		(0.447)	(0.007)	(0.214)	(0.039)
	Irrigated	-4.076	0.025	-2.388**	0.037
		(4.414)	(0.052)	(1.082)	(0.226)
Precipitation					
BHM[a] Percent GDP/capita (Meters)	All	1.445	-0.475*	-0.195	-0.755*
		(1.003)	(0.255)	(0.427)	(0.418)
	Poor	2.580*	-0.744**	-0.347	-1.162***
		(1.415)	(0.366)	(0.505)	(0.378)
	Rich	0.672	-0.269	-0.225	0.615
		(1.442)	(0.372)	(0.733)	(0.788)
D&G[b] Farm profit/acre (Inches)	All	-0.560	0.006	-0.217	-0.072***
		(0.679)	(0.018)	(0.167)	(0.028)
	Rainfed	-1.530***	0.022**	-0.496***	-0.092***
		(0.393)	(0.010)	(0.119)	(0.022)
	Irrigated	2.382	-0.029	1.351*	0.113
		(1.834)	(0.052)	(0.774)	(0.074)

Notes: Standard errors in parentheses. For columns 3 and 4, we adjust standard errors to account for the fact that these are two-stage estimates (i.e., variables are demeaned before estimation). *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

[a] The BHM specification is equation (2.21) with year and country fixed effects and linear and quadratic country trends. Countries are not weighted. Standard errors are clustered by country.

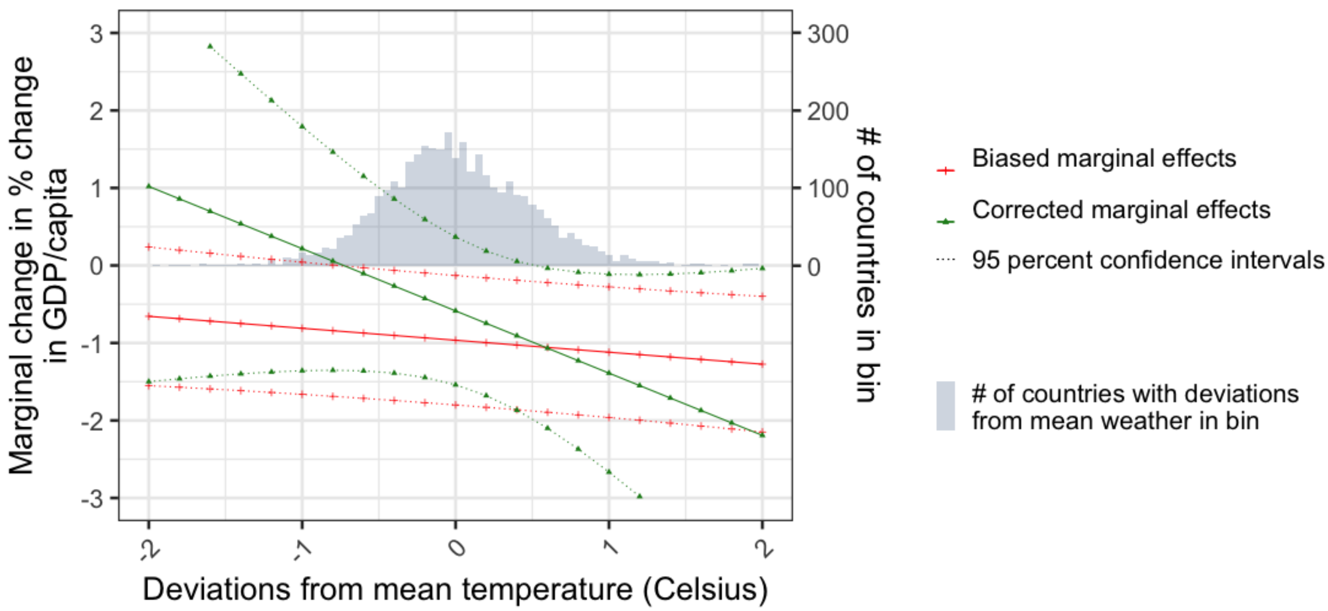
[b] The D&G specification is equation (2.20) with year and county fixed effects. Counties are weighted by farm acreage. Standard errors are clustered by county.

Figure 2.1 uses the estimated coefficients in Table 2.1 to compare the biased versus corrected annual marginal temperature effects on the percent GDP per capita for poor and rich countries. The biased estimates in BHM evaluated at the mean of the sample suggest that the marginal effect of temperature is almost constant around -1 percent/°C for poor countries and near zero for rich countries regardless of the weather. The corrected estimates for both poor and rich countries suggest the marginal effect of temperature goes from beneficial to harmful for temperature deviations at -0.7°C for poor countries and $+0.6$°C for rich countries. The 95 percent confidence intervals reveal the temperature effects are not quite significantly different from zero. Appendix F shows a similar figure for precipitation.

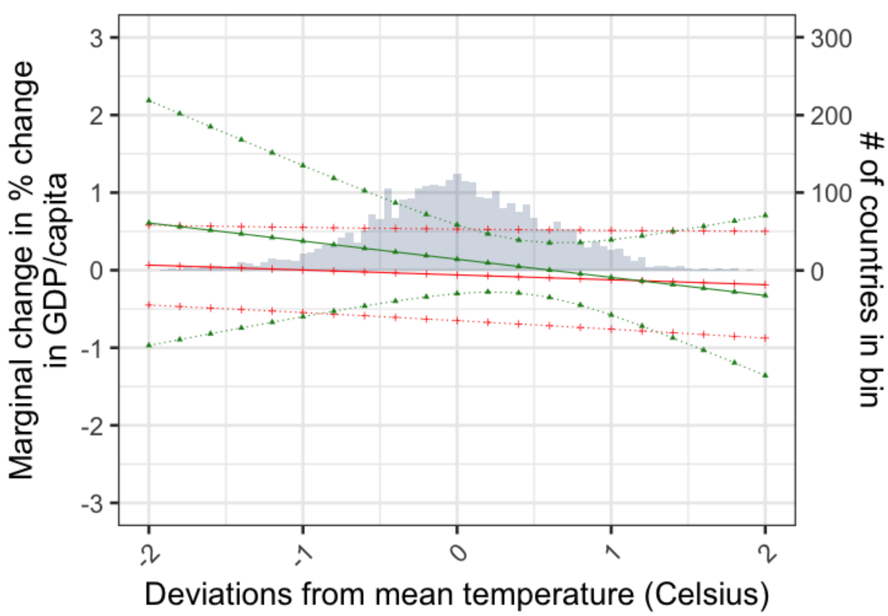
Figure 2.2 compares the biased and corrected marginal effect of temperature on farm profit/acre for rainfed and irrigated counties using the temperature coefficients in Table 2.1. The biased coefficients in D&G evaluated near the mean climate of the sample suggest the marginal temperature effects are near-constant and negative around $-\$5$/acre/°C for rainfed farms. The corrected results for rainfed farms suggest that the marginal temperature effect is negative only for cool deviations but becomes positive with warm deviations greater than 0.8°C. The biased and corrected marginal effects for irrigated farms both suggest a near-constant harmful effect between $-\$10$ and $-\$5$/acre/°C. The 95 percent confidence intervals show that the rainfed temperature results are significantly different from zero, but the irrigated temperature results are not significant.

Figure 2.1: BHM: Effect of Temperature on Percent GDP Per Capita

(a) Poor Countries



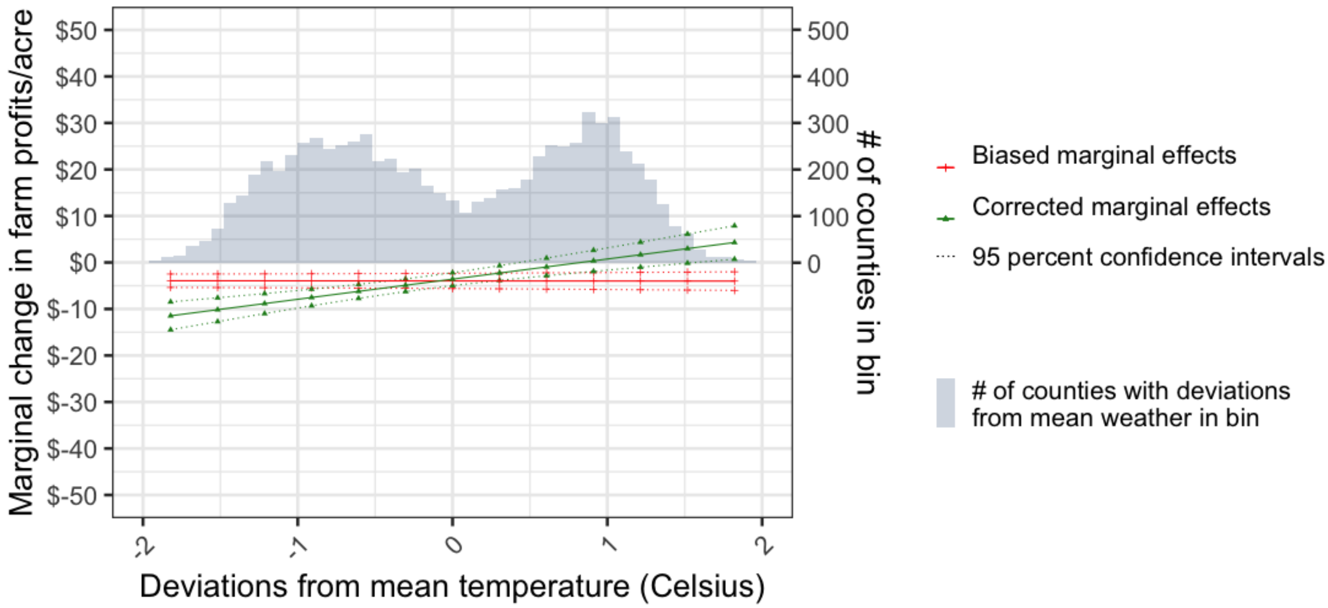
(b) Rich Countries



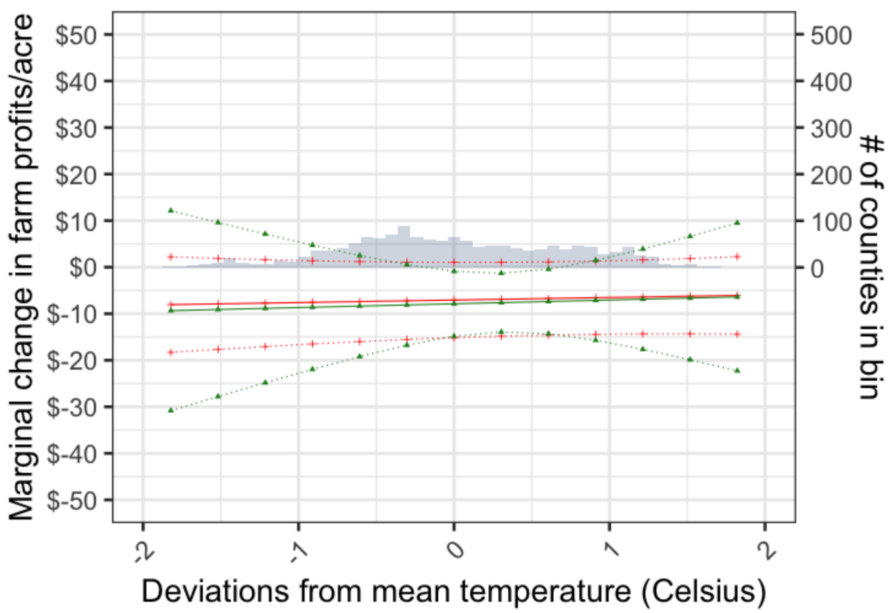
Notes: Biased (equation 2.16) and corrected (equation 2.19) marginal effects depend on whether fixed effects are incorrectly applied after transforming the variables (equation 2.16) or correctly applied before the variables are transformed (equation 2.19).

Figure 2.2: D&G: Effect of Temperature on Farm Profit Per Acre

(a) Rainfed Counties



(b) Irrigated Counties



Notes: Biased (equation 2.16) and corrected (equation 2.19) marginal effects depend on whether fixed effects are incorrectly applied after transforming the variables (equation 2.16) or correctly applied before the variables are transformed (equation 2.19). GDD in °F is converted to °C by: $°C = (GDD/183) * (5/9) + 8°C$.

The next set of results estimate the parameters of the climate-weather model when both climate and weather have quadratic effects, as in equation (2.15). This model includes an interaction term between weather and climate for both the farm profits per acre model of D&G by county and the percent GDP per capita model of BHM by country. Columns 1, 2 and 3 of Table 2.2 present the estimated coefficients of linear weather deviations ($\hat{\beta}_1$), the squared weather deviation ($\hat{\beta}_3$), and the interaction between climate and weather deviations ($\hat{\beta}_2$).

In the BHM context, the significant per capita income temperature results in Table 2.2 reveal that the linear coefficient is often positive while the coefficient on the interaction term is negative for poor countries. The negative interaction coefficients for temperature and precipitation coefficients imply that climate has a hill-shaped effect on per capita income. The marginal climate effect of temperature is positive when the expected temperatures of countries are cool and negative when the expected temperatures are hot. The marginal climate effect of temperature on GDP per capita switches sign at 13°C for the whole sample, 14°C for rich countries, and 16°C for poor countries. The quadratic temperature coefficient on weather deviations is not significant suggesting temperature deviations have a linear effect on GDP per capita. The quadratic coefficient on precipitation deviations, however, is negative and significant for the whole sample and poor countries suggesting rainfall deviations have a hill-shaped effect on GDP per capita. These results suggest that countries may want to have insurance against swings in precipitation but not against swings in temperature.

With the D&G data, the United States farm profit per acre results in Table 2.2 reveal that for rainfed farms, the temperature interaction coefficient is negative and the squared weather temperature coefficient is positive. The negative interaction term implies expected temperature has a hill-shaped effect on farm profit. The positive squared term suggests temperature fluctuations have a U-shaped effect. These weather results suggest that year-to-year variations in temperature lead to expected benefits to rainfed farms. Temperature does not have a significant effect on irrigated farms. The precipitation results suggest that rainfed and irrigated farms react in opposite ways to expected rainfall. Expected rainfall has a U-shaped effect on the profits of rainfed farms but a hill-shaped effect on the profits of irrigated farms. For rainfed farms, more rain above 15 inches is increasingly beneficial. For irrigated farms, the optimal amount of rain appears to be 21 inches. Less rain requires costly irrigation water and more rain makes irrigation less useful. Year-to-year variation in rainfall appears to be harmful to rainfed farms, but not irrigated farms. These results suggest that rainfed farmers may want to have weather insurance against swings in rainfall but not temperature.

Table 2.2: Coefficients of Nonlinear Climate Model

		[1] Linear Weather $\hat{\beta}_1$	[2] Climate Weather Interaction $\hat{\beta}_3$	[3] Weather Squared $\hat{\beta}_2$
Temperature				
BHM[a] Percent GDP/capita (°C)	All	1.209*** (0.357)	-0.094*** (0.023)	-0.161 (0.128)
	Poor	2.397 (1.751)	-0.148** (0.074)	-0.403 (0.267)
	Rich	0.861** (0.419)	-0.061 (0.037)	-0.106 (0.162)
D&G[b] Farm profit/acre (100 growing degree days °F)	All	-0.564 (0.574)	-0.019 (0.016)	0.194*** (0.047)
	Rainfed	-0.272 (0.486)	-0.024* (0.014)	0.197*** (0.039)
	Irrigated	-3.219 (3.800)	0.027 (0.093)	0.084 (0.225)
Precipitation				
BHM[a] Percent GDP/capita (Meters)	All	1.104 (1.080)	-0.759 (0.563)	-0.759* (0.404)
	Poor	1.544 (1.648)	-0.971 (0.854)	-1.156*** (0.357)
	Rich	0.886 (1.463)	-0.734 (0.774)	0.695 (0.800)
D&G[b] Farm profit/acre (Inches)	All	-0.375 (0.740)	0.007 (0.040)	-0.072*** (0.026)
	Rainfed	-1.602*** (0.408)	0.055** (0.021)	-0.092*** (0.022)
	Irrigated	2.596 (1.759)	-0.060 (0.098)	0.074 (0.074)

Notes: (1) Standard errors in parentheses. We adjust standard errors to account for the fact that these are two-stage estimates (i.e., variables are demeaned before estimation). $^*p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$. *(2) This model is based on equation (2.15) with an interaction term with climate.*

[a] BHM (equation 2.21) is estimated using the specification in equation (2.15). Model includes year and country fixed effects and linear and quadratic country trends. Countries are unweighted and standard errors are clustered by country.

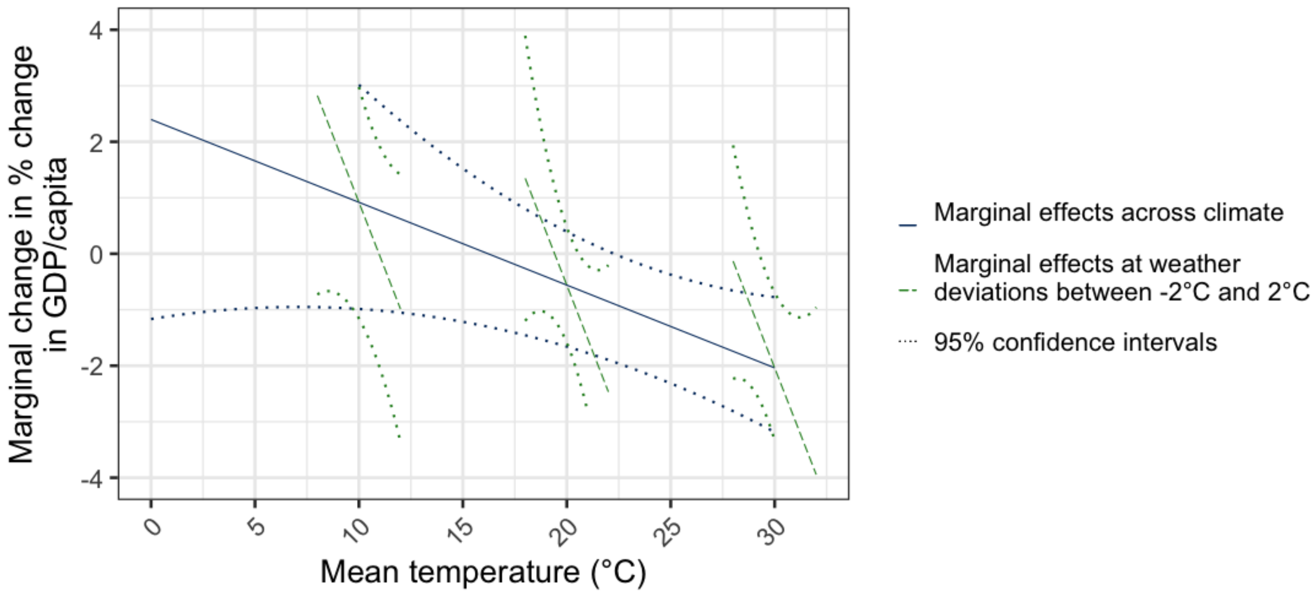
[b] D&G (equation 2.20) is estimated using the specification in equation (2.15). Model includes year and county fixed effects. Counties are weighted by farm acreage. Standard errors are clustered by county.

Figure 2.3 uses the coefficients from estimating equation (2.15) (as presented in Table 2.2) to compare the marginal temperature deviation effect of the weather versus the marginal expected temperature effect of climate on percent GDP per capita in poor and rich countries. The response is shown for climates from 0°C to 30°C. Note that there is a separate weather deviation function at each level of climate. The marginal climate and marginal weather effects are equal when the weather deviation is zero (the intersection of the lines). The marginal weather function is about 5 times steeper than the marginal climate function for poor countries. For poor countries, the marginal climate effect varies a great deal across this range from +2 percent/°C at 4°C to −2 percent/°C at 30°C . With rich countries, the climate response function to temperature is much flatter than the poor country response. The marginal climate effect varies from +0.9 percent/°C at 0°C to −1 percent/°C at 30°C. The marginal weather function for rich countries is about 4 times steeper than the marginal climate function. The weather (when it equals climate) provides an accurate forecast of the marginal effect of warming. However, the weather response function is far more concave than the climate response function. The nonlinear weather response is quite different from the nonlinear climate response.

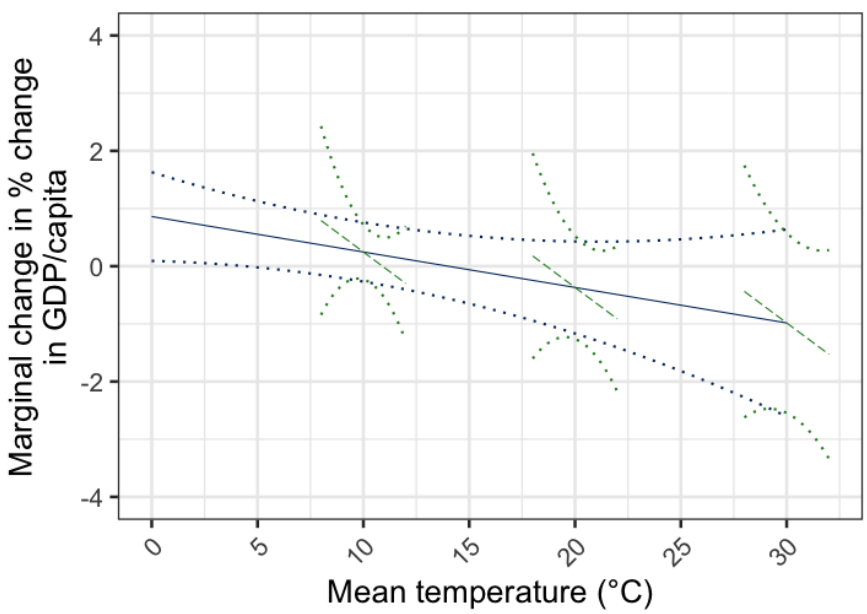
Figure 2.4 compares the marginal temperature weather deviation response (weather deviations) with the expected marginal temperature (climate) response of farm profits/acre for dryland and irrigated farmland again using equation (2.15). The range of climate responses is shown from growing season temperatures of 10°C to 30°C. Weather responses are shown for climates of 10°C, 20°C, and 30°C. The weather response is unique to each level of climate. For rainfed farms, the harmful marginal effect of warmer climates increases from −$1/acre/°C at 10°C to −$7/acre/°C at 30°C. For irrigated farms, the harmful marginal effect of warmer climates is shrinking from −$10/acre/°C at 10°C to −$4/acre/°C at 30°C. Perhaps the most surprising result in Figure 2.4 is the U-shaped effect of temperature deviations on farm net profits.

Figure 2.3: BHM: The Effect of Expected Temperature and Temperature Deviations on Percent GDP Per Capita

(a) Poor Countries



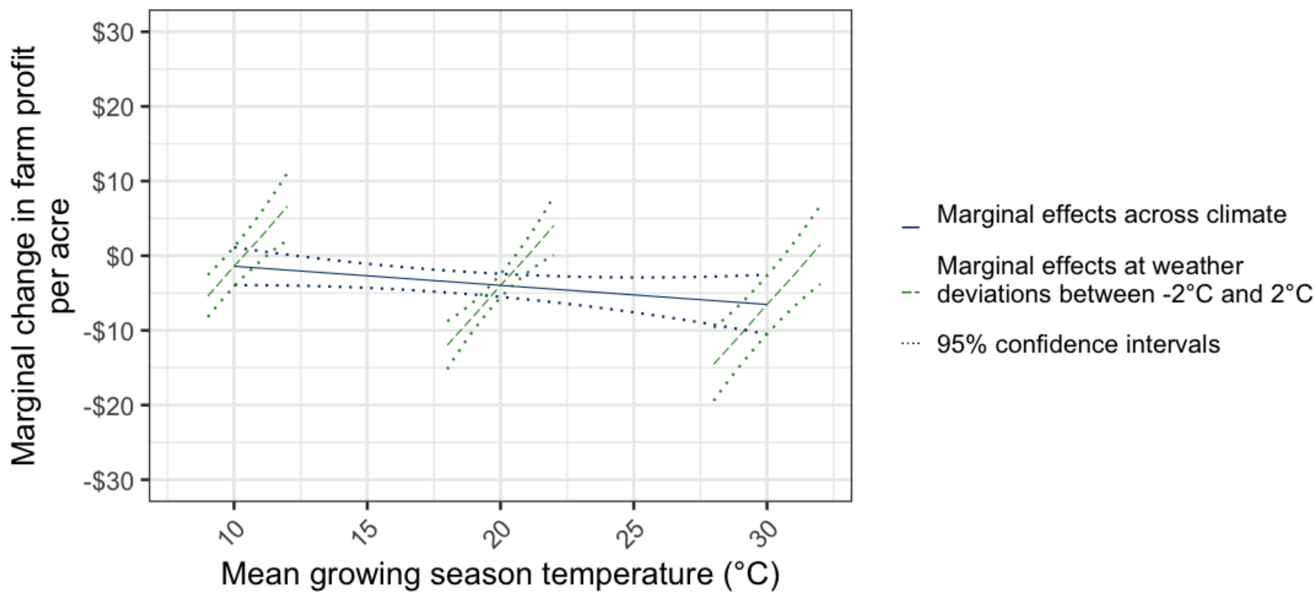
(b) Rich Countries



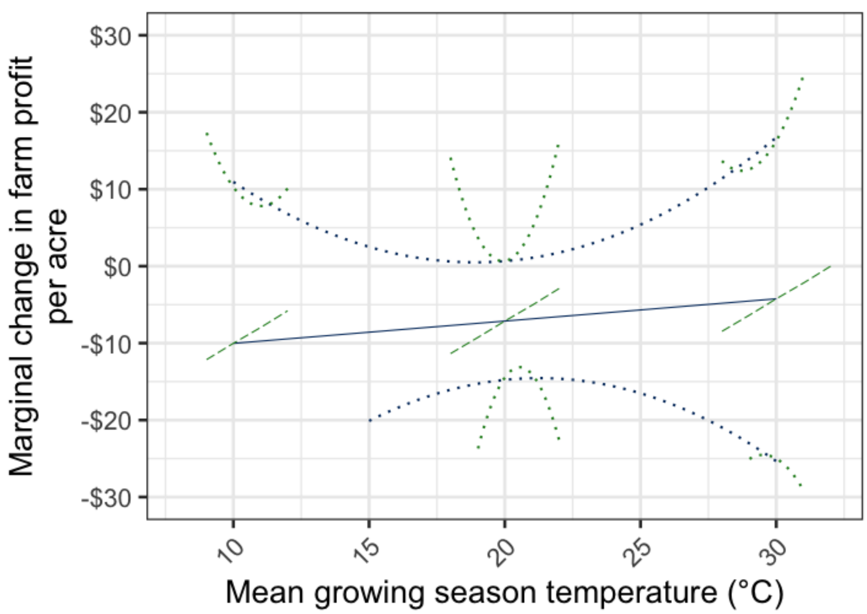
Notes: (1) Marginal weather effects and marginal climate effects are based on coefficients in Table 2.2. (2) The results predict the percent change in GDP per capita. (3) The marginal weather effects are estimated for temperature deviations between $-2^\circ C$ *and* $2^\circ C$ *at three different climates: mean temperatures of* $10^\circ C$, $20^\circ C$, *and* $30^\circ C$. *(4) See Appendix Figure F.3 for the precipitation effects.*

Figure 2.4: D&G: The Effect of Expected Temperature and Temperature Deviations on Farm Profit Per Acre

(a) Rainfed Counties



(b) Irrigated Counties



Notes: (1) Marginal weather effects and marginal climate effects are based on coefficients in Table 2.2. (2) The results predict the change in farm profits per acre. (3) The marginal weather effects are estimated for temperature deviations between $-2°C$ *and* $2°C$ *at three different climates: mean growing season temperatures of* $10°C$, $20°C$, *and* $30°C$. *(4) As in Figure 2.2, we have converted temperature units from growing season degree days to* $°C$. *(5) See Appendix Figure F.4 for the precipitation effects.*

Table 2.3 uses the corrected results in Tables 2.1 and 2.2 to examine a uniform 3°C climate warming scenario relative to year 1900 temperatures, implying that future temperatures warm another 2°C above current temperatures. We examine what this warming does to both the percent GDP per capita and farm profits per acre. Column 1 uses equation (2.16) to predict the D&G and the BHM marginal climate results using the linear weather coefficients in Table 2.1. Note that we are presenting stylized forecasts for a set of hypothetical locations. In contrast, in their published results, D&G and BHM aggregate their forecasts for each location in their sample based on each location's unique characteristics and various projections of future climate.

Column 2 of Table 2.3 uses equation (2.19) to predict the outcomes using the corrected linear coefficients in Table 2.1. The last two columns rely on the nonlinear climate model results from equation (2.15). Column 3 predicts the effect of a temporary 3°C warmer year using the weather coefficients in Table 2.2. Column 4 predicts the effect of a permanent 3°C warming using the climate coefficients in Table 2.2. Table 2.3 shows how the results change depending on whether the initial climate is 10°C or 25°C.[11]

The forecasted results in Table 2.3 vary a great deal across the different models. Using the original biased coefficients in the BHM model suggests a permanent 3°C warming would increase GDP per capita above current levels by +4.8 percent in poor countries and by +1.7 percent in rich countries. Only the rich country effect is statistically significant. Assuming a linear climate effect, the corrected coefficients in the BHM model suggest warming would lead to GDP per capita falling by −2.8 percent in poor countries and by −0.2 percent in rich countries. Only the poor country effect is significant. By implicit assumption, the climate effects in these models are the same regardless of the underlying climate of each country.

The most interesting result in Table 2.3, however, concerns the weather and climate effects with the nonlinear climate model. A temporary 3°C warmer year would increase GDP per capita by +0.2 percent in cool (10°C) poor countries but decrease GDP per capita by −4.2 percent in warm (25°C) poor countries. The same warm year increases GDP per capita in cool rich countries by +0.1 percent but decreases GDP per capita in warm rich countries by −1.8 percent. The predicted

[11] A temperature of 10°C is equivalent to 659 GDD and 25°C is equivalent to 5,600 GDD.

climate effect is less harmful. A +2°C permanent warming would cause GDP per capita in cool (10°C) poor countries to increase by +1.5 percent, but in warm (25°C) poor countries, GDP per capita falls by −2.9 percent. The warmer climate in cool rich countries causes GDP per capita to increase by +0.4 percent but in warm rich countries, to fall by −1.5 percent. The forecasted losses in warm poor countries are the only statistically significant climate and weather effects.

Table 2.3: Forecasted Change in Outcome with 3°C Warming Scenario

			[1] Original Biased Effect, Linear Climate	[2] Corrected Effect, Linear Climate	[3] Weather Effect, Nonlinear Climate	[4] Climate Effect, Nonlinear Climate
BHM Percent GDP/capita	Poor	10°C	4.78% (3.39)	-2.78%** (1.30)	0.23% (2.10)	1.54% (1.97)
		25°C	4.78% (3.39)	-2.78%** (1.30)	-4.21%*** (1.49)	-2.89%*** (0.94)
	Rich	10°C	1.65%** (0.82)	-0.19% (0.49)	0.07% (0.48)	0.37% (0.45)
		25°C	1.65%** (0.82)	-0.19% (0.49)	-1.77% (1.18)	-1.47% (1.30)
D&G Farm profit/acre	Rainfed	10°C	-$7.48** (3.10)	$1.48 (2.44)	$5.71 (3.58)	-$3.33 (2.40)
		25°C	-$7.48** (3.10)	$1.48 (2.44)	-$1.96 (3.06)	-$11.00*** (2.86)
	Irrigated	10°C	-$25.76 (31.06)	-$14.10 (9.37)	-$16.40 (17.32)	-$19.47 (19.30)
		25°C	-$25.76 (31.06)	-$14.10 (9.37)	-$7.72 (16.49)	-$10.79 (13.80)

*Notes: (1) Standard errors in parentheses. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. (2) The 3°C scenario from year 1900 is 2°C or 659 GDD warmer than the 1980-2010 temperature. BHM temperature is annual and D&G temperature is for the growing season. (3) Note that we are presenting stylized forecasts for a set of hypothetical locations. In contrast, in their published results, D&G and BHM aggregate their forecasts for each location in their sample based on each location's unique characteristics and various projections of future climate. (4) Column 1 uses equation (2.16). (5) Column 2 uses equation (2.19) to evaluate the warming assuming a linear climate model. (6) Columns 3 and 4 use equation (2.15) to evaluate the effect of a weather deviation and a change in climate assuming a quadratic climate model.*[12]

In order to determine the net effect of warming on global GDP per capita, one would have to sum the effects of future warming across all countries weighted by their population and GDP. Given that many countries who are "poor" today may be "rich" by the second half of the century, the overall global climate effect of the 3°C scenario on GDP per capita is likely to be less than −1 percent. Importantly, the model predicts that these effects will be highly heterogeneous. For example, the model implies that all of the net climate damage to GDP per capita will occur in relatively low latitude (warm) countries with the largest burden on those countries who remain poor. Of course, this is not a complete accounting of the effect of climate change. The analysis does not take into account nonmarket effects to health, ecosystems, and comfort. The analysis also does not account for changes in sea level.

The results of the 3°C climate scenario on United States farm profits are also shown in Table 2.3. Using the estimated linear weather coefficients from the original D&G approach in Table 2.1, column 1 predicts that rainfed farms would lose profits of −\$7/acre and irrigated farms would lose −\$26/acre. With the linear climate model and correcting for the estimation error in D&G, rainfed farms would gain \$1/acre and irrigated farms would lose −\$14/acre across the sample.

Using the nonlinear climate model reveals that weather effects of farm profits vary by initial climate. On a rainfed farm, a warmer year in a cool climate increases profit by \$6/acre, while in a warm climate, it decreases profit by −\$2/acre. The weather effect on irrigated farmland is larger than on rainfed farmland. With irrigated farmland in cooler climates (10°C), a warm year decreases profit by −\$16/acre and decreases profit in warmer climates (25°C) by −\$8/acre. None of the weather effects are statistically significant.

With the nonlinear climate model, the climate effects of the 3°C warming on farm profits are more harmful than the predicted weather effects. The climate effect on rainfed farmland in cooler climates is a loss of −\$3/acre. In warmer climates, the loss is −\$11/acre for rainfed farmland. The climate effect on irrigated farmland in cooler climates is a loss of −\$19/acre. The loss is −\$11/acre in warmer climates. The forecasted loss for rainfed farmland in warmer climates is the only statistically significant forecasted effect. The long-run climate effects may be larger than the short-run weather effects if climate leads to a reduction in irrigation capital and inputs.

2.4 Conclusion

An increasing number of studies use panel data to estimate nonlinear weather models on economic outcomes, but this literature has not been careful to distinguish between weather deviations and expected weather (climate). The result is a set of models that are not correctly specified. Many studies have assumed that fixed effects have removed the influence of climate from their weather estimation. However, this is valid only if the effect of climate is linear. Some studies have included a weather-climate interaction term, but they have not realized that this interaction term reveals a quadratic climate effect. In this paper, we develop a model that shows how the underlying climate can lead to different effects of weather on economic outcomes. The weather panel studies without a weather-climate interaction term effectively assume a linear climate effect. If climate has a nonlinear effect on outcomes, then the effect of weather depends on the climate, and the correct specification of the weather effect must include this interaction. In the simple case, where climate has a quadratic effect on outcomes, this involves including a single interaction term between weather and climate. Using this corrected specification, we show that the weather panel model can estimate the marginal effect of climate as well as the marginal effect of weather.

We have also discovered that many studies have used fixed effects incorrectly to estimate nonlinear weather effects. The literature has conditioned variables on the fixed effects after the weather variables have been transformed (raised to higher powers) rather than before they are transformed. This has led to biased coefficient estimates.

The empirical analysis in this study takes two prominent articles in the literature and reveals that both the theoretical and the econometric issues raised in this paper lead to very different results from what the authors report. Correcting the estimation error leads to important changes in the coefficients. There is a significant bias using existing fixed effects estimators on nonlinear variables. Second, correcting the specification makes an enormous difference in the results. If climate has a linear effect on economic outcomes, the model should only include transformations of weather deviations. The linear weather term in the model becomes the estimated climate effect. If climate has a nonlinear effect on outcomes, then climate also affects the nature of the weather effects.

The model must include weather-climate interaction terms to model the resulting weather effect correctly. The effect of climate can then be estimated using the nonlinearity of this between-group effect. The nonlinear effect of weather can be measured using the within-group variation of intertemporal weather deviations.

Third, many weather panel studies examine weather precisely to learn more about the effect of climate. With the correctly specified model, the model gives estimates of climate effects that are distinct from the weather effects.

The original results in D&G assuming a linear climate model suggest that a 3°C warming scenario would decrease rainfed farm profits by $-\$7$/acre and decrease irrigated farm profits by $-\$26$/acre. If climate has a linear effect on farm profit per acre, correcting the model for the biased coefficients suggests that the warming scenario would increase rainfed profits by \$1/acre and decrease irrigated profits by $-\$14$/acre. Allowing climate to have a nonlinear effect reveals that both the weather effect and climate effect will vary across climates. A temporary 3°C warming would increase rainfed profits in cooler climates by \$6/acre but decrease rainfed profits in warmer climates by $-\$2$/acre. The profits of irrigated farms in cooler climates would fall by $-\$16$/acre, and the profits of irrigated farms in warmer climates would fall by $-\$8$/acre. The nonlinear climate response is more harmful than the weather response. According to the climate response, a permanent 3°C warming reduces profits on rainfed farms in cooler climates by $-\$3$/acre and on rainfed farms in warmer climates by $-\$11$/acre. The climate response reduces profits on irrigated farms in cooler climates by $-\$19$/acre and on irrigated farms in warmer climates by $-\$11$/acre.

Using the linear coefficients estimated by BHM suggests that a +3°C climate scenario will be beneficial to both poor and warm countries. Assuming a linear climate model, correcting the estimation bias in these coefficients suggests that a 3°C warming would reduce poor country GDP per capita by -3 percent and rich country GDP per capita by -0.2 percent. However, if one assumes that climate has a nonlinear effect on GDP per capita, the results are more complicated. The weather effects of a warmer year would increase the GDP per capita of cooler poor countries by 0.2 percent but decrease the GDP per capita of warmer poor countries by -4.2 percent. The weather response would increase the GDP per capita of cooler rich countries by 0.1 percent but decrease the GDP per capita of warmer rich countries by -1.8 percent. In contrast, a permanent increase in temperature would lead to a less harmful outcome. The climate response to permanent warming would increase GDP per capita in cooler poor countries by 1.5 percent, but the GDP per capita of warmer poor countries would fall by -2.9 percent. The GDP per capita of cooler rich countries would rise by 0.4 percent and the GDP per capita of warmer rich countries would fall by -1.4 percent. Averaging the weather effects across countries suggests a warm year might lead to a -2 percent annual loss of global GDP per capita, while a permanent increase in temperature would lead to about a -1 percent annual loss in GDP per capita. The results suggest that the net damage to economies will be concentrated in countries with relatively warm current climates.

Note that these BHM results do not account for within-country heterogeneity. These effects only measure the direct effect of temperature on the economy. The economy will also be vulnerable to changes in sea level rise. The nonmarket impacts of warming are also not measured.

2.5 References

Blanc, Elodie and Wolfram Schlenker (2017) "The use of panel models in assessments of climate impacts on agriculture," *Review of Environmental Economics and Policy*, 10.1093/reep/rex016.

Burke, Marshall, Solomon M. Hsiang, and Edward Miguel (2015) "Global non-linear effect of temperature on economic production," *Nature*, 10.1038/nature15725.

——— (2017) "The marginal product of climate," *NBER Working Paper 24072.*

Deschênes, Olivier and Michael Greenstone (2007) "The economic impacts of climate change: Evidence from agricultural output and random fluctuations in weather," *American Economic Review*, 97 (1), 354–385, 10.1257/aer.97.1.354.

——— (2012) "The economic impacts of climate change: Evidence from agricultural output and random fluctuations in weather: Reply," *American Economic Review*, 102 (7), 3761–3773, 10.1257/aer.102.7.3761.

Greene, William H . (2003) *Econometric Analysis*, Upper Saddle River, NJ: Prentice Hall, 5th edition.

Hsiang, Solomon (2016) "Climate econometrics," *Annual Review of Resource Economics*, 8, 43–75, 10.1146/annurev-resource-100815-095343.

Kolstad, Charles D. and Frances C. Moore (2020) "Estimating the economic impacts of climate change using weather observations," *Review of Environmental Economics and Policy*, 14 (1), 10.1093/reep/rez024.

Lemoine, Derek (2020) "Estimating the consequences of climate change from variation in weather," *NBER Working Paper 25008*, 10.3386/w25008.

Mansur, Erin T. and Wendy Morrison (2008) "A discrete-continuous choice model of climate change impacts on energy," *Journal of Environmental Economics and Management*, 55, 175–193.

McIntosh, Craig T. and Wolfram Schlenker (2006) "Identifying non-linearities in fixed effects models," *Working Paper*.

——— (2016) "Measuring weather impacts using panel data," *2017 American Economic Association Conference Working Paper*, `https://www.aeaweb.org/conference/2017/preliminary/paper/fiHy249A`.

Mendelsohn, Robert and Ariel Dinar (2009) *Climate change and agriculture: An economic analysis of global impacts, adaptation and distributional effects*, Cheltenham: Edward Elgar Publishing, 256, 10.1111/j.1746-692x.2010.00181.x.

Mendelsohn, Robert and Emanuele Massetti (2017) "The use of cross-sectional analysis to measure climate impacts on agriculture: Theory and evidence," *Review of Environmental Economics and Policy*, 11 (2), `https://doi.org/10.1093/reep/rex017`.

Mendelsohn, Robert and James E. Neumann (2004) *The impact of climate change on the United States economy*, Cambridge: Cambridge University Press.

Mendelsohn, Robert, William D. Nordhaus, and Daigee Shaw (1994) "The impact of global warming on agriculture: A Ricardian analysis," *American Economic Review*, 84 (4), 753–771, 10.2307/2118029.

Mendelsohn, Robert and Michael E. Schlesinger (1999) "Climate-response functions," *Ambio*, 28, 362–366, 10.2307/4314910.

Mérel, P. and M. Gammans (2021) "Climate econometrics: Can the panel approach account for long-run adaptation?," *American Journal of Agricultural Economics, forthcoming.*

Moore, Frances C. and David B. Lobell (2014) "Adaptation potential of European agriculture in response to climate change," *Nature Climate Change*, 4, 610–614, 10.1038/nclimate2228.

Neyman, J. and Elizabeth L. Scott (1948) "Consistent estimates based on partially consistent observations," *Econometrica*, 16 (1), 1–32, 10.2307/1914288.

Ritchie, Joe T. and D. S. Nesmith (1991) "Temperature and Crop Development," in Hanks, John and J. T. Ritchie eds. *Agronomy Monographs: Modeling Plant and Soil Systems, Volume 31*, Chap. 2, 5–29: ASA CSSA SSSA, 10.2134/agronmonogr31.c2.

Schlenker, Wolfram and Michael J. Roberts (2009) "Nonlinear temperature effects indicate severe damages to U.S. crop yields under climate change," *Proceedings of the National Academy of Sciences of the United States of America*, 106 (37), 15594–98, 10.1073/pnas.0906865106.

Severen, Christopher, Christopher Costello, and Olivier Deschênes (2018) "A forward-looking Ricardian approach: Do land markets capitalize climate change forecasts?," *Journal of Environmental Economics and Management*, 89, 235–254, 10.1016/j.jeem.2018.03.009.

Smith, Joel B. and Robert Mendelsohn (2007) *The impact of climate change on regional systems: a comprehensive analysis of California*, United Kingdom: Edward Elgar Publishing, 304.

Chapter 3: The Effects of Price Shocks on New Vehicle Sales

with James Stock and Kenneth Gillingham[1]

Abstract– We estimate the effect of price changes on new vehicle sales using a structural vector autoregression (SVAR) with a hot-rolled steel price index as an instrument for vehicle transaction prices. Correctly identifying the price elasticity of new vehicle sales is especially important for estimating the impacts of fuel economy standards because changing fuel economy stringency is assumed to cause a shock to new vehicle prices. The resulting effect on new vehicle sales has broad implications beyond the immediate impact on the vehicle industry. In particular, new vehicles generally have the best safety features and pollution controls, so reducing replacement of used vehicles has consequences for public safety and pollution levels. This project is also a novel application of SVAR-IV, a relatively new methodology borrowed from the monetary policy literature. We compare the SVAR-IV with local projection IV (LP-IV), SVAR, and autoregressive distributed lag (ARDL) models, the latter of which has been considered in policymaking for fuel economy standards. Using the impulse responses in the first-year, we estimate a range of short-run price elasticities. The results are highly unstable, so we are unable to have confidence in any precise elasticity. In future work, we will pursue stronger instruments and more robust specifications.

[1]**Acknowledgments and Disclosures:** James Stock is a Professor in the Department of Economics at Harvard University. Kenneth Gillingham is a Professor in the School of the Environment at Yale University. The authors declare that they have no conflicts of interest and were not compensated for this work. Kenneth Gillingham has worked as a paid consultant for the California Air Resources Board on other analysis relating to fuel economy standards. James Stock has not yet reviewed the text and results in this version of the draft. Earlier work on this project appeared in "Comments on Notice of Proposed Rulemaking for The Safer Affordable Fuel-Efficient (SAFE) Vehicles Rule for Model Years 2021-2026 Passenger Cars and Light Trucks." Docket number EPA–HQ–OAR–2018–0283 and NHTSA–2018–0067. https://environment.yale.edu/gillingham/Comments%20on%20SAFE%20NPRM_Stock-Gillingham-Davis_102618.pdf

3.1 Introduction

Even small changes in the new vehicle market can affect the U.S. economy where new vehicle sales account for over three percent of U.S. GDP and over 1.5 million people are directly employed in vehicle-related manufacturing and dealerships.[2] New vehicle sales also have implications for public safety and environmental health, since new vehicles embody the latest technologies and regulations related to safety and pollution. Therefore, policymakers are particularly concerned with legislation and rulemaking that may affect new vehicle sales, such as standards for fuel economy and tailpipe emissions. To the extent that regulations lead to lower new vehicle adoption and lower replacement of used vehicles, the associated reductions in safety and increases in pollution will offset other benefits of the regulation.

Most recently, for example, regulators in the U.S. proposed to reduce the stringency of average fuel economy standards for new vehicles.[3] Proponents of the rule change argued that the decrease in fuel economy stringency will lead to higher utilization of new vehicles, thus displacing older vehicles with inferior safety features and pollution control. In particular, lower fuel economy standards will decrease new vehicle prices, increase demand for new vehicles, decrease demand for used vehicles, and increase used vehicle scrappage. This impact of new vehicle prices on scrappage decisions is known as the Gruenspecht effect (Gruenspecht 1982). Since scrapped vehicles are often the oldest, lowest-quality vehicles, they tend to have the least-effective safety features and the lowest fuel economy. Therefore, increased scrappage of older vehicles and increased utilization of new vehicles could reduce traffic fatalities and also offset some of the pollution increases from lowering fuel economy standards. Our research seeks to estimate the magnitude of these effects by modeling the

[2] New vehicle expenditures from: Bureau of Economic Analysis (BEA) Table 7.2.5U.

[3] In March 2020, the U.S. Department of Transportation (DOT) and U.S. Environmental Protection Agency (EPA) jointly issued modified Corporate Average Fuel Economy (CAFE) and CO2 standards for tailpipe emissions. The new rule increases the CAFE and CO2 standard stringency by 1.5 percent each year for model years 2021 to 2026 where stringency is measured in sales-weighted, firm-level, fleet average miles per gallon (MPG). The 2020 FRIA estimates that these new standards will reduce stringency for model year 2029 from 46.6 MPG under the 2012 rule to 40.5 MPG under test conditions and from 37.1 MPG to 33.2 MPG under real world conditions. Note that these estimates of the 2012 rule stringency are substantially lower than estimated in the 2012 FRIA, which estimated test condition CAFE stringency at 54.5 MPG in model year 2025. The 2020 FRIA explains that the decreased estimates of the 2012 rule stringency are the result of updated modeling assumptions, such as increased consumer preferences for SUVs.

See: EPA and DOT Final Regulatory Impact Analysis (FRIA) for "The Safer Affordable Fuel-Efficient (SAFE) Vehicles Rule for Model Year 2021-2026 Passenger Cars and Light Trucks," (March 2020).

relationship between fuel economy standards and new vehicle prices and sales.

For the particular question of estimating the effect of fuel economy standards, we have chosen to use time series methods with U.S.-level aggregate new vehicle sales data. Under this approach, we decompose the price increase from increased fuel economy stringency into two components: 1) the share that is valued by consumers as future fuel savings, and 2) the share that acts as a pure price shock. Before discussing the merits of this approach, we first describe two alternative approaches that have appeared in the literature. The first alternative is to use disaggregate vehicle data with hedonic and discrete choice methods that account for consumers' relative preferences across vehicle attributes (e.g., Griliches 1971; Ohta and Griliches 1976, 1986; Levinsohn 1988; Berry 1994; Berry et al. 1995; McCarthy 1996; Berry et al. 2004; Fischer et al. 2007; Leard 2020). The level of disaggregation in these studies varies substantially but can be as granular as vehicle make-model-market as in (e.g., Berry et al. 2004; Leard 2020).

When reported at all, the estimates of aggregate market-level new vehicle price elasticity also vary substantially. For example, Berry et al. (1995) (hereafter BLP) estimate own-price elasticities of demand for individual make-models, not the price elasticity of the entire new vehicle market with respect to the outside options. These elasticities are in the range of -3 to -7, which could be considered high upper bounds for the price elasticity of the aggregate new vehicle market as noted in Busse et al. (2013). However, a BLP model could also be used to actually estimate an elasticity for aggregate new vehicle sales. Taking this approach, Berry et al. (2004) estimate an aggregate market elasticity of -0.4. This estimate is consistent with Leard (2020), who also finds an aggregate new vehicle market price elasticity of -0.4 using a nested logit approach akin to Berry (1994), relating market shares to consumer preference parameters.[4] One other contribution of Leard is to improve the estimation using consumer survey data on second-choice preferences. Other examples of elasticity estimates from disaggregated make-model-level vehicle sales data include Levinsohn (1988) (short-run price elasticities ranging from -0.81 to -0.83), McCarthy (1996) (short-run price

[4]When Berry et al. (2004) consider an aggregate market elasticity of -1 suggested by General Motors staff, they estimate a mean semi-price elasticity of -10.56, which is well above the range in original BLP. Berry et al. also consider the case where there is no correlation between product-specific constant terms and price, which had been implicitly assumed in prior literature. This results in an estimate of aggregate market price elasticity of -0.2 and a mean semi-price elasticity of -0.75, well below the range in original BLP.

elasticity of -0.87), and Fischer et al. (2007) (long-run price elasticity of -0.36).

These models with disaggregated data are well-suited to explore substitution effects between vehicle types. Thus, these models would be useful for evaluating policies that differentially affect vehicle prices and features. However, we are interested in the case of a fuel economy standard, which affects all vehicles simultaneously. Therefore, we argue that modeling this policy requires an identification that is also based on market-level shocks. For example, the disaggregate models described above do not capture substitution between new and used vehicles and other outside options as well as deferment of purchases across time periods. While BLP and other nested logit models do include an outside option, these models do not capture time dynamics, which we believe to be an especially important margin of adjustment in the fuel economy context.[5]

A second approach to estimating the effects of fuel economy standards on new vehicle sales is to directly model the effect (e.g., Leard 2020; Dou and Linn 2020). For example, Dou and Linn (2020) use time-varying and vehicle-type variation in fuel economy standards to estimate that a 1 percent increase in fuel economy standards would reduce new vehicle sales by 0.4 percent. Using this estimate and assumptions about marginal costs and markups, Dou and Linn estimate an aggregate new vehicle price elasticity of -1.5 (short-run). Using the disaggregate model described above, Leard (2020) estimates an aggregate new vehicle cost per mile elasticity of demand of -0.02. This aggregate new vehicle cost per mile elasticity can be used together with the aggregate new vehicle market price elasticity to estimate the full sales effect of fuel economy standards. Under various assumptions, Leard estimates that a 1 percent tightening of fuel economy standards will decrease vehicle sales between 0.016 and 0.28 percent. Again, as with the disaggregate data models describe above, these models that directly estimate the effect of fuel economy standards on sales do not capture important time dynamics. Also, our model of the effect of pure price shocks is more generally applicable to other policies that affect new vehicle prices.

The third approach, which we adopt in this paper, is to use time series techniques with aggregate

[5]a) Disaggregated data could also be used in constructing an index of quality adjustments for use in our time series models, but this exercise would be somewhat redundant, since vehicle quality adjustment are already readily available from the U.S. Bureau of Labor Statistics.
b) Further information on the Bureau of Labor Statistics New Vehicle CPI: https://www.bls.gov/cpi/factsheets/new-vehicles.htm

vehicle sales data and macroeconomic covariates to estimate the price elasticity of new vehicle sales. New vehicle price elasticity estimates from this literature include a long-run estimate of -1.2 in Nerlove (1957); short-run estimates between -0.78 and -1.17 and long-run estimates between -0.30 and -0.46 in Hymans et al. (1970); a short-run estimate of -1.63 in Hess (1977); a short-run estimate of -0.79 and a long-run estimate of -0.61 in Center for Automotive Research (CAR) (2016);[6] and long-run estimates between -0.2 and -0.3 in the 2018 DOT/EPA PRIA.[7] To summarize, the range new vehicle price elasticities in our literature review of all three methodologies (disaggregate models, estimating the effect of fuel economy standards directly, and aggregate vehicle sales time series models) have been from -0.2 (long-run) to -1.63 (short-run).

In addition to producing a wide range of elasticity estimates, the time series literature also exhibits substantial disagreement about the choice of variables and model specification. In particular, there is disagreement over whether the price variable should be quality-adjusted; the differencing and log transformation of variables; the number of lags; and the choice of macroeconomic covariates. Most importantly, though, this literature has universally failed to resolve the classic challenge of endogeneity and simultaneity in the relationship between vehicle sales and price. To correct this shortcoming, we borrow a relatively new methodology from the monetary policy literature, the structural vector autoregression with instrumental variables (SVAR-IV).

Prior research has primarily used SVAR-IV in the context of Federal Reserve policy announcements as an instrument for interest rates. SVAR-IV is a two-step method combining vector autoregression and external instruments first introduced by Stock (2008) and then further popularized by Gertler and Karadi (2015) exploiting Federal Reserve policy surprises. We compare the SVAR-IV estimates to SVAR and autoregressive distributed lag (ARDL) models, a version of which was used in the 2018 DOT/EPA PRIA. While the ARDL model does not at all address the endogeneity between sales and price, the SVAR only imposes the assumption of contemporaneous exogeneity. That is, vehicle pricing events (e.g., dealership discounts) cannot be influenced by same-quarter inventories and vehicle sales.[8] Unlike ARDL, SVAR does not impose the exogeneity assumptions for the relationship

[6]See McAlinden et al. (2016).

[7]EPA and DOT Preliminary Regulatory Impact Analysis (PRIA) for "The Safer Affordable Fuel-Efficient (SAFE) Vehicles Rule for Model Year 2021-2026 Passenger Cars and Light Trucks," (July 2018).

[8]Note that we use the term "pricing events" to refer to dealership activities, such as discounts and promotions.

between price and previous quarters' sales. Under SVAR, dealership pricing events can be influenced by prior quarters' sales and inventories without violating the identification assumptions. The added advantage of instrumental variables with SVAR-IV is that even the contemporaneous exogeneity assumption is relaxed.

To explore the robustness of the SVAR-IV models, we also estimate local projection models with instrumental variables (LP-IV). Under certain conditions, SVAR-IV and LP-IV may actually identify identical impulse responses and elasticity estimates (see Plagborg-Møller and Wolf 2021). The advantages of local projections are that they have an easier implementation that VARs and sometimes require weaker assumptions. Local projection models were first introduced in Jordà (2005) and then augmented with instrumental variable techniques in Jordà et al. (2015). Stock and Watson (2018) provide an overview of the SVAR-IV approach and compare its efficiency to LP-IV using the Gertler and Karadi example. In this paper, Stock and Watson call for novel applications of SVAR-IV and LP-IV, a call we answer here. Thus, this paper contributes new applications of the SVAR-IV and LP-IV methods as well as more careful estimates of the relationship between new vehicle sales and price, an open and important policy question.

This paper proceeds as follows: Section 3.2 discusses the application of our elasticity estimates to the fuel economy setting; Section 3.3 describes our various time series models; Section 3.4 lays out our data; Section 3.5 presents the impulse responses and elasticities; and Section 3.6 concludes.

3.2 Fuel Economy Standards Are Not Pure Price Shocks

The time series models in this paper seek to identify the effect of exogenous price shocks on new vehicle sales. However, modeling the price increases that result from fuel economy standards as exogenous prices increases raises three issues: 1) we do not know the true costs of regulatory compliance (e.g., the 2020 DOT/EPA FRIA; National Research Council 2015; Dou and Linn 2020); 2) manufacturers may not pass through 100 percent of the costs of meeting the standards (e.g., Gron and Swenson 1996, 2000); and 3) consumers are likely to value at least some of the reduced

We reserve the term "sales" for quantity demanded (i.e., vehicles units sold).

operating costs from future fuel savings (e.g., Busse et al. 2013; Allcott and Wozny 2014; Sallee et al. 2016; Leard et al. 2017b).[9] Fortunately, there has been substantial literature on these topics (see citations). Thus, we leave it to future work and practitioners to rescale our impulse responses and elasticities to the particular price shock best-suited for the given analysis.

3.3 Models

Below, we describe four models for estimating the relationship between aggregate new vehicle transaction prices and sales: ARDL, SVAR, SVAR-IV, and LP-IV. We present these models in order from most restrictive assumptions (ARDL) to least restrictive (SVAR-IV and LP-IV). Our models abstract aways from the context of fuel economy standards and instead focus on identifying the effect of a one-time exogenous price increase. The purpose is to estimate the dynamic effect on sales of a price increase, that is, the causal effect of a price increase on current and future new vehicle demand. In contrast to a static causal effect, a dynamic causal effect allows the response to an intervention to evolve over time.

3.3.1 Autoregressive Distributed Lag (ARDL)

We begin with the ARDL model used in the 2018 DOT/EPA PRIA. While this model is deeply flawed, we present it here to as a matter of policy relevance and also as a basis of comparison for the SVAR, SVAR-IV, and LP-IV specifications. Comparing the results across various models also demonstrates the importance of relaxing the ARDL assumptions and carefully accounting for the endogeneity of sales and price.

Specifically, the 2018 PRIA estimates the dynamic response of new vehicle sales to an autonomous increase in price according to:

$$S_t = \theta_1 S_{t-1} + \theta_2 S_{t-2} + \beta[P_t - P_{t-1}] + \gamma_1 GDPgrowth_t + \gamma_2 E_t + \gamma_3 E_{t-1} + u_t \tag{3.1}$$

[9] Dou and Linn (2020) mentions a Leard et al. (forthcoming) estimate that each 1 percent increase in fuel economy raises production costs by $90 per vehicle.

where S_t is new vehicle sales in quarter t (seasonally adjusted at an annual rate (SAAR)); P_t is the average new vehicle transaction price; $GDPgrowth_t$ is the percentage growth of GDP from quarters $t-1$ to t (SAAR); E_t is number of persons employed (seasonally adjusted), and u_t is an error term. These data are fully described in Section 3.4.

The primary flaw in using an ARDL model for estimating the demand for new vehicles is the failure to account for the well-known fact that supply and demand are jointly determined. Thus, observed data on prices and quantities reflect variations in both supply and demand. Only if prices were set randomly— or, more generally, if prices are set in a way that is unrelated to demand— then a regression of quantity on price would estimate a demand curve. However, if prices respond to shifts in demand, then the regression of quantity on price will not provide an unbiased estimate of the demand curve or (in the logarithmic specification) of the demand elasticity. It is useful to distinguish two types of endogenous behavior of prices: the response of prices to past demand disturbances, and the response of prices to current demand disturbances. For the ARDL model to estimate the dynamic causal effect of a price change on demand, both these conditions need to be true, along with an additional assumption that prices not anticipate future demand disturbances. This trio of assumptions— past, contemporaneous, and future exogeneity— is referred to as strict exogeneity. Strict exogeneity is a key assumption for the validity of the ARDL model. Because the ARDL model includes lagged sales and prices as endogenous variables, current prices must be exogenous with respect to lagged sales and thus to lagged demand disturbances. The assumption of strict exogeneity of prices is not plausible in the context of car and light duty truck sales where manufacturers and dealerships can rapidly adjust prices in response to past, present, or expected demand. For example, a dealer who had disappointingly weak sales in the previous quarter and had thus built up an inventory on their lot, might decide to hold a sale in the current quarter. This behavior violates the past exogeneity condition. If the dealer holds a sale in the current quarter to counteract disappointing demand in that same quarter, then doing so violates the contemporaneous exogeneity condition.

The strict exogeneity condition can also be violated for reasons other than the feedback from demand disturbances to prices. In fact, strict exogeneity requires that any omitted factor that affects sales be uncorrelated with past, present, and future prices. It is easy to see how this requirement could be violated. For example, vehicle sales could be influenced by a major quality improvement in vehicles in a given year. But such a quality improvement in vehicles would also be correlated with vehicle prices, as automakers would be expected to raise their prices along with the quality improvement. Thus, vehicle prices would not be strictly exogenous. The ARDL model in equation 3.1 includes two drivers of demand, GDP growth and employment. For that model, the exogeneity condition is that prices must be strictly exogenous conditional on GDP growth and the current and first lag of employment. While these two variables control for some important drivers of demand, other demand disturbances remain, for example driven by unexpected changes in operating costs (e.g. gasoline prices), in tastes, or for economic reasons not fully captured by national GDP growth and employment.

Additional weaknesses of the ARDL model include its structural assumptions with respect to lags, the use of levels as opposed to logarithms, and its assumptions about cointegration between sales and employment. All of these assumptions need to be tested empirically, as we do in Section 3.5, not assumed in the initial statement of the model. As embodied in the number of lags of sales and prices in equation 3.1, the ARDL model assumes a specific structure for the dynamic effects of a price increase. The preferred approach would would flexibly test for the appropriate the number of lags. Furthermore, the levels specification implies a time-varying elasticity, but this should only be preferable if we can reject the constant elasticity assumption under the logarithmic specification. Also, the use of levels rather than differenced sales and employment assumes cointegration, which should also be tested.

3.3.2 Structural Vector Autoregression (SVAR)

The assumption that prices have no dynamic response to past demand shocks can be relaxed using an SVAR instead of an ARDL model by explicitly allowing feedback from past demand disturbances to prices. This is done by adding an equation that explicitly models current prices as a function of past demand shocks, which are in turn captured by past prices and sales data. Because cars and light trucks are a major part of the economy, it is plausible that there is some feedback from vehicle demand to overall macroeconomic conditions in the future. This logic suggests also including equations for other macroeconomic variables (e.g., GDP growth and employment) that allow for this feedback. The SVAR estimates a system of equations where each variable is regressed on lags of the other variables. Because of these multiple equations and the dynamic feedback in the SVAR, the SVAR resolves the problem of past exogeneity. However, the SVAR still needs an assumption similar to the ARDL model's contemporaneous exogeneity condition.

The assumption that prices have no dynamic response to past demand shocks can be relaxed using an SVAR instead of an ARDL model by explicitly allowing feedback from past demand disturbances to prices. This is done by adding an equation that explicitly models current prices as a function of past demand shocks, which are in turn captured by past prices and sales data. Because cars and light trucks are a major part of the economy, it is plausible that there is some feedback from vehicle demand to overall macroeconomic conditions in the future. This logic suggests also including equations for other macroeconomic variables (e.g., GDP growth and employment) that allow for this feedback. The SVAR estimates a system of equations where each variable is regressed on lags of the other variables. Because of these multiple equations and the dynamic feedback in the SVAR, the SVAR resolves the problem of past exogeneity. However, the SVAR still needs an assumption similar to the ARDL model's contemporaneous exogeneity condition.

We now follow Stock and Watson (2018) to describe SVAR estimation of the system:

$$A(L)Y_t = v_t \tag{3.2}$$

where v_t is an innovation (also known as a Wold-error) unobserved to the econometrician. $A(L) = I - \sum_{k=1}^{p} A_k L^k$ where I is the identity matrix, L is the lag operator, and p is the maximum number of lags in the SVAR system. As in the ARDL model, Y_t includes various transformations of vehicle sales S_t and price P_t as well as macro variables, such as an index of consumer sentiment, CS_t, GDP, GDP_t, an index of manufacturing production, M_t, and number of employed persons, E_t. For example, $Y_t = [CS_t \quad \Delta GDP_t \quad M_t \quad E_t \quad \Delta P_t \quad S_t]$. Since this is a structural VAR, we assume that the innovations v_t are linear combinations of shocks ϵ_t such that:

$$v_t = \Theta_0 \epsilon_t \tag{3.3}$$

where Θ_0 is some nonsingular matrix to be estimated. Also, Θ_0 is chosen such that v_t are uncorrelated giving:

$$E v_t v_t' = \Theta_0 \epsilon_t \epsilon_t' \Theta_0' = \Theta_0 \Sigma \Theta_0' = I \tag{3.4}$$

where Σ is the variance-covariance matrix. Combining Equations 3.2 and 3.3 and then inverting $A(L)$ yields:

$$Y_t = C(L)\Theta_0 \epsilon_t \tag{3.5}$$

where $C(L) = A(L)^{-1}$.[10]

A primary challenge with SVAR identification is that the Θ_0 that satisfies Equation 3.4 is not unique. Thus, Sims (1980) suggests a Cholesky decomposition to find the unique lower triangular square root of Σ. To ensure this lower triangularity, we must choose the order (i.e., a Wold causal ordering) in which each variable enters the SVAR so that each variable has zero contemporaneous correlation with the variables that enter after it and nonzero contemporaneous correlation with all other variables that entered before it. Thus, the first variable to enter the SVAR has zero contemporaneous correlation with all the other variables, and the last variable to enter the SVAR has nonzero contemporaneous correlation with all the other variables.[11] For example, if we assume that sales enters the ordering after price, then past sales as well as past and contemporaneous price can affect current price; past sales as well as past and contemporaneous price can affect current sales; but contemporaneous sales cannot effect current price.

To estimate this SVAR system, we first estimate the following system equation-by-equation by ordinary least squares (OLS):

$$Y_{t,i} = A_{0,i} Y_{t,1:i-1} + \sum_{k=1}^{p} A_{k,i} Y_{t-k} + \mu_i + v_{t,i} \tag{3.6}$$

[10]See Stock and Watson (2018) for further discussion of this "invertibility assumption."
[11]This ordering is known as a Wold causal ordering.

where μ_i is a vector of constants, and the first term implements the Wold causal ordering. When estimating the equation of the first variable in the ordering (i.e., when $i = 1$), the first term is omitted. The orthogonalized impulse response functions (OIRFs) are given by:

$$\frac{\partial Y_{t+h,i}}{\partial u_{t,j}} = \{\Phi_h \Theta_0\}_{i,j}$$

which is the effect of a one unit shock in variable j at time t on variable i at time $t + h$.[12] Φ_h are given by:

$$\Phi_0 = I \quad \text{and}$$
$$\Phi_h = \sum_{k=1}^{p} \Phi_{h-k} A_k$$

where terms $\Phi_{h-k} A_k$ are omitted if $k > h$.

3.3.3 Structural Vector Autoregression with Instrumental Variables (SVAR-IV)

Fully addressing the problem that prices are simultaneously determined requires relaxing the assumption of contemporaneous exogeneity. Relaxing this assumption that prices have no contemporaneous response to demand shocks requires an instrumental variable in the SVAR. In this case, we require an instrument that is correlated with price changes (relevance) but is uncorrelated with unobserved demand disturbances (exogeneity). Continuing to follow Stock and Watson (2018), estimation of SVAR-IV proceeds exactly as the estimation of SVAR described in Section 3.3.2 except that the estimate of $\Theta_{0,i \geq s,s}$ is replaced with $\Theta^{SVAR-IV}_{0,i \geq s,s}$ where s indicates the variable we are instrumenting (e.g., price). The subscript $i \geq s$ indicates that we are only instrumenting variable s and the variables that can be contemporaneously affected by variable s, as imposed by the assumptions of our Wold causal ordering.[13] To estimate $\Theta^{SVAR-IV}_{0,i \geq s,s}$, we introduce an instrument Z_t that satisfies the conditions of relevance and exogeneity (Stock and Watson (2018) refer to these as Condition

[12] https://blog.stata.com/2016/02/18/vector-autoregressionsimulation-estimation-and-inference-in-stata/

[13] This $i \geq s$ subscript is not in Stock and Watson (2018) where the variable to be instrumented, interest rates, $gs1$, is $s = 1$.

SVAR-IV):

(1) relevance: $E\epsilon_{t,s}Z_t' = \alpha' \neq 0$; and

(2) exogeneity with respect to other current shocks: $E\epsilon_{t,-s}Z_t' = 0$.

We use two-stage least squares equation-by-equation with instrument Z_t to estimate $\Theta_{0,s}^{SVAR-IV}$ from Equation 3.6 where the instrument may also include a number of lags.

From Condition SVAR-IV and Equation 3.3, it follows that:

$$E(v_t Z_t') = E(\Theta_0 \epsilon_t Z_t) = \Theta_0 E \begin{bmatrix} \epsilon_{t,s} Z_t' \\ \epsilon_{t,-s} Z_t' \end{bmatrix} = \Theta_0 \begin{bmatrix} \alpha' \\ 0 \end{bmatrix} = \begin{bmatrix} \Theta_{0,s,s}\alpha' \\ \Theta_{0,-s,s}\alpha' \end{bmatrix}. \tag{3.7}$$

Here, note that the matrix $\Theta_{t,i,j}$ corresponds to the effect of impulse j on variable i. Taking Equation 3.7 along with a unit effect normalization such that $\Theta_{0,s,s} = 1$, we see that:

$$\frac{E(v_{t,i}Z_t)}{E(v_{t,s}Z_t)} = \Theta_{0,i,s}. \tag{3.8}$$

Thus, $\Theta_{0,i,s}$ satisfies the IV regression:

$$v_{t,i} = \Theta_{0,i,s} v_{t,s} + \{\epsilon_{t,-s}\}. \tag{3.9}$$

Noting that $v_{t,i} = Y_{t,i} - Proj(Y_{t,i}|Y_{t-1}, Y_{t-2}, ...)$, we can rewrite Equation 3.9 as Equation 3.6 estimated by two stage least squares with instrument Z_t. Replacing $\Theta_{0,i\geq s,s}$ with $\Theta_{0,i\geq s,s}^{SVAR-IV}$, we proceed with SVAR estimation and calculation of OIRFs exactly as we did Section 3.3.2.

3.3.4 Local Projection IV (LP-IV)

We estimate LP-IV analogously to SVAR-IV (see Stock and Watson 2018 for more details). LP-IV has the nice feature that elasticities are estimated directly by the coefficient on the price term, β_t (when all variables are in logs). However, we also estimate an impulse response functions in terms of unit sales and calculate an alternative set of elasticities that are constructed analogously to those calculated under SVAR-IV. In SVAR-IV, we account for log-transformation using the adjustment factor:

$$qscale = (\$1,000/\$32,000) * 17.8/sschol(i_x, i_x) \tag{3.10}$$

where *sschol* is the Cholesky decomposition of the covariance matrix and i_x is the index of the price variable. We then adjust vehicle sales by *qscale* before estimation. In LP-IV, we construct an impulse response function in terms of unit sales effects according to:

$$\Delta S_t = (17.8 * (\frac{\$32,000 + \$1,000}{\$32,000})^{\beta_t} - 17.8) \tag{3.11}$$

where $\$32,000$ is the baseline price and 17.8 million are the baseline annual vehicle sales.

3.4 Data and Instrument

We primarily use quarterly data from the federal agencies, Bureau of Economic Analysis (BEA) and Bureau of Labor Statistics (BLS). Data is quarterly from 1967 to the first quarter of 2020 as shown in Figure 3.1.[14] Data include quarterly new motor vehicle sales, S_t, and transaction prices, P_t, for the U.S..[15] The sales data does not include imported light trucks prior to 1976, a period when imported truck sales were very low. To create the price series, we aggregate all new vehicle expenditures and investments across sectors (consumer, private, and government) and divide by total new vehicle sales.[16] We adjust the price series by an all-item CPI.[17] Before 1987, the BEA

[14]Employment data only extends back to the first quarter of 1968, but since GDP and vehicle prices are differenced, we do use the 1967 data.

[15]New vehicle sales data is from the BEA via the St. Louis Federal Reserve (FRED) where we sum sales from the following series: DAUTOSAAR, FAUTOSAAR, DLTRUCKSSAAR, and FLTRUCKSSAAR. Data is monthly, SAAR, and includes auto and light truck sales across all sectors (consumer, business, and government). We take means across months to aggregate to quarters.

[16]We take the sum of total new vehicle expenditures and investments in the consumer, private, and government sectors from BEA Table 7.2.5U. In particular, consumer sector expenditures are from line 6, "Final sales of domestic product: Personal consumption expenditures: New motor vehicles;" private sector investments are line 22, "Final sales of domestic product: Private fixed investment: New motor vehicles;" government sector auto investments are line 40, "Final sales of domestic product: Gross government investment: Autos: New autos;" and government sector truck investments are line 44, "Final sales of domestic product: Gross government investment: Trucks: New trucks (including utility vehicles)." These expenditures are quarterly and SAAR.

[17]The all-item CPI is from the BEA via FRED series CPIAUCSL, "Consumer Price Index for All Urban Consumers," rebased to 2016 dollars (2016=100).

data does not distinguish between light and heavy truck private fixed investment, so we assume, that private fixed investment on light trucks were 43 percent of all truck private fixed investment for this period. Similarly, the BEA data does not distinguish between light and heavy truck government investments before 2003, so we assume, that government investment on light trucks were 36 percent of all government truck investment for this period. We recognize that these assumptions introduce measurement error, which may attenuate our results.

The macroeconomic covariates are plotted in Figure 3.2. All covariate data is quarterly, including real GDP, GDP_t, and number of persons employed, E_t (both SAAR).[18] Note that the 2018 NHTSA PRIA ARDL model uses GDP growth rate,[19] but they update their model to use log-differenced GDP in the 2020 FRIA. We also include the University of Michigan consumer sentiment index, CS_t, which is not seasonally adjusted.[20] In consideration of our instrument (discussed below), we also include a index of manufacturing production output, M_t (seasonally adjusted).[21]

3.4.1 Hot-Rolled Steel as an Instrument

Relevance: We have identified hot-rolled steel as a highly relevant instrument candidate because of its importance in vehicle manufacturing. For instance, hot-rolled coil steel is made into sheets used for car bodies, while other types of hot-rolled steel are used in car frames.[22] In particular, we use the BLS quarterly producer price index by industry for iron and steel mills producing sheet and strip hot-rolled steel, including tin mill products, which extends back to 1965.[23] This series is

[18]a) The GDP data is from BEA via FRED series GDPC1, real gross domestic product, chained 2012 dollars, SAAR, quarterly.
b) Employment data is from the BLS via FRED series LNS12500000, full-time employed persons (in thousands) age 16 and over, SAAR. We took quarterly means across months.

[19]GDP growth data is from the BEA via FRED series A191RL1Q225SBEA, which is percent change in real GDP from the preceding quarter, SAAR.

[20]FRED series UMCSENT, University of Michigan, consumer sentiment index, not seasonally adjusted.

[21]a) The manufacturing index is from the Board of Governors via FRED series IPB00004SQ, industrial production: manufacturing (SIC), seasonally adjusted, quarterly.
b) We use the Standard Industrial Classification Code (SIC) rather than the more modern NAIC index because the SIC index extends back to 1967.

[22]a) Reuters. "Hot-rolled mess: China's steelmakers hit the skids as car sales slow." February 27, 2019.
b) Done, Brad. MachineDesign. "What's the difference between hot and cold rolled steel?" June 7, 2016.
c) Reuters. "EU investigates Chinese hot-rolled steel, puts duties on road wheels." October 10, 2019.

[23]The hot-rolled steel price index is from the BLS via FRED series PCU3311103311105, producer price index by industry: iron and steel mills: hot rolled steel sheet and strip, including tin mill products, available from 1965, not seasonally adjusted. We take quarterly means of the monthly data.

presented in Figure 3.3.

The hot-rolled steel producer price index series is not seasonally adjusted, so we have explored seasonal adjustments using the U.S. Census Bureau's X-13ARIMA-SEATS Seasonal Adjustment Program.[24] Seasonal adjustments tend to dampen the quarter-to-quarter variation, so we leave the series unadjusted. For example, the largest quarter-to-quarter decrease and increase are approximately -32 and 29 percent, respectively, in the unadjusted data. In the seasonally adjusted data, these values are -30 and 24 percent. All of these large price changes occur during the 2008 financial crisis. The mean absolute quarter-to-quarter change is 4 percent for the unadjusted series and 3.9 percent for the adjusted series.

Our primary alternative to the domestic industry hot-rolled steel price index is a quarterly series of U.S. flat hot-rolled steel import prices from the U.S. International Trade Commission (ITC). Specifically, we use Harmonized Tariff Schedule (HTS) commodity code 7208, "flat-rolled products of iron or nonalloy steel, of a width of 600 mm or more, hot-rolled, not clad, plated or coated." We divide import values (nominal USD) by quantities (kilograms) where import values include the value of the goods as well as the cost of duties, freight, insurance, and other charges. This imports category is specifically imports for consumption, which exclude into steel sent to bonded warehouses or foreign-trade zones (FTZs).[25] Figure 3.3 compares hot-rolled steel import prices to the domestic price index. Appendix Figure G.1 decomposes import quantities and values by country.

We also consider several alternative steel price indices. For example, producer price indices by industry are provided for structural hot-rolled steel, steel and iron, and cold-rolled steel. All of these indices are also available by commodity instead of industry for which the BLS has a different data collection process. Our preference for the industry index is due only to the fact that it is available from 1965, while the commodity indices are only available from 1982.[26] We also tried a BLS steel import price index available since 1979.[27] For all of these BLS series, we take quarterly

[24]We use the program's default settings. https://www.census.gov/srd/www/x13as/

[25]U.S. International Trade Commission. Downloaded 6-22-2020. https://dataweb.usitc.gov/

[26]BLS via FRED series WPU101703, producer price index by commodity for metals and metal products: hot rolled steel sheet and strip, including tin mill products, available from 1982, not seasonally adjusted. See also FRED series WPU101 and WPU101707.

[27]U.S. Bureau of Labor Statistics, Import Price Index (End Use): Iron and steel mill products [IR141], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/IR141, downloaded June 21, 2020.

means across monthly data. A notable feature of the SVAR-IV approach is that the instrument time series need not extend as long as the other variables, so the limited extent of the of the commodity indices do not preclude their use.

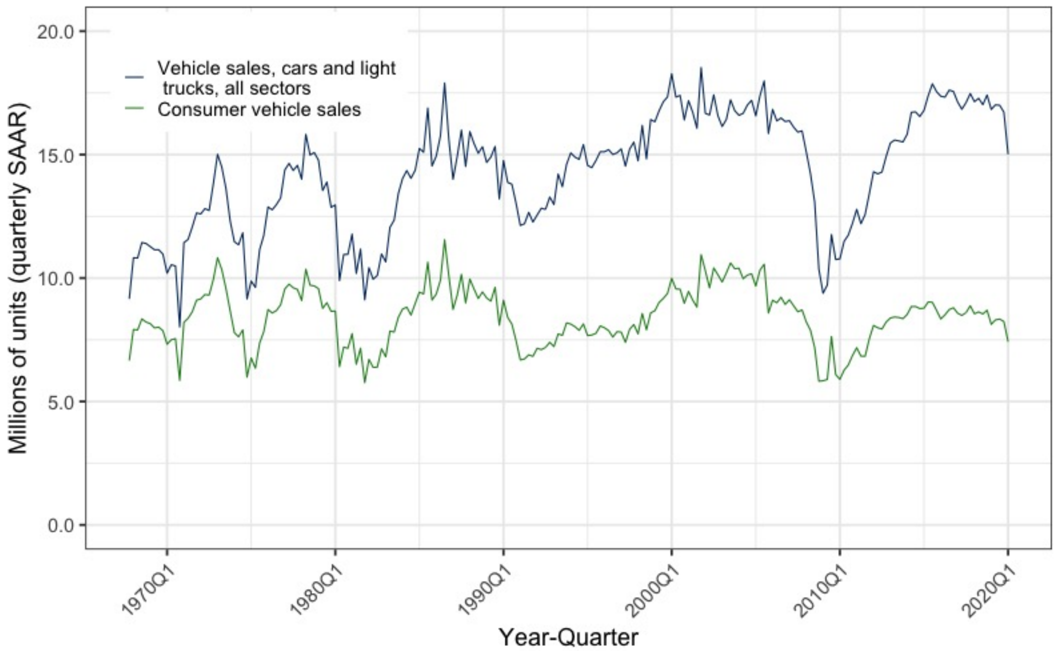
Exogeneity: Using steel as our instrument is potentially problematic because omitted variables that affect vehicle sales may also affect steel prices. As the vehicle market is a primary steel consumer, there is even the possibility that steel prices and vehicle sales are jointly determined (just as vehicle sales and prices are jointly determined). Financial news articles frequently suggest that vehicle markets can have a substantial impact on steel prices.[28] However, these articles also suggest several sources of steel price shocks that are plausibly exogenous to the U.S. auto market, including trade wars (e.g., the U.S.-China trade war), government dumping and anti-dumping policies (e.g., China and the European Union), and slumping demand in the Chinese auto market (to the extent that the Chinese auto mark is independent of the U.S. market). To help control for the effects of U.S. markets on steel prices, we include a Board of Governors index of manufacturing output (described above).

Quarterly means taken across monthly data. Index 2000=100, Not Seasonally Adjusted. Updated: Jun 12, 2020. Available after 1979.

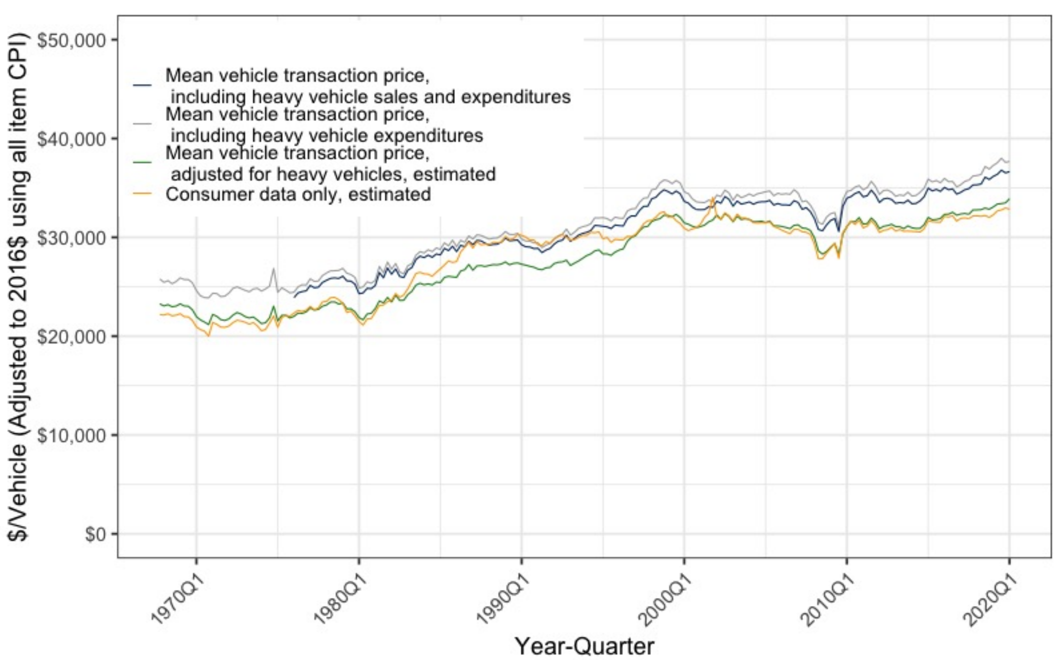
[28] a) Reuters. "China steel, iron ore edge lower in wobbly trade amid tepid demand." February 27, 2019.
b) Reuters. "Hot-rolled mess: China's steelmakers hit the skids as car sales slow." Februrary 27, 2019.

Figure 3.1: Vehicle Sales and Prices, Quarterly

(a) U.S. Vehicle Sales, SAAR

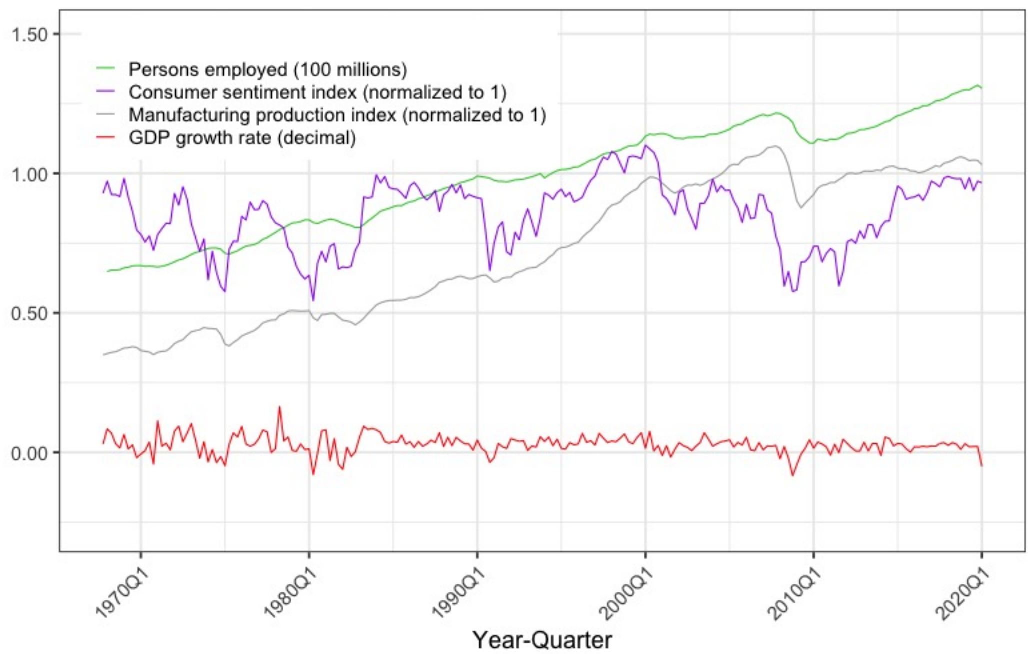


(b) U.S. Mean Vehicle Prices



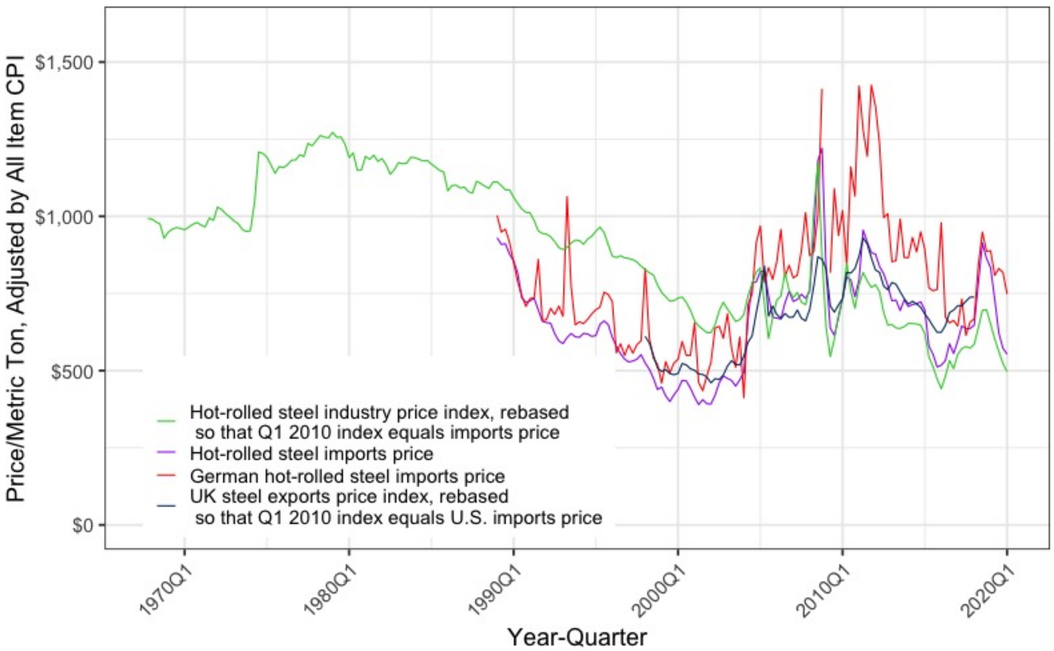
Notes: (1) Quarterly sales data are from FRED and include all light duty trucks and autos. (2) We create transaction prices by dividing quarterly BEA expenditures on new vehicles across sectors by the sales series and adjusting by an all-item CPI. (3) See Section 3.4 for more details.

Figure 3.2: Macro Covariates: GDP Growth, Employment, Manufacturing Index, and Consumer Sentiment



Notes: (1) GDP percent growth is from the BEA. (2) Number of persons employed is from the BLS. (3) Manufacturing production output index is from the Board of Governors and based on SIC codes. (4) Consumer sentiment index is from the University of Michigan. (5) See Section 3.4 for more details.

Figure 3.3: Hot-rolled Steel Price Indices and Import Prices



Notes: (1) The hot-rolled steel industry price index is from the BLS. (2) Hot-rolled steel import prices are from from the U.S. International Trade Commission (ITC). Specifically, we use Harmonized Tariff Schedule (HTS) commodity code 7208, "flat-rolled products of iron or nonalloy steel, of a width of 600 mm or more, hot-rolled, not clad, plated or coated." We divide quarterly import values (nominal USD) by quantities (kilograms) where import values include the value of the goods as well as the cost of duties, freight, insurance, and other charges. (3) The price of German imports to the U.S. is constructed using the same data as the total import price. (4) The U.K. exports price index is from the U.K. Office for National Statistics. (5) Both price indices have been re-based so that the index value in the first quarter of 2010 is the same value as the mean U.S. import price in that quarter. (6) All series are adjusted by an all-item CPI re-based to 2016 USD. (7) Not seasonally adjusted.

3.5 Estimation and Results

3.5.1 Cointegration

Prior to estimation, we first consider non-stationarity and cointegration to choose the appropriate transformations of the variables. This analysis leads us to applying our various models to relate the level of vehicle sales, S_t, the first difference of vehicle prices, ΔP_t, differenced GDP, ΔGDP_t, the level of employment, E_t, the level of the manufacturing index, M_t, and the level of the consumer sentiment index, CS_t. We difference vehicle prices and GDP to account for non-stationarity. Consumer sentiment does not exhibit a unit root (see Figure 3.2). The use of the levels of sales, manufacturing, and employment is justified only if these variables are cointegrated; if they are not cointegrated, then this regression will yield a spurious correlation between the three variables (Engle and Granger 1987). We therefore undertook Engle-Ganger Augmented Dickey Fuller tests of the null hypothesis of non-cointegration, and we rejected the null hypothesis, against the alternative of cointegration, at the one percent significance level for both the levels and log specifications. Visual inspection of the variables (Figures 3.1 and 3.2), as well as of their estimated cointegrating residual (error correction term), confirms the presence of a shared stochastic trend between the logs of these variables. This analysis especially strongly confirms the use of the levels (or logs) of sales and employment, not differences, in our models. As the instrumented variable, vehicle prices, is differenced, we also difference the instrument, a hot-rolled steel price index.

3.5.2 Wold Causal Ordering

As described in Section 3.3.2, we need to assume a Wold causal ordering to impose the lower triangular square root of Σ to ensure that the Θ_0 that satisfies Equation 3.4 is unique. In this ordering, variables enter the SVAR successively so that each variable has zero contemporaneous correlation with the variables that enter after it and nonzero contemporaneous correlation with all other variables that entered before it. Thus, we assume that (a) consumer sentiment, GDP growth, manufacturing output, and number of persons employed do not respond to unexpected movements in vehicle sales or prices within a quarter; (b) conditional on the consumer sentiment, GDP growth, manufacturing output, and employment, prices are contemporaneously exogenous for the purpose of estimating demand; and (c) sales can respond to price changes within a quarter. Together, these imply a Wold causal ordering for the series in the order of consumer sentiment, GDP growth, manufacturing output, number of persons employed, the change in price, and sales, (i.e., $\{CS_t, \Delta GDP_t, M_t, E_t, \Delta P_t, S_t\}$). Note that for the purpose of estimating the dynamic response of sales to a price disturbance, it does not matter the order of consumer sentiment, GDP growth, manufacturing output, or number of persons employed in the causal ordering.

3.5.3 Evidence of Constant Elasticity (Logs Versus Levels)

In the case of the ARDL model, we can also calculate the short-run price elasticity of vehicle sales according to $\eta = \beta \frac{\bar{P}}{\bar{S}}$, where the bars over the variables refer to the mean taken over all values in the dataset. Here, the levels specification of Equation 3.1 implies a time-varying elasticity, as opposed to the more standard logarithmic specification, which would imply a constant elasticity. The short-run elasticity from estimating the logarithmic version of Equation 3.1 is simply the price coefficient $\eta = \beta$. Thus, we perform two Wald tests of the hypothesis that the elasticity is constant, against the alternative that it varies with economic conditions.

Our first test for time-varying elasticity allows the elasticity to vary with the state of the economy through the growth rate of GDP and/or employment (two interaction terms). The second test allows the elasticity to vary with the four-quarter growth rate of GDP and/or the level of employment. Both tests fail to reject the null hypothesis that the elasticity is constant at the 10 percent significance level. In markets that are subject to large qualitative changes, such as the market for smart phones, one might think that the elasticity changes over time as the uses of the product changes. For vehicles, however, the primary use has long been to provide personal transportation services through driving. Absent theoretical or empirical reasons to treat the elasticity as time-varying, we prefer the standard approach of assuming the elasticity is constant. Therefore, we run all of our main specifications with log transformations of the variables (i.e., $\{\log(CS_t), \Delta \log(GDP_t), \log(M_t), \log(E_t), \Delta \log(P_t), \log(S_t)\}$). However, we have performed several sensitivities with variables in levels.

3.5.4 Choosing the Number of Lags: AIC

Table 3.1 shows the Akaike information criterions (AICs) for SVAR and SVAR-IV models with various lag lengths. In the instrumented models (row 2), the instrument is the domestic hot-rolled steel price index, differenced, with the same number of lags as the variables in the SVAR. While some models have AICs which are minimized with more than two lags, we are concerned that these models include attempt to estimate too many parameters. Therefore, we consider the 2-lag SVAR-IV as our preferred model.

Table 3.1: AIC, SVAR-IV,

		(1)	(2)	(3)	(4)	(5)
Levels or logs	Instrument	1 lag	2 lags	3 lags	4 lags	5 lags
Logs	No	-25.36	-25.71	-25.63	-25.61	-25.57
Logs	Yes	-21.09	-22.09	-21.67	-22.09	-22.02

Notes: See notes for Table 3.2.

3.5.5 Instrument Strength and First Stage

Table 3.2 shows the first-stage F-statistics as we estimate the SVAR-IV and LP-IV with variously 2, 3, or 4 lags; log transformations of the variables; the instrument (hot-rolled steel) in differences; and the instrument with the same number of lags as the variables in the SVAR/LP. Note that the preferred Olea and Pflueger (2013) first-stage F-statistics are the heteroscedasticity-robust F-statistics in column 4, since we have exactly one instrument for one endogenous regressor (i.e., $k = 1$). Both hot-rolled steel price indices, domestic and imported, are weak instruments for new vehicle prices.

Table 3.2: First-Stage F-Statistics for SVAR-IV

(1) Number of lags	(2) Levels or logs	(3) First stage F-statistic	(4) Heteroscedasticity-robust first stage F-statistic
SVAR-IV: Hot-rolled steel industry price index, CPI adjusted:			
2	Logs	9.9	7.3
3	Logs	9.4	7.5
4	Logs	14.1	13.5
SVAR-IV: Hot-rolled steel import prices, CPI adjusted:			
2	Logs	4.5	4.4
3	Logs	6.0	7.5
4	Logs	5.7	9.2
LP-IV: Hot-rolled steel industry price index, CPI adjusted:			
2	Logs	8.8	5.4
3	Logs	7.9	5.6
4	Logs	12.7	10.4
LP-IV: Hot-rolled steel import prices, CPI adjusted:			
2	Logs	5.3	5.1
3	Logs	7.5	8.4
4	Logs	7.2	9.4

Notes: (1) The instrument is the price index for hot-rolled steel, which is differenced (i.e., $\Delta \log(Z_t)$). (2) We use the same number of lags of the instrument as in the SVAR as indicated in column 1. (3) All lags are quarterly. (4) We estimate the log SVARs using a Wold causal ordering $\{\log(CS_t), \Delta \log(GDP_t), \log(M_t), \log(E_t), \Delta \log(P_t), \log(S_t)\}$. (5) Column 3 is the standard (conditional homoscedasticity, no serial correlation) first-stage F-statistic. (6) Column 4 is the heteroscedasticity-robust first-stage F-statistic (no lags).

3.5.6 Impulse Response Functions and Elasticities

Table 3.3 shows the impulse response functions (IRFs) and elasticities for new vehicle sales resulting from a \$1,000 one-time price increase. We calculate SVAR-IV elasticities using a \$32,000 base average price and 17.8 million vehicles of quarterly sales at an annual rate following:

$$\epsilon_t = \frac{\Delta S_t}{17,800} \bigg/ \frac{\$1,000}{\$32,000} \tag{3.12}$$

where ΔS_t is the impulse response of sales in period t. As discussed in Section 3.3.4, the LP-IV elasticities are estimated directly, but we calculate the LP-IV sales effects according to equation 3.11. The sales effects and elasticity estimates are all highly unstable and extremely sensitive to specification. The LP-IV estimates are especially sensitive. A similar story is evident in the impulse responses of Figures 3.4 to 3.7 where the impulse responses show large fluctuations before dissipating towards zero. Note that the LP-IV impulse responses are less smooth than those estimated by SVAR-IV. This is because the coefficients for each time horizon (lag length) are estimated separately under LP-IV. Given the lack of robustness, we have no confidence that we have precisely estimated any elasticities or impulse responses.

Table 3.3: Effect on Sales of a $1,000 One-Time Price Increase (Thousands of Units)

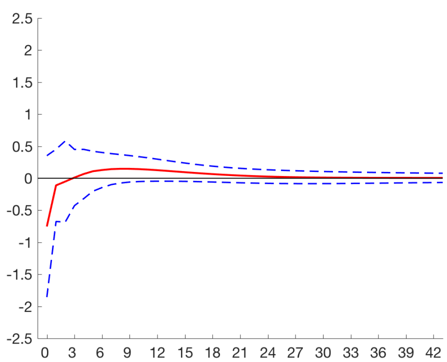
(1)	(2)	(3)	(4) First quarter (at annual rate)	(5) First quarter SE	(6) First year	(7) Total, years 1-10	(8) First quarter elasticity	(9) First year elasticity	(10) Q1 elasticity	(11) Q1 SE	(12) Mean Q1-4 elasticity	(13) Mean Q1-40 elasticity
ARDL model:			-31	145	-19	-38	-0.06	-0.03				
SVAR, SVAR-IV, and LP-IV models:												
Number of lags	Levels or logs	Instrument										
SVAR:												
2	Logs	No	-339	145	12	717	-0.61	0.02				
3	Logs	No	-348	149	149	1,226	-0.63	0.27				
4	Logs	No	-356	159	64	803	-0.64	0.11				
SVAR-IV: Hot-rolled steel industry price index, CPI adjusted:												
2	Logs	Yes	-753	672	-226	293	-1.35	-0.41				
3	Logs	Yes	-841	688	-125	723	-1.51	-0.22				
4	Logs	Yes	-751	704	-170	332	-1.35	-0.31				
SVAR-IV: Hot-rolled steel import prices, CPI adjusted:												
2	Logs	Yes	516	766	503	1,589	0.93	0.9				
3	Logs	Yes	-149	772	260	1,430	-0.27	0.47				
4	Logs	Yes	-772	792	-183	307	-1.39	-0.33				
LP-IV: Hot-rolled steel industry price index, CPI adjusted:												
2	Logs	Yes	-796	-330	741	10,347	-1.43	1.33	-1.49	0.9	1.25	1.77
3	Logs	Yes	-1,026	-411	479	12,000	-1.84	0.86	-1.93	1.2	0.79	2.02
4	Logs	Yes	-723	-236	432	14,846	-1.3	0.78	-1.35	0.74	0.72	2.47
LP-IV: Hot-rolled steel import prices, CPI adjusted:												
2	Logs	Yes	-69	-43	1,387	10,579	-0.12	2.49	-0.13	0.66	2.37	1.83
3	Logs	Yes	-645	349	469	6,503	-1.16	0.84	-1.2	0.94	0.8	1.11
4	Logs	Yes	-1,101	-642	270	7,224	-1.98	0.48	-2.08	0.54	0.43	1.18

Notes on next page.

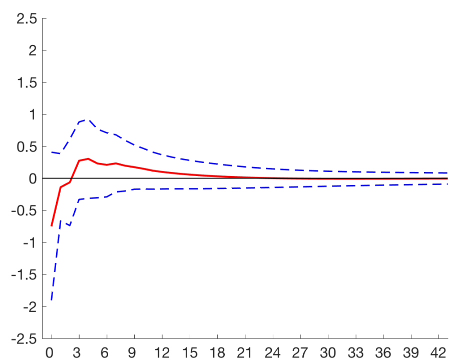
Notes for Table 3.3: (1) See notes for Table 3.2. (2) The ARDL model is estimated according to Equation 3.1, which is in levels and does not include the manufacturing index or consumer sentiment. These results differ from the 2018 PRIA, as we use our own data. (3) The logarithmic specification estimates impulse responses in percentages. We converted these responses to sales using a $32,000 base average price and 17.8 million vehicles of quarterly sales at an annual rate. (4) We calculate column 8 and 9 elasticities using Equation 3.12. (5) We interpret coefficients as elasticities for columns 10-13. These elasticities are only presented for LP-IV, since local projections estimate unique parameters for each lag length.

Figure 3.4: Response of Quarterly Sales (Millions, SAAR) to a $1,000 One-Time Permanent Price Increase, SVAR-IV, Instrument: Domestic Hot-Rolled Steel Price Index from 1968-2020

(a) SVAR-IV, 2 Lags



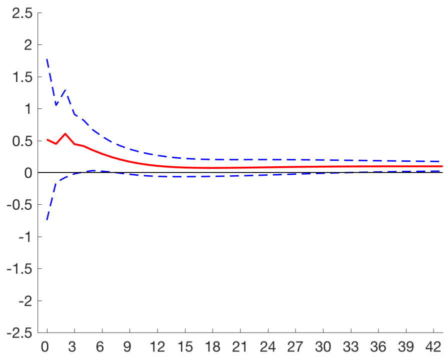
(b) SVAR-IV, 4 Lags



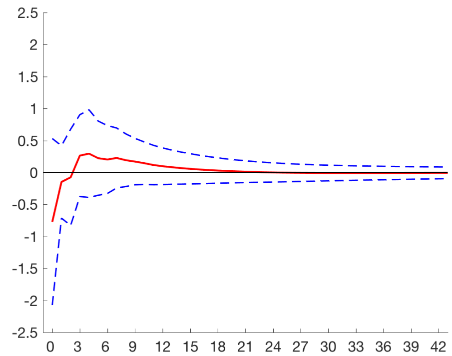
Notes: See notes for Table 3.3. The impulse response curves are plotted for the quarters following a one-time price shock. Quarters are indicated on the x-axis, while the y-axis is the change in new vehicle sales (millions, SAAR).

Figure 3.5: Response of Quarterly Sales (Millions, SAAR) to a $1,000 One-Time Permanent Price Increase, SVAR-IV, Instrument: Hot-Rolled Steel Imports Price from 1989-2020

(a) SVAR-IV, 2 Lags



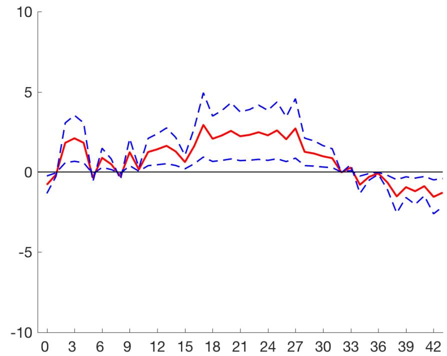
(b) SVAR-IV, 4 Lags



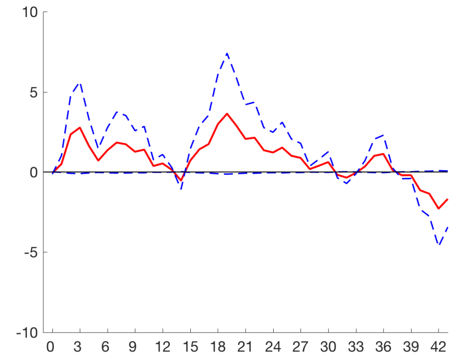
Notes: See notes for Table 3.3. The impulse response curves are plotted for the quarters following a one-time price shock. Quarters are indicated on the x-axis, while the y-axis is the change in new vehicle sales (millions, SAAR).

Figure 3.6: Response of Quarterly Sales (Millions, SAAR) to a $1,000 One-Time Permanent Price Increase, LP-IV, 2 Lags

(a) Instrument: Domestic Hot-Rolled Steel Price Index from 1968-2020



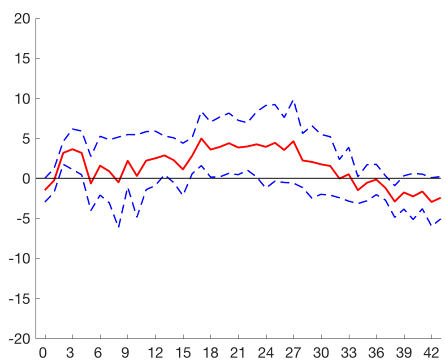
(b) Instrument: Hot-Rolled Steel Imports Price from 1989-2020



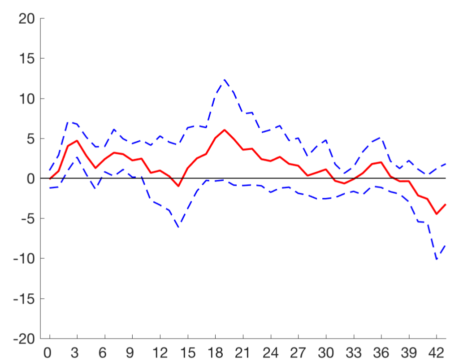
Notes: See notes for Table 3.3. The impulse response curves are plotted for the quarters following a one-time price shock. Quarters are indicated on the x-axis, while the y-axis is the change in new vehicle sales (millions, SAAR).

Figure 3.7: Response of Log Quarterly Sales to a One-Time One Percent Permanent Price Increase (i.e., Elasticities), LP-IV, 2 Lags

(a) Instrument: Domestic Hot-Rolled Steel Price Index from 1968-2020



(b) Instrument: Hot-Rolled Steel Imports Price from 1989-2020



Notes: See notes for Table 3.3. The impulse response curves are plotted for the quarters following a one-time price shock. Quarters are indicated on the x-axis, while the y-axis is the change in log new vehicle sales.

3.6 Conclusion

Using novel applications of SVAR-IV and LP-IV, we have calculated impulse response functions and short-run (first-year) price elasticities for new vehicle sales. Our most preferred specification (a 2-lag SVAR-IV) estimates a first-year price elasticity of -0.41. However, the results are highly unstable, so we are unable to have confidence in any precise elasticity. Hot-rolled steel price indices are weak instruments for new vehicle prices, and the impulse response functions and elasticities we calculate are highly sensitive to the modeling assumptions. Therefore, future work should seek stronger instruments, more stable specifications, and more precise impulse response and elasticity estimates. With this paper, we have discussed possible solutions to the strong assumptions of alternative uninstrumented models, and we urge caution in use of price elasticities estimated with these methods. However, we hope that the future work in this area will identify the precise elasticities that will be most useful to policymakers.

3.7 References

Allcott, Hunt and Nathan Wozny (2014) "Gasoline prices, fuel economy, and the energy paradox," *Review of Economics and Statistics*, 96 (5), 779–795, 10.1162/REST_a_00419.

Berry, Steven, James Levinsohn, and Ariel Pakes (1995) "Automobile prices in market equilibrium," *Econometrica*, 63 (4), 841–890, 10.2307/2171802.

——— (2004) "Differentiated products demand systems from a combination of micro and macro data: The new car market," *Journal of Political Economy*, 112 (1), 68–105, 10.1086/379939.

Berry, Steven T. (1994) "Estimating discrete-choice models of product differentiation," *The RAND Journal of Economics*, 25 (2), 242–262, 10.2307/2555829.

Busse, Meghan R., Christopher R. Knittel, and Florian Zettelmeyer (2013) "Are consumers myopic? Evidence from new and used car purchases," *American Economic Review*, 103 (1), 220–56, 10.1257/aer.103.1.220.

Dou, Xiaoya and Joshua Linn (2020) "How do US passenger vehicle fuel economy standards affect new vehicle purchases?," *Journal of Environmental Economics and Management*, 102, 10.1016/j.jeem.2020.102332.

Engle, Robert F. and C. W. J. Granger (1987) "Co-integration and error correction: Representation, estimation, and testing," *Econometrica*, 55 (2), 251–276, 10.2307/1913236.

Fischer, Carolyn, Winston Harrington, and Ian W.H. Parry (2007) "Should automobile fuel economy standards be tightened?," *Energy Journal*, 28 (4), 1–29, 10.5547/ISSN0195-6574-EJ-Vol28-No4-1.

Gertler, Mark and Peter Karadi (2015) "Monetary policy surprises, credit costs, and economic activity," *American Economic Journal: Macroeconomics*, 7 (1), 44–76, 10.1257/mac.20130329.

Griliches, Zvi (1971) "Hedonic prices revisited," in Griliches, Zvi ed. *Price Indexes and Quality Change*: Harvard University Press, https://www.hup.harvard.edu/catalog.php?isbn=9780674592582.

Gron, Anne and Deborah L. Swenson (1996) "Incomplete exchange-rate pass-through and imperfect competition: The effect of local production," *American Economic Review, Papers and Proceedings*, 86 (2), 71–76, 10.2307/2118099.

——— (2000) "Cost pass-through in the U.S. automobile market," *Review of Economics and Statistics*, 82 (2), 316–324, 10.1162/003465300558704.

Gruenspecht, H. K. (1982) "Differentiated regulation: The case of auto emissions standards," *American Economic Review*, 72 (2), 328–331, 10.2307/1802352.

Hess, Alan C. (1977) "A comparison of automobile demand equations," *Econometrica*, 45 (3), 683–702.

Hymans, Saul H., Gardner Ackley, and F. Thomas Juster (1970) "Consumer durable spending: Explanation and prediction," *Brookings Papers on Economic Activity*, 1970 (2), 173–206, 10.2307/2534239.

Jordà, Òscar (2005) "Estimation and inference of impulse responses by local projections," *American Economic Review*, 95 (1), 161–182, 10.1257/0002828053828518.

Jordà, Òscar, Moritz Schularick, and Alan M. Taylor (2015) "Betting the house," *Journal of International Economics*, 96 (Supplement 1), S2–S18, 10.1016/j.jinteco.2014.12.011.

——— (2020) "Estimating consumer substitution between new and used passenger vehicles," *Resources for the Future Working Paper 19-01*.

Leard, Benjamin, Joshua Linn, and Yichen Christy Zhou (2017b) "How much do consumers value fuel economy and performance?,"Technical report, Resources for the Future.

Levinsohn, James (1988) "Empirics of taxes on differentiated products: The case of tariffs in the U.S. automobile industry," in *Trade Policy Issues and Empirical Analysis*, 9–44 BT – Trade Policy Issues and Empirical Analy: National Bureau of Economic Research, Inc, `https://econpapers.repec.org/RePEc:nbr:nberch:5846`.

McAlinden, Sean, Yen Chen, Michael Schultz, and David Andrea (2016) "The Potential Effects of the 2017-2025 EPA/NHTSA GHG/Fuel Economy Mandates on the U.S. Economy,"Technical report, Center for Automotive Research (CAR).

McCarthy, Patrick S. (1996) "Market price and income elasticities of new vehicle demands," *The Review of Economics and Statistics*, 10.2307/2109802.

National Research Council (2015) *Cost, effectiveness, and deployment of fuel economy technologies for light-duty vehicles*, Washington D.C.: National Academies Press, 10.17226/21744.

Nerlove, Marc (1957) "A note on long-run automobile demand," *Journal of Marketing*, 10.1109/EPE.2014.6839415.

Ohta, Makoto and Zvi Griliches (1976) "Automobile prices revisited: Extensions of the hedonic hypothesis," in Terleckyj, Nestor E. ed. *Household Production and Consumption*, p. 325 – 398: NBER, `http://www.nber.org/books/terl76-1`.

——— (1986) "Automobile prices and quality: Did the gasoline price increases change consumer tastes in the U.S.?," *Journal of Business and Economic Statistics*, 4 (2), 187–198, 10.1080/07350015.1986.10509514.

Olea, José Luis Montiel and Carolin Pflueger (2013) "A robust test for weak instruments," *Journal of Business and Economic Statistics*, 31 (3), 358–369, 10.1080/00401706.2013.806694.

Plagborg-Møller, Mikkel and Christian K Wolf (2021) "Local projections and VARs estimate the same impulse responses," *Econometrica, Forthcoming*.

Sallee, James M., Sarah E. West, and Wei Fan (2016) "Do consumers recognize the value of fuel economy? Evidence from used car prices and gasoline price fluctuations," *Journal of Public Economics*, 135, 61–73, 10.1016/j.jpubeco.2016.01.003.

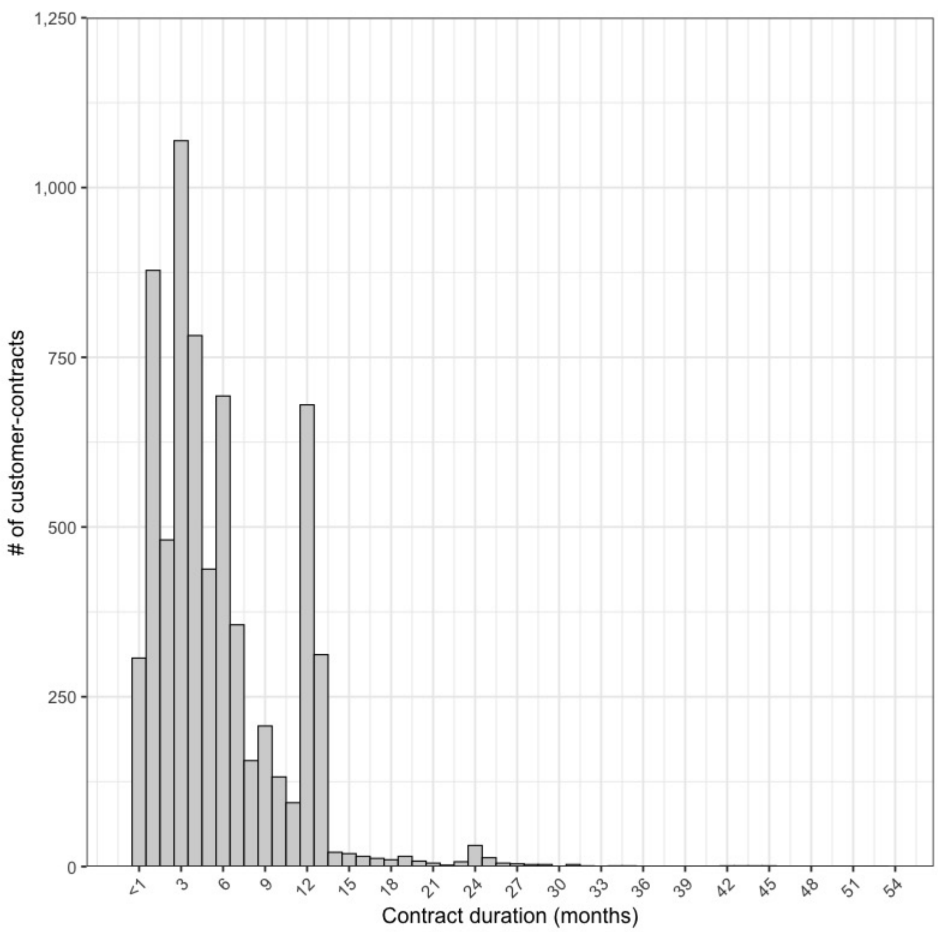
Sims, Christopher A. (1980) "Macroeconomics and reality," *Econometrica*, 48 (1), 10.2307/1912017.

Stock, James H. (2008) "What's new in econometrics: time series, lecture 7. Short course lectures," `https://www2.nber.org/minicourse{_}2008.html`.

——— (2018) "Identification and estimation of dynamic causal effects in macroeconomics using external instruments," *The Economic Journal*, 128 (610), 917–948, 10.1111/ecoj.12593.

Appendix A: Retail Choice: Additional Tables and Figures

Figure A.1: Distribution of Actual Contract Durations at Retailer A



Notes: (1) This histogram shows actual contract durations, which is the time a consumer spent on a particular contract. These are not the contract terms, which are shown in Figure 1.3. (3) Data includes contracts that are still ongoing, so the distribution is biased downwards. (4) Typically, Retailer A moves consumers to a default month-to-month contract shortly after the expiration of their termed contract if the consumer does not make an alternate selection. Month-to-month contracts may last indefinitely.

Table A.1: Consumer Switching Behavior at Retailer A

		# of customers	Share of customers
Retention and switching between retailers			
[1]	Left Retailer A by end of sample	2,428	57.7%
[2]	Left Retailer A and then returned	39	0.9%
[3]	Still on Retailer A at end of sample	1,791	42.6%
Retention and switching on multiple-month contracts			
[4]	Selected multiple-month contract	3,795	90.2%
[5]	Still on first contract at end of sample	1,099	26.1%
[6]	Left contract early to switch contracts with Retailer A	128	3.0%
[7]	Left contract early to leave Retailer A	579	13.8%
Behavior at end of multiple-month contracts			
[8]	Stayed full contract term	2,187	52.0%
[9]	Stayed full contract term and then left Retailer A	899	21.4%
[10]	Stayed full contract term and then stayed with Retailer A	1,385	32.9%
[11]	Stayed full contract term and had not made next choice by end of sample	907	21.6%
Contract choice after multiple-month contract if staying with Retailer A			
[12]	Switched to default month-to-month contract	1,222	29.0%
[13]	Switched to another contract	765	18.2%
[14]	Switched to another month-to-month contract	61	1.4%
Behavior after time on the default month-to-month contract			
[15]	Eventually switched off default contract to another contract with Retailer A	187	4.4%
[16]	Eventually left Retailer A after time on default contract	859	20.4%
[17]	Stayed on default for less than two months before switching or leaving	565	13.4%
[18]	Stayed on default for more than two months	669	15.9%
[19]	Still on default contract at end of sample	218	5.2%
Switching to and from time-varying contracts			
[20]	Switched from non-TOU to TOU	167	4.0%
[21]	Switched from TOU to non-TOU	12	0.3%
[22]	**Total**	4,208	

Notes continued on next page.

Notes for Table A.1: (1) Rows are not mutually exclusive because consumers can have complex switching patterns. For example, the same consumer could end a multiple-month contract and then switch to another multiple-month contract in one instance and then be switched to the default month-to-month contract upon the expiration of that second contract. (2) The default month-to-month contracts are the contracts consumers are most frequently switched to if they do not make an active contract selection at the end of their previous contract term. Consumers may actively choose these contracts, but this is rare. For example, only 15 consumers select a default rate as their first contract. (3) The median time on a default plan is 70 days, but this is biased downwards since 5 percent of consumers are still on a default rate.

Table A.2: Most Popular Contract Features at Retailer A

Contract feature	First contract at Retailer A		Subsequent contract	
	# of customers	Share of customers	# of customers	Share of customers
Not time-varying	4,049	96%	1,674	92%
Time-varying	159	4%	200	11%
Free weekends	98	2%	114	6%
Free nights	61	1%	94	5%
Green	1,018	24%	337	19%
Contract term				
Month-to-month (variable rate)	435	10%	1,408	78%
3 months	730	17%	264	15%
6 months	560	13%	239	13%
9 months	230	5%	109	6%
12 months	1,923	46%	372	20%
18 months	28	1%	14	1%
24 months	164	4%	75	4%
36 months	138	3%	79	4%
Total	4,208		1,816	

Notes: (1) Rows are not mutually exclusive because contracts may have multiple features and consumers may have been with Retailer A long enough to select multiple successive contracts. (3) The high share of consumers switching to monthly contracts is because many consumers are defaulted onto these contracts when they fail to sign a new contract upon the expiration of their existing contract. Consumer experiences with the default contract are described in Table A.1.

Table A.3: Contract Features of Ex Post Cost-Minimizing Versus Observed Contract Choices, $\beta = 1$

	Observed contract choices				Cost-minimizing contracts				Δ from observed to cost-minimizing		
Contract feature	Share of invoices	Share of customers	Median monthly kWh	Median monthly bill	Share of invoices	Share of customers	Median monthly kWh	Median monthly bill	Δ share of invoices	Δ share of customers	Δ median monthly kWh
Not time-varying	95%	97%	906	$87	89%	94%	918	$57	-6%	-3%	11
Time-varying	5%	8%	959	$95	11%	25%	768	$57	6%	17%	-191
Free weekends	3%	5%	956	$97	6%	15%	665	$55	2%	10%	-291
Free nights	2%	4%	976	$91	5%	11%	785	$56	3%	8%	-192
Green	21%	25%	993	$92	34%	65%	849	$52	13%	40%	-145
Default	15%	29%	782	$109	0%	0%	498	$104	-15%	-29%	-284
Contract term											
Month-to-month	17%	31%	769	$103	15%	63%	902	$77	-2%	32%	132
3 months	8%	20%	784	$67	27%	48%	860	$47	19%	28%	76
6 months	11%	16%	796	$71	11%	21%	763	$44	0%	5%	-33
9 months	5%	8%	1,014	$99	25%	37%	1,071	$70	20%	30%	57
12 months	48%	49%	951	$88	16%	26%	874	$45	-32%	-23%	-77
18 months	1%	1%	1,210	$121	2%	6%	1,317	$80	1%	5%	107
24 months	6%	6%	1,255	$127	3%	4%	1,103	$64	-3%	-2%	-152
36 months	4%	5%	1,511	$145	1%	2%	567	$48	-3%	-3%	-944
Sample summary											
	# of invoices	# of customers	Median monthly kWh	Median monthly bill	# of invoices	# of customers	Median monthly kWh	Median monthly bill			
	41,356	4,208	918	$88	41,356	4,208	918	$57			

Notes: (1) This table compares features of consumers' chosen contracts to my model estimates of their cost-minimizing contracts. (2) These results assume consumers had perfect information, which is equivalent to the ex post analysis. I also assume no discounting of the future. These assumptions are relaxed in Appendix Tables A.3 to A.6. (3) See notes for Figure 1.3.

Table A.4: Contract Features of Ex Post Cost-Minimizing Versus Observed Contract Choices, $\beta = 0.95$

	Observed contract choices				Cost-minimizing contracts				Δ from observed to cost-minimizing		
Contract feature	Share of invoices	Share of customers	Median monthly kWh	Median monthly bill	Share of invoices	Share of customers	Median monthly kWh	Median monthly bill	Δ share of invoices	Δ share of customers	Δ median monthly kWh
Not time-varying	95%	97%	906	$87	89%	94%	918	$57	-6%	-3%	11
Time-varying	5%	8%	959	$95	11%	25%	769	$57	6%	17%	-190
Free weekends	3%	5%	956	$97	6%	15%	665	$55	2%	10%	-291
Free nights	2%	4%	976	$91	5%	11%	785	$56	3%	8%	-192
Green	21%	25%	993	$92	34%	64%	851	$52	13%	39%	-142
Default	15%	29%	782	$109	0%	0%	411	$102	-15%	-29%	-371
Contract term											
Month-to-month	17%	31%	769	$103	15%	63%	901	$77	-2%	32%	131
3 months	8%	20%	784	$67	27%	48%	860	$47	18%	28%	76
6 months	11%	16%	796	$71	11%	21%	760	$44	0%	5%	-37
9 months	5%	8%	1,014	$99	25%	37%	1,075	$70	20%	29%	61
12 months	48%	49%	951	$88	16%	26%	874	$45	-32%	-22%	-77
18 months	1%	1%	1,210	$121	2%	6%	1,317	$80	1%	5%	107
24 months	6%	6%	1,255	$127	3%	4%	1,103	$64	-3%	-2%	-152
36 months	4%	5%	1,511	$145	1%	2%	567	$48	-3%	-3%	-944
Sample summary											
	# of invoices	# of customers	Median monthly kWh	Median monthly bill	# of invoices	# of customers	Median monthly kWh	Median monthly bill			
	41,356	4,208	918	$88	41,356	4,208	918	$57			

Notes: (1) This table compares features of consumers' chosen contracts to my model estimates of their cost-minimizing contracts. (2) These results assume consumers had perfect information, which is equivalent to the ex post analysis. In contrast to Appendix Tables A.3 to A.6, I assume that consumers discount their future savings. (3) See notes for Figure 1.3.

Table A.5: Contract Features of Imperfect Information Cost-Minimizing Versus Observed Contract Choices, $\beta = 1$

	Observed contract choices				Cost-minimizing contracts				Δ from observed to cost-minimizing		
Contract feature	Share of invoices	Share of customers	Median monthly kWh	Median monthly bill	Share of invoices	Share of customers	Median monthly kWh	Median monthly bill	Δ share of invoices	Δ share of customers	Δ median monthly kWh
Not time-varying	95%	97%	906	$87	89%	94%	930	$59	-6%	-2%	23
Time-varying	5%	8%	959	$95	11%	27%	780	$61	6%	20%	-179
Free weekends	3%	5%	956	$97	6%	16%	628	$53	3%	11%	-328
Free nights	2%	4%	976	$91	5%	12%	862	$64	3%	9%	-114
Green	21%	25%	993	$92	36%	72%	915	$61	15%	47%	-78
Default	15%	29%	782	$109	0%	0%	1,163	$239	-15%	-29%	381
Contract term											
Month-to-month	17%	31%	769	$103	24%	73%	1,075	$86	7%	41%	306
3 months	8%	20%	784	$67	28%	49%	808	$47	19%	30%	24
6 months	11%	16%	796	$71	12%	23%	803	$51	1%	7%	7
9 months	5%	8%	1,014	$99	19%	34%	959	$67	14%	26%	-55
12 months	48%	49%	951	$88	13%	25%	899	$50	-35%	-23%	-52
18 months	1%	1%	1,210	$121	2%	7%	1,296	$79	1%	6%	86
24 months	6%	6%	1,255	$127	2%	3%	1,109	$61	-4%	-3%	-147
36 months	4%	5%	1,511	$145	1%	3%	493	$41	-3%	-3%	-1,018
Sample summary											
	# of invoices	# of customers	Median monthly kWh	Median monthly bill	# of invoices	# of customers	Median monthly kWh	Median monthly bill			
	41,356	4,208	918	$88	41,356	4,208	918	$58			

(1) This table compares features of consumers' chosen contracts to my model estimates of their cost-minimizing contracts. (2) In contrast to Appendix Tables A.3 to A.6, these results feature an imperfect information case where consumers choose cost-minimizing contracts believing that their choice set will remain the same in the future. I assume no discounting of the future. (3) See notes for Figure 1.3.

Table A.6: Contract Features of Ex Post Cost-Minimizing Versus Observed Contract Choices, Expanded Choice Set, $\beta = 1$

	Observed contract choices				Cost-minimizing contracts				Δ from observed to cost-minimizing		
Contract feature	Share of invoices	Share of customers	Median monthly kWh	Median monthly bill	Share of invoices	Share of customers	Median monthly kWh	Median monthly bill	Δ share of invoices	Δ share of customers	Δ median monthly kWh
Not time-varying	95%	97%	906	$87	92%	95%	927	$49	-3%	-1%	20
Time-varying	5%	8%	959	$95	8%	20%	471	$37	3%	12%	-488
Free weekends	3%	5%	956	$97	5%	13%	414	$34	1%	8%	-542
Free nights	2%	4%	976	$91	4%	8%	568	$38	2%	5%	-408
Green	21%	25%	993	$92	45%	78%	846	$44	24%	53%	-147
Default	15%	29%	782	$109	0%	0%	0	$77	-15%	-29%	-782
Contract term											
Month-to-month	17%	31%	769	$103	15%	64%	872	$72	-1%	33%	103
3 months	8%	20%	784	$67	33%	54%	891	$42	24%	35%	107
6 months	11%	16%	796	$71	9%	18%	690	$37	-2%	2%	-106
9 months	5%	8%	1,014	$99	22%	37%	850	$39	17%	29%	-165
12 months	48%	49%	951	$88	11%	13%	657	$32	-38%	-36%	-294
18 months	1%	1%	1,210	$121	2%	4%	1,490	$94	1%	3%	280
24 months	6%	6%	1,255	$127	8%	20%	1,263	$66	2%	14%	7
36 months	4%	5%	1,511	$145	0%	1%	397	$34	-3%	-4%	-1,113
Sample summary											
	# of invoices	# of customers	Median monthly kWh	Median monthly bill	# of invoices	# of customers	Median monthly kWh	Median monthly bill			
	41,356	4,208	918	$88	41,356	4,208	918	$49			

Notes: (1) This table compares features of consumers' chosen contracts to my model estimates of their cost-minimizing contracts. (2) These results assume consumers had perfect information, which is equivalent to the ex post analysis. I also assume no discounting of the future. In contrast to Appendix Tables A.3 to A.5, I expand the choice set to all contracts on Retailer A's books in a given period. (3) See notes for Figure 1.3.

Appendix B: Nonlinear Weather and Climate: Data Appendix

D&G Data:

To estimate equation (2.20), D&G use county-level data on agricultural production, temperature, precipitation, and soil quality in the United States. Agricultural data come from Agricultural Census of 1987, 1992, 1997, and 2002. The main dependent variable is agricultural profits per acre of farmland.[1] Soil quality data was collected from the National Resource Inventory (NRI), which has data on 800,000 sites in census years. Rainfed counties are those with less than 10 percent of irrigated farmland. Weather data come from two sources: (i) monthly data are collected by the Parameter-Elevation Regressions on Independent Slopes Model (PRISM), that generates estimates of precipitation and temperature for the entire U.S.; (ii) daily data are collected by the National Climatic Data Center (NCDC) Summary of the Day Data. Their final sample is a balanced panel of 2,262 counties for a total 9,048 county-year observations. Table B.1 presents the main statistics of the D&G data (mean and standard deviation) over the full sample.

[1] The variable was calculated by the ratio between the market value of agricultural products sold minus total production expenses across all farms over the acres devoted to crops, pasture, and grazing.

Table B.1: County-level summary statistics, D&G data

	Mean	Std. Dev.
Profits per acre of farmland (2002 values)	41.92	74.82
Salinity	0.03	0.07
Fraction flood-prone	0.16	0.21
Wetlands	0.06	0.09
K-factor	0.31	0.07
Slope length	264.21	209.68
Fraction sand	0.05	0.16
Fraction clay	0.24	0.27
Moisture capacity	0.17	0.03
Permeability	2.34	2.29
Cumulative 100 GDD	38.21	11.12
Cumulative 100 GDD, rainfed	38.04	11.12
Cumulative 100 GDD, irrigated	38.94	11.11
Growing season precipitation (in inches)	15.81	7.81
Growing season precipitation (in inches), rainfed	16.52	7.26
Growing season precipitation (in inches), irrigated	12.87	9.20

Notes: Data from D&G replication files. Summary statistics calculated using the estimating sample and are weighted by acres of farmland. D&G's cumulative GDD variable (growing season degree-days, defined with a base of 46.4°F and ceiling of 89.6°F) was multiplied by 100 to reflect change in 100 GDD.

BHM Data:

To estimate equation (2.21), BHM use data on economic production and weather for 166 countries in the period between 1960 and 2010. The dependent variable is the change in annual GDP per capita. The variable *Poor* is an indicator for whether or not the country is below the global median GDP per capita of the sample. Weather data comes from the University of Delaware, which put together data from stations, both from the GHCN2 (Global Historical Climate Network) and from the archive of Legates & Willmott, spanning 1900 to 2010.[2] Their final sample is an unbalanced panel of 166 countries for a total 6.584 country-year observations. Table B.2 presents the main statistics of the BHM data (mean and standard deviation) over the full sample.

[2] See the website `https://psl.noaa.gov/data/gridded/data.UDel_AirT_Precip.html` for more details.

Table B.2: Country-level summary statistics, BHM data

	Mean	Std. Dev.
Annual percent change in GDP per capita (in %)	1.81	6.15
Percentage of poor countries	52.42	49.95
Avg. temp. in $^{\circ}C$	18.92	7.46
Avg. temp. in $^{\circ}C$, poor countries	22.72	5.16
Avg. temp. in $^{\circ}C$, rich countries	15.06	7.23
Avg. prp. in meters	1.16	0.74
Avg. prp. in meters, poor countries	1.32	0.75
Avg. prp. in meters, rich countries	1.00	0.70
Avg. number of years by country	39.58	

Notes: Data from BHM replication files. Averages are calculated for the estimating sample (unbalanced sample of 166 countries over 1960-2010). BHM's GDP growth variable was multiplied by 100 to reflect percent change in GDP per capita. "Avg. number of years by country" denotes the average total years that a country is observed in the sample. 86 countries have balanced panel data over the period (1960-2010).

Appendix C: Nonlinear Weather and Climate: Accounting for Multiple Fixed Effects

Our analysis has focused on the case of a single set of fixed effects, county fixed effects in D&G and country fixed effects in BHM. However, nonlinear terms may also arise in settings with multiple fixed effects. In fact, both the D&G and BHM specifications also include year fixed effects. D&G also include a fixed effect for whether a county is irrigated or dry. Recall from our notation that the deviation of a variable $V_{i,t}$ in location i from the location mean over time is:

$$v_{i,t} \coloneqq V_{i,t} - \bar{V}_i \quad \text{where} \quad \bar{V}_i = E_t[V_{i,t}|\gamma_i].$$

Thus:

$$v_{i,t}^2 = V_{i,t}^2 - 2\bar{V}_i V_{i,t} + E_t[V_{i,t}^2|\gamma_i].$$

Therefore, we argue that the nonlinear transformations should be applied after demeaning variables by their conditional expectations, thus fully removing the conditional mean from affecting estimation. In the case of multiple fixed effects, we can demean variables by:

$$v_{i,t} \coloneqq V_{i,t} - E[V_{i,t}|\phi_{i,t}]$$

where $\phi_{i,t}$ is the set of fixed effects. Thus, we obtain $v_{i,t}$ from the residuals of the OLS regression:

$$V_{i,t} = \phi_{i,t} + \varepsilon_{i,t}.$$

So in the case of both location fixed effects, γ_i, and year fixed effects, α_t, we can demean variables according to:

$$v_{i,t} \coloneqq V_{i,t} - E[V_{i,t}|\gamma_i, \alpha_t], \tag{C.1}$$

using the OLS regression:

$$V_{i,t} = \gamma_i + \alpha_t + \varepsilon_{i,t}.$$

Up to this point, our discussion has only concerned the demeaning variables with respect to their expectations conditional on fixed effects, a particular class of binary covariates. But there may be some contexts where it is desirable to demean each variable V by *all* of the other covariates $\mathbf{X}_{-z}$ as in:

$$v_{i,t} \coloneqq V_{i,t} - E[V_{i,t}|\phi_{i,t}, \mathbf{X}_{i,t,-z}].$$

While we do not estimate a specification with this type of demeaning, Table C.1 does show a version of the results where we demean each variable by its expectation conditional on both year and county fixed effects as in equation (C.1).

Column 1 of Table C.1. reproduces the results of the Table 2.1 misspecification (i.e., standard fixed effects as in equation 2.16). Column 2 shows that centering variables around county and year conditional expectations after applying the nonlinear transformations yields identical results (i.e., equations 2.16 and 2.18). Then, column 3 shows the corrected results from Table 2.1 where variables are demeaned by expectations conditional on county means before applying the quadratic (equation 2.19). Finally, column 4 shows the corrected results (equation C.1) where variables are demeaned by expectations conditional on both county and year means before applying the quadratic.

The similarity of columns 3 and 4 suggests that centering with respect to county is more important than centering with respect to year in correcting the bias of equation (2.16). There is one intuitive reason that centering variables around county means is most important. The spatial variation in the data is much greater than the interannual variation. Controlling for county to county differences has a much bigger effect than controlling for national differences from year to year. The purpose of the D&G and BHM models is to use random weather fluctuations to predict impacts from climate changes. Given the specification of equation (2.19), a given change in climate should have the same impact across every climate. In the misspecification of equation (2.16) (or 2.18), the impact varies across climates (county mean weather). Table 2.1 shows that whether we include county mean weather (as a result of the misspecification of equation 2.19) or not (equation 2.16) has a substantial impact on the estimates. In contrast, year fixed effects capture the deviation of U.S.-wide weather from the U.S.-wide mean over time. Table C.1 shows that accounting for U.S.-wide weather trends is just not as important in this model as accounting for location-specific climate.

Table C.1: Account for Multiple Fixed Effects: Biased and Corrected Fixed Effect Coefficients of Temperature and Precipitation

Specification:	[1] Biased Linear Coefficient [Eq. 2.16 or 2.18]	[2] Biased Quadratic Coefficient [Eq. 2.16 or 2.18]	[3] Correct Linear Coefficient [Eq. 2.19]	[4] Correct Quadratic Coefficient [Eq. 2.19 and C.1]
Demeaning:	Standard FE	Conditioning on location and year FEs	Conditioning on location FEs only	Conditioning on location and year FEs
BHM, Percent GDP/capita, (°C)				
Linear temperature coefficient	1.272*** (0.379)	1.272*** (0.378)	-0.100 (0.206)	-0.100 (0.207)
Quadratic temperature coefficient	-0.049*** (0.012)	-0.049*** (0.012)	-0.211 (0.131)	-0.204 (0.179)
D&G, Farm profit/acre, (100 growing degree days °F)				
Linear temperature coefficient	-1.453** (0.594)	-1.453** (0.594)	-1.279*** (0.301)	-1.220*** (0.282)
Quadratic temperature coefficient	0.002 (0.008)	0.002 (0.008)	0.192*** (0.047)	0.236** (0.100)

Notes: Standard errors in parentheses. In columns 3 and 4, we adjust standard errors to account for the fact that these are two-stage estimates (i.e., variables are demeaned before estimation). $*p < 0.1$, $**p < 0.05$, $***p < 0.01$. *Column 1 contains the biased estimates from equation (2.16) where coefficients do not vary by poor/rich (or rainfed/irrigated). Column 2 shows the results of conditioning terms on country (county) and year fixed effects using equation (C.1) and then estimating equation (2.18). Column 2 estimates are identical to the biased estimates of column 1. Column 3 shows the corrected results of conditioning terms on country (county) fixed effects and then estimating equation (2.16). Column 4 shows the corrected results of conditioning terms on country (county) and year fixed effects using equation (C.1) and then estimating equation (2.16). Consistent with Stata's xtreg and areg programs, we accounted for a constant when calculating the degrees of freedom for estimates based on equation (2.19), but not equation (2.16). BHM estimates were multiplied by 100 to reflect percent change in GDP per capita.*

Appendix D: Nonlinear Weather and Climate: Modifications from Original D&G Specification

In contrast to our version of D&G's specification shown in equation (2.20), D&G's actual specification is:

$$Y_{i,t} = \beta_0 + \gamma_i + \delta_t + \lambda D_{i,t} + \mathbf{X'}_{i,t}\alpha + \Sigma_{j=1}^{2}[\beta_{1,j} D_{i,t}\mathbf{W}_{j,i,t} + \beta_{2,j} D_{i,t}\mathbf{W}_{j,i,t}^{2}] + \varepsilon_{i,t} \tag{D.1}$$

where the difference is that D&G use an irrigation status $D_{i,t}$ that can vary over years, while we estimate our version of their results using an irrigation status D_i that does not vary over time. This change is necessary because an irrigation status that changes over time will bias county mean weather toward 0, thus overstating the magnitude of weather deviations and attenuating the estimated effects. Columns 1-4 of Table D.1 compare D&G's original approach (equation D.1) with our modified version (equation 2.16) using standard fixed effects. Columns 5-8 compare the estimates under our equation 2.19 correction where the quadratic transformation is applied after demeaning the variables. The estimates tend to be greater in magnitude with higher significance levels when irrigation status is held constant.

Table D.1: Comparison of Biased Coefficients of Temperature and Precipitation With Modifications from Original D&G Specification

		[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
		Biased Linear Coefficient [Eq. 2.16 or 2.18]		Biased Quadratic Coefficient [Eq. 2.16 or 2.18]		Corrected Linear Coefficient [Eq. 2.19]		Corrected Quadratic Coefficient [Eq. 2.19]	
		Original D&G	D&G Modified	Original D&G	D&G Modified	Original D&G	D&G Modified	Original D&G	D&G Modified
Temperature									
D&G Farm profit/acre (100 growing degree days °F)	All	-1.440**	-1.453**	0.001	0.002	-1.286***	-1.279***	0.276***	0.192***
		(0.597)	(0.594)	(0.008)	(0.008)	(0.300)	(0.301)	(0.065)	(0.047)
	Rainfed	-1.403***	-1.129**	0.002	-0.001	-1.228***	-1.089***	0.018	0.199***
		(0.489)	(0.447)	(0.007)	(0.007)	(0.269)	(0.214)	(0.020)	(0.039)
	Irrigated	-1.445	-4.076	-0.004	0.025	-1.680***	-2.388**	-0.015	0.037
		(2.654)	(4.414)	(0.031)	(0.052)	(0.533)	(1.082)	(0.019)	(0.226)
Precipitation									
D&G Farm profit/acre (Inches)	All	-0.600	-0.560	0.007	0.006	-0.212	-0.217	-0.047*	-0.072***
		(0.681)	(0.679)	(0.018)	(0.018)	(0.166)	(0.167)	(0.028)	(0.028)
	Rainfed	-0.706	-1.530***	0.005	0.022**	-0.473***	-0.496***	-0.023	-0.092***
		(0.568)	(0.393)	(0.013)	(0.010)	(0.141)	(0.119)	(0.018)	(0.022)
	Irrigated	-0.001	2.382	0.013	-0.029	0.534	1.351*	0.006	0.113
		(1.868)	(1.834)	(0.051)	(0.052)	(0.565)	(0.774)	(0.024)	(0.074)

Notes on next page.

Notes for Table D.1: Standard errors in parentheses. In columns 5-8, we adjust standard errors to account for the fact that these are two-stage estimates (i.e., variables are demeaned before estimation). *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. *Biased estimates (columns 1-4) reflect equations (2.16) and (2.18). Corrected estimates (columns 5-8) reflect equation (2.19) with D&G covariates and weather variables. The odd columns (1,3,5,7) reflect the original D&G estimates with an irrigation status indicator $D_{i,t}$ that can vary over time. The even columns (2,4,6,8) from Table 2.1 reflect our modification to the D&G estimates used in the body of this paper where the irrigation status indicator E_i is constant within a county.*

The All estimates do not decompose the effects by rainfed and irrigated counties. The rainfed/irrigated estimates are from the specifications where the weather coefficients vary by whether a county is rainfed/irrigated. D&G dependent variable is farm profits per farmland acre. Specification includes time-varying covariates such as soil quality, and the county observations are weighted by farmland. Standard errors are clustered by county.

Appendix E: Nonlinear Weather and Climate: Equivalence of the Estimates from Equations (2.16) and (2.18)

Columns 1 and 2 of Table E.1 verify that the estimate of the quadratic coefficient from equation 2.16 is equal to the coefficient on the sum of the terms $(2\bar{w}_i w_{i,t} + w_{i,t}^2 - E_t[w_{i,t}^2|\gamma_i])$ as shown in equation 2.18. Columns 3-5 show the coefficients for each of the individual terms $2\bar{W}_i w_{i,t}$, $\delta w_{i,t}^2$ and $-E_t[w_{i,t}^2|\gamma_i]$ when we allow them to be estimated without restriction as in equation (E.1):

$$y_{i,t} = \beta_1 w_{i,t} + \beta_2 \bar{W}_i w_{i,t} + \beta_3 w_{i,t}^2 + \beta_4 E_t[w_{i,t}^2|\gamma_i] + \mathbf{x}_{i,t}'\alpha + \varepsilon_{i,t}. \tag{E.1}$$

An interesting result is that the coefficients for $w_{i,t}^2$ and $-E_t[w_{i,t}^2|\gamma_i]$ are equal, although these terms are not collinear. Since an OLS coefficient measures the average effect of a variable, the intuition for this result is that the average effect of $w_{i,t}^2$ is equal to the average effect of the mean of $w_{i,t}^2$, which is $-E_t[w_{i,t}^2|\gamma_i]$. However, the coefficient on the term $2\bar{W}_i w_{i,t}$ is quite different from the other two, so we see that the constraint that the three terms have the same coefficient is clearly binding. There is also no economic justification for this constraint. Interestingly, the coefficients on the interaction terms $2\bar{W}_i w_{i,t}$ are quite similar to the coefficients on the quadratic terms (i.e., the constrained terms). This suggests that the estimated effect of the quadratic term in the biased equation is largely influenced by the interaction term.

Table E.1: Biased Estimates of Quadratic Temperature and Precipitation Coefficients

		[1] Biased Quadratic Coefficient [Eq. 2.16]	[2] Coefficient on $[2\bar{W}_i w_{i,t} + w_{i,t}^2 - E[w_{i,t}^2\|\gamma_i]]$ [Eq. 2.18]	[3] Coefficient on $2\bar{W}_i w_{i,t}$ [Eq. E.1]	[4] Coefficient on $w_{i,t}^2$ [Eq. E.1]	[5] Coefficient on $-E[w_{i,t}^2\|\gamma_i]$ [Eq. E.1]
Temperature						
BHM[a] Percent GDP/capita (°C)	All	-0.049*** (0.012)	-0.049*** (0.012)	-0.046*** (0.011)	-0.311 (0.218)	-0.311 (0.218)
	Poor	-0.077** (0.037)	-0.077** (0.037)	-0.073** (0.037)	-0.812* (0.440)	-0.812* (0.440)
	Rich	-0.032* (0.018)	-0.032* (0.018)	-0.030 (0.019)	-0.168 (0.273)	-0.168 (0.273)
D&G[b] Farm profit/acre (100 growing degree days °F)	All	0.002 (0.008)	0.002 (0.008)	-0.016** (0.008)	0.492*** (0.107)	0.492*** (0.107)
	Rainfed	-0.001 (0.007)	-0.001 (0.007)	-0.017** (0.007)	0.465*** (0.079)	0.465*** (0.079)
	Irrigated	0.025 (0.052)	0.025 (0.052)	0.006 (0.044)	0.378 (0.594)	0.378 (0.594)
Precipitation						
BHM[a] Percent GDP/capita (Meters)	All	-0.475* (0.255)	-0.475* (0.255)	-0.395 (0.284)	-1.223** (0.595)	-1.223** (0.595)
	Poor	-0.744** (0.366)	-0.744** (0.366)	-0.528 (0.433)	-1.975*** (0.515)	-1.975*** (0.515)
	Rich	-0.269 (0.372)	-0.269 (0.372)	-0.371 (0.385)	0.836 (1.069)	0.836 (1.069)
D&G[b] Farm profit/acre (Inches)	All	0.006 (0.018)	0.006 (0.018)	-0.002 (0.020)	-0.073* (0.039)	-0.073* (0.039)
	Rainfed	0.022** (0.010)	0.022** (0.010)	0.021** (0.011)	-0.099*** (0.031)	-0.099*** (0.031)
	Irrigated	-0.029 (0.052)	-0.029 (0.052)	-0.021 (0.052)	0.189 (0.141)	0.189 (0.141)

Notes on next page.

*Notes for Table E.1: Standard errors in parentheses. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. Column 1 reflects the BHM/D&G versions of equation (2.16), while column 2 refers to equation (2.18) with the decomposed form of the quadratic term. We use BHM/D&G covariates and weather variables. Columns 3-5 present the separate estimates of each term of the decomposed quadratic term as shown in equation (E.1).*

The All estimates do not decompose the results into poor and rich countries, or rainfed and irrigated counties. The poor/rich (rainfed/irrigated) estimates are from the specifications where the weather coefficients vary by whether a country (county) is poor/rich (rainfed/irrigated). Consistent with Stata's xtreg and areg programs, we accounted for a constant when calculating the degrees of freedom for estimates based on equations (2.16), (2.18), and (E.1).

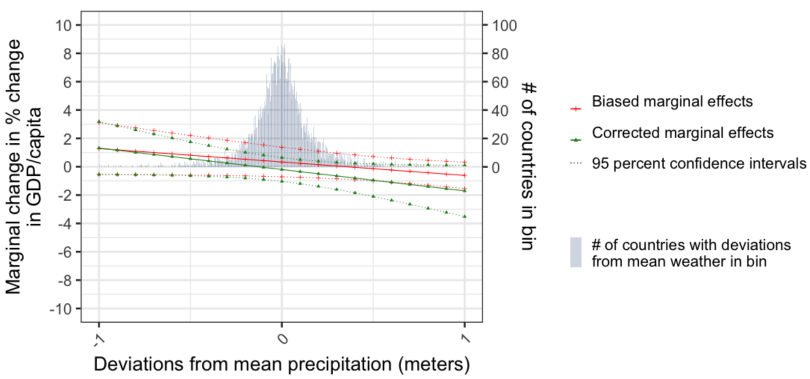
[a] *BHM dependent variable is GDP per capita growth. As shown in equation (2.21), the BHM specification includes year and country fixed effects and linear and quadratic country trends. Countries are not weighted. Standard errors are clustered by country. BHM estimates were multiplied by* 100 *to reflect percent change in GDP per capita.*

[b] *D&G dependent variable is farm profits per farmland acre. As shown in equation (2.20), the specification includes time-varying covariates such as soil quality, and county observations are weighted by farmland acreage. Standard errors are clustered by county.*

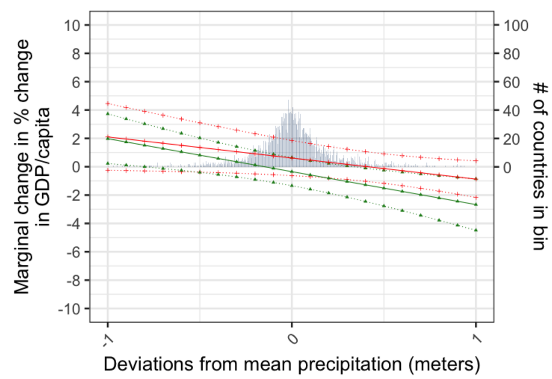
Appendix F: Nonlinear Weather and Climate: Additional Figures and Tables

Figure F.1: BHM: Marginal Effect of Precipitation on Percent GDP Per Capita

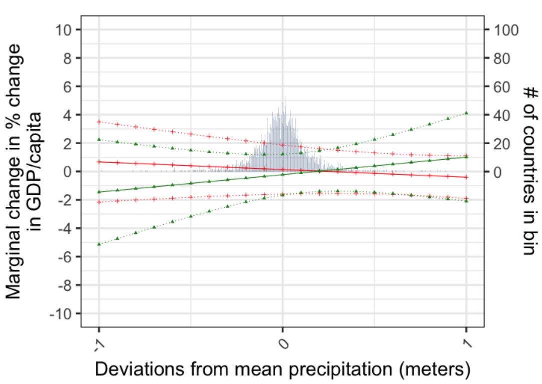
(a) All Countries



(b) Poor Countries

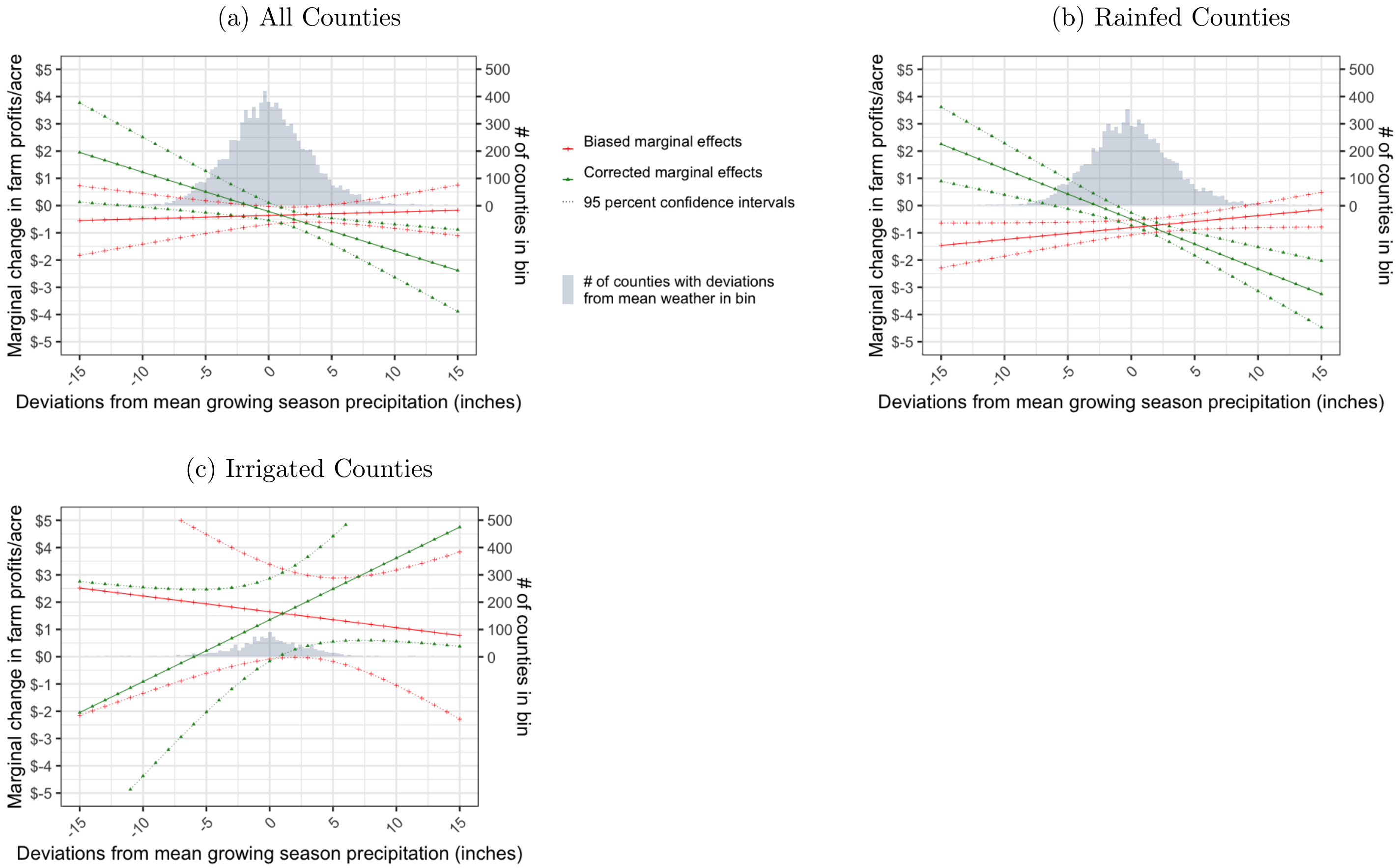


(c) Rich Countries



*Notes: Biased and corrected marginal effects reflect the BHM versions of equations (2.16) and (2.19), respectively, with BHM covariates and weather variables. Histograms show the distribution of deviations from country mean precipitations over years. Curves are estimated at sample mean precipitations (*1.2 *meters for all countries,* 1.3 *meters for poor countries, and* 1.0 *meters for rich countries). BHM estimates were multiplied by* 100 *to reflect percent change in GDP per capita.*
Panel a): The "All Countries" specification does not estimate different coefficients by poor and rich countries.

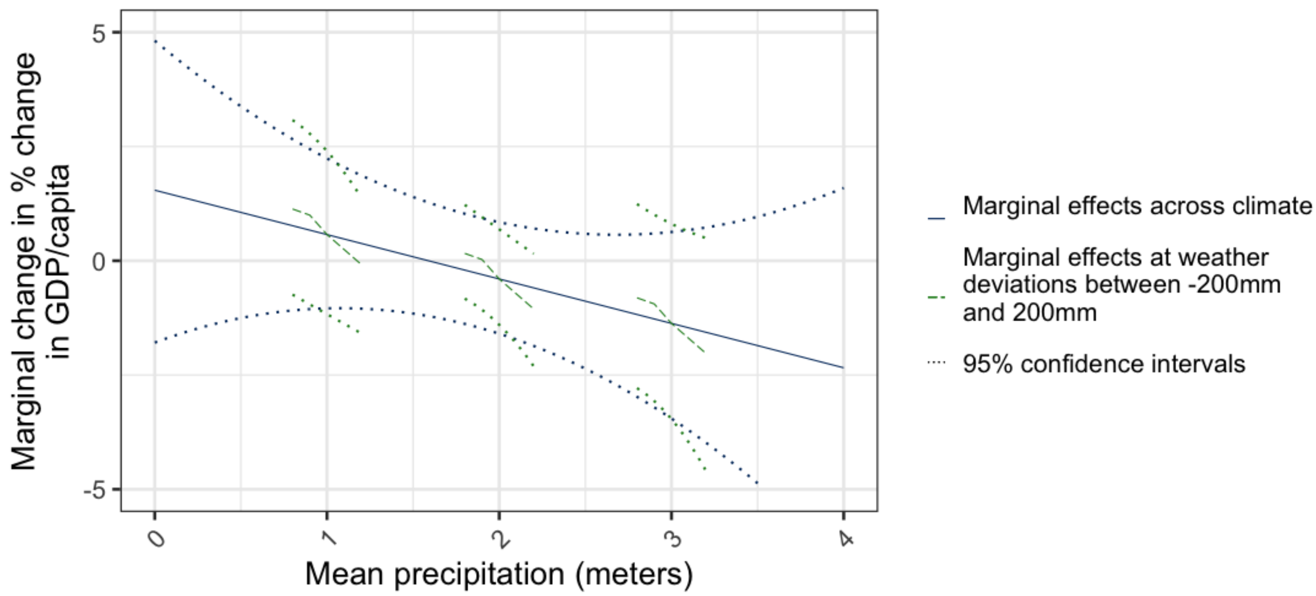
Figure F.2: D&G: Marginal Effect of Precipitation on Farm Profit Per Acre



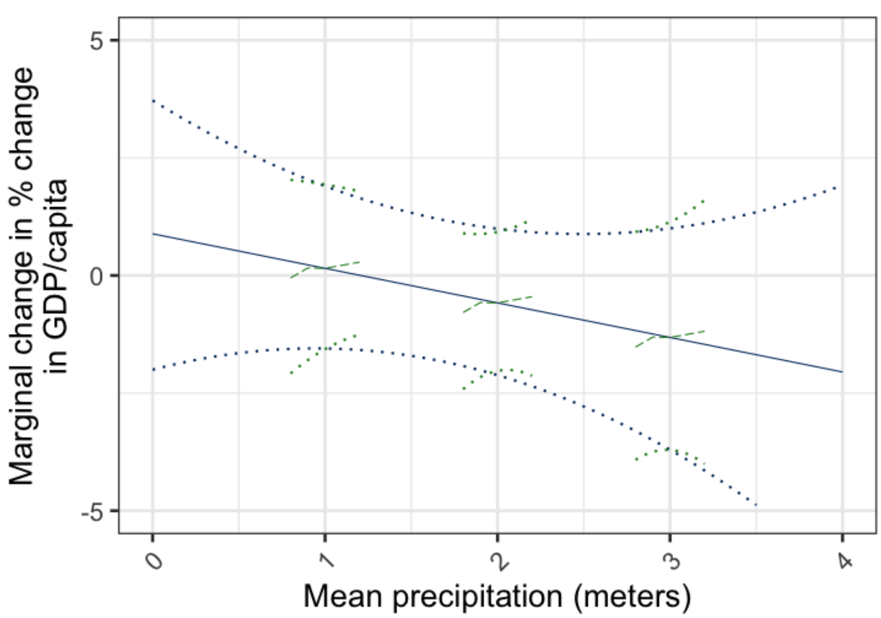
*Notes: Biased and corrected marginal effects reflect the D&G versions of equations (2.16) and (2.19), respectively, with D&G covariates and weather variables. Histograms show the distribution of deviations from county mean growing season precipitations over years. Curves are estimated at sample mean precipitations (*16.0 *inches for all counties,* 16.5 *inches for rainfed counties, and* 12.7 *inches for irrigated counties).*
Panel a): The "All Counties" specification does not estimate different coefficients by rainfed and irrigated counties.

Figure F.3: BHM: The Effect of Expected Precipitation and Precipitation Deviations on Percent GDP Per Capita

(a) Poor Countries



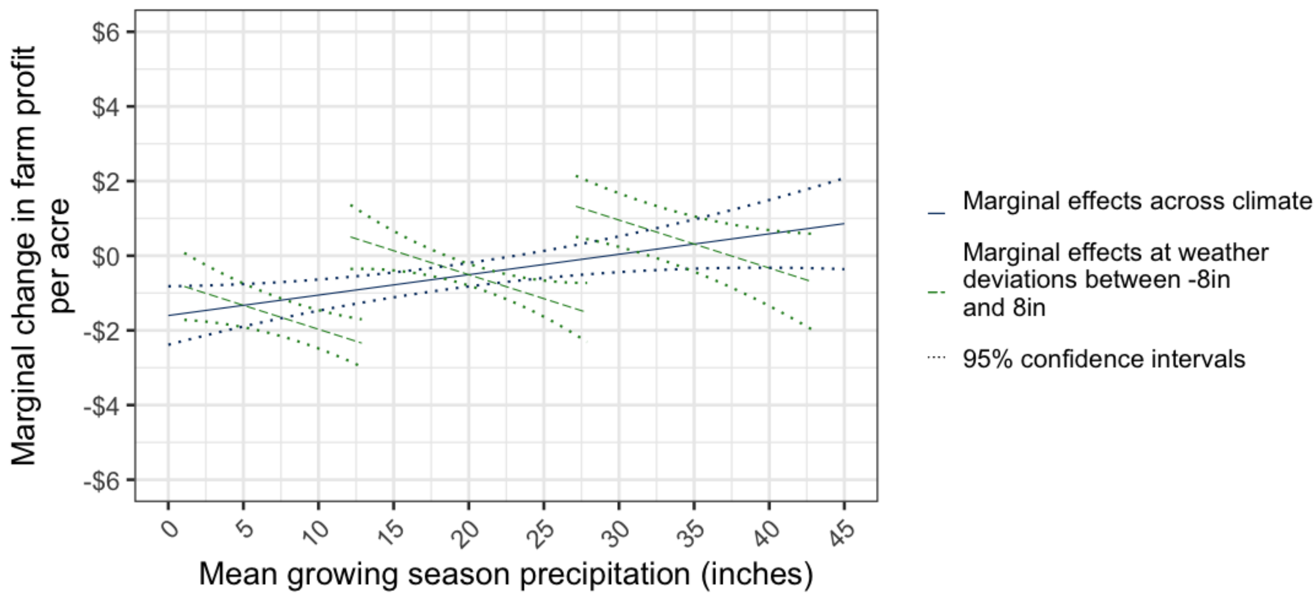
(b) Rich Countries



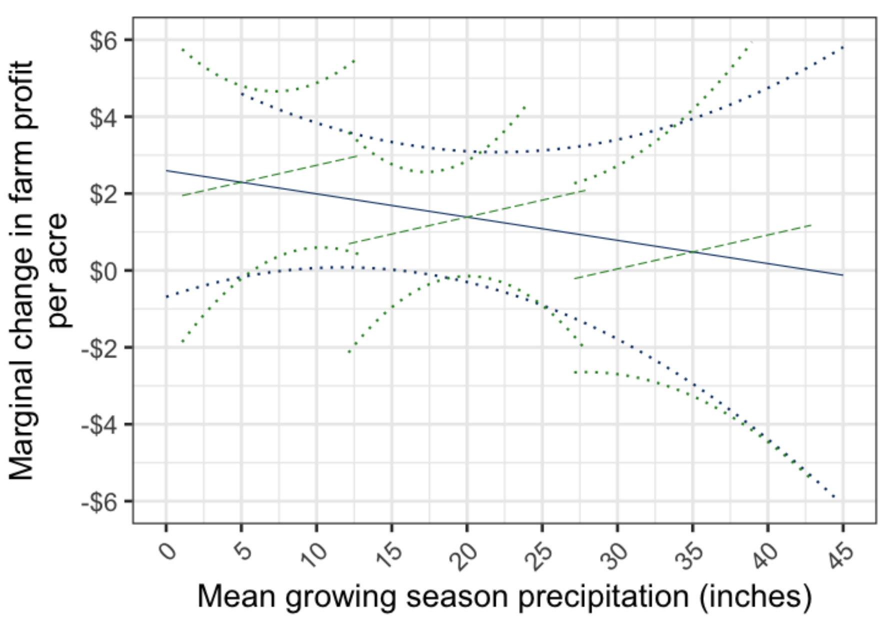
Notes: Marginal weather effects and marginal climate effects are based on coefficients in Table 2.2,. The results predict the percent change in GDP per capita.

Figure F.4: D&G: The Effect of Expected Precipitation and Precipitation Deviations on Farm Profit Per Acre

(a) Rainfed Counties



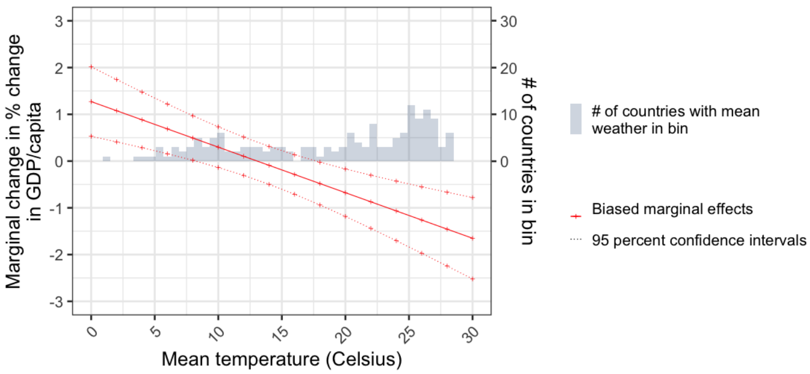
(b) Irrigated Counties



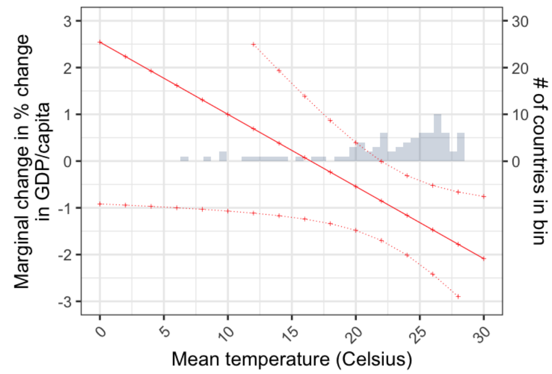
Notes: Marginal weather effects and marginal climate effects are based on coefficients in Table 2.2. The results predict the change in farm profits per acre.

Figure F.5: BHM: Incorrect (Varying) Forecast of the Marginal Effect of Temperature on Percent GDP Per Capita

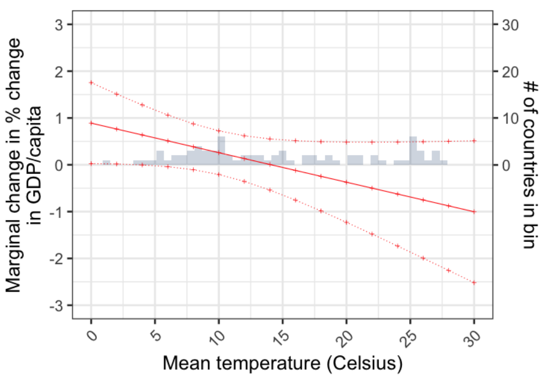
(a) All Countries



(b) Poor Countries



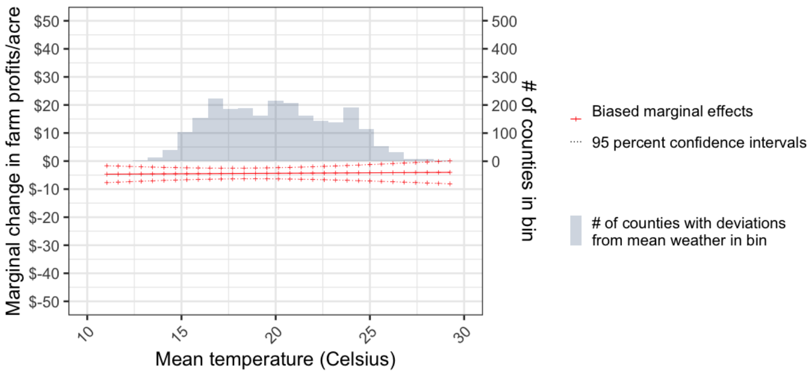
(c) Rich Countries



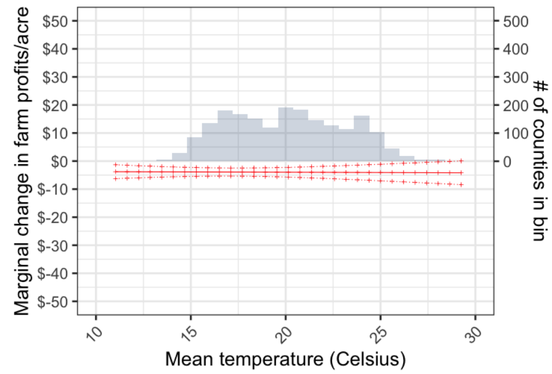
Notes: Forecast is inconsistent with the underlying model that assumes weather has the same effect regardless of climate. Biased marginal effects from the BHM version of equation (2.16) with BHM covariates and weather variables. Histograms show the distribution of country mean temperatures over years.
Panel a): The "All Countries" specification does not estimate different coefficients by poor and rich countries.

Figure F.6: D&G: Incorrect (Varying) Forecast of the Marginal Effect of Temperature on Farm Profit Per Acre

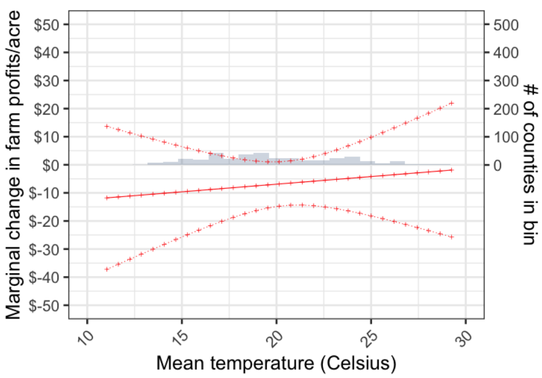
(a) All Counties



(b) Rainfed Counties



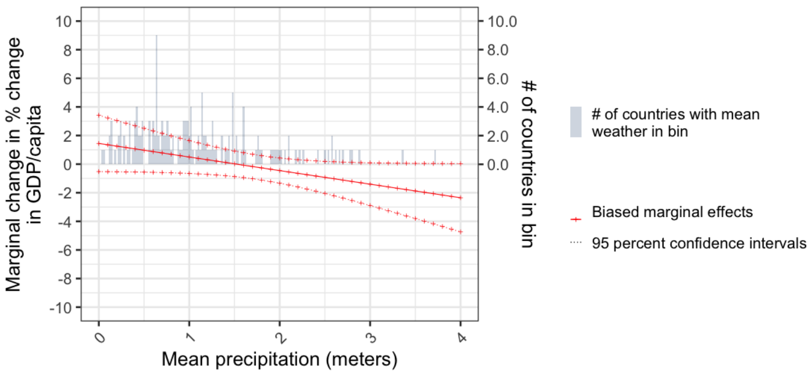
(c) Irrigated Counties



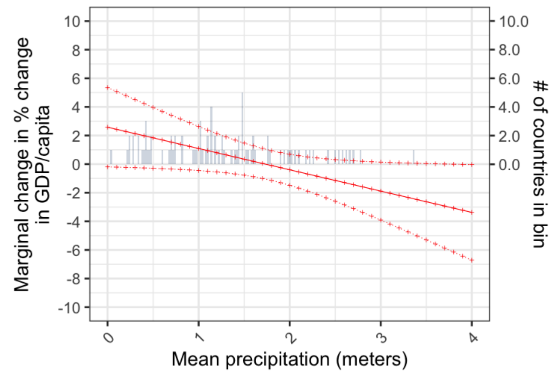
Notes: Forecast is inconsistent with the underlying model that assumes weather has the same effect regardless of climate. Biased marginal effects from the D&G version of equation (2.16) with D&G covariates and weather variables. Histograms show the distribution of county mean temperatures over years. Since D&G use GDD in degrees Fahrenheit as their temperature variable, we convert GDD to growing season mean temperature (Celsius) by: ($°C = (GDD/183) * (5/9) + 8°C$) where $8°C$ is the base temperature above which growing degrees are measured.
Panel a: The "All Counties" specification does not estimate different coefficients by rainfed and irrigated counties.

Figure F.7: BHM: Incorrect (Varying) Forecast of the Marginal Effect of Precipitation on Percent GDP Per Capita

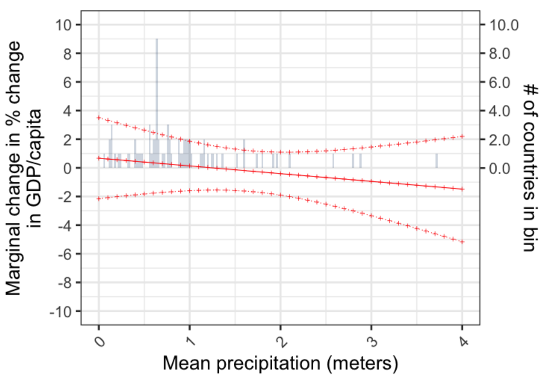
(a) All Countries



(b) Poor Countries

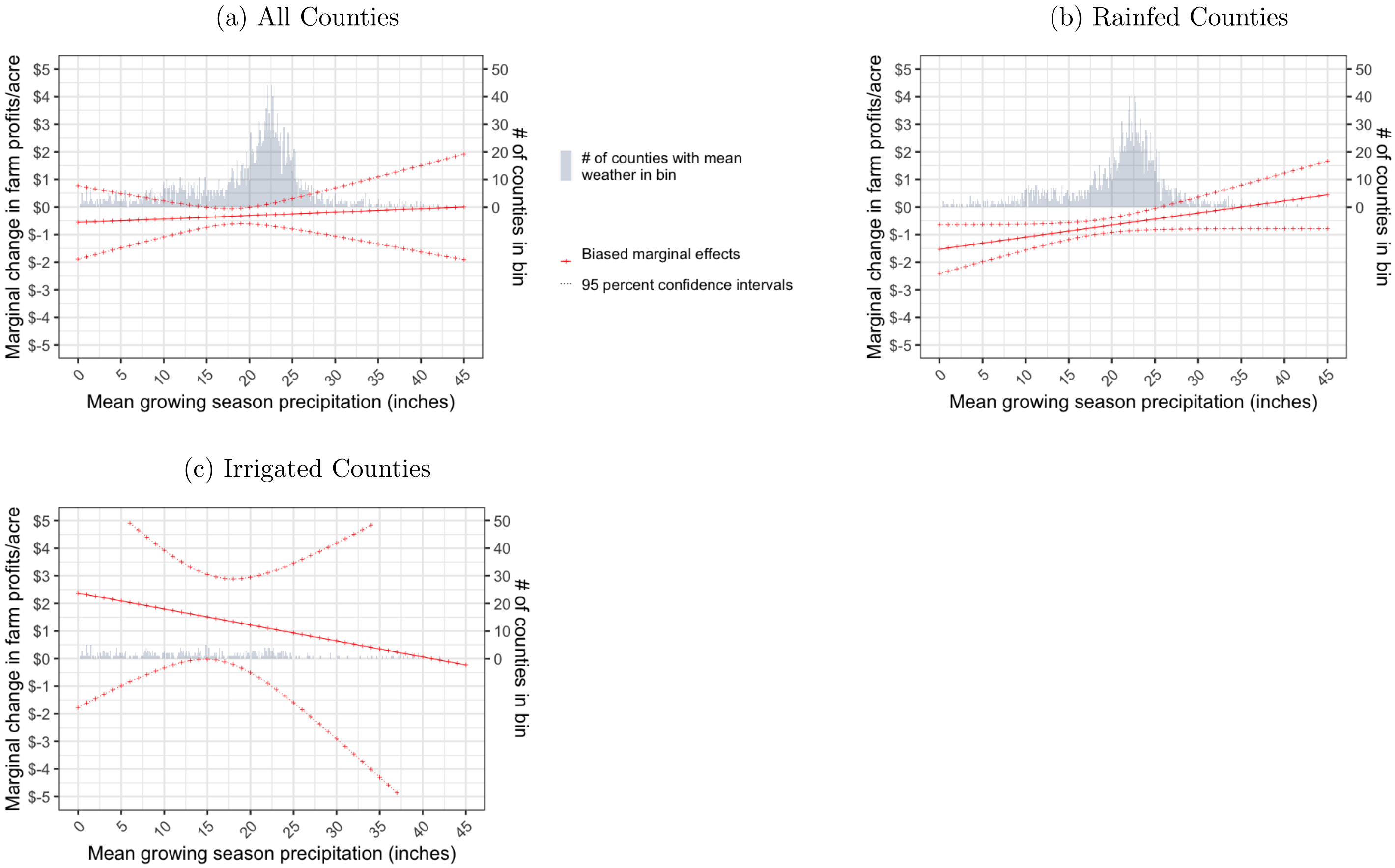


(c) Rich Countries



Notes: Forecast is inconsistent with the underlying model that assumes weather has the same effect in each location. Biased marginal effects reflect the BHM version of equation (2.16) with BHM covariates and weather variables. Histograms show the distribution of country mean precipitations over years.
Panel a): The "All Countries" specification does not estimate different coefficients by poor and rich countries.

Figure F.8: D&G: Incorrect (Varying) Forecast of the Marginal Effect of Precipitation on Farm Profit Per Acre

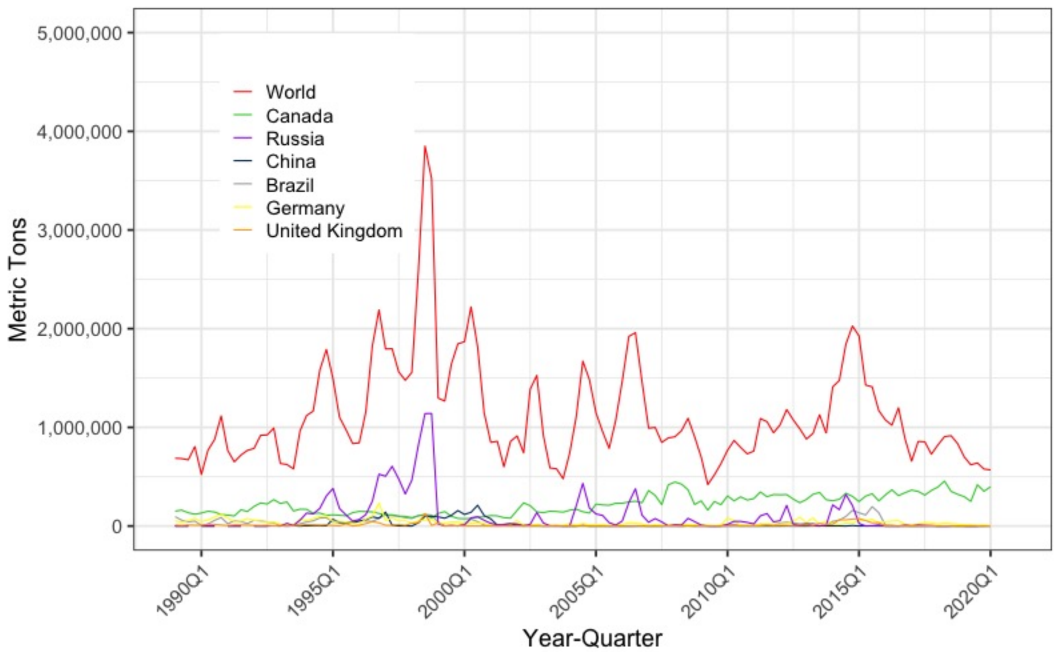


Notes: Forecast is inconsistent with the underlying model that assumes weather has the same effect in each location. Biased marginal effects reflect the D&G version of equation (2.16) with D&G covariates and weather variables. Histograms show the distribution of county mean precipitations over years.
Panel a): The "All Counties" specification does not estimate different coefficients by rainfed and irrigated counties.

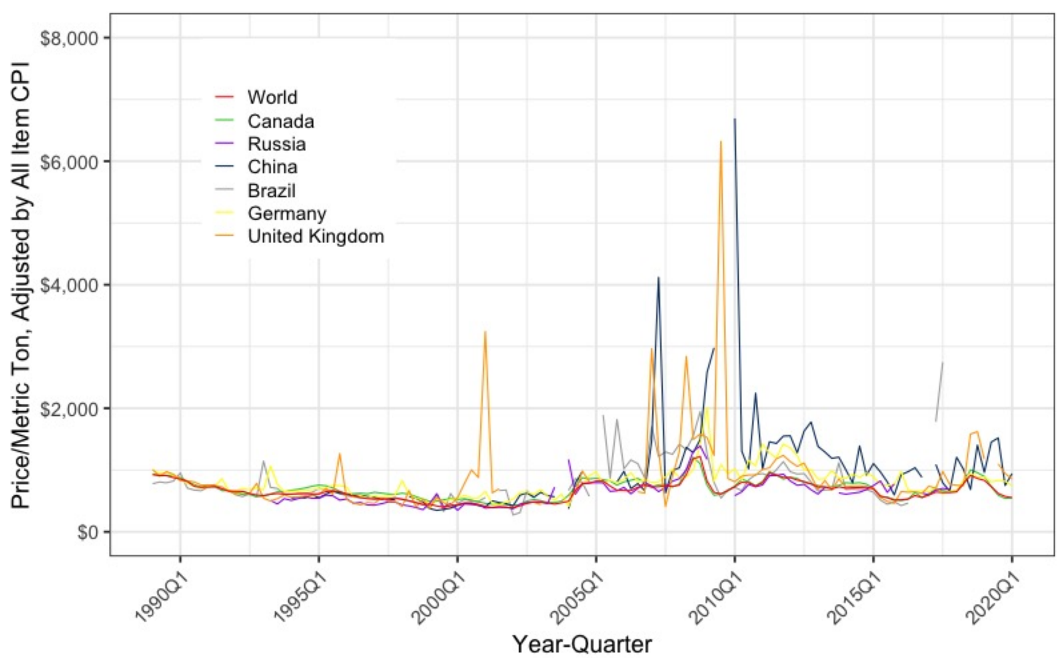
Appendix G: New Vehicle Sales: Data Appendix

Figure G.1: U.S. Imports of Flat Hot-Rolled Steel From the World and Selected Countries, Quarterly

(a) U.S. Import Quantities (metric tons)



(b) U.S. Import Prices, CPI Adjusted



Notes: (1) Based on data from from the U.S. International Trade Commission (ITC). Specifically, we use Harmonized Tariff Schedule (HTS) commodity code 7208, "flat-rolled products of iron or nonalloy steel, of a width of 600 mm or more, hot-rolled, not clad, plated or coated." We divide quarterly import values (nominal USD) by quantities (in kilograms) where import values include the value of the goods as well as the cost of duties, freight, insurance, and other charges. (2) Adjusted by all-item CPI re-based to 2016 USD. (3) Not seasonally adjusted.

ProQuest Number: 28322074

INFORMATION TO ALL USERS
The quality and completeness of this reproduction is dependent on the quality
and completeness of the copy made available to ProQuest.

Distributed by ProQuest LLC (2021).
Copyright of the Dissertation is held by the Author unless otherwise noted.

This work may be used in accordance with the terms of the Creative Commons license
or other rights statement, as indicated in the copyright statement or in the metadata
associated with this work. Unless otherwise specified in the copyright statement
or the metadata, all rights are reserved by the copyright holder.

This work is protected against unauthorized copying under Title 17,
United States Code and other applicable copyright laws.

Microform Edition where available © ProQuest LLC. No reproduction or digitization
of the Microform Edition is authorized without permission of ProQuest LLC.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346 USA