Dynamic heterogeneity: rational habits and the heterogeneity of household responses to gasoline prices

Aurélien Saussay[a,1]

[a] *Grantham Research Institute, London School of Economics, London*
[b] *OFCE, Sciences Po, Paris*

Abstract

The heterogeneity of household response to gasoline prices has key implications for the distributional impacts of gasoline taxation. However, this heterogeneity has mostly been assessed in a static framework, which ignores the dynamic nature of gasoline consumption. I contribute to this debate by developing a simple rational habits model of gasoline consumption, which allows to assess both rigidities on households' response to contemporaneous gasoline prices and forward-looking behavior *vis-à-vis* future gasoline prices. The parsimonious nature of this model makes it amenable to estimation on long-run household panel data, which allows the analysis of long-term responses. Estimation on the PSID panel dataset for the period 1999-2015 yields a long-term price elasticity of -0.88. Interactions with quintiles of income reveals significant heterogeneity in households long-term response across the income distribution: poorer households' gasoline consumption exhibits stronger habits, while richer households are more forward-looking. These findings suggest that policies fostering gasoline price increases should be complemented with measures facilitating the adaptation of poorer households' gasoline demand.

Keywords: Gasoline demand, Rational habits, Price elasticity, Heterogeneity

December 15th, 2019

1. Introduction

Personal private vehicles accounted for 19% of total U.S. CO_2 emissions in 2016. This is the third largest source of GHG emissions in the U.S., rivaled only by the power and industrial sectors. Tackling this source of emissions is therefore critical to climate change mitigation in the U.S. The overwhelming majority of these emissions results from the combustion of gasoline. The imposition of a Pigouvian tax on gasoline consumption has very often been suggested in the literature as an effective policy tool to reduce its associated GHG emissions (Sterner, 2007; Ross et al., 2017).

However, increasing the price of gasoline – either through carbon taxation, gasoline specific taxation or reduction in fossil fuel subsidies – raises a number of policy issues. In particular, gasoline price increases may affect poorer households disproportionately. Households on the lower end of the income scale dedicate a larger share of their budget to gasoline expenditure than wealthier households, leading to a tax burden inversely proportional to household's income. This effect would make gasoline taxation regressive (Poterba, 1991). Obviously this reasoning only applies to the gross distributional impacts of gasoline taxation. Recycling of the tax receipts, which I will not discuss in the present article, can significantly reduce this regressivity of gasoline taxation, or even make it progressive under certain schemes – see Combet et al. (2010) and Berry (2019) for recent discussions.

The distributional consequences of gasoline price increases have been examined repeatedly in the literature (Poterba, 1991; Metcalf, 1999; West, 2004; Sterner, 2012; West and Williams, 2012), and involve identifying the heterogeneity of households' responses to gasoline price variations. When this heterogeneity is taken into account, the regressivity of gasoline taxation appears more limited (West, 2004). Further, if permanent income is considered, it may even be close to inexistent (Sterner, 2012).

However, most of these studies have been conducted in the frame of a static model. By construction, this type of model ignores the dynamic nature of gasoline consumption, which involves a double decision: first a discrete choice to invest in a vehicle bundle, then a continuous choice to consume gasoline, given the vehicle(s) available to produce the personal transportation service (Mannering and Winston, 1985). This discrete-continuous process implies an intertemporal dependence in gasoline consumption.

The dynamics of this process can be a source of additional heterogeneity. It is well established that household's gasoline price elasticity vary with income (Yatchew and No, 2001; West, 2004; Wadud et al., 2010b; Blundell et al., 2012). In addition, households may face rigidities in adapting their gasoline consumption to changing price conditions – rigidities which may also be heterogeneous across the income distribution. For instance, modifying their vehicle bundle through the purchase of a new car may be difficult for liquidity or credit constrained households (Attanasio et al., 2008). More generally, parameters such as the distance between home and workplace or the availability of other transportation modes cannot be changed easily in response to gasoline price variations.

This article seeks to address the following research questions: are rigidities to gasoline consumption adaption heterogeneous across households? How does this affect household's long term response to gasoline price increases? How does this potential dynamic heterogeneity affect the regressivity of gasoline price increases?

Answering the first question requires to first construct a dynamic model of gasoline consumption. One possible venue explored in the literature is the use of a discrete-continuous model (West, 2004; Fang, 2008; Gillingham, 2011; Spiller, 2012). Unfortunately, these models are very data-intensive – in particular they require detailed knowledge of households' vehicle bundles over time. The characteristics of interest include *e.g.* their make, model, size or number of cylinders. While this can be obtained over specific regions[1] in panel form over a short period of time, this level of detail is not available over decadal time spans. Further, there is currently no household-level panel dataset that collects vehicle bundle information with the

[1]In the US, see for example Gillingham (2011) and Spiller (2012) for such datasets in the state of California.

required level of precision at the national level over a prolonged time span. This poses a problem for the second question: estimating long-term responses necessitates a commensurately long period of observation.

I therefore develop a simple dynamic model of gasoline consumption. Following the contributions of Scott (2012) and Filippini et al. (2018), I build on the rational habits model of Becker et al. (1994). This model captures the intertemporal dimension of gasoline demand through a simple functional form linking present consumption to its past and future levels.

One key advantage of this model is that it is amenable to estimation on a long-run household panel, provided the panel surveys gasoline consumption. This type of panels is increasingly available around the world, which makes this contribution particularly relevant. In the U.S., I can make use of the 1999-2015 period from the longest running household panel survey in the world, the Panel Study of Income Dynamics (PSID), to estimate the model.

I then estimate a modified specification including income quintiles interaction terms to analyze how habits formation, forward looking behavior and response in gasoline consumption to gasoline price vary across the income distribution. I then derive long-term price elasticity estimates from these results.

In the last section, I apply parameters estimates for each quintile to the households sample to conduct a micro-simulation of the distributional impacts of an increase in gasoline price corresponding to a \$50/tCO_2 carbon tax. I derive these impacts both in terms of consumer surplus variation and tax burden by quintile, and use the latter to compute a Suits index to measure its level of regressivity. I then use the dynamic features of the gasoline consumption model to explore whether this level of regressivity has changed over time over the observation period.

The article is organized as follows: I first present the theoretical framework in which the rational habits model of gasoline consumption is derived. I then present the consumption data obtained from the PSID and the localized gasoline price variable constructed from the

Council for Community and Economic Research's (C2ER) Cost of Living Index (COLI). Section 4 outlines the econometric challenges raised by the estimation of the rational habits model and how to address them. In section 5, I present results from the model estimation, both in aggregate and by quintile of income, in addition to a number of robustness checks. Then, using the parameters estimates obtained for each quintile, I conduct an analysis of the distributional impact of a gasoline price increase in section 6, and discuss the interaction between its level of regressivity and dynamic household behavior. Finally I conclude.

2. Theoretical framework

2.1. Rational habits model of gasoline consumption

In this section, I introduce a simple dynamic model of household gasoline consumption. Static household-level models of gasoline demand, which have been commonly used in the literature examining distributional aspects of gasoline price increases (Metcalf, 1999; Wadud et al., 2010b; Sterner, 2012), make the implicit assumption that households can adjust their gasoline consumption quickly to gasoline price shocks, thereby focusing on households' short-term response.

Yet, a number of potential rigidities may slow down households' responses. First is the composition of the household's vehicles bundle. Replacing an older model with a newer, potentially smaller or more efficient one cannot be achieved instantly and may be hampered by liquidity constraints or limited access to credit (Attanasio et al., 2008). Households who own multiple vehicles may suffer less stringently from this rigidity, since they can adjust their demand immediately when faced with higher gasoline prices by increasing the number of miles driven on their most efficient vehicle (Bomberg and Kockelman, 2007).

Households' existing vehicle stock need not be the only factor affecting the swiftness of their response. Other factors can include the distance between home and workplace, the availability of other transportation modes, or even the driving style of household members – in particular, it has been estimated that so-called "eco-driving", an assortment of driving best practices, can reduce gasoline consumption on any given car by up to 10% (Barkenbus, 2010).

Dynamic partial adjustment models of gasoline consumption take into account this possible delay in households' reactions to gasoline price fluctuation (Alberini and Filippini, 2011). However households in this class of model are myopic: anticipations about future gasoline prices are not considered to play a role in households' consumption decisions, which thus cannot by construction exhibit any forward-looking behavior.

In the following, I introduce a simple dynamic model of household gasoline consumption incorporating these two features, dynamic consumption adjustment and forward-looking behavior. Following Scott (2012)'s initial foray into this approach and Filippini et al. (2018)'s analysis of residential electricity demand, I build on Becker et al. (1994) to specify a rational habits model of gasoline consumption.

In order to represent the dynamics of gasoline consumption over time, it is necessary to relax the time-separability of household utility. In the following, I consider a simple specification u for non-time-separable utility amenable to display habit persistence (Constantinides, 1990):

$$u = f(g_t, g_{t-1}, c_t, x_t) \tag{1}$$

In equation (1), u is a function of present and past gasoline consumption g_t and g_{t-1}, consumption of a composite good c_t and household characteristics x_t. Households are assumed to maximize the sum of their discounted utility over their lifetime and to be infinite-lived.

Considering the composite good as numeraire, households thus solve:

$$\begin{aligned} &\max \sum_{t=1}^{\infty} \delta^{t-1} u(g_t, g_{t-1}, c_t; x_t) \\ &s.t. \quad \begin{cases} g_0 = G_0 \\ \sum_{t=1}^{\infty} \delta^{t-1}(c_t + p_t g_t) = W \end{cases} \end{aligned} \tag{2}$$

with $\delta < 1$ the rate of preference for the present ($\delta = \frac{1}{1+r}$ with r the discount rate), G_0 the initial level of gasoline consumption, p_t the price of gasoline and W the household's present value of wealth.

Problem (2) implies that households' gasoline consumption path (and therefore their consumption at any given time t) depends on the complete sequence of gasoline prices. Note however, that following Filippini et al. (2018) I do not assume perfect foresight. More realistically, I simply consider that households will take into account their expectations about future consumption and future prices when deciding their current level of gasoline consumption. This does not preclude them from being boundedly rational or even myopic. Indeed, as will be shown below, the solution to this optimization problem encompasses the full gamut of forecast horizons, from myopic to perfect foresight.

The formulation of problem (2) is similar to Becker et al. (1994)'s model of rational addiction, and can therefore be solved in the same way. Becker et al. (1994) show that if we now consider a quadratic specification in c_t, g_t, g_{t-1} and x_t for utility u, the first order conditions of problem (2) yield:

$$g_t = \theta g_{t-1} + \delta\theta g_{t+1} + \theta_1 p_t + \theta_2 x_t + \theta_3 x_{t+1} \tag{3}$$

Following Baltagi and Griffin (2001, 2002) and Laporte et al. (2010), I further simplify equation (3) to yield the canonical formulation of the rational habits model of gasoline consumption[2]:

$$g_t = \alpha_p g_{t-1} + \alpha_f g_{t+1} + \beta p_t + \gamma x_t \tag{4}$$

[2]In practice, the omission of x_{t+1} does not modify the interpretation of the model. Equation (7) shows that the model implies that present consumption is a function of the sequence of anticipated future prices and household characteristics. The inclusion of x_{t+1} would simply change the summation term from $\sum_{i=1}^{\infty} \beta p_{t+i} + \gamma x_{t+i}$ to $\sum_{i=1}^{\infty} \beta p_{t+i} + \gamma x_{t+i} + \gamma' x_{t+i+1}$. This would not change the conclusions, while burdening the notations unnecessarily.

This model implies that current gasoline consumption is a function of past and future gasoline consumption, and of present gasoline price and household characteristics. The dependence on past consumption α_p represents the impact that habits formation may have on gasoline consumption. As such, habits provide a parsimonious device to capture the various rigidities on gasoline consumption adjustments highlighted previously in this section.

Conversely, the dependence on future gasoline consumption α_f results from the rational forward-looking behavior of households in problem (2). Implicitly, since future gasoline consumption depends on the future trajectory of gasoline prices, this implies that households take into account their expectations about future gasoline prices when choosing their current gasoline consumption. Note however that as mentioned above, this model encompasses a large range of household forward-looking behaviors. In particular, as $\alpha_f \to 0$, households become increasingly myopic. In the limit case, equation (4) is reduced to a fully myopic dynamic partial adjustment model.

2.2. Short and long-term gasoline price elasticities

Additional manipulation of equation (4) yields further understanding of the model. The rational habits model is a second-order difference equation, and can be rewritten using the lag operator L:

$$g_t = \alpha_p L\, g_t + \alpha_f L^{-1}\, g_t + \beta p_t + \gamma x_t \tag{5}$$

This in turn can be formulated as a second order lag polynomial (Laporte et al., 2010):

$$\left(1 - \frac{1}{\alpha_f} L + \frac{\alpha_p}{\alpha_f} L^2\right) g_t = -\frac{1}{\alpha_f} L \left(\beta p_t + \gamma x_t\right) \tag{6}$$

After factorization, this can be rewritten in the following form[3]:

[3]See proof in Appendix A.1.

$$g_t = \phi_1 g_{t-1} + \frac{1}{\alpha_f \phi_2} \sum_{\mathrm{i}\,=\,0}^{\infty} \frac{1}{\phi_2^i} \left(\beta p_{t+i} + \gamma x_{t+i}\right) \tag{7}$$

where $0 < \phi_1 < 1, \phi_2 > 1$ are the roots of the polynomial defined on the left-hand side of equation (6):

$$\phi_1 = \frac{1 - \sqrt{1 - 4\alpha_p \alpha_f}}{2\alpha_f}, \quad \phi_2 = \frac{1 + \sqrt{1 - 4\alpha_p \alpha_f}}{2\alpha_f} \tag{8}$$

Equation (7) provides several useful insights on model (4). First, it allows to derive formulas for the short-run gasoline price elasticity, σ_{short}, defined as the short-term response to permanent change in gasoline price[4]:

$$\sigma_{short} = \frac{2\beta}{1 - 2\alpha_f + \sqrt{1 - 4\alpha_p \alpha_f}} \frac{\overline{p}}{\overline{g}} \tag{9}$$

and its long-run counterpart, σ_∞, defined as the long-term response to a permanent change in gasoline price:

$$\sigma_\infty = \frac{\beta}{1 - \alpha_p - \alpha_f} \frac{\overline{p}}{\overline{g}} \tag{10}$$

where $\overline{p}$ and $\overline{g}$ are evaluated at their respective sample mean.

2.3. Gasoline demand response to a price change

Equation (7) yields a demand equation for a constant trajectory of gasoline prices and household characteristics. Holding p_t and x_t constant for all $t' \geq t$, we get[5]:

$$\begin{aligned} g_t &= \phi_1 g_{t-1} + \frac{\beta p_t + \gamma x_t}{\alpha_f \left(\phi_2 - 1\right)} \\ g_\infty &= \frac{\beta p_t + \gamma x_t}{\alpha_f \left(\phi_2 - 1\right)\left(1 - \phi_1\right)} \end{aligned} \tag{11}$$

These demand equations allow us to derive both the short-term and long-term responses

[4]See proof in Appendix A.3.
[5]See proof in Appendix A.2.

of gasoline demand to a one-time gasoline price increase Δp occurring at time $t+1$ and sustained thereafter, while still holding x_t constant. If we write $\Delta g_{t+1} \equiv g_{t+1} - g_t$ and $\Delta g_\infty \equiv g_\infty - g_t$ (with some abuse of notation), we have:

$$\begin{aligned} \Delta g_{t+1} &= \phi_1 \Delta g_t + \frac{\beta \Delta p}{\alpha_f (\phi_2 - 1)} \\ \Delta g_\infty &= \frac{\Delta g_{t+1}}{1 - \phi_1} \end{aligned} \tag{12}$$

To guide intuition in analyzing these results, it is useful to note that if α_p is sufficiently small[6], the first order Taylor expansion of ϕ_1 and ϕ_2 in α_p is:

$$\phi_1 \approx \alpha_p, \quad \phi_2 \approx \frac{1}{\alpha_f} - \alpha_p \tag{13}$$

Substituting in (12), this gives:

$$\begin{aligned} \Delta g_{t+1} &\approx \alpha_p \Delta g_t + \frac{\beta \Delta p}{1 + \alpha_f (\alpha_p - 1)} \\ \Delta g_\infty &\approx \frac{\Delta g_{t+1}}{1 - \alpha_p} \end{aligned} \tag{14}$$

The short-term response is thus composed of an inertia term and and a price response term. The magnitude of the inertia term $\alpha_p \Delta g_t$ is a function of habits strength α_p. Interestingly, its sign is not related to the sign of Δp. Hence if a sudden increase occurs after a period of moderate prices that encourage growing gasoline consumption, households with strong habits will carry on some of the momentum in their gasoline demand, which will dampen their short-term response to the price increase.

The price response term is a function of the dynamic coefficients and of gasoline price coefficient β. Unsurprisingly, it is directly proportional to β. Further, since we have $\alpha_p + 1 > 0$, the magnitude of the response term is also proportional to α_f: the more forward looking a household, the stronger its short-term price response will be. Interestingly, stronger habits

[6]The Taylor expansion of ϕ_1 and ϕ_2 in α_p at zero is convergent iif $|\alpha_p| < 1$, which is necessarily the case to ensure that the second-order difference equation defining g_t, equation (4), is not divergent.

(larger α_p) also increase its magnitude. The price response is therefore a function of the joint effects of forward looking behavior and habits. We also note that as expected, Δg_{t+1} converges towards its myopic equivalent as $\alpha_f \to 0$.

The long-term response is proportional to its short-term counterpart. The multiplier is proportional to the habits strength. In the edge case, a hypothetical household that would not experience any rigidity on its gasoline demand ($\alpha_p = 0$) would adjust its consumption instantly to its long-term level, and would exhibit $\Delta g_{t+1} = \Delta g_\infty$.

I now turn to the datasets I will use to bring the rational habits model to the data.

3. Data

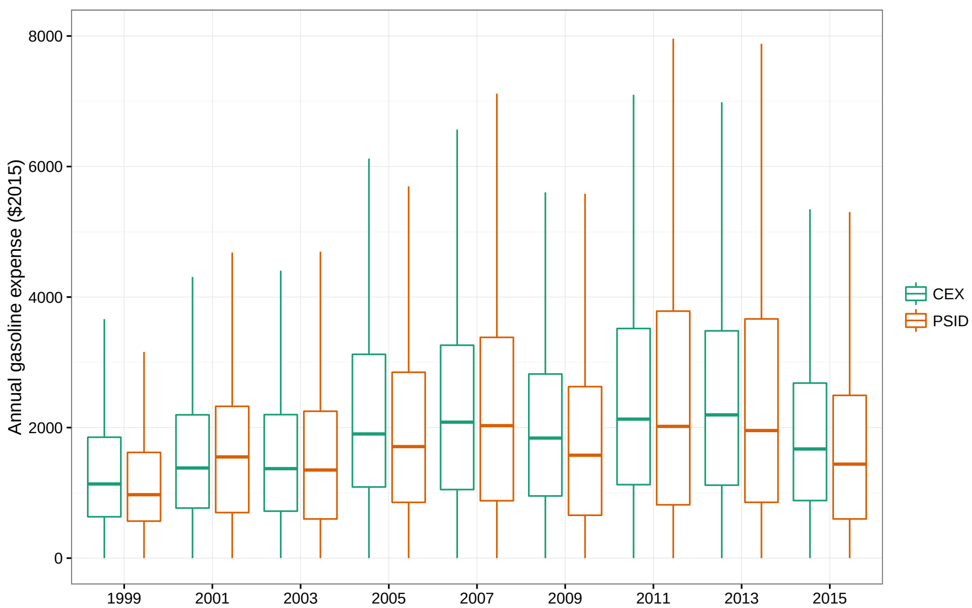
3.1. Gasoline consumption

To measure gasoline consumption data at the household level over a long period of time, I use data from the Panel Study of Income Dynamics. The PSID is the longest running panel study in the world. Started in 1968, it collects a large range of socio-economic data at both the household and individual levels. From a starting sample of 5,000 families comprising 18,000 individuals, the PSID has grown to survey close to 10,000 households and 24,000 individuals in their 2017 wave. Initially conceived as an annual survey, the PSID has been conducted biennially since 1997. In the following, I construct the panel by including every household surveyed in at least five consecutive waves of the PSID from 1999 until 2015. This design decision is discussed in detail in section 5.2.

While not originally part of the data surveyed, the PSID started to collect detailed annual data on household expenses in 1999, including gasoline consumption. In benchmarks conducted against the Bureau of Labor Statistics' Consumer Expenditure Survey (CEX), Li et al. (2010) and Andreski et al. (2014) find that the PSID's reported expenses track the CEX reference data very closely, both in the aggregate and across distinct expense categories. Besides, even though the CEX provides a useful reference point, it is not suited to the purposes of the present inquiry: each participating household is only surveyed at most four quarters, which precludes any analysis of long-term household-level responses[7].

I conduct a similar comparison to assess the reliability of the gasoline consumption data reported in the dataset. Specifically, PSID wave is compared with its corresponding CEX wave: for example, the 1999 PSID households are compared with CEX households surveyed throughout the four quarters of 1999[8]. I then confront the distribution of gasoline consumption reported by both surveys for each year in the sample, as illustrated in Figure 1. Table B.1 also provides a year by year comparison of summary statistics between gasoline expenditure reported in the PSID and the CEX.

Figure 1: Comparison of gasoline consumption between the PSID and the CEX (1999-2015)[9]



I also find that gasoline expenditure reported in the PSID is very close to its CEX coun-

[7]It should be noted that the CEX has been used repeatedly to build long-run pseudo-panels (Attanasio and Weber, 1995; Fernández-Villaverde and Krueger, 2007). Yet this entails abandoning households as the unit of analysis in favor of cohorts. The large variance of gasoline consumption – even across similar individuals and families – emphasizes the importance of household-level analysis, which makes the pseudo-panel approach ill-suited for my purposes.

[8]The CEX has a rolling survey design. Each participating households is surveyed at most four consecutive quarters. Therefore, to recover households that were surveyed over a period corresponding directly to a calendar year, we need to select those who began their participation to the CEX on the first quarter of that given year.

[9]These comparisons are based on the unweighted samples from the PSID and CEX for each of the years considered.

terpart: over the whole period, the average reported PSID gasoline consumption is 0.2% higher than in the CEX. However, the PSID estimates are slightly more dispersed: the PSID interquartile range is 12% larger than that of the CEX. In addition, the imputation procedure used by the PSID sometimes leads to aberrant gasoline expenditure that lie significantly outside the corresponding range reported by the CEX: to avoid any risk of biasing the estimations, these outliers are excluded in the final dataset[10].

Households' total annual income is also included in the dataset. In addition to being an important covariate of gasoline expenditure, knowledge of households' income will be necessary to analyze the distributional impacts of gasoline price increases in section 6. Regarding the inclusion of this variable, it should be noted that analyses of consumer demand sometimes use total expenditure in lieu of total income, either explicitly or implicitly through the use of budget shares (Banks et al., 1997). The explicit use of total expenditure instead of income is motivated by the permanent income hypothesis (Friedman, 1957), under which households are assumed to smooth their consumption over their life-cycle, thereby reducing the impact of transitory income shocks.

However, the PSID does not collect data on the whole range of consumer spending – Andreski et al. (2014) find that expenditure data surveyed in the PSID account for approximately 70% of total household expenditure as measured by the CEX. In contrast, total household income is fully accounted for in the dataset. To balance this limitation with the possible interest of considering total expenditure, I will focus in the remainder of this article on household income in the main estimations, but report results of alternate specifications including total expenditure in Appendix D.

In addition, I consider a number of time-varying household characteristics found to be relevant in the literature examining households' gasoline consumption patterns. I control in particular for the number of vehicles owned. This implies that long-term household response

[10] In effect, this entails excluding close to the top percentile of gasoline consumers as reported by the PSID in each wave.

estimates will only take into account changes in the composition of the household's vehicles bundle, but not its size. While this limitation must be acknowledged, it is unlikely to be a large source of bias, as less than 15% of households observed reduce the number of cars they own between PSID surveys.

Table 1: Summary statistics

	Mean	Std. dev.	25th perc.	Median	75th perc.
Gasoline expenditure	806.78	682.83	323.32	676.02	1129.90
Income	64,624	51,415	26,988	51,699	87,930
Total expenditure	38,031	31,836	20,263	32,845	49,404
Number of vehicles	1.70	1.12	1.00	2.00	2.00
Household size	2.69	1.48	2.00	2.00	4.00
Non-metropolitan county	0.21	0.41	0.00	0.00	0.00
Age of head	47.81	15.65	35.00	46.00	58.00
Education	15.78	15.11	12.00	13.00	16.00

Other controls include the size of the household, the age of the household's head as defined by the PSID[11], his or her educational attainment, and whether the household's home is located in a metropolitan or non-metropolitan area. The summary statistics for all PSID variables over the 6,074 households included in the dataset are provided in Table 1.

For this latter characteristic, the PSID reports the U.S. Department of Agriculture's Rural-Urban Continuum Code (RUCC, also called the Beale code), which is defined by the U.S. Department of Agriculture at the county level and identifies its level of urbanization among 9 possible values. While I make use of that level of detail to construct the gasoline price variable (see next section), in the subsequent regressions the RUCC indicator is aggregated to a single binary variable indicating whether the household considered resides in a metropolitan or non-metropolitan country as defined by the USDA. This is reported in Table 2.

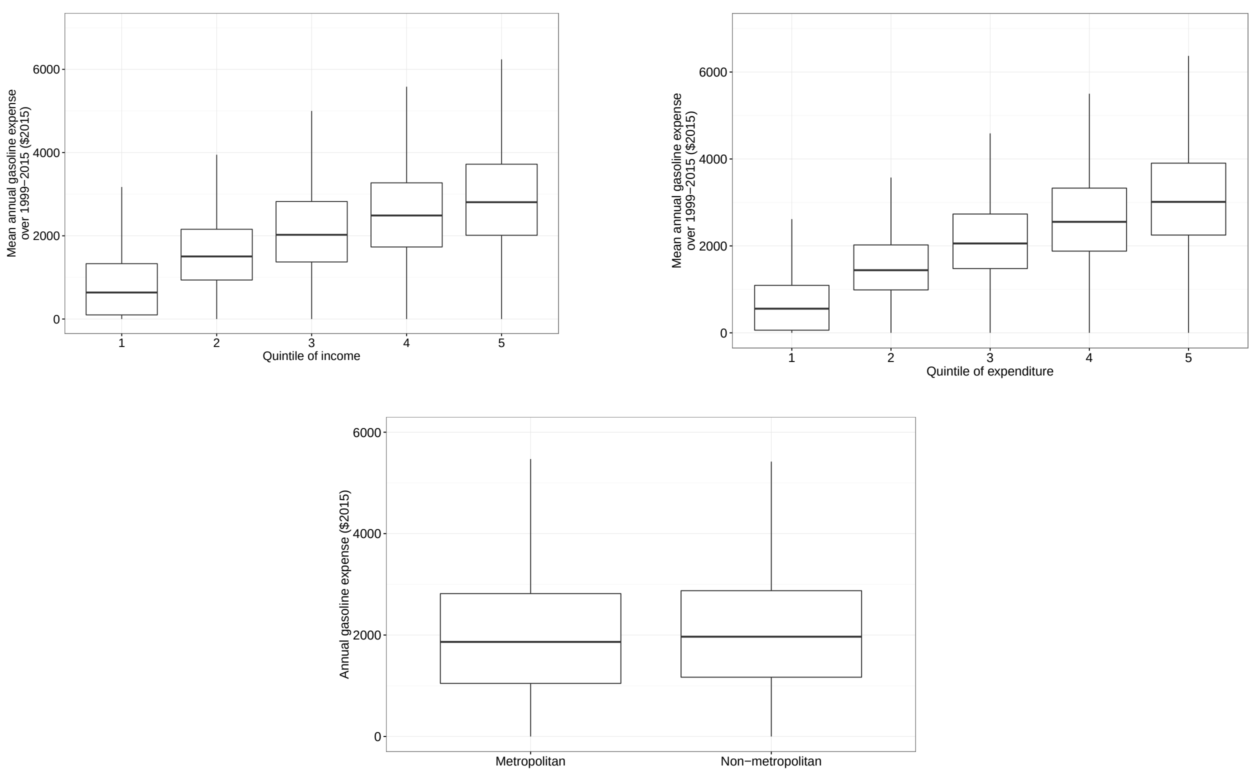
Further, gasoline consumption presents significant variance across each of these dimen-

[11]From the PSID survey design reference: "The Head of the FU must be at least 18 years old and the person with the most financial responsibility for the FU. If this person is female and she has a spouse or partner in the FU, then he or she is designated as Head. If she has a boyfriend with whom she has been living for at least one year, then he is Head. However, if the husband or boyfriend is incapacitated and unable to fulfill the functions of Head, then the FU will have a female Head."

Table 2: Levels of urbanization in the Rural-Urban Continuum Code

	RUCC	Definition
Metropolitan	1	Counties in metro areas of 1 million population or more
	2	Counties in metro areas of 250,000 to 1 million population
	3	Counties in metro areas of fewer than 250,000 population
	4	Urban population of 20,000 or more, adjacent to a metro area
	6	Urban population of 2,500 to 19,999, adjacent to a metro area
	8	Rural or $<$ 2,500 urban population, adjacent to a metro area
Non-metropolitan	5	Urban population of 20,000 or more, not adjacent to a metro area
	7	Urban population of 2,500 to 19,999, not adjacent to a metro area
	9	Rural or $<$ 2,500 urban population, not adjacent to a metro area

Figure 2: Gasoline consumption as a function of household characteristics

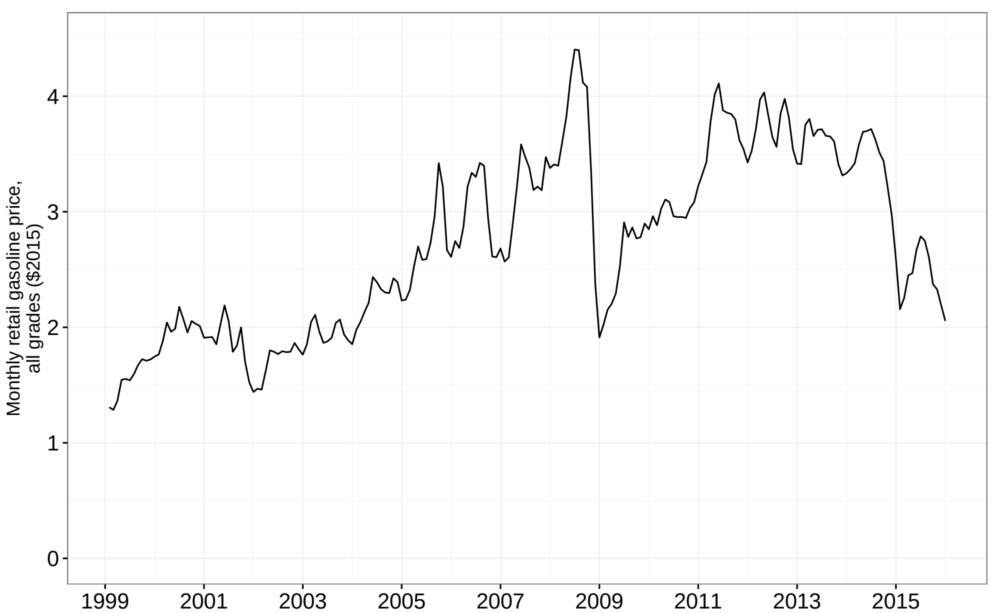


sions. Figure 2 illustrates how gasoline consumption varies with income and expenditure quintile, and location. These simple charts present the expected unconditional univariate correlations: gasoline consumption generally increases with income and total expenditure and is slightly higher for rural households. This latter finding is in contrast with results commonly found in Europe, which highlight a more sizeable difference in gasoline consumption between urban and rural settings – see Berry (2019) for a recent example in France. This may result from the higher availability of public transportation options in European cities, which reduces the need for private motorized vehicles and their associated gasoline consumption (Buehler et al., 2017).

3.2. Gasoline prices

Gasoline prices in the U.S. have been remarkably volatile over the period of interest covering 1999 to 2015. The households observed have been exposed to gasoline price per gallon ranging from \$1.28 to \$4.40 (monthly average retail gasoline price, all grades, constant \$2015). The evolution of monthly U.S. gasoline price is summarized in Figure 3.

Figure 3: Monthly average retail gasoline price in the U.S., all grades (1999-2015)[12]



However, the price of gasoline is not homogeneous across U.S. territories, and may vary substantially both across and within state boundaries. According to the U.S. Energy Information Administration, while gasoline prices have been dominated by the cost of crude

[12]Source: U.S. Energy Information Agency.

oil over most of the period of interest (see Table 3), federal and state taxes, refining and production costs, and distribution and marketing have comprised between 34% and 57% of the price of a gallon of gasoline in the U.S. from 1999 to 2015.

Table 3: Gasoline retail price components (1999-2015)

	1999-2003	2004-2007	2008-2011	2012-2015
Crude Oil	43%	54%	66%	62%
Federal and state taxes	29%	19%	14%	14%
Refining and production costs	15%	17%	9%	13%
Distribution and marketing	13%	10%	10%	11%

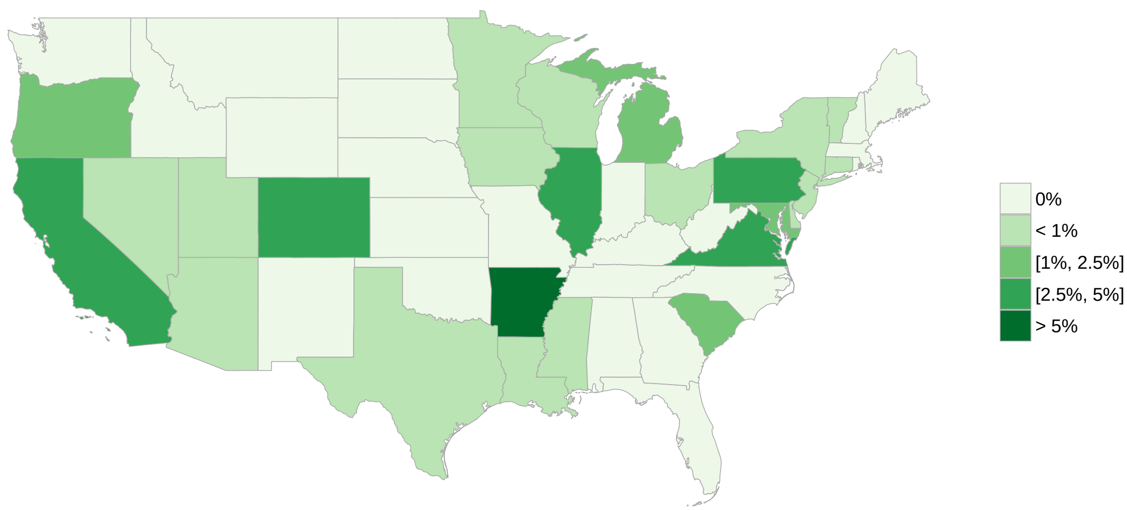
Source: U.S. Energy Information Agency

All three of these components are spatially heterogeneous across the U.S. territory. Sources of spatial variation in gasoline prices thus include differences in distribution costs – which relates in particular to the distance between point of sale and refinery; differences in environmental policies across states which impact refining costs; and differing levels of taxation across localities. These last two items should not be confused: environmental policies refer to requirements on the formulation of gasoline, notably to achieve cleaner combustion. For example, refineries targeting the Californian market are required to reduce the sulphur, benzene or aromatic hydrocarbon content of the gasoline, which entails further refining steps and increased costs (State of California, 2012). Conversely, gasoline taxation heterogeneity simply results from varying levels of specific and *ad valorem* excise tax enacted by local governments.

I now examine the exogeneity of this spatial heterogeneity. By regressing annual changes in nominal state gasoline taxes on first-differenced economic and political variables, Li et al. (2014) show conclusively that state gasoline taxes are not correlated with observable socioeconomic, political and industrial variables. However, while the topology of gasoline distribution networks remains mostly outside of households' control, the last factor, state-level environmental regulations, is still determined by state legislatures. The composition of these parliaments can be influenced by households' voting behavior on environmental policy issues, which in turn reflect their general beliefs about environmental preservation. This raises a concern for endogeneity, as environmentally conscious voters would be expected to both moderate their gasoline consumption (*e.g.* through the purchase of more efficient vehicles) and vote for more environmentally demanding regulations on gasoline refining – which would result into a spurious observation of higher prices seemingly leading to lower gasoline consumption.

To mitigate this potential source of endogeneity, I collect data on every U.S. congressional elections for the House of Representatives from 1994 until 2014[13] and identify the share of the vote apportioned to a candidate affiliated to the Green Party of the United States[14] in each congressional district. These shares are then aggregated at the state level.

Figure 4: Green share of the vote in congressional elections (mean of the 2006, 2008 and 2010 elections)



Further, in order to smooth election-to-election erratic variations, I compute a moving average over the three preceding elections, thereby capturing the magnitude of the green vote over a rolling four-year period. For example, the value used for the 2011 wave in the sample is an average of the green shares of the vote observed in the 2006, 2008 and 2010 elections. Figure 4 includes a map of this indicator for that wave and reveals large disparities in the

[13]All 435 seats of the U.S. House of Representatives are renewed every two years.

[14]The Green Party of the United States is a federation of state-level parties that gained national recognition from the Federal Election Commission in 2001 and ranked in 2014 as the 4$^{\text{th}}$ largest party in the US with 248,189 registered voters.

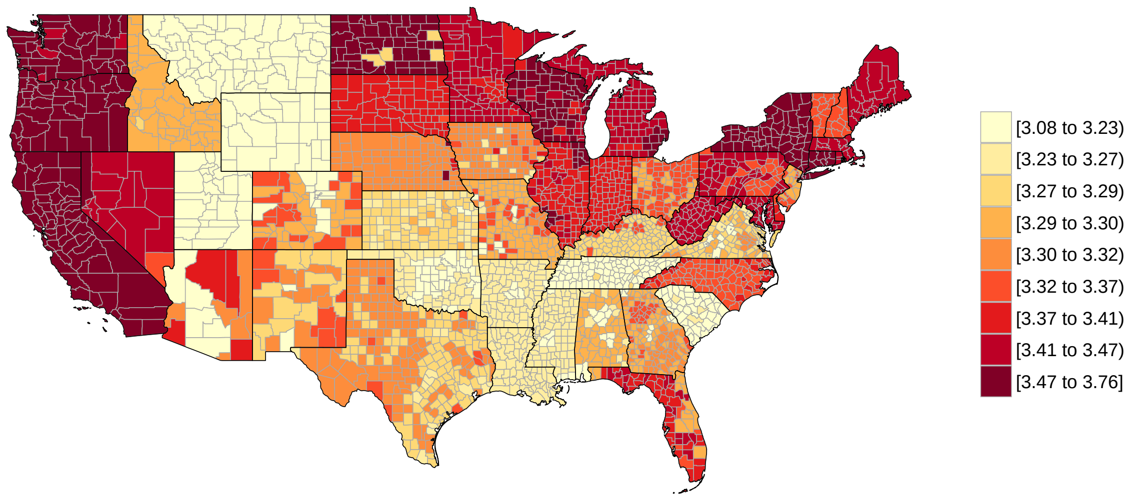
share of green votes across the U.S., with many states not featuring a single green candidate. In complement, maps for all years, which document a general increase in the electoral success of the Green Party over the period considered, are provided in Figures B.3 through B.11. I use this indicator as a proxy for households' idiosyncratic sensitivity to environmental topics and policies and integrate it as an additional control in all specifications.

To maximize identifying variance, I seek to use gasoline prices with the highest possible spatial resolution given the household dataset. I obtain gasoline prices at the city and county levels from the Cost of Living Index (COLI) collected by the Council for Community and Economic Research (C2ER). Conducted since 1968, the COLI surveys quarterly prices for over 60 goods and services, including gasoline, in more than 250 individual locations around the continental United States and Alaska.

To preserve their participants' anonymity, the PSID only releases the state and RUCC code (see previous section) of the county in which they reside. I therefore construct an annual gasoline price index by state and RUCC code from the COLI price data. Each COLI urban area is ascribed its corresponding RUCC code. For each state-RUCC code combination, a price index is then calculated as follows: for a given year all observations within a single state-RUCC combination are averaged; for state-RUCC combinations where no gasoline price was observed, the metropolitan or non-metropolitan state average is used. In some rare edge cases, only a single observation is reported for the state - in that case, this is the only value used. These occurrences are limited to sparsely populated states.

This procedure allows to significantly improve the spatial variance of the gasoline price variable, as shown for example for the year 2011 in Figure 5. It should be noted that in high-price states, the gasoline price scale used on the national map can obscure state-level heterogeneity – this is particularly the case for California and New York, for which close-up maps are provided in Figures B.1 and B.2 (in addition, localized gasoline price maps are provided for every PSID survey year included in the sample in Figures B.12 through B.20).

Figure 5: Localized gasoline prices in the continental U.S. (2011)



For example in 2011, annual average gasoline prices varied across locations in the U.S. from \$3.08 to \$3.76 per gallon – the most expensive retail gasoline was 22% more expensive than its cheapest counterpart.

4. Econometric approach

In an empirical setting, the rational habits model yields two testable hypotheses: first, households will exhibit habits in their gasoline consumption, and therefore rigidity and their adjustment to changing prices if $\alpha_p > 0$ and is statistically significant. Second, households will present evidence of forward-looking behavior if $\alpha_f > 0$ and is statistically significant. The resulting estimates for the model parameters, along with equations (9) and (10) then allow to derive estimates for the short and long term household responses to gasoline price variations.

To estimate the rational habits model of gasoline consumption, I specify the following dynamic panel model:

$$g_{it} = \alpha_p g_{it-1} + \alpha_f g_{it+1} + \beta p_{it} + \gamma x_{it} + \epsilon_{it} \tag{15}$$

where g_{it} is the gasoline consumption of household i (in gallons), p_{it} is the price paid by household i (in \$ per gallons) and x_{it} are the socio-economic covariates of household i. It should be noted that since the PSID is conducted every two years, each of the household-specific variables are measured every two years. Hence $t-1$ actually refers to a household state observed *two years* prior to year t.

Estimating model (15) raises a number of econometric challenges. First is the common issue of controlling for unobserved household heterogeneity. Second is the endogeneity concern of lag and lead gasoline consumption. Third is the inclusion of households who consume zero gasoline in the sample. I examine each of these difficulties and address them in turn.

4.1. Unobserved household heterogeneity

While the PSID provides us with several socio-economic household covariates that have been identified in the literature as having an impact on gasoline demand patterns, a number of unobserved household characteristics remain – including, but not limited to, household preferences for vehicle size and type and driving style. In a panel setting, repeated observations of the same household allows us to control for time-invariant unobserved heterogeneity using a normally distributed random variable – random effects, or a household-specific intercept – fixed effects (Wooldridge, 2010).

A number of previous analyses of household-level gasoline consumption in a panel setting implement a random effects model (Wadud et al., 2010a,b; Frondel et al., 2012). However, the random effects panel estimator is only consistent under the assumption that individual effects are uncorrelated with explanatory variables (that is $E(X_{it}\mu_i) = 0$ with $X_{it} \equiv [g_{it-1}, g_{it+1}, p_{it}, x_{it}]$ and μ_i the household-specific intercept). I opt to lift this restrictive assumption by resorting to fixed effects to model unobserved household heterogeneity.

To control for the aggregate macroeconomic cycle, which can contribute to an underlying trend in both gasoline prices and households' spending patterns, time fixed effects are also included in the econometric specification. The identification strategy thus relies on gasoline price variations across location. This reinforces the importance of the localized gasoline price instrument constructed in section 3.2. The idiosyncratic error ϵ_{it} of equation (15) is therefore disaggregated as follows:

$$\nu_{it} = \mu_i + \delta_t + u_{it} \tag{16}$$

4.2. Endogeneity of lag and lead consumption

The endogeneity of lag and lead consumption constitutes the main difficulty in estimating a rational habits model (Baltagi and Griffin, 2001). Indeed, per equation (15), we have the following relationships:

$$\begin{aligned} g_{it-1} &= \alpha_p g_{it-2} + \alpha_f g_{it} + \beta p_{it-1} + \gamma x_{it-1} + \mu_i + \delta_{t-1} + \nu_{it-1} \\ g_{it+1} &= \alpha_p g_{it} + \alpha_f g_{it+2} + \beta p_{it+1} + \gamma x_{it+1} + \mu_i + \delta_{t+1} + \nu_{it+1} \end{aligned} \tag{17}$$

Therefore, if $\alpha_p > 0$ and $\alpha_f > 0$, we have $E(g_{it-1}\nu_{it}) > 0$ and $E(g_{it+1}\nu_{it}) > 0$. Obviously, these correlations imply that a simple OLS estimation of equation (15) would be biased upwards.

However they also preclude the use of the usual within transformation to cancel out the households fixed effects, since it does not eliminate the dynamic panel bias (Wooldridge, 2010). If we call ν^*_{it} the within-transformed error term, we have:

$$\nu^*_{it} = \nu_{it} - \frac{1}{T-2}\sum_{k=2}^{T-1} \nu_{ik} \tag{18}$$

Equation (18) implies that $E(g_{it-1}\nu^*_{it}) < 0$ and $E(g_{it+1}\nu^*_{it}) < 0$, and that the within estimator would be biased downwards. Even then, the pooled and within estimators, while biased, can still provide lower and upper bounds on the true value of the dynamic parameters.

Thus the dynamic panel bias applies fully to this setting. It should be noted that estimators traditionally used to overcome this issue, Arellano and Bond's (1991) first-differenced GMM and Blundell and Bond's (1998) system GMM, are not applicable in this case. These estimators only use internal instruments for the lags of the dependent variable[15]. Given that gasoline consumption is linked by construction to its lags and leads in all periods in the rational habits models, its estimation will require external instruments.

I return to the theoretical model to drive the choice of instruments. Equation (7) implies that future consumption results from the sequence of future gasoline prices and household covariates. A symmetric factorization of equation (6) would yield the same implication for past consumption and lags of past prices and household characteristics. I therefore instrument the two-year lag (resp. lead) of gasoline consumption by the one and two-year lags (resp. leads) of gasoline price, and the two-year lag (resp. lead) of household income, size and metropolitan status. This then yields a specification amenable to estimation by fixed effects 2-stage least squares (FE2SLS). This instrumentation scheme is comparable to previous designs used in the literature to bring rational addiction or rational habits models to the data, notably by Baltagi and Griffin (2001), Laporte et al. (2010) or Filippini et al. (2018).

5. Results

5.1. Main results

Table 4 presents the estimation results of the rational habits model. These first results are estimated with FE2SLS, and include a full set of household and time fixed effects. Two-year lag and lead gasoline consumption, g_{t-1} and g_{t+1}, are instrumented by the one and two-year lags and leads of gasoline price, the two-year lag and lead of income, household size and metropolitan status respectively. All standard errors are clustered at the household level, to control for possible auto-correlation (Wooldridge, 2010).

I find that gasoline price is as expected negatively correlated with gasoline consumption in all specifications. Further, households' present gasoline consumption is dependent on their past consumption, and therefore exhibits habits formation. The value of α_p is fairly

[15]In first-differenced GMM, first-differences of the dependent variable are instrumented by further lags of the dependent variable. In system GMM, lags of the dependent variable are also instrumented by further lags of the first differences

Table 4: Rational habits model of gasoline consumption: estimation results

	(1)	(2)	(3)	(4)
g_{t-1}	0.178***	0.173***	0.171***	0.169***
	(0.0422)	(0.0419)	(0.0418)	(0.0418)
g_{t+1}	0.136***	0.128***	0.127***	0.128***
	(0.0485)	(0.0476)	(0.0477)	(0.0476)
p_t	-198.9***	-197.6***	-197.4***	-195.8***
	(39.64)	(39.72)	(39.73)	(42.61)
Income	0.000518***	0.000533***	0.000532***	0.000533***
	(0.000157)	(0.000155)	(0.000155)	(0.000155)
Number of vehicles	196.0***	196.3***	196.4***	196.4***
	(6.879)	(6.868)	(6.865)	(6.859)
Household size	41.87***	42.21***	42.21***	42.19***
	(5.153)	(5.130)	(5.130)	(5.124)
Non-metropolitan county		28.33	28.45	26.86
		(27.00)	(27.04)	(27.70)
Age of head			30.35***	30.31***
			(11.37)	(11.34)
Education			-0.0257	-0.0231
			(0.244)	(0.244)
Green vote	-138.1	-122.1	-119.0	-127.2
	(331.3)	(333.1)	(333.6)	(341.3)
Time FE	Yes	Yes	Yes	Yes
Census division FE	No	No	No	Yes
F-test	192.9	178.4	156.2	104.7
Hansen's J	4.78	7.09	7.03	7.34
Hansen's J *p-value*	0.57	0.53	0.53	0.50
Observations	31,308	31,308	31,308	31,308
2-year elasticity	-0.743***	-0.730***	-0.728***	-0.722***
	(0.157)	(0.155)	(0.155)	(0.166)
Long-term elasticity	-0.909***	-0.886***	-0.882***	-0.874***
	(0.197)	(0.193)	(0.192)	(0.204)

Standard errors in parentheses.

Standard errors for elasticities estimates computed using the delta method.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

consistently estimated between 0.166 and 0.178 across specifications.

Importantly, I also find that α_f is positive and statistically significant across all variants of the model, comprised between 0.127 and 0.136. This provides evidence that expectations about future gasoline demand have an impact on present consumption decisions, thereby implying forward-looking behavior in the household's gasoline demand. As an aside, although left out of the derivation, the original formulation of the rational habits model by Becker et al. (1994) predicts that the ratio of α_f to α_p provides an estimate of the household's rate of time of preference, and by extension their private discount rate. The estimates imply an annual discount rate of around 17%, which is remarkably consistent with empirical measurements of households' private discount rate.[16].

The controls enter with the expected signs. Gasoline consumption increases with income, the size of the household and the number of vehicles owned. It also increases with the age of the household head. As hypothesized, gasoline consumption is negatively correlated with the share of green vote in each state, even though the coefficient is not significant. Non-metropolitan households also tend to consume more gasoline than their metropolitan counterparts, although again the point estimate was not statistically significant in these specifications. However, this coefficient can only be identified on the 2.3% of households in the sample which on average move from metropolitan to non-metropolitan counties and *vice-versa* in any given wave. There is likely too little variance to reliably identify this parameter.

Given my interest in the forward-looking behavior of households, I also control for education as measured by the number of years of education completed by the household head in specification (3). The inclusion of this covariate does not change the estimate of α_f nor of α_p. In specification (4) we also include Census divisions fixed effects to control for regional specificities (geography, transport infrastructure, economic activities) which may impact gasoline

[16]Previous estimations of rational addiction models in the literature have often lead to implausibly large (above 100%) or even negative discount rates. See Baltagi and Griffin (2001) and Baltagi and Geishecker (2006) for further discussion. For a recent review of the empirical literature on the measurement of households' time preferences, see Cohen et al. (2016).

consumption[17], and find again that the results are robust to these additional controls.

The econometric validity of the FE2SLS estimates is assessed by a number of tests ensuring that assumptions on the instruments are verified. The complete first stage results for the main specification, column (3), are provided in Table C.1. I test against the possibility that the model is underidentified using a Kleibergen-Paap rk test (Kleibergen and Paap, 2006), which is rejected with a *p-value* $< 10^{-5}$. We further need to ensure that the instruments are not weak. To this end, I perform a Cragg-Donald test (Cragg and Donald, 1993) and determine that all specifications pass the Stock-Yogo 5% maximal instrument variable relative bias critical value (Stock and Yogo, 2005). Finally, I also report the value of the Hansen's J statistic (Hansen, 1982) and its associated *p-value* to test for the joint validity of the instruments used. A rejection of the Hansen's J overidentification test would question the validity of the instruments. With *p-values* comprised between 0.52 and 0.57 across all specifications, the null hypothesis clearly cannot be rejected, comforting the joint validity of the instrumentation scheme.

These results and equations (9) and (10) allow to estimate short-term and long-term gasoline price elasticities at the sample mean. The standard errors on these estimates are estimated using the delta method (Hoef, 2012). It should be noted that given the temporal resolution of the data, the "short-term" elasticity actually measures a 2-year response. The point estimates are consistent across models (1) through (4), with a 2-year elasticity comprised between -0.72 and -0.74, while the long term elasticity is found between -0.87 and -0.91. As expected the long-term response is larger than its medium-term counterpart, yet I observe that most of the household response to gasoline price variation is already completed after two years.

These values are within the upper end of the range reported in the literature. Most

[17] These regional fixed effects are not included in the households fixed effects, since PSID households are tracked when they move across state boundaries. Indeed, on average, around 15% of the sample moves to a new state between each wave of the PSID.

interestingly, they are comparable in magnitude to those estimated on discrete-continuous models, lending credence to the idea that the rational habits model captures the dynamics of household gasoline demand accurately while being simultaneously simpler and significantly more parsimonious in terms of data requirements. This latter advantage makes it amenable to estimation on long run household-level panels.

In complement, equations (9) and (10), which allowed us to compute 2-year and long-term gasoline price elasticities, can be easily adapted to calculate income elasticities. To this end, I simply substitute β by the point estimate of the income coefficient, and $\overline{p}$ by $\overline{y}$, the household income sample mean. Applying these adapted formulas to the main specification, column (3), yields an income elasticity of 0.05 and 0.06 at the two year and long-term horizons respectively. This is on the lower end of a recent meta-analysis conducted by Havranek and Kokes (2015), who report mean income elasticities of 0.1 in the short run and 0.23 in the long run.

Comparison with the myopic and static models

As was mentioned in section 2, the rational habits model encompasses both the myopic backward-looking (with $\alpha_f = 0$) and the static model (with $\alpha_p = \alpha_f = 0$). This feature allows to compare estimates of gasoline price elasticities across these three models. The myopic model is estimated through a FE2SLS setup similar to that of the main rational habits model, while the static model is estimated using a simple within estimator. Results are reported in Table 5.

While the ratio between medium and long term elasticities is preserved, the myopic model produces estimates close to three times smaller to the rational habits model. These estimates, -0.25 and -.30 respectively, are in effect very close to that of the static model (-0.27). The failure to allow for household intertemporal dependence in gasoline consumption – in effect neglecting the joint effects of habits formation and forward-looking behavior – thus leads to a significant under-estimation of the magnitude of households' response to gasoline price

Table 5: Comparison of the rational habits, myopic and static models

	(1) Rational habits	(2) Myopic	(3) Static
g_{t-1}	0.167***	0.164***	
	(0.0420)	(0.0387)	
g_{t+1}	0.127***		
	(0.0475)		
p_t	-197.6***	-79.88***	-90.87***
	(39.72)	(16.87)	(6.836)
Income	0.000532***	0.000740***	0.00106***
	(0.000155)	(0.000136)	(0.000116)
Number of vehicles	195.7***	197.9***	211.9***
	(6.863)	(6.155)	(5.637)
Household size	42.59***	47.22***	51.31***
	(5.148)	(4.375)	(3.930)
Non-metropolitan county	28.67	30.14	37.17**
	(27.05)	(19.53)	(18.51)
Age of head	30.22***	6.148	0.430
	(11.35)	(7.976)	(0.947)
Education	0.00139	0.000960	0.00139**
	(0.000888)	(0.000791)	(0.000707)
Green vote	-124.2	262.4	-58.80
	(334.0)	(300.0)	(304.4)
Time FE	Yes	Yes	Yes
F-test	156.3	192.3	226.8
Hansen's J	6.85	4.86	
Hansen's J *p-value*	0.55	0.30	
Observations	31,308	37,570	43,900
2-year elasticity	-0.728***	-0.250***	-0.271***
	(0.155)	(0.053)	(0.02)
Long-term elasticity	-0.878***	-0.299***	
	(0.191)	(0.064)	

Standard errors in parentheses.

Standard errors for elasticities estimates computed using the delta method.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

variations.

5.2. *Robustness*

Comparison of estimators

To assess the robustness of the results, I submit them to a number of checks. I first test a number of alternative estimators. As discussed in section 4, despite their bias a simple pooled and within estimators can provide a helpful interval within which the true lead and lag coefficients should be found. Table 6 reports the comparison between these estimators and the FE2SLS used in the previous subsection.

I also implement the feasible efficient GMM (FEGMM) estimator. According to theory, the FEGMM should bring efficiency improvements against the FE2SLS, since it exploits more moment conditions in the data. However, estimating the optimal weighting matrix necessary to implement FEGMM requires recovering 4^{th} order moments from the data. While this is impractical in datasets with few panel units – notably when observing states over time, as in Baltagi and Geishecker (2006), Scott (2012) or Filippini et al. (2018) –, a relatively large sample of more than 6,000 households makes FEGMM estimation applicable.

Table 6 reports that as expected, the FE2SLS estimates (1) for the dynamic coefficients α_f and α_p are found between the estimates resulting from the within (3) and pooled (2) estimators respectively. The FEGMM estimates (4) for these dynamic parameters are very close to those obtained from FE2SLS. Further, the magnitude of the gasoline price coefficient is also remarkably consistent across estimators. These findings reinforce the robustness of the main results, particularly for the estimates of the 2-year and long-term household elasticities.

Households that do not consume gasoline

One of the issues involved in the analysis of household-level gasoline consumption patterns is that a small yet significant portion of households (around 10% of the sample) does not consume gasoline, generally because they do not own a car and haven't rented any personal vehicles during the year surveyed.

Table 6: Comparison of estimators

	(1) FE2SLS	(2) Pooled	(3) Within	(4) FEGMM
g_{t-1}	0.167***	0.232***	-0.0728***	0.164***
	(0.0420)	(0.00575)	(0.00792)	(0.0419)
g_{t+1}	0.127***	0.266***	-0.0669***	0.132***
	(0.0475)	(0.00597)	(0.00875)	(0.0468)
p_t	-197.6***	-203.0***	-217.8***	-198.2***
	(39.72)	(17.85)	(41.62)	(39.40)
Income	0.000532***	0.000759***	0.000998***	0.000571***
	(0.000155)	(0.0000716)	(0.000142)	(0.000154)
Number of vehicles	195.7***	162.2***	207.7***	196.0***
	(6.863)	(4.830)	(6.764)	(6.850)
Household size	42.59***	25.26***	53.94***	41.76***
	(5.148)	(2.008)	(5.171)	(5.109)
Non-metropolitan county	28.67	-5.879	48.64	21.64
	(27.05)	(6.909)	(29.80)	(26.46)
Age of head	30.22***	-2.337***	32.71***	30.09***
	(11.35)	(0.156)	(10.80)	(11.35)
Education	0.00139	0.00177***	0.00222**	0.00136
	(0.000888)	(0.000580)	(0.000974)	(0.000887)
Green vote	-124.2	-22.19	-118.8	-101.9
	(334.0)	(182.3)	(402.4)	(332.5)
Time FE	Yes	Yes	Yes	Yes
F-test	156.3	1484.8	138.6	157.2
Hansen's J	6.85			6.85
Hansen's J *p-value*	0.55			0.55
Observations	31,308	31,781	31,781	31,308
2-year elasticity	-0.728***	-0.952***	-0.642***	-0.735***
	(0.155)	(0.083)	(0.123)	(0.154)
Long-term elasticity	-0.878***	-1.267***	-0.599***	-0.883***
	(0.191)	(0.11)	(0.115)	(0.19)

Standard errors in parentheses.

Standard errors for elasticities estimates computed using the delta method.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Gasoline consumption enters linearly in the rational habits model, which implies that these zero observations can be included within the framework of the model. Still, their inclusion could be a source of bias. To test against this possibility, I perform two tests. First, the main specification is estimated on a subset of the sample which excludes observations of households that did not consume gasoline in a given year[18] – thereby voluntarily creating a selection bias. Second, considering gasoline consumption as a left-truncated dependent variable, an instrumental variable Tobit estimator is applied to the main specification.

This second strategy brings another layer of complexity, as the non-linearity of the Tobit estimator precludes the use of fixed effects (Greene, 2004). Instead, I resort to the Chamberlain device to model household unobserved heterogeneity (Chamberlain, 1982): household fixed effect are projected onto all the realizations of the k household covariates over the observation sample.

$$\mu_i = \sum_{k=0}^{K} \sum_{t=0}^{T} x_{ikt} \tag{19}$$

Beyond this modification in the modelling of household unobserved heterogeneity, I maintain the same instrumentation strategy. Results are presented in Table 7. Excluding households that do not consume gasoline from the sample produces a downward bias of around 19% on elasticities estimates, as expected. However, the Tobit estimates are remarkably consistent with the main FE2SLS results. These results confirm that the inclusion of zeros among the gasoline consumption observations is not a source of bias in the main results.

Panel design

As mentioned in section 3, the panel dataset is constructed through the inclusion of households that have been observed over at least five consecutive waves of the PSID. This design decision has to balance countervailing objectives. To capture a more faithful measurement of households' dynamic response over a long period of time, it is useful to maximize the

[18]The years in which the household did consume gasoline are maintained in the sample.

Table 7: Inclusion of households that do not consume gasoline

	(1) FE2SLS $g_t \geq 0$	(2) FE2SLS $g_t > 0$	(3) Tobit $g_t \geq 0$
g_{t-1}	0.167***	0.205***	0.207***
	(0.0420)	(0.0459)	(0.0348)
g_{t+1}	0.127***	0.147***	0.241***
	(0.0475)	(0.0530)	(0.0353)
p_t	-197.6***	-250.5***	-185.5***
	(39.72)	(46.45)	(44.97)
Income	0.000532***	0.000542***	0.00110***
	(0.000155)	(0.000165)	(0.000112)
Number of vehicles	195.7***	119.6***	224.1***
	(6.863)	(7.527)	(10.65)
Household size	42.59***	48.35***	19.12***
	(5.148)	(5.963)	(6.755)
Non-metropolitan county	28.67	26.39	4.057
	(27.05)	(31.65)	(21.68)
Age of head	30.22***	37.37***	-2.928***
	(11.35)	(13.27)	(0.653)
Education	0.00139	0.00252***	0.000821
	(0.000888)	(0.000909)	(0.000639)
Green vote	-124.2	-187.1	55.77
	(334.0)	(347.1)	(241.7)
Time FE	Yes	Yes	Yes
Chamberlain's device	No	No	Yes
F-test	156.3	106.0	
Hansen's J	6.85	10.5	
Hansen's J *p-value*	0.55	0.23	
Observations	31,308	26,854	31,313
2-year elasticity	-0.728***	-0.824***	-0.710***
	(0.155)	(0.165)	(0.169)
Long-term elasticity	-0.878***	-1.046***	-0.908***
	(0.191)	(0.217)	(0.215)

Standard errors in parentheses.

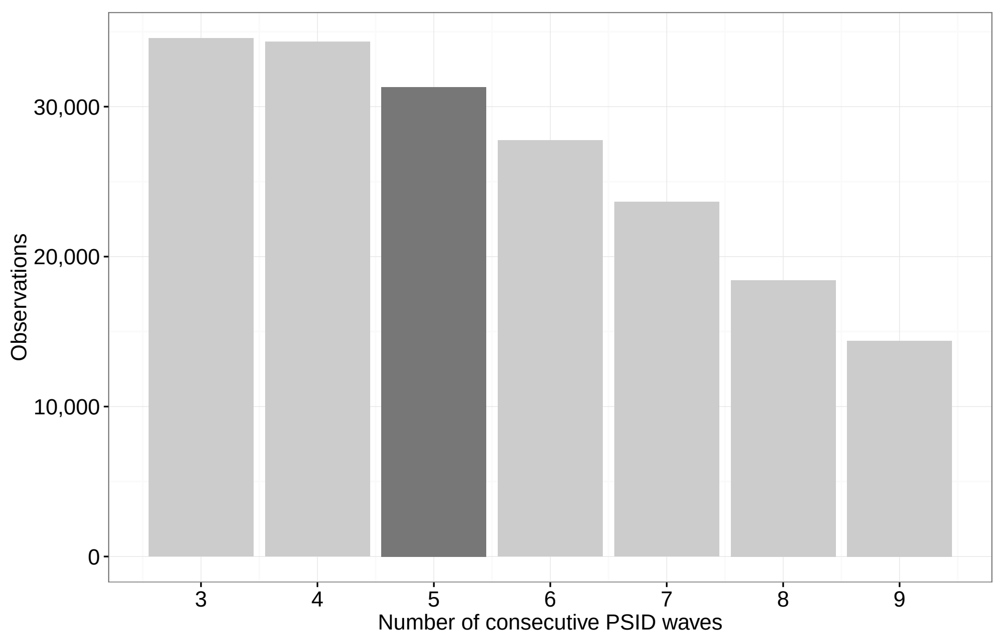
Standard errors for elasticities estimates computed using the delta method.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

minimum number of *consecutive* waves over which they are observed. Conversely, despite its enduring quality, the PSID is still subject to attrition. This rapidly reduces the number of households observed as one increases the constraint on the minimum number of consecutive observations. Finally, given the dynamic nature of the model, it is preferable to avoid unbalancing the panel excessively.

A household must be observed in at least three consecutive periods to be included in an estimation of the rational habits model, which requires an observation of one lead and one lag of gasoline consumption. Besides, the full period of observation from 1999 to 2015 covers nine waves of the PSID. Figure 6 presents the number of observations available in the panel as a function of the constraint on the minimum number of consecutive observations required to include a household, with the chosen 5-wave design highlighted in dark grey.

Figure 6: Number of observations as a function of panel design



It is apparent that the total number of observations decreases very quickly as the panel design progresses towards a fully balanced panel. A panel built only from households observed in all nine PSID waves would contain less than half as many observations as the chosen design. Conversely, relaxing the constraint to four or even the minimum three consecutive waves only increases the number of observations from the design by less than 10%, at the cost of a significantly unbalanced panel.

These considerations have driven the choice of panel design. Yet, I still test the robustness of the results to this design decision by estimating the main specification on all possible panel designs described in Figure 6. Results are reported in Table C.2. The elasticities point estimates remain remarkably stable across all designs, with 2-year elasticity varying between -0.67 and -0.79, and long-term elasticity being comprised between -0.80 and -0.97. Considering only the three designs with more than 30,000 observations, these ranges are much narrower, at [-0.73, -0.70] and [-0.88, -0.86] respectively. This confirms that previous results are not driven by the choice of panel design.

5.3. Estimation by quintile of income

This section examines whether households' dynamic response varies across the income distribution. To this end, the main specification, equation (15), is modified to interact lag and lead gasoline, gasoline price and income (y_{it}) with a time-invariant household income quintile indicator variable.

However, since the sample covers more than a decade and a half, households are likely to move across the income distribution between waves. To overcome this issue, I compute for each household the mean income over all waves in which it is observed. Each household is then ascribed to a time-invariant income quintile on the basis of that mean income, which is fixed over time.

If we write $\mathbb{1}_q$ the quintile indicator variable, y_{it} the annual income of household i in year t and $x_{-y,it}$ all other household covariates except income, the estimating equation is modified as follows:

$$g_t = \sum_{q=1}^{5} \mathbb{1}_q \left[\alpha_{p,q} g_{it-1} + \alpha_{f,q} g_{it+1} + \beta_q p_{it} + \gamma_{y,q} y_{it}\right] + \gamma_{-y} x_{-y,it} + \epsilon_{it} \tag{20}$$

The full estimation results are provided in Table C.3, including an estimate of 2-year and long-run elasticity by quintile of income. The results for α_p and α_f are provided in Figures 7 and 8 respectively.

Figure 7: α_p as a function of income quintile

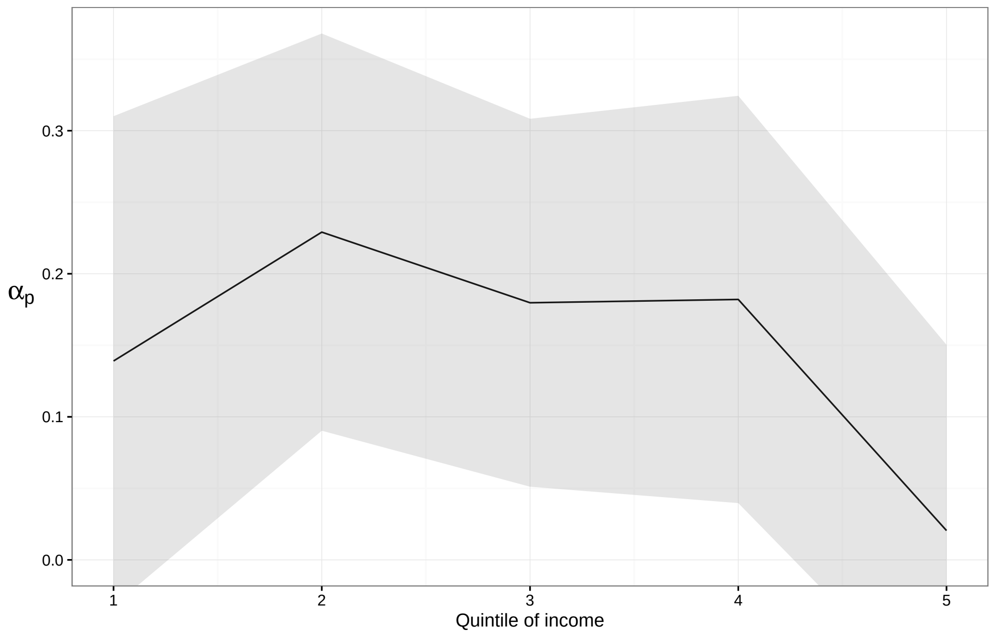
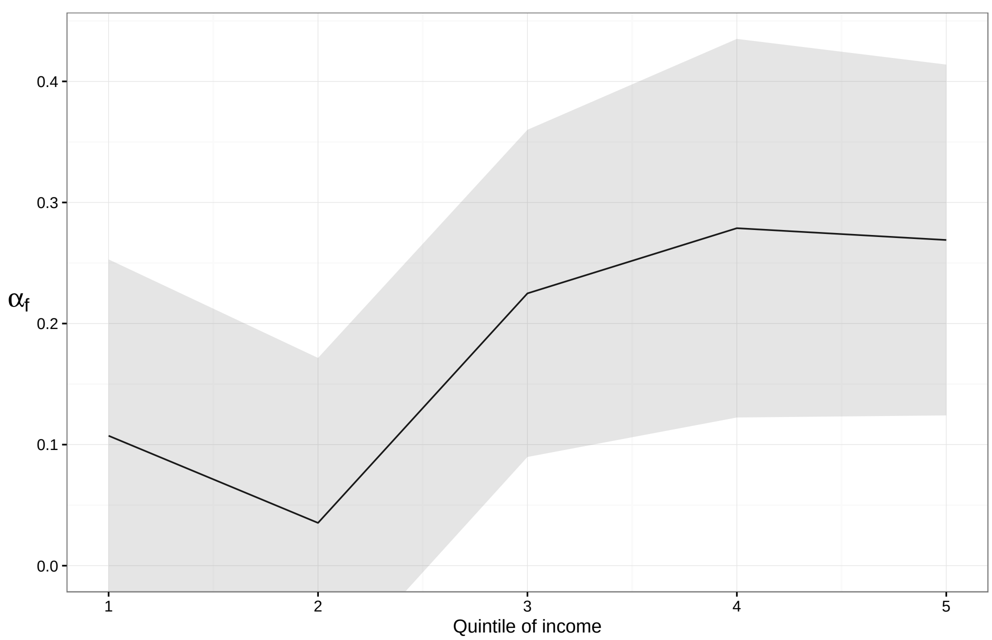


Figure 8: α_f as a function of income quintile



I find that the bottom 80% of the income distribution has stronger habits than the top income quintile, which does not exhibits statistically significant habits at all. These habits are particularly strong in the bottom half of the distribution. Conversely, forward-looking behavior grows stronger with the level of income, while it is not statistically significant for the lowest two income quintiles. This latter effect deserves more scrutiny, as a growing body of evidence (Hilgert et al., 2003; Lusardi and Mitchell, 2011; Lusardi and Tufano, 2015) suggests that financial literacy is poor among less affluent segments of the population. Thus the increased level of forward-looking behavior among the top two quintiles may simply stem from improved financial education among richer households. To test this hypothesis, the same specification is estimated without including educational attainment of the household head among the covariates (see Table C.4). The findings are robust to this removal, indicating that the greater propensity of richer households to be forward-looking results from a different cause.

Table 8: Gasoline price elasticities by quintile of income

	Quintile of income				
	1	2	3	4	5
2-year elasticity	-1.241***	-0.673***	-0.751***	-0.765***	-0.700***
	(0.123)	(0.0275)	(0.0335)	(0.0345)	(0.0204)
Long-term elasticity	-1.443***	-0.872***	-0.921***	-0.943***	-0.709***
	(0.198)	(0.0517)	(0.0646)	(0.0709)	(0.0251)

Standard errors in parentheses estimated using the delta method.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 8 reports gasoline price elasticities by quintile of income. Estimates for medium-term gasoline price elasticity range from -1.24 for the lowest income quintile to -0.70 for the highest quintile. The long-term response ranges from -1.44 to -0.71 respectively. These results confirm two common findings in the literature: household response to gasoline price is heterogeneous across households, and decreases with income.

However, applying a dynamic model also reveals that the spread in gas price sensitivity between the bottom and the top of the income distribution is large, with poorer households twice as responsive as richer households. These results also confirm the fact that households on the lower end of the income distribution take longer to adjust their response fully, while households from the top income quintile have virtually completely adjusted their response after two years.

6. Distributional impacts of a gasoline price increase

In this section, I combine the estimates of households responses by income quintiles obtained in the previous section with the demand equations derived in section 2.3 to perform a micro-simulation of the distributional impacts of a gasoline price increase.

6.1. Consumer surplus variation

First, I estimate the change in consumer surplus resulting from an increase in gasoline price corresponding to the enactment of a \$50/tCO_2 carbon tax. Using the average carbon content of a gallon of gasoline as reported by the U.S. Energy Information Agency, this corresponds to an increase of 44 cents per gallon. This can be compared to the current federal excise tax on gasoline of 18.4 cents per gallon, and the average state tax of 34.1 cents per gallon as of 2018 (American Petroleum Institute, 2018).

The model does not allow to distinguish between household responses to changes in taxation and their responses to price variations resulting from supply or demand shocks. Recognizing this limitation, in the following, the carbon tax will be modelled as a price increase. Using equations (11), the short-term and long-term change in consumer surplus resulting from a price change Δp can be derived:

$$\begin{aligned} \Delta CS_{t+1} &= g_t \Delta p - \frac{|\Delta g_{t+1}| \Delta p}{2} \\ \Delta CS_{\infty} &= g_t \Delta p - \frac{|\Delta g_{\infty}| \Delta p}{2} \end{aligned} \tag{21}$$

Table 9 presents the results of that micro-simulation conducted over the entire sample[19]. I find that a price increase corresponding to a \$50/tCO_2 carbon tax would reduce gasoline consumption by 10% after 2 years, and 12% in the long term. Considering 2015 data on CO_2

[19]Excluding the 1999 wave, for which g_{t-1} is not observed. Lagged gasoline consumption is necessary to the computation of both Δg_{t+1} and Δg_{∞} per equation (12).

emissions attributable to gasoline consumption in the U.S., this would imply reductions of 110 $MtCO_2$ and 125 $MtCO_2$ respectively. To take into account differences in standards of living, I report the ratio of consumer surplus variation to household income: poorer households experience a larger loss of consumer surplus relative to their income. This loss is also slightly reduced in the long term, as households complete their adjustment to the increase in gasoline price. Note that as expected from the estimates of medium and long-term elasticities by quintile in the previous section, richer households in the top income quintile complete their demand adjustment in less than two years, leading to an equal loss of consumer surplus over the 2-year and long-term horizons.

Table 9: Change in consumer surplus for a 44 cts increase in gasoline price

		Quintile of income				
		1	2	3	4	5
Ratio of consumer surplus	2-year	0.72%	0.67%	0.56%	0.46%	0.31%
variation to income	Long-term	0.71%	0.65%	0.55%	0.45%	0.31%

In addition to the variation in consumer surplus, I also calculate how the ratio between a household's tax burden and its income varies across the income distribution. This allows us to assess whether a \$50/tCO_2 carbon tax applied to gasoline is regressive (exhibiting a decreasing ratio of tax to income with increasing income) or progressive (increasing ratio with increasing income).

Suits (1977) proposed an index to quantify the degree of progressivity or regressivity of a given tax, constructed on the same principle as the Gini index. The Suits index is bounded by -1 and 1 and measures the deviation from proportionality of a tax's incidence, with 0 being perfectly proportional, +1 absolutely progressive – the whole tax is paid by the single richest household – and -1 absolutely regressive – the whole tax is paid by the single poorest household. Defining s as the share of cumulative household income and τ as the share of cumulative tax receipt, the Suits index can be expressed as follows:

$$S = \frac{1}{2} \int_{s=0}^{1} \left(\tau(s + ds) - \tau(s) \right) \, ds \tag{22}$$

Table 10 presents the ratio of tax burden to household income by quintile of income at the 2-year mark and in the long run, along with estimates of the Suits index for both time horizons[20]. These findings confirm the regressivity of gasoline taxation in the US, with a tax burden ratio to household income more than twice as high for the lowest quintile than for the top one, leading to a Suits index of -0.240 after two years and -0.236 in the long run. A gasoline price increase is slightly less regressive in the long-term, as households have enough time to adjust their response completely.

Table 10: Regressivity of a 44 cts increase in gasoline price (along the income distribution)

		Quintile of income					Suits index
		1	2	3	4	5	
Ratio of tax burden	2-year	0.70%	0.66%	0.53%	0.43%	0.29%	-0.240
to income	Long-term	0.68%	0.64%	0.51%	0.41%	0.29%	-0.236

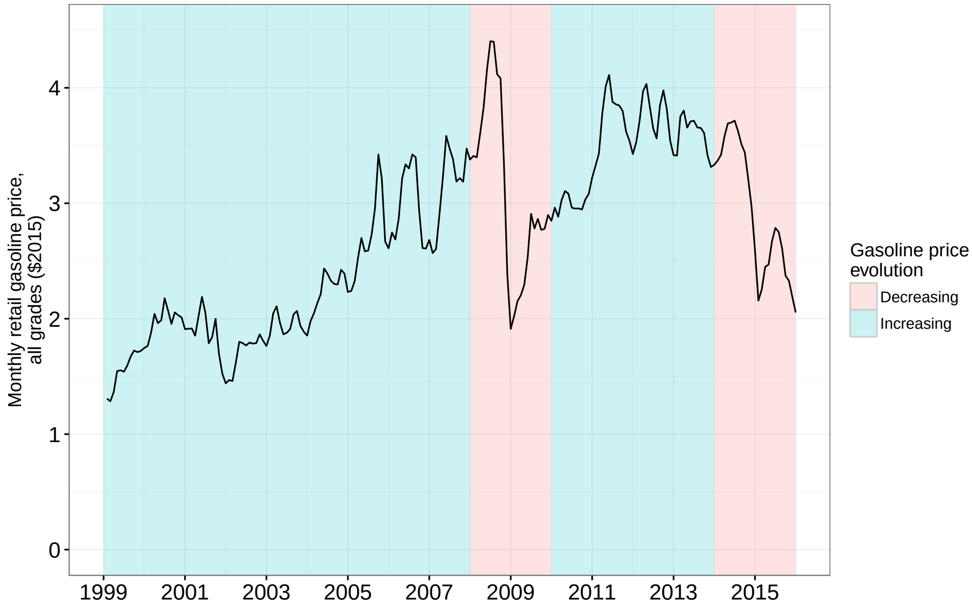
6.2. Dynamic heterogeneity and regressivity

The results of the previous section were estimated over the entire sample, comprising households surveyed in all waves from 1999 until 2015. This approach requires the implicit hypothesis that households will present the same response to a given increase in gasoline price throughout the period. While this assumption is valid in a static model, it is not necessarily verified in a dynamic framework.

More specifically, equations (12) governing the short and long-term changes in gasoline consumption (see section 2.3) showed that households response can be decomposed into an inertia component and a price response component. In the model, the price response term is solely a function of the magnitude of the price variation and the model's parameters: it will therefore be identical across the entire period of simulation.

[20]Given that the sample covers multiple PSID waves, I report the mean Suits index across all waves.

Figure 9: Phases of gasoline price movements in the U.S. (1999-2015)



However, the inertia term depends on the past dynamic of gasoline consumption, and will thus vary over time. In particular, as noted above, households with strong habits will carry some of the momentum of their gasoline demand, even if it is in a direction opposite to that of the price shock. In this latter case, this will reduce the magnitude of their response, which would lead to a larger gasoline consumption after the price increase – and in the case of a tax, a larger tax burden. Provided this phenomenon is larger on the lower end of the income distribution – which would be expected, given the stronger habits measured for the first three quintiles –, this could have an impact on the regressivity of the tax.

To explore this hypothesis, each two-year period covered by a PSID wave is categorized as an *increasing* or *decreasing* gasoline price period[21]. This characterizes the following four phases, as depicted in Figure 9: increasing over 1999-2007, decreasing over 2008-2009, increasing again over 2010-2013 and finally decreasing over 2014-2015.

I then estimate the 2-year Suits index resulting from the gasoline price increase corre-

[21]These two directions are simply defined using the difference between the gasoline price at the beginning and end of each two-year period.

Table 11: Evolution of regressivity over time

	1999 - 2007	2008-2009	2010-2013	2014-2015
Suits index	-0.238	-0.248	-0.239	-0.246
Tax burden ratio by quintile of income				
(1)	0.71	0.88	0.71	0.88
	[0.68, 0.75]	[0.79, 0.96]	[0.67, 0.76]	[0.79, 0.98]
(2)	0.69	0.82	0.66	0.74
	[0.67, 0.72]	[0.77, 0.87]	[0.63, 0.70]	[0.69, 0.80]
(3)	0.59	0.62	0.51	0.59
	[0.57, 0.61]	[0.59, 0.66]	[0.49, 0.53]	[0.55, 0.63]
(4)	0.48	0.49	0.41	0.46
	[0.47, 0.49]	[0.46, 0.51]	[0.40, 0.43]	[0.43, 0.49]
(5)	0.33	0.32	0.27	0.30
	[0.32, 0.34]	[0.30, 0.33]	[0.26, 0.28]	[0.28, 0.31]

Tax burden presented as a percentage of total household income.
Confidence intervals estimated using 500 bootstrap iterations.

sponding to a \$50/tCO_2 carbon tax in each phase separately. As anticipated, the regressivity of the price increase appears higher when it occurs after a period of decreasing gasoline price. Examining the results by quintile reveals that this finding is driven by the response of the first two quintiles, which experience a higher tax burden ratio when the tax is enacted after a phase of decreasing gasoline price.

To evaluate the significance of this result, I estimate a confidence interval around each of the tax burden ratio estimates by performing a bootstrap. The 95% confidence intervals around the tax burden ratios of the first two quintiles are disjointed across periods, implying that their difference is statistically significant.

This cursory evidence suggests that gasoline consumption dynamics do play a part in the regressivity of gasoline taxation, particularly on the lower end of the income distribution. This finding warrants further investigation of the interplay between dynamic heterogeneity in household responses and regressivity. It implies that compensatory policies should take into account the recent history of gasoline prices and trends in gasoline consumption, and should be reinforced when gasoline price increases are enacted after phases of growing gasoline consumption.

7. Conclusion

This article seeks to investigate the existence and magnitude of rigidities and forward-looking behavior in the response of households gasoline consumption to changes in gasoline price. Ultimately, my purpose is to identify the possible heterogeneity of these dynamic features among households, and to analyze how this dynamic heterogeneity can affect the distributional impacts of gasoline price increases.

To this end, I develop a simple dynamic model of household gasoline consumption using the rational habits framework of Becker et al. (1994). This model allows to capture the intertemporal dimension of gasoline demand through a parsimonious functional form linking present consumption to its past and future levels. Importantly, this paucity of data requirements makes it amenable to estimation on long-run household-level panel datasets.

This model is estimated on a large panel of 6,074 U.S. households obtained from the PSID and covering the years 1999 to 2015, a period marked by a high level of gasoline price variance. To complement this household-level data, I construct a localized gasoline price index at the state-RUCC level using city and county-level data gathered by the C2ER Cost of Living Index.

I then estimate a rational habits model of gasoline consumption using FE2SLS. Households exhibit habits formation and forward-looking behavior in their gasoline consumption, with relatively strong price elasticities of -0.73 after two years and -0.88 in the long-term in my preferred specification. These findings are robust to a number of robustness checks, in particular to different choices of estimator and to the inclusion of households that do not consume gasoline.

I further find evidence of dynamic heterogeneity among households. In particular, habit formation in gasoline consumption is stronger on the bottom half of the income distribution, while conversely the top two income quintiles exhibit stronger forward-looking behavior. This contributes to a large heterogeneity in household responses: households in the lowest income quintile are twice as sensitive to gasoline prices as their counterparts in the top one.

After conducting a micro-simulation of a gasoline price increase commensurate with a \$50/tCO_2 carbon tax, I estimate the change in consumer surplus and tax burden experienced along the income distribution and find that gasoline price increases are regressive, with a Suits index of -0.236.

Interestingly, I also find suggestive evidence of interactions between dynamic heterogeneity and the regressivity of gasoline price increases. Due to the greater inertia of their gasoline consumption, households in the bottom two quintiles of the income distribution experience a larger tax burden ratio after periods of falling gasoline prices. This implies that a gasoline tax implemented after a period of lenient prices would be more regressive.

These findings have important policy implications. Beyond a confirmation of the heterogeneity of households' responses to gasoline prices, the results demonstrate the existence and importance of dynamic heterogeneity. Households do not adjust their response at the same rate along the income distribution, which has an effect on the distributional impacts of increases in gasoline price. This makes the case for a reinforcement of compensatory policies targeting households on the lower end of the income distribution. In addition, these policies could be further strengthened in the short run, to help these households adjust their gasoline consumption faster.

This article also opens a number of venues for future research. The rational habits model could in particular provide insights on the puzzle of higher consumer sensitivity to gasoline tax increases when compared with price movements of similar magnitude (Davis and Kilian, 2011; Coglianese et al., 2017). In this framework, an announced gasoline tax can be perceived by the household as being a fairly certain future price component – for which adopting a stronger forward-looking behavior is more justified, thereby increasing its price response.

In addition, the dynamic heterogeneity finding could be explored further along other dimensions of gasoline consumption variability, notably across locational characteristics of households. A quantile regression approach could also shed new insights on the variability of households' response along the gasoline consumption distribution. More generally, the parsimonious nature of this model and its associated data requirements makes it applicable to a wide range of countries, both developed and emerging.

References

Alberini, A. and Filippini, M. (2011). Response of residential electricity demand to price: The effect of measurement error. *Energy Economics*, 33(5):889–895.

American Petroleum Institute (2018). State motor fuel taxes: Notes summary.

Andreski, P., Li, G., Samancioglu, M. Z., and Schoeni, R. (2014). Estimates of annual consumption expenditures and its major components in the psid in comparison to the ce. *American Economic Review*, 104(5):132–35.

Arellano, M. and Bond, S. (1991). Some tests of specification for panel data: Monte carlo evidence and an application to employment equations. *The Review of Economic Studies*, 58(2):277–297.

Attanasio, O. P., Koujianou Goldberg, P., and Kyriazidou, E. (2008). Credit constraints in the market for consumer durables: Evidence from micro data on car loans. *International Economic Review*, 49(2):401–436.

Attanasio, O. P. and Weber, G. (1995). Is consumption growth consistent with intertemporal optimization? evidence from the consumer expenditure survey. *Journal of Political Economy*, 103(6):1121–1157.

Baltagi, B. H. and Geishecker, I. (2006). Rational alcohol addiction: evidence from the russian longitudinal monitoring survey. *Health Economics*, 15(9):893–914.

Baltagi, B. H. and Griffin, J. M. (2001). The econometrics of rational addiction: the case of cigarettes. *Journal of Business & Economic Statistics*, 19(4):449–454.

Baltagi, B. H. and Griffin, J. M. (2002). Rational addiction to alcohol: panel data analysis of liquor consumption. *Health Economics*, 11(6):485–491.

Banks, J., Blundell, R., and Lewbel, A. (1997). Quadratic engel curves and consumer demand. *Review of Economics and Statistics*, 79(4):527–539.

Barkenbus, J. N. (2010). Eco-driving: An overlooked climate change initiative. *Energy Policy*, 38(2):762–769.

Becker, G. S., Grossman, M., and Murphy, K. M. (1994). An empirical analysis of cigarette addiction. *The American Economic Review*, 84(3):396–418.

Berry, A. (2019). The distributional effects of a carbon tax and its impact on fuel poverty: A microsimulation study in the french context. *Energy Policy*, 124:81 – 94.

Blundell, R. and Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87(1):115–143.

Blundell, R., Horowitz, J. L., and Parey, M. (2012). Measuring the price responsiveness of gasoline demand: Economic shape restrictions and nonparametric demand estimation. *Quantitative Economics*, 3(1):29–51.

Bomberg, M. and Kockelman, K. M. (2007). Traveler response to the 2005 gas price spike. In *86th Annual Meeting of the Transportation Research Board, Washington, DC*.

Buehler, R., Pucher, J., Gerike, R., and Götschi, T. (2017). Reducing car dependence in the heart of europe: lessons from germany, austria, and switzerland. *Transport Reviews*, 37(1):4–28.

Chamberlain, G. (1982). Multivariate regression models for panel data. *Journal of Econometrics*, 18(1):5–46.

Coglianese, J., Davis, L. W., Kilian, L., and Stock, J. H. (2017). Anticipation, Tax Avoidance, and the Price Elasticity of Gasoline Demand. *Journal of Applied Econometrics*, 32(1):1–15.

Cohen, J. D., Ericson, K. M., Laibson, D., and White, J. M. (2016). Measuring time preferences. Technical report, National Bureau of Economic Research.

Combet, E., Ghersi, F., Hourcade, J.-C., and Thubin, C. (2010). La fiscalité carbone au risque des enjeux d'équité. *Revue française d'économie*, 25(2):59–91.

Constantinides, G. M. (1990). Habit formation: A resolution of the equity premium puzzle. *Journal of Political Economy*, 98(3):519–543.

Cragg, J. G. and Donald, S. G. (1993). Testing identifiability and specification in instrumental variable models. *Econometric Theory*, 9(2):222–240.

Davis, L. W. and Kilian, L. (2011). Estimating the effect of a gasoline tax on carbon emissions. *Journal of Applied Econometrics*, 26(7):1187–1214.

Fang, H. A. (2008). A discrete–continuous model of households' vehicle choice and usage, with an application to the effects of residential density. *Transportation Research Part B: Methodological*, 42(9):736–758.

Fernández-Villaverde, J. and Krueger, D. (2007). Consumption over the life cycle: Facts from consumer expenditure survey data. *The Review of Economics and Statistics*, 89(3):552–565.

Filippini, M., Hirl, B., and Masiero, G. (2018). Habits and rational behaviour in residential electricity demand. *Resource and Energy Economics*, 52:137–152.

Friedman, M. (1957). The permanent income hypothesis. In *A theory of the consumption function*, pages 20–37. Princeton University Press.

Frondel, M., Ritter, N., and Vance, C. (2012). Heterogeneity in the rebound effect: Further evidence for germany. *Energy Economics*, 34(2):461–467.

Gillingham, K. (2011). How do consumers respond to gasoline price shocks? heterogeneity in vehicle choice and driving behavior. *Manuscript: Yale University.*

Greene, W. (2004). Fixed effects and bias due to the incidental parameters problem in the tobit model. *Econometric Reviews*, 23(2):125–147.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica*, pages 1029–1054.

Havranek, T. and Kokes, O. (2015). Income elasticity of gasoline demand: A meta-analysis. *Energy Economics*, 47:77–86.

Hilgert, M. A., Hogarth, J. M., and Beverly, S. G. (2003). Household financial management: The connection between knowledge and behavior. *Federal Reserve Bulletin*, 89:309.

Hoef, J. M. V. (2012). Who invented the delta method? *The American Statistician*, 66(2):124–127.

Kleibergen, F. and Paap, R. (2006). Generalized reduced rank tests using the singular value decomposition. *Journal of Econometrics*, 133(1):97–126.

Laporte, A., Karimova, A., and Ferguson, B. (2010). Quantile regression analysis of the rational addiction model: investigating heterogeneity in forward-looking behavior. *Health Economics*, 19(9):1063–1074.

Li, G., Schoeni, R. F., Danziger, S., and Charles, K. K. (2010). New expenditure data in the psid: Comparisons with the ce. *Monthly Labor Review*, 133(2):29–39.

Li, S., Linn, J., and Muehlegger, E. (2014). Gasoline taxes and consumer behavior. *American Economic Journal: Economic Policy*, 6(4):302–42.

Lusardi, A. and Mitchell, O. S. (2011). Financial literacy and planning: Implications for retirement wellbeing. Technical report, National Bureau of Economic Research.

Lusardi, A. and Tufano, P. (2015). Debt literacy, financial experiences, and overindebtedness. *Journal of Pension Economics & Finance*, 14(4):332–368.

Mannering, F. and Winston, C. (1985). A dynamic empirical analysis of household vehicle ownership and utilization. *The RAND Journal of Economics*, pages 215–236.

Metcalf, G. E. (1999). A distributional analysis of green tax reforms. *National Tax Journal*, pages 655–681.

Poterba, J. M. (1991). Is the gasoline tax regressive? *Tax policy and the economy*, 5:145–164.

Ross, M. L., Hazlett, C., and Mahdavi, P. (2017). Global progress and backsliding on gasoline taxes and subsidies. *Nature Energy*, 2(1):16201.

Scott, K. R. (2012). Rational habits in gasoline demand. *Energy Economics*, 34(5):1713–1723.

Spiller, E. (2012). Household vehicle bundle choice and gasoline demand: A discrete-continuous approach.

State of California (2012). California procedures for evaluating alternative specifications for phase 3 reformulated gasoline using the california predictive model.

Sterner, T. (2007). Fuel taxes: An important instrument for climate policy. *Energy Policy*, 35(6):3194–3202.

Sterner, T. (2012). Distributional effects of taxing transport fuel. *Energy Policy*, 41:75–83.

Stock, J. H. and Yogo, M. (2005). Testing for weak instruments in linear iv regression.

Suits, D. B. (1977). Measurement of tax progressivity. *American Economic Review*, 67(4):747–752.

Wadud, Z., Graham, D. J., and Noland, R. B. (2010a). Gasoline demand with heterogeneity in household responses. *The Energy Journal*, pages 47–74.

Wadud, Z., Noland, R. B., and Graham, D. J. (2010b). A semiparametric model of household gasoline demand. *Energy Economics*, 32(1):93–101.

West, S. E. (2004). Distributional effects of alternative vehicle pollution control policies. *Journal of Public Economics*, 88(3):735–757.

West, S. E. and Williams, R. C. (2012). Estimates from a consumer demand system: implications for the incidence of environmental taxes. In *Fuel Taxes and the Poor*, pages 98–125. RFF Press.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

Yatchew, A. and No, J. A. (2001). Household gasoline demand in Canada. *Econometrica*, 69(6):1697–1709.

Appendix A Proofs

A.1 Factorization of the lag polynomial

Starting from equation (5):

$$\begin{aligned} g_t &= \alpha_p L\, g_t + \alpha_f L^{-1}\, g_t + \beta p_t + \gamma x_t \\ -\left(\alpha_f L^{-1} - 1 + \alpha_p L\right) g_t &= \beta p_t + \gamma x_t \\ \left(1 - \frac{1}{\alpha_f} L + \frac{\alpha_p}{\alpha_f} L^2\right) g_t &= -\frac{1}{\alpha_f} L(\beta p_t + \gamma x_t) \end{aligned}$$

The roots of the left-hand side polynomial are:

$$\phi_1 = \frac{1 - \sqrt{1 - 4\alpha_p \alpha_f}}{2\alpha_f}, \quad \phi_2 = \frac{1 + \sqrt{1 - 4\alpha_p \alpha_f}}{2\alpha_f}$$

which lets us factorize the second-order lag polynomial into:

$$(1 - \phi_1 L)(1 - \phi_2 L)\, g_t = -\frac{1}{\alpha_f} L(\beta p_t + \gamma x_t)$$

The following geometric series expansion then yields equation (7):

$$\frac{-L}{1 - \phi_2 L} = \frac{1}{\phi_2} \frac{1}{1 - \frac{1}{\phi_2} L^{-1}} = \frac{1}{\phi_2} \sum_{i=0}^{\infty} \frac{1}{\phi_2^i} L^{-i}$$

A.2 Short and long-term responses

In this section, we examine the response to a price shock at time t, sustained for all $t' \geq t$, and considering all other household characteristics x_t constant for all $t' \geq t$.

Under the assumption of a constant path for p_t and x_t for all $t' \geq t$, equation (7) can be rewritten as follows:

$$
\begin{aligned}
g_t &= \phi_1 g_{t-1} + \frac{\beta p_t + \gamma x_t}{\alpha_f \phi_2} \sum_{\text{i = 0}}^{\infty} \frac{1}{\phi_2^i} \\
&= \phi_1 g_{t-1} + \frac{\beta p_t + \gamma x_t}{\alpha_f \phi_2} \frac{\frac{1}{\phi_2}}{1 - \frac{1}{\phi_2}} \\
&= \phi_1 g_{t-1} + \frac{\beta p_t + \gamma x_t}{\alpha_f \phi_2} \frac{\phi_2}{\phi_2 - 1}
\end{aligned}
\tag{A.1}
$$

Which gives equation the short-term member of (11). To determine long-term demand, we use the fact that g_∞ verifies:

$$
g_\infty = \phi_1 g_\infty + \frac{\beta p_t + \gamma x_t}{\alpha_f \left(\phi_2 - 1\right)} \tag{A.2}
$$

A.3 Short and long-term elasticities

By definition,

$$
\sigma_{short} = \frac{\partial g_t}{\partial p_t} \frac{p_t}{g_t} \tag{A.3}
$$

From (11), we have:

$$
\begin{aligned}
\frac{\partial g_t}{\partial p_t} &= \frac{\beta}{\alpha_f \left(\phi_2 - 1\right)} \\
&= \frac{\beta}{\alpha_f \left(\frac{1+\sqrt{1-4\alpha_p\alpha_f}}{2\alpha_f} - 1\right)} \\
&= \frac{2\beta}{1 - 2\alpha_f + \sqrt{1 - 4\alpha_p\alpha_f}}
\end{aligned}
\tag{A.4}
$$

Which gives equation (9). Similarly, we have:

$$
\begin{aligned}
\frac{\partial g_\infty}{\partial p_t} &= \frac{\beta}{\alpha_f\left(1-\phi_1\right)\left(\phi_2-1\right)} \\
&= \frac{\beta}{\alpha_f\left(1-\frac{1-\sqrt{1-4\alpha_p\alpha_f}}{2\alpha_f}\right)\left(\frac{1+\sqrt{1-4\alpha_p\alpha_f}}{2\alpha_f}-1\right)} \\
&= \frac{4\alpha_f\beta}{\left(2\alpha_f-1+\sqrt{1-4\alpha_p\alpha_f}\right)\left(1-2\alpha_f+\sqrt{1-4\alpha_p\alpha_f}\right)} \\
&= \frac{4\alpha_f\beta}{1-4\alpha_p\alpha_f-\left(4\alpha_f^2-4\alpha_f+1\right)} \\
&= \frac{4\alpha_f\beta}{4\alpha_f\left(1-\alpha_p-\alpha_f\right)}
\end{aligned}
\tag{A.5}
$$

Which yields equation (10).

Appendix B Complementary data on gasoline expenditure and prices

B.1 Comparison between the PSID and the CEX

Table B.1: Summary statistics for gasoline expenditure in the PSID and the CEX (1999-2015)

	PSID			CEX		
	Mean	Median	Std. dev.	Mean	Median	Std. dev.
1999	1,277.52	971.84	1,072.61	1,353.90	1,136.77	1,032.78
2001	1,720.45	1,550.43	1,495.77	1,638.04	1,381.11	1,399.63
2003	1,596.79	1,350.20	1,347.19	1,611.06	1,370.93	1,290.11
2005	2,110.33	1,708.77	1,784.44	2,295.90	1,946.18	1,758.02
2007	2,500.90	2,029.47	2,115.19	2,427.53	2,106.80	1,903.36
2009	1,923.02	1,576.53	1,682.69	2,099.14	1,846.50	1,639.17
2011	2,524.01	2,207.51	2,177.55	2,523.02	2,178.57	1,882.20
2013	2,408.46	1,954.80	2,080.88	2,482.65	2,230.14	1,842.07
2015	1,886.23	1,440.00	1,659.62	1,920.32	1,682.49	1,462.27

B.2 *Close-up map by state of gasoline retail prices*

Figure B.1: Localized gasoline prices in the state of California (2011)

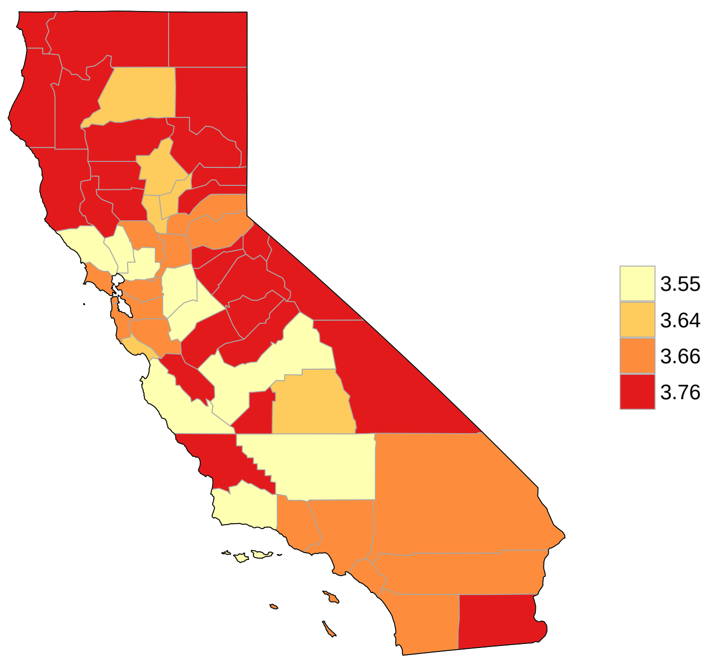
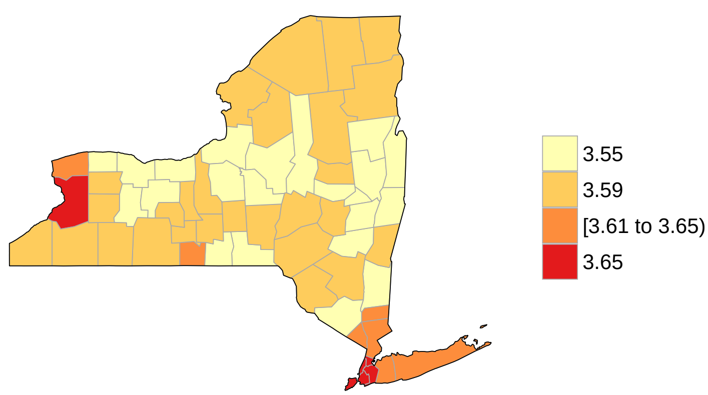


Figure B.2: Localized gasoline prices in the state of New York (2011)



B.3 Green share of the vote in congressional elections (1999-2015)

Figure B.3: Green share of the vote in the three previous congressional elections (1999)

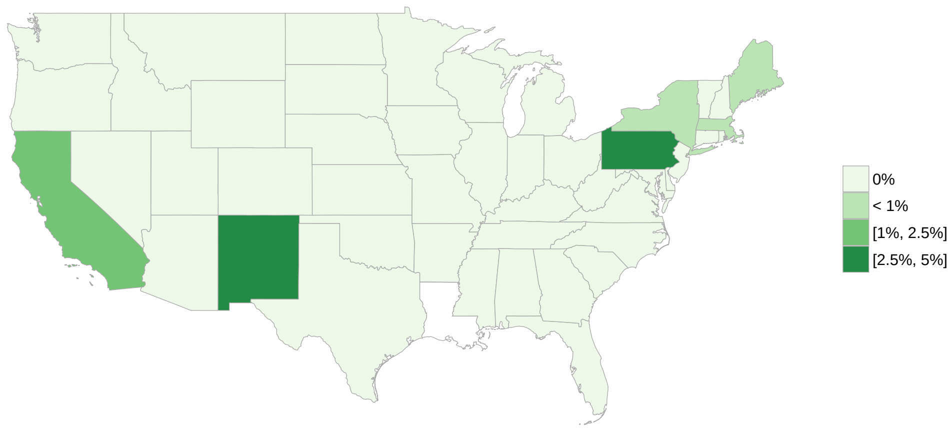


Figure B.4: Green share of the vote in the three previous congressional elections (2001)

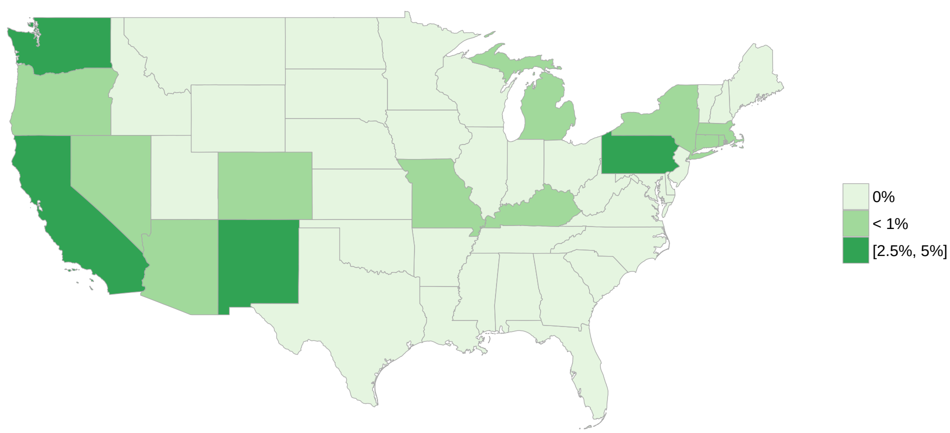


Figure B.5: Green share of the vote in the three previous congressional elections (2003)

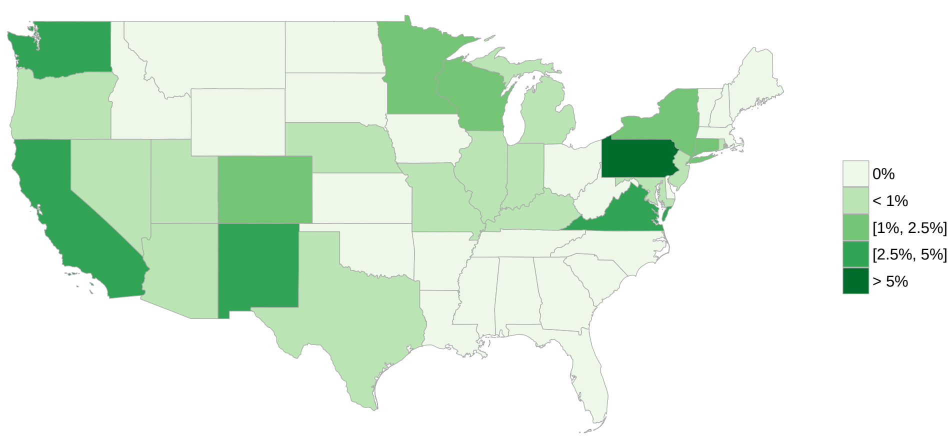


Figure B.6: Green share of the vote in the three previous congressional elections (2005)

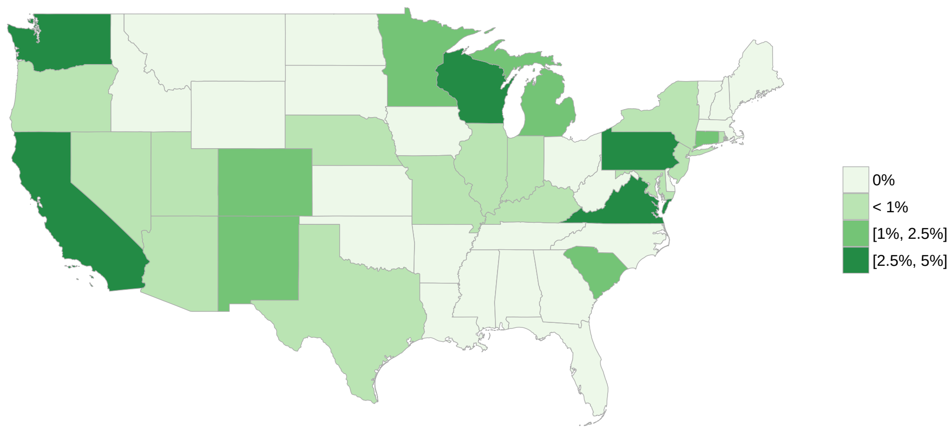


Figure B.7: Green share of the vote in the three previous congressional elections (2007)

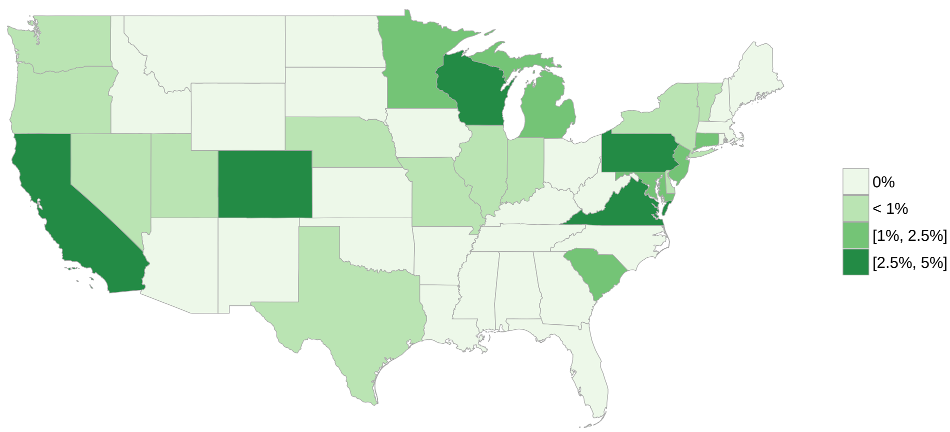


Figure B.8: Green share of the vote in the three previous congressional elections (2009)

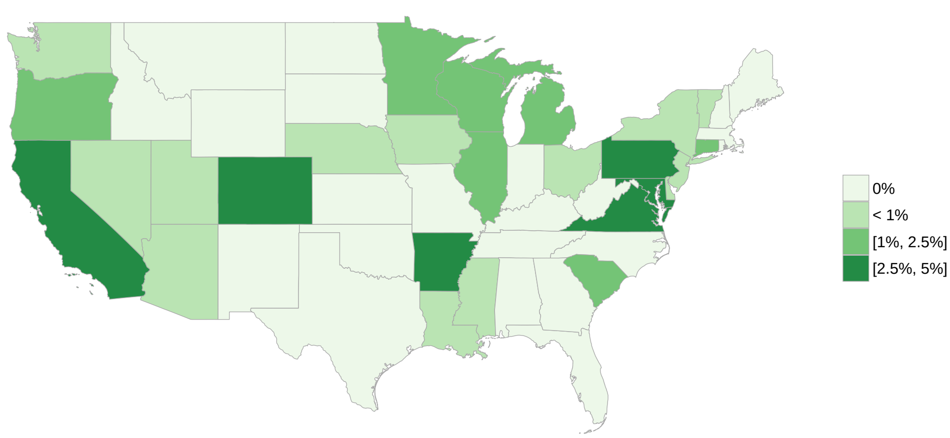


Figure B.9: Green share of the vote in the three previous congressional elections (2011)

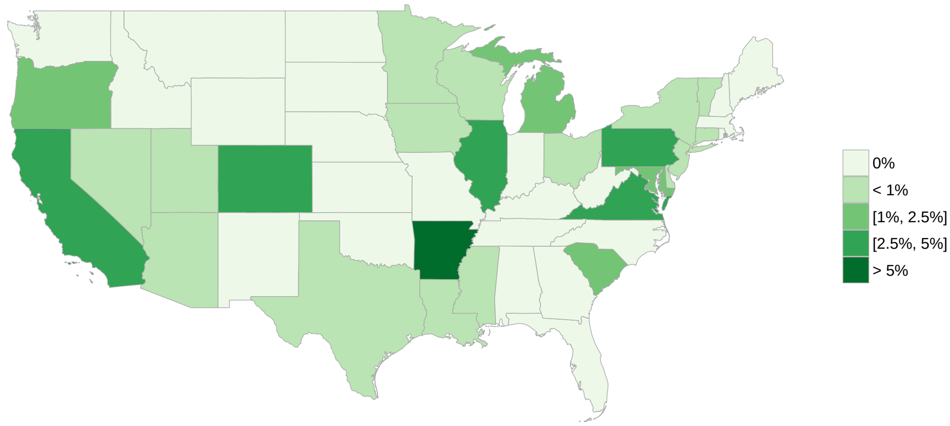


Figure B.10: Green share of the vote in the three previous congressional elections (2013)

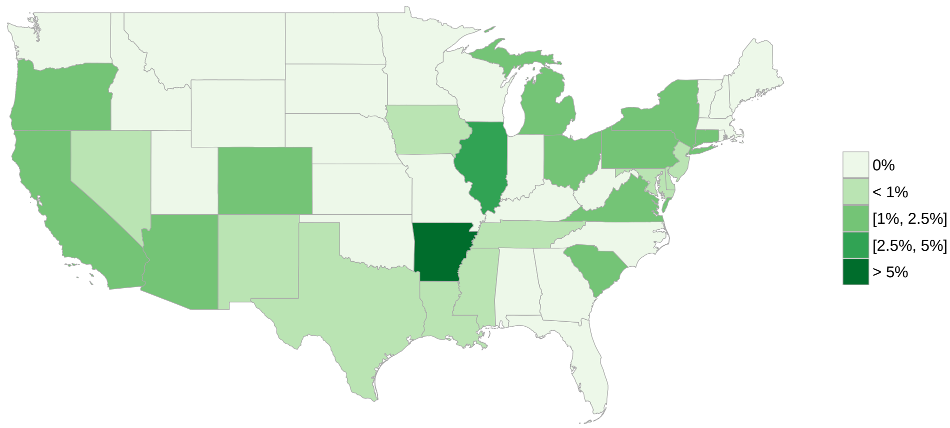
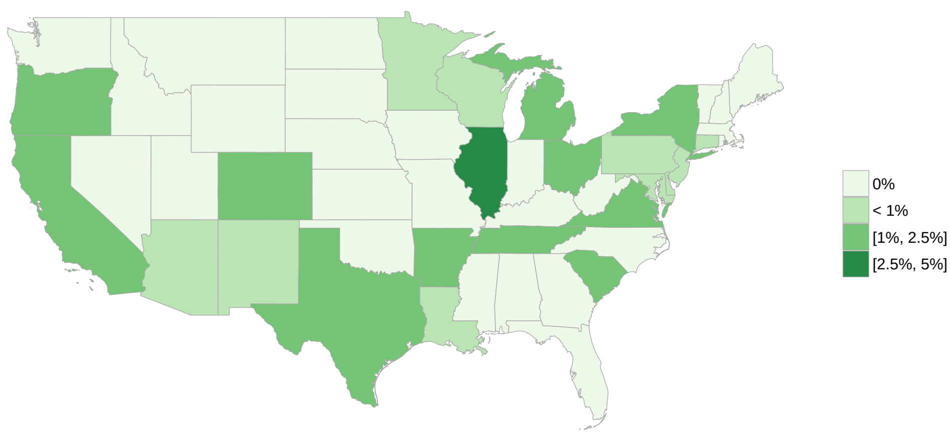


Figure B.11: Green share of the vote in the three previous congressional elections (2015)



B.4 Annual localized gasoline retail prices map (1999-2015)

Figure B.12: Localized retail gasoline prices in the continental U.S. (1999)

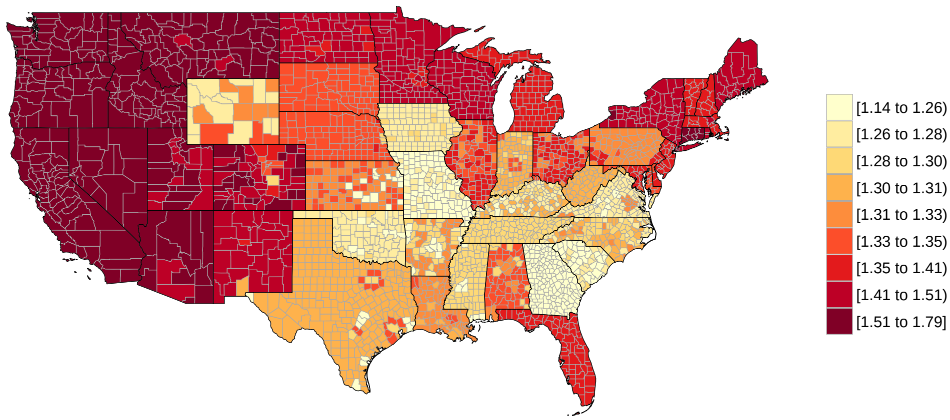


Figure B.13: Localized retail gasoline prices in the continental U.S. (2001)

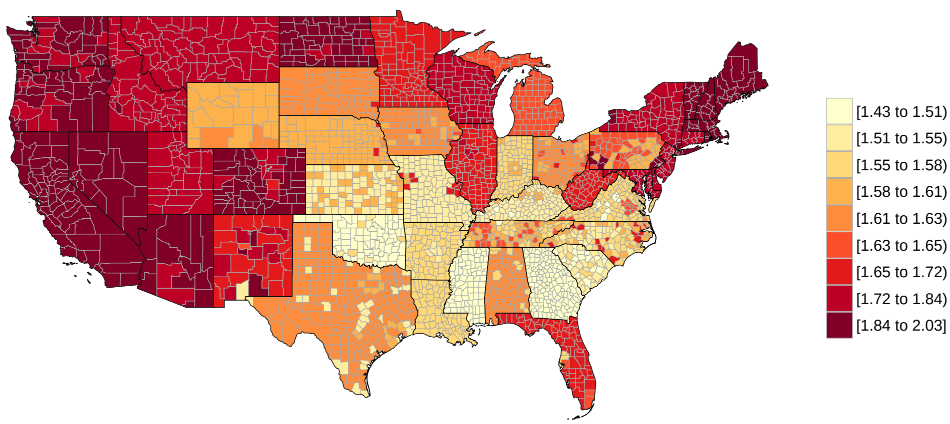


Figure B.14: Localized retail gasoline prices in the continental U.S. (2003)

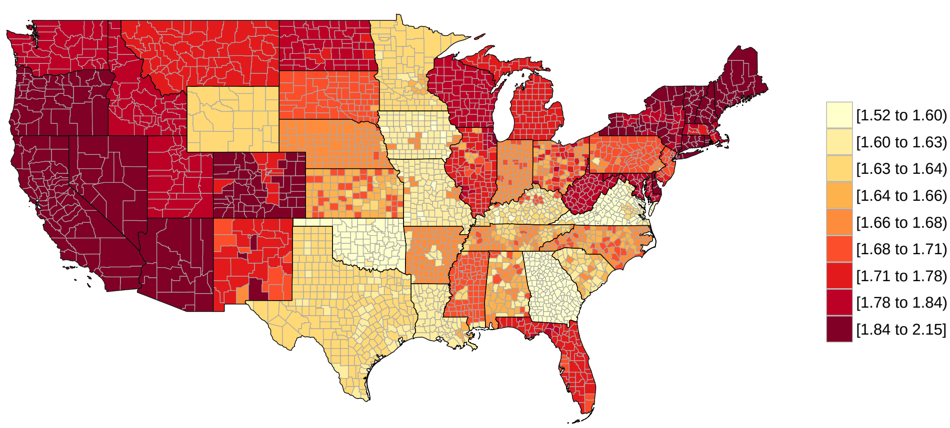


Figure B.15: Localized retail gasoline prices in the continental U.S. (2005)

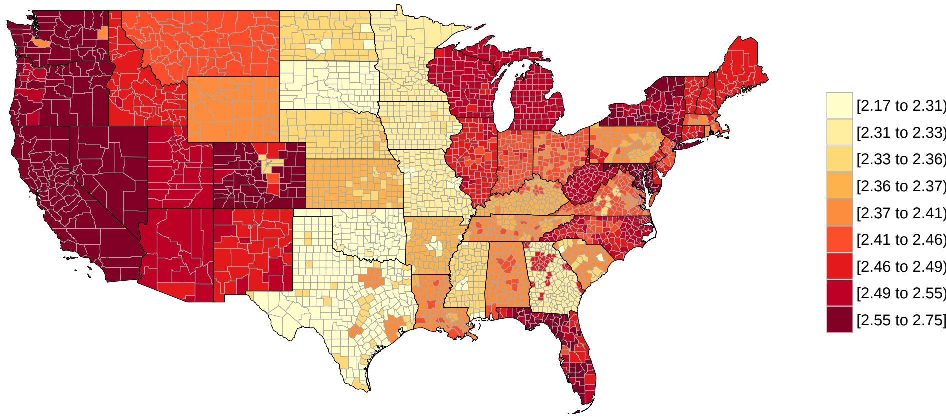


Figure B.16: Localized retail gasoline prices in the continental U.S. (2007)

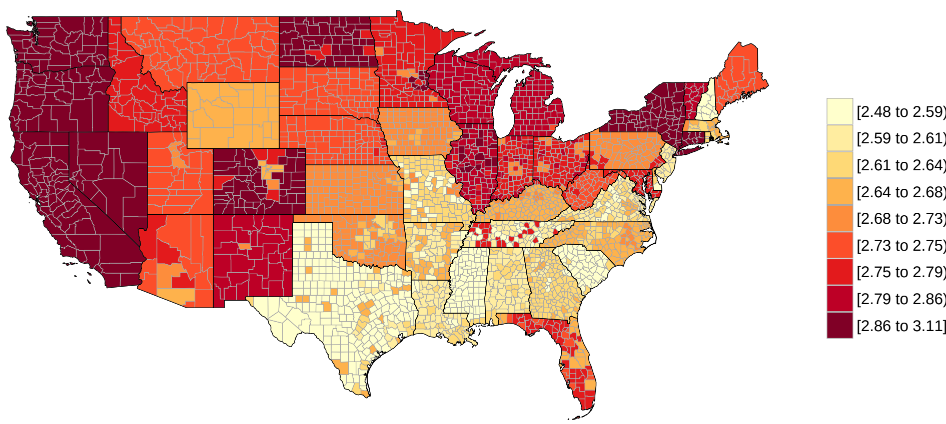


Figure B.17: Localized retail gasoline prices in the continental U.S. (2009)

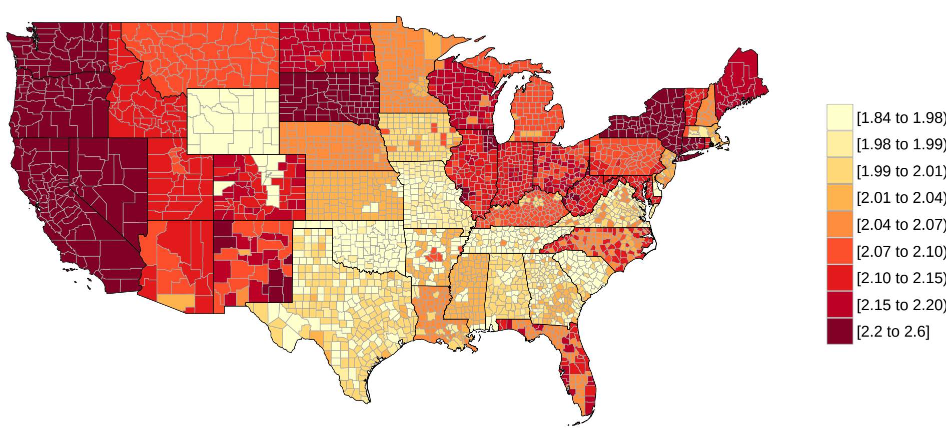


Figure B.18: Localized retail gasoline prices in the continental U.S. (2011)

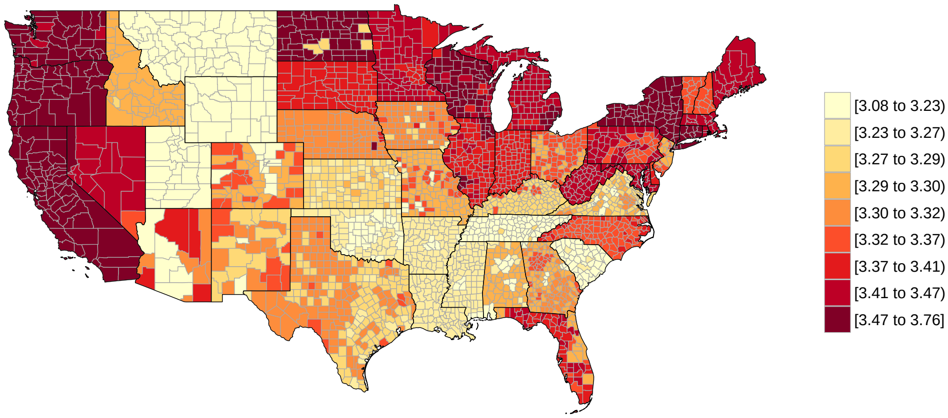


Figure B.19: Localized retail gasoline prices in the continental U.S. (2013)

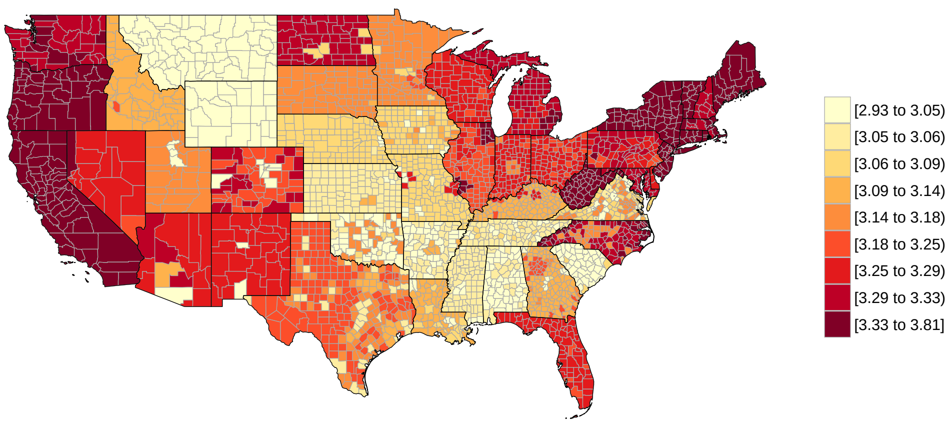
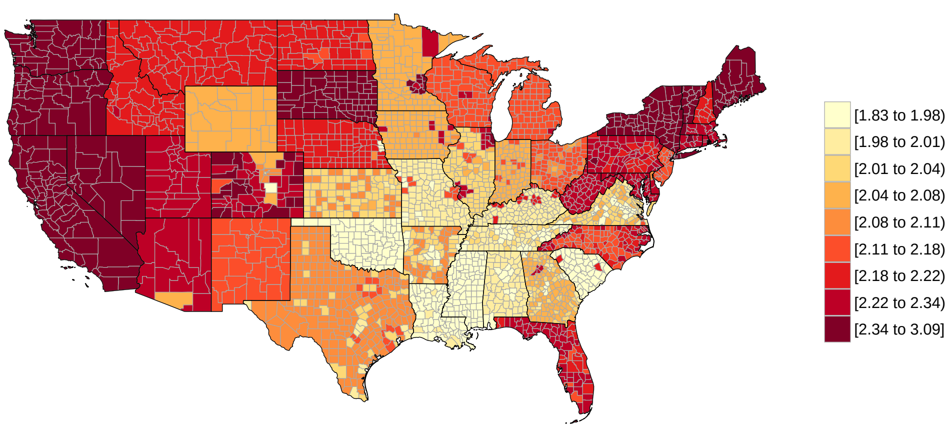


Figure B.20: Localized retail gasoline prices in the continental U.S. (2015)



Appendix C Complementary results

Table C.1: First stage results for the main specification of the rational habits model

	g_{t-1}	g_{t+1}
Income$_{t-1}$	0.00149***	0.000140
	(0.000150)	(0.000131)
Household size$_{t-1}$	87.45***	11.14*
	(5.680)	(4.919)
Non-metropolitan county$_{t-1}$	67.17*	74.31**
	(30.37)	(27.76)
p_{t-1}	67.01	90.57
	(59.21)	(51.57)
p_{t-2}	-216.7***	68.25
	(56.89)	(53.40)
Income$_{t+1}$	0.000526***	0.00142***
	(0.000149)	(0.000147)
Household size$_{t+1}$	-13.65*	81.14***
	(5.313)	(5.485)
Non-metropolitan county$_{t+1}$	16.01	58.82*
	(24.41)	(24.00)
p_{t+1}	-70.84	-60.39
	(54.55)	(52.22)
p_{t+2}	4.405	-166.2***
	(39.40)	(36.71)
p_t	23.44	-10.74
	(52.49)	(50.45)
Income	0.000798***	0.000390**
	(0.000157)	(0.000137)
Number of vehicles	27.12***	23.77***
	(6.025)	(5.857)
Household size	-14.38*	-7.781
	(5.936)	(5.552)
Non-metropolitan county	22.88	-20.59
	(33.82)	(30.56)
Age of head	-2.617	15.50
	(11.69)	(11.09)
Education	-0.0125	-0.399
	(0.317)	(0.289)
Green vote	-85.25	-66.49
	(409.5)	(391.8)
Time FE	Yes	Yes
F-test	41.87	40.03
Observations	31,308	31,308

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.2: Comparison of panel designs

	Minimum number of consecutive observations						
	(1) 3 waves	(2) 4 waves	(3) 5 waves	(4) 6 waves	(5) 7 waves	(6) 8 waves	(7) 9 waves
g_{t-1}	0.177***	0.173***	0.167***	0.157***	0.171***	0.188***	0.174***
	(0.0413)	(0.0416)	(0.0420)	(0.0442)	(0.0480)	(0.0552)	(0.0661)
g_{t+1}	0.126***	0.120**	0.127***	0.127**	0.129**	0.144**	0.182**
	(0.0464)	(0.0465)	(0.0475)	(0.0502)	(0.0553)	(0.0651)	(0.0752)
p_t	-186.9***	-189.4***	-197.6***	-185.5***	-197.1***	-205.2***	-209.8***
	(38.60)	(38.81)	(39.72)	(41.90)	(45.09)	(50.57)	(58.53)
Income	0.000479***	0.000496***	0.000532***	0.000564***	0.000532***	0.000441**	0.000440*
	(0.000152)	(0.000152)	(0.000155)	(0.000165)	(0.000180)	(0.000203)	(0.000230)
Number of vehicles	198.9***	200.3***	195.7***	194.0***	191.5***	182.9***	181.4***
	(6.660)	(6.640)	(6.863)	(7.237)	(7.952)	(8.857)	(10.07)
Household size	41.86***	41.73***	42.59***	42.49***	41.18***	35.74***	33.42***
	(4.993)	(5.016)	(5.148)	(5.370)	(5.862)	(6.491)	(7.645)
Non-metropolitan county	9.742	12.11	28.67	33.78	23.91	43.03	19.68
	(27.18)	(27.60)	(27.05)	(29.19)	(32.38)	(38.14)	(43.97)
Age of head	25.15**	25.06**	30.22***	33.89***	46.56***	46.39***	38.25**
	(10.33)	(10.28)	(11.35)	(12.08)	(13.64)	(15.60)	(18.03)
Education	0.00155*	0.00145*	0.00139	0.00196**	0.00121	0.00131	0.000363
	(0.000857)	(0.000859)	(0.000888)	(0.000971)	(0.00102)	(0.00120)	(0.00134)
Green vote	-12.53	-52.70	-124.2	-281.9	-335.0	-173.5	-451.6
	(326.1)	(330.2)	(334.0)	(344.3)	(357.4)	(374.6)	(417.7)
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F-test	169.1	164.6	156.3	146.2	130.3	110.3	95.8
Hansen's J	7.65	7.86	6.85	8.58	8.24	7.04	7.80
Hansen's J *p-value*	0.47	0.45	0.55	0.38	0.41	0.53	0.45
Observations	34,564	34,319	31,308	27,761	23,640	18,426	14,389
2-year elasticity	-0.702***	-0.701***	-0.728***	-0.673***	-0.71***	-0.742***	-0.793***
	(0.154)	(0.152)	(0.155)	(0.16)	(0.173)	(0.197)	(0.241)
Long-term elasticity	-0.858***	-0.852***	-0.878***	-0.801***	-0.861***	-0.919***	-0.967***
	(0.191)	(0.188)	(0.191)	(0.193)	(0.213)	(0.247)	(0.297)

Standard errors in parentheses.
Standard errors for elasticities estimates computed using the delta method.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table C.3: Estimation of the rational habits model by quintile of income

	Quintile of income				
	1	2	3	4	5
g_{t-1}	0.138	0.227***	0.177**	0.179**	0.0116
	(0.104)	(0.0844)	(0.0781)	(0.0866)	(0.0795)
g_{t+1}	0.110	0.0371	0.226***	0.281***	0.265***
	(0.0887)	(0.0828)	(0.0822)	(0.0949)	(0.0883)
p_t	-156.6***	-175.2***	-198.2***	-219.1***	-241.5***
	(40.60)	(41.02)	(41.07)	(42.18)	(42.04)
Income	0.00299***	0.00134**	0.000934*	0.000280	0.000354*
	(0.000762)	(0.000682)	(0.000499)	(0.000369)	(0.000205)
Household size			37.95***		
			(5.175)		
Non-metropolitan county			38.03		
			(26.98)		
Age of head			28.99**		
			(11.67)		
Number of vehicles			193.8***		
			(6.830)		
Education			0.00166		
			(0.00166)		
Green vote			-132.9		
			(333.0)		
F-test			83.2		
Observations			31,308		

Standard errors in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table C.4: Estimation of the rational habits model by quintile of income, excluding educational attainment

	Quintile of income				
	1	2	3	4	5
g_{t-1}	0.138	0.228***	0.179**	0.182**	0.0215
	(0.104)	(0.0844)	(0.0781)	(0.0865)	(0.0789)
g_{t+1}	0.108	0.0348	0.225***	0.279***	0.270***
	(0.0886)	(0.0828)	(0.0822)	(0.0950)	(0.0881)
p_t	-156.5***	-175.3***	-198.0***	-218.8***	-240.9***
	(40.60)	(41.02)	(41.07)	(42.19)	(42.03)
Income	0.00299***	0.00132*	0.000926*	0.000280	0.000358*
	(0.000761)	(0.000682)	(0.000499)	(0.000370)	(0.000205)
Household size			37.49***		
			(5.149)		
Non-metropolitan county			37.86		
			(26.96)		
Age of head			29.27**		
			(11.69)		
Number of vehicles			194.7***		
			(6.832)		
Green vote			-130.8		
			(332.6)		
F-test			85.8		
Observations			31,308		

Standard errors in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Appendix D Alternate specifications including household expenditure

Table D.1: Rational habits model of gasoline consumption: estimation results with total expenditure

	(1)	(2)	(3)	(4)
g_{t-1}	0.140***	0.138***	0.136***	0.135***
	(0.0382)	(0.0379)	(0.0379)	(0.0379)
g_{t+1}	0.122***	0.118***	0.117***	0.117***
	(0.0389)	(0.0382)	(0.0384)	(0.0384)
p_t	-200.6***	-198.9***	-198.8***	-196.0***
	(39.52)	(39.60)	(39.62)	(42.36)
Total expenditure	0.00147**	0.00147**	0.00147**	0.00147**
	(0.000668)	(0.000669)	(0.000668)	(0.000670)
Number of vehicles	192.7***	192.9***	193.0***	192.9***
	(7.016)	(7.024)	(7.019)	(7.019)
Household size	42.03***	42.23***	42.23***	42.21***
	(5.022)	(5.019)	(5.018)	(5.015)
Non-metropolitan county		32.07	32.27	30.52
		(27.11)	(27.14)	(27.76)
Age of head			30.33***	30.28***
			(11.26)	(11.23)
Education			0.0227	0.0254
			(0.250)	(0.250)
Green vote	-173.9	-156.4	-152.8	-158.3
	(335.2)	(336.3)	(336.8)	(344.4)
Time FE	Yes	Yes	Yes	Yes
Census division FE	No	No	No	Yes
F-test	186.8	173.1	151.5	101.5
Hansen's J	1.83	3.63	3.46	3.75
Hansen's J *p-value*	0.93	0.89	0.90	0.88
Observations	31,308	31,308	31,308	31,308
2-year elasticity	-0.732***	-0.721***	-0.719***	-0.709***
	(0.148)	(0.148)	(0.148)	(0.158)
Long-term elasticity	-0.853***	-0.838***	-0.834***	-0.822***
	(0.179)	(0.178)	(0.177)	(0.189)

Standard errors in parentheses.

Standard errors for elasticities estimates computed using the delta method.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table D.2: Estimation of the rational habits model by quintile of expenditure

	Quintile of expenditure				
	1	2	3	4	5
g_{t-1}	-0.0213	0.0156	0.00306	0.0743	0.123**
	(0.0401)	(0.0383)	(0.0407)	(0.0480)	(0.0625)
g_{t+1}	-0.0273	0.0224	-0.00145	0.0855	0.194***
	(0.0408)	(0.0394)	(0.0453)	(0.0548)	(0.0731)
p_t	-159.4***	-199.0***	-237.9***	-245.2***	-236.5***
	(40.09)	(40.52)	(41.74)	(41.14)	(41.18)
Total expenditure	0.0132***	0.0120***	0.00855***	0.00695***	0.000538**
	(0.000995)	(0.000802)	(0.000971)	(0.000612)	(0.000271)
Household size			30.71***		
			(4.783)		
Non-metropolitan county			52.66*		
			(28.28)		
Age of head			26.16**		
			(11.30)		
Number of vehicles			174.7***		
			(6.620)		
Education			-0.000897		
			(-0.000897)		
Green vote			-378.6		
			(348.1)		
F-test			95.7		
Observations			31,308		

Standard errors in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$