Title

Essays in Energy and Urban Economics

Author

Kadish, Jonathan Noah

Publication Date

2018

Peer reviewed|Thesis/dissertation

Essays in Energy and Urban Economics

by

Jonathan Noah Kadish

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Agricultural & Resource Economics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Meredith Fowlie, Co-chair
Professor Solomon Hsiang, Co-chair
Professor Maximilian Auffhammer
Assistant Professor Victor Couture

Spring 2018

Essays in Energy and Urban Economics

Abstract

Essays in Energy and Urban Economics

by

Jonathan Noah Kadish

Doctor of Philosophy in Agricultural & Resource Economics

University of California, Berkeley

Professor Meredith Fowlie, Co-chair

Professor Solomon Hsiang, Co-chair

Most human decisions are made without consideration for the associated energy use. But our actions, from choices about where to live and how to commute, to the routine decision about when to sleep or whether to go to church, have energy implications. Collectively, these decisions form cities filled with gasoline-consuming cars and dictate when power plants turn on and off. Choices, even when made subconsciously, are made subject to constraints formed by past societies' decisions or attempts to coordinate with family, friends, and colleagues. This dissertation broadly investigates how technologies, norms, and incentives affect human behavior, energy use, and, ultimately, climate change. I develop novel and large datasets to investigate unanswered questions in energy and urban economics.

In Chapters 1 and 2, I ask how transportation technology affects urban growth. A broad set of interacting factors, such as physical features and land use policy, cause particular spatial organizations of households and firms in rapidly growing urban areas. Transportation costs and real estate prices drive individuals' decisions about where to work and live. These choices have tremendous welfare implications, costing time and energy, and resulting in externalities including air pollution and traffic congestion. Despite high social costs, there is little empirical evidence about the effect of changes in transportation costs on city structure. I estimate the effects of two transportation innovations - (1) a speed limit increase, and (2) ridesharing services - on residential real estate prices and development. I find that prices respond quickly and significantly to transportation cost changes. Consistent with my theoretical model, an increase in speed limits decreases housing prices by over 3% on average, with the largest effect in the city center. Subsequent housing development is farther from the central city. In contrast, the launch of Uber increases housing prices by over 2% after the first year, with a larger immediate effect in the central city. Housing development occurs closer to the central city after treatment. Both treatments change the ability of households to access surrounding markets. Applying the concept of "market access" from the trade literature, I show that the distribution of business establishments around a property dramatically changes the magnitude of each effect.

Chapter 3, co-authored with Solomon Hsiang and Terin Mayer, shows that electricity use can be predicted from human activities. System operators predict electricity loads in order to schedule power plants and allocate transmissions resources to ensure grid reliability. In the long term, forecasting dictates whether new power plants will be contracted or built. We combine hourly data on electricity load with the American Time Use Survey and show that, with just three time-use variables, we can predict over 90% of variation in electricity use. In an increasingly data-rich world, we know more about what individuals' locations and activities, making our finding a potentially valuable tool for improving prediction as well as offering ways to reduce electricity usage by shifting human activity.

Chapter 4, also co-authored with Solomon Hsiang, explores energy use on holidays and weekends. Government policies that coordinate labor and leisure have profound economic implications. However, manipulating the structure of our workweek, weekend, and holiday calendar in order to improve economic outcomes is a policy lever that has been largely unused. Setting optimal coordinating mechanisms and allocations may present a useful tool for reducing carbon emissions, particularly if individuals are constrained in the number of days they have available to take coordinated leisure. We provide evidence that there is an environmental externality associated with labor relative to leisure. We empirically estimate the effect of weekends and holidays on electricity loads, vehicle travel, and air travel in the U.S. We observe large reductions in electricity load, air travel, and vehicle travel on many holidays, as well as reductions on surrounding days. Using time use data and exogenous variation in when a holiday is observed, we provide evidence that, beyond labor being less carbon intensive than leisure, agents enjoy less carbon-intensive activities on holidays. Holidays during the summer all result in significant savings. New holidays may be more economically valuable if they are scheduled for hot days that are likely to require high marginal cost generators to meet electricity demand.

Contents

List of Figures

List of Tables

Acknowledgments

This dissertation would not have been possible without the support of many people and organizations. I am incredibly grateful to Solomon Hsiang and Meredith Fowlie who have been amazingly helpful throughout my graduate career. Victor Couture, Maximilian Auffhammer, Reed Walker, James Sallee, and Michael Anderson have all provided useful feedback on the works contained in this dissertation.

All members of Global Policy Laboratory have been a necessary source of support and companionship. I am in particular grateful to Tamma Carleton and Jonathan Proctor for their help and encouragement.

I am grateful for financial support from the National Science Foundation, Giannini Foundation, Institute for Research on Labor and Employment, and Fisher Center for Real Estate & Urban Economics.

Finally, I thank my family and friends. I could not have done this without the support and patience of my parents, Josh and Lisa, my brothers, Nathan and Seth, my partner, Connie, and our dog, Zucchini.

Chapter 1

Speed and Sprawl: How highway speeds limits cause suburban development

1.1 Introduction

The history of urban growth is closely linked to transportation innovation. Households' decisions about where to live, work, and consume depend on available transportation technology and infrastructure. These decisions have economically significant implications. In the U.S., we collectively spend over 22,000 person-years driving everyday[1] and vehicle transportation results in externalities including local air pollution, congestion, accidents, greenhouse gases, and fatalities.[2] Despite the welfare implications, there is little empirical evidence on how falling transportation costs have shaped the spatial organization and growth of cities.

In this paper, I measure the causal effect of speed limits on real estate prices and housing development. I leverage the 1995 repeal of the National Maximum Speed Law (NMSL), which allowed states to increase the speed limit on urban freeways above 55 mph. The abrupt and staggered timing of this change allows me to estimate the effect on both the level and growth rate of real estate prices and new construction. I describe the spatial structure of each effect while controlling for high-resolution spatial- and temporal-controls.

Understanding how innovation affects marginal growth in cities is important as cities continue to grow and new cities develop. For example, when New York City was settled in 1625 as the Dutch city of *Nieuw-Amsterdam*, walking was the most common form of transportation. A trip to present-day Harlem (*Nieuw Haarlem*) would take over three hours each way, a time-intensive endeavor compared to the half hour subway or 20 minute (traffic-permitting) car trip today. In contrast, when Las Vegas was set-

[1] 210 million drivers (Federal Highway Administration, 2009) average 56 minutes of driving per day (2009 National Household Travel Survey)

[2] For a summary, see Parry, Walls, and Harrington (2007)

tled in 1905, the horseless buggy was already gaining popularity, and by 1915 Henry Ford had built 1 million cars. Using a modern analogue, this paper asks whether the different spatial structure of New York and Las Vegas, where population densities are 27,788 versus 4,223 people per square mile, respectively, might be in part due to the differing technologies available at the time of these cities' development. I find evidence that spatial structure does indeed depend on the widespread availability of transportation technologies, although it is *marginal* development at a moment in time that is most dependent on the current technology. Much like a tree grows in concentric rings whose thickness reflects rainfall availability in the current year, cities grow outward, adding homes, establishments, and infrastructure, in a manner that reflects current transportation technologies.

To ground these findings in a generalizable theoretical framework, I develop a model of urban housing and transportation. This model integrates the concept of market access (a measure of the transportation cost to a particular market), which was developed in the trade literature, into a classic urban economics framework. The theory provides intuition for how each of the transportation shocks in my study affect home values, and how the innovations interact with distance to the central business district and nearby consumption amenities.

To test the predictions of the model, I create a novel dataset, measuring routed distances between over 12.5 million residential transactions and 7.5 million business establishments across 177 U.S. cities. I use this dataset to implement the most spatially resolved market access analysis to date. After the NMSL was repealed in 1995, some states adopted higher speed limits over the next 4 years. I argue that the timing of the adoption, which depended on legislative priorities, was exogenous to real estate prices and growth.

This paper provides the first quasi-experimental evidence of how diffuse transportation innovation changes real estate values, urban development, and population in cities. Analyzing real estate prices and development within each city before and after the speed limit change, interacted with the distances to nearby markets, I find large impacts. Consistent with theory, a speed limit increase causes housing prices to fall near the center of the city, and increases the rate of suburban development. Integrating market access in an urban setting is important: I find that distant business establishments attenuate the negative effect of a speed limit increase on real estate prices.

Previous research on transportation in the urban literature has focused on the effect of large infrastructure investments on city-level outcomes (Baum-Snow, 2007a; Duranton and Turner, 2012) rather than diffuse changes on marginal development. More detailed within-city analyses focus on a single city or event, such as Ahlfeldt et al. (2015) on the construction and removal of the Berlin Wall, and Anderson (2014) on subway strikes in Los Angeles. Market access, applied mostly in the trade and development literature, has bolstered our understanding of the value of integrated economic markets (Donaldson, 2010; Donaldson and Hornbeck, 2016), but has not been linked to analyses of urban growth.

This paper builds on the current literature in urban, trade, and transportation economics, implementing a market access approach using granular data within many cities. The importance of understanding the dynamics of real estate, jobs, amenities, and transportation in governing city structure is underscored by a rapid transformation in the transportation sector: autonomous vehicles are already being deployed, and marginal innovations such as bike-sharing and electric-skateboards are fundamentally changing the way we move around cities. The extent of future changes and whether they will ameliorate or exacerbate the externalities associated with our automobile-centric transportation system is a larger question that I begin to address in this paper.

The paper proceeds as follows: Section 1.2 integrates market access into an urban theoretical framework to provide intuition about the effects of speed limits on urban real estate. In section 1.3, I summarize the existing empirical literature and describe my contributions. Section 1.4 describes the speed treatment in detail. Section 1.5 summarizes the datasets and how I linked 12.5 million residential real estate transactions to 7.5 million businesses. Section 1.6 presents my empirical design, which I apply in section 1.7, empirically estimating the relationship between transportation innovation and real estate prices and growth. Section 1.8 concludes and provides directions for future research.

1.2 Theoretical Framework

There is a rich theoretical literature on city structure. This literature has evolved from the monocentric city model, developed by Alonso (1964), Mills (1967), and Muth (1969), which proposes a linear city with a single job location - the central business district (CBD) - where all workers earn the same wage. Commute costs increase with distance and, in equilibrium, all workers have the same level of utility. Even in its simplest form, the model describes a fundamental relationship between urban land use, transportation, and population. Many authors have expanded the original model by adding, for example, travel time and multiple transportation modes (Anas and Moses, 1979; Baum-Snow, 2007b), income heterogeneity (Duranton and Puga, 2013), durable housing (Glaeser and Gyourko, 2005), or endogenous job growth (Ogawa and Fujita, 1980; Imai, 1982). For a review of the vast theoretical literature on urban land use, see Duranton and Puga (2015).

I build on the existing urban theory literature by proposing a flexible framework for including fixed and (distance- and time-) varying transportation costs in a monocentric city model. I add (exogenously located) consumption amenities to the model. In equilibrium, the value of a house depends on access to both the CBD and these local amenities.

Consider a closed, linear city with one unit land available at each location x. Assume, without loss of generality, that the center of the city is at $x = 0$. Individuals consume one unit of land at location x which costs $h(x)$, and land and housing can be thought of as interchangeable by assuming capital costs are zero. Also assume the cost of marginal land is zero. Everyone works at the CBD and earns the same wage, w.

Let s^i be the average speed of traveling by mode i, τ^i_f be a fixed cost of mode i, and τ^i_v be the price per mile travelled using mode i (which I call the variable cost of the mode). Walking would have zero fixed and variable cost. A bus ticket that is not distance dependent would have a positive fixed cost and no variable cost. A car, compared to a bus or walking, would have a higher average speed, a high fixed cost (which would include the car purchase amortized over the car life, insurance, and parking) and a high variable cost (the cost of fuel and distance-based-depreciation). Finally, I also include a time cost, which depends on the speed of the mode. The time for a trip to the CBD is x/s, and time is valued at the wage, w.

Initially, I ignore the mode choice, i, to derive the effect of a speed change. I reintroduce i in chapter 2 to derive the effect of ridesharing. The cost of traveling any distance, d, is $T(d, s) = \tau_f + \tau_v d + wd/s$. Therefore, the cost of commuting to the CBD from a house at location x is $T(x, s) = \tau_f + \tau_v x + wx/s$. The cost includes a the fixed amount, τ_f, a distance-dependent amount, τ_v, and includes the time cost, wx/s.

Suppose there are J consumption locations where consumers buy goods. Let c be a vector of length J containing consumption of each good, and I be a vector of length J indicating whether or not an individual consumes any of the good ($I_j = \mathbb{1}\{c_j > 0\}$). Therefore the j^{th} entry of c, c_j equals the quantity of good consumed and $I_j = 1$ if the quantity is greater than zero. Given the vectors (also indexed by j) of prices, p, and locations, y, let c^* be the solution to the individuals consumption maximization problem at house x. Utility is strictly quasi-concave and increasing in c.

$$c^*(x) = \arg\max_c \sum_j [u_j(c_j) - \underbrace{T(d(x, y_j), s)}_{\text{travel cost from } x \text{ to establishment } j} * \underbrace{I_j}_{\text{whether or not hh visits } j}]$$
$$\text{s.t. } w - h(x) - T(x, s) - c' \cdot p - \sum_j [T(d(x, y_j), s) * I_j] \geq 0$$

The household spends the remaining wage after paying for housing and commuting costs on a bundle of good, c, that maximizes utility. A home's value is driven both by the proximity to work and by the proximity to consumption amenities. Because each cost is additive, we can consider the willingness to pay for the two location components separately.

In equilibrium, since everyone makes the same wage, utility is also equalized. Set a threshold utility, $\underline{\text{u}}$. Individuals achieve this utility by consuming $c^*(x)$. Hence, the expenditure minimization problem is:

$$x = \arg\min_x h(x) + T(x, s) + c^*(x)' \cdot p + \sum_j [T(d(x, y_j), s) * I_j]$$
$$\text{s.t. } u(c) \geq \underline{\text{u}}$$

This defines the bid-rent function for housing, which is the maximum price an individual will pay for housing at location x conditional on achieving $\underline{\text{u}}$:

$$\Psi(x, \underline{\text{u}}) = w - T(x, s) - c^*(x)' \cdot p - \sum_j [T(d(x, y_j), s) * I_j]$$

In equilibrium, $h(x) = \Psi(x, \underline{\text{u}})$, competition drives price up to the willingness to pay and at the edge of the city $\Psi(x, \underline{\text{u}}) = h(x) = 0$. Since each component is additive, I consider the distance to the CBD and the distance to amenities separately. First, willingness to pay is decreasing in distance from the city, as distance from the CBD increases. Assuming the city has N workers, since each house occupies one unit of land, the linear city extends $N/2$ land units in each direction from the CBD. Since it is symmetrical, I show the gradient only on one side of the CBD.

Figure 1.1 shows the partial equilibrium gradient of housing prices with respect to distance to CBD, where $\Psi^{CBD}(x, \underline{u}) = w - T(x, s)$.

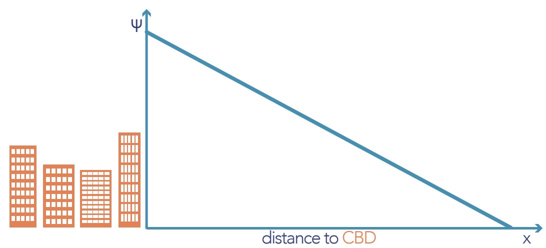


Figure 1.1: Decreasing housing prices with distance to CBD.

Prices also rise with proximity to amenities, $\Psi^c(x, \underline{u}) = \sum_j [u_j(c_j^*) - T(d(x, y_j), s) * I_j]$. In a simple example with two amenities that are perfect substitutes, willingness to pay for housing decreases with distance to each amenity. This is illustrated in Figure 1.2.

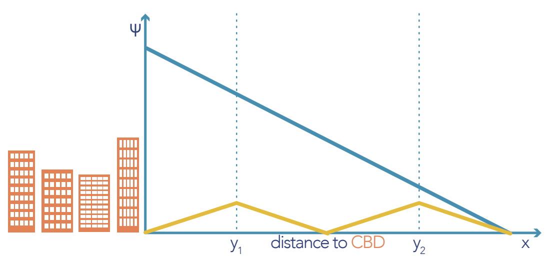


Figure 1.2: Decreasing housing prices with distance to amenities.

Adding these bid rent curves illustrates the equilibrium price gradient in the city. This is shown in Figure 1.3.

Speed Limit Change

What happens to the equilibrium prices when speed limits increase? Again, we can analyze the commuting effect and the amenity effect separately. The slope of the bid rent curve as distance from the CBD increases is $\Psi_x^{CBD} = -\tau_v - w/s$. Therefore, increasing s decreases the magnitude of the slope of the bid-rend function since $\partial\Psi_x^{CBD}/\partial s = -w/s^2$. The change in slope of the bid-rent function and the resulting willingness to pay for proximity to the CBD is illustrated in Figure 1.4.

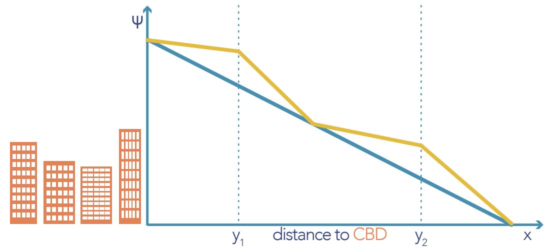


Figure 1.3: Totaled bid-rent function with CBD and amenities.
Housing price is the sum of the willingness to pay for proximity to CBD and willingness to pay for proximity to nearby amenities.

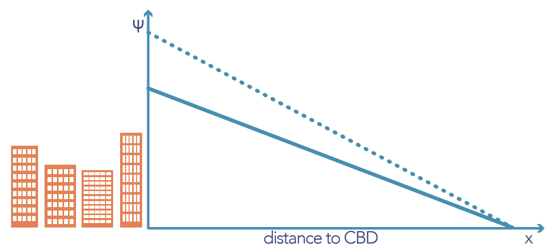


Figure 1.4: Increasing speed decreases the willingness to pay near the CBD.

Intuitively, living near the CBD has become less valuable, making the suburbs relatively more attractive. A similar effect occurs with proximity to businesses, particularly in the simplified example shown in Figure 1.5.

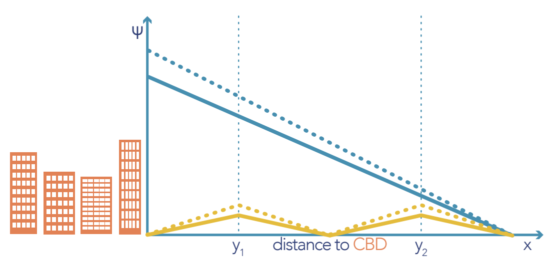


Figure 1.5: Increasing speed decreases the value of nearby amenities.

Again, to get the total effect the results can be added, as shown in Figure 1.6.

However, the result is less intuitive when amenities are generalized. This is because a change in speed, s, can change the optimal set of amenities visited, c^*. To see this, consider the set of businesses visited by a given household at location x, $c*(x)$. Without loss of generality, order the businesses by their net utility contribution to Ψ^c. Then for

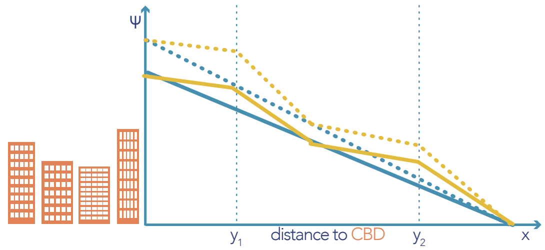


Figure 1.6: Increasing speed decreases willingness to pay overall in the simple example.

the m consumption locations that the household at location x visits, it must be true that the m^{th} business provides more net utility than the $(m+1)^{th}$ business. Formally,

$$u_m(c_m) - T(d(x, y_m), s) > u_{m+1}(c_{m+1}) - T(d(x, y_{m+1}), s) \iff$$
$$u_m(c_m) - u_{m+1}(c_{m+1}) > T(d(x, y_m), s) - T(d(x, y_{m+1}), s).$$

The difference in utility must be greater than the difference in travel cost. Suppose s increases to s'. Then, because $\partial T/\partial s < 0$, the household at location x switches from visiting c_m to visiting c_{m+1} if

$$u_m(c_m) - T(d(x, y_m), s') < u_{m+1}(c_{m+1}) - T(d(x, y_{m+1}), s') \iff$$
$$u_m(c_m) - u_{m+1}(c_{m+1}) < T(d(x, y_m), s') - T(d(x, y_{m+1}), s')$$

That is, the change in travel cost to c_{m+1} is greater than the change in travel cost to c_m by enough to overcome the difference in utility. If a change in s causes such a change in $c*$, then this results in an increase in Ψ^c, all else equal. As a result, the presence of amenities and their spatial distribution causes the effect of a speed limit change on Ψ^c to be uncertain rather than always negative.

Predictions

The theory informs the following predictions. Increasing speed limits should reduce real estate prices, with the largest magnitude effect near the CBD and a decreasing effect with distance. Having more distant amenities attenuates the impact of a speed limit increase on prices because distant establishments are now more accessible. Nearby amenities matter less.

Development is not included in this model (because, for tractability, housing stock and population are fixed). However, I can make predictions based on relative prices. A speed limit increase causes suburbs to become more valuable relative to the central city. Therefore, I expect development to become more sprawled and occur farther from the city center.

1.3 Empirical Literature

While the theory on urban land use and transportation is well developed, the empirical literature is more limited. Summarizing this literature, Duranton and Puga (2015) conclude that we do not have basic answers to many first-order questions. Our knowledge of urban sprawl is based on only a few articles, many of which investigate aggregated data in a single city rather than granular data from a broad cross-section of cities. To improve upon this existing body of evidence, I use household level attribute and transaction data from 177 cities matched to zip code level establishment data to estimate the effect of a plausibly exogenous, diffuse transportation innovation on prices and development outcomes.

The major challenge to identifying the causal effects of transportation changes on many outcomes of interest (such as population density, land rents, and economic output) is that selection into treatment is usually nonrandom. Assessment of any transportation improvement requires constructing an accurate counterfactual. Before this paper, the primary strategy that has been used to identify a causal effect of transportation on an urban outcome of interest has been to find an instrument that plausibly satisfies the exclusion restriction. Baum-Snow (2007a) pioneered the use of a planned route as an instrument for road infrastructure, measuring the effect of highways on population decentralization in cities. He instruments for the number of highway rays (road segments) through a city's CBD using the number of rays in the 1947 national interstate highway plan, estimating that an additional highway through a central city cause the population within a central district to fall by 18%. This important finding opens many questions about the process and mechanisms that lead to a new population distribution. Because he uses decennial census data and the I.V. approach to measure a long difference, it is not possible to get more detail about the immediate effect on suburban growth or how any change in structure occurs. The same approach has since been used in multiple countries, including China (Baum-Snow et al., 2015) and Spain (Garcia-Lopez, 2012), revealing similar results to those in the U.S. Duranton and Turner (2012) use a similar instrument to show that highways increase total employment in a city. These focus on city level outcomes, such as total employment, rather than within city organization.

Three studies use random or quasi-random approaches to evaluate changes within cities. Ahlfeldt et al. (2015) leverage the construction and removal of the Berlin Wall as a natural experiment, finding that population densities and land prices change based on shifting access to West Berlin.This paper proposes the idea of market access in an urban setting, but uses only one city and an empirical context with limited policy relevance. Gonzalez-Navarro and Quintana-Domeque (2015) use a randomized control trial to estimate the effect of road paving on property values in peripheral neighborhoods in Mexico. They find that paving increased home values by 17%. Heblich, Redding, and Strum (2017) develop a structural estimation method for commuting flows, and apply it to the development of the steam locomotive and its effect on London. They show that the model can be used to predict the impact of the railways construction or removal. While these studies are useful, their limited spatial scope and isolated treatments make it difficult to draw general conclusions or policy implications.

There are many studies on the effect of public transportation, specifically subway stations, on real estate prices. Most of these studies cannot be considered causal because they do not address the endogeneity problem that public transportation will be expanded where it it most valuable and needed. One of the most credible articles, Gibbons and Machin (2005), use a difference-in-differences approach to estimate the effect of a subway extension in London. They find that a 1 km reduction in the distance to a subway station, for properties within 2 km of a station, increased real estate price by 2%.

Couture, Duranton, and Turner (2016) use survey data from National Household Travel Survey to estimate travel speeds across U.S. cities. They find several determinants of city speed: those cities that are more centralized are slower, while cities designed with ring roads allow cars to move more quickly. Generally, adding roads increases speeds but adding vehicles decreases them.

Finally, there are several articles that broadly try to understand the factors that cause urban sprawl. Glaeser and Kahn (2004), use a cross section of international cities to show a strong correlation between car ownership, gasoline taxes, and cities with low urban density. To address the endogeneity of gasoline taxes, they use a country's legal origin as an instrument. In places where gasoline is cheaper, population density is lower. This suggests, they argue, that cars are responsible for sprawl. Burchfield et al. (2006) measure sprawl from satellite images, and correlate basic measures of development density with possible explanatory factors. They find sprawl is associated with dispersed employment, automobile ownership, and physical geography such as rugged, non-mountainous terrain. Bento et al. (2005) ask essentially the opposite question of how urban form affects travel demand. They show that in denser cities, fewer individuals commute by car and total vehicle travel is lower. While these studies are useful for thinking about the universe of factors that may influence the spatial expansion of cities, endogeneity concerns prevent any causal interpretation.

This paper moves the empirical urban literature forward by applying a market access approach to a cutting-edge dataset. Below, I describe the setting and dataset in detail.

1.4 Empirical Setting

The final repeal of the National Maximum Speed Law (NMSL) changed the speed at which drivers on urban freeways could travel. The law was a provision of the Federal 1974 Emergency Highway Energy Conservation Act, which was created in response to the 1973 oil crisis. NMSL prohibited speed limits higher than 55 mph on all roads, and compliance was required in order for states to receive highway funds. U.S.

In April of 1987 Congress voted to allow states to raise the speed limit on rural interstates to 65 mph. In November of 1995 Congress removed all limits, allowing states to dictate the appropriate speed limits. Between 1995 and 1999, states passed legislation to change urban speed limits. Eleven states kept an urban speed limit of 55 mph, 33 states increased the limit to 60-65 mph, and 6 states increased the limit to or above 70 mph. Retting and Greene (1997) measured speeds in two cities before and after the policy change, and found that average highway speed increased 2-3 mph and the proportion of cars traveling over 70 mph increased dramatically.

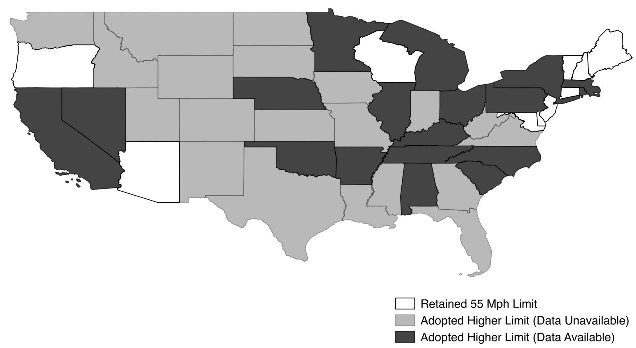


Figure 1.7: Urban speed limit adoption status and data availability by State.

Several studies have used these changes to measure the effect of speed limits on vehicle fatalities, finding increases in fatalities after each speed limit increase (Friedman, Hedeker, and Richter, 2009; Farmer, Retting, and Lund, 1999). Two papers in economics have used this empirical setting. Ashenfelter and Greenstone (2004) estimate the implicit value of a statistical life that states accept when they adopted a higher rural speed limit in 1987. They find that fatality rates increased by 35%, implying time savings of $1.54 million per fatility (in 1997 dollars). Van Benthem (2015) uses both the 1987 and 1995 events to show that the increase in speed resulted in more accidents, fatalities, and higher air pollution.

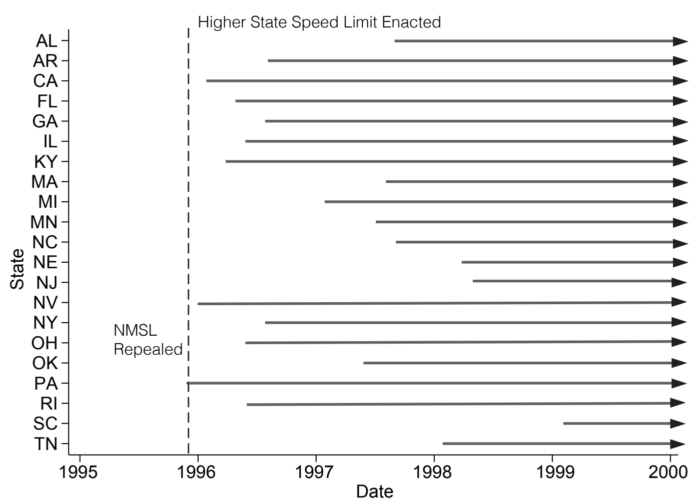


Figure 1.8: Urban speed limit change date by State.

I leverage the fact that states chose when to enact legislation to raise the speed limit by estimating a trend-break model. I compare similar households before and after a speed limit change, relative to households in states that still have a lower speed limit but will adopt a higher speed limit. The richness of my data allows me to control for spatially- and temporally-fine fixed effects.

1.5 Data

I create a housing transaction dataset containing over 7 million residential real estate transactions from 28 states between 1994 and 2000. I link these transactions with 7.5 million business establishments using routed road distances. To my knowledge, this has never been done in the economics literature. I describe the dataset collection and cleaning process below.

Housing Prices and Attributes

I web scrape the universe of publicly-available and internet-accessible U.S. residential real estate transactions. For each transaction, I collect the sale date and price. For each property, I collect the most recent attributes according to public records, when available. I use age (by subtracting year built from year of transaction), size (square footage), and type (single-family, multi-family, condo, etc.) as house-specific covariates in my analysis. I show my results are robust to the choice of covariates. When using covariates, I ignore homes that were built after the date of transaction (and are therefore guaranteed to have inaccurate attributes) which represents less than 1% of the sample.

Business Establishments

The U.S. Censuses' County Business Patters (CBP) records the number of business establishments by zip code from 1994-2016. Some observations are censored to maintain confidentiality. In these cases, I replace the establishment count with the minimum value based on the flag code, which is the most conservative assumption. It leads to an underestimate of the number business establishments. I match each zip code with its corresponding zip code tabulation area (ZCTA). ZCTAs, which were created by the Census Bureau, represent spatial units while zip codes sometimes lack a spatial component.

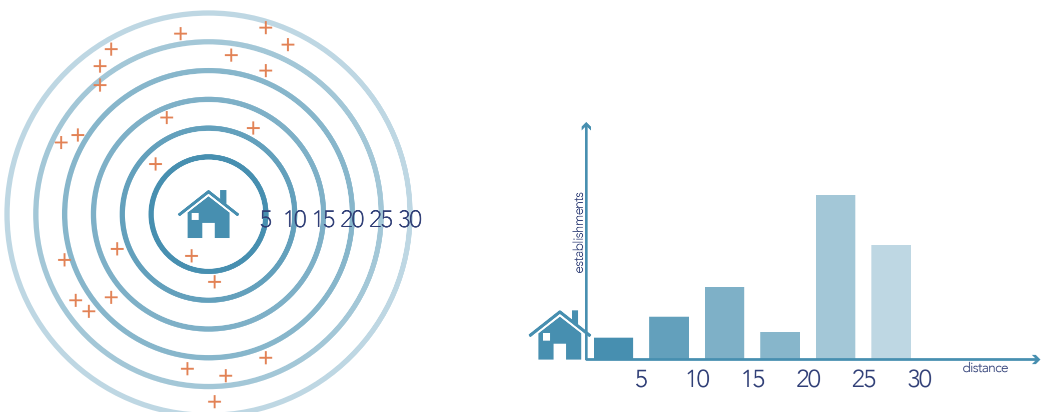
Distances

For each property, I measure the minimum routed distance along roadways in Open Street Maps to the CBD. Distances are calculated using djikstra's algorithm. This method is also used by websites like Google Maps. Following Baum-Snow (2007), I take the CBD definition from the *1982 Economic Censuses' Geographic Reference Manual*, which is based on the tracts that are identified by local businesspeople as the central place of business. Cities are defined as Core Based Statistical Areas (CBSAs), which are defined by the Office of Management and Budget.

Matching each household to local business establishments which represent locations of consumption c for each good c_j from my theoretical framework, I measure the distance from each home to every ZCTA within 40 routed miles. I create these measurements for all 12.5 million transactions, resulting in over a billion transportation network measurements.

Binning Establishments

In order to represent the spatial structure of the establishments around a home, I create bins with counts of the number of establishments within concentric annuli. I choose 5-miles to be the bandwidth. The first bin would be the count of establishments within a 5-mile buffer, the second would be the count of establishments in the annulus that begins past 5-miles and ends at 10-miles, with this pattern continuing to 40-miles. Figure 1.9 illustrates how nearby establishments are counted and added to bins.



(a) An example house transaction with surrounding establishment bins

(b) Establishment counts by annulus bins for an example transaction

Figure 1.9: Establishments around a house summed into bins by distance.
All distances are shortest-path along a road network.

Speed Limits

I collect speed limit data from the Insurance Institute for Highway Safety. The dataset includes urban and rural maximum speed limit and date of speed limit change.

1.6 Empirical Design

To measure the effect of speed limits on urban real estate, I assume that exact timing of each event is exogenous to my outcomes of interest, and measure the change in level and slope of each outcome variable before and after the abrupt transportation innovation. This design is often called a "trend-break" model. In the ideal experiment, I would be able to compare an identical house in an identical city with only a difference in the speed limit. Because this comparison is impossible, I compare homes in the same census tract, before and after treatment, controlling for house-specific covariates, relative to changes in untreated cities that will receive treatment later. In practice, this means I limit my sample to transactions in treated cities, making the control homes those that will be treated but have not yet. The identifying variation is the level and rate of growth of the treated variables, at a specific location, before and after treatment, relative to the level and rate of growth of untreated variables over the same period. The identifying assumption is that the expectations of the level and rate of growth of each outcome variable, before and after treatment, conditional on covariates, are equal.

Econometric Model

A basic trend-break model regresses the outcome of interest on a treatment, trend, and treatment-trend interaction:

$$\begin{aligned} log(outcome_{ict}) = & \beta_1 \mathbb{1}\{t > T\}_{it} \\ & + \beta_2 \mathbb{1}\{t > T\}_{it} * (Trend_t - T) \\ & + Trend_t * \rho_c + \mu_t + \varepsilon_{ict} \end{aligned} \tag{1.1}$$

where T is time of ride-share launch or speed limit change, $Trend_t$ is number of time period since beginning of sample, c is a spatial unit larger than the unit of observation i.

In this case, if i is a property, then I could regress the price of property i at time t on a binary variable that indicates whether the speed limit has changed and a variable that interacts the trend with the binary treatment. The pre-trend can be a smaller spatial unit. I control for the pre-trend in each census tract to account for differential growth rates within cities. Figure 2.4 shows the graphical interpretation of a trend-break.

In my preferred specification shown in equation I add additional controls to account for observable and unobservable variation in the outcome of interest for property i in census tract c during month t:

$$\begin{aligned} log(outcome_{ict}) = & \beta_1 \mathbb{1}\{t > T\}_{it} \\ & + \beta_2 \mathbb{1}\{t > T\}_{it} * (Trend_t - T) \\ & + Trend_t \times \rho_c + \mu_t + CBSA_i \times month_t \\ & + \psi_1 age_i + \psi_2 age_i^2 + \theta_1 size_i + \theta_2 size_i^2 + \varepsilon_{ict} \end{aligned} \tag{1.2}$$

where T is date of speed limit change, $Trend_t$ is number of time period since beginning of sample, $Trend_t \times \rho_c$ are census-tract level trends, $CBSA_i \times month_t$ are city-specific

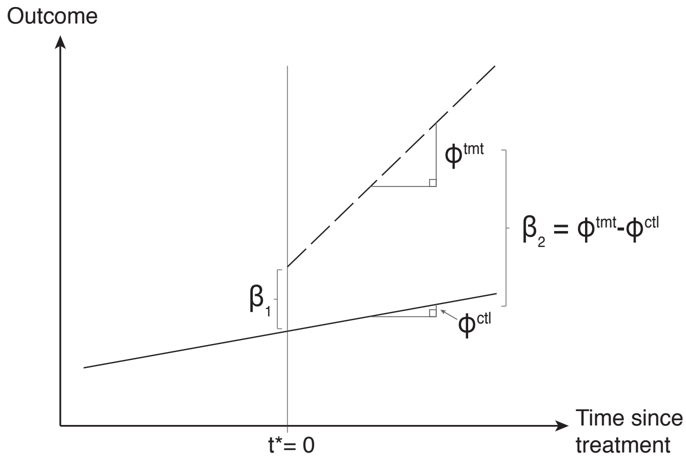


Figure 1.10: The trend break model measures a change in level and slope.

month dummies, age_i is the number of years since the house i was built, and $size_i$ is the square footage of house i.

Interacting the Trend (Treatment) parameter with a distance to CBD and establishment bins gives heterogeneous effects on the change in slope (level) at different locations within a city. For simplification, I define $treat_{it} = \mathbb{1}\{t > T\}_{it}$ and $post\_trend_{it} = \mathbb{1}\{t > T\}_{it} * (Trend_t - T)$.

$$
\begin{aligned}
&log(outcome_{ict}) = \\
&\quad \underbrace{\beta_1 treat_{it} + \beta_2 post\_trend_{it}}_{\text{baseline trend-break}} \\
&\quad + \underbrace{\sum_{k=1}^{n} Distance_i^k \times [\gamma_{0k} + \gamma_{1k} treat_{it} + \gamma_{2k} post\_trend_{it} + \gamma_{3k} \times trend_t]}_{\text{distance-to-CBD interactions}} \\
&\quad + \underbrace{\sum_{m=1}^{8} log(Establishments_{itm}) \times [\delta_{0m} + \delta_{1m} treat_{it} + \delta_{2m} treat_{it} \times post\_trend_{it} + \delta_{3m} \times trend_t]}_{\text{binned-establishments interactions}} \\
&\quad + \underbrace{trend_{ct} + \mu_t + CBSA_i \times month_t}_{\text{spatial- and temporal- controls}} \\
&\quad + \underbrace{\psi_1 age_i + \psi_2 age_i^2 + \theta_1 size_i + \theta_2 size_i^2}_{\text{home-specific attribute controls}} + \varepsilon_{ict}
\end{aligned} \tag{1.3}
$$

where T is date of speed limit change , $Trend_t$ is number of time period since beginning of sample, $trend_{ct}$ are census-tract level trends, $Distance_i^k$ is the distance from property i to the CBD to the k^{th} power, $Establishments_{itm}$ is the number of business establishments between $5 \times (m-1)$ and $5m$ miles from property i at time t, $CBSA_i \times month_t$ are city-specific month dummies, age_i is the number of years since the house i was built, and $size_i$ is the square footage of house i. I estimate equations 1.2 and 1.3 in the next section.

1.7 Results

Trend-Break

To illustrate that a trend-break model fits the empirical setting, I first estimate a flexible specification, creating dummy variables that represent the periods before and after treatment. I use four-month intervals, and start 20 months before treatment. Let p index the 4 month periods before and after treatment, with treatment occurring between zero and one. Then $running_{it}^{p} = 1$ if $p - 1 < (t - T)/4 \leq p$, where T is treatment date and t is transaction date.

I estimate the equation

$$\begin{aligned} log(prices_{ict}) = & \sum_{p=-5}^{max} \alpha_p \times running_{it}^{p} \\ & + Trend_t \times rho_c + \mu_t + \gamma CBSA_i \times month_t \\ & + \psi_1 age_i + \psi_2 age_i^2 + \theta_1 size_i + \theta_2 size_i^2 + \varepsilon_{ict} \end{aligned} \tag{1.4}$$

where max is 13. This specification corresponds to Equation 1.2, with treatment and trend variables replaced by dummy variables, allowing them to vary non-linearly. Figure 1.11 graphs the coefficient and standard error for each $running_{it}^{p}$, and a linear-trend is estimated through the pre- and post- treatment coefficients to represent the trend-break that is attributable to treatment.

The pre-trend is flat, suggesting that the tract-level trends, city-by-month fixed effects, and month-of-sample fixed effects, combined with housing covariates, remove variation that is unrelated to treatment. After treatment, there is a negative treatment term and a negative trend term for approximately 3 years.

Results from direct estimations of Equation 1.2 are shown in column 4 of Table 1.1. The speed limit treatment reduced prices by up to 4% immediately, and prices continued to trend downward by 3.7% per year relative to the counterfactual. This finding is robust to the choice of covariates. The estimates become most attenuated when tract-specific trends are ommitted, suggesting that controlling for spatially-specific trends is important for the empirical context.

The results are also robust to selection of different sample periods, as shown in Table 1.2.

Distance Interactions

Theory predicts that treatment to be heterogeneous by distance to the city center. The speed limit increase should cause prices to drop most near the center. I interact the treatment and trend variables with a 3^{rd}-order polynomial in distance to the CBD,

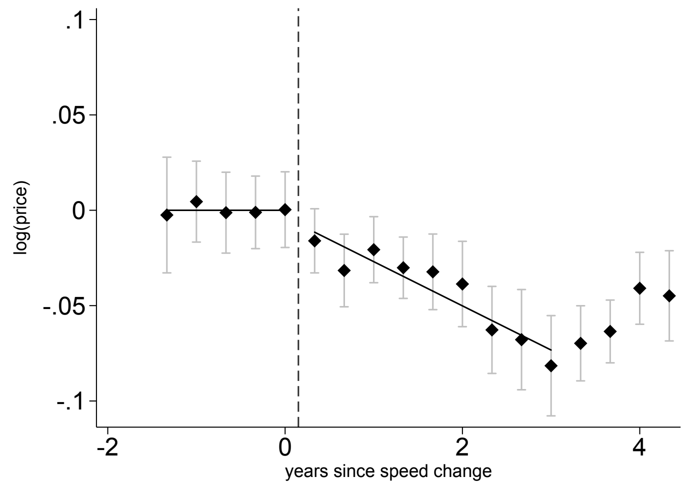


Figure 1.11: Effect of speed limit change on real estate prices: Trend-break model
Point estimates are dummy variables for 4 month periods before and after treatment. The estimation includes census-tract level trends, city-specific month dummies, house age, house age squared, house size, and house size squared. Standard errors are two-way clustered by city and month-of-sample. Linear trends are estimated through dummy estimates before and after treatment.

Table 1.1: The effect of a speed limit increase on real estate prices.

	(1) log(price)	(2) log(price)	(3) log(price)	(4) log(price)
$\mathbb{1}$\{Speed limit increase\}	-0.0264***	-0.0358***	-0.0368***	-0.0398***
	(-3.65)	(-5.63)	(-7.50)	(-8.01)
$\mathbb{1}$\{Speed...\}*Trend	-0.0111**	-0.0372***	-0.0357***	-0.0371***
	(-3.12)	(-5.50)	(-5.42)	(-5.44)
Observations	7912024	7912024	7902610	7250732
Year FE	X	X	X	X
Month-of-sample FE	X	X	X	X
Tract-specific trend		X	X	X
City-by-month FE			X	X
House covariates				X

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Standard errors are two-way clustered by city and month-of-sample. Column 4 is estimation of Equation 1.2, which includes month-of-sample fixed effects, tract-specific trends, city-by-month fixed effects, and house specific covariates.

Table 1.2: Effect of speed limit increase on real estate prices (sample robustness).

	(1) log(price)	(2) log(price)	(3) log(price)	(4) log(price)
𝟙{Speed limit increase}	0.00929	-0.0252***	-0.0159**	-0.0398***
	(0.92)	(-4.35)	(-2.82)	(-8.01)
𝟙{Speed...}*Trend	-0.162	-0.0326	-0.0429*	-0.0371***
	(-1.61)	(-1.13)	(-2.39)	(-5.44)
Observations	1332901	1996029	2657979	7250732
Months before and after treatment	12	18	24	

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Standard errors are two-way clustered by city and month-of-sample. All columns correspond to estimation of Equation 1.2 which include month-of-sample fixed effects, tract-specific trends, city-by-month fixed effects, and house specific covariates.

estimating the marginal effect of treatment for transactions at varying distance from the center. The treatment term effects are shown in Figure 1.12. They correspond to the marginal effects of β_1^k in Equation 1.3: the effects on prices just before and just after the treatment date at every distance from the CBD. Speed limits lead to a large drop in housing prices near the CBD. The effect, greater than 5%, suggests these houses are no longer as valuable since commute times from the suburbs have fallen. This effect returns to zero as homes move farther from the CBD. Figure 1.12 also shows the results of binned distance estimates, allowing for a non-parametric relationship between distance and effect. While the results are similar to the polynomial, the binned approach estimates some increase in housing prices at far distances, which suggests growth in the suburban housing market. This is discussed in the results subsection on development and sprawl.

The heterogeneous trend in prices, shown in Figure 1.13 continue in the same direction as the treatment effect. This ongoing trend could reflect the continued capitalization of the transportation innovation into housing prices. It could also be ongoing changes that are not captured by the treatment variables. It may be the case that businesses (and job locations) become more sprawled over time as well, causing proximity to the CBD to continually lose value for several years after the speed limit change. From Figure 1.11, it appears that the downward trend continues for about 3 years after the speed limit change, at which point prices begin to rebound.

Establishment Interactions

Figure 1.14 graphs the coefficients on the uninteracted establishment variables (β_4^{bin} in Eq. 1.3). These estimates cannot be considered causal, but are nevertheless interesting

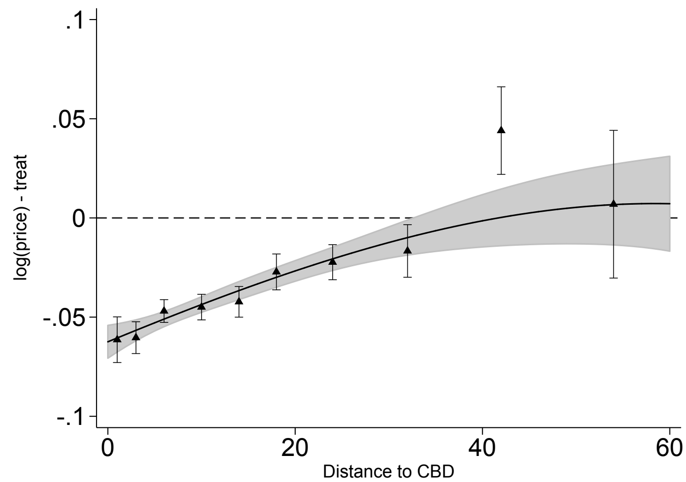


Figure 1.12: Treatment-term heterogeneity by distance from CBD.
Line is a 3^{rd}-order polynomial. Point estimates are based on a binned model where the estimate is at the centerpoint. The estimation includes census-tract level trends, city-specific month dummies, house age, house age squared, house size, and house size squared. Standard errors are two-way clustered by city and month-of-sample.

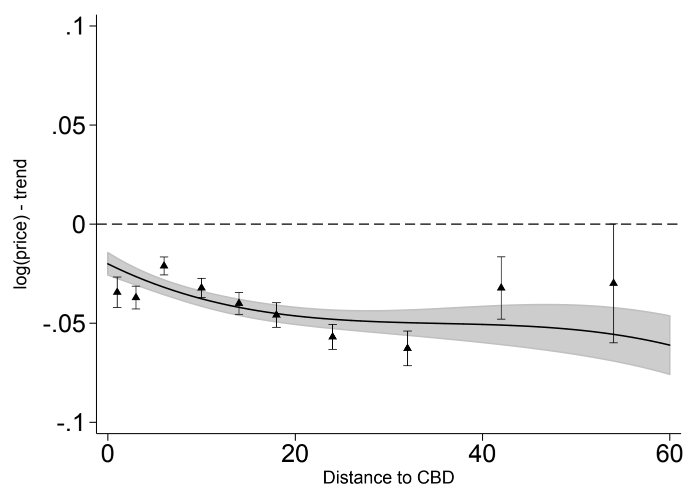


Figure 1.13: Trend-term heterogeneity by distance from CBD.
Line is a 3^{rd}-order polynomial. Point estimates are based on a binned model where the estimate is at the centerpoint. The estimation includes census-tract level trends, city-specific month dummies, house age, house age squared, house size, and house size squared. Standard errors are two-way clustered by city and month-of-sample.

in this early application market access to urban economics. They may represent the magnitude of a general equilibrium effect.

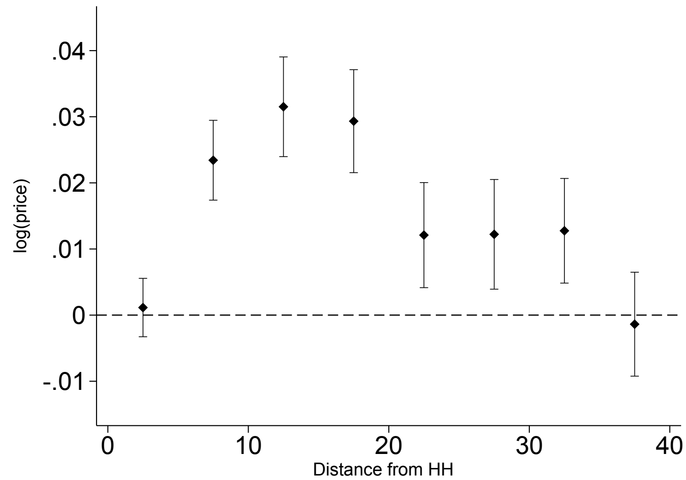


Figure 1.14: Coefficients on logged establishment count bins.
The estimation includes census-tract level trends, city-specific month dummies, house age, house age squared, house size, and house size squared. Standard errors are two-way clustered by city and month-of-sample.

Since the functional form is log-log, the interpretation is an elasticity of home prices to nearby business establishments. The general results for the two time periods are similar. A 1% increase in establishments between 10 and 20 miles away increases a property's price by about 0.03%. The effect is smaller for closer and more distant establishments.

Figure 1.15 graphs the coefficients on the treatment-interacted establishment variables (β_1^{bin} in Eq. 1.3). Here, I leverage the speed limit treatment to causally estimate the marginal effect of an additional business in each bin on treatment. When speed limits increased, a 10% increase in businesses between 15 and 20 miles increases the treatment effect by almost 0.1 percentage points. The magnitude is similar for the Uber treatment for businesses between 5 and 10 miles.

Consistent with theory, the market access effect with speed limits is for distant establishments. The intuition is that a higher speed limit makes traveling long distances less costly.

Housing Development

How does each treatment affect housing development within the city? I look at outcomes of interest for homes built just before and just after treatment. Outcomes include distance to CBD (as a measure of the rate-of-sprawl), house size, and lot size. Since I only have data on the year a house was built, rather than the specific date, some controls are no longer useful. The estimating equations returns to the basic trend-break (Equation 2.1).

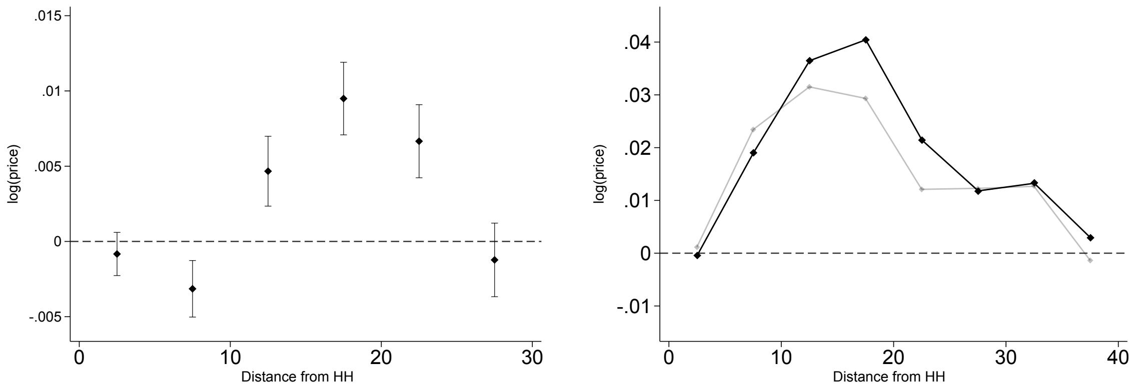


Figure 1.15: Coefficients on treatment interaction with logged establishment count bins.
The estimation includes census-tract level trends, city-specific month dummies, house age, house age squared, house size, and house size squared. Standard errors are two-way clustered by city and month-of-sample.

Table 1.3 shows the effect of the speed limit increase on development outcomes. Following an increase in relative prices of suburbs, the speed limit change spurs more distant development. The average distance from the CBD of new construction increases by 2.7% in the first year and an additional 3.6% in subsequent years. Consistent with a story of suburban sprawl, the new homes are larger and built on larger lots.

Table 1.3: Effect of speed limit increase on housing development outcomes.

	(1) log(distance to CBD)	(2) log(house size)	(3) log(lot size)
$\mathbb{1}$\{Speed limit increase\}	0.0270***	0.00266***	0.0359***
	(13.05)	(4.64)	(15.32)
$\mathbb{1}$\{Speed...\}*Trend	0.0363***	0.00324***	0.0314***
	(28.22)	(6.95)	(15.80)
Observations	2850869	3180394	3066163

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Note: Standard errors are two-way clustered by city and year.

To see whether the treatments change the number of houses being built, I collapse the data to the tract level, with a count of houses built in each year. I estimate the effect of treatment on the count of new houses. Table 2.3 shows that the speed limit increase also increased the number of houses being built.

Table 1.4: Effect of speed limits on rate of development.

	(1) log(homes built)
$\mathbb{1}$\{Speed limit increase\}	0.00361 (0.19)
$\mathbb{1}$\{Speed...\}*Trend	0.0529*** (3.59)
Observations	129874

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Standard errors are two-way clustered by city and year.

Figure 1.16 interacts the trend term with distance to CBD to see where in the city the increase in development occurs. For speed limits, there are two local maxima: one close to the city center and another 40 miles outside the center. Theory predicts the growth in suburban development due to the relative price change. The growth in central development does not have a theoretical grounding. One possible explanation that should be explored in future work is how the speed limit change affected traffic. If traffic near the central city got worse then there would be demand for housing close to the CBD.

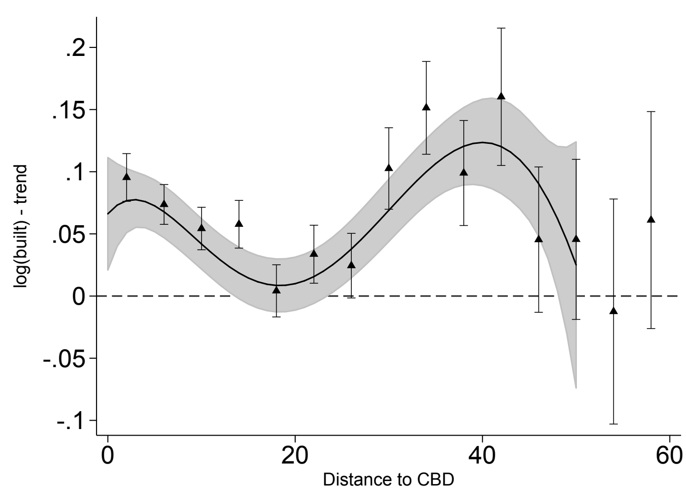


Figure 1.16: New Development Trend-term heterogeneity by distance from CBD. Line is a 5^{th}-order polynomial. Point estimates are based on a binned model where the estimate is at the centerpoint. The estimation includes census-tract level trends and year fixed effects. Standard errors are two-way clustered by city and year.

1.8 Conclusions

This paper uses a novel dataset to understand how urban real estate prices and development respond to a diffuse transportation innovation. An increase in speed limits equates to an economic technology shock that is immediately adoptable by millions of urbanites, making this paper unique from past papers that have estimated the effects of infrastructure projects. In addition, by observing the same treatment repeatedly across a large number of cities I am able to apply a rich set of controls. Increasing speed limits represents improvements in travel time, market access, and convenience. I find that increasing speed limits results in an abrupt, significant, and robust drop in urban housing prices. The result is largest closest to the CBD and goes to zero in the suburbs as suburbs become more attractive relative to the central city. This increases housing development, and the new housing is more sprawled, on average. Home sizes and lot sizes also increase, which are both associated with suburban housing. These result could help inform housing policy since there is little empirical economic evidence on how transportation interacts with urban real estate development.

I bring the concept of market access to the urban literature, applying it to detailed spatial data, and finding that surrounding spatial organization matters. Higher speed limits result in better access to distant businesses. A logical next step is to explore heterogeneity by business sectors and understand how much of the effect is attributable to better access to jobs versus consumption.

Given the rapid innovation in transportation technology, understanding how changes affect marginal development will be important for planning city growth and designing policies to control pollution and energy use. As global cities grow, so do the complications and externalities associated with automobile travel. This paper provides the first empirical evidence on the causal effect of diffuse transportation innovations. It represents a starting point for thinking about the future of cities under continually falling transportation costs.

Chapter 2

Transportation, Market Access, and Urban Development: Evidence from Uber

2.1 Introduction

Over the past decade, transportation technology has changed rapidly. Cities have seen the introduction of countless new forms of transportation. Today, an urban commuter can choose from a spectrum of transportation modes, from public buses to rentable electric kick-scooters (see BIRD and Lime-S Electric Scooters, or Scoot for Electric Mopeds), for her journey to work. Combining different transit types may be the most efficient for certain routes. Take a crowdsourced shuttle route (see Chariot) and then pick up a bike from a bikeshare station (companies differ by city, see Ford GoBike in the Bay Area). Or use your phone to find an unstationed bikeshare bike on a sidewalk (see ofo and JUMP). Start riding and leave it on the sidewalk for the next rider when you reach your destination. JUMP bikes even come with an electric motor to help you up any steep hills.

Perhaps the most pervasive new transportation technology has been "ridesharing". The two most popular services, Uber and Lyft, allow anyone with a smartphone to order a car, which arrives within minutes and takes you to your destination for a quoted price. Although the service is similar to a taxi, it is generally less expensive, more reliable, and more convenient. Each of these factors has probably contributed to the success of these Uber and Lyft, which, as of April 2018 have received estimated valuations at $72 billion and $11 billion, respectively[1]. Anecdotally, these services have allowed individuals and families to reduce their dependence on private cars, while helping to solve the "last mile logistics" problem for public transportation. However, little empirical evidence exists on how these services are used and their impacts, in part because Uber and Lyft keep data about usage private.

In this paper, I show, using publicly available data, that Uber is changing real estate

[1] see Uber value and Lyft value.

in cities. Transportation and urban growth are fundamentally linked. Theoretical models of urban growth and structure depend on a transportation technology. The workhorse model in urban economics, called the monocentric city model and developed by Alonso (1964), Mills (1967), and Muth (1969), includes a mode of transportation that moves workers from their homes to a central business district (CBD) where they work. An extended version of this model, presented in section 2.2, can provide insight into how a city changes with the introduction of a new mode of transportation.

I empirically measure the causal effect of Uber launching in a new city on real estate prices and housing development in that city. I use the exact timing of the rollout of Uber, which depends on the efficacy of localized management teams and city-specific policies, to estimate the effects. Although ridesharing companies likely employed strategies to determine where and when to begin service, much of the timing decision depends on a local management team and its ability to recruit drivers. The abrupt and staggered timing of these events allows me to estimate the change in both the level and growth rate of real estate prices and new construction. Using a novel dataset, I describe the spatial structure of the effect while controlling for high-resolution spatial- and temporal-controls.

In addition, I measure routed distances between over 2.8 million residential real estate transactions and 7.5 million business establishments across 177 U.S. cities in order to apply the concept of market access (a measure of the transportation cost to a particular market), which was developed in the trade literature. I show how the spatial structure of business establishments around a home changes the impact of a transportation innovation. This, alongside Chapter 1, is the most spatially resolved market access analysis to date.

This chapter provides further quasi-experimental evidence of how diffuse transportation innovation changes real estate values, urban development, and population in cities. Analyzing real estate prices and development within each city before and after each change, interacted with the distances to nearby markets, I find large impacts. The availability of Uber increases housing prices near the center of cities and leads to faster urban development. Integrating market access in an urban setting is important: I find that nearby business establishments make the Uber effect larger.

The paper proceeds as follows: Section 2.2 integrates market access into an urban theoretical framework to provide intuition about the effects of ridesharing on urban real estate. In section 2.3, I summarize the existing empirical literature on ridesharing and describe my contributions. Section 2.4 describes ridesharing in detail. Section 2.5 presents my empirical design. Section 2.6 explains that data that I use in section 2.7, empirically estimating the relationship between transportation innovation and real estate prices and growth. Section 2.8 concludes and provides directions for future research.

2.2 Theoretical Framework for Ridesharing

I add ridesharing into the theoretical framework introduced in Chapter 1 (see section 1.2). First, add mode choice back into the transportation cost equation. Recall that s^i is the average speed of traveling by mode i, τ_f^i is a fixed cost of mode i, τ_v^i is the price per mile travelled using mode i (which I call the variable cost of the mode), and time is valued at the wage, w. The total cost of traveling a distance d using mode i is $T(d, s^i) = \tau_f^i + \tau_v^i d + wd/s$. Individuals choose their choice of travel for all trips, and the choice depends on distance. Illustrating the intuition, Figure 2.1 displays the total cost of travel at different distances to the CBD for commuting. The agent chooses to adopt ridesharing if $\tau_f^{ride} + \tau_v^{ride}x + wx/s^{ride} < \tau_f^{car} + \tau_v^{car}x + wx/s^{car}$. Assuming the speed of a car and a rideshare are the same, the time-dependent terms cancel out, simplifying the choice to adopt if $\tau_f^{ride} + \tau_v^{ride}x < \tau_f^{car} + \tau_v^{car}x$.

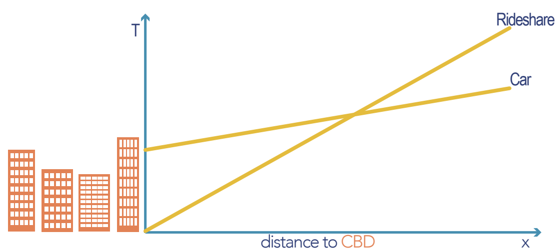


Figure 2.1: Transportation cost by distance to CBD using a private car or ridesharing
Homes closer to the CBD will find it optimal to adopt ridesharing as their new mode of transportation.

As illustrated in Figure 2.1, agents adopt ridesharing if they live closer to the CBD due to a shorter commute. However, those that adopt now face a higher marginal cost to commute. Therefore, the bid-rent function increases more near the CBD.

The result for amenities follows similar intuition. Those who live closer to amenities now face a steeper cost gradient to reach them. This may cause rideshare adopters to switch to closer amenities, in general, than the set that they visited when using a private vehicle. Recall that $\Psi^c(x, \underline{u}) = \sum_j [u_j(c_j^*) - T(d(x, y_j), s) * I_j]$, where c is a vector of length J containing consumption of each good, and I is a vector of length J indicating whether or not an individual consumes any of the good ($I_j = \mathbb{1}\{c_j > 0\}$). Given a higher marginal transportation cost, an individual choosing Uber over owning a car will benefit from additional nearby amenities. Therefore, homes with nearby amenities should see a larger price increase than similar homes distant nearby amenities.

Predictions

The theory informs the following predictions. Ridesharing should increase real estate prices, with the largest magnitude effect near the CBD and a decreasing effect with

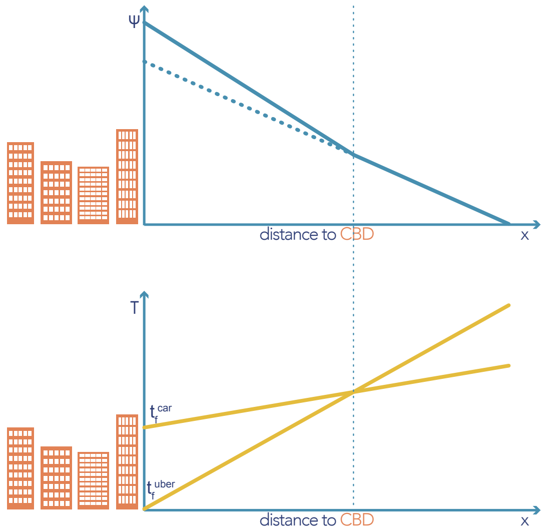


Figure 2.2: Housing prices increase near the CBD where households adopt ridesharing.

distance. Having more nearby amenities increases the impact of ridesharing on prices because ridesharing has a higher marginal cost of travel, making proximity more valuable.

Since development is not included in this iteration of the model (because, for tractability, housing stock and population are fixed) I make predictions based on relative prices. Uber causes the central-city to become more valuable relative to suburbs. I expect development to become less sprawled and occur closer to the city center.

2.3 Empirical Literature

Many papers have estimated relationships between urban areas and transportation technologies. This literature is summarized in Chapter 1, Section 1.3. Increasingly, there are papers that study Uber to understand a variety of economic phenomenon. In particular, labor economists have used Uber to understand new employment arrangements. Katz and Krueger (2016) describes the rise in more flexible work arrangements. Drivers for Uber and Lyft are free to make their own schedules, and often have other part- or full-time jobs. Hall and Krueger (2016) also focuses on the labor aspects of Uber, describing survey results in which drivers are drawn to Uber in order to smooth income fluctuations and work additional hours. Drivers, on average, are more like the general workforce than like taxi-drivers in terms of their demographics.

Another paper studying driver behavior and demand for flexible work is Angrist, Caldwell, and Hall (2017). They examine an experiment where Uber drivers in Boston were given the opportunity to work based on the taxi-model where they are not required to pay a proportion of their fares to Uber but instead pay a fixed cost for a medallion. The authors find that drivers tend to prefer the ridesharing model, which provides more flexibility, with some drivers declining the medallion offer even when they would have been better off by purchasing it. This result suggest there are gains for drivers to the ridesharing model, which doesn't involve a fixed-cost purchase like a taxi medallion.

There has also been some research on how Uber affects consumers. Cohen et al. (2016) estimate the consumer surplus from Uber by reconstructing the demand curve using pricing fluctations (known as "surge" pricing at the time of the paper's publication - Uber no longer uses this term). They estimate that UberX, the service that I focus on here, generated approximately $2.9 billion in consumer surplus in just four cities in 2015. This implies a total consumer surplus of $6.8 billion. The implications of my paper, discussed later, suggest that some of this surplus may be captured in housing markets. Less attention has been focused on the social costs of Uber. One exception is Li, Hong, and Zhang (2016). They find that cities that have adopted Uber have also witnessed an increase in traffic congestion.

There have been no quantitative studies on how Uber changes decisions about where to live and how to move around cities. This is the first paper about how Uber has changed real estate prices and urban structure.

2.4 Empirical Setting

Ridesharing, which allows anyone with a smart phone to order a ride within minutes, has emerged in cities across the U.S. over the last 5 years. Uber and Lyft, the largest ridesharing providers, have changed the way many urbanites move around cities. These "rideshare services" allow smartphone users to order a ride quickly and easily. Ridesharing is more convenient than a taxi and often far less expensive. Uber, which originally provided a more expensive towncar service, introduced a cheaper product called "UberX", which allows individuals to drive users requesting a ride to and from a requested destination. Lyft, a smaller company in most markets, primarily provides a similar service to "UberX". I estimate the effect of UberX launch in U.S. cities and the growth of real estate prices. Figure 2.3 shows the cities in which UberX operated by the end of 2016.

When Uber launches service in a new city, the company appoints a city manager and team to recruit drivers and promote the transportation mode. Uber and Lyft are strategic about when to launch in a city. However, the companies are also constrained by the available staff and resources within in the city, creating some randomness to the actual launch dates. In addition, during the early stages of rideshare companies, the analytics teams were small relative to the number of data scientists employed at ridesharing companies today. Given my empirical design, discussed below, the company would have to make the rollout decision based on an expectation of faster housing growth, or a metric correlated with faster future price growth, in order to violate the identifying assumption. The cities in my data and their UberX launch dates are shown in a table in the appendix table A.1.

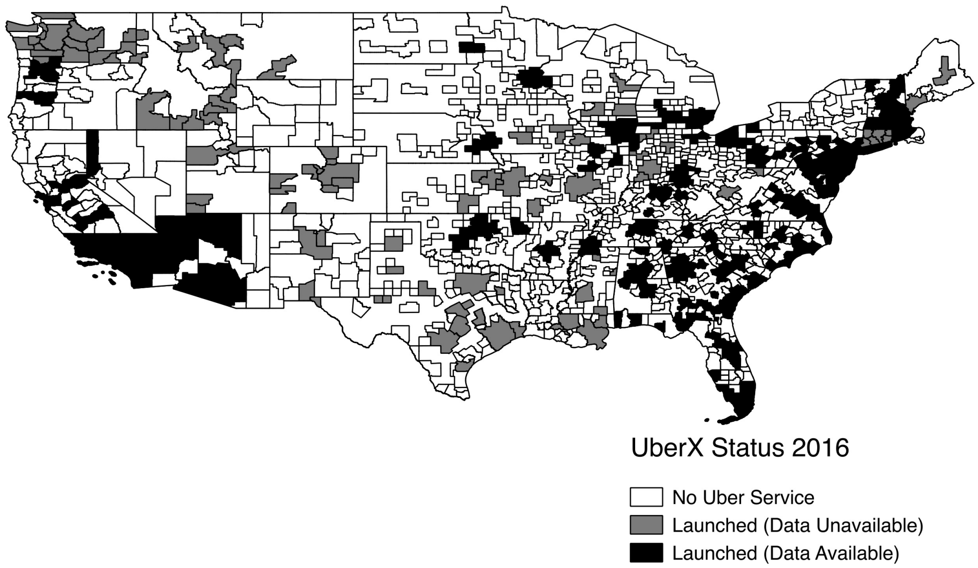


Figure 2.3: Uber launch status as of 2016 and data availability by Core Based Statistical Area.

2.5 Empirical Design

I use two strategies to measure the effect of ridesharing on urban real estate. The first strategy assumes that exact timing of each event is exogenous to my outcomes of interest, and measure the change in level and slope of each outcome variable before and after an abrupt change in transportation technology. This design is often called a "trend-break" model. In the ideal experiment, I would be able to compare an identical house in an identical city with only a difference in the availability of ridesharing. Because this comparison is impossible, I compare homes in the same census tract, before and after treatment, controlling for house-specific covariates, relative to changes in untreated cities that will receive treatment later. In practice, this means I limit my sample to transactions in treated cities, making the control homes those that will be treated but have not yet. The identifying variation is the level and rate of growth of the treated variables, at a specific location, before and after treatment, relative to the level and rate of growth of untreated variables over the same period. The identifying assumption is that the expectations of the level and rate of growth of each outcome variable, before and after treatment, conditional on covariates, are equal.

In the second empirical strategy, included in the appendix, I replace the treatment variable with google trends data indicating the popularity of the search terms "Uber" and "Lyft" in a given city. This method accounts for the relative popularity of ridesharing within a city over time. Katz and Krueger (2016) verify that Google Trends data is closely correlated with the labor market for alternative work arrangements, which include Uber and Lyft drivers. Since the market equilibrium includes both drivers and riders, the data should also be correlated with ridership. The advantage of this strategy is that it allows time for the treatment to take effect. When "Uber" is first launched in a city, it is likely to take time for people to learn about the new service, find the best way to use it, and change their transportation behavior.

Econometric Model

A basic trend-break model regresses the outcome of interest on a treatment, trend, and treatment-trend interaction:

$$\begin{aligned} log(outcome_{ict}) = &\ \beta_1 \mathbb{1}\{t > T\}_{it} \\ &+ \beta_2 \mathbb{1}\{t > T\}_{it} * (Trend_t - T) \\ &+ Trend_t * \rho_c + \mu_t + \varepsilon_{ict} \end{aligned} \tag{2.1}$$

where T is time of ride-share launch or speed limit change, $Trend_t$ is number of time period since beginning of sample, c is a spatial unit larger than the unit of observation i.

In this case, if i is a property, then I could regress the price of property i at time t on a binary variable that indicates whether the speed limit has changed and a variable that interacts the trend with the binary treatment. The pre-trend can be a smaller spatial unit. I control for the pre-trend in each census tract to account for differential growth rates within cities. Figure 2.4 shows the graphical interpretation of a trend-break.

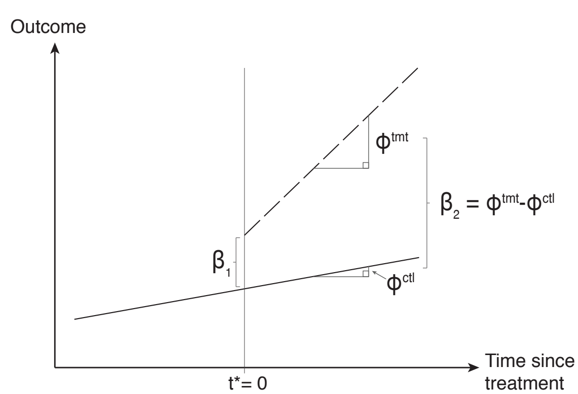


Figure 2.4: The trend break model measures a change in level and slope.

In my preferred specification shown in equation I add additional controls to account for observable and unobservable variation in the outcome of interest for property i in census tract c during month t:

$$\begin{aligned} log(outcome_{ict}) = & \beta_1 \mathbb{1}\{t > T\}_{it} \\ & + \beta_2 \mathbb{1}\{t > T\}_{it} * (Trend_t - T) \\ & + Trend_t \times \rho_c + \mu_t + CBSA_i \times month_t \\ & + \psi_1 age_i + \psi_2 age_i^2 + \theta_1 size_i + \theta_2 size_i^2 + \varepsilon_{ict} \end{aligned} \tag{2.2}$$

where T is date of Uber launch, $Trend_t$ is number of time period since beginning of sample, $Trend_t \times \rho_c$ are census-tract level trends, $CBSA_i \times month_t$ are city-specific month dummies, age_i is the number of years since the house i was built, and $size_i$ is the square footage of house i.

Interacting the Trend (Treatment) parameter with a distance to CBD and establishment bins gives heterogeneous effects on the change in slope (level) at different locations within a city. For simplification, I define $treat_{it} = \mathbb{1}\{t > T\}_{it}$ and $post\_trend_{it} = \mathbb{1}\{t > T\}_{it} * (Trend_t - T)$.

$$\begin{aligned} & log(outcome_{ict}) = \\ & \underbrace{\beta_1 treat_{it} + \beta_2 post\_trend_{it}}_{\text{baseline trend-break}} \\ & + \underbrace{\sum_{k=1}^{n} Distance_i^k \times [\gamma_{0k} + \gamma_{1k} treat_{it} + \gamma_{2k} post\_trend_{it} + \gamma_{3k} \times trend_t]}_{\text{distance-to-CBD interactions}} \\ & + \underbrace{\sum_{m=1}^{8} log(Establishments_{itm}) \times [\delta_{0m} + \delta_{1m} treat_{it} + \delta_{2m} treat_{it} \times post\_trend_{it} + \delta_{3m} \times trend_t]}_{\text{binned-establishments interactions}} \\ & + \underbrace{trend_{ct} + \mu_t + CBSA_i \times month_t}_{\text{spatial- and temporal- controls}} \\ & + \underbrace{\psi_1 age_i + \psi_2 age_i^2 + \theta_1 size_i + \theta_2 size_i^2}_{\text{home-specific attribute controls}} + \varepsilon_{ict} \end{aligned} \tag{2.3}$$

where T is date of Uber launch, $Trend_t$ is number of time period since beginning of sample, $trend_{ct}$ are census-tract level trends, $Distance_i^k$ is the distance from property i to the CBD to the k^{th} power, $Establishments_{itm}$ is the number of business establishments between $5 \times (m-1)$ and $5m$ miles from property i at time t, $CBSA_i \times month_t$ are city-specific month dummies, age_i is the number of years since the house i was built, and $size_i$ is the square footage of house i. I estimate equations 2.2 and 2.3 in the next section.

2.6 Data

To estimate the effect of ridesharing on real estate prices and development, I use over 3 million residential real estate transactions from 28 states spanning 2010 - 2016. I link these transactions with 7.5 million business establishments using routed road distances. To my knowledge, this has never been done in the economics literature. For a detailed description of the creation of this dataset, see Section 1.5, as the dataset creation is identical to the methods for speed limits but applied to a later time period. The subsections titled "Housing Prices and Attributes", "Business Establishments", "Distances", and "Binning Establishments" all apply.

Uber and Lyft

Data on Uber and Lyft are combined from three sources. First, Uber provided the launch dates of UberX for each city in the U.S. Second, I find UberX service area as of March 2017 by querying Uber's API for property. Third, I download city-specific Google Trends for each metropolitan area for which the trends are available. Google Trends data and results are described in detail in the Appendix

2.7 Results

Trend-Break

To illustrate that a trend-break model fits the empirical setting, I first estimate a flexible specification, creating dummy variables that represent the periods before and after treatment. I use four-month intervals, and start 20 months before treatment. Let p index the 4 month periods before and after treatment, with treatment occurring between zero and one. Then $running_{it}^{p} = 1$ if $p - 1 < (t - T)/4 \leq p$, where T is treatment date and t is transaction date.

I estimate the equation

$$\begin{aligned} log(prices_{ict}) = &\sum_{p=-5}^{max} \alpha_p \times running_{it}^{p} \\ &+ Trend_t \times rho_c + \mu_t + \gamma CBSA_i \times month_t \\ &+ \psi_1 age_i + \psi_2 age_i^2 + \theta_1 size_i + \theta_2 size_i^2 + \varepsilon_{ict} \end{aligned} \tag{2.4}$$

where max for 9 periods after Uber rollout due to the end of my data sample. This specification corresponds to Equation 2.2, with treatment and trend variables replaced by dummy variables, allowing them to vary non-linearly. Figure 2.5 graphs the coefficient and standard error for each $running_{it}^{p}$, and a linear-trend is estimated through the pre- and post- treatment coefficients to represent the trend-break that is attributable to treatment.

The pre-trend is flat, suggesting that the tract-level trends, city-by-month fixed effects, and month-of-sample fixed effects, combined with housing covariates, remove variation that is unrelated to treatment. After treatment there is a positive treatment term and a positive trend term.

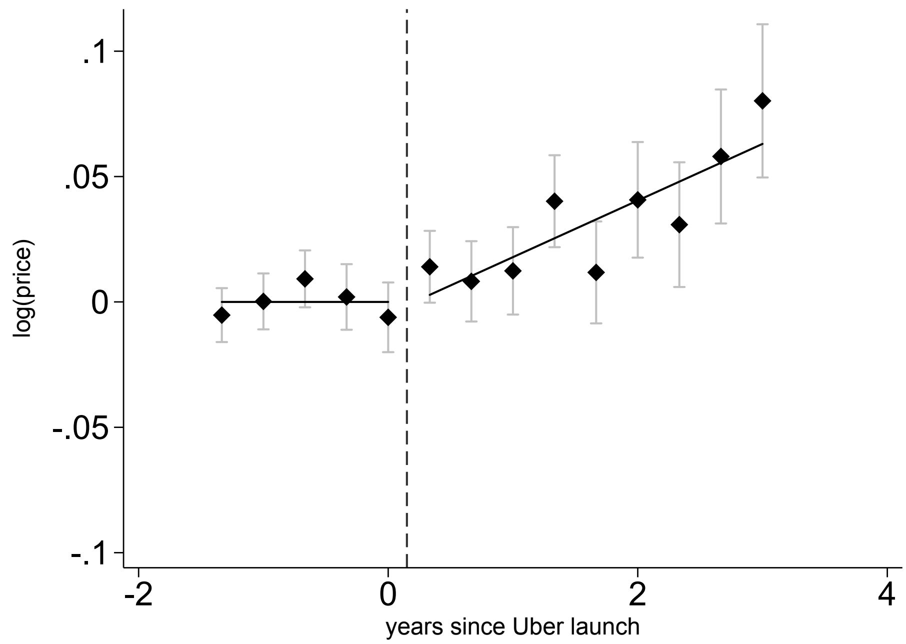


Figure 2.5: Effect of uber on real estate prices: Trend-break model
Point estimates are dummy variables for 4 month periods before and after treatment. The estimation includes census-tract level trends, city-specific month dummies, house age, house age squared, house size, and house size squared. Standard errors are two-way clustered by city and month-of-sample. Linear trends are estimated through dummy estimates before and after treatment

The results are robust to covariates, with tract-specific trends removing unobserved variation that is correlated with treatment. In the preferred specification that includes all covariates, the treatment effect is positive but insignificant, while the trend effect is 2.3%. This is consistent with the there being a window of time for consumers to adopt the new technology and adjust their transportation habits optimally. The results are also robust to selection of different sample periods, as shown in appendix table 1.2.

Distance Interactions

Theory predicts treatment to be heterogeneous by distance to the city center. Ridesharing should cause prices to increase most near the center. I interact the treatment and trend variables with a 3^{rd}-order polynomial in distance to the CBD, estimating the marginal effect of treatment for transactions at varying distance from the center. The treatment term effects are shown in Figure 2.6. They correspond to the marginal effects of β_1^k in Equation 2.3: the effects on prices just before and just after the treatment date at every distance from the CBD. Figure 2.6 also shows the results of binned distance estimates, allowing for a non-parametric relationship between distance and effect. While the results are similar to the polynomial, the binned approach estimates some in-

Table 2.1: The effect of Uber on real estate prices.

	(1) log(price)	(2) log(price)	(3) log(price)	(4) log(price)
𝟙{Uber launched}	0.0253***	0.000337	-0.00104	0.00423
	(9.33)	(0.12)	(-0.36)	(1.43)
𝟙{Uber...}*Trend	0.0539***	0.0138***	0.0129***	0.0229***
	(28.97)	(3.75)	(3.49)	(5.97)
Observations	2881729	2881729	2881729	2564156
Year FE	X	X	X	X
Month-of-sample FE	X	X	X	X
Tract-specific trend		X	X	X
City-by-month FE			X	X
House covariates				X

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Standard errors are two-way clustered by city and month-of-sample. Column 4 is estimation of Equation 1.2, which includes month-of-sample fixed effects, tract-specific trends, city-by-month fixed effects, and house specific covariates.

crease in housing prices at far distances, which suggests growth in the suburban housing market. This is discussed in the results subsection on development and sprawl.

The treatment effect is positive near the center of the city, but becomes negative at farther distances. The increase in the center of the city is consistent with theory that rideshares are useful for short rides but not for long commutes due to their high marginal cost. The negative effect on suburbs may be a mechanical effect due to a misclassification of the exact timing that rideshares moved into different areas of the city. When Uber and Lyft launched, the service territory in each city was very limited. It has expanded over time to serve more suburban areas. By treating the entire city at once, the actual treatment effect may be captured in the trend, which I discuss next.

The heterogeneous trends in prices, shown in Figure 2.7 continue in the same direction as the treatment effect. This ongoing trend could reflect the continued capitalization of the Uber into housing prices. It could also be ongoing changes that are not captured by the treatment variables. The increasing trend effect with distance from CBD may reflect an expansion of service territory over time. As of September 2017 (after the end of my housing sample), Lyft announced that they would provide coverage across 40 entire states. This makes the service territory endogenous, since it is still necessary to have a driver in the vicinity in order to get a ride. Because I cannot verify the service territory before March 2017, the positive trend, which is higher in the initially untreated areas, may be due to capitalization in more communities. However, in the areas that are known to be treated, close to the city center, the effect is positive

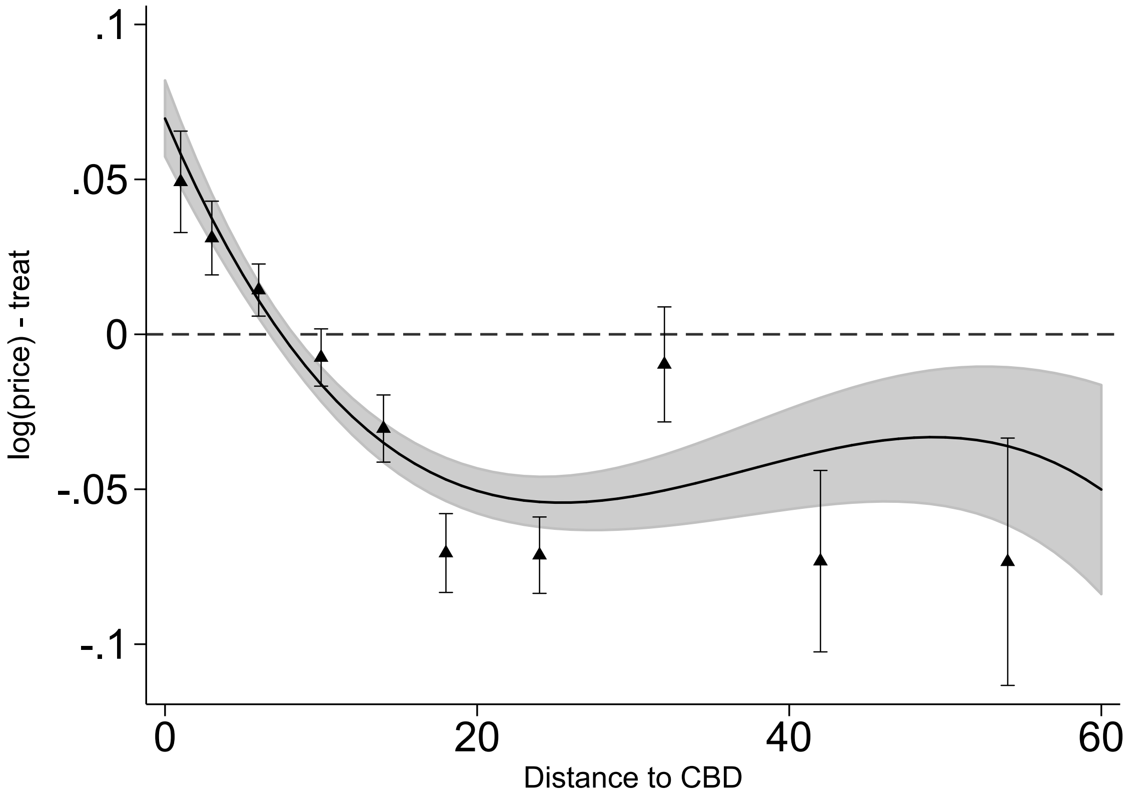


Figure 2.6: Treatment-term heterogeneity by distance from CBD. Line is a 3^{rd}-order polynomial.

Point estimates are based on a binned model where the estimate is at the centerpoint. The estimation includes census-tract level trends, city-specific month dummies, house age, house age squared, house size, and house size squared. Standard errors are two-way clustered by city and month-of-sample.

and significant. Most Uber launches have been too recent to estimate the continual price change past two years, but the estimates will be extended in future work.

Establishment Interactions

How does the network of business establishments around a house effect the magnitude of the treatment effect? Figure 2.8a graphs the coefficients on the treatment-interacted establishment variables (β_1^{bin} in Eq. 1.3). Here, I leverage Uber treatment to causally estimate the marginal effect of an additional business in each bin on treatment. When Uber was introduced, a 10% increase in businesses between 5 and 10 miles increases the treatment effect by almost 0.1 percentage points. Figure 2.8b graphs the coefficients on the uninteracted establishment variables (β_4^{bin} in Eq. 2.3) in grey and the baseline value plus the Uber effect in black. Uber has the largest effect on houses where there are more establishments within 15 miles. This is consistent with theory: the market

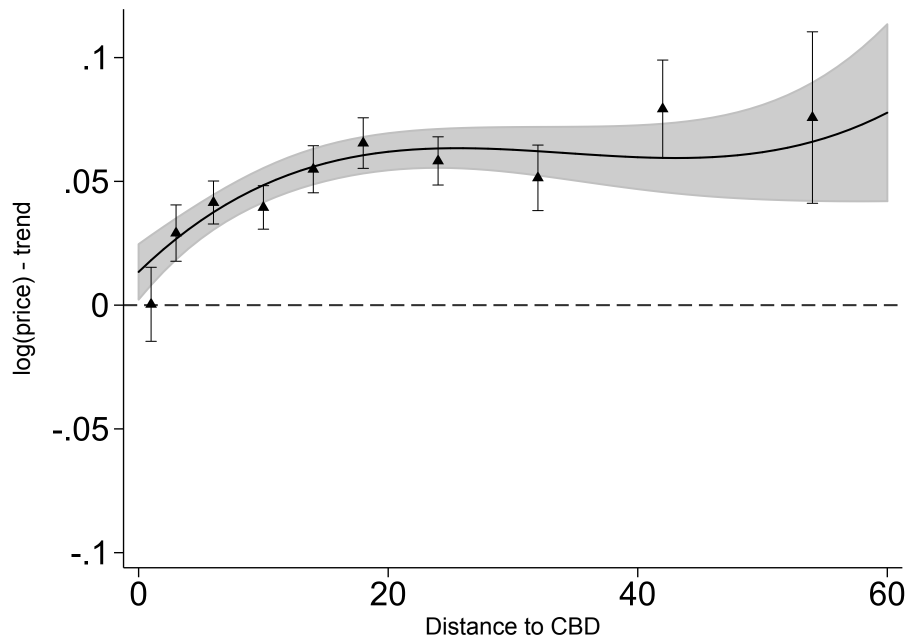


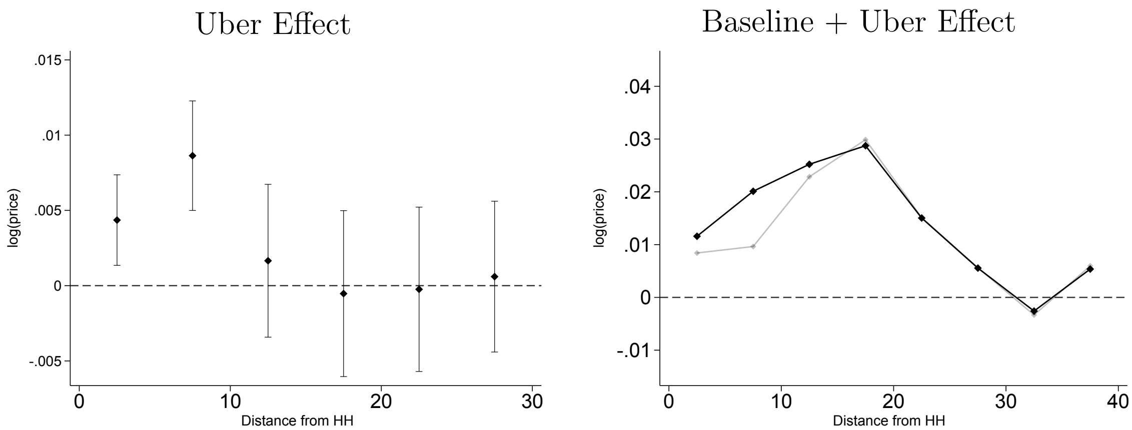
Figure 2.7: Trend-term heterogeneity by distance from CBD. Line is a 3^{rd}-order polynomial.
Point estimates are based on a binned model where the estimate is at the centerpoint. The estimation includes census-tract level trends, city-specific month dummies, house age, house age squared, house size, and house size squared. Standard errors are two-way clustered by city and month-of-sample.

access effect is for nearby establishments because Uber makes traveling short distances less costly overall.

The baseline coefficients shown in Figure 2.8b (β_4^{bin} in Eq. 2.3) should not be considered causal because their estimation is not based on a natural experiment and they may represent unobserved housing and location attributes. However, they are interesting and represent the first estimation of nearby amenities with this level of detail. Since the functional form is log-log, the interpretation is an elasticity of home prices to nearby business establishments. A 1% increase in establishments between 10 and 20 miles away is associated with a higher property price of about 0.02-0.03%. The effect is smaller for closer and more distant establishments.

Housing Development

How does each treatment affect housing development within the city? I look at outcomes of interest for homes built just before and just after the Uber launch. Outcomes include distance to CBD (as a measure of the rate-of-sprawl), house size, and lot size. Since I only have data on the year a house was built, rather than the specific date, some controls are no longer useful. The estimating equations returns to the basic trend-



(a) Coefficients on trend interaction with logged establishment count bins.

(b) Coefficients on baseline establishment bin counts plus effect after one year.

Figure 2.8: Effect of treatment after one year on the value of a marginal business establishment.
The estimation includes census-tract level trends, city-specific month dummies, house age, house age squared, house size, and house size squared. Standard errors are two-way clustered by city and month-of-sample.

break (Equation 2.1).

Table 2.2 shows the effect of Uber on development outcomes. The average distance from the CBD of new construction decreases by 14% in the first year and an additional 3.1% in subsequent years. Consistent with a story of urban redevelopment, the new homes are smaller and built on smaller lots. Average house size fell by 1.5% and lot size by 8% per year, on average, since Uber launched.

Table 2.2: Effect of ridehsharing on housing development outcomes.

	(1) log(distance to CBD)	(2) log(house size)	(3) log(lot size)
𝟙{Uber launched}	-0.137***	0.000974	-0.00920
	(-15.99)	(0.46)	(-1.15)
𝟙{Uber...}*Trend	-0.0311**	-0.0152***	-0.0838***
	(-2.93)	(-5.54)	(-8.89)
Observations	217013	211846	198651

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Note: Standard errors are two-way clustered by city and year.

To see whether the treatments change the number of houses being built, I collapse the data to the tract level, with a count of houses built in each year. I estimate the effect of each treatment on the count of new houses. Table 2.3 shows that both treatments have increased the number of houses being built.

Table 2.3: Effect of ridesharing on rate of development.

	(1)
	log(homes built)
$\mathbb{1}\{\text{Uber launched}\}$	0.0483***
	(3.91)
$\mathbb{1}\{\text{Uber...}\}$*Trend	0.0621*
	(1.97)
Observations	27547

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Standard errors are two-way clustered by city and year.

Figure 2.9 interacts the trend term with distance to CBD to see where in the city the increase in development occurs. New development attributable to ridesharing occurs closer to the CBD where prices also increase. Uber's launch spurs development within 20 miles of the CBD.

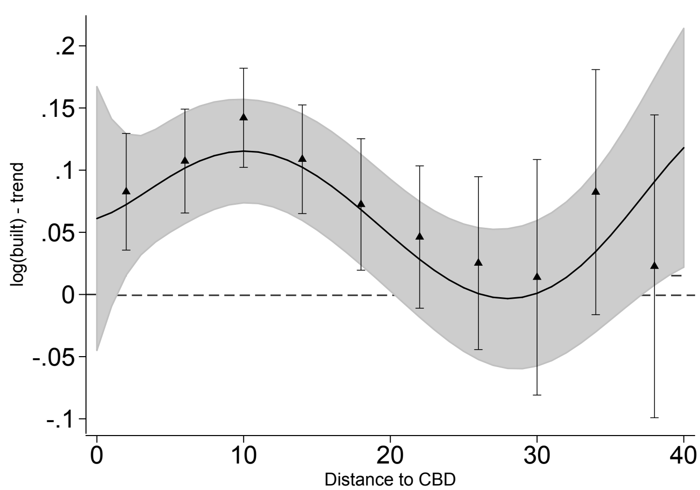


Figure 2.9: New Development Trend-term heterogeneity by distance from CBD. Line is a 5^{th}-order polynomial. Point estimates are based on a binned model where the estimate is at the centerpoint. The estimation includes census-tract level trends and year fixed effects. Standard errors are two-way clustered by city and year.

2.8 Discussion

This paper investigates how transportation technology is changing urban real estate prices and development. New transportation technologies are changing the way people move around cities. Today's new technologies are often immediately adoptable by millions of urbanites, making this paper unique from past papers that have estimated the effects of infrastructure projects and other transportation innovations.

While the rapid changes make identification challenging, observing the same treatment repeatedly across a large number of cities increases the likelihood that I have estimated unbiased, causal effects of Uber on real estate. The results suggest that access to Uber is a valuable amenity, which supports the finding in previous work by Cohen et al. (2016). The launch of ridesharing corresponds to a significant increase in urban housing prices. The result is largest closest to the CBD. Prices continue to trend upwards after treatment. Housing development increases closer to the CBD, and properties are, on average, smaller. Some of the new housing is probably redevelopment and urban infill. These result could help inform housing policy since there is little empirical economic evidence on how transportation interacts with urban real estate development.

In addition, my implementation of a market access approach in an urban context finds that the network of businesses and amenities around a property matters. The magnitude depends dramatically on transportation technology. Ridesharing improved access to nearby businesses. A logical next step is to explore heterogeneity by business sectors and understand how much of the effect is attributable to better access to jobs versus consumption. The spatial detail of this data is unlike any previous work in urban economics, and could also apply to problems that may arise from transportation innovation, including traffic, air pollution, and gentrification.

Chapter 3

Time use and Energy Consumption

3.1 Introduction*

Electricity production is the largest source of greenhouse gas emissions (EPA, 2018). Reducing electricity consumption must be part of any strategy to avoid catastrophic climate change. While economists have claimed that reductions can be achieved by pricing electricity based on the marginal cost of generation (Joskow and Wolfram, 2012), recent research suggests consumers respond little to price changes (Ito, Ida, and Tanaka, 2018; Gillan, 2017). In addition, energy efficiency measures have been proposed as an inexpensive way to curb electricity consumption. However, investing in new, expensive assets also appears to have little success at reducing usage (Fowlie et al., 2017; Allcott and Greenstone, 2012).

Less attention has been given to the behaviors that lead to the current quantities and patterns of electricity use. While electronic devices are a necessary part of our energy consumption, how and when we interact with these devices also matters. These decisions and behaviors can be conscious or unconscious, and may be seemingly unrelated to electricity use.

In this paper, we analyze over 4 million hourly observations of U.S. electricity consumption and time-use activities. We find that the percentage of the U.S. population engaged in three time-use variables - sleep, work, and leisure - explains over 90% of the variation in national electricity consumption. Moreover, modeling the relationship between time use and electricity demand allows us to precisely estimate the electricity elasticities with respect to major activity categories, revealing the percentage change in national electricity consumption attributable to a percentage of the population shifting from one activity to another.

3.2 Motivational Framework

We propose that each individual, i, consumes electricity at a given time, t. The consumption depends on her activity at time t and the set of electric devices that he or

*The material in this chapter is coauthored with Solomon Hsiang and Terin Mayer.

she interacts with. For simplicity, we assume the set of devices does not change with time, although this assumption is not strictly necessary:

$$electricity\_use_{it} = f(activity_{it}, devices_i).$$

Summing over individuals gives aggregate electricity consumption as the outcome. Because devices are unchanging over time, if we assume the set of devices is equivalent for all individuals[2], then there is another function, $g_a(\cdot)$, that gives electricity use at a given time t as a function of the number of individuals engaged in activity a. Let act_{at} be the number of individuals engaged in activity a at time t. Then:

$$\sum_t electricity\_use_{it} = \sum_t f(activity_{it}, devices_i)$$
$$electricity\_use_t = \sum_a g_a(act_{at})$$

We focus on variation in electricity consumption and time use that occurs at the day-type-by-hour-of-day level. For simplicity, we refer to this as day type by hour from here on. Day type is day-of-week (Monday through Sunday) with a separate category for federal holidays regardless of their day-of-week. Aggregating by day type, d, and hour, h, yields total electricity use for a given day-type hour. Example units of observation would be Mondays at 4:00 PM, Saturdays at 1:00 AM, or an aggregate of all holidays at 8:00 AM. We can sum across day types, d, and hours h:

$$\sum_{t \in d} \sum_{t \in h} electricity\_use_t = \sum_{t \in d} \sum_{t \in h} \sum_a g_a(act_{at})$$
$$electricity\_use_{dh} = \sum_a g_a(act_{adh}).$$

Our goal is to empirically estimate this relationship between activities and electricity use. Below, we describe the data and empirical model used to do so.

3.3 Data

Electricity

Hourly electricity load data comes from Federal Energy Regulatory Commision (FERC) Form 714 and from two Independent System Operators (ISOs): ISONE and ERCOT. The data from the ISOs are provided at the state level. The data from FERC is in load zones, often defined by utility. Each is matched to the state level for aggregation, described in detail below. Our electricity dataset contains 4,399,295 hourly observations from 47 states.

[2] This assumption is not unreasonable given a large sample

Time Use

Time-use data are from the American Time Use Survey (ATUS) conducted by the Bureau of Labor Statistics (BLS). In this survey, respondents are asked to keep a 24-hour diary of how they use their time. Interviewers subsequently code these diaries into a set of nested categories. Our analysis works with the most general 18 time-use activities: (1) "Personal Care Activities", (2) "Household (HH) Activities", (3) "Caring For & Helping HH Members", (4) "Caring For & Helping NonHH Members", (5) "Work & Work-Related Activities", (6) "Education", (7) "Consumer Purchases", (8) "Professional & Personal Care Services", (9) "Household Services", (10) "Government Services & Civic Obligations", (11) "Eating and Drinking", (12) "Socializing, Relaxing, and Leisure", (13) "Sports, Exercise, & Recreation, (14) "Religious and Spiritual Activities", (15) "Volunteer Activities", (16) "Phone Calls", (16) "Traveling", and (18) "Data Codes". The final category (18) refers to when the surveyor is unable to code the activity or the surveyee cannot remember his or her activity. The 181,335 survey responses spanning the period 2006-2016 are aggregated to create shares of the population engaged in each of 18 time-use categories during each time unit. All aggregations are performed using the frequency weights supplied by the BLS to make the sample representative of the U.S. population in terms of demographics and employment status. Results do not change significantly when weights are used or not. The proportion of the population engaged in each of the 18 time-use categories by hour of each day type (this aggregation is explained in the next section) is available in the appendix figure B.1.

Aggregation

In our primary specification, we aggregate to day type by hour at the national level. We use local time when aggregating, which is not problematic because electricity supply is rarely transmitted across time zones.

Three additional datasets with different levels of spatial and temporal resolution are also created for additional analysis. First, we create a version which maintains seasonal variation, so each observation is a season by day type by hour. The other two datasets are created for cross validation: one maintains separate data for each year - year by day type by hour - and the other maintains data by state - state by day type by hour. Electricity data is summed across states and averaged across years as needed to match the dimensions of the four datasets. The time use and electricity datasets are then merged by day type and hour, as well as season, year and state when applicable.

3.4 Model Selection

For ease of interpretation, we take the logarithm of electricity use and divide act_{adh} by the total population to make the variables population shares. Therefore, we intend to estimate a function, $h_a(\cdot)$ for each of our time-use variables of interest:

$$\log(electricity\_use_{dh}) = \sum_a h_a(prop\_act_{adh})$$

where $electricity\_use_{dh}$ is total electricity use on day type d at hour h, $prop\_act_{adh}$ is the proportion of the U.S. population engaged in activity a on day type d at hour h, and h_a is the activity specific function that describes the relationship between electricity use and activity.

To avoid over-fitting a model, we select the three time-use variables that result in the best in-sample fit. The variables that explain the most variation in electricity consumption are (1) "Personal Care Activities", (5) "Work & Work-Related Activities", and (12) "Socializing, Relaxing, and Leisure". Respecitively, we refer to these as Sleep, Work, and Leisure, since the three categories are primarily represented by the three (less cumbersomely named) subcategories. Figure 3.2 show the proportion of the population engaged in each of these activities for every hour during each day type. Seasonal versions of this figure can be found in the appendix (figure B.2)

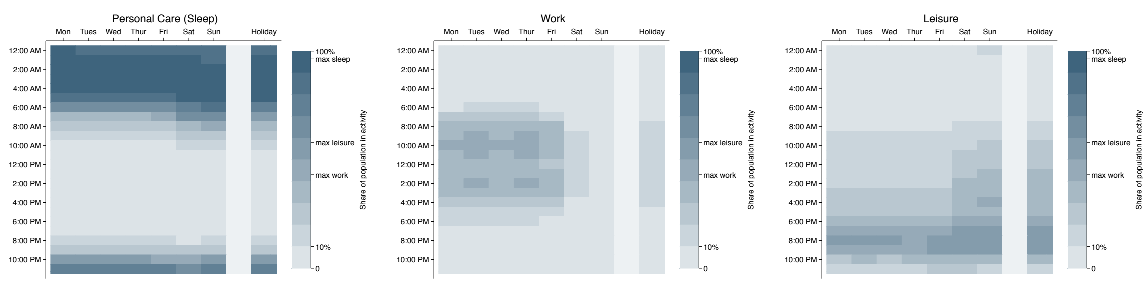


Figure 3.1: Sleep, work, leisure by hour and day type in the U.S.
Data is collapsed from the American Time Use Survey from the years 2006-2016 to create population shares. The "Holiday" day type includes all 10 federal holidays.

Thus far our model has been agnostic to the functional form of $h_a(\cdot)$, which describes the relationship between electricity use and each activity. We perform cross validation across years and states in order to select functional form (linear, cubic, or quadratic) for each of the four categories. The estimating model that maximizes out-of-sample fit is:

$$\begin{aligned}\log(electricity\_load_{dh}) = \alpha &+ \sum_{p=1}^{3} \beta^p_{sleep} * sleep^p_{dh} \\ &+ \sum_{p=1}^{2} \beta^p_{work} * work^p_{dh} \\ &+ \sum_{p=1}^{3} \beta^p_{lesiure} * leisure^p_{dh} + \varepsilon_{dh} \qquad (3.1)\end{aligned}$$

where d is day type, h is hour, and $\{sleep, work, leisure\}$ represent the proportion of the U.S. population engaged in each activity. Based on our cross validation, the function $h_a(\cdot)$ is cubic in sleep (that is, when $a = sleep$), quadratic in work, and cubic in leisure. In addition, we show some results that estimate separate parameters for each season. The regressions equation for the seasonally interacted version can be respresented by adding a seasonal subscript to each term. This is available in the appendix equation B.1.

3.5 Results

Actual and predicted electricity consumption from estimating equation 3.1 are shown in figure 3.2. Electricity consumption changes dramatically both across day types and within days. Our model, containing only three time-use variables, has excellent in-sample fit. For our primary specification, aggregated across the full year, the R^2 value exceeds 90% for each day of the week. For the aggregate holiday, we over-predict load and prediction accuracy is worse.

Electricity patterns differ by season, so we show the seasonally interacted model model in which we allow for different seasonal averages and for the electricity elasticities of the time-use activities to also differ by season (see equation B.1). Generally, the prediction remains very accurate for Spring, Summer, and Fall, and is only slightly worse during the winter. Seasonal estimates of holidays are improved, probably due to heterogeneity in time use by holiday. Chapter 4 explores this fact in detail.

Figure 3.3 depicts a set of response functions captured by our model, showing the predicted change in electricity consumption across the full extent of the population's recorded participation in the activity, holding fixed the share of the population engaged in the other activities. We find that the effects on electricity use of participation in each activity are often non-linear, suggesting that the effect of the first individual changing to a given activity is different from the last. For example, national electricity consumption increases as a growing share of the population goes to work, but that the marginal effect decreases. This is consistent with workplaces having an initial fixed electricity demand from turning on HVAC systems and lighting, spread over a greater and greater number of workers.

We estimate the same model independently for each year and then for each state to illustrate the heterogeity of the response across time and space. By and large, sleep is an energy-saver relative to work, leisure, and the non-estimated time-use categories. The response functions appear to be fairly stable across years. The state-by-state picture, however, indicates that further calibration of the model may be necessary to more accurately reflect the relationship between time use and electricity demand in some states. Our model does not include climate variables, which also represent an important predictor of electricity use and are likely to interact with activity choice. Therefore, it is not surprising that state-by-state response functions vary substantially. Understanding whether the heterogeneity can be explained by climate variables is an important direction for future research.

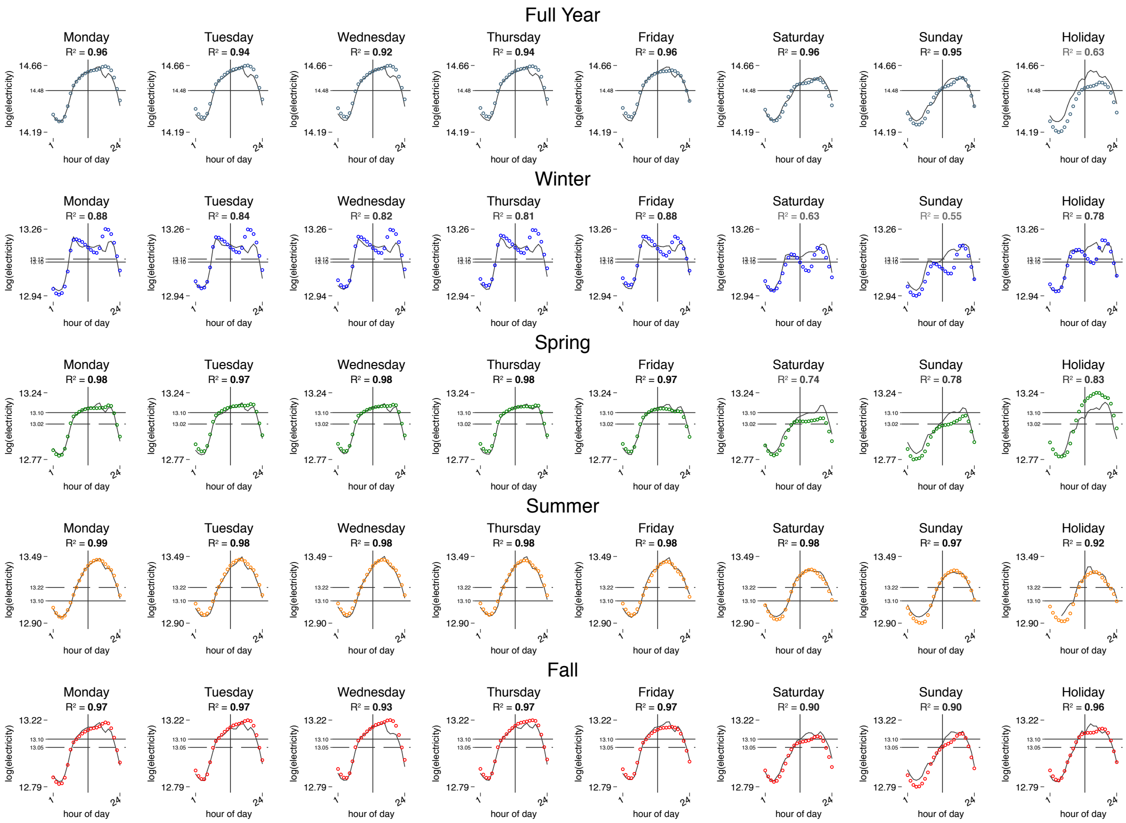
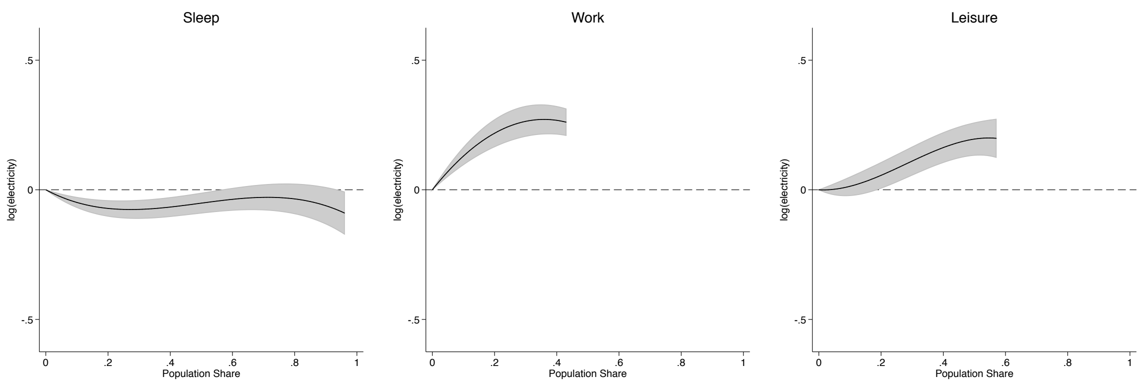


Figure 3.2: Actual and predicted electricity use by hour and day type in the U.S. Circles represent the actual 11-year national average of an hour of week of the given season, with the gray line representing our model's prediction. The "Holiday" day type includes all 10 federal holidays. The black horizontal line represents the average log of electricity demand across all seasons, with the colored dashed lines indicating the seasonal average for comparison. The goodness of fit measure R^2 is calculated per-panel, depicting the share of the total variation of electricity consumption explained by the model.

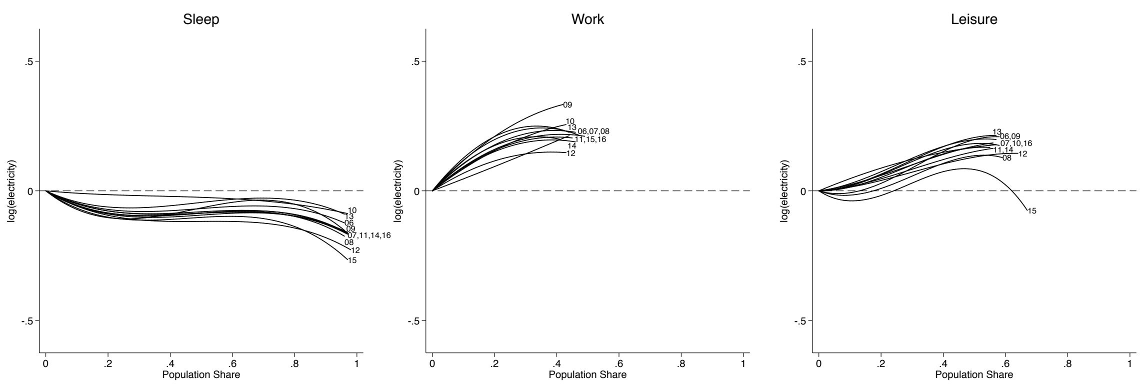
We test the predictive power of our model through two types of cross validation that bear a common logical structure. If we have uncovered a durable relationship between types of human behavior and electricity consumption, we should be able to estimate the elasticity coefficients of our model using a subset of the data available to us and predict with some accuracy the electricity demand in an hour knowing only the share of the population engaged in three activities for that hour.

We execute the cross validation exercise by using subsets of our sample defined by years (Figure 3.4a) and by states (Figure 3.4b). In the first case, we pick a year whose hourly electricity and time-use observations are used to train the model and another year as the validation data, whose time-use observations are fed to our model to make predictions, allowing for a direct comparison between the actual and predicted electricity demand. The state by state cross validation operates on a similar principle.

(a) Full Sample



(b) Year-by-year



(c) State-by-State

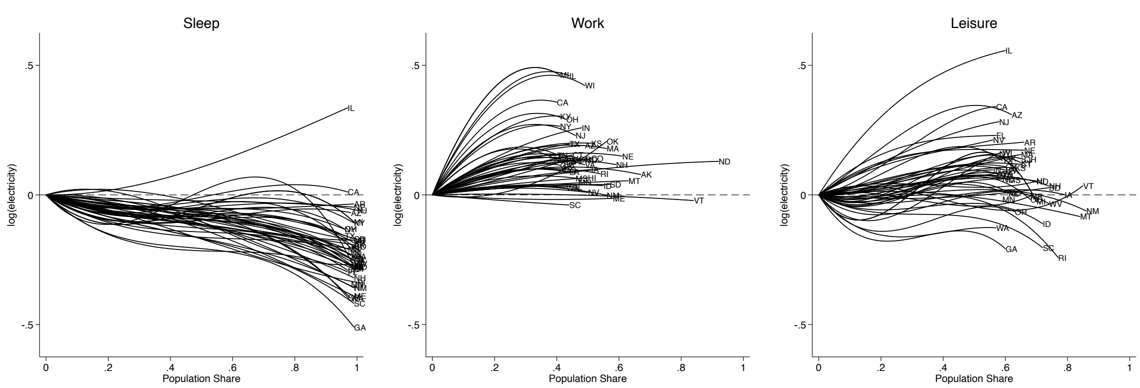
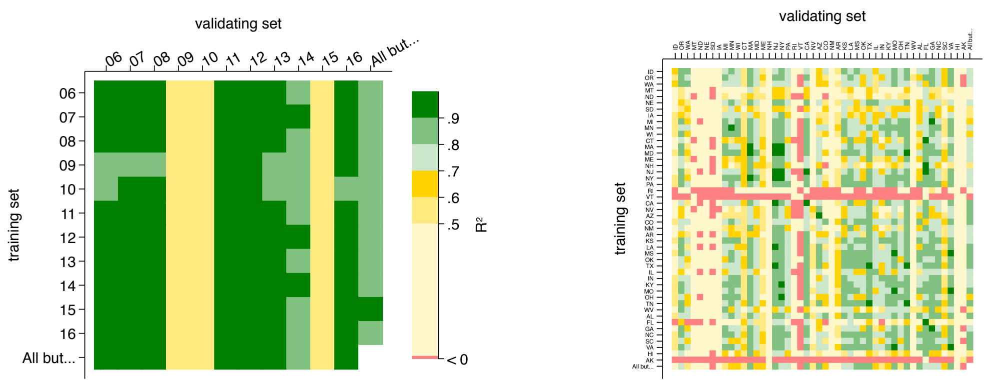


Figure 3.3: Time Use by season, hour, and day type in the U.S.
Data is collapsed from the American Time Use Survey from the years 2006-2016. The "Holiday" day type includes all 10 federal holidays.

Both the actual and predicted electricity values have their averages subtracted before the R^2 calculation occurs, since in this application we are focused on the power of the model to predict hour-by-hour changes in demand and do not wish to penalize the model's performance if it fails to predict the overall level of electricity consumption which is different across years and states due to unobserved changed. In both cases we also consider a scenario where data composed of all but a given year or state serves as the training dataset and also the validating dataset. Along the diagonal of these figures we depict the within-sample fit of the model when estimated using observations only from that year or state.



(a) Year-by-year

(b) State-by-state

Figure 3.4: Cross validation of predicted electricity use across years and states.

In general, we find that our model of electricity demand predicts much of the variation, even out sample. This is especially true when the cross validation is across years. However, particularly in the same climate-zone, state-by-state cross validation does quite well. Cross validation exercises for the seasonally interacted model are included in the appendix figure B.3.

To understand how changes in time use could change electricity use, we calculate the marginal effect of an additional participant in each of our time-use categories at each hour for weekdays and weekends. The results are shown in figure 3.5. Suppose a policy maker wanted to decrease electricity load at 6 PM on weekdays. Our results suggest that the best course of action would be to encourage activity shift from work to sleep. However, shifting from work to leisure or work to any other time-use variables (since the ommitted categories are represented by zero) would also reduce electricity use.

Several other aspects of this graph are also worth noting. First, during the middle of a weekday when many people are already working, shifting from work to leisure would actually increase energy use. This is in contrast to weekends when the marginal effect of working is always higher than the other time-use categories, because few people are working and therefore the marginal effect is high. Seasonal versions of this graph are available in the appendix, figures B.4, B.5, B.6, and B.7.

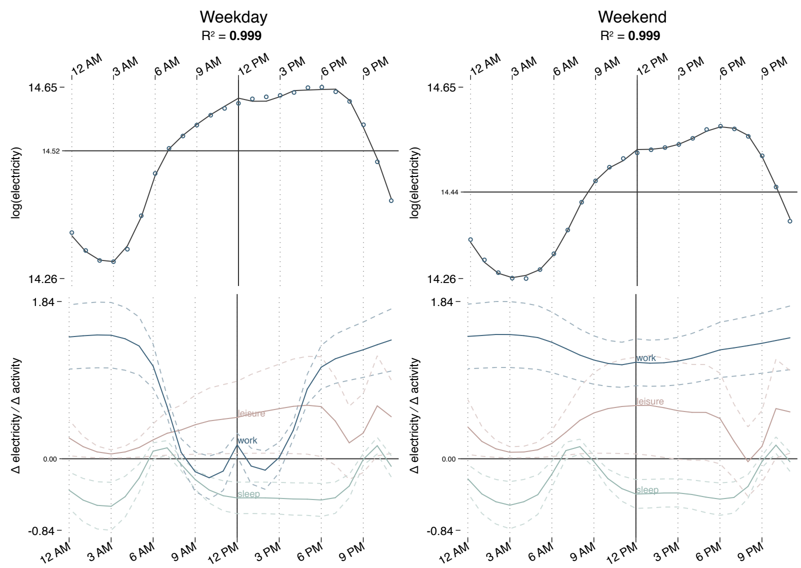


Figure 3.5: Marginal effects of participation in sleep, work, and leisure by hour of the day for weekdays and weekends.
The top panel is prediction for each hour. The bottom panel gives the marginal effect on electricity use of each time-use activity.

3.6 Discussion

Our results show that activities can be used to accurately predict electricity use. Using just three time-use variables and a simple regression, we are able to predict most of the hourly variation across a typical week. We view this as a starting point in how activities and behavior could be used to improve prediction of electricity use, which is economically valuable for utilities and may lead to novel ideas for reducing carbon emissions. Our results suggest that, insofar as demand management policies envision different behavior modification goals, some goals promise greater demand reduction than do others. Consider, for example, that a percentage shift in the population working into virtually any other activity seems to generate a net reduction in electricity demand. One could use our analysis to bound the maximum electricity demand achievable by a demand management policy whose aim is to generate such behavioral shifts.

As we seek to manage supply and demand in the electrical grid and investigate efficiency improvements that could help us mitigate and adapt to climate change, we willl need to keep an open mind on the kind of innovation that we imagine. Technology and durable good structure our electricity use, but so do habits, culture and custom. Our hope is that other researchers will now investigate what insight such models can bring to specific regional contexts and whether their predictions can be verified by experimental designs focused on the electricity elasticities of these time-use activities and the net reductions possible from achievable behavioral modification.

Chapter 4

Population Scale Coordination of Leisure Reduces Energy Consumption

4.1 Introduction*

The most recent report from the Intergovernmental Panel on Climate Change warns policymakers that continued greenhouse gas emissions will dramatically increase the likelihood of future damage and irreversible impacts (IPCC, 2014). Despite a call for reductions in emissions and some jurisdictions voluntarily adopting cap-and-trade policies, many large polluting nations are unwilling to agree to systematic solutions due to political economy obstacles. Where action has been taken, questions remain about the efficacy of these policies, and more countries will need to adopt carbon cutting measures in order avoid the worst climate change impacts.

Finding ways to achieve meaningful carbon reductions and mitigate damages may necessitate new and more politically attractive ideas. A policy dimension that has not been discussed is the allocation and coordination of labor and leisure time. In the U.S., the status-quo is a 5-day workweek, 2-day weekend, and 10 federal holidays. However, this structure is largely arbitrary. The 7-day week was probably developed in Babylon about 4,000 years ago and has no basis in nature (in contrast to years and months). There is substantial variation in the workweeks across countries, with the structure often set to accommodate the predominant religion. Figure 4.1 shows how hours worked in the U.S. has changed over time. Generally, average hours worked has fallen steadily, from nearly 70 hours per week in 1830. The 5-day workweek was popularized in 1926 by Henry Ford, but not adopted nationwide until the Fair Standards Labor Act was instated in 1940. Currently, the U.S. labor force mostly works 40 hours per week. But average hours worked has gone relatively unchanged over the past 45 years, in contrast to the steady decline in working hours throughout earlier U.S. history.

Changing the structure of coordinated leisure time may present opportunities to

*This chapter is coauthored with Solomon Hsiang.

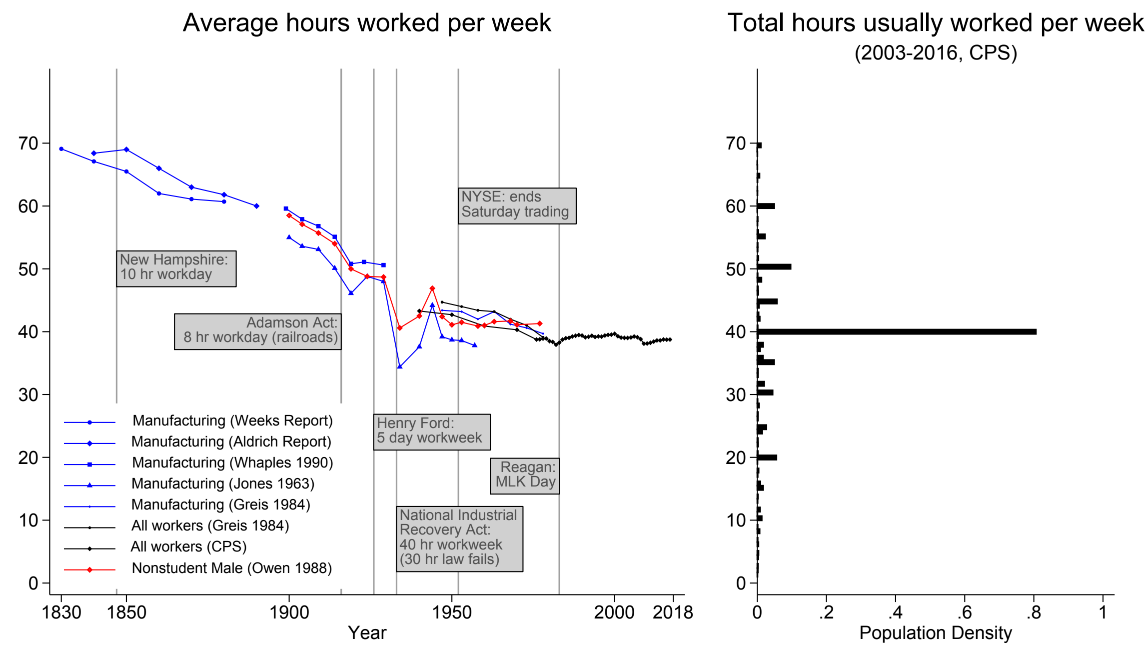


Figure 4.1: History of average hours worked per week in the U.S.

address pollution externalities in politically feasible ways. Yet, to our knowledge, questions about how we could optimize our labor and leisure coordinating mechanisms to reach a desired economic outcome or increase efficiency have gone mostly unasked. One exception is daylight savings time (DST), which was created around 1900, and has been deployed with a stated goal of conserving energy. In the U.S., the Energy Policy Act of 2005 expanded DST four additional weeks beginning in 2007. Evidence about the effect of DST on energy consumption is mixed, with large variation in savings estimates. Estimates of savings are generally small and rely on engineering models (Kandel, Metz, and Commission, 2001; Kandel, Sheridan, and Commission, 2007). Two economic studies call these findings into question. Kotchen and Grant (2011) present evidence that DST actually increased residential electricity consumption in Indiana. Kellogg and Wolff (2008) use a quasi-experiment in Australia to show no effect of DST on electricity use.

In this paper, we present a simple labor model in which there are two types of leisure, "coordinated" and "uncoordinated", where coordinated leisure is taken with members of one's social network, and requires a coordinating mechanism. When each activity (labor, coordinated leisure, and uncoordinated leisure) has a different emissions intensity, policies that encourage the lower-emissions choice or discourage the higher emissions choice could be used to reduce emissions. In addition, even if we can adopt a first-best policy (a carbon tax or cap and trade system), if agents are constrained in their ability to substitute toward the lower-emissions activity, then the policy can be made more efficient by loosening this constraint.

We use a set of natural experiments in leisure coordination to understand the effect of coordinated leisure on U.S. emissions. The U.S. government sets 10 federal holidays, which serve as day that the government is recommending citizens choose coordinated leisure. Weekends also present a coordinating mechanism for leisure. Individuals, along with employers, choose whether to forgo work and choose leisure on each weekend, holiday, and surrounding days. We empirically estimate the effect of these holidays on electricity, vehicle travel, and air travel. We focus on the electricity and transportation sectors because they each accounted for 28% of greenhouse gas (GHG) emissions in the U.S. in 2016 (EPA, 2018), representing the two largest GHG emitting sectors. We develop a novel method to estimate the counterfactual usage in each energy sector, flexibly controlling for season, temperature, and energy displacement before and after holidays. Our results imply that leisure days results in large energy savings on many holidays, particularly during the summer. We use the American Time Use Survey (ATUS) to show that the reductions are explained by individuals substituting sleep and leisure for work time.

4.2 Theory

We motivate our question with a model in which a representative agent chooses to allocate each day of the year to one of three activities: labor, uncoordinated leisure, and coordinated leisure. Coordinated leisure is leisure taken at the same time as members of the agent's social network. Let L, l_u and l_c be the number of days of labor, uncoordinated leisure, and coordinated leisure, respectively. We can write $L + l_u + l_c = 365$.

On days the agent chooses to work, she receives w with which she can buy a composite good c, where the price of c is normalized to 1. In the case where no constraints are imposed on how many days of each activity an agent can choose, given an additively separable utility function $u(c, l_c, l_u) = u(wL, l_c, l_u) = u(w(365 - l_c - l_u), l_c, l_u)$, the agent solves:

$$\begin{aligned}
&\max_{l_c, l_c} u(w(365 - l_c - l_u), l_c, l_u)\\
F.O.C.s : \; & wu_c(\cdot) = u_{l_c}(\cdot)\\
& wu_c(\cdot) = u_{l_u}(\cdot)\\
\implies & wu_c(\cdot) = u_{l_u}(\cdot) = u_{l_c}(\cdot)
\end{aligned}$$

The agent equates the marginal utility from each type of leisure with the marginal utility of working more in order to consume more c. But this model assumes that individuals can easily substitute between labor and both types of leisure. In reality, individuals may face constraints. While many jobs offer a fixed number of holidays and vacation days, which could make reaching the optimal impossible, optimizing individuals may be able to seek out jobs with more or less flexibility, or take multiple jobs which allow the agent to work on additional days. It may be less feasible to take coordinate leisure days beyond days that have been set explicitly as non-workdays because coordinating a network could be prohibitively costly. Hence, coordinated leisure days outside those encouraged by government policies and existing social norms may be infeasible for most individuals. Importantly, we do not believe coordinated leisure occurs only on holidays and weekends. Rather, these serve as coordinating mechanisms, substantially lowering the costs coordinating leisure. It is likely that individuals also coordinate on days before and after the government's announced coordinated leisure days, which we will be discussed later in the paper.

We can explore the implications of a binding constraint on coordinated leisure: $l_c \leq \bar{l}_c$. Although we won't provide quantitative evidence of this constraint here, we believe that for at least some agents, this is a reasonable assumption. One rationale is that we anecdotally observe most individuals taking all available coordinated leisure days. While this might be due to frictions in job markets, it seems more likely that, when offered an additional coordinated leisure day, many individuals would choose to take it, even at some implied wage cost. Whether this is empirically founded should be explored in future work.

For constrained individuals, $l_c \leq \bar{l}_c$ implies that at the optimal $wu_c(\cdot) = u_{l_u}(\cdot) < u_{l_c}(\cdot)$. The agent would like to consume more l_c, but doesn't have sufficient opportunities to do so. Relative to the optimal, she works more and spends more uncoordinated leisure time, equating the marginal utility of the two unconstrained activities. This equilibrium is shown in Figure 4.2.

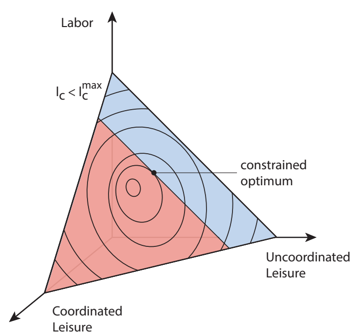


Figure 4.2: Optimal time allocation with constrained coordinated leisure.

Relaxing the constraint on l_c by increasing $\bar{l}_c$ would allow substitution into coordinated leisure and reductions in both uncoordinated leisure and labor. The agent increases each activity based on the the ratio of marginal utility, decreasing L and l_u in a proportion amount to maintain $wu_c(\cdot) = u_{l_u}(\cdot)$. Increasing $\bar{l}_c$ will have this effect until enough $\bar{l}_c$ is high enough for the agent to reach $wu_c(\cdot) = u_{l_u}(\cdot) = u_{l_c}(\cdot)$. Relaxing the constraint will decrease income but is strictly welfare improving for the agent, assuming the individual only earns wages and not a portion of profits or returns to capital.

Now suppose there are emissions intensities of labor, uncoordinated leisure, and coordinated leisure, e_L, e_{l_u}, and e_{l_c}, respectively. Then an individual's emissions are:

$$E = e_L * L + e_{l_c} * l_c + e_{l_u} * l_u$$
$$\iff E = 365 * e_L - [e_L - e_{l_c}] * l_c - [e_L - e_{l_u}] * l_u$$

Total emissions are baseline emissions if the individual chose only labor, $365 * e_L$, minus the choice of each type of leisure times the net emissions intensity of each type of leisure relative to the emissions intensity of labor. Suppose a damage function of emissions, $d(E)$, with $d'(E) > 0$. Then a social planner maximizing utility will take into account the damages from emissions and solve:

$$\begin{aligned} &\max_{l_c, l_c} u(w(365 - l_c - l_u), l_c, l_u) - d(E) \\ F.O.C.s: &\, wu_c(\cdot) = u_{l_c}(\cdot) - d'(E)[e_{l_c} - e_L] \\ &\, wu_c(\cdot) = u_{l_u}(\cdot) - d'(E)[e_{l_u} - e_L] \\ \implies &\, wu_c(\cdot) - d'(E)e_L = u_{l_u}(\cdot) - d'(E)e_{l_u} = u_{l_c}(\cdot) - d'(E)e_{l_c} \end{aligned}$$

The externality is internalized by subtracting the marginal damage from each activity. Therefore, the social planner equates the marginal utility of each activity with the associated damage internalized. The pigovian tax is $\tau = d'(E)$, which will correct the externality in this unconstrained world. To reach the socially optimal allocation, agents reoptimize based on the relative value of e_L, e_{l_u} and e_{l_c}, shifting toward the less carbon intensive activity. Alternatively, a government may be able to use policy instruments that encourage or discourage each activity (for example, by providing coordinating mechanisms) in order to reach the same allocation.

Returning to the constrained world in which the individual is not internalizing the emissions externality (where the private equilibrium is $wu_c(\cdot) = u_{l_u}(\cdot) < u_{l_c}(\cdot)$), the inefficiency caused by this constraint could ameliorate or exacerbate the externality depending on the values of e_L, e_{l_u}, and e_{l_c}. To see this, we explore the effect of a small change in coordinated leisure dl_c. We can quantify the effect on emissions of this change:

$$\begin{aligned} dE &= e_L * \frac{\partial L}{\partial l_c} dl_c + e_{l_u} * \frac{\partial l_u}{\partial l_c} dl_c + e_{l_c} dl_c \\ \iff dE &= (e_L * \frac{\partial L}{\partial l_c} + e_{l_u} * \frac{\partial l_u}{\partial l_c} + e_{l_c}) dl_c \end{aligned}$$

The constraint on total allocation, $L + l_u + l_c = 365$ implies $\frac{\partial L}{\partial l_c} + \frac{\partial l_u}{\partial l_c} + 1 = 0$, which is equivalent to $-(\frac{\partial L}{\partial l_c} + \frac{\partial l_u}{\partial l_c}) = 1$. As discussed above, if we change the amount of coordinated leisure, it changes the choice of both labor and uncoordinated leisure based on the relative marginal utility of each. Therefore, on the margin, we can estimate this by a fixed proportion based on the marginal utilities, and it will be approximately equal to the observed equilibrium proportion. Call the proportional shift to labor α, so $\alpha = -\frac{\partial L}{\partial l_c}$. This implies $\frac{\partial l_u}{\partial l_c} = -(1 - \alpha)$. We can rewrite the above expression:

$$\begin{aligned} dE &= [-\alpha e_L + -(1 - \alpha)e_{l_u} + e_{l_c}] * dl_c \\ \implies dE &= [e_{l_c} - (\alpha e_L + (1 - \alpha)e_{l_u})] * dl_c \end{aligned}$$

Therefore, the change in emissions from a change in coordinated leisure will be the difference between the emissions intensity of coordinated leisure e_{l_c} and the equilibrium weighted-average of the emissions intensity of labor and uncoordinated leisure, $\alpha e_L + (1 - \alpha)e_{l_u}$. If we implement a policy that affects the choice of l_c, we can calculate the associated change in emissions. There are two cases to consider. First, if coordinated leisure is more emission intensive than the counterfactuals, $e_{l_c} > \alpha e_L + (1-\alpha)e_{l_u}$, then the externality brings the constrained individual closer to socially optimal allocation. In contrast, if coordinated leisure is less emissions intensive $e_{l_c} < \alpha e_L + (1-\alpha)e_{l_u}$, then the externality moves the constrained individual farther from the socially optimal allocation.

The primary contribution of this paper is to provide empirical support for the second case: that e_{l_c} is lower than the equilibrium weighted-average of e_L and e_{l_u} in the electricity sector. We demonstrate this by devising a strategy to estimate the effect on electricity consumption of days when we know many agents are choosing coordinated leisure: weekends, holidays, and days before and after holidays.

The implications for a climate policy from the model are as two-fold. First, it may be possible to use policies that encourage coordinate leisure in order to reduce emissions. While we cannot quantify the overall welfare implications because we have not modeled the impact on firms, non-firm owners would be able to increase overall utility in cases where there is a binding constraint on coordinated leisure. Second, a carbon tax, which will impose a cost on firms, may be made more efficient by combining it with leisure policy if there is a binding constraint on coordinated leisure.

4.3 Holidays as a Policy Instrument in the U.S.

We use weekends, federal holidays, and the days around federal holidays to represent exogenous changes in in the number of individuals choosing coordinated leisure. A history of each federal holiday is given in Table 4.1. These holidays represent mandatory leisure for federal employees, although they could seek out other work on these days. However, private firms are not required to provide holidays off for workers (the U.S. is the only advanced economy that does not mandate paid vacation or holidays). Many firms do provide paid holidays for employees, while others provide flexibility if workers wish to take holidays as leisure.

Congress has the power to declare holidays, and has used that power recently: the first five holidays were approved in 1870, and the most recent (Martin Luther King (MLK), Jr. Holiday) was created in 1983. In 1968, Congress approved a law moving four holidays to Mondays, citing "substantial benefits to both the spiritual and economic life of the Nation" (Stathis, 1999). This serves as an example of congress enacting holiday-related leisure policy with the goal of improving welfare.

We code each holiday as a a dummy variable in the data, which takes value 1 on a specific holiday and 0 on every other day of the year. For example, the variable $NewYears$ is 1 on January 1^{st} (New Year's Day) of each year, and 0 otherwise. Days around holidays probably also have low coordinating costs for many agents, therefore we measure the effects of these days too. For example, the day before New Year's Day is New Year's Eve, and while it is not a federal holiday, it is anecdotally more likely to be chosen as coordinated leisure day. Hence, the variable $NewYears\_1DayBefore$ is

Table 4.1: Federal Holidays in the U.S.

Holiday	Description	History
New Year's Day	January 1	Created in 1870
Martin Luther King Jr. Day	3rd Monday in January	Created in 1983
President's Day	The 3rd Monday in February	Created in 1870 as George Washington's Birthday, moved in 1968 to fall on a Monday
Memorial Day	The last Monday in May	Created in 1888 as Decoration Day, moved in 1968 to fall on a Monday
Independence Day	July 4	Created in 1870
Labor Day	The 1st Monday in September	Created in 1894, moved in 1968 to fall on a Monday
Columbus Day	The 2nd Monday in October	Created in 1968
Veteran's Day	November 11	Established as Armistice Day in 1938, moved in 1968 to fall on a Monday, returned to November 11 in 1975
Thanksgiving Day	The 4th Thursday in November	Created in 1870 but date not designated until 1941
Christmas Day	December 25	Created in 1870

1 on December 31^{st} of each year, and 0 otherwise. We create this variable for 8 days before and 8 days after each holiday, and test for robustness in the appendix.

Four holidays fall on the same date every year (Christmas, New Year's, Independence, and Veterans Day), while the other 6 holidays fall on a specific day of the week (e.g. Thanksgiving Day is always the 4^{th} Thursday in November). When a date-specific holiday falls on a Saturday or Sunday, the federal holidays is observed on Friday or Monday, respectively. For example, in 2010, December 25^{th} fell on a Saturday. Therefore, the federal holiday is moved to Friday, December 24^{th}. For these four holidays, we include a dummy variable both for the holiday date and the observed date. So the variable $Christmas$ has value 1 on December 25^{th} in all years, while the variable $Christmas\_Observed$ has value 1 on Decemeber 25^{th} for years 2006 through 2009 and on December 24^{th}, 2010. We find the total effect of a given holiday by taking the sum of the two coefficients.

4.4 Materials and Methods

We aggregate electricity load data from form 714 Federal Energy Regulatory Commision (FERC), ISONE, and ERCOT to create a state-level, hourly electricity dataset from 2006-2014. The data is combined with population weighted minimum and maximum temperature data for each day. Daily air travel data was compiled from the on-time flight records from the Bureau of Transportation Statistics from 1991 - 2014. Vehicle travel data was from highway sensors in California from Caltrans PeMS from 2001-2014.

In our ideal experiment, we would have data on each individual's choice of coordinated leisure, uncoordinated leisure, and labor for each day, and the associated energy use from that individual, on that day. In reality, we only observe certain dates that are treated as holidays. We believe that on these days, many individuals choose coordinated leisure because coordination costs are low. We also observe separate dates that are not treated as holidays. Therefore, the empirical challenge is to estimate counterfactual energy consumption on treated days.

Here we describe the method for estimating the electricity consumption effects in detail, with notes on how the estimations for vehicle travel and air travel differ. Figures from travel estimations, as well as exact specification, are included in the appendix. In order to estimate counterfactual electricity consumption, we control for variation that is unrelated to holiday treatment, including day of week, temperature, and seasonal variation. In addition, we control for location using state-by-year level fixed effects[2]. Figure 4.3 shows the variation in $\log(electricity\_load)$ for a single year. These are means across all states of state-demeaned data, representing the spatial average of $\log(electricity\_load)$ after removing state-level fixed effects.

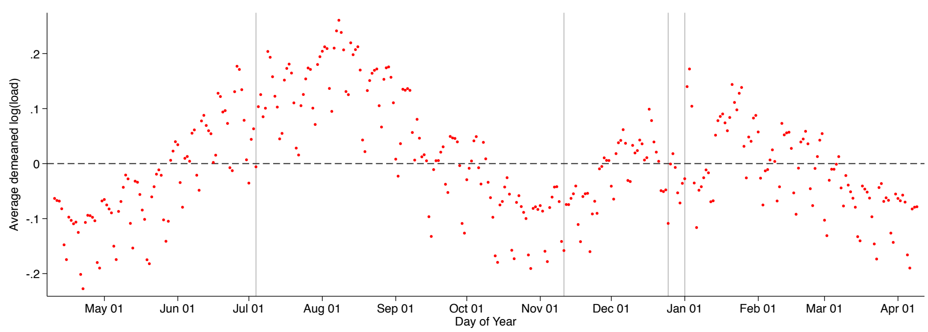
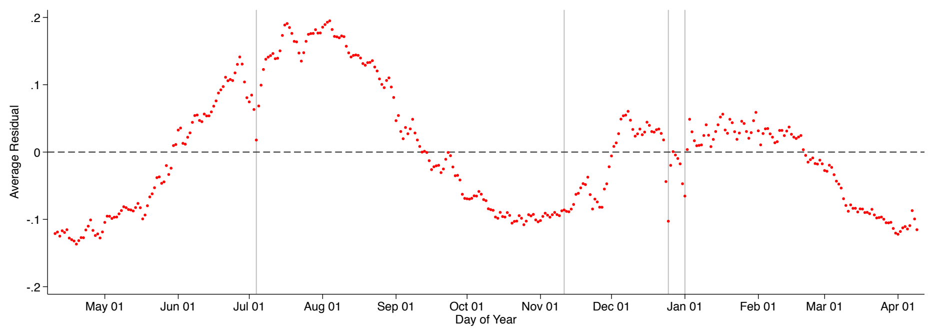


Figure 4.3: Average demeaned $\log(electricity\_load)$ by day of year, single year. Data for each state is demeaned, and the mean across all states is taken for each day. The time period is April 10, 2007 through April 9, 2008. Vertical lines are July 4 (Independence Day), November 11 (Veterans Day), December 25 (Christmas Day), and January 1 (New Year's Day).

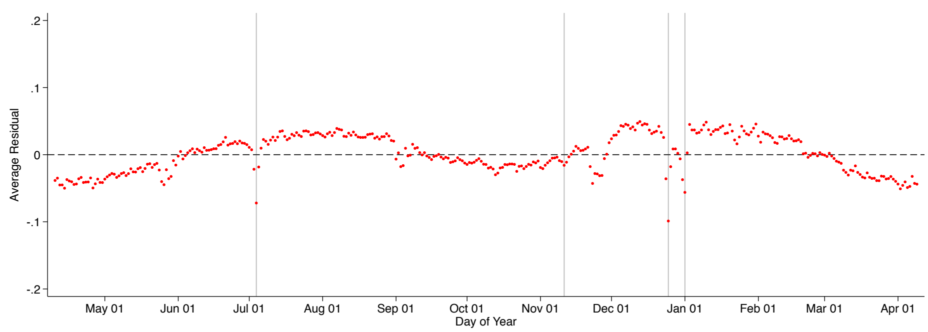
We observe large variation throughout a single year, noticing in particular a day-of-week effect and local maxima in August and January. We can control for day-of-week effects, and use all years of data, over which time individual dates fall on multiple days of the week. The remaining variation after removing day-of-week and year fixed effects is shown in Figure 4.4a.

Removing day-of-week variation and averaging over years makes the seasonal trend clear. In addition, we can distinctly see that there is lower consumption on holidays, particularly those that fall on a single day. Christmas Day, Independence Day, and New Year's Day all appear to be large diversions from the overall trend.

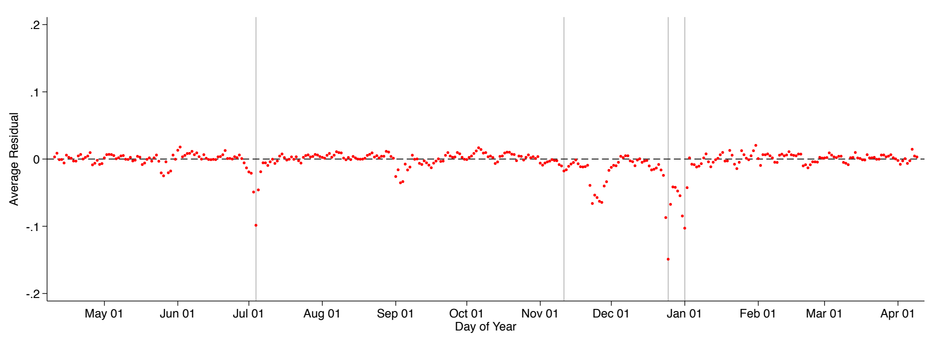
[2]In order to accurately measure the effect of holidays that fall near the year change, we indicate years as starting on April 10. This has no effect on the results, but removes the discontinuity in the model's predicted values to April 10, which does not have any nearby holidays.



(a) Day-of-week effects removed



(b) Day-of-week, temperature effects removed



(c) Day-of-week, temperature, and seasonal-trend effects removed

Figure 4.4: Average residuals by day of year after controlling for covariates. These are mean residuals across all states from a regressions with state-level fixed effects and (a) day-of-week fixed effects, (b) day-of-week fixed effects and 5^{th} degree polynomials of minimum and maximum daily temperature, and (c) day-of-week fixed effects, 5^{th} degree polynomials of minimum and maximum daily temperature, and a 15^{th} degree polynomial of day of year. Vertical lines are July 4 (Independence Day), November 11 (Veterans Day), December 25 (Christmas Day), and January 1 (New Year's Day).

Next we control for minimum and maximum temperature. Average residuals are shown in Figure 4.4b. Controlling for temperature greatly reduces the unexplained variation.

To capture remaining seasonal variation, we fit a 15^{th} order day-of-year polynomial to all "normal" days, which are days for which we are not interested in estimating the impact. That is, we are interested in the effect of holidays and surrounding days, so we exclude these days when estimating the polynomial. Here, and in our discussed specifications, we include 8-days before and after holidays as days of interest, although we show specifications with 2, 5, and 10 day windows as well. Figure 4.5 depicts how this method relies on days outside of the 8-day window to estimate counterfactual energy consumption.

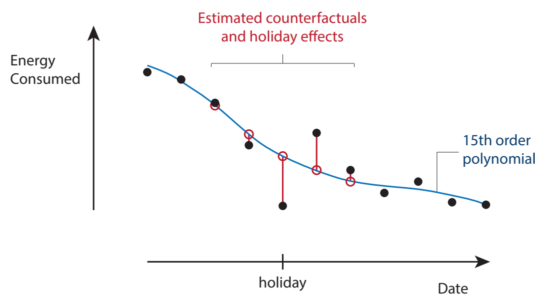


Figure 4.5: Counterfactual energy consumption estimation.
A 15^{th} order day-of-year polynomial is fitted to all days that are not of interest, after controlling for observables. This removes unobservable variation that is unrelated to treatment.

This method is similar to one used to measure counterfactual densities at kink points, developed in Saez (1999) and employed by Saez (2010), Chetty et al. (2011), and Ito and Sallee (2014). The method uses a flexible polynomial specification to estimate a density based on observations that are near a discontinuity, then applies and adjustment factor based on the size of the discontinuity to represent the population shift. Since we are not estimating a density and therefore are unsure as to whether displacement occurs, there is no need to apply the adjustment factor described in these papers. Instead, our method should capture any displacement that occurs within the surrounding days that are not used to estimated the polynomial. This also allows conjecture as to the extent to which individuals choose coordinated leisure on the days surrounding federal holidays. All results are robust to the choice of polynomial order, and specifications with 7^{th}, 9^{th}, 11^{th}, and 13^{th} order day-of-year polynomials are shown in the appendix (tables C.3, C.6, and C.9). The choice does not appear to introduce a systematic bias, and overfitting is not a concern because the polynomial is only fit on data outside the selected window of surrounding days. We select the 15^{th} order polynomial because it effectively removes the remaining seasonal trends after we have controlled for temperature and minimizes mean-squared-error when we perform 5-fold cross validation.

Having removed variation that is unrelated to treatment, we can estimate the impact of a holiday relative to our constructed counterfactual. In fact, the estimate for holidays which do not change dates is the unexplained residual shown in the Figure 4.4c. Our identifying assumption is that, after controlling for these covariates and fixed effects, holidays are as good as randomly assigned. Our methodology is simply implemented for all holidays with a single regression specification:

$$\begin{aligned}\log(electricity\_load_{ity}) = \alpha + \sum_{h \in H} \beta_h * DayofInterestDummy_h \\ + \mu_i + \delta_y + \sum_{d=1}^{6} \psi_d * DayofWeekDummy_d \\ + f(t) + g_1(MaxTemp_{ity}) + g_2(MinTemp_{ity}) + \varepsilon_{ity} \end{aligned} \quad (4.1)$$

where H is the set of holidays and surrounding days, $i \in I$ is the set of states, $t \in T$ is the set of days of the year $1, 2, ..., 365$, $y \in Y$ is the set of years 2006 to 2014, $f(\cdot)$ is a 15^{th} order polynomial, and $g_1(\cdot)$ and $g_2(\cdot)$ are 5^{th} order polynomials.

The equation above applies to the daily regression for estimating the effect of holidays on electricity loads. A similar estimating equation is used to measure the effect of holidays on vehicle travel and air travel. The primary difference is that weather is not included in the travel specifications because demand depends less on temperature and weather, and the flexible polynomial captures seasonal variation sufficiently. The exact equations as well as the decomposition for vehicle travel and air travel are available in the appendix.

For electricity, we are identifying the within-state effect of a holiday versus a non-holiday, controlling for day-of-week, temperature, and seasonal trends, assuming $E[\varepsilon_{ity}|X_{ity}] = 0$, where X_{ity} are all covariates included above. For vehicle travel, the identifying assumption is the same except there is no control for temperature and the identification is within highway sensor. For air travel, the idenfitifcation is within airport.

We believe this method is superior to a more traditional implementation that uses month-by-year fixed effects for several reasons. First, we see clear within-month trends in Figure 4.3, which would not be easy to capture. Second, our method makes it simpler to estimate the impact of holidays when the surrounding days occur in a different month. For example, when estimating Christmas, New Year's, and the surrounding days, we are producing estimates in both December and January. Month-by-year fixed effects would use a separate counterfactual for estimates in different months. We could solve this by moving our month designation (as we have done for the year designation); however our method is superior because it includes this benefit and allows the counterfactual to include trends in the surrounding days rather than a simple average.

For hourly versions of the regressions, we run the same specification 24 times: once for each hour of the day. This allows each covariate and effect to be heterogeneous by hour-of-day. We do the same to measure the effect of holidays on the proportion of the U.S. population engaged in each time use activity.

4.5 Results

Electricity use, vehicle travel, and air travel, in general, drop significantly on holidays and weekend. Figure 4.6 shows the effect of weekends and each federal holiday on consumption in the electricity and transportation sectors. Energy consumption falls significantly on Saturday and Sundays relative to weekdays, which tend to have similar usage across each day. Christmas, New Year's, Memorial, Independence, and Thanksgiving Days all result in large reductions in energy consumption on the holidays. However, certain holidays exhibit spillovers to surrounding days, particularly in terms of travel. Christmas, New Year's, and Thanksgiving result in increased air and vehicle travel before and after the holidays, counteracting some of the overall effect. In contrast, the summer holidays result in no spillovers, resulting in large net reductions in electricity and travel emissions. In the sections that follow, I explore these results and the mechanisms that cause them in detail.

Individual Holiday Effects

Table 4.2 shows the effect of each holiday on electricity load, air miles, and vehicle flow. We find significant reductions in electricity use, air travel, and vehicle travel on Christmas, New Year's, Memorial, Independence, Labor, and Thanksgiving Days. These findings are robust to different specifications, and all robustness tables are included in the appendix tables C.1 - C.9. Christmas effects are the largest, with reductions of 15, 14, and 24 log points for electricity, air miles, and vehicle flow, respectively. The magnitudes on Thanksgiving and independence day are similar.

Day of Week Effects

Table 4.3 shows the average day-of-week effects on electricity consumption, air travel, and vehicle travel. Day-of-week effects are relative to the omitted day (Wednesday), and Monday through Thursday are are similar to each other, though not identical. Friday exhibits a small but significant reduction in electricity use, but an increase in air and vehicle travel. This is likely due to increased travel demand from weekend trips. Saturdays and Sundays both exhibit large reductions in all three sectors. Sunday's reduction of about 9 log points is significantly larger than Saturday reduction of about log point for electricity. Air Travel reduces by over 10 log points on Saturday, while vehicle travel falls most on Sunday - it drops by over 21 log points and remains lower on Monday.

Total Effects and Emissions

Table 4.5 reports the sum of effects from each holiday and 8-days before and after. The cumulative effects are large, suggesting that Christmas and New Year's combined result in load reductions equivalent to 56% of a counterfactual day's total electricity

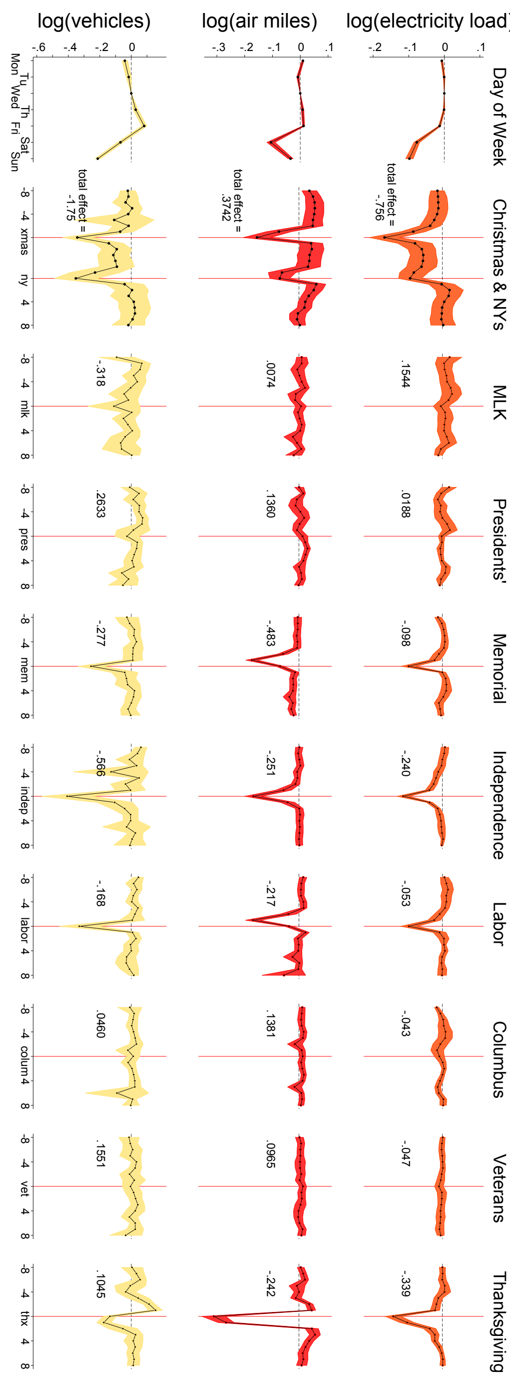


Figure 4.6: Daily changes in energy use on weekends and holidays the U.S. Estimates are from regression equation 4.1 for electricity, and corresponding equations C.1 and C.2 for air and vehicle travel, respectively, where the primary difference is the sample and the use of temperature controls.

Table 4.2: Individual Federal Holiday Energy Effects

	(1) log(electricity)	(2) log(air miles)	(3) log(vehicles)
Christmas	-0.1542***	-0.1439***	-0.3428***
	(0.0200)	(0.0237)	(0.0468)
New Year's	-0.0860***	-0.0656***	-0.3522***
	(0.0153)	(0.0172)	(0.0686)
Martin Luther King, Jr.	-0.00513	0.00739	-0.110
	(0.00991)	(0.00945)	(0.0846)
Presidents	-0.00329	0.0126	-0.0270
	(0.00439)	(0.00910)	(0.0393)
Memorial	-0.0915***	-0.0651***	-0.256***
	(0.01000)	(0.00695)	(0.0419)
Independence	-0.1044***	-0.1565***	-0.4072***
	(0.0079)	(0.0146)	(0.0817)
Labor	-0.0899***	-0.0339***	-0.330***
	(0.0106)	(0.00720)	(0.0639)
Columbus	-0.00784	0.0130**	0.0128
	(0.00655)	(0.00508)	(0.0206)
Veterans	-0.00909	0.0140	-0.00512
	(0.00580)	(0.00617)	(0.0264)
Thanksgiving	-0.134***	-0.292***	-0.136***
	(0.0106)	(0.0205)	(0.0214)
Observations	140,019	2,541,130	19,801,068
R-squared	0.996	0.986	0.885
Year-by-Location FE	X	X	X
Min. / Max. Temp. Polynomial	X		
Day-of-Year Polynomial	X	X	X
8 Days Before / After Holidays	X	X	X

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Estimates are from regression equation 4.1 for electricity, and corresponding equations C.1 and C.2 for air and vehicle travel, respectively, where the primary difference is the sample and the use of temperature controls. All specifications include include 8-day dummy variables before and after each holiday to control for spillovers. Other controls include year-by-location fixed effects and a $15^{th}-order$ polynomial in day-of-year

load. The Thanksgiving holiday and Independence holiday are 41% and 28% savings, respectively.

In order to compare holidays totat energy savings, we convert each effect their emissions magnitudes in 2014 by attributing the emissions throughout the year and measuring the difference from our counterfactual energy use versus the actual energy use on holiday.

Table 4.3: Day of Week Energy Effects

	(1) log(electricity)	(2) log(air miles)	(3) log(vehicles)
Monday	-0.00541***	0.00953***	-0.0409***
	(0.00125)	(0.00182)	(0.00936)
Tuesday	-0.000139	-0.00838***	-0.0163**
	(0.00118)	(0.00211)	(0.00827)
Wednesday	*omitted*	*omitted*	*omitted*
Thursday	-0.000750	0.00835***	0.0281***
	(0.00112)	(0.00224)	(0.00789)
Friday	-0.0139***	0.0120***	0.0822***
	(0.00143)	(0.00183)	(0.00869)
Saturday	-0.0794***	-0.105***	-0.0687***
	(0.00412)	(0.00735)	(0.00940)
Sunday	-0.0993***	-0.0348***	-0.213***
	(0.00526)	(0.00458)	(0.00896)
Observations	140,019	2,541,130	19,801,068
R-squared	0.996	0.986	0.885
Year-by-Location FE	X	X	X
Min. / Max. Temp. Polynomial	X		
Day-of-Year Polynomial	X	X	X
8 Days Before / After Holidays	X	X	X

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Estimates are from regression equation 4.1 for electricity, and corresponding equations C.1 and C.2 for air and vehicle travel, respectively, where the primary difference is the sample and the use of temperature controls. All specifications include include 8-day dummy variables before and after each holiday to control for spillovers. Other controls include year-by-location fixed effects and a $15^{th} - order$ polynomial in day-of-year

Hourly Time Use

To understand the heterogeneity in energy savings across leisure days, we estimate changes in hourly electricity consumption on holidays and compare them with changes in activity choice. Reductions in electricity consumption are consistent with reductions in work time and increases in sleep and leisure time. Figure 4.7 shows the estimated changes in electricity consumption by hour on each federal holiday in red. In blue, from top to bottom are deviations in sleep, leisure, and work. Generally, the population shifts work time to sleep and leisure time on holidays, and these changes align closely with reductions in electricity usage. Linear regression of time use estimates on electricity usage for the set of holidays confirms that these three time use variables explain over 75% of the reductions in hourly electricity usage.

Table 4.4: Summed Effects of Holidays, Observed Holidays, and Surrounding Days

Holiday	Electricity	95% CI	Air Travel	95% CI	Vehicle Travel	95% CI
Christmas & New Year's	-0.765	[-1.08,-0.45]	0.374	[-0.095,0.843]	-1.753	[-2.75,-0.756]
Martin Luther King, Jr.	0.152	[-0.027,0.331]	0.007	[-0.171,0.186]	-0.318	[-1.035,0.399]
Presidents'	0.017	[-0.07,0.103]	0.136	[0.027,0.245]	0.263	[-0.166,0.693]
Memorial	-0.096	[-0.18,-0.013]	-0.483	[-0.602,-0.365]	-0.277	[-0.766,0.211]
Independence	-0.197	[-0.288,-0.105]	-0.234	[-0.328,-0.14]	-0.369	[-0.95,0.212]
Labor	-0.055	[-0.15,0.04]	-0.217	[-0.368,-0.066]	-0.168	[-0.544,0.208]
Columbus	-0.039	[-0.107,0.029]	0.138	[0.043,0.233]	0.046	[-0.324,0.416]
Veterans	-0.043	[-0.146,0.06]	0.093	[-0.04,0.227]	0.055	[-0.454,0.564]
Thanksgiving	-0.342	[-0.458,-0.226]	-0.243	[-0.397,-0.089]	0.105	[-0.379,0.588]

This table reports the sums of the estimated savings for each holiday and the 8-days before and after the holiday in log points, and the associated 95% confidence interval. Estimates are from regression equation 4.1 for electricity, and corresponding equations C.1 and C.2 for air and vehicle travel, respectively, where the primary difference is the sample and the use of temperature controls.

Table 4.5: Total Emissions Change from Summed Effects

	Emissions (Million Metric Tons CO_2 Equivalent)			
Holiday	Electricity	Air Travel	Vehicles	Total
Christmas & New Year's	-3.674	0.258	-5.674	-9.089
Martin Luther King, Jr.	0.786	0.006	-1.475	-0.683
Presidents'	0.118	0.056	1.227	1.401
Memorial	-0.668	-0.194	-0.984	-1.846
Independence	-0.932	-0.041	-0.863	-1.836
Labor	-0.511	-0.074	-0.820	-1.405
Columbus	-0.174	0.060	0.157	0.043
Veterans	-0.181	0.034	0.756	0.608
Thanksgiving	-1.914	-0.057	0.582	-1.389

This table reports the sums of the estimated savings for each holiday and the 8-days before and after the holiday in log points, and the associated 95% confidence interval. Estimates are from regression equation 4.1 for electricity, and corresponding equations C.1 and C.2 for air and vehicle travel, respectively, where the primary difference is the sample and the use of temperature controls.

Temperature Responses

To understand how temperature affects electricity load, and whether there is a different response to temperature on leisure days versus workdays, we present temperature response functions for each days type. We generate different temperature response functions for weekends versus weekdays and holidays versus non-holidays by estimating the the impact of temperature on electricity loads using a 5^{th} order maximum temperature[3]

[3]We perform the same exercise for minimum temperature which yields similar results.

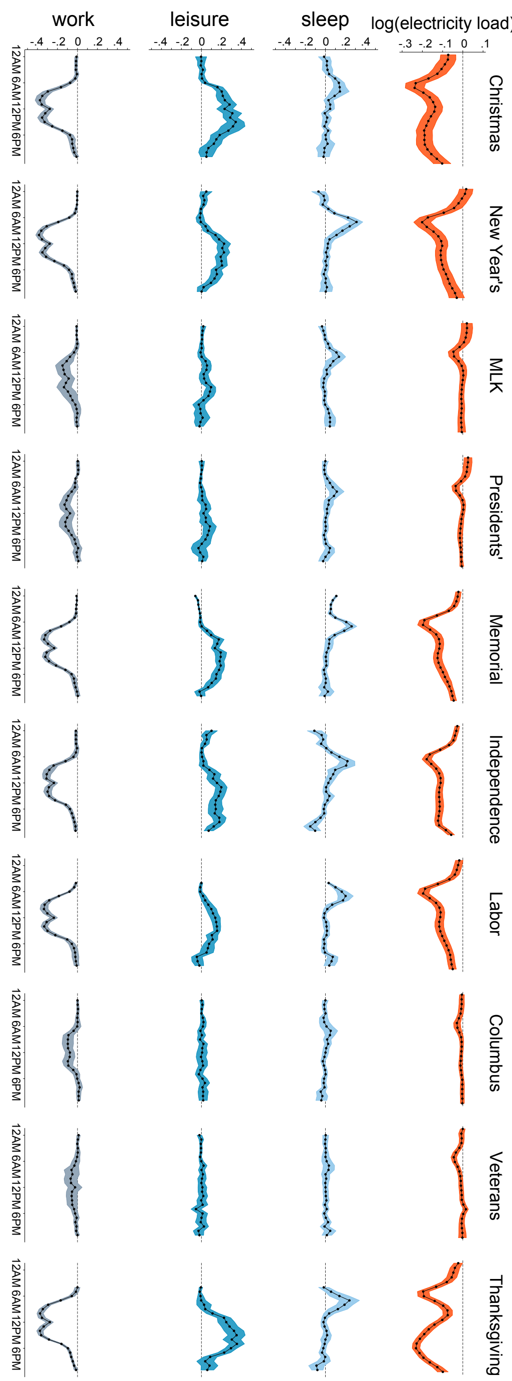


Figure 4.7: Hourly changes in electricity load and time use on holidays the U.S. Estimates are from hourly regressions (separate estimation for each hour) using equation 4.1 for electricity, and corresponding equations C.3 for time use estimation, respectively, where the primary difference is the sample.

polynomial. The results are shown in Figure 4.8.

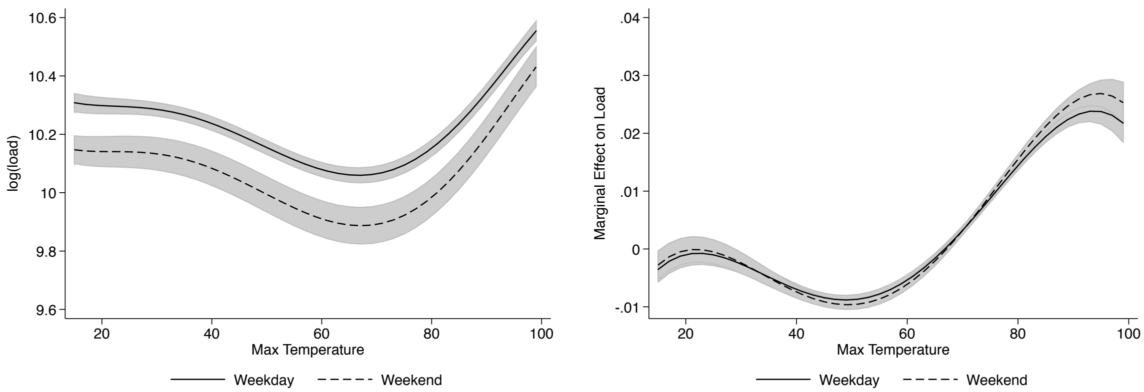


Figure 4.8: Electricity Temperature Response Functions on Weekends and Weekdays. These are from a regressions where $g(\cdot)$ is allowed to vary on weekends versus weekdays.

While we observe a large difference in levels between day-types, differences in the marginal effect of temperature on load are small and mostly insignificant. It is informative, however, to compare the magnitude of these results to our holiday savings estimates. Very hot days increase load by up to about 3%, which is less than half the average treatment effect of a federal holiday, and small compared to the savings we see on major holidays, which are presented below.

Observed Holidays

For holidays which can be observed on an alternative date, we sum the effects of the holiday and the observed holiday in table 4.2, because the total effects are the savings that occur on a holiday in a normal year when the federal holiday is observed on the traditional holiday date. However, examining the differences in the estimated coefficient for a holiday versus an observed holidays, shown in table 4.6, supports the causal mechanism discussed above. Although the variation is not completely randomly because holidays and observed holidays only fall on separate days when the holiday occurs on a weekend, it is likely that the variation is as good as random once we control for day-of-week effects. On New Year's and Independence Day, we see negative and significant coefficients on both terms. In each case, the "Holiday" dummy variable is larger in magnitude than the "Observed Holiday" dummy variable, even though the observed holiday is the day many people get to take off work. This provides suggestive evidence that the effect we see on holidays is not just the effect of individuals substituting between labor and leisure. There is an additional effect that may be due to the activities that people choose on holidays. Some examples that would save electricity are outdoor recreation and cooking collectively rather than individually.

Table 4.6: Individual Federal Holiday Energy Effects

	(1) log(electricity)	(2) log(air miles)	(3) log(vehicles)
Christmas	-0.0985***	-0.103***	-0.208***
	(0.0203)	(0.0237)	(0.0525)
Observed Christmas	-0.0584***	-0.0406***	-0.135***
	(0.0161)	(0.0138)	(0.0305)
New Year's	-0.0431***	-0.0229	-0.145**
	(0.0158)	(0.0220)	(0.0684)
Observed New Year's	-0.0450***	-0.0427**	-0.207***
	(0.0130)	(0.0172)	(0.0501)
Independence	-0.0599***	-0.140***	-0.210***
	(0.00510)	(0.0126)	(0.0623)
Observed Independence	-0.0470***	-0.0170	-0.198***
	(0.00321)	(0.0122)	(0.0657)
Veterans	-0.00592	0.0109	-0.105
	(0.00565)	(0.0104)	(0.0885)
Observed Veterans	-0.00322	0.00315	0.100
	(0.00433)	(0.00927)	(0.0734)

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

We estimate a separate effect of the holiday and the observed federal holiday. The "Holiday" dummy variable is 1 on the holiday date, while the "Observed Holiday" variable is 1 on the federally observed holiday date. Estimates are from regression equation 4.1 for electricity, and corresponding equations C.1 and C.2 for air and vehicle travel, respectively, where the primary difference is the sample and the use of temperature controls. All specifications include include 8-day dummy variables before and after each holiday to control for spillovers. Other controls include year-by-location fixed effects and a $15^{th} - order$ polynomial in day-of-year

4.6 Conclusions

Our results suggest that coordinated leisure could provide a useful policy tool for reducing greenhouse gas emissions. We observe significant reductions in electricity loads and travel on weekends, many holidays, and days surrounding holidays. We also provide evidence that holidays which create three-day weekends result in large load reductions during the summer, but smaller electricity savings in other seasons. Evidence from time use surveys and exogenous variation in the day on which a holiday is observed suggests that individuals choose less electricity intensive activities on holidays, and the result is not entirely driven by differences in workplace electricity consumption and home electricity consumption.

Another way to think about our electricity results is to compare them to other electricity savings mechanisms. Days with high electricity demand are problematic for system operators because capacity is limited by existing infrastructure. Therefore, utilities have a variety of tools to reduce demand when they face the possibility of demand that is above generation capacity. When this occurs, reductions in demand are valuable because the marginal cost of electricity is extremely high. Critical peak pricing (CPP) programs charge higher rates for customers during these “critical” periods. Many studies have estimated the savings from CPP programs. Ito, Ida, and Tanaka (2015) find that CPP programs reduce electricity demand 14-17%, depending on the price that consumers face. They compare this effect to moral suasion, in which consumers are asked to save electricity but do not face a higher price. Moral suasion achieves 3.1% savings. Therefore, the reduction in load that we observe on large holidays are similar to their CPP estimates, and planning additional coordinated leisure in times when loads are likely to be high presents an economically valuable opportunity. While it would not be possible to perfectly plan a holiday to coincide with a critical peak event, more concretely identifying the mechanisms causing the savings identified here might lead to new ideas for behavioral conservation.

We find emissions reductions of over 14 million metric tons of CO_2 from the ten federal holidays and surrounding days. This is equivalent to over 5 gallons of gasoline or 106 vehicle miles travelled per person in the U.S.[4] It is also equivalent to over 350,000 additional flights[5].

[4] Conversion rates from the EPA, http://www.epa.gov/cleanenergy/energy-resources/refs.html: 0.008887 metric tons CO_2/gallon of gasoline or 0.00042 metric tons CO_2/mile.

[5] Based on 0.000271 metric tons CO_2/passenger-mile, with an average of 150 passengers per flight and 1000 miles per flight

Bibliography

Ahlfeldt, Gabriel M., Stephen J. Redding, Daniel M. Sturm, and Nikolaus Wolf. 2015. “The Economics of Density: Evidence From the Berlin Wall.” *Econometrica* 83 (6):2127–2189. URL http://onlinelibrary.wiley.com/doi/10.3982/ECTA10876/abstract.

Allcott, Hunt and Michael Greenstone. 2012. “Is There an Energy Efficiency Gap?” *The Journal of Economic Perspectives* 26 (1):3–28. URL http://www.jstor.org/stable/41348804.

Alonso, William. 1964. “Location and land use. Toward a general theory of land rent.” *Location and land use. Toward a general theory of land rent.* .

Anas, Alex and Leon N. Moses. 1979. “Mode choice, transport structure and urban land use.” *Journal of Urban Economics* 6 (2):228–246. URL http://www.sciencedirect.com/science/article/pii/009411907990007X.

Anderson, Michael L. 2014. “Subways, Strikes, and Slowdowns: The Impacts of Public Transit on Traffic Congestion.” *American Economic Review* 104 (9):2763–96. URL http://www.aeaweb.org/articles.php?doi=10.1257/aer.104.9.2763.

Angrist, Joshua D., Sydnee Caldwell, and Jonathan V. Hall. 2017. “Uber vs. Taxi: A Drivers Eye View.” Working Paper 23891, National Bureau of Economic Research. URL http://www.nber.org/papers/w23891.

Ashenfelter, Orley and Michael Greenstone. 2004. “Using Mandated Speed Limits to Measure the Value of a Statistical Life.” *Journal of political economy* 112 (S1):S226–S267. URL http://www.safetylit.org/citations/index.php?fuseaction=citations.viewdetails&citationIds[]=citjournalarticle_233842_38.

Baum-Snow, Nathaniel. 2007a. “Did Highways Cause Suburbanization?” *The Quarterly Journal of Economics* 122 (2):775–805. URL http://qje.oxfordjournals.org/content/122/2/775.

———. 2007b. “Suburbanization and transportation in the monocentric model.” *Journal of Urban Economics* 62 (3):405–423. URL http://www.sciencedirect.com/science/article/pii/S0094119006001161.

Baum-Snow, Nathaniel, Loren Brandt, J Vernon Henderson, Matthew A Turner, and Qinghua Zhang. 2015. "Roads, railroads and decentralization of Chinese cities." Working Paper.

Bento, Antonio M., Maureen L. Cropper, Ahmed Mushfiq Mobarak, and Katja Vinha. 2005. "The Effects of Urban Spatial Structure on Travel Demand in the United States." *Review of Economics and Statistics* 87 (3):466–478. URL http://dx.doi.org/10.1162/0034653054638292.

Burchfield, Marcy, Henry G. Overman, Diego Puga, and Matthew A. Turner. 2006. "Causes of Sprawl: A Portrait from Space." *The Quarterly Journal of Economics* 121 (2):587–633. URL http://qje.oxfordjournals.org/content/121/2/587.

Chetty, Raj, John N. Friedman, Tore Olsen, and Luigi Pistaferri. 2011. "Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records." *The Quarterly Journal of Economics* 126 (2):749–804. URL http://qje.oxfordjournals.org/content/126/2/749.

Cohen, Peter, Robert Hahn, Jonathan Hall, Steven Levitt, and Robert Metcalfe. 2016. "Using Big Data to Estimate Consumer Surplus: The Case of Uber." Working Paper 22627, National Bureau of Economic Research. URL http://www.nber.org/papers/w22627.

Couture, Victor, Gilles Duranton, and Matthew A. Turner. 2016. "Speed." Working Paper.

Donaldson, Dave. 2010. "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure." Working Paper 16487, National Bureau of Economic Research. URL http://www.nber.org/papers/w16487.

Donaldson, Dave and Richard Hornbeck. 2016. "Railroads and American Economic Growth: A Market Access Approach." *The Quarterly Journal of Economics* 131 (2):799–858. URL http://qje.oxfordjournals.org/content/131/2/799.

Duranton, Gilles and Diego Puga. 2013. "The Growth of Cities." SSRN Scholarly Paper ID 2309234, Social Science Research Network, Rochester, NY. URL http://papers.ssrn.com/abstract=2309234.

———. 2015. "Chapter 8 - Urban Land Use." In *Handbook of Regional and Urban Economics*, *Handbook of Regional and Urban Economics*, vol. 5, edited by J. Vernon Henderson and William C. Strange Gilles Duranton. Elsevier, 467–560. URL http://www.sciencedirect.com/science/article/pii/B9780444595171000088.

Duranton, Gilles and Matthew A. Turner. 2012. "Urban Growth and Transportation." *The Review of Economic Studies* 79 (4):1407–1440. URL http://restud.oxfordjournals.org/content/79/4/1407.

EPA, US. 2018. "Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990-2016." Reports and Assessments. URL https://www.epa.gov/ghgemissions/inventory-us-greenhouse-gas-emissions-and-sinks-1990-2016.

Farmer, Charles M, Richard A Retting, and Adrian K Lund. 1999. "Changes in motor vehicle occupant fatalities after repeal of the national maximum speed limit." *Accident Analysis & Prevention* 31 (5):537–543. URL http://www.sciencedirect.com/science/article/pii/S000145759900010X.

Fowlie, Meredith, Catherine Wolfram, C. Anna Spurlock, Annika Todd, Patrick Baylis, and Peter Cappers. 2017. "Default Effects and Follow-On Behavior: Evidence from an Electricity Pricing Program." Working Paper 23553, National Bureau of Economic Research. URL http://www.nber.org/papers/w23553.

Friedman, Lee S., Donald Hedeker, and Elihu D. Richter. 2009. "Long-Term Effects of Repealing the National Maximum Speed Limit in the United States." *American Journal of Public Health* 99 (9):1626–1631. URL http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2724439/.

Garcia-Lopez, Miquel-ngel. 2012. "Urban spatial structure, suburbanization and transportation in Barcelona." *Journal of Urban Economics* 72 (23):176–190. URL http://www.sciencedirect.com/science/article/pii/S0094119012000356.

Gibbons, Stephen and Stephen Machin. 2005. "Valuing rail access using transport innovations." *Journal of Urban Economics* 57 (1):148–169. URL http://www.sciencedirect.com/science/article/pii/S0094119004001020.

Gillan, James. 2017. "Dynamic Pricing, Attention, and Automation: Evidence from a Field Experiment in Electricity Consumption." Working Paper.

Glaeser, Edward L. and Matthew E. Kahn. 2004. "Chapter 56 Sprawl and urban growth." Elsevier, 2481–2527. URL http://www.sciencedirect.com/science/article/pii/S1574008004800130.

Glaeser, EdwardL. and Joseph Gyourko. 2005. "Urban Decline and Durable Housing." *Journal of Political Economy* 113 (2):345–375. URL http://www.journals.uchicago.edu/doi/10.1086/427465.

Gonzalez-Navarro, Marco and Climent Quintana-Domeque. 2015. "Paving Streets for the Poor: Experimental Analysis of Infrastructure Effects." *Review of Economics and Statistics* 98 (2):254–267. URL http://dx.doi.org/10.1162/REST_a_00553.

Hall, Jonathan V. and Alan B. Krueger. 2016. "An Analysis of the Labor Market for Ubers Driver-Partners in the United States." Working Paper 22843, National Bureau of Economic Research. URL http://www.nber.org/papers/w22843.

Heblich, Stphan, Stephen J. Redding, and Daniel M. Strum. 2017. "The Making of a Modern Metropolis: Evidence from London." Working Paper.

Imai, Haruo. 1982. "CBD hypothesis and economies of agglomeration." *Journal of Economic Theory* 28 (2):275–299. URL http://www.sciencedirect.com/science/article/pii/002205318290062X.

IPCC. 2014. "Summary for Policymakers." In *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by C. B. Field, V. R. Barros, D. J. Dokken, K. J. Mach, M. D. Mastrandrea, T. E. Bilir, M. Chatterjee, K. L. Ebi, Y. O. Estrada, R. C. Genova, B. Girma, E. S. Kissel, A. N. Levy, S. MacCracken, P. R. Mastrandrea, and L. L. White. Cambridge, United Kingdom, and New York, NY, USA: Cambridge University Press, 1–32.

Ito, Koichiro, Takanori Ida, and Makoto Tanaka. 2015. "The Persistence of Moral Suasion and Economic Incentives: Field Experimental Evidence from Energy Demand." Working Paper 20910, National Bureau of Economic Research. URL http://www.nber.org/papers/w20910.

———. 2018. "Moral Suasion and Economic Incentives: Field Experimental Evidence from Energy Demand." *American Economic Journal: Economic Policy* 10 (1):240–267. URL https://www.aeaweb.org/articles?id=10.1257/pol.20160093.

Ito, Koichiro and James M. Sallee. 2014. "The Economics of Attribute-Based Regulation: Theory and Evidence from Fuel-Economy Standards." Working Paper 20500, National Bureau of Economic Research. URL http://www.nber.org/papers/w20500.

Joskow, Paul L. and Catherine D. Wolfram. 2012. "Dynamic Pricing of Electricity." *American Economic Review* 102 (3):381–385. URL https://www.aeaweb.org/articles?id=10.1257/aer.102.3.381.

Kandel, Adrienne, Daryl Metz, and California Energy Commission. 2001. *Effects of Daylight Saving Time on California Electricity Use: Staff Report*. California Energy Commission.

Kandel, Adrienne, Margaret Sheridan, and California Energy Commission. 2007. *The Effect of Early Daylight Saving Time on California Electricity Consumption: a Statistical Analysis: Staff Report*. California Energy Commission.

Katz, Lawrence F. and Alan B. Krueger. 2016. "The Rise and Nature of Alternative Work Arrangements in the United States, 1995-2015." Working Paper 22667, National Bureau of Economic Research. URL http://www.nber.org/papers/w22667.

Kellogg, Ryan and Hendrik Wolff. 2008. "Daylight time and energy: Evidence from an Australian experiment." *Journal of Environmental Economics and Management* 56 (3):207–220. URL http://www.sciencedirect.com/science/article/pii/S0095069608000661.

Kotchen, Matthew J. and Laura E. Grant. 2011. "Does Daylight Saving Time Save Energy? Evidence from a Natural Experiment in Indiana." *Review of Economics and Statistics* 93 (4):1172–1185. URL http://dx.doi.org/10.1162/REST_a_00131.

Li, Ziru, Yili Hong, and Zhongju Zhang. 2016. "An Empirical Analysis of On-Demand Ride Sharing and Traffic Congestion." SSRN Scholarly Paper ID 2843301, Social Science Research Network, Rochester, NY. URL https://papers.ssrn.com/abstract=2843301.

Mills, Edwin S. 1967. "An Aggregative Model of Resource Allocation in a Metropolitan Area." *The American Economic Review* 57 (2):197–210. URL http://www.jstor.org/stable/1821621.

Muth, Richard F. 1969. "CITIES AND HOUSING; THE SPATIAL PATTERN OF URBAN RESIDENTIAL LAND USE." .

Ogawa, Hideaki and Masahisa Fujita. 1980. "Equilibrium Land Use Patterns in a Nonmonocentric City*." *Journal of Regional Science* 20 (4):455–475. URL http://onlinelibrary.wiley.com/doi/10.1111/j.1467-9787.1980.tb00662.x/abstract.

Parry, Ian W. H., Margaret Walls, and Winston Harrington. 2007. "Automobile Externalities and Policies." *Journal of Economic Literature* 45 (2):373–399. URL http://www.jstor.org/stable/27646797.

Retting, Richard A. and M. A. Greene. 1997. "Traffic Speeds Following Repeal of the National Maximum Speed Limit." *ITE journal* 67 (5):42–46. URL http://www.safetylit.org/citations/index.php?fuseaction=citations.viewdetails&citationIds[]=citjournalarticle_78769_13.

Saez, Emmanuel. 1999. "Do Taxpayers Bunch at Kink Points?" Working Paper 7366, National Bureau of Economic Research. URL http://www.nber.org/papers/w7366.

———. 2010. "Do Taxpayers Bunch at Kink Points?" *American Economic Journal: Economic Policy* 2 (3):180–212. URL http://www.aeaweb.org/articles.php?doi=10.1257/pol.2.3.180.

Stathis, Stephen W. 1999. "Federal Holidays: Evolution and Application." Tech. rep., Congressional Research Service.

Van Benthem, Arthur. 2015. "What is the optimal speed limit on freeways?" *Journal of Public Economics* 124:44–62. URL http://www.sciencedirect.com/science/article/pii/S0047272715000146.

Appendix A

Transportation, Market Access, and Urban Development: Evidence from Uber - Appendix

A.1 Empirical Setting

Table A.1 show Uber launch dates for cities in which I have housing data.

A.2 Sample Robustness

Table A.2 shows sample length robustness for my primary specification.

A.3 Google Trends

Data

Google Trends calculates the relative popularity of a search term over time. Google trends measure the relative popularity of a search term, normalizing the level to 100 for the peak search popularity. I match each Google metropolitan areas to the appropriate CBSA(s).

Google Trends are correlated with Uber labor supply, according to Katz and Krueger (2016). Figure A.1 shows the trends for searches "uber" and "lyft" in the "San Francisco - Oakland - San Jose CA" metropolitan area. Figure A.2 shows "uber" + "lyft" over time with the launch dates in the corresponding metropolitan areas.

Results

The figure and tables below replace the uber launch treatment and trend variables with the location-specific Google trend value for "Uber" and "Lyft". The variable is

Table A.1: UberX Launch Dates with Housing Data Available

CBSA	Launch Date	CBSA	Launch Date
San Francisco-Oakland-Fremont, CA	7/4/2012	Keene, NH	10/17/2014
Santa Barbara-Santa Maria-Goleta, CA	8/27/2012	Laconia, NH	10/17/2014
New York-Northern New Jersey-Long Island, NY-NJ-PA	8/27/2012	Lebanon, NH-VT	10/17/2014
Los Angeles-Long Beach-Santa Ana, CA	3/14/2013	Manchester-Nashua, NH	10/17/2014
Boston-Cambridge-Quincy, MA-NH	3/15/2013	Dayton, OH	10/18/2014
Chicago-Joliet-Naperville, IL-IN-WI	4/25/2013	Lincoln, NE	10/24/2014
San Diego-Carlsbad-San Marcos, CA	5/10/2013	Las Vegas-Paradise, NV	10/24/2014
Atlanta-Sandy Springs-Marietta, GA	6/28/2013	Kalamazoo-Portage, MI	10/25/2014
Washington Court House, OH	8/8/2013	Burlington-South Burlington, VT	11/1/2014
Phoenix-Mesa-Glendale, AZ	9/5/2013	Little Rock-North Little Rock-Conway, AR	11/6/2014
Charlottesville, VA	9/21/2013	Chattanooga, TN-GA	11/14/2014
Minneapolis-St. Paul-Bloomington, MN-WI	9/28/2013	Reno-Sparks, NV	11/15/2014
Providence-New Bedford-Fall River, RI-MA	10/5/2013	Portland-Vancouver-Hillsboro, OR-WA	11/22/2014
Sacramento–Arden-Arcade–Roseville, CA	10/19/2013	Cape Coral-Fort Myers, FL	12/6/2014
Baltimore-Towson, MD	10/26/2013	Naples-Marco Island, FL	12/6/2014
Tucson, AZ	11/1/2013	North Port-Bradenton-Sarasota, FL	12/12/2014
Detroit-Warren-Livonia, MI	11/2/2013	Deltona-Daytona Beach-Ormond Beach, FL	12/13/2014
Oklahoma City, OK	11/16/2013	Toledo, OH	12/27/2014
Nashville-Davidson–Murfreesboro–Franklin, TN	12/14/2013	Dover, DE	12/31/2014
Pittsburgh, PA	2/12/2014	Seaford, DE	12/31/2014
Fresno, CA	2/22/2014	Pensacola-Ferry Pass-Brent, FL	12/31/2014
Savannah, GA	3/15/2014	Akron, OH	12/31/2014
Hilton Head Island-Beaufort, SC	3/15/2014	Fayetteville, NC	1/10/2015
Cincinnati-Middletown, OH-KY-IN	3/22/2014	Springfield, IL	1/12/2015
Tulsa, OK	4/5/2014	Pittsfield, MA	1/31/2015
Cleveland-Elyria-Mentor, OH	4/12/2014	Springfield, MA	1/31/2015
Riverside-San Bernardino-Ontario, CA	4/19/2014	Harrisburg-Carlisle, PA	1/31/2015
Ann Arbor, MI	4/22/2014	State College, PA	2/6/2015
Atlantic City-Hammonton, NJ	4/23/2014	Scranton–Wilkes-Barre, PA	2/13/2015
Ocean City, NJ	4/23/2014	Champaign-Urbana, IL	2/18/2015
Trenton-Ewing, NJ	4/23/2014	Panama City-Lynn Haven-Panama City Beach, FL	2/28/2015
Vineland-Millville-Bridgeton, NJ	4/23/2014	Rockford, IL	3/14/2015
Allentown-Bethlehem-Easton, PA-NJ	4/23/2014	Peoria, IL	3/21/2015
Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	4/23/2014	Lancaster, PA	3/28/2015
Raleigh-Cary, NC	4/25/2014	Augusta-Richmond County, GA-SC	4/6/2015
Louisville/Jefferson County, KY-IN	4/26/2014	Key West, FL	4/18/2015
Virginia Beach-Norfolk-Newport News, VA-NC	5/3/2014	Erie, PA	4/18/2015
Jacksonville, IL	5/9/2014	Stillwater, OK	5/2/2015
Memphis, TN-MS-AR	5/10/2014	Mobile, AL	5/15/2015
Miami-Fort Lauderdale-Pompano Beach, FL	6/4/2014	Fargo, ND-MN	5/15/2015
Orlando-Kissimmee-Sanford, FL	6/4/2014	Kill Devil Hills, NC	5/23/2015
San Luis Obispo-Paso Robles, CA	6/13/2014	Greenville, NC	6/20/2015
Omaha-Council Bluffs, NE-IA	6/14/2014	Jacksonville, NC	6/20/2015
Wilmington, NC	6/28/2014	Goldsboro, NC	6/20/2015
Greensboro-High Point, NC	7/4/2014	Morehead City, NC	6/20/2015
Winston-Salem, NC	7/4/2014	New Bern, NC	6/20/2015
Charleston-North Charleston-Summerville, SC	7/10/2014	Reading, PA	8/29/2015
Columbia, SC	7/12/2014	Bowling Green, KY	9/12/2015
Greenville-Mauldin-Easley, SC	7/12/2014	Flint, MI	10/2/2015
Bakersfield-Delano, CA	7/19/2014	Albany, GA	10/29/2015
Cambridge, MD	7/19/2014	Bainbridge, GA	10/29/2015
Easton, MD	7/19/2014	Brunswick, GA	10/29/2015
Ocean Pines, MD	7/19/2014	Douglas, GA	10/29/2015
Salisbury, MD	7/19/2014	Fitzgerald, GA	10/29/2015
Oxnard-Thousand Oaks-Ventura, CA	7/26/2014	Hinesville-Fort Stewart, GA	10/29/2015
Grand Rapids-Wyoming, MI	7/26/2014	Jesup, GA	10/29/2015
Eugene-Springfield, OR	8/2/2014	St. Marys, GA	10/29/2015
Richmond, VA	8/7/2014	Thomasville, GA	10/29/2015
Lansing-East Lansing, MI	8/22/2014	Tifton, GA	10/29/2015
Asheville, NC	8/22/2014	Valdosta, GA	10/29/2015
Modesto, CA	8/23/2014	Vidalia, GA	10/29/2015
Lexington-Fayette, KY	8/23/2014	Waycross, GA	10/29/2015
Gainesville, FL	8/28/2014	Birmingham-Hoover, AL	12/29/2015
Tallahassee, FL	8/28/2014	Ocala, FL	1/1/2016
Tuscaloosa, AL	8/29/2014	Huntsville, AL	3/4/2016
Knoxville, TN	8/29/2014	Lake Havasu City-Kingman, AZ	3/11/2016
Roanoke, VA	8/29/2014	Yuma, AZ	3/17/2016
Myrtle Beach-North Myrtle Beach-Conway, SC	8/30/2014	Youngstown-Warren-Boardman, OH-PA	6/25/2016
Athens-Clarke County, GA	9/12/2014	Gettysburg, PA	8/5/2016
Worcester, MA	9/13/2014	York-Hanover, PA	8/5/2016
Harrisonburg, VA	9/13/2014	Altoona, PA	8/20/2016
Montgomery, AL	9/27/2014	Johnstown, PA	8/20/2016
Flagstaff, AZ	10/4/2014	Lewisburg, PA	11/24/2016
Fayetteville-Springdale-Rogers, AR-MO	10/11/2014	Lewistown, PA	11/24/2016
Berlin, NH-VT	10/17/2014	Selinsgrove, PA	11/24/2016
Claremont, NH	10/17/2014	Williamsport, PA	11/24/2016
Concord, NH	10/17/2014		

$log(ride\_trend)$ where $ride\_trend = Uber + Lyft$. The results should be interpreted as elasticities.

Table A.2: Effect of ridesharing on real estate prices (sample robustness).

	(1) log(price)	(2) log(price)	(3) log(price)	(4) log(price)
𝟙{Uber launched}	0.0381**	-0.00539	-0.00189	0.00423
	(2.63)	(-1.76)	(-0.63)	(1.43)
𝟙{Uber...}*Trend	0.946***	0.00555	0.0620***	0.0229***
	(3.46)	(0.42)	(6.35)	(5.97)
Observations	704569	1044748	1336384	2564156
Months before and after treatment	12	18	24	

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Standard errors are two-way clustered by city and month-of-sample. All columns correspond to estimation of Equation 1.2 which include month-of-sample fixed effects, tract-specific trends, city-by-month fixed effects, and house specific covariates.

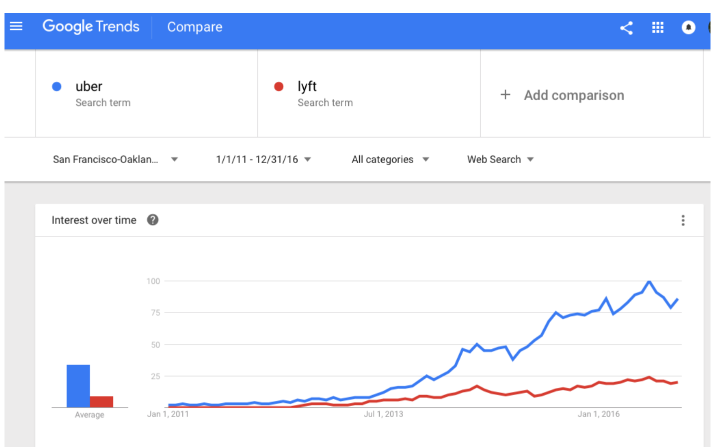


Figure A.1: Google Trends search for "uber" and "lyft" in the "San Francisco - Oakland - San Jose CA" metropolitan area.

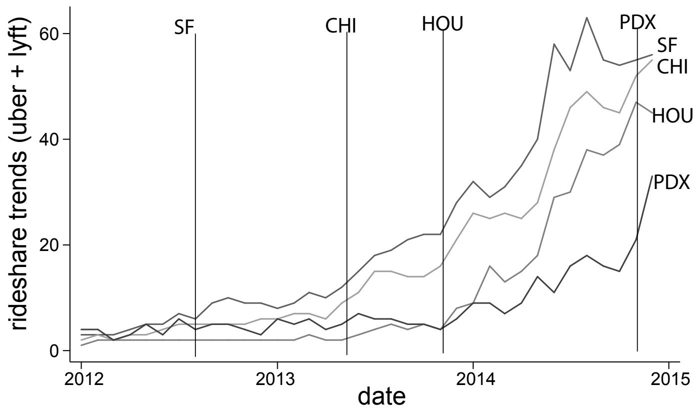


Figure A.2: Google Trends and UberX launch date for 4 US cities.
Cities included here are San Francisco, CA, Chicago, IL, Houston, TX, and Portland, OR.

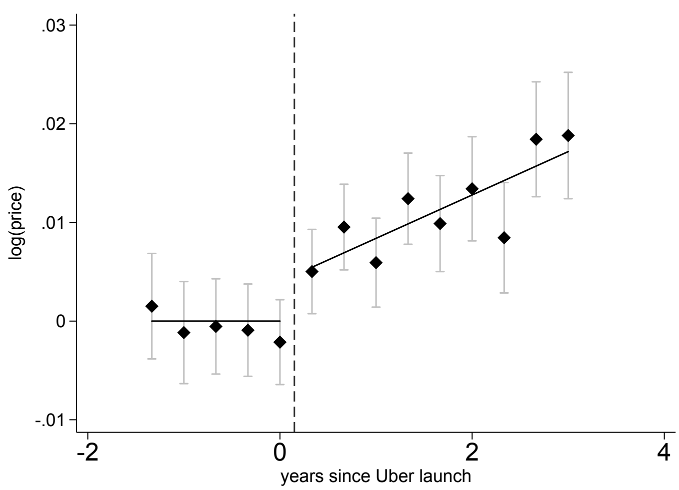


Figure A.3: Effect of uber on real estate prices: Trend-break model using Google Trends
Point estimates are dummy variables for 4 month periods before and after treatment and normalize all estimates relative to the estimated pre-trend. The estimation includes census-tract level trends, city-specific month dummies, house age, house age squared, house size, and house size squared. Standard errors are two-way clustered by city and month-of-sample. Linear trends are estimated through dummy estimates.

Table A.3: Effect of ridesharing on real estate prices (Google Trends treatment, model robustness.

	(1) log(price)	(2) log(price)	(3) log(price)	(4) log(price)	(5) log(price)
log(ride trend)	0.0986***	0.0662***	0.0689***	0.0698***	0.0743***
	(47.10)	(29.81)	(29.74)	(29.96)	(31.48)
Observations	2749030	2749030	2749030	2740682	2452134
Year FE	X	X	X	X	X
Month-of-sample FE	X	X	X	X	X
Tract-specific trend		X	X	X	X
City-by-month FE			X	X	X
Uber service area				X	X
House covariates					X

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Standard errors are two-way clustered by city and month-of-sample.

Table A.4: Effect of ridesharing on real estate prices (Google Trends treatment, sample robustness.

	(1) log(price)	(2) log(price)	(3) log(price)	(4) log(price)
log(ride trend)	0.0129*	0.0142***	0.0338***	0.0743***
	(2.01)	(3.76)	(10.37)	(31.48)
Observations	674162	1000570	1280333	2452134
Months before and after treatment	12	18	24	

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.5: Standard errors are two-way clustered by city and month-of-sample.

Appendix B

Time-use and Energy Consumption - Appendix

B.1 Data

Time Use

Figure B.1 shows the proportion of the population engaged in each of the 18 times use categories from the American Time Use Survey (ATUS).

Seasonal time use for each of our selected categories, corresponding to figure 3.2, but broken down seasonally, are shown below in figure B.2

B.2 Model

The seasonally interacted version of the model is:

$$\begin{aligned}\log(electricity\_load_{dhs}) = \sum_{s\in\{winter,spring,summer,fall\}} [\alpha_s \\ + \sum_{p=1}^{3} \beta^p_{sleep,s} * sleep^p_{dhs} \\ + \sum_{p=1}^{2} \beta^p_{work,s} * work^p_{dhs} \\ + \sum_{p=1}^{3} \beta^p_{lesiure,s} * leisure^p_{dhs} + \varepsilon_{dhs}]\end{aligned} \tag{B.1}$$

where d is day-type, h is hour-of-day, s is the season ($\{winter, spring, summer, fall\}$) and $\{sleep, work, leisure\}$ represent the proportion of the U.S. population engaged in each activity. Based on our cross-validation, the function $h_a(\cdot)$ is cubic in sleep (when $a = sleep$), quadratic in work, and cubic in leisure.

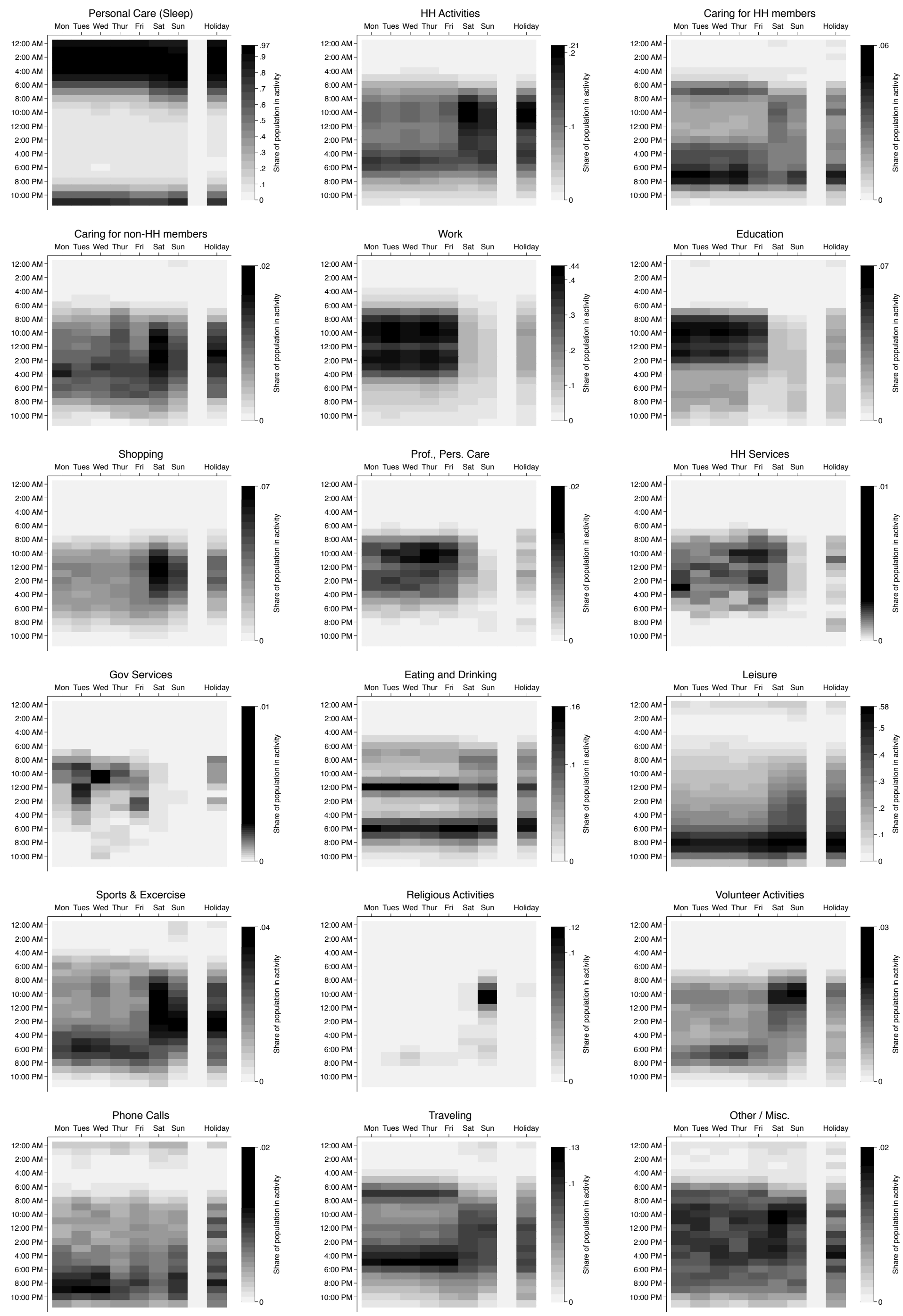


Figure B.1: Time Use by hour and day-type in the U.S.
Data is collapsed from the American Time Use Survey from the years 2006-2016. The "Holiday" day-type includes all 10 federal holidays.

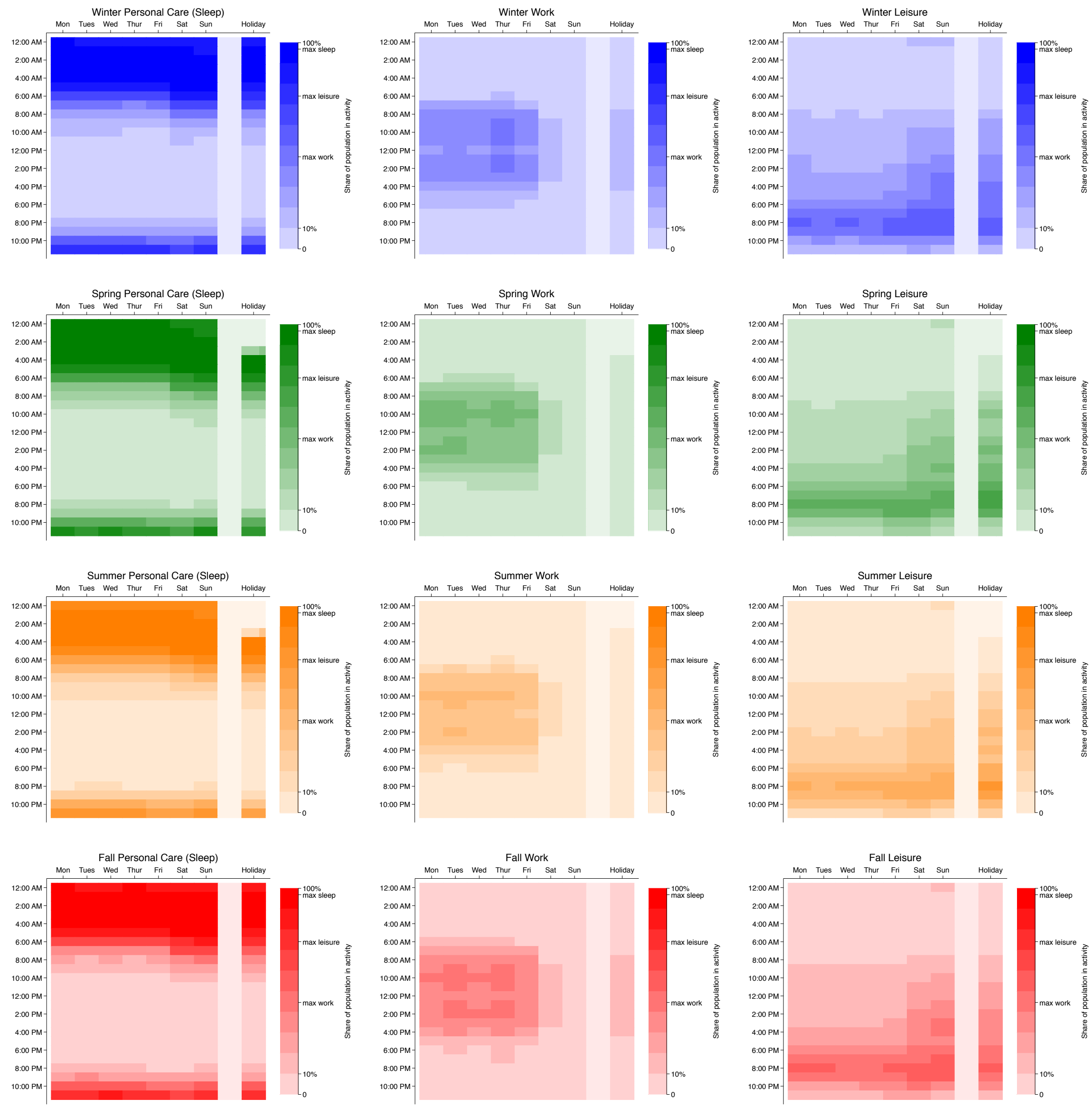
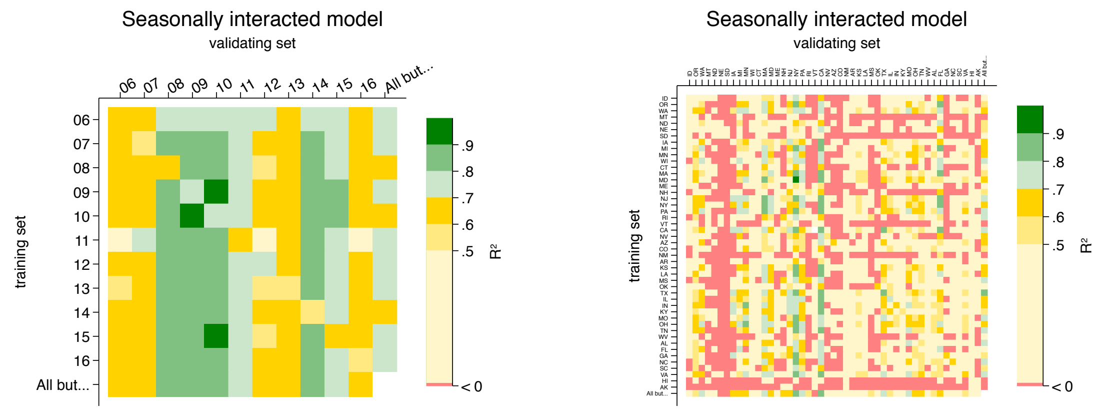


Figure B.2: Sleep, work, leisure by hour, day-type, and season in the U.S. Data is collapsed from the American Time Use Survey from the years 2006-2016 to create population shares. The "Holiday" day-type includes all 10 federal holidays.

B.3 Results

Seasonal Cross Validation

Cross-validation by year and by state, performed using the seasonal specification, is shown in figures B.3a and B.3, respectively.



(a) Year-by-year **(b)** State-by-state

Figure B.3: Cross validation of predicted electricity use across years and states using seasonal model.

Seasonal Marginal Effects

Marginal effect results for winter, spring, summer, and fall are shown in figures B.4, B.5, B.6, and B.7, respectively.

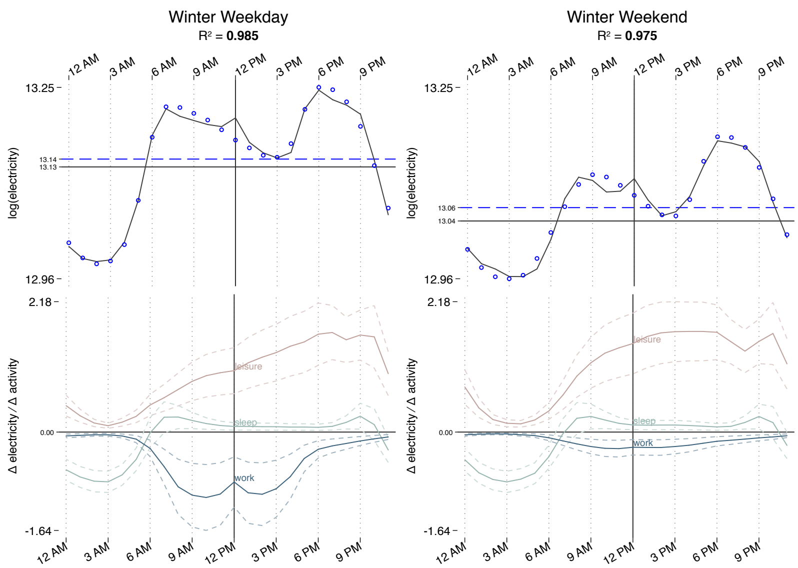


Figure B.4: Marginal effects of participation in sleep, work, and leisure by hour of the day for weekdays and weekends during winter.
The top panel is prediction for each hour. The bottom panel gives the marginal effect on electricity use of each time use activity.

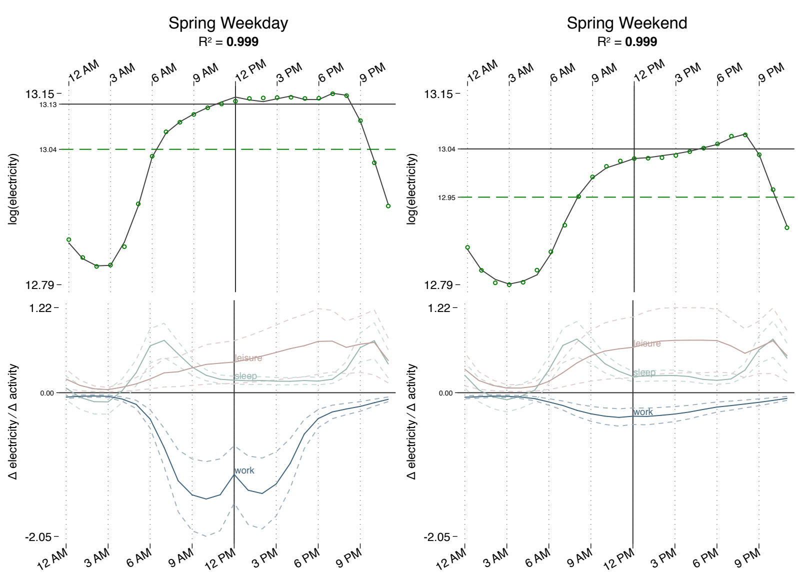


Figure B.5: Marginal effects of participation in sleep, work, and leisure by hour of the day for weekdays and weekends during spring.
The top panel is prediction for each hour. The bottom panel gives the marginal effect on electricity use of each time use activity.

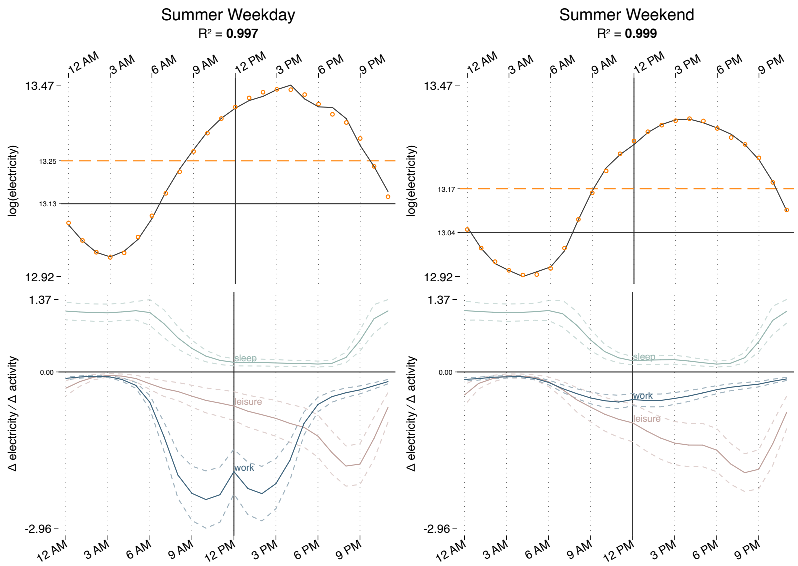


Figure B.6: Marginal effects of participation in sleep, work, and leisure by hour of the day for weekdays and weekends during summer.
The top panel is prediction for each hour. The bottom panel gives the marginal effect on electricity use of each time use activity.

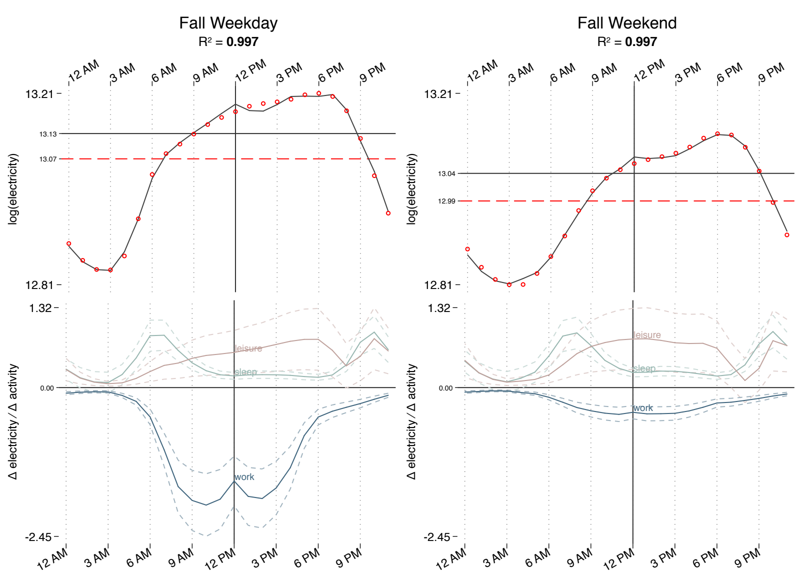


Figure B.7: Marginal effects of participation in sleep, work, and leisure by hour of the day for weekdays and weekends during fall.
The top panel is prediction for each hour. The bottom panel gives the marginal effect on electricity use of each time use activity.

Appendix C

Population Scale Coordination of Leisure Reduces Energy Consumption - Appendix

C.1 Data

The raw data for a single year of air travel is shown in figure C.1 and the raw data for a single year of vehicle travel is shown in figure C.2. These figures coorespond to the electricity data plot in figure 4.3.

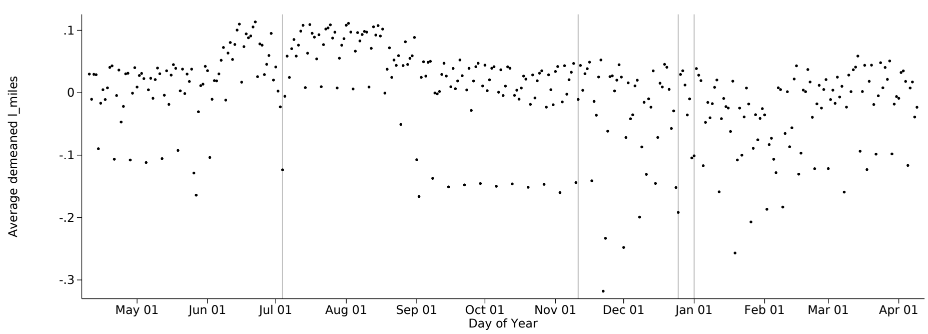


Figure C.1: Average demeaned log(air_miles) by day of year, single year. Data for each flight origin is demeaned, and the mean across all airports is taken for each day. The time period is April 10, 2007 through April 9, 2008. Vertical lines are July 4 (Independence Day), November 11 (Veterans Day), December 25 (Christmas Day), and January 1 (New Year's Day).

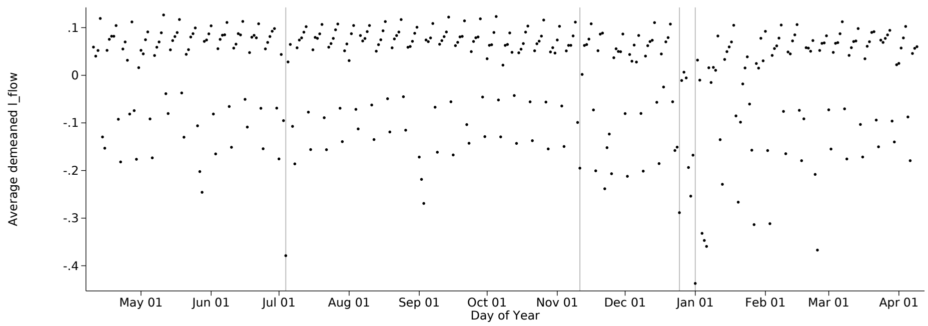


Figure C.2: Average demeaned $\log(vehicle\_flow)$ by day of year, single year. Data for each highway sensor is demeaned, and the mean across all highway sensors is taken for each day. The time period is April 10, 2007 through April 9, 2008. Vertical lines are July 4 (Independence Day), November 11 (Veterans Day), December 25 (Christmas Day), and January 1 (New Year's Day).

C.2 Materials and Methods - Air and Vehicle Travel, Timeuse

The preferred specification for the air travel estimation is

$$\begin{aligned}\log(air\_miles_{ity}) = \alpha + \sum_{h \in H} \beta_h * DayofInterestDummy_h \\ + \mu_i + \delta_y + \sum_{d=1}^{6} \psi_d * DayofWeekDummy_d \\ + f(t) + \varepsilon_{ity}\end{aligned} \tag{C.1}$$

where H is the set of holidays and surrounding days, $i \in I$ is the set of airports, $t \in T$ is the set of days of the year $1, 2, ..., 365$, $y \in Y$ is the set of years 1991 to 2014, $f(\cdot)$ is a 15^{th} order polynomial, and $g(\cdot)$ is a 5^{th} order polynomial.

The preferred specification for the vehicle travel estimation is

$$\begin{aligned}\log(vehicle\_flow_{ity}) = \alpha + \sum_{h \in H} \beta_h * DayofInterestDummy_h \\ + \mu_i + \delta_y + \sum_{d=1}^{6} \psi_d * DayofWeekDummy_d \\ + f(t) + \varepsilon_{ity}\end{aligned} \tag{C.2}$$

where H is the set of holidays and surrounding days, $i \in I$ is the set of vehicle sensors in California, $t \in T$ is the set of days of the year $1, 2, ..., 365$, $y \in Y$ is the set of years 2001 to 2014, $f(\cdot)$ is a 15^{th} order polynomial, and $g(\cdot)$ is a 5^{th} order polynomial.

The equation to estimate hourly timeuse results also correspond closely to the electricity equation. The model is run separately for each hour of the day:

$$\begin{aligned} activity_{ity} = \alpha + \sum_{h \in H} \beta_h * DayofInterestDummy_h \\ + \mu_i + \delta_y + \sum_{d=1}^{6} \psi_d * DayofWeekDummy_d \\ + f(t) + g(MaxTemp_{ity}) + g(MinTemp_{ity}) + \varepsilon_{ity} \end{aligned} \tag{C.3}$$

where $activity_{ityh}$ is the proportion of the population engaged in an activity $\{sleep, work, leisure\}$, H is the set of holidays and surrounding days, $i \in I$ is the set of states, $t \in T$ is the set of days of the year $1, 2, ..., 365$, $y \in Y$ is the set of years 2006 to 2014, $f(\cdot)$ is a 15^{th} order polynomial, and $g(\cdot)$ is a 5^{th} order polynomial.

C.3 Results - Robustness

Below we show tables with three types of sensitivities. First, we show different models. Second, we show our preferred model but with different numbers of days before and after holidays included as days of interest. Third, we show our preferred model but with different orders of polynomials for removing seasonal variation. These tables also support the assertion that, because coordinated leisure spills over into days around holidays, our estimates would be biased downward by not including surrounding days as days of interest.

Electricity

Table C.1 shows the effect of each holiday on electricity load using different sets of fixed effects and model components.

Table C.2 shows the effect of each holiday on electricity load using different number of days-before and days-after the holiday as days of interest.

Finally, we show robustness tables for our choice of polynomial order in table C.3.

C.4 Air Travel

Table C.4 shows the effect of each holiday on air miles traveled using different sets of fixed effects and model components.

Table C.5 shows the effect of each holiday on air miles traveled using different number of days-before and days-after the holiday as days of interest.

Finally, we show robustness tables for our choice of polynomial order in table C.6.

Table C.1: Holiday and Weekend Electricity Effects - Model Sensitivity

	(1)	(2)	(3)	(4)
Christmas	-0.0492***	-0.0474***	-0.0377**	-0.0985***
	(0.00447)	(0.00917)	(0.0171)	(0.0203)
Observed Christmas	-0.0837***	-0.0819***	-0.0572***	-0.0584***
	(0.00594)	(0.0103)	(0.0162)	(0.0161)
New Year's	-0.0483***	-0.0534***	-0.00777	-0.0431***
	(0.00474)	(0.0103)	(0.0113)	(0.0158)
Observed New Year's	-0.0618***	-0.0602***	-0.0435***	-0.0450***
	(0.00706)	(0.0123)	(0.0138)	(0.0130)
Martin Luther King, Jr.	-0.00956	-0.00860	0.0398***	-0.00513
	(0.00875)	(0.00850)	(0.00965)	(0.00991)
Presidents	-0.00405	-0.00451	0.0258***	-0.00329
	(0.00330)	(0.00428)	(0.00629)	(0.00439)
Memorial	-0.0718***	-0.0915***	-0.116***	-0.0915***
	(0.00927)	(0.00972)	(0.0109)	(0.01000)
Observed Independence	-0.0697***	-0.0640***	-0.0478***	-0.0470***
	(0.00449)	(0.00472)	(0.00329)	(0.00321)
Independence	-0.0494***	-0.0435***	-0.0265***	-0.0599***
	(0.00320)	(0.00482)	(0.00648)	(0.00510)
Labor	-0.0595***	-0.0916***	-0.0763***	-0.0899***
	(0.0127)	(0.0106)	(0.0110)	(0.0106)
Columbus	-0.00594	-0.00486	-0.0439***	-0.00784
	(0.00440)	(0.00649)	(0.00961)	(0.00655)
Veterans	-0.00538**	-0.00435	-0.0170**	-0.00592
	(0.00262)	(0.00438)	(0.00794)	(0.00565)
Observed Veterans	-0.00356	-0.00332	-0.00355	-0.00322
	(0.00219)	(0.00332)	(0.00514)	(0.00433)
Thanksgiving	-0.110***	-0.131***	-0.120***	-0.134***
	(0.00979)	(0.0105)	(0.0109)	(0.0106)
Monday	-0.00718***	-0.00655***	-0.00504*	-0.00541***
	(0.00131)	(0.00125)	(0.00251)	(0.00125)
Tuesday	-0.000769	-0.000858	-0.000346	-0.000139
	(0.00113)	(0.00106)	(0.00244)	(0.00118)
Thursday	-0.000497	-0.000199	-0.000529	-0.000750
	(0.00117)	(0.00105)	(0.00239)	(0.00112)
Friday	-0.0149***	-0.0151***	-0.0134***	-0.0139***
	(0.00150)	(0.00150)	(0.00257)	(0.00143)
Saturday	-0.0802***	-0.0802***	-0.0781***	-0.0794***
	(0.00411)	(0.00408)	(0.00467)	(0.00412)
Sunday	-0.100***	-0.100***	-0.0984***	-0.0993***
	(0.00524)	(0.00521)	(0.00564)	(0.00526)
Observations	140,019	140,019	140,019	140,019
R-squared	0.996	0.996	0.995	0.996
Month-of-Sample FE	X		X	
State-by-Year FE	X	X	X	X
Min. / Max. Temp. Polynomial	X	X	X	X
Day-of-Year Polynomial		X		X
8 Days Before / After Holidays			X	X

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

This table presents four specifications for estimating the effect of federal holidays on electricity loads. Specifications (1) and (2) do not include the days before and after the holiday as days of interest. Specifications (3) and (4) include 8-days before and after each holiday to control for spillovers. Specifications (1) and (3) use Month-by-Year fixed-effects to control for seasonal variation, whereas specifications (2) and (4) use a 15^{th} order polynomial.

C.5 Vehicle Travel

All vehicle travel sensitivities use a 5% random sample due to computing constraints.

Table C.2: Holiday and Weekend Electricity Effects - Days Before and After Sensitivity

	(1) 2 days	(2) 5 days	(3) 8 days	(4) 10 days
Christmas	-0.0882***	-0.0884***	-0.0946***	-0.110***
	(0.0160)	(0.0167)	(0.0184)	(0.0207)
Observed Christmas	-0.0583***	-0.0582***	-0.0584***	-0.0587***
	(0.0159)	(0.0160)	(0.0161)	(0.0162)
New Year's	-0.0538***	-0.0519***	-0.0452***	-0.0231
	(0.0116)	(0.0124)	(0.0152)	(0.0180)
Observed New Year's	-0.0447***	-0.0447***	-0.0450***	-0.0453***
	(0.0129)	(0.0130)	(0.0130)	(0.0131)
Martin Luther King, Jr.	-0.00865	-0.00681	-0.00507	-0.000826
	(0.00857)	(0.00888)	(0.00994)	(0.0104)
Presidents	-0.00169	-0.00318	-0.00361	-0.00401
	(0.00435)	(0.00437)	(0.00438)	(0.00480)
Memorial	-0.0916***	-0.0895***	-0.0913***	-0.0928***
	(0.00991)	(0.0101)	(0.01000)	(0.00994)
Independence	-0.0582***	-0.0607***	-0.0601***	-0.0586***
	(0.00460)	(0.00493)	(0.00511)	(0.00555)
Observed Independence	-0.0469***	-0.0467***	-0.0470***	-0.0473***
	(0.00308)	(0.00317)	(0.00321)	(0.00321)
Labor	-0.0911***	-0.0912***	-0.0902***	-0.0881***
	(0.0107)	(0.0107)	(0.0107)	(0.0107)
Columbus	-0.00595	-0.00595	-0.00752	-0.0104
	(0.00662)	(0.00663)	(0.00661)	(0.00684)
Veterans	-0.00487	-0.00699	-0.00670	-0.00302
	(0.00489)	(0.00503)	(0.00548)	(0.00605)
Observed Veterans	-0.00302	-0.00295	-0.00323	-0.00110
	(0.00433)	(0.00430)	(0.00433)	(0.00466)
Thanksgiving	-0.132***	-0.135***	-0.135***	-0.127***
	(0.0103)	(0.0102)	(0.0104)	(0.0108)
Monday	-0.00674***	-0.00558***	-0.00541***	-0.00548***
	(0.00119)	(0.00119)	(0.00125)	(0.00129)
Tuesday	-0.000935	-0.000215	-0.000149	-0.000323
	(0.00107)	(0.00108)	(0.00117)	(0.00124)
Thursday	-0.000416	-0.000727	-0.000759	-0.00112
	(0.00101)	(0.00104)	(0.00112)	(0.00119)
Friday	-0.0135***	-0.0139***	-0.0139***	-0.0145***
	(0.00137)	(0.00141)	(0.00143)	(0.00152)
Saturday	-0.0798***	-0.0794***	-0.0794***	-0.0803***
	(0.00401)	(0.00410)	(0.00412)	(0.00416)
Sunday	-0.0997***	-0.0988***	-0.0993***	-0.1000***
	(0.00516)	(0.00520)	(0.00526)	(0.00527)
Observations	140,019	140,019	140,019	140,019
R-squared	0.996	0.996	0.996	0.996
State-by-Year FE	X	X	X	X
Min. / Max. Temp. Polynomial	X	X	X	X
Day-of-Year Polynomial	X	X	X	X
2 Days Before / After Holidays	X			
5 Days Before / After Holidays		X		
10 Days Before / After Holidays				X
8 Days Before / After Holidays			X	

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

All specifications include fixed effects from the primary specification. The columns vary the number of days estimated before and after each holiday.

Table C.7 shows the effect of each holiday on vehicle flow using different sets of fixed effects and model components.

Table C.5 shows the effect of each holiday on vehicle miles traveled using different number of days-before and days-after the holiday as days of interest.

Table C.3: Holiday and Weekend Electricity Effects - Polynomial Sensitivity

	(1) 9th order	(2) 11th order	(3) 13th order	(4) 15th order
Christmas	-0.112***	-0.106***	-0.0986***	-0.0946***
	(0.0167)	(0.0176)	(0.0180)	(0.0184)
Observed Christmas	-0.0584***	-0.0584***	-0.0584***	-0.0584***
	(0.0161)	(0.0161)	(0.0161)	(0.0161)
New Year's	-0.0328***	-0.0420***	-0.0481***	-0.0452***
	(0.0116)	(0.0139)	(0.0150)	(0.0152)
Observed New Year's	-0.0449***	-0.0449***	-0.0450***	-0.0450***
	(0.0130)	(0.0130)	(0.0130)	(0.0130)
Martin Luther King, Jr.	-0.00539	-0.00755	-0.00691	-0.00507
	(0.00907)	(0.00969)	(0.00984)	(0.00994)
Presidents	-0.00191	-0.00139	-0.00361	-0.00361
	(0.00440)	(0.00454)	(0.00434)	(0.00438)
Memorial	-0.0919***	-0.0932***	-0.0930***	-0.0913***
	(0.00971)	(0.00968)	(0.00981)	(0.01000)
Independence	-0.0612***	-0.0598***	-0.0595***	-0.0601***
	(0.00521)	(0.00539)	(0.00525)	(0.00511)
Observed Independence	-0.0469***	-0.0469***	-0.0470***	-0.0470***
	(0.00321)	(0.00322)	(0.00322)	(0.00321)
Labor	-0.0879***	-0.0878***	-0.0890***	-0.0902***
	(0.0108)	(0.0109)	(0.0108)	(0.0107)
Columbus	-0.0114*	-0.0104	-0.00899	-0.00752
	(0.00671)	(0.00664)	(0.00662)	(0.00661)
Veterans	0.000651	-0.00172	-0.00424	-0.00670
	(0.00553)	(0.00548)	(0.00541)	(0.00548)
Observed Veterans	-0.00320	-0.00321	-0.00323	-0.00323
	(0.00433)	(0.00433)	(0.00433)	(0.00433)
Thanksgiving	-0.126***	-0.129***	-0.132***	-0.135***
	(0.0106)	(0.0107)	(0.0107)	(0.0104)
Monday	-0.00528***	-0.00530***	-0.00532***	-0.00541***
	(0.00125)	(0.00125)	(0.00124)	(0.00125)
Tuesday	-1.28e-05	-4.76e-05	-7.12e-05	-0.000149
	(0.00116)	(0.00117)	(0.00116)	(0.00117)
Thursday	-0.000748	-0.000768	-0.000775	-0.000759
	(0.00113)	(0.00113)	(0.00113)	(0.00112)
Friday	-0.0139***	-0.0139***	-0.0139***	-0.0139***
	(0.00144)	(0.00144)	(0.00144)	(0.00143)
Saturday	-0.0793***	-0.0794***	-0.0794***	-0.0794***
	(0.00414)	(0.00414)	(0.00413)	(0.00412)
Sunday	-0.0992***	-0.0992***	-0.0992***	-0.0993***
	(0.00528)	(0.00528)	(0.00527)	(0.00526)
Observations	140,019	140,019	140,019	140,019
R-squared	0.996	0.996	0.996	0.996
State-by-Year FE	X	X	X	X
Min. / Max. Temp. Polynomial	X	X	X	X
Day-of-Year Polynomial	X	X	X	X
8 Days Before / After Holidays	X	X	X	X

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

All specifications include fixed effects from the primary specification. The columns vary the order of polynomial that is used to remove seasonal variation.

Finally, we show robustness tables for our choice of polynomial order in table C.9.

Table C.4: Holiday and Weekend Air Travel Effects - Model Sensitivity

	(1)	(2)	(3)	(4)
Christmas	-0.105***	-0.108***	-0.177***	-0.103***
	(0.0160)	(0.0166)	(0.0202)	(0.0237)
Observed Christmas	-0.0586***	-0.0621***	-0.0412***	-0.0406***
	(0.0176)	(0.0185)	(0.0138)	(0.0138)
New Year's	-0.0287	-0.0311	-0.0665***	-0.0229
	(0.0202)	(0.0214)	(0.0181)	(0.0220)
Observed New Year's	-0.0574**	-0.0551**	-0.0446**	-0.0427**
	(0.0223)	(0.0235)	(0.0173)	(0.0172)
Martin Luther King, Jr.	0.000845	-0.00170	-0.0270***	0.00739
	(0.00802)	(0.00836)	(0.00825)	(0.00945)
Presidents	0.00714	0.00842	-0.0122	0.0126
	(0.00884)	(0.00874)	(0.00898)	(0.00910)
Memorial	-0.0371***	-0.0522***	-0.0579***	-0.0651***
	(0.00507)	(0.00628)	(0.00674)	(0.00695)
Observed Independence	-0.0459***	-0.0453***	-0.0176	-0.0170
	(0.0134)	(0.0112)	(0.0123)	(0.0122)
Independence	-0.120***	-0.119***	-0.0932***	-0.140***
	(0.0144)	(0.0121)	(0.0150)	(0.0126)
Labor	-0.00435	-0.0311***	-0.0364***	-0.0339***
	(0.0109)	(0.00661)	(0.00843)	(0.00720)
Columbus	0.00656	0.00702	-0.0308***	0.0130**
	(0.00405)	(0.00468)	(0.00689)	(0.00508)
Veterans	0.0111*	0.00875	-0.0382***	0.0109
	(0.00623)	(0.00945)	(0.0121)	(0.0104)
Observed Veterans	0.00626	0.00348	0.00254	0.00315
	(0.00555)	(0.00921)	(0.00933)	(0.00927)
Thanksgiving	-0.309***	-0.300***	-0.358***	-0.292***
	(0.0176)	(0.0201)	(0.0215)	(0.0205)
Monday	0.00878***	0.00940***	0.0107***	0.00953***
	(0.00159)	(0.00175)	(0.00228)	(0.00182)
Tuesday	-0.00967***	-0.00974***	-0.00701***	-0.00838***
	(0.00200)	(0.00215)	(0.00257)	(0.00211)
Thursday	0.00883***	0.00865***	0.00845***	0.00835***
	(0.00194)	(0.00207)	(0.00257)	(0.00224)
Friday	0.00774***	0.00771***	0.0122***	0.0120***
	(0.00182)	(0.00194)	(0.00223)	(0.00183)
Saturday	-0.110***	-0.110***	-0.107***	-0.105***
	(0.00732)	(0.00736)	(0.00752)	(0.00735)
Sunday	-0.0409***	-0.0412***	-0.0340***	-0.0348***
	(0.00457)	(0.00463)	(0.00485)	(0.00458)
Observations	2,541,130	2,541,130	2,541,130	2,541,130
R-squared	0.986	0.986	0.986	0.986
Month-of-Sample FE	X		X	
Origin-by-Year FE	X	X	X	X
Day-of-Year Polynomial		X		X
8 Days Before / After Holidays			X	X

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

This table presents four specifications for estimating the effect of federal holidays on air miles traveled. Specifications (1) and (2) do not include the days before and after the holiday as days of interest. Specifications (3) and (4) include 8-days before and after each holiday to control for spillovers. Specifications (1) and (3) use Month-by-Year fixed-effects to control for seasonal variation, whereas specifications (2) and (4) use a 15^{th} order polynomial.

Table C.5: Holiday and Weekend Air Travel Effects - Days Before and After Sensitivity

	(1) 2 days	(2) 5 days	(3) 8 days	(4) 10 days
Christmas	-0.135***	-0.131***	-0.116***	-0.0939***
	(0.0158)	(0.0170)	(0.0204)	(0.0241)
Observed Christmas	-0.0418***	-0.0409***	-0.0406***	-0.0395***
	(0.0138)	(0.0137)	(0.0138)	(0.0137)
New Year's	-0.0268	-0.0144	-0.0161	-0.0354
	(0.0190)	(0.0195)	(0.0220)	(0.0245)
Observed New Year's	-0.0441**	-0.0431**	-0.0427**	-0.0417**
	(0.0173)	(0.0172)	(0.0172)	(0.0172)
Martin Luther King, Jr.	0.000804	0.00516	0.00711	0.00835
	(0.00830)	(0.00840)	(0.00945)	(0.0103)
Presidents	0.0115	0.0127	0.0136	0.0119
	(0.00879)	(0.00887)	(0.00908)	(0.00924)
Memorial	-0.0587***	-0.0620***	-0.0656***	-0.0682***
	(0.00630)	(0.00658)	(0.00695)	(0.00717)
Independence	-0.139***	-0.139***	-0.139***	-0.141***
	(0.0127)	(0.0126)	(0.0126)	(0.0125)
Observed Independence	-0.0182	-0.0174	-0.0169	-0.0159
	(0.0124)	(0.0123)	(0.0122)	(0.0121)
Labor	-0.0334***	-0.0326***	-0.0331***	-0.0369***
	(0.00675)	(0.00702)	(0.00719)	(0.00697)
Columbus	0.00848*	0.00912*	0.0119**	0.0151***
	(0.00467)	(0.00475)	(0.00507)	(0.00524)
Veterans	0.00897	0.00986	0.0135	0.00966
	(0.00982)	(0.00987)	(0.0104)	(0.0111)
Observed Veterans	0.00205	0.00279	0.00315	0.000676
	(0.00928)	(0.00925)	(0.00927)	(0.00943)
Thanksgiving	-0.301***	-0.295***	-0.289***	-0.299***
	(0.0201)	(0.0202)	(0.0206)	(0.0207)
Monday	0.0103***	0.00999***	0.00954***	0.00854***
	(0.00173)	(0.00175)	(0.00182)	(0.00185)
Tuesday	-0.0101***	-0.0105***	-0.00835***	-0.00909***
	(0.00220)	(0.00220)	(0.00211)	(0.00215)
Thursday	0.00982***	0.00862***	0.00834***	0.00878***
	(0.00205)	(0.00220)	(0.00225)	(0.00202)
Friday	0.0135***	0.0123***	0.0120***	0.0105***
	(0.00169)	(0.00181)	(0.00184)	(0.00195)
Saturday	-0.108***	-0.106***	-0.105***	-0.104***
	(0.00734)	(0.00735)	(0.00735)	(0.00733)
Sunday	-0.0338***	-0.0348***	-0.0348***	-0.0353***
	(0.00449)	(0.00449)	(0.00458)	(0.00457)
Observations	2,541,130	2,541,130	2,541,130	2,541,130
R-squared	0.986	0.986	0.986	0.986
Origin-by-Year FE	X	X	X	X
Day-of-Year Polynomial	X	X	X	X
2 Days Before / After Holidays	X			
5 Days Before / After Holidays		X		
10 Days Before / After Holidays				X
8 Days Before / After Holidays			X	

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

All specifications include fixed effects from the primary specification. The columns vary the number of days estimated before and after each holiday.

Table C.6: Holiday and Weekend Air Travel Effects - Polynomial Sensitivity

	(1) 9th order	(2) 11th order	(3) 13th order	(4) 15th order
Christmas	-0.0905***	-0.0842***	-0.0832***	-0.116***
	(0.0200)	(0.0193)	(0.0200)	(0.0204)
Observed Christmas	-0.0406***	-0.0406***	-0.0406***	-0.0406***
	(0.0138)	(0.0138)	(0.0138)	(0.0138)
New Year's	-0.0450**	-0.0539**	-0.0394*	-0.0161
	(0.0194)	(0.0209)	(0.0215)	(0.0220)
Observed New Year's	-0.0429**	-0.0429**	-0.0430**	-0.0427**
	(0.0172)	(0.0172)	(0.0172)	(0.0172)
Martin Luther King, Jr.	0.00806	0.00633	0.0111	0.00711
	(0.00916)	(0.00936)	(0.00932)	(0.00945)
Presidents	0.00840	0.00878	0.00734	0.0136
	(0.00891)	(0.00901)	(0.00902)	(0.00908)
Memorial	-0.0578***	-0.0594***	-0.0686***	-0.0656***
	(0.00691)	(0.00684)	(0.00705)	(0.00695)
Independence	-0.154***	-0.152***	-0.144***	-0.139***
	(0.0124)	(0.0124)	(0.0125)	(0.0126)
Observed Independence	-0.0170	-0.0170	-0.0170	-0.0169
	(0.0122)	(0.0122)	(0.0122)	(0.0122)
Labor	-0.0345***	-0.0340***	-0.0344***	-0.0331***
	(0.00718)	(0.00719)	(0.00712)	(0.00719)
Columbus	0.0123**	0.0133***	0.0149***	0.0119**
	(0.00496)	(0.00499)	(0.00505)	(0.00507)
Veterans	0.0159	0.0127	0.00767	0.0135
	(0.00990)	(0.0100)	(0.0102)	(0.0104)
Observed Veterans	0.00310	0.00308	0.00308	0.00315
	(0.00928)	(0.00928)	(0.00928)	(0.00927)
Thanksgiving	-0.297***	-0.300***	-0.302***	-0.289***
	(0.0202)	(0.0202)	(0.0202)	(0.0206)
Monday	0.00930***	0.00928***	0.00939***	0.00954***
	(0.00182)	(0.00182)	(0.00182)	(0.00182)
Tuesday	-0.00862***	-0.00866***	-0.00855***	-0.00835***
	(0.00211)	(0.00211)	(0.00211)	(0.00211)
Thursday	0.00829***	0.00827***	0.00830***	0.00834***
	(0.00226)	(0.00226)	(0.00225)	(0.00225)
Friday	0.0119***	0.0118***	0.0119***	0.0120***
	(0.00185)	(0.00185)	(0.00184)	(0.00184)
Saturday	-0.106***	-0.106***	-0.106***	-0.105***
	(0.00735)	(0.00735)	(0.00735)	(0.00735)
Sunday	-0.0350***	-0.0350***	-0.0349***	-0.0348***
	(0.00458)	(0.00458)	(0.00458)	(0.00458)
Observations	2,541,130	2,541,130	2,541,130	2,541,130
R-squared	0.986	0.986	0.986	0.986
Origin-by-Year FE	X	X	X	X
Day-of-Year Polynomial	X	X	X	X
8 Days Before / After Holidays	X	X	X	X

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

All specifications include fixed effects from the primary specification. The columns vary the order of polynomial that is used to remove seasonal variation.

Table C.7: Holiday and Weekend Vehicle Travel Effects - Model Sensitivity

	(1)	(2)	(3)	(4)
Christmas	-0.126***	-0.124***	-0.168***	-0.208***
	(0.0361)	(0.0359)	(0.0350)	(0.0525)
Observed Christmas	-0.155***	-0.152***	-0.128***	-0.135***
	(0.0310)	(0.0305)	(0.0313)	(0.0305)
New Year's	-0.137***	-0.121***	-0.242***	-0.145**
	(0.0428)	(0.0407)	(0.0422)	(0.0684)
Observed New Year's	-0.265***	-0.271***	-0.213***	-0.207***
	(0.0439)	(0.0399)	(0.0459)	(0.0501)
Martin Luther King, Jr.	-0.0964	-0.0974	-0.166**	-0.110
	(0.0731)	(0.0749)	(0.0750)	(0.0846)
Presidents	-0.0303	-0.0381	-0.0593*	-0.0270
	(0.0314)	(0.0330)	(0.0334)	(0.0393)
Memorial	-0.250***	-0.258***	-0.252***	-0.256***
	(0.0368)	(0.0385)	(0.0380)	(0.0419)
Observed Independence	-0.249***	-0.241***	-0.213***	-0.198***
	(0.0482)	(0.0522)	(0.0612)	(0.0657)
Independence	-0.188***	-0.180***	-0.211***	-0.210***
	(0.0475)	(0.0511)	(0.0546)	(0.0623)
Labor	-0.327***	-0.348***	-0.339***	-0.330***
	(0.0545)	(0.0559)	(0.0553)	(0.0639)
Columbus	0.0207	0.0289	0.00516	0.0128
	(0.0170)	(0.0179)	(0.0178)	(0.0206)
Veterans	-0.126	-0.132	-0.149*	-0.105
	(0.0772)	(0.0805)	(0.0857)	(0.0885)
Observed Veterans	0.112*	0.108	0.114	0.100
	(0.0639)	(0.0667)	(0.0723)	(0.0734)
Thanksgiving	-0.162***	-0.166***	-0.180***	-0.136***
	(0.0135)	(0.0143)	(0.0139)	(0.0214)
Monday	-0.0436***	-0.0429***	-0.0389***	-0.0409***
	(0.00675)	(0.00708)	(0.00898)	(0.00936)
Tuesday	-0.0148**	-0.0150**	-0.0137*	-0.0163**
	(0.00591)	(0.00632)	(0.00813)	(0.00827)
Thursday	0.0281***	0.0278***	0.0276***	0.0281***
	(0.00596)	(0.00623)	(0.00787)	(0.00789)
Friday	0.0811***	0.0817***	0.0832***	0.0822***
	(0.00655)	(0.00692)	(0.00863)	(0.00869)
Saturday	-0.0733***	-0.0730***	-0.0701***	-0.0687***
	(0.00731)	(0.00757)	(0.00921)	(0.00940)
Sunday	-0.223***	-0.222***	-0.209***	-0.213***
	(0.00711)	(0.00732)	(0.00887)	(0.00896)
Observations	987,628	987,628	987,628	19,801,068
R-squared	0.896	0.894	0.894	0.885
Month-of-Sample FE	X		X	
Sensor-by-Year FE	X	X	X	X
Day-of-Year Polynomial		X		X
8 Days Before / After Holidays			X	X

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

This table presents four specifications for estimating the effect of federal holidays on air miles traveled. Specifications (1) and (2) do not include the days before and after the holiday as days of interest. Specifications (3) and (4) include 8-days before and after each holiday to control for spillovers. Specifications (1) and (3) use Month-by-Year fixed-effects to control for seasonal variation, whereas specifications (2) and (4) use a 15^{th} order polynomial.

Table C.8: Holiday and Weekend Vehicle Travel Effects - Days Before and After Sensitivity

	(1) 2 days	(2) 5 days	(3) 8 days	(4) 10 days
Christmas	-0.168***	-0.172***	-0.203***	-0.184***
	(0.0423)	(0.0396)	(0.0464)	(0.0557)
Observed Christmas	-0.132***	-0.131***	-0.128***	-0.125***
	(0.0320)	(0.0319)	(0.0315)	(0.0314)
New Year's	-0.158***	-0.158***	-0.138**	-0.191**
	(0.0458)	(0.0475)	(0.0616)	(0.0765)
Observed New Year's	-0.217***	-0.214***	-0.211***	-0.208***
	(0.0466)	(0.0467)	(0.0461)	(0.0462)
Martin Luther King, Jr.	-0.105	-0.105	-0.105	-0.119
	(0.0750)	(0.0753)	(0.0780)	(0.0817)
Presidents	-0.0302	-0.0225	-0.0244	-0.0188
	(0.0331)	(0.0335)	(0.0343)	(0.0353)
Memorial	-0.261***	-0.256***	-0.259***	-0.266***
	(0.0386)	(0.0391)	(0.0406)	(0.0416)
Independence	-0.200***	-0.206***	-0.202***	-0.204***
	(0.0561)	(0.0560)	(0.0561)	(0.0565)
Observed Independence	-0.218***	-0.216***	-0.213***	-0.210***
	(0.0618)	(0.0617)	(0.0613)	(0.0613)
Labor	-0.345***	-0.347***	-0.340***	-0.345***
	(0.0560)	(0.0561)	(0.0565)	(0.0568)
Columbus	0.0259	0.0289	0.0227	0.0277
	(0.0180)	(0.0184)	(0.0191)	(0.0198)
Veterans	-0.128	-0.127	-0.124	-0.134
	(0.0840)	(0.0842)	(0.0859)	(0.0909)
Observed Veterans	0.109	0.110	0.114	0.113
	(0.0707)	(0.0708)	(0.0719)	(0.0764)
Thanksgiving	-0.161***	-0.160***	-0.147***	-0.163***
	(0.0146)	(0.0150)	(0.0191)	(0.0223)
Monday	-0.0412***	-0.0396***	-0.0405***	-0.0436***
	(0.00738)	(0.00798)	(0.00890)	(0.00929)
Tuesday	-0.0144**	-0.0131*	-0.0153*	-0.0193**
	(0.00686)	(0.00739)	(0.00806)	(0.00844)
Thursday	0.0292***	0.0287***	0.0277***	0.0249***
	(0.00662)	(0.00747)	(0.00775)	(0.00832)
Friday	0.0870***	0.0844***	0.0835***	0.0762***
	(0.00727)	(0.00821)	(0.00856)	(0.00938)
Saturday	-0.0715***	-0.0693***	-0.0691***	-0.0694***
	(0.00821)	(0.00900)	(0.00917)	(0.00980)
Sunday	-0.220***	-0.218***	-0.211***	-0.212***
	(0.00782)	(0.00832)	(0.00871)	(0.00897)
Observations	987,628	987,628	987,628	987,628
R-squared	0.894	0.894	0.894	0.894
Sensor-by-Year FE	X	X	X	X
Day-of-Year Polynomial	X	X	X	X
2 Days Before / After Holidays	X			
5 Days Before / After Holidays		X		
10 Days Before / After Holidays				X
8 Days Before / After Holidays			X	

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

All specifications include fixed effects from the primary specification. The columns vary the number of days estimated before and after each holiday.

Table C.9: Holiday and Weekend Vehicle Travel Effects - Polynomial Sensitivity

	(1) 9th order	(2) 11th order	(3) 13th order	(4) 15th order
Christmas	-0.169***	-0.153***	-0.146***	-0.203***
	(0.0378)	(0.0390)	(0.0424)	(0.0464)
Observed Christmas	-0.128***	-0.128***	-0.128***	-0.128***
	(0.0315)	(0.0315)	(0.0315)	(0.0315)
New Year's	-0.154***	-0.176***	-0.164***	-0.138**
	(0.0496)	(0.0568)	(0.0603)	(0.0616)
Observed New Year's	-0.212***	-0.212***	-0.212***	-0.211***
	(0.0461)	(0.0461)	(0.0461)	(0.0461)
Martin Luther King, Jr.	-0.0965	-0.100	-0.0949	-0.105
	(0.0768)	(0.0777)	(0.0777)	(0.0780)
Presidents	-0.0316	-0.0306	-0.0338	-0.0244
	(0.0336)	(0.0338)	(0.0343)	(0.0343)
Memorial	-0.245***	-0.249***	-0.259***	-0.259***
	(0.0391)	(0.0393)	(0.0400)	(0.0406)
Independence	-0.225***	-0.221***	-0.211***	-0.202***
	(0.0552)	(0.0554)	(0.0559)	(0.0561)
Observed Independence	-0.213***	-0.213***	-0.213***	-0.213***
	(0.0613)	(0.0613)	(0.0613)	(0.0613)
Labor	-0.344***	-0.343***	-0.344***	-0.340***
	(0.0558)	(0.0560)	(0.0563)	(0.0565)
Columbus	0.0257	0.0281	0.0309*	0.0227
	(0.0187)	(0.0187)	(0.0188)	(0.0191)
Veterans	-0.122	-0.130	-0.138	-0.124
	(0.0854)	(0.0855)	(0.0857)	(0.0859)
Observed Veterans	0.114	0.114	0.114	0.114
	(0.0719)	(0.0719)	(0.0719)	(0.0719)
Thanksgiving	-0.159***	-0.166***	-0.172***	-0.147***
	(0.0151)	(0.0156)	(0.0171)	(0.0191)
Monday	-0.0410***	-0.0410***	-0.0409***	-0.0405***
	(0.00889)	(0.00889)	(0.00888)	(0.00890)
Tuesday	-0.0158**	-0.0159**	-0.0158**	-0.0153*
	(0.00804)	(0.00805)	(0.00803)	(0.00806)
Thursday	0.0278***	0.0277***	0.0278***	0.0277***
	(0.00778)	(0.00778)	(0.00776)	(0.00775)
Friday	0.0837***	0.0836***	0.0836***	0.0835***
	(0.00857)	(0.00857)	(0.00858)	(0.00856)
Saturday	-0.0693***	-0.0695***	-0.0695***	-0.0691***
	(0.00917)	(0.00918)	(0.00917)	(0.00917)
Sunday	-0.211***	-0.211***	-0.211***	-0.211***
	(0.00871)	(0.00871)	(0.00871)	(0.00871)
Observations	987,628	987,628	987,628	987,628
R-squared	0.894	0.894	0.894	0.894
Sensor-by-Year FE	X	X	X	X
Day-of-Year Polynomial	X	X	X	X
8 Days Before / After Holidays	X	X	X	X

Robust standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

All specifications include fixed effects from the primary specification. The columns vary the order of polynomial that is used to remove seasonal variation.