NBER WORKING PAPER SERIES

FORMATIVE EXPERIENCES AND THE PRICE OF GASOLINE

Christopher Severen
Arthur van Benthem

Working Paper 26091
http://www.nber.org/papers/w26091

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2019

The authors declare that they have no relevant or material financial interests or any other conflicts of interest that relate to the research described in this paper. We are grateful to Ken Gillingham, Gabriel Kreindler, Jeff Lin, Kyle Meng, Alejandro Molnar, Christopher Palmer, Andrew Plantinga, Steven Puller, Kate Vyborny, and Corey White for their comments and suggestions, and to PJ Elliott for excellent research assistance. We also thank seminar participants at the 2018 Urban Economics Association conference, the 2019 Association of Environmental and Resource Economists conference, UC Berkeley, and the University of Pennsylvania. This paper represents research that is being circulated for discussion purposes. The views expressed in this paper are solely those of the authors and do not necessarily reflect those of the Federal Reserve Bank of Philadelphia, the Federal Reserve System, or the National Bureau of Economic Research. All errors or omissions are the responsibility of the authors.

Formative Experiences and the Price of Gasoline
Christopher Severen and Arthur van Benthem
NBER Working Paper No. 26091
July 2019
JEL No. D12,D90,L91,Q41,R41

ABSTRACT

An individual's initial experiences with a common good, such as gasoline, can shape their behavior for decades. We first show that the 1979 oil crisis had a persistent negative effect on the likelihood that individuals that came of driving age during this time drove to work in the year 2000 (i.e., in their mid 30s). The effect is stronger for those with lower incomes and those in cities. Combining data on many cohorts, we then show that large increases in gasoline prices between the ages of 15 and 18 significantly reduce both (i) the likelihood of driving a private automobile to work and (ii) total annual vehicle miles traveled later in life, while also increasing public transit use. Differences in driver license age requirements generate additional variation in the formative window. These effects cannot be explained by contemporaneous income and do not appear to be only due to increased costs from delayed driving skill acquisition. Instead, they seem to reflect the formation of preferences for driving or persistent changes in the perceived costs of driving.

Christopher Severen
Research Department
Federal Reserve Bank of Philadelphia
Ten Independence Mall
Philadelphia, PA 19106
chris.severen@gmail.com

Arthur van Benthem
The Wharton School
University of Pennsylvania
1354 Steinberg Hall - Dietrich Hall
3620 Locust Walk
Philadelphia, PA 19104
and NBER
arthurv@wharton.upenn.edu

1 Introduction

Behavior is often modeled as an economic choice that depends on preferences and the current economic environment, yet observable characteristics of agents and of the economic environment often fail to explain significant variation in behavior. In many consumer choice models, remaining variation is cast as idiosyncratic preference for particular goods or outcomes (e.g., Manski and McFadden 1981; Berry, Levinsohn, and Pakes 1995). Where do these observably idiosyncratic tastes come from? A growing literature in economics argues that cumulative and recent experiences exert a significant influence on preferences (Bronnenberg, Dubé, and Gentzkow 2012; Malmendier and Nagel 2011; Simonsohn 2006).

We add to this narrative by suggesting that 'formative experiences' with an economic activity shape individual behavior for decades to come. We document a striking fact: commuters in the United States who experience a large positive shock to the price of gasoline while coming of driving age—and thus having their first experiences with driving—are less likely to drive to work in a private automobile later in life. Encounters with gasoline prices when first entering a driving environment have path-dependent effects on transportation choices decades later. In many economic models, such path dependence would manifest itself as individual preference heterogeneity.

We first show that individuals who experienced the 1979 oil crisis during their formative driving years (around age 16) are 0.2-0.5 percentage points less likely than preceding cohorts to drive to work 20 years later (at the time of the 2000 census). Because gasoline prices increased dramatically during the oil crisis, we use an event study (or regression discontinuity-in-time) framework to identify this effect. The effect is stronger in urban settings with plausible alternative transportation options, and for lower-income workers. We also find that the decrease in driving is counteracted by an increase in public transit usage.

We then expand the scope of the analysis to carefully identify the effect of gasoline price movements during formative years on two aspects of driving behavior: mode of commuting (an extensive margin) and miles traveled (an intensive margin). We study many cohorts across many cross sections of the U.S. population and exploit variation in gasoline prices across states and time over a long horizon to flexibly control for time-invariant differences across locations, life-cycle (age) effects, and contemporaneous factors that influence transportation choices (e.g., current gasoline prices). Individuals respond to price changes during their formative driving years much more so than to price levels. A doubling of the real price of gasoline between the ages of 15 and 17 leads to a 0.3-0.4 percentage point reduction in the probability of driving to work later in life and a 0.2-0.3 percentage point increase in transit usage; this corresponds well to the estimated effect of the 1979 oil crisis. This extensive margin effect is large when workers are young (between 25 and 34), but is also present in older workers. There is some evidence of a smaller effect on household access to a vehicle, indicating that durable goods consumption may respond to price shocks far in the past. Effects are specific to price movements during this formative period and cannot be explained by channels such as income, educational attainment, or marital status, suggesting that preferences are indeed at work.

Drivers that experience a doubling of real gasoline prices between ages 15 and 17 also drive 3.4-8.2 percent fewer annual miles as adults. Using a framework similar to the extensive margin analysis, we combine five waves of national travel survey data to determine whether there is a later-life, intensive-margin effect of gasoline price movements during formative years. This effect corresponds to roughly 900-1,100 fewer average annual miles traveled for drivers in affected cohorts, conditional on having access to a vehicle. Furthermore, we find that drivers that were exposed to gas price hikes early in life are somewhat less likely to own fuel-inefficient light-duty trucks.

Supplementary analysis strengthens these results. First, it is precisely the effect of gasoline price shocks between ages 15 and 18 that are acting upon later-life behavior—gas price movements earlier or later do not matter. We also allow the formative window to vary according to state-specific driving age restrictions. An effect is present from 1 year before to 2 years after the minimum age at which teenagers can obtain a full-privilege license, which covers ages 15 to 18 in many states. Variation in these restrictions across states and over time strengthen identification and lead to slightly larger effects of gasoline price movements during formative driving years.

We rule out a primary role for two obvious explanatory mechanisms: income and costly skill acquisition. Gasoline price movements (and the 1979 oil crisis in particular) are associated with recessions, and entering the labor market during a recession can decrease permanent income and, therefore, change later-life driving behavior. Our results are robust to including contemporaneous income as a control, and more formal analysis reveals that income mediates at most 24% of the observed effect, but likely much less. We also study the effects of changes in minimum driving age to determine whether restrictions on learning to drive during formative years discourage take up of driving; it does not appear to. Nor does the 1979 oil crisis lead to a delay in driver license take up at the national level. Further, delayed skill acquisition is unlikely to lead to the intensive-margin effect. Together, these results indicate that driving behavior is imprinted by formative experiences with gasoline prices, either through a change in preferences or a change in the perceived cost of driving and volatility of gas prices.

Our paper contributes to several literatures in economics. A growing empirical literature connects cumulative and recent experiences to preference formation and behavior.[1] Recent experience of violence or cumulative experience of poor economic conditions can impact risk attitudes later in life (Callen et al. 2014; Malmendier and Nagel 2011).[2] Malmendier and Shen (2018) show that exposure to severe unemployment spells leads to consumption deviations inconsistent with the permanent income hypothesis, and Fujiwara, Meng, and Vogl (2016) find evidence that rainy election days decrease voter turnout in future elections. Another strand of the literature links later-life purchasing decisions of migrants to brand availability or relative prices in places of origin (Bronnenberg, Dubé, and Gentzkow 2012; Logan and Rhode 2010). Anderson et al. (2015) document an interesting source of brand loyalty by showing that preferences for particular automobile brands are transmitted across generations.

Our paper contributes to these findings in several ways. First, we show that experiences during a relatively narrow window in which agents are first interacting with a good, and thus forming initial impressions and information sets, have long-run consequences for behavior. The lack of an effect of gasoline price shocks after this formative window suggests initial contact may be more important than cumulative experience for some behaviors. We also demonstrate that relatively mundane experiences (interactions with gasoline prices) can be important; formative experiences need not be life-changing or extreme. Finally, our results show that price volatility (rather than price level) can imprint later behavior.

A smaller literature studies the behavioral determinants of location choice, commuting, and vehicle choice.[3] One line of research focuses on the short-run impact of recent

[1]Theoretical foundations for 'habit formation' go back to seminal work by Pollak (1970). With 'habit-forming' goods, an individual's current preferences depend on past consumption levels to which s/he has become accustomed, and thus past prices.

[2]Related work focuses on the impact of childhood experiences on later-life outcomes. Giuliano and Spilimbergo (2013) show that growing up during a recession promotes redistributionary preferences, while Brown, Cookson, and Heimer (2019) find that children growing up in financially underserved Native American reservations later make less use of financial products and have lower credit scores than children growing up in better-served reservations.

[3]There is a large literature on path dependence based on the supply of transportation infrastructure—it influences the location of cities (Davis and Weinstein 2008; Bleakley and Lin 2012; Michaels and Rauch

experiences. Busse et al. (2015) find that weather at the time of purchase influences vehicle choice, while Simonsohn (2006) and Simonsohn and Loewenstein (2006) show that people maintain reference dependence with respect to their former city of residence regarding housing prices and commuting when moving to a new city. Another set of papers finds short-run benefits of forced experimentation, studying the effects of incentivizing households to use public transit shortly after they move to a new location (Bamberg 2006; Yang and Lim 2017) or of commuting behavior after being forced to travel along new routes (Larcom, Rauch, and Willems 2017). These papers find some short- or medium-run path dependence and rely on the experience of recent events. In contrast, our results show persistence over the life cycle from the price variation in a three-year period.

We also speak to a literature in urban and environmental economics that seeks to understand the short-run relationship between gasoline prices and driving behavior. Much of this literature was spawned by volatile oil prices in the 1970s (Espey 1998). A part of this literature studies how fleet composition responds to changes in gasoline prices (Puller and Greening 1999) or efficiency mandates (West et al. 2017), and whether this is consistent with standard models of rational expectations about gasoline prices and consumer valuation of future fuel costs (Allcott and Wozny 2014; Busse, Knittel, and Zettelmeyer 2013; Gillingham, Houde, and Benthem 2019; Jacobsen and van Benthem 2015; Li, Timmins, and von Haefen 2009). Hughes, Knittel, and Sperling (2008) show that gasoline usage has become more inelastic in the twenty-first century; Small and van Dender (2007) attribute this to rising incomes.[4] Gillingham, Jenn, and Azevedo (2015) explore heterogeneity in fuel price elasticities by geography and the fuel economy and age of the vehicle. However, there is little evidence on how differences in behavior and preferences arise. Our paper documents a new, long-run response to macroeconomic energy price shocks which uncovers one channel that gives rise to heterogeneous driving behavior.

The next section describes relevant information about the research setting and the data sources that we use. We then analyze how the 1979 oil crisis continues to influence driving patterns into the twenty-first century in an event-study framework in Section 3. Section 4 uses a more comprehensive empirical strategy to quantify the long-run effects of gasoline prices across many cohorts of drivers. This section also separately estimates intensive and

2018), urban form (Brooks and Lutz 2019), and regional growth (Donaldson and Hornbeck 2016).

[4]Recent papers investigate if millennials have different preferences for driving than the previous generation (Klein and Smart 2017; Knittel and Murphy 2019; Leard, Linn, and Munnings 2019). These papers find that vehicle ownership and driving patterns are relatively similar once income and demographics are accounted for. In contrast, our paper documents a source for preference heterogeneity that persists over the life cycle and does not depend on income.

extensive margin effects. We then examine the effects of changes in the minimum age of full and provisional licensing of young drivers in Section 5. Section 6 offers conclusions and implications.

2 Context and Data

2.1 Context: Driving in the United States

The United States is a notably automobile-friendly nation: about 76 percent of workers commute alone in a private vehicle (85 percent including carpoolers), compared with 56 percent (64 percent) in the United Kingdom. Laws regulating driving tend to provide few barriers, and people start driving at younger ages than in most other nations. In 30 U.S. states, it is still possible to obtain an unrestricted (full-privilege) driver's license before the age of 18, the standard minimum unrestricted age in most of Europe. In 1980, only in seven states was the minimum full-privilege driving age greater than 16, and in five states it was less than 16. Learner's permits have traditionally been granted between the ages of 14 and 16; in 1980, eight states allowed those 14 years old to begin supervised driving. Since 1970, more than half of states have permitted those under 16 to begin supervised driving (see Appendix Table A.3 for details).

Many teenagers begin driving soon after obtaining the minimum legal age. In 1980, roughly 44 percent of those aged 16, 66 percent of those aged 17, and 77 percent of those aged 18 had a driver's license. These numbers have been falling as states started implementing graduated licensing programs that delay full-privilege licenses until age 17 or 18. By 2010, only 28 percent of those aged 16, 46 percent of those aged 17, and 61 percent of those aged 18 were licensed.[5]

This age distribution matters: there are several reasons to believe it is substantially easier to learn to drive when young than when older (at least in the United States). First, young people often have access to vehicles, as well as training and supervision, while living at home. It is the norm in most non-urban and many urban communities to learn to drive during one's teenage years. Further, many high schools have traditionally offered subsidized driver training programs. Finally, the opportunity cost of time for this age group is likely lower than for older people.

[5] Data from the 2016 release of the Highway Statistics published by the Federal Highway Administration, Table DL-220; see Appendix A.1 for details.

2.2 Data

We draw our primary data from several sources, and discuss each below. See Appendix A.1 for additional data details.

Commuting Behavior and Vehicle Ownership. The decennial census asks questions about commuting mode and time. These 'Journey to Work' questions appear in the 1980, 1990, and 2000 censuses, and more recently in the American Community Survey (ACS). We use data from these three censuses, as well as the 2006/10, 2011/15, 2016, and 2017 ACS. Key variables of interest are: (i) the primary mode of commute for each worker in the household; (ii) whether a household keeps a vehicle at home for use by members of the household; and (iii) public transit ridership for each worker in the household. The census data include rich demographic and economic controls.

Age plays a central role in our analysis, but interpreting age in census data requires qualification. Census microdata report both age and birth year for each person in the household. Age is understood in terms of a particular *reference day*, which is April 1 of the enumeration year. Birth year is defined as sample year less age. Thus, someone born in May reporting 36 years of age in the 2000 census was born in 1963, whereas someone born in March reporting an age of 36 in 2000 was born in 1964. We use birth year to define cohorts at age 16, recognizing that there is some spillover across years. The ACS is conducted on a rolling basis, meaning that there is no constant reference day. Although the sampling year is reported in the multi-year ACS data, it is difficult to precisely recover the birth year because the sampling date is not reported. However, in both census and ACS data, errors should be consistent from year to year.

Travel Surveys and Vehicle Data. We draw on five waves of the National Household Travel Survey (NHTS) and its predecessors from 1990, 1995, 2001, 2009, and 2017. Our main variable of interest is miles driven in each vehicle, combined with information on which household member is the main driver of each vehicle. We aggregate vehicles across primary drivers to develop a person-specific measure of annual miles traveled. There are fewer demographic details available in these data; we use sex, race, urban/rural status, and family size. The data also contain vehicle-level information on make, model, and vintage which allows us to identify a vehicle's type (passenger car or light-duty truck). We merge make, model, and vintage information with EPA data on fuel economy from Allcott and Knittel (2019).

Gasoline Prices. The Energy Information Administration reports nominal tax-inclusive state-level gasoline price data starting in 1983. For the years 1966-1982, we use data from annual Federal Highway Administration (FHWA) Highway Statistics publications (Small and van Dender 2007; Li, Linn, and Muehlegger 2014).[6] Appendix Figure A.1 plots these data.

Driver Licensing. We develop a database of driver licensing requirements by drawing from several sources, including the FHWA's publication Driver License Administration Requirements and Fees, records from the Insurance Institute of Highway Safety, records from several state Departments of Motor Vehicles, and newspapers (see Appendix A.1 for more details). We also use aggregate data on driver licensing published by the FHWA in Highway Statistics that list the number of driver's licenses held by people of each age from 16 to 24 in each year. To estimate rates of driver license adoption, we construct age-specific population estimates from the National Cancer Institute's SEER Population database.

3 The 1979 Oil Crisis and Later-Life Driving

The price of gasoline in the United States was relatively stable over the 1950s and 1960s. Beginning in the 1970s, the global oil market entered a phase of increased price volatility. The United States experienced a shock to gasoline prices in late 1973 and 1974, in response to oil production limits by OPEC. However, a larger price shock was in store. Iranian oil production fell 94 percent in late 1978 during the Iranian Revolution, and the United States began to experience gasoline shortages and rising prices in early 1979. Prices continued to rise until the early 1980s, at which point they briefly leveled off (in nominal terms). The Iran-Iraq war and removal of U.S. price controls by the Reagan administration in late 1980 and early 1981 led to another increase in prices. Nominal prices stayed at this unprecedented level through 1986, though high inflation devalued the real price.[7] Figure 1 highlights the period of dramatically increasing gasoline prices.

The high gasoline prices of the early 1980s are notable for three reasons. First, the increase was large in magnitude, sudden and well-publicized in the media.[8] The nominal

[6] We thank Erich Muehlegger for sharing these data with us.

[7] See Hamilton (1985) for a discussion of the proximate causes of oil shocks.

[8] See, for instance, `https://books.google.com/ngrams/graph?content=gas+price&year_start=1960&year_end=2000&corpus=15&smoothing=0`.

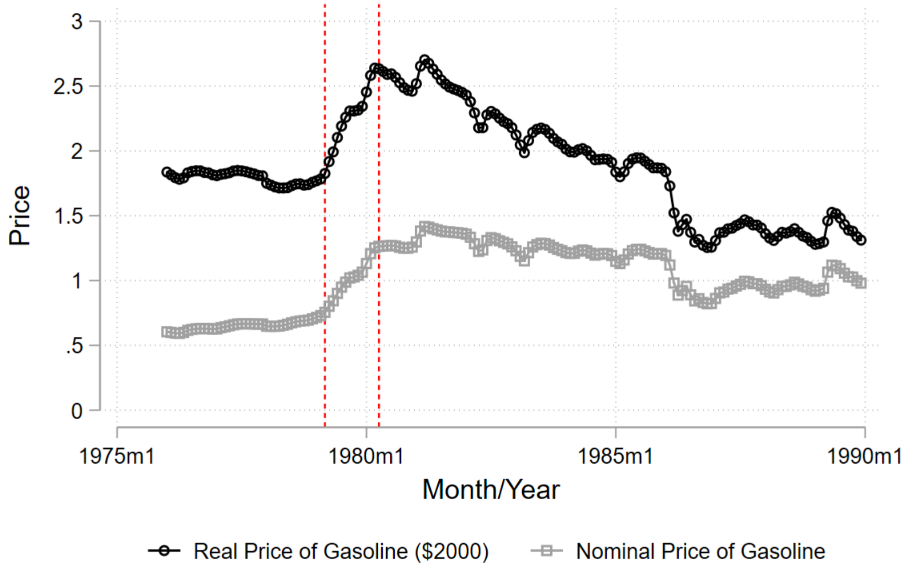

Figure 1: Gasoline prices from 1975 to 1990.

price doubled over the course of a year. This increase was unexpected, and was exacerbated by unpredictable demand-side responses.[9] In fact, average consumer beliefs are often best reflected by a no-change forecast, so shocks to prices can be modeled as unexpected (Anderson, Kellogg, and Sallee 2013). Second, nominal prices had never been so high, and real prices had not seen such levels since the 1930s. This was the first time since the U.S. became an automobile-dependent society that nominal gasoline prices exceeded $1 per gallon and real gasoline prices were higher than $3 per gallon (in 2015 dollars). The $1 price level may have been particularly salient. Third, not only was the cost of gasoline high, queuing at the gasoline pump meant that an additional time expenditure was required to obtain gasoline. These queues could be quite substantial: Frech and Lee (1987) and Deacon and Sonstelie (1989) highlight the negative consequences of time wasted by queuing; Deacon and Sonstelie (1985) use this event to estimate drivers' value of time. High prices led to a reduction in driving during the early 1980s. Total vehicle miles traveled had been increasing by 4-4.5 percent annually from 1976 to 1979, but then declined

[9] While the traditional explanation for the 1979 oil crisis was a supply-side shock, a more recent economic literature mainly attributes the price spike to two types of demand shocks—increased inventory demand in anticipation of future shortages triggered by widespread panic about the political unrest in Iran, and 'traditional' demand shocks reflecting a strong global economy (Baumeister and Kilian 2016). Such demand shocks are unpredictable by nature.

over the next two years before slowly recovering.

The price shock was likely most consequential for those just coming of driving age. There are two consequences of this price episode that may be related to long-run driving behavior: (i) the perceptions of the price of driving may have changed[10] and (ii) learning to drive may have become more expensive. Both factors could plausibly lead to a reduction in driving later in life. If particular cohorts perceive driving as more costly (even though everyone faces the same contemporaneous prices)—for example, because gasoline prices are particularly salient when deciding to take driving lessons and learning how to drive—they will be less likely to drive and, if they drive, travel fewer miles. Given that most drivers in the U.S. learn to drive before the age of 18, those who fail to learn to drive early in life may face more difficulties and higher opportunity costs learning to drive later in life and even forego driving altogether.

3.1 Later-Life Driving

Figure 2 plots driving, public transit use, and household vehicle access more than two decades after the 1979 oil crisis, as reported in the 2000 census. Commuters are matched to cohorts at age 16 (displayed along the horizontal axis). For example, the behavior of those born in 1964 in 2000 (at age 36) is indexed to the year 1980 (when they turned 16), while the behavior of those born in 1968 in 2000 (at age 32) is indexed to cohort year 1984. Vertical bars bound the period of most rapid increases in gasoline prices shown in Figure 1 (prices continued to climb for another year afterwards).

Driving behavior changes for those who came of age after 1980. The probability that a commuter drives to work decreases and the probability that a commuter takes mass transit increases. The decrease appears to be slightly less than one-half of a percentage point, and marks a jump in behavior from cohorts turning 16 after 1980, and those that came before. Furthermore, the bottom panel illustrates that individuals in these cohorts are less likely to have access to a vehicle.[11]

Treatment is not precise: gasoline prices increased over the course of 1979-80, and then bumped up further in 1981. We do not have a strong prior for the precise age at which gasoline prices become salient; in most states, learner's permits could be awarded at 15

[10] Media coverage of the oil crisis was ubiquitous. Even though some teens may not have paid directly for gasoline or driving, media coverage and family interactions discussing gasoline price fluctuations likely meant that prices were salient and played a role shaping perceptions and expectations.

[11] For vehicle access, the sample is not limited to workers.

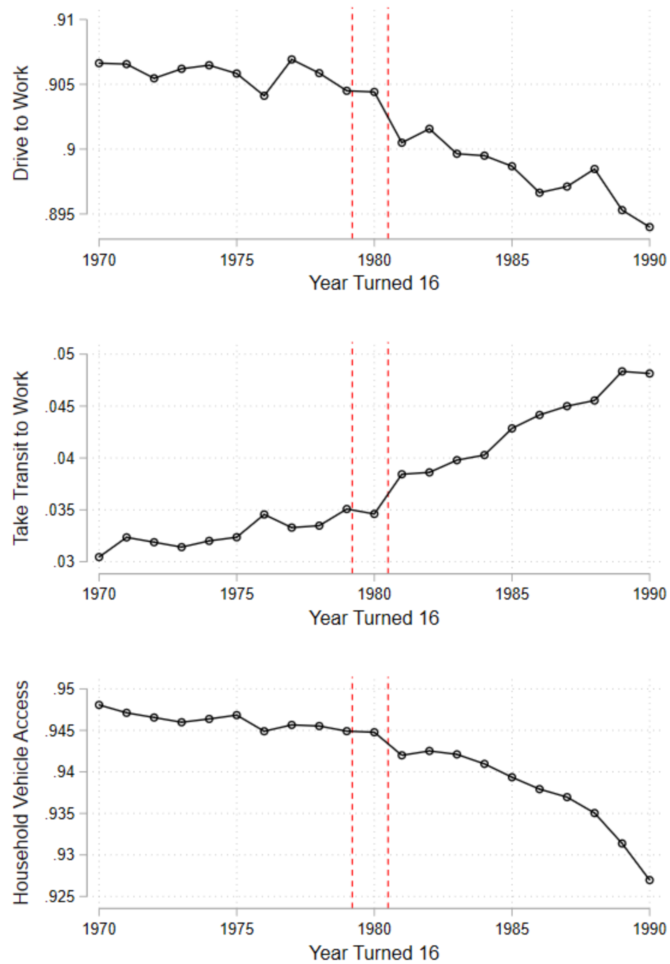

Figure 2: Commuting behavior and vehicle ownership by cohort in 2000.

in 1980. Nonetheless, it is clear from Figure 2 that a break, if it exists, most likely occurs between the 1980 and 1981 cohorts.[12] We therefore turn to an event study framework to quantify the size of this break. (The empirical design resembles a 'regression discontinuity in time' with birth year as the running variable; we refer to this setup as 'RD' for simplicity). We discuss under what conditions these RD estimates below have a causal interpretation, and adopt a more comprehensive repeated cross-section approach that lets us characterize which ages are formative in Section 4.

[12]Note that the time series are trending upward or downward in 'year turned 16', reflecting that later cohorts are younger during the 2000 census and transportation behavior more generally exhibits life-cycle trends. The analysis in Section 4 controls for such life-cycle patterns.

Specifically, we quantify the break by estimating variants of the following equation:

$$Y_i = \alpha + g(S_i) + \tau D_i + X_i'\lambda + \varepsilon_i \tag{1}$$

where Y_i is an outcome of interest for individual i in the 2000 census, S_i is the year that i turned 16, and X_i are other characteristics of i. Treatment is the binary variable D_i, which is equal to one if i turned 16 after 1980. The function $g(\cdot)$ captures trends in driving behavior; we experiment with linear and quadratic functions that are allowed different slopes before and after 1980.[13] Data are limited to a symmetric bandwidth around the treatment year.[14]

Panel A of Table 1 presents RD estimates of τ using linear and quadratic trends. Estimates with linear trends are shown over a bandwidth of two to ten years, while those with quadratic trends are shown over a bandwidth of five to ten years. Results indicate a sharp decrease in the likelihood of driving to work of 0.2 to 0.5 percentage points that persists roughly twenty years after turning 16. The quadratic results are less precise, but less prone to bias by accommodating more response curvature along the running variable. Point estimates are relatively similar across both linear and quadratic specifications.

The relationship between the gasoline price jump in 1979-80 and reduced later-life driving behavior of cohorts coming of age after 1980 can be assigned a causal interpretation if no other observable or unobservable confounding factors experience a discontinuous break at the same point in time. We demonstrate that the observable covariates are smooth in Appendix Figures A.2, A.3, and A.4 across a range of demographic, employment, and housing characteristics. There are no obvious discontinuities in these graphs, though some display more curvature than those in Figure 2. We also report results from 'donut' regression discontinuity tests that omit the 1980 cohort in Appendix Table A.1 and alleviate measurement error due to the gradual change of prices throughout the years 1979-80. Results are similar in magnitude, though slightly less significant.

[13] We report heteroskedasticity-robust standard errors throughout this section. Kolesár and Rothe (2018) caution against clustering standard errors by the running variable, and present simulation evidence that shows the heteroskedasticity-robust standard errors outperform standard errors clustered by the running variable with small or moderate window widths. Furthermore, clustering on the annual running variable here would lead to a few-clusters problem.

[14] We define the treatment time as just after 1980. Thus, a bandwidth of two includes cohorts that turn 16 in 1979, 1980, 1981, and 1982.

Table 1: Discontinuity in turning 16 after 1980 on commuting behavior in 2000.

Model	Poly. order	Bandwidth (years)								
		2	3	4	5	6	7	8	9	10
Panel A: Effect on driving, no controls										
	1	-0.0050*	-0.0029+	-0.0026+	-0.0032**	-0.0026*	-0.0027**	-0.0032**	-0.0032**	-0.0029**
		(0.0022)	(0.0016)	(0.0014)	(0.0012)	(0.0011)	(0.0010)	(0.0009)	(0.0009)	(0.0008)
	2				-0.0033	-0.0039*	-0.0032+	-0.0021	-0.0027+	-0.0032*
					(0.0022)	(0.0019)	(0.0016)	(0.0015)	(0.0014)	(0.0013)
Panel B: Effect on driving, controls:										
+ demographics	1	-0.0046*	-0.0025	-0.0023+	-0.0029*	-0.0025*	-0.0024*	-0.0028**	-0.0026**	-0.0021*
		(0.0022)	(0.0016)	(0.0014)	(0.0012)	(0.0011)	(0.0010)	(0.0009)	(0.0009)	(0.0008)
	2				-0.0028	-0.0035+	-0.0030+	-0.0020	-0.0026+	-0.0034**
					(0.0022)	(0.0018)	(0.0016)	(0.0015)	(0.0014)	(0.0013)
Panel C: Effect on driving, controls:										
+ demographics, state of birth FEs	1	-0.0046*	-0.0023	-0.0019	-0.0025*	-0.0020+	-0.0019+	-0.0022*	-0.0020*	-0.0014+
		(0.0022)	(0.0016)	(0.0013)	(0.0012)	(0.0011)	(0.0010)	(0.0009)	(0.0009)	(0.0008)
	2				-0.0027	-0.0031+	-0.0027+	-0.0019	-0.0024+	-0.0030*
					(0.0021)	(0.0018)	(0.0016)	(0.0015)	(0.0014)	(0.0013)
Panel D: Effect on driving, controls:										
+ demographics, state of birth FEs	1	-0.0046*	-0.0022	-0.0018	-0.0024*	-0.0019+	-0.0017+	-0.0021*	-0.0019*	-0.0013
+ $\ln(income)$		(0.0022)	(0.0016)	(0.0013)	(0.0012)	(0.0011)	(0.0010)	(0.0009)	(0.0009)	(0.0008)
	2				-0.0027	-0.0030+	-0.0026	-0.0018	-0.0023	-0.0029*
					(0.0021)	(0.0018)	(0.0016)	(0.0015)	(0.0014)	(0.0013)
N		545k	811k	1075k	1343k	1614k	1888k	2148k	2398k	2642k

Regression discontinuity estimates of the effect of turning 16 after 1980 on a binary indicator of whether the respondent drove to work, as reported in the 2000 Census. Bandwidth is symmetric around 1980.5. Sample includes all native-born persons actively working in the Census, and excludes farm workers and those coded N/A for transportation mode. Demographic controls include sex, race, and educational attainment. Observations weighted by person sample weights. Standard errors are robust to heteroskedasticity (see text). Sample sizes are 1-2% smaller in panels B through D. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$.

The effect cannot be explained away by controlling for observable, contemporaneous characteristics. Panels B through D of Table 1 progressively add more controls to the specification in Equation (1). Panel B adds demographic controls we take as exogenous (sex and race), as well as educational attainment (which could be endogenous). Panel C adds state of birth fixed effects to control for differential commuting behavior in different places. We include state of birth, rather than state of residence, because it is exogenous with respect to later-life commuting decisions. Panel D adds contemporaneous income, but we recognize this may not be an appropriate control if later-life income is both influenced by graduating high school into a recession and income influences vehicle purchasing (see Section 4.1.2 and Appendix A.2 for more discussion).

These covariates decrease point estimates by about a quarter, but do not completely explain behavior. State of birth plays an important role, but estimates are still significant after accounting for differences across locations. Contemporaneous income also influences estimates, but the effect is still present in many specifications. This suggests that there are persistent effects of gasoline prices while coming of age that cannot be explained by earnings.

It is apparent from the preceding analysis that cohorts turning 16 after 1980 exhibit different driving behavior later in life. This cannot be due to gasoline prices immediately leading up to 2000: everyone faced the same price profile in the preceding years. The analysis centers on adults 36 years old, so life-cycle trends have mostly smoothed out. Further, there is a logical link between the high gas prices and reduced driving. However, this research design ultimately confirms a gasoline price effect only for a single cohort. We therefore approach this question with a different and more comprehensive repeated cross-section research design in Section 4, but first we examine other, related outcome variables and the responses of different subgroups to the 1979 oil crisis. The intuitive responses we find add some credibility to the RD results presented above.

3.2 Transit and Vehicle Accessibility

The negative effect on driving is largely compensated by an increase in transit use, as shown by Panel A of Table 2. Those coming of age just after 1980 are 0.2-0.4 percentage points more likely to take transit to work than their counterparts coming of age a bit earlier. The absolute magnitude of this effect is between 50 and 100 percent of the effect size in Panel A of Table 1, suggesting that transit is the primary substitute for driving (relative to working at home, carpooling, or self-powered means).

Consistent with the effects on driving to work and public transit, those coming of age after 1980 are also less likely to have access to a vehicle. Panel B shows RD estimates of Equation (1) on vehicle access for all prime-age adults (not just workers). Linear results are dubious at larger bandwidths, as Panel C of Figure 2 shows greater curvature in vehicle access for cohorts coming of age in the late 1980s. The estimates from the quadratic specifications, 0.2-0.3pp, are generally in line with the transit results. Those coming of age after 1980 are less likely to drive to work, more likely to take transit, and less likely to have access to a private vehicle.

Table 2: Discontinuity in turning 16 after 1980 on transit usage and vehicle access in 2000.

	Bandwidth (years)								
Poly. order	2	3	4	5	6	7	8	9	10
Panel A: Transit usage									
1	0.0036*	0.0027*	0.0027**	0.0023**	0.0017*	0.0016*	0.0016**	0.0015**	0.0018**
	(0.0015)	(0.0011)	(0.0009)	(0.0008)	(0.0007)	(0.0007)	(0.0006)	(0.0006)	(0.0005)
2				0.0038**	0.0037**	0.0030**	0.0023*	0.0024**	0.0018*
				(0.0014)	(0.0012)	(0.0011)	(0.0010)	(0.0009)	(0.0009)
N	545k	811k	1075k	1343k	1614k	1888k	2148k	2398k	2642k
Panel B: No vehicle access									
1	0.0033*	0.0026*	0.0020*	0.0016+	0.0009	0.0007	0.0005	-0.0002	-0.0012*
	(0.0016)	(0.0011)	(0.0010)	(0.0008)	(0.0008)	(0.0007)	(0.0007)	(0.0006)	(0.0006)
2				0.0037*	0.0034**	0.0027*	0.0023*	0.0028**	0.0034**
				(0.0015)	(0.0013)	(0.0012)	(0.0011)	(0.0010)	(0.0009)
N	698k	1038k	1376k	1717k	2061k	2409k	2739k	3058k	3370k

Regression discontinuity estimates of the effect of turning 16 after 1980 on a binary indicator of transit usage or vehicle access as reported in the 2000 Census. Bandwidth is symmetric around 1980.5. Panel A includes all native-born persons actively working in the Census, and excludes farm workers and those coded N/A for transportation mode. Panel B includes non-workers. Observations weighted by person sample weights. Standard errors are robust to heteroskedasticity (see text). $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$.

3.3 Variation by Location and Income

Whether or not commuters are able to substitute away from the automobile depends on the choices available to them. Therefore, we expect effects to be stronger in urban settings where there are plausible alternatives to driving (public transit, walking, etc.). We first examine the RD effect for commuters who reside in the 'principal city' of an MSA.[15] The choice of location is potentially endogenous, however, so we interpret subgroup analysis on location as suggestive evidence. Estimates, shown in Appendix Table A.2, are larger than the effect in the whole population and are largely robust to bandwidth and trend specification. For urban dwellers, someone 36 or under is -0.6 to -1.9 percentage points less likely to drive to work than someone 37 or older in 2000. Conversely, there is little effect on workers who live outside of metropolitan areas, as Panel B reveals. Point estimates are small, mostly positive, and insignificant. Taken together with the results in Tables 1 and 2, these results suggest that the persistent effect of the 1979 oil price shock is largely concentrated in cities where viable transportation alternatives are available.

Panels C and D of Table A.2 report RD estimates for two other groups. Panel C limits the sample to black workers, and shows evidence of significant and negative effects. The linear specification loses significance at higher bandwidths; this is likely due to greater curvature in the running variable. Panel D limits the sample to workers without a college education. Results are smaller in magnitude and significance, and point estimates are generally smaller than those reported in Table 1.

Finally, Appendix Figure A.5 examines the effect of being in a post-1980 cohort on driving across the income distribution. We divide the population of commuters both into centile and into decile bins, and then run the RD estimator using the linear specification with a bandwidth of five years within each bin. Estimates for the lowest decile are negative (about -1.4 percentage points) and significant. The third decile is, unexpectedly, positive, but otherwise the first eight deciles are negative and significant. There is a positive or no effect for the two highest deciles. Estimates for each centile are smoothed and shown with a dotted line, and generally conform to the decile estimates.

Taken together, these results suggest that the post-1980 cohorts are different in ways that support a difference in transportation habits: lower-income workers, and workers in cities, show stronger effects. Further, transit use compensates the loss in driving.

4 Long-Run Driving Effects of Gasoline Prices

To more precisely link long-run behavior with formative exposure to gasoline prices, we use all available variation in gasoline prices across time and states with a fixed effects

[15] There are several MSAs for which principal city status may violate disclosure rules and therefore not reported in the 2000 census. This is why sample sizes are lower than in Table 1.

research design that uses repeated cross sections over nearly four decades. We tie the price of gasoline that someone likely experiences during their formative driving years to later-life driving behavior. We show results for different treatment age windows (such as between ages 15 and 17), and the years around a state's minimum driving age. We use all available public-use census/ACS microdata since 1980 on commuting mode to study the extensive margin of driving behavior, and all NHTS data since 1990 to study the intensive-margin response and vehicle choice.

Our primary specification models outcome Y_{icst} for person i in cohort c born in state s observed in sample year t as:

$$Y_{icst} = \theta T_{cs} + \kappa_s + \delta_t + \eta_a + X'_{it}\lambda + \varepsilon_{icst} \tag{2}$$

where the treatment variable T_{cs} is either the price of gasoline or change in the price of gasoline during formative driving years for cohort c in state s. θ is the parameter of interest and measures the response in the outcome variable to gasoline price exposure at early driving ages. Different states may exhibit different behavior on average due to different provision of infrastructure, social norms, etc.; state fixed effects, κ_s, capture these differences. Sample year fixed effects, δ_t, control for current gas prices, business-cycle trends in employment, etc.[16] We also add individual age fixed effects η_a to capture important life-cycle trends in transportation behavior. We include a vector of individual and household characteristics, X_{it}. We limit the sample to prime-age (25- to 54-year-old) native-born adults, and exclude residents of Alaska and Hawaii as we do not have a complete series of gasoline prices for these states. We discuss other sample adjustments that are particular to each dataset below. Standard errors are clustered by state.

Equation (2) is conceptually similar to a fixed effects difference-in-differences estimator. We observe cohorts, who have different formative experiences in the years around their initial driving age, at several points in time as they progress throughout the life cycle. Thus we are comparing, e.g., the driving behavior of a 36-year-old person in a state at particular time with other 36-year-old people in the same state at earlier or later points in time.

We experiment with different definitions of treatment based on exposure to gasoline prices during one's formative driving years. We use a data-driven approach to establish which ages are most formative, i.e. have the largest effects on later behavior. We consider

[16]We include state-by-sample year fixed effects in some specifications to flexibly control for local differences in these contemporaneous factors.

both absolute 'calendar' age and age relative to the minimum driving age in state s. Finally, we test if later-life driving behavior is best explained by the *level* of gasoline prices during formative driving years, or by *changes* in gasoline prices. Specifically, we let P_{cs}^{a} be the price of gasoline that a cohort c in state s experiences at age a, and $P_{cs}^{m_{cs}}$ the gasoline price that people in cohort c faced when they reached the minimum full-privilige driving age in state s. We show results using several definitions of T_{cs}:

$$(i)\ P_{cs}^{\Delta a,(a-h)} = \frac{P_{cs}^{a} - P_{cs}^{a-h}}{P_{cs}^{a-h}}, \quad (ii)\ P_{cs}^{a},$$

$$(iii)\ P_{cs}^{\Delta(m_{cs}+j),(m_{cs}-k)} = \frac{P_{cs}^{m_{cs}+j} - P_{cs}^{m_{cs}-k}}{P_{cs}^{m_{cs}-k}} \quad (iv)\ P_{cs}^{m_{cs}}.$$

where we choose different combinations of $h \in \{1,2\}$ and $j,k \in \{0,1,2\}$. Choice (i) gives the percentage change in price between age a and $a-h$ and choice (ii) gives the price at age a. Choice (iii) represents the percentage price change during a window of ages around the minimum driving age; choice (iv) gives the price at the minimum driving age. Results (in Tables 3 through 6 below) will reveal that gasoline price shocks over an age range between 15 to 18 (or between one year before and two years after the minimum driving age) matter much more than levels.

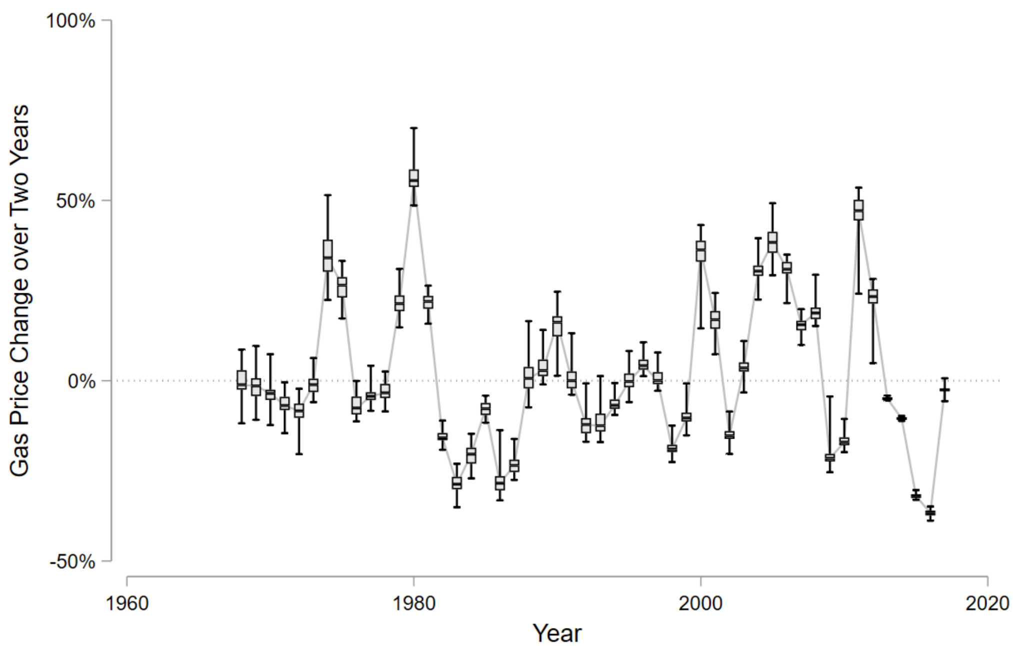

Figure 3: Box plot of 2-year lagged percentage changes in state-level gasoline prices.

Figure 3 plots gasoline price exposure variable (i) for $h = 2$, showing the price changes for each calendar year rather than by cohort-age. The boxes and whiskers in Figure 3 indicate the quartiles of variation across states. Appendix Figure A.1 adds plots for price exposure variables (i) with $h = 1$ and (ii). These figures showing annual and biennial gasoline price changes highlight the variation used to identify long-run effects. Our preferred specification uses $P_{cs}^{\Delta 17,15}$, as the window from 15 to 17 covers most driver licensing uptake (see Section 2 and Section 5). It is also the age range over which the treatment effect is the strongest (see Table 6).

Ideally, we would observe everyone's residential location in these formative years. However, census data only provide information on state of birth and current residence. We therefore define a sample of *stayers* who currently reside in their state of birth (about 64 percent of the full sample) as our primary population of analysis. However, we provide robustness analysis showing that our results are not sensitive to using the full sample (merged either on state of birth or state of current residence). NHTS data do not contain information about place of birth or prior migration decisions, so we merge on current state of residence.[17]

4.1 Extensive Margin

We first explicitly incorporate gasoline prices into the analysis of commuting mode. We merge data on driving behavior from all public use census/ACS microdata available between 1980 and 2017 with gasoline prices based on respondent age and birth state. In most specifications, we only include those who still reside in the state of their birth. Our state-level gasoline price data begins in 1966, so our sample includes those for whom we can calculate our primary definition of treatment: $P_{cs}^{\Delta 17,15}$. This corresponds to cohorts born between 1951 and 1992.

[17] We are concerned about incorrectly matching people to gasoline prices experienced earlier in life. In our data, there is no way to track people through time and space. Given the life cycle of migration decisions, we believe including only those who still reside in their state of birth to be reasonable (Kaplan and Schulhofer-Wohl 2017). One concern is that those who leave their birth state to attend college may move back, but face a different gas price environment during early adulthood. While we cannot rule this out, even students applying to elite universities are significantly more likely to matriculate if they attended high school in the same state (Griffith and Rothstein 2009; Bostwick 2016), limiting concern about return migration.

4.1.1 Driving to Work

Estimates of Equation (2) on an indicator of driving to work are shown in Table 3. Note that each regression coefficient is from a separate linear probability model. Rows correspond to different definitions of treatment. We show specifications with gasoline price levels vs. changes, where the change is taken over either the ages 15 to 17 ($P_{cs}^{\Delta 17,15}$) or the two-year window around the minimum full-privilege driving age ($P_{cs}^{\Delta(m_{cs}+1),(m_{cs}-1)}$). Appendix Table A.4 shows results for various alternative specifications of treatment. Column (1) uses census year, state, and age fixed effects to capture life-cycle trends in commuting for those who reside in their state of birth when observed in the census (*stayers*). Columns (2) and (3) alter the sample. In Column (2), all workers are included regardless of current state of residence; gasoline prices reflect those in the state of birth. In Column (3), all workers are again included, but gasoline prices reflect current state of residence. Columns (4) to (7) progressively add more controls, again restricted to the *stayers* sample. Column (4) adds demographic controls for sex, marital status, educational attainment, and race. Column (5) adds log household income, which could influence a number of transportation and residential location margins. Column (6) adds state-by-sample year fixed effects to control for contemporaneous differences in driving conditions (current gas prices, parking restrictions, etc.) Finally, Column (7) includes quadratic birth year trends to control for smooth, secular time trends in preferences or the driving environment.

First, note that driving to work responds to changes in gasoline prices, but hardly to levels. For the specifications using gasoline price changes in Columns (1) to (3), all coefficients are negative and significant, and the effect size varies from -0.3 to -0.4 percentage points. Estimates are similar both for the 15 to 17 calendar age range and for the window one year before to one year after the minimum driving age. The results in Column (1) indicate that a doubling in the price of gasoline between the ages of 15 and 17 ($P_{cs}^{\Delta 17,15} = 1$) leads to a 0.38 percentage point or 0.43 percent reduction in driving to work later in life.[18] For the two-year window around the minimum driving age, the effect is -0.41 percentage points (or 0.46 percent). The precise magnitude slightly varies across specifications, but is always statistically significant.

Columns (4) through (7) condition on contemporaneous covariates or fixed effects. These factors, such as income, educational attainment, or current state, could be influenced by economic conditions during formative years (which might be correlated with gasoline price shocks). If such channels are important, including these controls could lead

[18] In our sample, 88.2 percent of workers drive to work.

Table 3: The effect of formative gasoline price on driving to work using the census/ACS 1980-2017.

	1[drive] (1)	1[drive] (2)	1[drive] (3)	1[drive] (4)	1[drive] (5)	1[drive] (6)	1[drive] (7)
$P_{cs}^{\Delta 17,15}$	-0.0038***	-0.0028**	-0.0031***	-0.0037***	-0.0039***	-0.0039***	-0.0043***
	(0.0010)	(0.0008)	(0.0009)	(0.0010)	(0.0010)	(0.0010)	(0.0009)
P_{cs}^{16}	-0.0007	0.0012+	-0.0029***	-0.0009	-0.0011	-0.0011	-0.0011
	(0.0010)	(0.0006)	(0.0007)	(0.0008)	(0.0009)	(0.0008)	(0.0008)
$P_{cs}^{\Delta(m_{cs}+1,m_{cs}-1)}$	-0.0041***	-0.0038***	-0.0040***	-0.0040***	-0.0040***	-0.0042***	-0.0045***
	(0.0010)	(0.0008)	(0.0008)	(0.0011)	(0.0010)	(0.0011)	(0.0010)
$P_{cs}^{m_{cs}}$	-0.0012	0.0006	-0.0012	-0.0013	-0.0015	-0.0015+	-0.0015+
	(0.0010)	(0.0006)	(0.0010)	(0.0009)	(0.0009)	(0.0008)	(0.0008)
Census year FEs	Y	Y	Y	Y	Y	-	-
State of birth FEs	Y	Y	Y	Y	Y	-	-
Age FEs	Y	Y	Y	Y	Y	Y	Y
Demographics	-	-	-	Y	Y	Y	Y
ln HH income	-	-	-	-	Y	Y	Y
State-X-year FEs	-	-	-	-	-	Y	Y
Quad. birth year	-	-	-	-	-	-	Y
Price in state of	Birth	Birth	Res	Birth	Birth	Birth	Birth
Sample	Stay	All	All	Stay	Stay	Stay	Stay

Each row and column represents the results from a different regression, for twenty-eight total. Dependent variable is a binary indicator of whether the respondent drove to work, as reported in the Census. Sample includes all native-born persons actively working in the Census between the ages of 25-54, and excludes farm workers and those coded N/A for transportation mode. Demographics include sex, marital status, educational attainment, and race. Observations weighted by person sample weights. Standard errors clustered by state of birth. + p<0.10, * p<0.05, ** p<0.01, *** p<0.001.

to biased estimates of the effect. However, these controls could also account for other effects of coming of age during a period of increasing gasoline prices. For completeness, we include them in Columns (4) to (7). They make little difference, strongly suggesting that income or education are not a primary mechanism for our results. Furthermore, the magnitude of these estimates corresponds to the effect measured in Section 3 (reflecting the doubling of prices in between 1979 and 1981).

Appendix Table A.4 shows that exposure to gasoline price changes during the 15-17 age range (or the window from one year before to one year after the minimum driving age) is associated with the strongest impact on later-life driving, but effects remain significant when using $P_{cs}^{\Delta(m_{cs}+2),(m_{cs}+1)}$ or $P_{cs}^{\Delta 18,17}$ (for the latter case, significance remains for only some of the specifications). Together these results suggest that the formative effect is strongest between the ages of 15 and 17.[19]

4.1.2 Income Mediation

Gasoline price movements can be associated with recessions, and recessions experienced when entering labor markets can have long-run effects on earnings (Oreopoulos, von Wachter, and Heisz 2012; Stuart 2017). If a gasoline price shock during the formative window impacts income (through cohort effects, labor market entry timing, educational attainment, etc.), and income influences transportation behavior, then some portion of the effect shown in Table 3 may be *indirectly* due to a shift in income rather than *directly* from experiencing the price shock. Thus, the negative effects in Columns (1) to (4) in Table 3 could in principle represent effects mediated through income channels.

Three pieces of evidence argue against income mediation as a main explanation of the effect we find. (Though we focus on income, this finding holds for other characteristics highly correlated with income, such as education.) First, Columns (5) to (7) in Table 3 control for contemporaneous income (and Columns (4) through (7) include controls for educational attainment). The stability of the estimates suggests that income cannot explain the observed effect. Second, results are similar when we exclude cohorts born in 1963 to 1965 (those that came of driving age during the 1979 oil crisis).[20] Finally, we conduct mediation analysis to quantify the indirect, income-mediated effect of these gasoline price shocks on later-life driving.

We formalize our mediation model in Appendix A.2. The analysis consists of modeling later-life driving behavior as a function of formative-year gasoline price shocks and contemporaneous income, and separately modeling contemporaneous income as a function of the formative-year gasoline price shocks. Results are reported in Appendix Table A.10 and show that income is indeed negatively associated with upward gasoline price shocks. However, the mediation analysis reveals that only between 2% and 24% of the total effect of the gasoline price shocks can be explained through the income channel (our average estimate is 13%). Reduced income due to the experience of recessions during formative years is not the primary component driving the results in Table 3.

[19] We discuss the use of cohort fixed effects in Appendix A.3. Most gasoline price variation is temporal rather than cross sectional, and the use of cohort fixed effects absorbs much of this variation. Nonetheless, extensive-margin results, while less precise, are similar to those presented here. Intensive-margin results are noisy.

[20] More precisely, when we estimate the same model as in Column (1) of Table 3, but exclude the 1963 to 1965 birth cohorts, the coefficient on $P_{cs}^{\Delta 17,15}$ is -0.0042** (s.e. 0.0015), slightly less precise but larger in magnitude than when these cohorts are included.

4.1.3 Other Extensive Margins (Transit Use, Vehicle Ownership)

We also investigate if those who experience positive gasoline price shocks during formative years substitute commutes to transit, and if they are less likely to have access to a vehicle. Table 4 summarizes the results for the same definitions of treatment as in Table 3. In Columns (1) and (2), the outcome variable is transit usage; in Columns (3) to (6) it is access to a vehicle. We find that about one-half to three-quarters of the extensive-margin effect is accounted for by a shift to transit. There may also be a small, negative effect of increases in gas prices on vehicle ownership (about one-quarter to one-half of the extensive-margin effect). The effect is not extremely robust, although more so in the sample that includes all household members than in the employed sample.

Table 4: The effect of formative gasoline price on other outcomes using the census/ACS 1980-2017.

	Transit usage		Vehicle available			
	1[transit] (1)	1[transit] (2)	1[vehicle] (3)	1[vehicle] (4)	1[vehicle] (5)	1[vehicle] (6)
$P_{cs}^{\Delta 17,15}$	0.0029***	0.0024**	-0.0014	-0.0009	-0.0019*	-0.0018**
	(0.0007)	(0.0009)	(0.0008)	(0.0006)	(0.0009)	(0.0006)
P_{cs}^{16}	0.0001	0.0004	0.0004	0.0007	-0.0007	-0.0001
	(0.0007)	(0.0005)	(0.0007)	(0.0005)	(0.0009)	(0.0007)
$P_{cs}^{\Delta(m_{cs}+1,m_{cs})}$	0.0028*	0.0021	-0.0025	-0.0023+	-0.0019	-0.0022
	(0.0012)	(0.0013)	(0.0016)	(0.0013)	(0.0016)	(0.0013)
$P_{cs}^{m_{cs}}$	0.0006	0.0008	0.0001	0.0003	-0.0008	-0.0005
	(0.0007)	(0.0005)	(0.0007)	(0.0005)	(0.0008)	(0.0006)
Census year FEs	Y	-	Y	-	Y	-
State of birth FEs	Y	-	Y	-	Y	-
Age FEs	Y	Y	Y	Y	Y	Y
Demographics	-	Y	-	Y	-	Y
ln HH income	-	Y	-	Y	-	Y
State-X-year FEs	-	Y	-	Y	-	Y
Quad. birth year	-	Y	-	Y	-	Y
Sample	Empl	Empl	Empl	Empl	All	All

Each row and column represents the results from a different regression, for twenty-four total. Dependent variable is either an indicator for transit usage or whether a vehicle is present in the household. Sample includes all native-born persons actively working in the Census between the ages of 25-54 still living in their state of birth, and excludes farm workers and, for transit use, those coded N/A for transportation mode. Demographics include sex, marital status, educational attainment, and race. Observations weighted by person sample weights. Standard errors clustered by state of birth. + p<0.10, * p<0.05, ** p<0.01, *** p<0.001.

4.2 Intensive Margin

We next perform a similar exercise to study the intensive margin: miles of driving and the fuel economy of the vehicles owned. We merge data on vehicle miles traveled (VMT) from all NHTS waves since 1990 with gasoline prices based on respondent age and current state. There are five waves in this dataset: 1990, 1995, 2001, 2009, and 2017. The NHTS also reports the vehicle's make and model, which we use to identify the type of vehicle (car vs. light-duty truck) and fuel economy. Household income is reported in binned values; we standardize these to correspond to quintiles of the income distribution and interact them with sample year dummies to flexibly capture changes in income patterns.

4.2.1 Miles of Driving

One might expect that drivers who were exposed to large gasoline price increases during their initial driving years perceive driving as more costly throughout their lives and drive fewer miles. Estimates of Equation (2) on the log of miles traveled are reported in Table 5. Column (1) uses NHTS sample year, state of residence, and age fixed effects. Column (2) adds demographic controls for race, urban/rural status, and family size. Column (3) adds the household income bins interacted with observation year. Columns (4) and (5) add more flexible state-by-year fixed effects to control for contemporaneous differences in driving conditions (current gas prices, parking restrictions, etc.) Column (5) adds a quadratic birth year trend.

These estimates represent the elasticity of long-run driving behavior to formative gasoline price changes. As with the extensive-margin results, gasoline price changes during early driving years matter while levels do not. Coefficients for specifications using changes are negative and significant whether controls or additional fixed effects are included or not. The effect varies from -3.4 to -8.2 percent. A doubling in the price of gasoline between the ages of 15 and 17 reduces later-life miles traveled by 6.2 to 8.2 percent; for the age window one year before to one year after the minimum driving age, the effect of doubling gasoline prices on miles traveled is -3.4 to -5.7 percent.[21] As with the extensive-margin analysis, including controls measured contemporaneously in Columns (2) to (5) could introduce bias in the estimated effect. However, these controls also capture other important differences between drivers. Regardless, the effect size is not greatly impacted by their inclusion. The specification with a quadratic trend in birth year in Column

[21] In our sample, respondents drive about 14,080 miles per year on average across the vehicles for which they are described as the main driver.

Table 5: The effect of formative gasoline price on log miles traveled using NHTS 1990-2017.

	ln(VMT) (1)	ln(VMT) (2)	ln(VMT) (3)	ln(VMT) (4)	ln(VMT) (5)
$P_{cs}^{\Delta 17,15}$	-0.0786** (0.0264)	-0.0822** (0.0260)	-0.0771** (0.0261)	-0.0773** (0.0259)	-0.0624* (0.0255)
P_{cs}^{16}	0.0213+ (0.0109)	0.0202+ (0.0110)	0.0190+ (0.0109)	0.0198+ (0.0111)	0.0032 (0.0096)
$P_{cs}^{\Delta(m_{cs}+1,m_{cs}-1)}$	-0.0502* (0.0193)	-0.0567** (0.0197)	-0.0470* (0.0201)	-0.0478* (0.0204)	-0.0344+ (0.0196)
$P_{cs}^{m_{cs}}$	0.0147 (0.0120)	0.0127 (0.0120)	0.0108 (0.0117)	0.0108 (0.0118)	-0.0027 (0.0107)
Sample year FEs	Y	Y	Y	-	-
State FEs	Y	Y	Y	-	-
Age FEs	Y	Y	Y	Y	Y
Controls	-	Y	Y	Y	Y
Income-by-year bin FEs	-	-	Y	Y	Y
State-X-year FEs	-	-	-	Y	Y
Quad. birth year	-	-	-	-	Y

Each row and column represents the results from a different regression, for twenty total. Dependent variable is log person VMT. Sample includes all respondents aged 25-54 with positive person VMT. Demographics include race, urbanization, and family size. Observations weighted by person sample weights. Standard errors clustered by state. + p<0.10, * p<0.05, ** p<0.01, *** p<0.001.

(5) shows a slightly attenuated effect on miles traveled.

Appendix Table A.6 shows the results for a broader range of treatment ages. Results behave similarly for these other treatment definitions, and effects remain significant for most alternatives. Compared with the extensive-margin results, the effects are strong for the entire 15 to 18 age range, with the largest treatment effect for $P_{cs}^{17,16}$. Together, these effects point to the importance of experiences with gasoline prices between the ages of 15 and 18 on later-life driving habits—both on the intensive and the extensive margin.

4.2.2 Fuel Economy and Vehicle Choice

Besides miles traveled, another margin of response is the fuel efficiency of the type of vehicle chosen by a driver. Appendix Table A.9 reports two additional outcome variables: the fuel-consumption rate in gallons-per-mile (GPM), and whether a driver owns a more fuel-efficient passenger car or a less fuel-efficient light-duty truck (pickup, SUV, or minivan). As the NHTS reports data at both the person and the vehicle level, we use the average fuel-consumption rate per driver (across the vehicles to which the person is assigned as the main driver) as the outcome variable in Columns (1) and (2); Columns (3) and (4) use GPM at the vehicle level. Columns (5) through (8) have a similar structure, now looking at the effect of formative gasoline price changes on large-vehicle ownership. All specifications include a rich set of fixed effects and controls. We also include vehicle age and quadratic vehicle model-year fixed effects to isolate variation within drivers of a vehicle of a certain vintage and age.

We find a zero effect on the fuel-consumption rating. This is not entirely surprising given that fuel economy is measured with considerable error.[22] We do find a negative effect on the large-vehicle indicator, suggesting that those who experienced a doubling of gasoline prices during their initial driving years are 1.1-2.7 percentage points less likely to drive a light-duty truck. The results are statistically significant in half of the specifications. All in all, we interpret this as modest suggestive evidence that gasoline price shocks have long-term effects on the types of vehicles that people drive, but more precisely measured fuel-economy data would be required to estimate the effect with a higher degree of confidence.

4.3 The Formative Window and Persistence

To show that the timing of gasoline price shocks matters, and to identify at what ages the formative experiences are most powerful, we merge later-life driving behavior to gasoline prices from ages 13 to 22. We similarly merge later-life driving behavior to gasoline prices several years before and after state-level minimum driving ages. For this test, we use Equation (2) with observation year, state, and age fixed effects, similar to Column (1) in both Table 3 and Table 5. Results are shown in Table 6.

Estimates of the long-run extensive-margin effect of driving to work are shown in the top panel. Using either $P_{cs}^{\Delta a,a-1}$ with ages $a \in \{13, 14, \ldots, 21, 22\}$ or $P_{cs}^{\Delta(m_{cs}+\tau),(m_{cs}+\tau-1)}$ with years relative to the state-cohort specific minimum driving age $\tau \in \{-3, -2, \ldots, 5, 6\}$ as the treatment definitions, estimates generally show that gasoline price shocks between the ages of 15 and 18 are influential for later-life driving, whereas shocks before or after play much less of a role. The largest coefficients are on $P_{cs}^{\Delta 16,15}$ and $P_{cs}^{\Delta(m_{cs}),(m_{cs}-1)}$. Estimates of the long-run intensive-margin effect on miles traveled are reported in the

[22]Fuel-economy ratings vary substantially within make and model, but we are unable to match this given the coarseness of the NHTS data.

Table 6: The effect of gasoline price changes at different ages.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$a =$	13	14	15	16	17	18	19	20	21	22
$\tau =$	-3	-2	-1	0	1	2	3	4	5	6
Panel A: Extensive margin (1[drive])										
$P_{cs}^{\Delta(a,a-1)}$	-0.0005	0.0012	-0.0001	-0.0054**	-0.0036**	-0.0023	-0.0009	0.0001	0.0005	0.0022*
	(0.0016)	(0.0012)	(0.0015)	(0.0016)	(0.0014)	(0.0016)	(0.0014)	(0.0017)	(0.0013)	(0.0011)
$P_{cs}^{\Delta(m_{cs}+\tau,m_{cs}+\tau-1)}$	0.0009	-0.0015	-0.0029	-0.0048***	-0.0044*	-0.0036*	0.0004	0.0012	0.0002	-0.0011
	(0.0012)	(0.0013)	(0.0019)	(0.0013)	(0.0018)	(0.0017)	(0.0020)	(0.0013)	(0.0014)	(0.0019)
Panel B: Intensive margin (ln(person VMT))										
$P_{cs}^{\Delta(a,a-1)}$	-0.0567	0.0263	0.0211	-0.0949*	-0.1125**	-0.0954*	-0.0395	0.0080	-0.0253	-0.0169
	(0.0498)	(0.0374)	(0.0403)	(0.0428)	(0.0401)	(0.0374)	(0.0422)	(0.0378)	(0.0412)	(0.0366)
$P_{cs}^{\Delta(m_{cs}+\tau,m_{cs}+\tau-1)}$	-0.0571	-0.0120	-0.0204	-0.0606+	-0.0618+	-0.0678+	-0.0583	-0.0077	-0.0213	0.0198
	(0.0379)	(0.0428)	(0.0445)	(0.0350)	(0.0343)	(0.0346)	(0.0399)	(0.0376)	(0.0379)	(0.0406)

Each row and column represents the results from a different regression, for forty total. Dependent variable is a binary indicator of whether the respondent drove to work in the Census data, and log person VMT in the NHTS sample. Regressions include state (or state of birth), sample year, and age fixed effects. Observations weighted by person sample weights. Standard errors clustered by state of birth. + p<0.10, * p<0.05, ** p<0.01, *** p<0.001.

bottom panel; the largest coefficients are about a year later than for the extensive margin. The most responsive age range is generally between 15 and 18 years, or between one year before and two or three years after the minimum driving age.[23]

We next ask whether these effects only last for a few years (say, for people into their early 30s), or whether they are persistent throughout the life cycle. Table 7 allows for heterogeneous treatment effects by 10-year age bins. Results reveal some differences in persistence across the intensive and extensive margins and specifications of the formative window. Generally, we conclude that the effects are stronger for younger (age 25 to 34) and older (age 45 to 54) drivers, with little discernible effect for those age 35 to 44. On the intensive margin, conclusions are broadly similar, but estimates are less precise.

5 Interpretation and Mechanisms

Empirical results presented so far demonstrate that gasoline price shocks have an effect on driving behavior in the long run. Specifications including income rule out one confounding explanation: gasoline price increases are often associated with recessions, which have path-dependent effects on earnings and potentially transportation behavior.

[23]Table A.8 shows the same effects but now using a two-year treatment window. The same conclusions apply.

Table 7: Persistence in the effect of formative gasoline price on driving.

	Extensive margin		Intensive margin	
	1[drive]	1[drive]	ln(VMT)	ln(VMT)
	(1)	(2)	(3)	(4)
$P_{cs}^{\Delta 17,15} \times$				
1[25-34]	-0.0050**	-0.0054***	-0.0890*	-0.0552
	(0.0018)	(0.0013)	(0.0433)	(0.0425)
1[35-44]	-0.0001	0.0006	-0.0529	-0.0328
	(0.0014)	(0.0014)	(0.0578)	(0.0524)
1[45-54]	-0.0050***	-0.0054***	-0.0925+	-0.1111*
	(0.0014)	(0.0013)	(0.0516)	(0.0497)
$P_{cs}^{\Delta(m_{cs}+1, m_{cs}-1)} \times$				
1[25-34]	-0.0031*	-0.0039*	-0.0464	-0.0279
	(0.0015)	(0.0015)	(0.0341)	(0.0323)
1[35-44]	-0.0038*	-0.0019	-0.0595	-0.0581
	(0.0019)	(0.0014)	(0.0479)	(0.0474)
1[45-54]	-0.0056**	-0.0069**	-0.0445	-0.0406
	(0.0019)	(0.0020)	(0.0427)	(0.0425)
Sample year FEs	Y	Y	Y	Y
State FEs	Y	Y	Y	Y
Age FEs	Y	Y	Y	Y
Demographics	-	Y	-	Y
Income	-	Y	-	Y
State-X-year FEs	-	Y	-	Y
Quad. birth year	-	Y	-	Y

Dependent variable in Columns (1) and (2) is a binary indicator of whether the respondent drove to work; demographics include sex, marital status, educational attainment, and race; and income is log household income. Dependent variable in Columns (3) and (4) is log person VMT; demographics include race, urbanization, and family size; and income are income bins interacted with sample year. Observations weighted by person sample weights. Standard errors clustered by state. + p<0.10, * p<0.05, ** p<0.01, *** p<0.001.

We now discuss if costly skill acquisition is a potential mechanism that can explain our results. If learning to drive has high time, vehicle and/or fuel costs, increases in these costs may have long-run impacts on driving adoption. Given that most drivers in the U.S. have historically learned to drive before they turn 18, parental inputs are also important. With a binding constraint on gas expenses, households may delay or completely avoid teenage driver training in favor of other necessary commuting expenses. Do high gasoline prices keep people from learning to drive in the long run?

We provide several pieces of evidence showing that costly skill acquisition is unlikely to wholly explain the results in Sections 3 and 4. First, it is difficult to assert that the intensive-margin effect for those who own a vehicle is due to a high cost of learning to drive. Hence, the presence of an intensive-margin effect on miles driven is highly suggestive that delayed skill acquisition cannot be the sole or dominant explanation for the long-run effects on driving behavior. Second, there does not appear to be a reduction in the take up of teen driver's licenses in response to the 1979 oil crisis. Finally, we show that regulations that explicitly restrict teenage driving through minimum driver licensing age requirements do not have negative effects on later-life driving rates.

Together, these null results suggest a role for formative experiences that shape preferences for the long run. Teen's preferences for driving or perceptions of the costs of driving are impacted by gasoline price shocks, and these differences persist. We cannot distinguish between a shift in a deep preference parameter, or how preferences are filtered by updated perceptions of the cost of driving.

5.1 Evidence from Driver Licensing Counts

We return to the gasoline price shock induced by the 1979 oil crisis and examine the response in driver licensing. We compute the percentage of each cohort that has a license by a certain age. This statistic is not directly observable in the data, but the Federal Highway Administration (FWHA) publishes annual data on the number of drivers at each age from 15 to 24. We combine this data with supplemental estimates of the age distribution of the population to generate percentage licensed by age. In general, other factors (such as changes in minimum driver licensing age) were mostly constant during the late 1970s and early 1980s.

There is not a noticeable change in driver licensing following the 1979 oil crisis, although the data are somewhat noisy. Figure 4 shows the percentage of each age (16, 17,

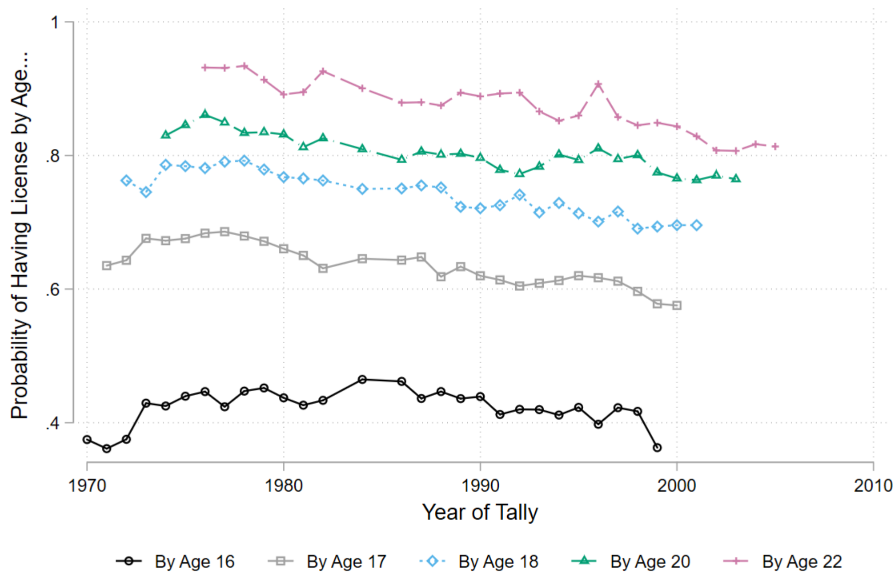

Figure 4: The probability of having a driver's license by calendar year and age.

18, 20, and 22) that has a license in each year.[24] Timing for ages is based on calendar years rather than individual birthdays, so the lines should be read as "the percentage of those aged 16 at the end of the year who received a license by the end of the year." Driver saturation is generally smooth and slightly decreasing, though there may be a slight depression in some of the series in 1981 and 1982 but a slight uptick for others. All things considered, we do not find clear evidence for a delay in driver license uptake.

5.2 Evidence from Driver License Minimum Age Requirements

Legislative restrictions provide another avenue that potentially limit driver training. If high gas prices delay driving skill acquisition and reduce later-life driving because of this delay, it is likely that directly delaying driver skill acquisition through driving age restrictions will also reduce later-life driving. We combine data from several sources to develop a panel of teenage driver's license requirements covering 1967 to 2017 to test this channel.[25]

[24] Figure 4 omits 1983 and 1985 because the driver's license counts by age in those years was extrapolated from prior years. These data are noisy, but show a marked increase in driver's licenses of 18 year olds in 1983.

[25] Our two primary sources are the FHWA's Driver License Administration Requirements and Fees report and a database of graduated driver license (GDL) adoptions from the Insurance Institute of Highway Safety.

Table 8: Do youth driving restrictions affect later driving behavior?

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Extensive margin (1[drive])						
Minimum Full Privilege Age	0.0078	0.0048	0.0071	0.0072	0.0082+	0.0092
	(0.0052)	(0.0040)	(0.0047)	(0.0048)	(0.0048)	(0.0056)
Minimum Intermediate License Age	-0.0107	-0.0088	-0.0091	-0.0097	-0.0137	-0.0124
	(0.0147)	(0.0122)	(0.0136)	(0.0138)	(0.0127)	(0.0121)
Panel B: Intensive margin (ln(person VMT))						
Minimum Full Privilege Age	0.0012		0.0010	-0.0030	-0.0108	0.0196
	(0.0129)		(0.0132)	(0.0159)	(0.0182)	(0.0143)
Minimum Intermediate License Age	-0.0269		-0.0239	-0.0270	-0.0007	0.0239
	(0.0651)		(0.0565)	(0.0592)	(0.0699)	(0.0588)
Sample year FEs	Y	Y	Y	Y	-	-
State FEs	Y	Y	Y	Y	-	-
Age FEs	Y	Y	Y	Y	Y	Y
Dem. controls	-	-	Y	Y	Y	Y
Income controls	-	-	-	Y	Y	Y
State-X-year FEs	-	-	-	-	Y	Y
Quad. birth year	-	-	-	-	-	Y
Sample	Stay	All	Stay	Stay	Stay	Stay

Each panel and column contains the results from a different regression, for eleven total. Dependent variable in first panel is a binary indicator of whether the respondent drove to work; demographics include sex, marital status, educational attainment, and race; and income is log household income. Dependent variable in second panel is log person VMT; demographics include race, urbanization, and family size; and income are income bins interacted with sample year. See text for description of driver license requirements. Observations weighted by person sample weights. Standard errors clustered by state. + p<0.10, * p<0.05, ** p<0.01, *** p<0.001.

We test for extensive- and intensive-margin effects of changing minimum driving age restrictions in Table 8. We first construct two measures of minimum driving age. The first measure, minimum full privilege age, gives the minimum age at which a suitably trained teenager can obtain an unrestricted license (i.e., with no restrictions on time of use, purpose, destination, or passengers). This is very similar to the age used to merge gas price relative to driver license age in Section 4.[26] The other measure, minimum intermediate license age, captures the minimum age at which drivers can make unaccompanied trips, but with some restrictions.[27] Summary information on these measures is shown in Appendix Table A.3. We include both measures of restrictiveness in each specification.

[26]We use minimum age in years and months to define treatment here, whereas in Section 4 this variable is rounded to the nearest year to facilitate matching with coarser age data.

[27]That is, they need not be accompanied by a parent or older driver (as for a learner's permit).

Columns (1) and (2) show the effects from models with sample year, state, and age fixed effects. Column (2) includes the full sample for the census data, which is otherwise restricted to the *stayers* sample. Column (3) adds the respective demographic controls as in Table 3 and 5. Column (4) adds income controls, while Column (5) adds state-by-sample year fixed effects and Column (6) includes a quadratic trend in birth year. We find little evidence of a long-run effect of these regulations on later-life driving behavior. Estimates in Table 8 show the effect of a one-year increase in either measure of minimum age, and are never significant in the expected direction.

We suggest some caution in interpretation: magnitudes cannot be directly compared to results from analysis in the prior sections because the treatments are fundamentally different (and measured in incompatible units). Our primary analysis studies the effects of gasoline price changes, while the analysis here studies the long-run impacts of driving age restrictions. However, with these caveats in mind, this analysis suggests that age restrictions for teenagers on learning to drive do not inhibit the long-run adoption of driving. As high gasoline prices are less extreme than legal restrictions on driving, it is reasonable to conjecture that gasoline price shocks do not impact driver license uptake either.

6 Conclusion

Early experiences frame how people perceive different goods and activities. These formative periods can drive later-life behaviors, expectations, and norms. In the case of driving, we find that individuals who experience large price increases during their formative driving years behave differently than those who did not experience such shocks: they drive to work less often, take transit more, are less likely to have access to a vehicle, and drive fewer miles if they own a vehicle. Early-life experiences are thus one source of path dependence in transportation demand.

These results highlight that macroeconomic price shocks can give rise to long-lived preference heterogeneity. Combining the intensive-margin (miles traveled) and extensive-margin (vehicle ownership) effects of a doubling of gasoline prices, we find a combined long-run, path-dependent driving reduction of 3.6-8.7 percent. The literature has reported short- and medium-run estimates that are generally in the range of -5 to -39 percent and decreasing in magnitude over time (Small and van Dender 2007; Li, Linn, and Muehlegger 2014). Our long-run effect is smaller but in the same order of magnitude of the more recent short-run miles-traveled elasticities in the literature, although it operates only on cohorts that are exposed to large gasoline price movements during teenage driving years.

We show that these long-run effects are most likely due to the formation of preferences for the driving experience and its perceived associated costs, rather than through long-run income effects or a reduction in the number of people who end up learning to drive. These results show that formative experiences that determine later-in-life behavior need not be 'extreme'; everyday mundane experiences with market prices can have long-lasting impacts on preferences and behavior.

References

Allcott, Hunt, and Christopher R Knittel. 2019. "Are consumers poorly informed about fuel economy? Evidence from two experiments." *American Economic Journal: Economic Policy* 11 (1): 1–37.

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

Anderson, Soren T, Ryan Kellogg, Ashley Langer, and James M Sallee. 2015. "The intergenerational transmission of automobile brand preferences." *The Journal of Industrial Economics* 63 (4): 763–793.

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

Bamberg, Sebastian. 2006. "Is a residential relocation a good opportunity to change people's travel behavior? Results from a theory-driven intervention study." *Environment and Behavior* 38 (6): 820–840.

Baumeister, Christiane, and Lutz Kilian. 2016. "Forty years of oil price fluctuations: why the price of oil may still surprise us." *Journal of Economic Perspectives* 30 (1): 139–160.

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

Bleakley, Hoyt, and Jeffrey Lin. 2012. "Portage and path dependence." *The Quarterly Journal of Economics* 127 (2): 587–644.

Bostwick, Valerie. 2016. "Signaling in higher education: The effect of access to elite colleges on choice of major." *Economic Inquiry* 54 (3): 1383–1401.

Bronnenberg, Bart J, Jean-Pierre H Dubé, and Matthew Gentzkow. 2012. "The evolution of brand preferences: Evidence from consumer migration." *American Economic Review* 102 (6): 2472–2508.

Brooks, Leah, and Byron F Lutz. 2019. "Vestiges of transit: Urban persistence at a microscale." *The Review of Economics and Statistics* 101 (3): 385–399.

Brown, James R, J Anthony Cookson, and Rawley Z Heimer. 2019. "Growing up without finance." *Journal of Financial Economics* Forthcoming.

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

Busse, Meghan R, Devin G Pope, Jaren C Pope, and Jorge Silva-Risso. 2015. "The psychological effect of weather on car purchases." *The Quarterly Journal of Economics* 130 (1): 371–414.

Callen, Michael, Mohammad Isaqzadeh, James D Long, and Charles Sprenger. 2014. "Violence and risk preference: Experimental evidence from Afghanistan." *American Economic Review* 104 (1): 123–148.

Davis, Donald R, and David E Weinstein. 2008. "A search for multiple equilibria in urban industrial structure." *Journal of Regional Science* 48 (1): 29–65.

Deacon, Robert T, and Jon Sonstelie. 1985. "Rationing by waiting and the value of time: Results from a natural experiment." *Journal of Political Economy* 93 (4): 627–647.

———. 1989. "The welfare costs of rationing by waiting." *Economic Inquiry* 27 (2): 179–196.

Donaldson, Dave, and Richard Hornbeck. 2016. "Railroads and American economic growth: A "market access" approach." *The Quarterly Journal of Economics* 131 (2): 799–858.

Espey, Molly. 1998. "Gasoline demand revisited: An international meta-analysis of elasticities." *Energy Economics* 20 (3): 273–295.

Frech, HE, and William C Lee. 1987. "The welfare cost of rationing-by-queuing across markets: Theory and estimates from the U.S. gasoline crises." *The Quarterly Journal of Economics* 102 (1): 97–108.

Fujiwara, Thomas, Kyle Meng, and Tom Vogl. 2016. "Habit formation in voting: Evidence from rainy elections." *American Economic Journal: Applied Economics* 8 (4): 160–188.

Gillingham, Kenneth, Sébastien Houde, and Arthur van Benthem. 2019. "Consumer myopia in vehicle purchases: Evidence from a natural experiment." *NBER Working Paper* 25845.

Gillingham, Kenneth, Alan Jenn, and Inês ML Azevedo. 2015. "Heterogeneity in the response to gasoline prices: Evidence from Pennsylvania and implications for the rebound effect." *Energy Economics* 52 (S1): S41–S52.

Giuliano, Paola, and Antonio Spilimbergo. 2013. "Growing up in a recession." *Review of Economic Studies* 81 (2): 787–817.

Griffith, Amanda L, and Donna S Rothstein. 2009. "Can't get there from here: The decision to apply to a selective college." *Economics of Education Review* 28 (5): 620–628.

Hamilton, James D. 1985. "Historical causes of postwar oil shocks and recessions." *The Energy Journal* 6 (1): 97–116.

Hughes, Jonathan, Christopher R Knittel, and Daniel Sperling. 2008. "Evidence of a shift in the short-run price elasticity of gasoline demand." *The Energy Journal* 29 (1): 113–134.

Jacobsen, Mark R, and Arthur A van Benthem. 2015. "Vehicle scrappage and gasoline policy." *American Economic Review* 105 (3): 1312–1338.

Kaplan, Greg, and Sam Schulhofer-Wohl. 2017. "Understanding the long-run decline in interstate migration." *International Economic Review* 58 (1): 57–94.

Klein, Nicholas J, and Michael J Smart. 2017. "Millennials and car ownership: Less money, fewer cars." *Transport Policy* 53 (C): 20–29.

Knittel, Christopher R, and Elizabeth Murphy. 2019. "Generational trends in vehicle ownership and use: Are millennials any different?" *NBER Working Paper* 25674.

Kolesár, Michal, and Christoph Rothe. 2018. "Inference in regression discontinuity designs with a discrete running variable." *American Economic Review* 108 (8): 2277–2304.

Larcom, Shaun, Ferdinand Rauch, and Tim Willems. 2017. "The benefits of forced experimentation: Striking evidence from the London underground network." *The Quarterly Journal of Economics* 132 (4): 2019–2055.

Leard, Benjamin, Joshua Linn, and Clayton Munnings. 2019. "Explaining the evolution of passenger vehicle miles traveled in the United States." *The Energy Journal* 40 (1): 25–54.

Li, Shanjun, Joshua Linn, and Erich Muehlegger. 2014. "Gasoline taxes and consumer behavior." *American Economic Journal: Economic Policy* 6 (4): 302–342.

Li, Shanjun, Christopher Timmins, and Roger H von Haefen. 2009. "How do gasoline prices affect fleet fuel economy?" *American Economic Journal: Economic Policy* 2 (2): 113–137.

Logan, Trevon D, and Paul W Rhode. 2010. "Moveable feasts: A new approach to endogenizing tastes." *Working Paper.*

Malmendier, Ulrike, and Stefan Nagel. 2011. "Depression babies: Do macroeconomic experiences affect risk taking?" *The Quarterly Journal of Economics* 126 (1): 373–416.

Malmendier, Ulrike, and Leslie Sheng Shen. 2018. "Scarred consumption." *NBER Working Paper* 24696.

Manski, Charles F, and Daniel McFadden. 1981. "Alternative estimators and sample designs for discrete choice analysis." In *Structural analysis of discrete data with econometric applications,* edited by Charles F Manski and Daniel McFadden. Cambridge, MA: MIT Press.

Michaels, Guy, and Ferdinand Rauch. 2018. "Resetting the urban network: 117–2012." *The Economic Journal* 128 (608): 378–412.

Oreopoulos, Philip, Till von Wachter, and Andrew Heisz. 2012. "The short-and long-term career effects of graduating in a recession." *American Economic Journal: Applied Economics* 4 (1): 1–29.

Pollak, Robert A. 1970. "Habit formation and dynamic demand functions." *Journal of Political Economy* 78 (4 (Part 1)): 745–763.

Puller, Steven L, and Lorna A Greening. 1999. "Household adjustment to gasoline price change: An analysis using 9 years of US survey data." *Energy Economics* 21 (1): 37–52.

Simonsohn, Uri. 2006. "New Yorkers commute more everywhere: Contrast effects in the field." *Review of Economics and Statistics* 88 (1): 1–9.

Simonsohn, Uri, and George Loewenstein. 2006. "Mistake #37: The effect of previously encountered prices on current housing demand." *The Economic Journal* 116 (508): 175–199.

Small, Kenneth A, and Kurt van Dender. 2007. "Fuel efficiency and motor vehicle travel: The declining rebound effect." *The Energy Journal* 28 (1): 25–51.

Stuart, Bryan. 2017. "The long-run effects of recessions on education and income." *Center for Economic Studies, U.S. Census Bureau Working Paper* 17-52.

West, Jeremy, Mark Hoekstra, Jonathan Meer, and Steven L Puller. 2017. "Vehicle miles (not) traveled: Fuel economy requirements, vehicle characteristics, and household driving." *Journal of Public Economics* 145 (January): 65–81.

Yang, Nan, and Yong Long Lim. 2017. "Temporary incentives change daily routines: Evidence from a field experiment on Singapore's subways." *Management Science* 64 (7): 3365–3379.

Appendix

A.1 Data Notes

Census Data

We draw data on individual commuting behavior in part for the United States Census and American Community Survey (ACS). In particular, we collect the 5% state samples for 1980 and 1990, the 5% sample for 2000, the 2006-10 5-year ACS, the 2011-15 5-year ACS, and the 2016 and 2017 1-year ACS data abstracts from the IPUMS website. We focus on 'Journey to Work' variables, but also draw on a variety of demographic and economic characteristics. These variables have generally been harmonized by IPUMS.

The primary outcome variables from the Census/ACS that we use are derived from the following variables (along with IPUMS descriptions):

> TRANWORK is asked similarly from 1980 on, and "reports the respondent's primary means of transportation to work . . . over the course of the previous week . . . The primary means of transportation was that used on the most days or to cover the greatest distance." This variable varies by person, and is only available for employed persons who are currently working.
>
> VEHICLES is available from 1990 on, and "reports the number of cars, vans, and trucks of one-ton capacity or less kept at home for use by household members," including "company cars regularly kept at home and used for non-business purposes." This variable is available for households.
>
> AUTOS is available in 1980, and "reports the number of automobiles owned or used regularly by any household member. It includes company cars kept at home and available for personal use." This variable is available for households.
>
> TRUCKS is available in 1980, and "reports the number of trucks and vans regularly kept at home for use by members of the household, including company vehicles. It excludes trucks with more than one-ton capacity, those permanently out of working order, and those used only for business purposes." This variable is available for households.

We then define our primary outcome variables from these as follows:

1[drive] is equal to 1 if TRANWORK takes codes 10-15 (for auto, truck, or van conveyance), and 0 otherwise.

1[transit] is equal to 1 if TRANWORK takes codes 30-34 or 36 (public transit conveyance excluding taxis), and 0 otherwise.

1[vehicle] is equal to 1 if VEHICLES is greater than or equal to 1 (1990 on) or the sum of AUTOS and TRUCKS is greater than or equal to 1 (1980), and 0 otherwise.

A couple of other variables play key roles in our analysis, as we use them to merge Census/ACS data with measures of gasoline price variation (i.e., treatment).

AGE is asked in all years, and is used to create a variable BIRTHYR by subtraction from the survey year. Ostensibly, AGE is meant to be relative to Census day (in early April) for Census samples, and relative to the day of survey for ACS samples. Thus, BIRTHYR is not necessarily an accurate measure of the year of birth. For example, someone born on March 15, 1964, would respond to the 2000 Census that their AGE is 36, and BIRTHYR would be recorded as 1964. However, someone born on April 15, 1964, would respond that their AGE is 35, and BIRTHYR would be recorded as 1965. Hence, there is some measurement error in birth year, and results should be interpreted with this caveat in mind. Respondents from ACS years also suffer from some measurement error, but it should be zero on average.

BPL reports the respondent's state or country of birth, and STATEFIP gives a respondent's current state of residence. We use these variables to merge a respondent to gasoline prices in their formative years in their likely state of residence at that time. We merge on BPL, on STATEFIP, and on both if a respondent currently resides in their state of birth. 63.8% of the whole sample and 63.3% of the commuting sample currently reside in their state of birth.

We also use a variety of other variables as controls, for sample selection, or in robustness exercises. These include sex, marital status, educational attainment (separate indicators for high school and college completion), race and ethnicity (indicators for African American and Hispanic), household income (inflation adjusted using CPI), employment and labor force participation, wage, housing tenure, rent and house value, and measures of travel time for commuting.

NHTS Data

We use five waves of the National Household Travel Survey and its predecessor, the Nationwide Personal Transportation Survey, (collectively NHTS) from 1990, 1995, 2001, 2009, and 2017. These data document details at the household, person, vehicle, and trip level. Our analysis focuses on the person level. We use the following variables to generate our primary outcomes:

> ANNMILES is a self-reported annualized miles estimate given per vehicle.
>
> WHOMAIN describes which person in the household drives each vehicle the most.

We then use these to create a person-specific measure of total vehicle miles traveled (VMT) by adding together all ANNMILES across vehicles for which WHOMAIN is the primary driver. We top code this value at 115,000 miles annually (alternative top codes at 50,000 or 200,000 miles make little difference).

We use the variable R_AGE to determine a respondent's birth year and to perform merges to gasoline prices. Information on when (generally the month) the interview was conducted is also used. The variable HHSTATE captures the household's current state of residence; no historical detail on location of nativity or migration is provided.

For some specifications, we use other details associated with particular vehicles. MAKE, MODEL, VEHYEAR, HYBRID, and FUELTYPE (or near variants) give make, model, vehicle year, hybrid status, and gas/diesel/electric information about vehicles. The make and model information is relatively coarse, and codes roughly align with those used by the National Highway Traffic Safety Administration (NHTSA).

We also derive a number of other control variables including race (an indicator for white), urbanization (an indicator for urban residential environment), family size, and bins of household income (we harmonize bins measures across years and adjust for inflation, resulting in five bins which are then interacted with year).

Gasoline Price Data

The Energy Information Administration reports nominal tax-inclusive state-level gasoline price data starting in 1983. For the years 1966-1982, we use the Highway Statistics Annuals (Li, Linn, and Muehlegger 2014; Small and van Dender 2007).

Driver License Regulation Data

We develop a data set of driver licensing requirements from several sources. Our primary sources are (i) the FHWA *Driver License Administration Requirements and Fees* booklet and (ii) the Insurance Institute for Highway Safety (IIHS) database on graduated driver license (GDL) programs. We supplement these materials with various newspaper articles, legal database queries, and inquiries to reference desks at state libraries.

The FHWA booklet has been published roughly biannually since the 1960s. We found and use years 1967, 1972, 1980, 1982, 1984, 1986, 1988, 1994, and 1996. IIHS data cover 1995 to 2017, and report some information starting in 1990. In general, if we do not observe a change between two periods, we assume no legislative change. If we do observe a change, we perform additional inquiries to determine the date of change.

We define two measures of driver license minimum age. Our primary measure is the minimum age at which a teenager can obtain a full-privilege driver license. Our definition allows for teenagers to have taken driver education classes and be enrolled in school (often requirements for receiving a license before the age of 17 or 18). We exclude hardship rules, farm licenses, and other types of specialty licenses (e.g., motorcycle). The second measure captures the minimum age at which a teenager can obtain an intermediate license. These licenses permit unaccompanied driving, but place some restrictions on when a license holder may drive alone (e.g., daytime only), or who they may drive with (e.g., one non-family member).

Driver Licensing Counts

The Federal Highway Administration (FHWA) publishes data on driver licensing. We use Table DL-220 "Licensed Drivers, by Sex and Age Group, 1963-2016" from *Highway Statistics* (2016), which lists the number of driver licenses held by people of each age from 16 to 24 in each year. The FHWA did not require states to report counts by age in 1983 and 1985, and instead extrapolated these data. We exclude these years.

To estimate rates of driver license adoption, we require age-specific estimates of population. We construct this data from the National Cancer Institute's SEER Population data, which provides population estimates by age from 1969 to 2017. We sum county-level population estimates across all counties for each age and year.

Fuel-Efficiency Data

We use EPA fuel-economy data from Allcott and Knittel (2019). These data report fuel-economy data by make, model, year, trim, fuel type, and engine size. In the NHTS data, we only observe make, model, year, and fuel type. We therefore create an average fuel efficiency by make, model, year, and fuel type class, measured in gallons per mile (GPM), and use this as our measure of vehicle efficiency.

A.2 Mediation Analysis

We perform mediation analysis to explicitly account for the *indirect* effect gasoline price shocks experienced during formative years could have on later-life driving through an income effect. If a gasoline price shock during this formative window impacts wages (through cohort effects, labor market entry timing, educational attainment, etc.), and wages influence transportation behavior, the indirect portion of the effect is not due to a shift in preferences.

We build a simple mediation model (Baron and Kenny 1986; MacKinnon 2012).[28] We mostly retain notation from Section 4: Later-life driving, Y, is modeled as a function of the gasoline price shock experienced during formative driving years, T, and contemporaneous income, M. However, the gasoline price shock may also have an effect on income, and so mediate later-life driving indirectly through income. The mediation model can be expressed by the stacked equation (suppressing subscripts for exposition):

$$\begin{pmatrix} Y \\ M \end{pmatrix} = \begin{pmatrix} \theta^Y \\ \theta^M \end{pmatrix} T + \begin{pmatrix} \gamma \\ 0 \end{pmatrix} M + \begin{pmatrix} \delta^Y \\ \delta^M \end{pmatrix} X + \begin{pmatrix} \epsilon^Y \\ \epsilon^M \end{pmatrix}. \tag{A.1}$$

In Equation (A.1), θ^Y is the effect of T on Y, while γ is the effect of M on Y. T is permitted to have its own effect on M via θ^M. The direct effect of formative gasoline price shocks on later-life driving is captured by θ^Y, the indirect effect mediated through income is the product $\gamma\theta^M$, and the total effect sums these two together: $\theta^Y + \gamma\theta^M$.

We implement Equation (A.1) in a similar manner to Equation (2), and include in δ age, state of birth, sample year fixed effects, and exogenous demographics (sex and race). The fixed effects and demographic covariates are allowed to vary across outcomes, (hence δ^Y and δ^M). We cluster standard errors by state of birth across outcomes. Finally, we assume that T and M are exogenous conditional on the fixed effects and covariates we include as well as autonomy (that is, γ does not vary with T).[29] Alternatively, mediation analysis

[28]Mediators are a class of what are denoted as 'bad controls' by (Angrist and Pischke 2008). They are 'bad' in the sense that they can confound estimation of average treatment effects. Recent literature has begun to explicitly explore these estimators (e.g., Dippel et al. 2017; Heckman and Pinto 2015). In particular, Heckman and Pinto (2015) reviews early econometric mediation analysis.

[29]One set of assumptions under which this model is identified is labeled sequential exogeneity: (i)

could proceed using estimates of γ and θ^M drawn from the literature (e.g., the literature on graduating during a recession provides proxies for θ^M).

Appendix Table A.10 presents several specifications for different combinations of treatment and income. Columns (1), (3) and (5) report effects using the absolute calendar age measure of treatment $P_{cs}^{\Delta 17,15}$, while Columns (2), (4), and (6) use the measure based on minimum full-privilege driver license age $P_{cs}^{\Delta(m_{cs}+1),(m_{cs}-1)}$. Columns (1) and (2) use household income, Columns (3) and (4) use wage income, and Columns (5) and (6) use personal income.

Estimates of θ^Y are similar to those in the main text. Income has a positive relationship with driving a private vehicle to work (driving is a normal good); estimates of γ indicate that a 10% increase in income is associated with roughly a 0.2 percentage point increase in driving to work. Finally, the various definitions of gasoline price shocks have a negative relationship with income, though the strength of this relationship varies with the definition of income used. Estimates of θ^M indicate that experiencing a doubling of gasoline prices during formative years may be associated with up to a -4.9 percent decrease in income.[30]

These results show that most the vast majority of the gasoline price shock effect does not come through the indirect income channel. The ratio of the direct effect to the total effect indicates that between 76 and 98 percent of the total effect cannot be explained by income. This analysis suggests that income-based explanations are, for the most part, unimportant in understanding the long-run relationship between gasoline price shocks during formative years and later-life driving.

A.3 Additional Specifications

We have also run models with cohort fixed effects. The use of cohort fixed effects absorbs most of the variation in gasoline prices, as gasoline prices vary much more over time than space (see Appendix Figure A.1). This should signal caution to taking estimates from models with cohort fixed effects too seriously. Though θ in Equation (2) is still identified when we control for cohort (birth year) fixed effects, the source of identifying variation

$(Y, M) \perp\!\!\!\perp T|X$, (ii) $Y \perp\!\!\!\perp M|T, X$ and (iii) common support (Imai, Keele, and Yamamoto 2010; Imai et al. 2011). Heckman and Pinto (2015) argues that such conditions may be strong, and Dippel et al. (2017) provide an approach to identification with endogenous variation.

[30] This suggests, for example, that the 1979 Oil Crisis is associated with income loss of about 2.5 percent later in life. This measure is smaller than in Kahn (2010) and Stuart (2017), reflecting the fact that gasoline price shocks and recessions are not always correlated.

changes. The variation that remains after conditioning on cohort fixed effects (in addition to the state fixed effects) are differential changes in T_{cs} across states, e.g., a larger increase in Georgia than in Alabama in a given year. These movements are only a small piece of the observable variation facing agents.

We present results with cohort fixed effects in Appendix Table A.5 (outcome: driving) and Appendix Table A.7 (outcome: miles traveled). In general, estimates from these models are statistically significant only for a subset of specifications because of a loss of power due to much less variation. Using $P_{cs}^{\Delta(m_{cs}+2,m_{cs})}$ as the treatment definition, we find that estimates of the extensive-margin effect are very similar in magnitude to the results in Table 3. Additional identifying variation in this case is coming from within state changes in driver license age requirements over time. The estimated effect is now concentrated in the 16 to 18 age range. Intensive-margin results on miles traveled are noisy and almost always lack significance when cohort fixed effects are included.

Appendix References

Allcott, Hunt, and Christopher R Knittel. 2019. "Are consumers poorly informed about fuel economy? Evidence from two experiments." *American Economic Journal: Economic Policy* 11 (1): 1–37.

Angrist, Joshua D, and Jörn-Steffen Pischke. 2008. *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton University Press.

Baron, Reuben M, and David A Kenny. 1986. "The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations." *Journal of Personality and Social Psychology* 51 (6): 1173.

Dippel, Christian, Robert Gold, Stephan Heblich, and Rodrigo Pinto. 2017. *Instrumental variables and causal mechanisms: Unpacking the effect of trade on workers and voters.* Technical report. National Bureau of Economic Research.

Heckman, James J, and Rodrigo Pinto. 2015. "Econometric mediation analyses: Identifying the sources of treatment effects from experimentally estimated production technologies with unmeasured and mismeasured inputs." *Econometric Reviews* 34 (1-2): 6–31.

Imai, Kosuke, Luke Keele, Dustin Tingley, and Teppei Yamamoto. 2011. "Unpacking the black box of causality: Learning about causal mechanisms from experimental and observational studies." *American Political Science Review* 105 (4): 765–789.

Imai, Kosuke, Luke Keele, and Teppei Yamamoto. 2010. "Identification, inference and sensitivity analysis for causal mediation effects." *Statistical science:* 51–71.

Kahn, Lisa B. 2010. "The long-term labor market consequences of graduating from college in a bad economy." *Labour Economics* 17 (2): 303–316.

Li, Shanjun, Joshua Linn, and Erich Muehlegger. 2014. "Gasoline taxes and consumer behavior." *American Economic Journal: Economic Policy* 6 (4): 302–342.

MacKinnon, David. 2012. *Introduction to Statistical Mediation Analysis.* Routledge.

Small, Kenneth A, and Kurt van Dender. 2007. "Fuel efficiency and motor vehicle travel: The declining rebound effect." *The Energy Journal* 28 (1): 25–51.

Stuart, Bryan. 2017. "The long-run effects of recessions on education and income." *Center for Economic Studies, U.S. Census Bureau Working Paper* 17-52.

A.4 Appendix Figures and Tables

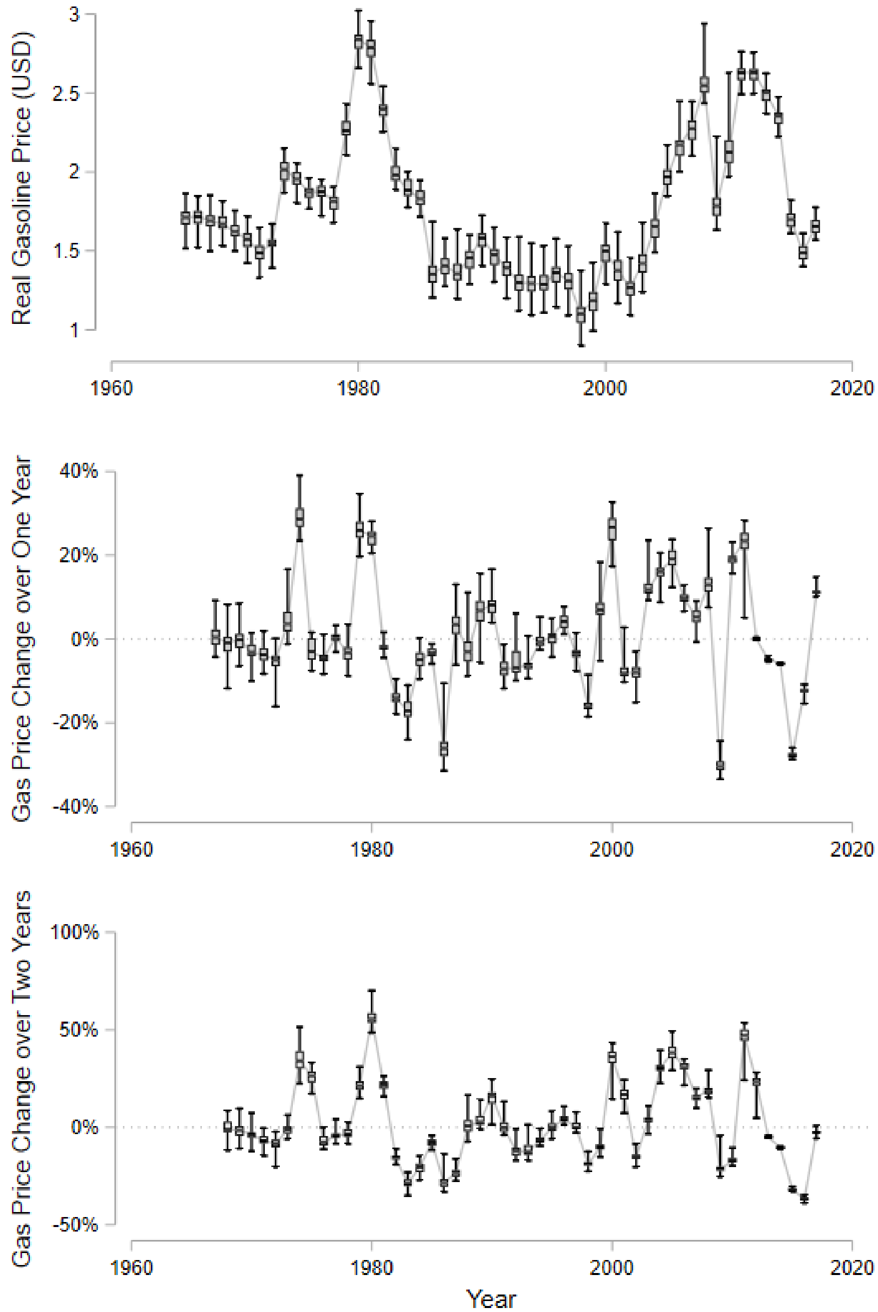

Figure A.1: Box plots of state gas prices and 1- and 2-year percentage changes; minimum, maximum and quartiles.

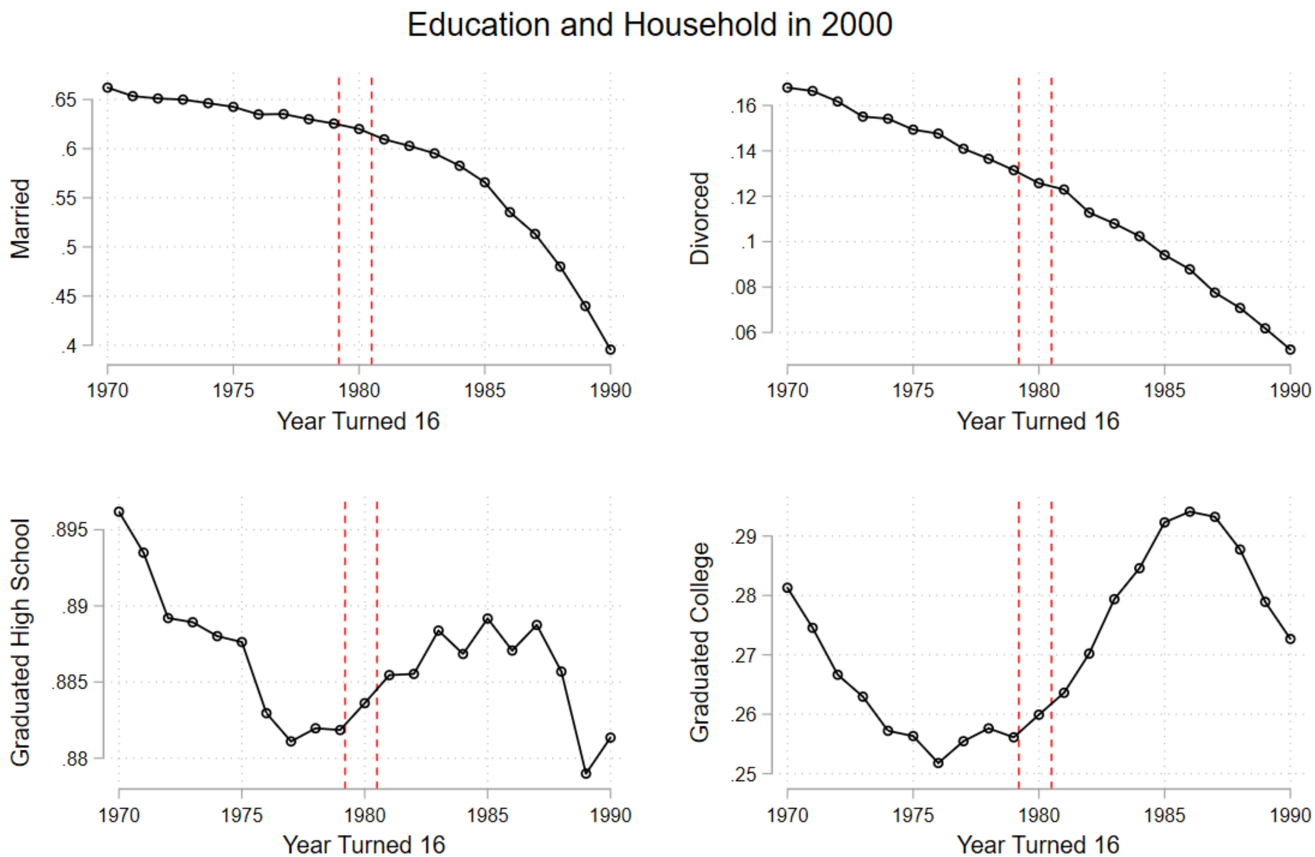

Figure A.2: Demographic characteristics in 2000 by year turned 16.

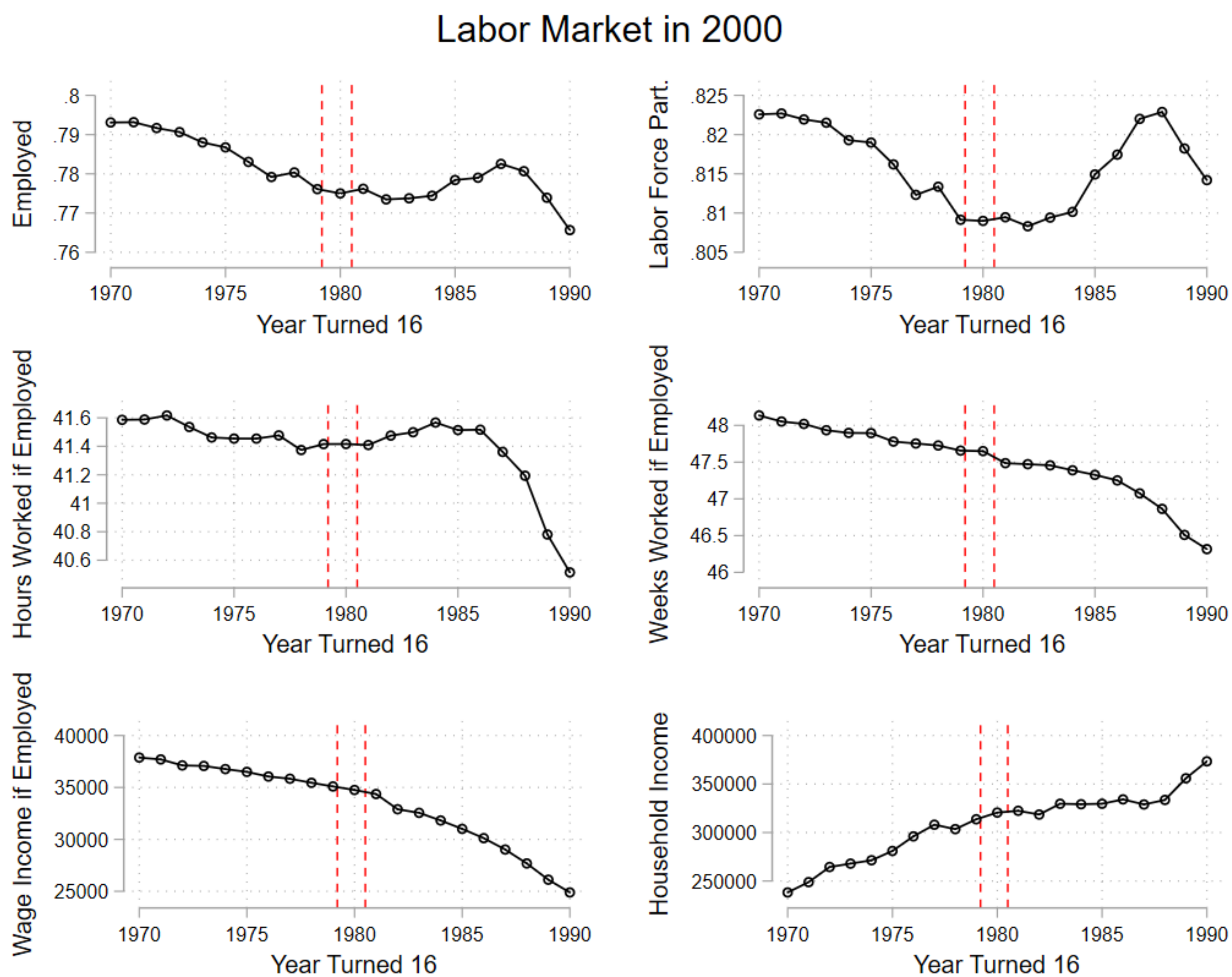

Figure A.3: Labor market characteristics in 2000 by year turned 16.

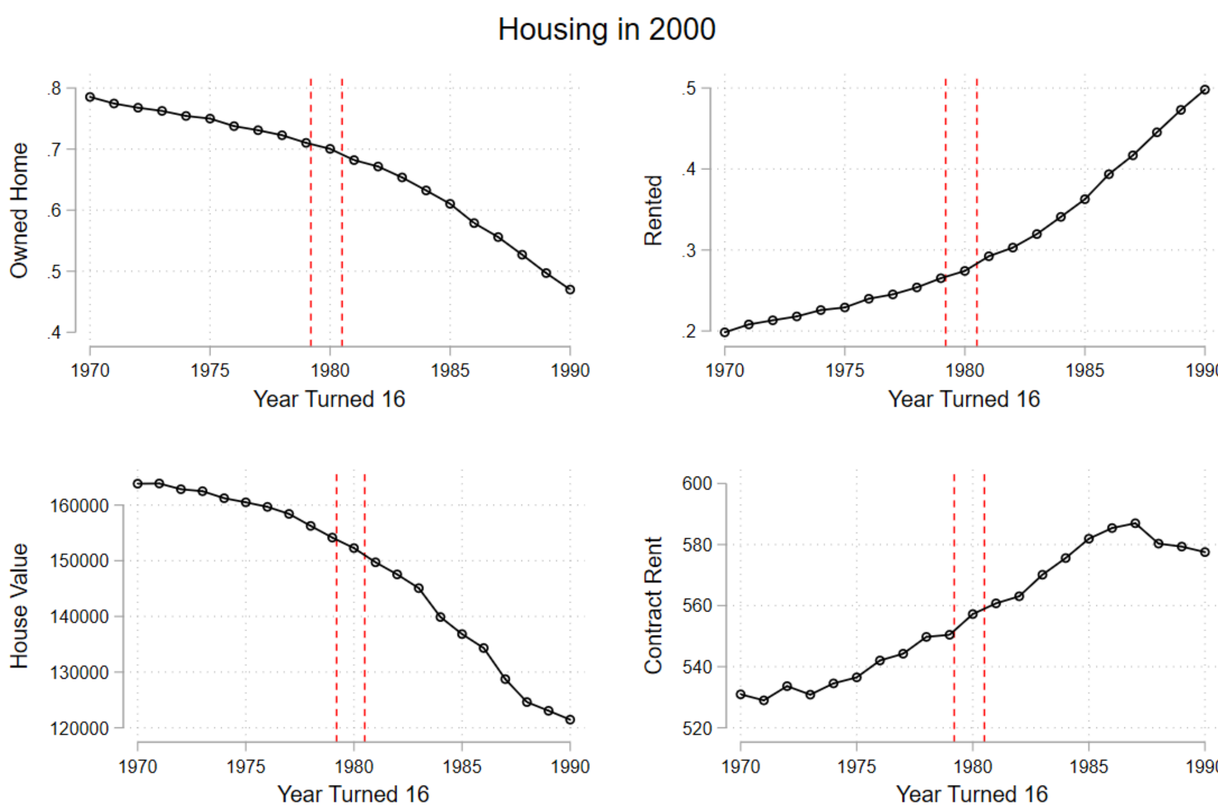

Figure A.4: Housing characteristics in 2000 by year turned 16.

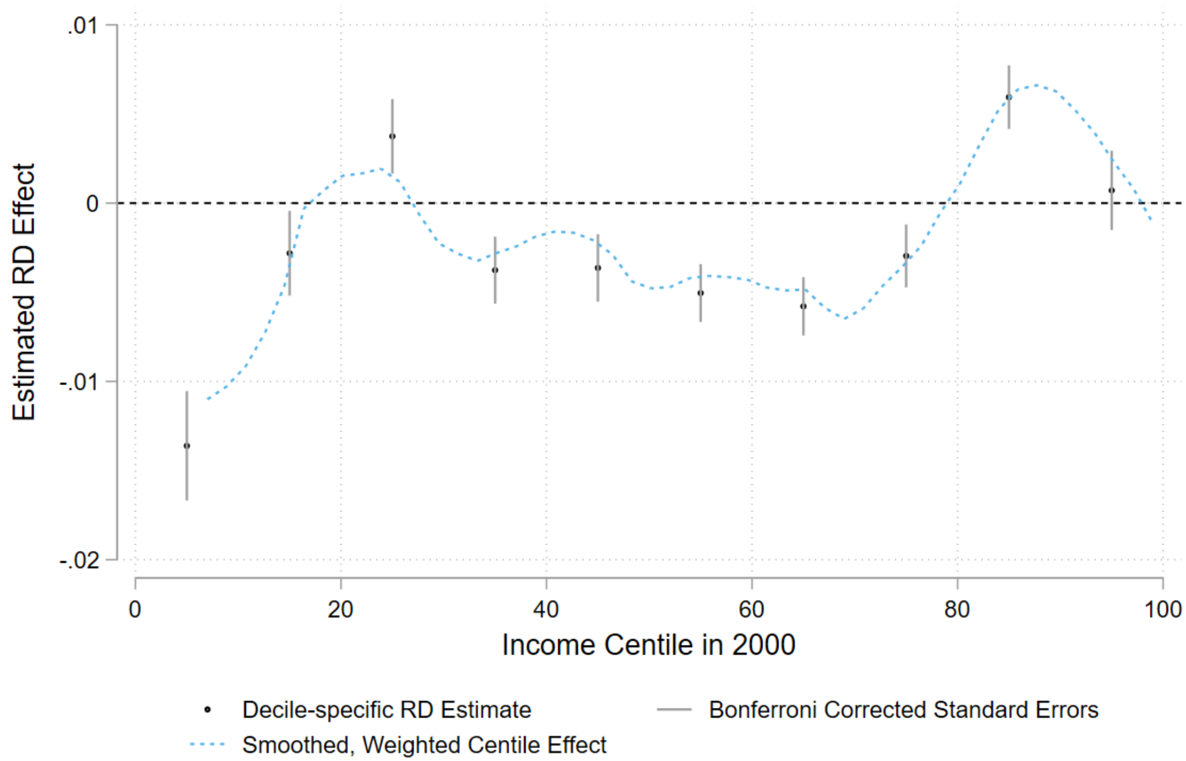

Figure A.5: Regression discontinuity estimates of the 1979-80 gas price shock on driving in 2000 by income decile and (smoothed) centile. Decile estimates shown as dots with Bonferroni-corrected 95 percent confidence intervals represented by the vertical bars (corrected for ten tests).

Table A.1: Discontinuity in turning 16 after 1980 on transportation behavior in 2000 – Donut RD omitting those who turn 16 in 1980.

		Bandwidth (years)								
Model	Poly. order	2	3	4	5	6	7	8	9	10
Panel A: Effect on driving, no controls										
	1	-0.0037	-0.0020	-0.0039*	-0.0036**	-0.0028*	-0.0029**	-0.0037**	-0.0034**	-0.0031**
		(0.0029)	(0.0020)	(0.0016)	(0.0013)	(0.0012)	(0.0011)	(0.0010)	(0.0009)	(0.0009)
	2				-0.0030	-0.0045+	-0.0035+	-0.0021	-0.0031+	-0.0037*
					(0.0028)	(0.0023)	(0.0020)	(0.0018)	(0.0016)	(0.0015)
Panel B: Effect on driving, controls:										
+ demographics	1	-0.0037	-0.0018	-0.0037*	-0.0034*	-0.0025*	-0.0025*	-0.0031**	-0.0027**	-0.0022*
		(0.0028)	(0.0019)	(0.0016)	(0.0013)	(0.0012)	(0.0011)	(0.0010)	(0.0009)	(0.0009)
	2				-0.0028	-0.0043+	-0.0036+	-0.0023	-0.0033*	-0.0039*
					(0.0027)	(0.0023)	(0.0020)	(0.0018)	(0.0016)	(0.0015)
Panel C: Effect on driving, controls:										
+ demographics, state of birth FEs	1	-0.0035	-0.0015	-0.0032*	-0.0028*	-0.0021+	-0.0020+	-0.0025*	-0.0021*	-0.0016+
		(0.0028)	(0.0019)	(0.0015)	(0.0013)	(0.0012)	(0.0011)	(0.0010)	(0.0009)	(0.0009)
	2				-0.0026	-0.0037+	-0.0031	-0.0019	-0.0028+	-0.0033*
					(0.0027)	(0.0023)	(0.0020)	(0.0018)	(0.0016)	(0.0015)
Panel D: Effect on driving, controls:										
+ demographics, state of birth FEs	1	-0.0035	-0.0015	-0.0031*	-0.0026*	-0.0020+	-0.0019+	-0.0024*	-0.0020*	-0.0015+
+ ln(*income*)		(0.0028)	(0.0019)	(0.0015)	(0.0013)	(0.0012)	(0.0011)	(0.0010)	(0.0009)	(0.0009)
	2				-0.0027	-0.0036	-0.0031	-0.0018	-0.0027+	-0.0032*
					(0.0027)	(0.0023)	(0.0020)	(0.0018)	(0.0016)	(0.0015)
N		550k	818k	1085k	1349k	1622k	1892k	1250k	2401k	2642k

Regression discontinuity estimates of the effect of turning 16 after 1980 on a binary indicator of whether the respondent drove to work, as reported in the 2000 Census. Bandwidth is symmetric around 1980, but excludes 1980 (e.g., a bandwidth of two includes 1978, 1979, 1981, and 1982). Sample includes all native-born persons actively working in the Census, and excludes farm workers and those coded N/A for transportation mode. Demographic controls include sex, race, and educational attainment. Observations weighted by person sample weights. Standard errors are robust to heteroskedasticity (see text). Sample sizes are 1-2% smaller in panels B through D. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$.

Table A.2: Discontinuity in turning 16 after 1980 on commuting behavior in 2000 – Subgroup analysis.

		Bandwidth (years)								
Model	Poly. order	2	3	4	5	6	7	8	9	10
Panel A: Effect on driving										
Sample: Principal city	1	-0.0185*	-0.0120+	-0.0108*	-0.0124**	-0.0092*	-0.0061	-0.0090*	-0.0096**	-0.0094**
		(0.0089)	(0.0065)	(0.0054)	(0.0047)	(0.0043)	(0.0039)	(0.0037)	(0.0035)	(0.0033)
	2				-0.0157+	-0.0167*	-0.0163*	-0.0087	-0.0085	-0.0096+
					(0.0085)	(0.0073)	(0.0065)	(0.0059)	(0.0055)	(0.0051)
	N	62k	92k	122k	154k	187k	220k	252k	283k	313k
Panel B: Effect on driving										
Sample: Not in metro	1	-0.0030	0.0004	0.0000	0.0013	0.0008	0.0014	0.0002	0.0003	0.0006
		(0.0042)	(0.0030)	(0.0025)	(0.0022)	(0.0020)	(0.0019)	(0.0017)	(0.0017)	(0.0016)
	2				-0.0016	0.0003	-0.0002	0.0022	0.0013	0.0006
					(0.0041)	(0.0035)	(0.0031)	(0.0028)	(0.0026)	(0.0024)
	N	114k	170k	225k	280k	336k	393k	447k	500k	552k
Panel C: Effect on driving										
Sample: Black	1	-0.0168*	-0.0099	-0.0107*	-0.0107*	-0.0067+	-0.0052	-0.0048	-0.0019	0.0002
		(0.0083)	(0.0061)	(0.0050)	(0.0045)	(0.0040)	(0.0037)	(0.0035)	(0.0033)	(0.0031)
	2				-0.0145+	-0.0176*	-0.0144*	-0.0118*	-0.0135**	-0.0136**
					(0.0080)	(0.0068)	(0.0061)	(0.0056)	(0.0052)	(0.0048)
	N	57k	84k	111k	139k	166k	193k	220k	245k	270k
Panel D: Effect on driving										
Sample: No college	1	-0.0037	-0.0017	-0.0022	-0.0027*	-0.0020+	-0.0023*	-0.0028**	-0.0023*	-0.0016+
		(0.0025)	(0.0018)	(0.0015)	(0.0014)	(0.0012)	(0.0011)	(0.0011)	(0.0010)	(0.0009)
	2				-0.0021	-0.0033	-0.0022	-0.0016	-0.0027+	-0.0036*
					(0.0025)	(0.0021)	(0.0019)	(0.0017)	(0.0016)	(0.0015)
	N	394k	585k	774k	965k	1157k	1350k	1534k	1711k	1883k

Regression discontinuity estimates of the effect of turning 16 after 1980 on a binary indicator of whether the respondent drove to work, as reported in the 2000 Census. Bandwidth is symmetric around 1980.5. Sample includes all native-born persons actively working in the Census, and excludes farm workers and those coded N/A for transportation mode. Observations weighted by person sample weights. Standard errors are robust to heteroskedasticity (see text). $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$.

Table A.3: Minimum driver licensing ages across states

Year	[14,14.5)	[14.5,15.5)	[15.5,16.5)	[16.5,17.5)	[17.5,18]	Average minimum age
Panel A: Minimum full privilege license age						
1970	1	5	38	4	3	16.37
1980	0	5	39	5	2	16.29
1990	0	5	39	5	2	16.27
2000	0	2	24	18	7	16.83
2010	0	0	4	32	15	17.23
Panel B: Minimum provisional license age						
1970	2	7	39	3	0	16.00
1980	2	7	40	2	0	15.97
1990	1	7	41	2	0	15.98
2000	1	4	41	5	0	16.05
2010	1	2	39	9	0	16.10
Panel C: Learner's permit minimum age						
1972	8	18	24	1	0	
1980	8	21	22	0	0	
1988	7	22	22	0	0	
1994	6	24	21	0	0	
2010	6	25	20	0	0	

Frequency of states (and DC) with minimum driver age in each bin for the years listed. Provisional licenses allow unaccompanied driving, but limit time of use or number of passengers. Average minimum age is weighted by state population. Learner's Permit Minimum Age is less accurately recorded and reported in FHWA data, and states vary widely in the privileges it accords. Source: see description in text and Appendix.

Table A.4: The effect of formative gasoline price on driving to work using the census/ACS 1980-2017, one year price changes, various other definitions of treatment.

	1[drive] (1)	1[drive] (2)	1[drive] (3)	1[drive] (4)	1[drive] (5)	1[drive] (6)	1[drive] (7)
$P_{cs}^{\Delta(18,16)}$	-0.0027* (0.0011)	-0.0030*** (0.0008)	-0.0030*** (0.0008)	-0.0024* (0.0010)	-0.0024* (0.0010)	-0.0023* (0.0010)	-0.0027** (0.0009)
$P_{cs}^{\Delta(18,17)}$	-0.0024 (0.0016)	-0.0038** (0.0011)	-0.0041*** (0.0011)	-0.0017 (0.0015)	-0.0017 (0.0015)	-0.0016 (0.0015)	-0.0020 (0.0014)
$P_{cs}^{\Delta(17,16)}$	-0.0038** (0.0014)	-0.0030* (0.0012)	-0.0036** (0.0012)	-0.0036* (0.0013)	-0.0037** (0.0013)	-0.0037** (0.0013)	-0.0041** (0.0012)
$P_{cs}^{\Delta(16,15)}$	-0.0054** (0.0016)	-0.0039** (0.0013)	-0.0046** (0.0014)	-0.0053** (0.0016)	-0.0056*** (0.0016)	-0.0056** (0.0017)	-0.0061*** (0.0015)
$P_{cs}^{\Delta(m_{cs}+2,m_{cs})}$	-0.0034** (0.0012)	-0.0043*** (0.0009)	-0.0041*** (0.0010)	-0.0031** (0.0012)	-0.0032** (0.0011)	-0.0032** (0.0012)	-0.0037** (0.0011)
$P_{cs}^{\Delta(m_{cs}+2,m_{cs}+1)}$	-0.0036* (0.0017)	-0.0049** (0.0014)	-0.0050** (0.0014)	-0.0030+ (0.0016)	-0.0033* (0.0016)	-0.0029+ (0.0016)	-0.0035* (0.0015)
$P_{cs}^{\Delta(m_{cs}+1,m_{cs})}$	-0.0044* (0.0018)	-0.0051*** (0.0014)	-0.0054*** (0.0015)	-0.0041* (0.0018)	-0.0042* (0.0018)	-0.0044* (0.0018)	-0.0048** (0.0017)
$P_{cs}^{\Delta(m_{cs},m_{cs}-1)}$	-0.0048*** (0.0013)	-0.0038** (0.0013)	-0.0046*** (0.0012)	-0.0048** (0.0014)	-0.0047** (0.0014)	-0.0049** (0.0014)	-0.0052*** (0.0013)
Census year FEs	Y	Y	Y	Y	Y	-	-
State of birth FEs	Y	Y	Y	Y	Y	-	-
Age FEs	Y	Y	Y	Y	Y	Y	Y
Demographics	-	-	-	Y	Y	Y	Y
ln HH income	-	-	-	-	Y	Y	Y
State-X-year FEs	-	-	-	-	-	Y	Y
Quad. birth year	-	-	-	-	-	-	Y
Price in state of	Birth	Birth	Res	Birth	Birth	Birth	Birth
Sample	Stay	All	All	Stay	Stay	Stay	Stay

Each row and column represents the results from a different regression, for fifty-six total. Dependent variable is a binary indicator of whether the respondent drove to work, as reported in the Census. Sample includes all native-born persons actively working in the Census between the ages of 25-54, and excludes farm workers and those coded N/A for transportation mode. Demographics include sex, marital status, educational attainment, and race. Observations weighted by person sample weights. Standard errors clustered by state of birth. + p<0.10, * p<0.05, ** p<0.01, *** p<0.001.

Table A.5: The effect of formative gasoline price on driving to work using the census/ACS 1980-2017, with cohort FEs.

	1[drive] (1)	1[drive] (2)	1[drive] (3)	1[drive] (4)
2-year price change				
$P_{cs}^{\Delta(m_{cs}+2,m_{cs})}$	-0.0041+	-0.0039+	-0.0038+	-0.0037+
	(0.0023)	(0.0021)	(0.0021)	(0.0020)
$P_{cs}^{\Delta(m_{cs}+1,m_{cs}-1)}$	-0.0016	-0.0016	-0.0012	-0.0017
	(0.0019)	(0.0019)	(0.0019)	(0.0019)
1-year price change				
$P_{cs}^{\Delta(m_{cs}+2,m_{cs}+1)}$	-0.0057*	-0.0053*	-0.0054*	-0.0048*
	(0.0024)	(0.0022)	(0.0021)	(0.0021)
$P_{cs}^{\Delta(m_{cs}+1,m_{cs})}$	-0.0019	-0.0018	-0.0016	-0.0019
	(0.0025)	(0.0025)	(0.0025)	(0.0025)
$P_{cs}^{\Delta(m_{cs},m_{cs}-1)}$	-0.0009	-0.0009	-0.0004	-0.0008
	(0.0024)	(0.0023)	(0.0024)	(0.0024)
Levels				
$P_{cs}^{m_{cs}}$	-0.0013	-0.0015	-0.0020	-0.0022
	(0.0026)	(0.0024)	(0.0024)	(0.0019)
Census year FEs	Y	Y	Y	Y
State of birth FEs	Y	Y	Y	Y
Age FEs	Y	Y	Y	Y
Birth year FEs	Y	Y	Y	Y
Demographics	-	Y	Y	Y
ln HH income	-	-	Y	Y
State-X-year FEs	-	-	-	Y

Each row and column represents the results from a different regression, for twenty-four total. Dependent variable is a binary indicator of whether the respondent drove to work, as reported in the Census. Sample includes all native-born persons actively working in the Census between the ages of 25-54, and excludes farm workers and those coded N/A for transportation mode. Demographics include sex, marital status, educational attainment, and race. Observations weighted by person sample weights. Standard errors clustered by state of birth. + p<0.10, * p<0.05, ** p<0.01, *** p<0.001.

Table A.6: The effect of formative gasoline price on log miles traveled using NHTS 1990-2017, one year price changes, various other definitions of treatment.

	ln(VMT) (1)	ln(VMT) (2)	ln(VMT) (3)	ln(VMT) (4)	ln(VMT) (5)
$P_{cs}^{\Delta(18,16)}$	-0.0807**	-0.0851***	-0.0709**	-0.0719**	-0.0524*
	(0.0236)	(0.0237)	(0.0246)	(0.0249)	(0.0226)
$P_{cs}^{\Delta(18,17)}$	-0.0954*	-0.1023**	-0.0802+	-0.0781+	-0.0484
	(0.0374)	(0.0380)	(0.0404)	(0.0406)	(0.0366)
$P_{cs}^{\Delta(17,16)}$	-0.1125**	-0.1171**	-0.1060*	-0.1097**	-0.0837*
	(0.0401)	(0.0403)	(0.0409)	(0.0407)	(0.0398)
$P_{cs}^{\Delta(16,15)}$	-0.0949*	-0.0995*	-0.0953*	-0.0920*	-0.0760+
	(0.0428)	(0.0415)	(0.0409)	(0.0401)	(0.0402)
$P_{cs}^{\Delta(m_{cs}+2,m_{cs})}$	-0.0525*	-0.0540*	-0.0403	-0.0405	-0.0212
	(0.0229)	(0.0231)	(0.0244)	(0.0248)	(0.0228)
$P_{cs}^{\Delta(m_{cs}+2,m_{cs}+1)}$	-0.0678+	-0.0693+	-0.0561	-0.0541	-0.0229
	(0.0346)	(0.0350)	(0.0393)	(0.0396)	(0.0361)
$P_{cs}^{\Delta(m_{cs}+1,m_{cs})}$	-0.0618+	-0.0644+	-0.0462	-0.0484	-0.0279
	(0.0343)	(0.0357)	(0.0353)	(0.0358)	(0.0347)
$P_{cs}^{\Delta(m_{cs},m_{cs}-1)}$	-0.0606+	-0.0743*	-0.0699+	-0.0690+	-0.0531
	(0.0350)	(0.0338)	(0.0352)	(0.0347)	(0.0349)
Sample year FEs	Y	Y	Y	-	-
State FEs	Y	Y	Y	-	-
Age FEs	Y	Y	Y	Y	Y
Controls	-	Y	Y	Y	Y
Income-by-year bin FEs	-	-	Y	Y	Y
State-X-year FEs	-	-	-	Y	Y
Quad. birth year	-	-	-	-	Y

Each row and column represents the results from a different regression, for twenty total. Dependent variable is log person VMT. Sample includes all respondents aged 25-54 with positive person VMT. Demographics include race, urbanization, and family size. Observations weighted by person sample weights. Standard errors clustered by state. + p<0.10, * p<0.05, ** p<0.01, *** p<0.001.

Table A.7: The effect of formative gasoline price on log miles traveled using NHTS 1990-2017, with cohort FEs.

	ln(VMT) (1)	ln(VMT) (2)	ln(VMT) (3)	ln(VMT) (4)
2-year price change				
$P_{cs}^{\Delta(m_{cs}+2,m_{cs})}$	0.0363	0.0432	0.0588+	0.0608+
	(0.0341)	(0.0344)	(0.0330)	(0.0323)
$P_{cs}^{\Delta(m_{cs}+1,m_{cs}-1)}$	0.0475	0.0383	0.0394	0.0377
	(0.0396)	(0.0393)	(0.0369)	(0.0355)
1-year price change				
$P_{cs}^{\Delta(m_{cs}+2,m_{cs}+1)}$	0.0088	0.0167	0.0343	0.0369
	(0.0401)	(0.0406)	(0.0438)	(0.0432)
$P_{cs}^{\Delta(m_{cs}+1,m_{cs})}$	0.0604	0.0639	0.0741	0.0752
	(0.0533)	(0.0539)	(0.0530)	(0.0508)
$P_{cs}^{\Delta(m_{cs},m_{cs}-1)}$	0.0338	0.0129	-0.0003	-0.0028
	(0.0487)	(0.0453)	(0.0436)	(0.0418)
Levels				
$P_{cs}^{m_{cs}}$	0.0091	0.0012	-0.0076	-0.0129
	(0.0326)	(0.0316)	(0.0323)	(0.0313)
Sample year FEs	Y	Y	Y	-
State FEs	Y	Y	Y	-
Age FEs	Y	Y	Y	Y
Birth year FEs	Y	Y	Y	Y
Controls	-	Y	Y	Y
Income-by-year bin FEs	-	-	Y	Y
State-X-year FEs	-	-	-	Y

Each row and column represents the results from a different regression, for twenty total. Dependent variable is log person VMT. Sample includes all respondents aged 25-54 with positive person VMT. Demographics include race, urbanization, and family size. Observations weighted by person sample weights. Standard errors clustered by state. + p<0.10, * p<0.05, ** p<0.01, *** p<0.001.

Table A.8: The effect of gasoline prices changes at different ages (two-year difference).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$a =$	13	14	15	16	17	18	19	20	21	22
$\tau =$	-3	-2	-1	0	1	2	3	4	5	6
Panel A: Extensive margin (1[drive])										
$P_{cs}^{\Delta a,(a-2)}$	-0.0018*	0.0004	0.0004	-0.0022*	-0.0038***	-0.0026*	-0.0016	-0.0003	0.0002	0.0014+
	(0.0009)	(0.0010)	(0.0008)	(0.0010)	(0.0010)	(0.0010)	(0.0010)	(0.0011)	(0.0008)	(0.0008)
$P_{cs}^{\Delta(m_{cs}+\tau,m_{cs}+\tau-2)}$	-0.0005	-0.0002	-0.0020+	-0.0030*	-0.0041***	-0.0034**	-0.0014	0.0008	0.0006	-0.0004
	(0.0007)	(0.0007)	(0.0010)	(0.0012)	(0.0010)	(0.0012)	(0.0014)	(0.0010)	(0.0009)	(0.0011)
Panel B: Intensive margin (ln(person VMT))										
$P_{cs}^{\Delta a,(a-2)}$	-0.0253	-0.0156	0.0222	-0.0287	-0.0786**	-0.0807**	-0.0532*	-0.0124	-0.0087	-0.0174
	(0.0316)	(0.0215)	(0.0264)	(0.0260)	(0.0264)	(0.0236)	(0.0218)	(0.0272)	(0.0227)	(0.0234)
$P_{cs}^{\Delta(m_{cs}+\tau,m_{cs}+\tau-2)}$	-0.0309	-0.0289	-0.0096	-0.0309	-0.0502*	-0.0525*	-0.0511*	-0.0261	-0.0146	0.0011
	(0.0232)	(0.0205)	(0.0278)	(0.0256)	(0.0193)	(0.0229)	(0.0228)	(0.0255)	(0.0233)	(0.0228)

Each row and column represents the results from a different regression, for forty total. Dependent variable is a binary indicator of whether the respondent drove to work in the Census data, and log person VMT in the NHTS sample. Regressions include state (or state of birth), sample year, and age fixed effects. Observations weighted by person sample weights. Standard errors clustered by state of birth. + p<0.10, * p<0.05, ** p<0.01, *** p<0.001.

Table A.9: The effect of formative gasoline price on vehicle efficiency and type.

	Gallons per mile				Truck, SUV, etc.			
	Average GPM (1)	Average GPM (2)	GPM (3)	GPM (4)	Any Big (5)	Any Big (6)	1[Big] (7)	1[Big] (8)
$P_{cs}^{\Delta(18,16)}$	-0.0000	-0.0001	-0.0001	-0.0001	-0.0265**	-0.0245*	-0.0193*	-0.0194+
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0095)	(0.0101)	(0.0092)	(0.0097)
$P_{cs}^{\Delta(17,15)}$	0.0000	-0.0002	-0.0002	-0.0003	-0.0213+	-0.0173	-0.0155	-0.0141
	(0.0003)	(0.0003)	(0.0002)	(0.0002)	(0.0111)	(0.0112)	(0.0106)	(0.0104)
$P_{cs}^{\Delta(m_{cs}+2,m_{cs})}$	0.0001	0.0001	-0.0001	-0.0000	-0.0203*	-0.0169+	-0.0141	-0.0110
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0090)	(0.0085)	(0.0094)	(0.0085)
$P_{cs}^{\Delta(m_{cs}+1,m_{cs}-1)}$	-0.0002	-0.0003	-0.0003	-0.0004	-0.0238+	-0.0209	-0.0193	-0.0179
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0126)	(0.0125)	(0.0117)	(0.0116)
Sample year FEs	Y	-	Y	-	Y	-	Y	-
State FEs	Y	-	Y	-	Y	-	Y	-
Age FEs	Y	Y	Y	Y	Y	Y	Y	Y
Demographics	-	Y	-	Y	-	Y	-	Y
Income-by-year bin FEs	-	Y	-	Y	-	Y	-	Y
State-X-year FEs	-	Y	-	Y	-	Y	-	Y
Vehicle age	-	-	Y	Y	-	-	Y	Y
Quad. vehicle year	-	-	Y	Y	-	-	Y	Y
Sample	Person	Person	Vehicle	Vehicle	Person	Person	Vehicle	Vehicle
Mean of dep. var.	0.0508	0.0508	0.0509	0.0509	0.4681	0.4681	0.4422	0.4422

Each row and column represents the results from a different regression, for thirty-two total. Dependent variable in Columns (1) to (4) is a measure of fuel economy in gallons per mile, and in Columns (5) to (8) is an indicator for a large vehicle (larger than a station wagon). Columns (1), (2), (5) and (6) treat people as the level of observation; other columns treat vehicles as the level of observation. Demographics include race, urbanization, and family size. Observations weighted by person sample weights. Standard errors clustered by state. + p<0.10, * p<0.05, ** p<0.01, *** p<0.001.

Table A.10: Mediation analysis of joint effect of gasoline price shocks and income.

	(1)	(2)	(3)	(4)	(5)	(6)
Effects of M and T on Y	1[drive]	1[drive]	1[drive]	1[drive]	1[drive]	1[drive]
θ^Y	-0.0038***	-0.0041***	-0.0032**	-0.0037**	-0.0031**	-0.0037**
	(0.0010)	(0.0011)	(0.0009)	(0.0010)	(0.0011)	(0.0012)
γ	0.0223***	0.0223***	0.0170***	0.0170***	0.0216***	0.0216***
	(0.0024)	(0.0024)	(0.0045)	(0.0045)	(0.0044)	(0.0045)
Effect of T on M	ln(Y)	ln(Y)	ln(Y)	ln(Y)	ln(Y)	ln(Y)
θ^M	-0.0053	-0.0062+	-0.0488***	-0.0371***	-0.0460***	-0.0335***
	(0.0034)	(0.0036)	(0.0034)	(0.0034)	(0.0035)	(0.0033)
Direct effect (θ^Y)	-0.0038***	-0.0041***	-0.0032**	-0.0037**	-0.0031**	-0.0037**
	(0.0010)	(0.0011)	(0.0009)	(0.0010)	(0.0011)	(0.0012)
Indirect effect ($\gamma\theta^M$)	-0.0001	-0.0001	-0.0008**	-0.0006**	-0.0010***	-0.0007***
	(0.0001)	(0.0001)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Total effect ($\theta^Y + \gamma\theta^M$)	-0.0040***	-0.0042***	-0.0040***	-0.0043***	-0.0041***	-0.0044***
	(0.0010)	(0.0010)	(0.0008)	(0.0043)	(0.0010)	(0.0010)
Treatment definition (T)	$P_{cs}^{\Delta 17,15}$	$P_{cs}^{\Delta(m_{cs}\pm 1)}$	$P_{cs}^{\Delta 17,15}$	$P_{cs}^{\Delta(m_{cs}\pm 1)}$	$P_{cs}^{\Delta 17,15}$	$P_{cs}^{\Delta(m_{cs}\pm 1)}$
Income definition (M)	Household income		Wage income		Personal income	

See Appendix A.2 for details. Dependent variable is a binary indicator of whether the respondent drove to work, as reported in the Census. Sample includes all native-born persons actively working in the Census between the ages of 25-54, and excludes farm workers and those coded N/A for transportation mode. All models include age, state of birth, and sample year fixed effects. Demographics include sex and race. Observations weighted by person sample weights. Standard errors clustered by state of birth. Income is modeled in logs. $P_{cs}^{\Delta(m_{cs}\pm 1)}$ is equivalent to $P_{cs}^{\Delta(m_{cs}+1,m_{cs}-1)}$. + p<0.10, * p<0.05, ** p<0.01, *** p<0.001.