Gasoline demand and car choice: estimating gasoline demand using household information

Hilke A. Kayser*

Hamilton College, 198 College Hill Road, Clinton, NY 13323, USA

Abstract

In recent years, calls for carbon taxes as a policy tool to combat global warming have kept a discussion of the price and income elasticities of gasoline demand alive. To date, gasoline demand elasticity estimates are almost exclusively based on aggregate data that are subject to aggregation problems and make distributional concerns impossible to address. By using household-level data from the Panel Study of Income Dynamics (PSID) and imputed fuel efficiency measures and gasoline prices, I estimate household demand for gasoline and the corresponding price and income demand elasticities. An attempt is made to include data on the car stock in the estimation since car-portfolio and gasoline demand decisions are closely related. Empirical results from a selection corrected gasoline demand regression suggest low short-run price and income elasticities and clear differences in gasoline demand across the population. These results suggest that a gasoline tax is not likely to result in large decreases in gasoline consumption while potentially imposing hardship on identifiable segments of the population. © 2000 Elsevier Science B.V. All rights reserved.

JEL classification: H3; Q41

Keywords: Gasoline consumption; Energy; Price elasticity

1. Introduction

Understanding the determinants of gasoline demand has been of interest to economists for almost three decades. Initially, studies mainly addressed concerns...
about the availability of depletable resources and national security concerns raised by the oil supply shocks of the 1970s. Lately, studies increasingly address the various environmental consequences of gasoline consumption particularly with respect to the emission of greenhouse gases. A large number of studies have produced greatly varying estimates of the price elasticity depending on the specification and the data used. Almost all of the studies use data at the aggregate level for reasons including the limited nature of data at the household level. It is well known that aggregate data may not be appropriate for an estimation of household responses to higher gasoline prices because of aggregation problems. Thus, the effects of higher gasoline prices on the demand for gasoline by households in the United States are still not well understood. The effectiveness and equity of a number of current policy suggestions depend on a clear understanding of private driving and the responsiveness of gasoline demand to income and price changes, warranting a further investigation of household demand for gasoline.

In this paper, I provide estimates of the demand for gasoline and the associated demand elasticities by using household level data from the Panel Study of Income Dynamics (PSID). The PSID contains information about households’ annual miles traveled. Some standards by the Environment Protection Agency (EPA) for regulating tailpipe emissions of pollutants such as nitrogen oxides, carbon monoxide and hydrocarbons are set per mile and tend to be directly related to miles driven. However, the emission of other pollutants — most notably carbon dioxide — is directly related to gasoline consumption, not miles driven. In order to estimate the gasoline consumption we need to account for systematic differences in the fuel efficiency of the car fleet across the population. I therefore impute household-specific fuel efficiencies using data from the Survey of Consumer Finances (SCF) and information from the EPA. Gasoline demand is then obtained by dividing each household’s miles traveled by the household’s fuel efficiency estimate.

Estimating gasoline demand using household level data rather than aggregate data provides estimates that reflect more closely how individual consumers respond to changes in gasoline prices or household income, while allowing specific characteristics of a household to affect the demand for gasoline. As we will see, a household’s burden from higher gasoline prices is closely linked to observable characteristics of the household.

Besides using new and to a certain degree richer data sources, an important contribution of this research is the inclusion of data on the car stock. Since

---

1 For excellent surveys of the existing literature and the estimated price and income elasticities see Dahl (1986) and Dahl and Sterner (1991).

2 See for example the work by Khazzoom (1991) and Espey (1997) for a detailed discussion of the link between EPA standards as per mile requirement and emissions of these pollutants in response to gasoline taxes. Since manufacturers in an attempt to minimize costs will cut back on the use of emissions reduction equipment when their cars become more fuel efficient, the EPA requirements should lead to emissions per mile that just meet the requirement — regardless of the fuel efficiency of the car.
consumers receive utility from the transportation services that gasoline in combination with an automobile determine, gasoline demand is an indirect demand. Car-portfolio decisions and gasoline demand decisions are therefore related. Previous work using household data to estimate gasoline demand has taken the households’ car stock as given, focusing only on short-run responses to price changes in the form of reduced miles driven. It is of interest, however, to allow for changes in the car stock to affect the demand for gasoline. Though the data at hand does not provide much information about the car stock, a household’s decision to own a car or not will be included in the specification.

In Section 2, I discuss the conceptual framework and the estimation methods used in the empirical work before describing the data. Estimation results for gasoline demand and miles traveled are presented and discussed in Section 4. The paper ends with a concluding summary.

2. Conceptual framework and empirical methods

In this paper, gasoline demand and the demand for automobiles are modeled as a joint decision. Gasoline demand as a function of the price of gasoline is a consumer choice decision that in principle is made simultaneously with the consumer’s choice about the car portfolio. Similarly, the choice of car hinges on the household’s need for transportation services provided uniquely or most cheaply by a car. Two specification issues follow. First, the unit of observation in the empirical work is the household since driving is inherently a household decision. Second, the underlying consumer choices suggest the simultaneous estimation of two related decisions by estimating (a) a discrete choice model for the automobile portfolio; and (b) a continuous gasoline demand function that is conditional on the automobiles owned.

A small number of studies have used household-level data to investigate gasoline demand over the past two decades. Archibald and Gillingham (1980) estimate households’ gasoline demand in the United States using the 1973/1974 Consumer Expenditure Survey (CES), a five-quarter panel data set. In their analysis, Archibald and Gillingham explicitly estimate the short-run demand for households living in 23 larger Metropolitan Statistical Areas (MSAs). Gasoline consumption is calculated as the ratio of the reported gasoline expenditures divided by an average gasoline price from the Bureau of Labor Statistics for the 23 MSAs. Archibald and Gillingham’s price elasticity estimates are problematic because their dependent variable is measured with error. As they point out, the gasoline prices used to deflate gasoline expenditure and whose coefficients give rise to the estimated price elasticity are average gasoline prices. A household’s gasoline consumption is thus measured with error, and this error is obviously related to the price variable in the

3See also the work by Holmes (1976) and Hill (1980) who use data from the PSID to identify population groups that are affected most by higher gasoline prices.
vector of regressors, leading to biased estimates. The income elasticity is also somewhat problematic since Archibald and Gillingham use total expenditures as a proxy for lifetime income. Annual expenditures have been shown to be a poor proxy for permanent income (Hall, 1978; Menchik and David, 1982; Zeldes, 1989) such that the expenditure elasticity may not give a clear indication of the income elasticity of gasoline demand.

In another relevant study, Mannering and Winston (1985) use household data to simultaneously estimate automobile choice and usage. Price elasticities of the demand for miles traveled are estimated in the process of estimating households’ demand for specific automobile attributes, and the importance of brand preferences on the vehicle selection. For the time period between January 1979 and June 1980 they estimate a short-run price elasticity of \(-0.228\) and an insignificant short-run income elasticity of \(0.049\).

In this paper I build on previous work, using richer household data to simultaneously estimate the binary decision about whether or not to own a car, and of the continuous demand for gasoline.\(^4\) The underlying model best derives gasoline consumption from a household production of transportation services. Together with the household’s automobile stock and time, gasoline enters as an input into the production of the economic good of transportation services. This leads to two equations. First it gives rise to an indirect demand for gasoline that depends on the household’s economic situation, household characteristics that influence the taste or need for transportation services, car characteristics as captured by the fuel efficiency of the car fleet, and prices. Second, it gives rise to a binary choice equation that relates the household’s car ownership to household characteristics, the household’s economic situation, and prices.

For the gasoline demand equation, a Box–Cox test does not provide conclusive evidence as to the use of a log-linear specification or a linear one. Likelihood ratio tests present no preference for either a log-linear specification or a linear specification.\(^5\) For ease of comparison, this paper will thus follow the log-linear specification used by Archibald and Gillingham. The estimated equation for the short-run demand of gasoline is as follows:

\[
\ln(g_i) = \beta_0 + \beta_1 \ln(p_i) + \beta_2 \ln(m_i) + \beta_3 \ln(m_i)^2 + \beta_4 \ln(p_i) \ln(m_i) \\
+ \sum_{j=5}^{J} \beta_j W_{ij} + e_i \equiv X_i \beta + e_i
\]

where \(\ln(g)\) is the natural logarithm of gasoline consumption which is calculated dividing reported number of miles by the estimated fuel efficiency of the house-

\(^4\)Unfortunately, the data are not rich enough with respect to car stock information to allow the estimation of a more structural model.

\(^5\)The Box–Cox regression produces an estimate for lambda of 0.33. Likelihood ratio test for linearity and for log linearity lead to rejections of both the hypothesis that lambda is 0 or 1, respectively, at any significance level.
hold's car fleet and thus includes car characteristics, \( \ln(p) \) is the logarithm of the price of gasoline, \( \ln(m) \) is the logarithm of income, \( \ln(m)^2 \) is the square of this quantity, \( \ln(p) \ln(m) \) is an interaction term between the logarithm of income and the logarithm of price, and \( W_j \) is a vector of dummy variables to account for taste differences.\(^6\)

Included in the specification is an interaction term between income and the price of gasoline, following Archibald and Gillingham. By including the interaction term, I allow the price elasticity to vary across the income distribution. Poorer households who drive tend to spend a larger share of their income on gasoline and thus gasoline price changes have a larger impact for these households. Though it is possible that the relatively larger impact on their income makes the less well off more responsive to price changes, the fact that gasoline expenditure is such a large component of their overall expenditure may also be an indication that poorer households have very little choice in the amount of driving. It is probably fair to assume that the amount of driving they undertake, even though it imposes a large burden on them economically, must be necessary driving and cheaper than any available alternative, such as flying, taking the train or a bus. Wealthier households have less of a need to respond substantially to price changes, but they have more of an ability to do so. The price elasticity may therefore vary across the income distribution. For similar reasons, the income elasticity may also vary with the level of prices.\(^7\)

The automobile choice equation is discrete with the following general form:

\[
C_i = \gamma_0 + \gamma_1 p_i + \gamma_2 m_i + \gamma_3 p c_i + \sum_{k=4}^{K} \gamma_k W_{ik} + \nu_i \equiv Z_i' \gamma + \nu_i
\]  

(2)

where \( C_i \) is the unobserved car ownership demand variable which takes on the value one if the household owns at least one car, and zero if not, \( p c_i \) is the estimated price of a car for household \( i \), and \( W_{ik} \) is a vector of household characteristics to account for different preferences in the demand for cars.

In the empirical model I want to account for the fact that households may have a positive demand for gasoline but report no gasoline demand because they do not own a car, i.e. that car-ownership and gasoline consumption decisions are made jointly. This suggests using a two-stage estimation approach that was first discussed

---

\(^6\)Appendix A presents a discussion of biases from measurement error in the dependent variable of the sort that we would expect for the imputed gasoline demand variable in this study.

\(^7\)Car characteristics such as the vehicle stock or the fuel efficiency of the households' car fleet are not in the vector of explanatory variables. The unit of observation in the empirical work is the household since I consider the underlying demand for transportation services to be inherently a household decision. I am therefore estimating the demand for gasoline by a household with specific characteristics. The number of cars in a household is not included because it is rather highly correlated with the number of adults in the household and with the household's income. The fuel efficiency of the car fleet is not included as an explanatory variable because it is accounted for in the left-hand side variable. Gasoline demand is calculated as the reported miles driven divided by the household-specific fuel efficiency. I will discuss alternative specifications as part of the discussion of the empirical results later in the paper.
by Heckman (1976) and is known as the Heckman selection correction model. Gasoline demand is only observed when the household owns a car. Assume that \( \epsilon_i \) and \( \nu_i \) have a bivariate normal distribution with zero means and correlation \( \rho \). For identification, \( \sigma^2_{\nu} \) is normalized to one. Then the expected gasoline demand can be expressed as:

\[
E[\ln(g_i)|C_i > 0] = E[X_i^\prime \beta + \epsilon_i | W_i^\prime \gamma > u_i] = X_i^\prime \beta + \rho \sigma_\epsilon \frac{\phi(Z_i^\prime \gamma)}{\Phi(Z_i^\prime \gamma)}
\]

and the demand for gasoline becomes:

\[
\ln(g_i) = X_i^\prime \beta + \nu_i \lambda_i + \epsilon_i \lambda_i = \frac{\phi(Z_i^\prime \gamma)}{\Phi(Z_i^\prime \gamma)}
\]

In the Heckman selection correction specification, a Probit model is used for the car ownership equation to obtain an estimate for the selection correction term \( \lambda_i \). An ordinary least squares regression for the gasoline demand equation which includes that estimated selection correction term will lead to consistent estimates for the coefficients, \( \beta \), even in the case where the error terms are correlated.

For identification of the parameters in the two equations, one of two conditions must hold. Either the error terms have to be uncorrelated, or if the error terms are correlated, there has to be at least one variable in the vector of explanatory variables in the car choice equation, \( W_i \), that is not included in the vector of explanatory variables in the gasoline demand equation, \( X_i \) (Maddala, 1983, pp. 231–234). I include an imputed average price of cars for each state. The price of a vehicle should affect the likelihood of owning a car without having an effect on the amount of driving. Given the specification in Eq. (4) I can determine the short-run elasticities as follows:

\[
\eta_p = \frac{\delta E[\ln(g_i)|X_i]}{\delta p} = \beta_1 + \beta_4 \ln(\bar{m})
\]

\[
\eta_m = \frac{\delta E[\ln(g_i)|X_i]}{\delta m} = \beta_2 + 2\beta_3 \ln(\bar{m}) + \beta_4 \ln(\bar{p})
\]

where \( \bar{m} \) is the mean of average income and \( \bar{p} \) is the average gasoline price in the sample.

3. Data

To estimate the gasoline demand and car-ownership decisions I will use household data for 1981 from the Panel Study of Income Dynamics (PSID). The data may appear outdated and are certainly not representative any more in terms of the
characteristics of the car fleet that travels the roads in the United States. However, data from 1981 are the most recent data for one year in which gasoline prices were changing rather substantially. The year 1981 is the last year of a period of substantially fluctuating prices that followed the oil supply shocks of the late 1970s. Ever since 1981, gasoline prices in the United States have fallen gradually in real terms. The 1981 data will therefore predict household responses to more substantial increases in the price of gasoline through a carbon tax or any other form of gasoline tax more accurate than would estimates from more recent data.

The PSID is a panel data set that was started in 1968 with approximately 5000 American households. Since 1968 the PSID has collected data annually from the households in the original sample as well as from any new household that was formed by members of the original families. Each sampled household responds to approximately 2000 questions. Table A1 in Appendix B presents means and standard deviations for the variables that enter the empirical model. The list of explanatory variables includes family composition dummies since one would expect households with a larger number of adults to drive more, while households with children may both require additional transportation services and drive less by staying more in the vicinity of the home. Dummy variables indicating whether the household is a one-adult household or one with more adults, and indicating whether the household has no children, one child or several children are thus part of the list of independent variables.

Also included are a number of demographic variables that serve as proxies for unobservable taste differences. Among these characteristics are ethnicity, gender, age, marital status, and educational attainment of the head. These variables are chosen because they either significantly affect gasoline demand in the works by Archibald and Gillingham (1980), and Greening and Jeng (1994), or because they are suggested as playing a role in household’s driving through differential treatment by automobile insurance policies.

The list of variables includes two variables that define the living environment of the household: whether a household lives in a rural environment, and whether public transportation is available for the household members to get to work. It is likely that households living outside of metropolitan areas have different driving patterns than households who live in a large city or in the vicinity of a large city. Traveling distances are presumably longer for rural households, and households in rural areas are assumed to drive larger, less fuel-efficient cars. Both higher annual mileage and lower fuel efficiency should increase the gasoline demand for rural households.

Economic factors are the household’s income, the price of gasoline for the household, and the employment status of the head of the household and the

---

8 Throughout the study I will use as head those designated head in the PSID. In a cohabitating couple this is generally the husband or male partner.
9 A dummy rather crudely classifies households as living in a rural setting if they do not live in or around one of the 40 largest Metropolitan Statistical Areas (MSA) in the United States and if the nearest city has fewer than 50,000 inhabitants.
spouse. The appropriate measure of household income is controversial. Since annual income may fluctuate substantially from year-to-year, annual income may not be the appropriate variable to measure a household's well-being. Rather, according to Friedman's permanent income hypothesis a better measure of a household's income would be a measure that smooths out annual fluctuation but maintains variations across different stages in a household's life-cycle (Friedman, 1957). Accordingly, in the empirical work I use an income measure that averages household's income over the 11-year period from 1976 to 1986, i.e. an average centered around the reference year of 1981.10

Since the PSID does not contain information on gasoline prices, the prices used in this study are those developed by Chernik and Reschovsky (1992). Using data from the Bureau of Labor Statistics, Chernik and Reschovsky assign average retail prices directly to households living in or around any of the 30 larger cities in the United States for which gasoline price data are readily available. The rest of the sample is stratified according to four US regions and three city sizes, resulting in 42 different prices net of taxes. For each state, these prices are adjusted to include local, state and federal taxes when these are applicable. Because of regional differences in the cost-of-living, all other prices are represented by a regional price index that reflects cost of living differences by state, as published by Fournier and Rasmussen (1986) for the year 1980.

The PSID also does not contain information on gasoline consumption or on the gas mileage which could be used to estimate households’ gasoline consumption. Consequently, prior studies have calculated each household's gasoline consumption by dividing the reported miles by the average national fuel efficiency. However, this is a crude measure that ignores any variation in the fuel efficiency of car fleets across the population. In this study I adopt the following procedure to estimate gas mileage. Using the 1983 Survey of Consumer Finances (SCF) and the annual Gas Mileage Guides for New Car Buyers from the Environmental Protection Agency (1974–1986), I impute household specific gas mileage.11 The Survey of Consumer Finances contains information on the number of cars as well as the make, model, and vintage of the first three cars in the household. I assign the fuel efficiency values for each make, model, and year of a car from the Gas Mileage Guide to the corresponding cars in the SCF and take a simple average across the cars in a household to arrive at an average fuel efficiency value of the household’s car fleet for all households in the SCF. In a second step I run an ordinary least squares regression of the imputed miles per gallon (mpg), on a vector of explanatory variables, X, that includes household income, gender, ethnicity, and age of the household head, employment information, and information about the residential location of the household. The coefficients from the regression using data from the SCF, \(b_{\text{SCF}}\), are used to calculate predicted fuel efficiency values for each household in the PSID according to the following equation:

---

10 Not all households are part of the survey for all 11 years, in which case the average is calculated from the years that are available.

11 Some descriptive statistics for this imputed fuel efficiency variable can be found in Appendix C.
mpg_{83,PSID} = b'_{83,SCF} X_{83,PSID} \tag{7}

Table A2 in Appendix B shows the coefficients and standard errors of the mpg regression from which the coefficients $b_{83,SCF}$ are taken to impute fuel efficiency value for households in the PSID according to Eq. (7). Using this more disaggregate measure of fuel efficiency, gasoline consumption is calculated by dividing the reported miles driven by the household-specific fuel efficiency. Since an imputed value from another data set is used, gasoline consumption will be subject to errors in measurement that can lead to biased estimates.\textsuperscript{12}

I also impute the estimated cost of a car using the SCF. The SCF contains information about the bluebook value for each of the first three cars in the household. For households with more than one car, the average car value was constructed. Regressing the average car value on household characteristics such as work status, household income, region of residence, number of adults, number of children, gender, race, and marital status in the SCF, I estimate coefficients that I can use to impute the average car value for all households in the PSID based on their household characteristic. Households that do not own a car thus also have a value of a car assigned to them. This value serves as a proxy for the amount of money a household with the given characteristics can be expected to spend were they to buy a car.

4. Estimation results

Table 1 shows the empirical results from the two-stage Heckman estimation of the model described above. The first three columns show the results using the natural logarithm of miles traveled divided by the imputed fuel efficiency as the dependent variable. The last three columns present results for a regression of the natural logarithm of the reported miles traveled on the same list of explanatory variables, also from a Heckman specification. The results from both Heckman models show that the selection correction term, $\lambda$, is statistically significant at the 5% significance level, indicating that households indeed self-select into the group of car-owning households, and that this self-selection has a significant influence on gasoline demand and on the number of miles traveled.\textsuperscript{13}

In this study I am mainly interested in the demand for gasoline so that the discussion of the results will primarily focus on the first set of regression results. The coefficients on the explanatory variables in the gasoline demand equation are mostly as predicted. Since the dependent variable is a natural logarithm of gasoline demand, the coefficients can be interpreted as percentage changes in gasoline

\textsuperscript{12} Refer to Appendix A for a more detailed discussion of the bias from measurement error.

\textsuperscript{13} Information regarding the related car choice equation estimation is available from the author upon request.
Table 1
Two-stage Heckman selection correction model results for gasoline demand and miles traveled

<table>
<thead>
<tr>
<th>Variable</th>
<th>Logarithm of gasoline demand</th>
<th>Logarithm of miles traveled</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\beta$</td>
<td>$t$-ratio</td>
</tr>
<tr>
<td>Intercept</td>
<td>5.156</td>
<td>21.56</td>
</tr>
<tr>
<td>Logarithm of income (* $10,000)</td>
<td>1.074</td>
<td>4.35</td>
</tr>
<tr>
<td>Logarithm of income squared</td>
<td>$0.060$</td>
<td>2.08</td>
</tr>
<tr>
<td>Logarithm of gasoline price</td>
<td>1.684</td>
<td>1.97</td>
</tr>
<tr>
<td>Interaction of log price and log income</td>
<td>$-1.822$</td>
<td>1.86</td>
</tr>
<tr>
<td>Female head of household</td>
<td>$-0.360$</td>
<td>6.88</td>
</tr>
<tr>
<td>Public transportation available</td>
<td>$-0.226$</td>
<td>5.46</td>
</tr>
<tr>
<td>Head and spouse not working</td>
<td>$-0.213$</td>
<td>4.24</td>
</tr>
<tr>
<td>Rural residence</td>
<td>0.138</td>
<td>3.19</td>
</tr>
<tr>
<td>Unmarried head</td>
<td>0.068</td>
<td>1.07</td>
</tr>
<tr>
<td>One adult in the household</td>
<td>$-0.112$</td>
<td>2.13</td>
</tr>
<tr>
<td>One child in the household</td>
<td>0.073</td>
<td>1.60</td>
</tr>
<tr>
<td>Several children in the household</td>
<td>$-0.047$</td>
<td>1.10</td>
</tr>
<tr>
<td>Non-white head of household</td>
<td>$-0.123$</td>
<td>2.89</td>
</tr>
<tr>
<td>Age of household head (* 100)</td>
<td>4.309</td>
<td>7.95</td>
</tr>
<tr>
<td>Age squared</td>
<td>$-5.214$</td>
<td>9.19</td>
</tr>
<tr>
<td>Household head did not finish high</td>
<td>$-0.090$</td>
<td>1.38</td>
</tr>
<tr>
<td>Household head has a high school</td>
<td>$-0.065$</td>
<td>1.00</td>
</tr>
<tr>
<td>Beyond high school, no advanced</td>
<td>0.038</td>
<td>0.61</td>
</tr>
<tr>
<td>(selection correction term)</td>
<td>$-0.392$</td>
<td>2.88</td>
</tr>
<tr>
<td>$N$</td>
<td>4944</td>
<td></td>
</tr>
<tr>
<td>Price elasticity</td>
<td>$-0.23$</td>
<td></td>
</tr>
<tr>
<td>Income elasticity</td>
<td>0.49</td>
<td></td>
</tr>
<tr>
<td>Log likelihood</td>
<td>$-6862$</td>
<td></td>
</tr>
</tbody>
</table>

Consumption. Halvorsen and Palmquist (1980) show that the percentage impact of the dummy variables on gasoline demand is $e^{\beta} - 1$. Generally, $e^{\beta} - 1$ and $\beta$ tend to be close in magnitude, but they do not have to be.\textsuperscript{14}

The empirical results suggest that households headed by a woman consume approximately 30% less gasoline than households headed by a man, ceteris paribus. Though this value appears to be high, the reader should keep in mind that households with both male and female adults present during the interviewing period are coded as having a male head of household, leaving as female heads of household essentially only households with one or more female adults. Such female-only households are likely to be different in a variety of unobservable dimensions that I do not control for other than just their gender. Other findings

\textsuperscript{14}For example, the coefficient on one adult of $-0.112$ is close to the percentage change of $-0.106$, while the coefficient on female head of $-0.36$ differs more substantially from its percentage change of $-0.30$. 
indicate that households with a non-white head of household consume on average 11.6% less gasoline than their white counterparts, and that age has a positive and significant effect on gasoline consumption though at a decreasing rate. Having both the head and the spouse — if present — be unemployed lowers gasoline consumption significantly as does having only one adult in the household.

The living environment has a significant effect on gasoline consumption. As predicted, living in the presence of good public transportation tends to significantly lower gasoline consumption by as much as 20%. On the other hand, living in a rural setting increases gasoline consumption. The level of education and the number of children in the household do not appear to significantly affect gasoline consumption. Though it is intuitively reasonable to assume that children affect a household’s driving pattern and car choice, it appears as though the various plausible behavioral responses to having children cancel each other out, leaving households with children with a demand for gasoline that is very similar to that of households who do not have children.

The coefficients involving income are also as expected. The results indicate that households with higher incomes tend to consume more gasoline but that the additional consumption comes at a decreasing rate. The estimated income and price elasticities are located at the bottom of the table. Recall from Eqs. (5) and (6) that the price elasticity for this specification is a function of average income, and the income elasticity is a function of both the price of gasoline and the average income. Evaluated at a mean average income of $28,667 and an average gasoline price of $1.29, the estimated short-run price elasticity is -0.23 and the estimated short-run income elasticity is 0.48. A Wald test for joint significance of the two coefficients involving the price of gasoline reveals that the coefficients and thus the price elasticity estimates are jointly marginally significant at the 10% significance level. A similar test for the income coefficients reveals that the income coefficients are significant at the 5% significance level.

The interaction term between the price of gasoline and income implies that the income elasticity is lower when prices are higher, and that the price elasticity is greater at higher levels of income. The coefficient confirms the suspicion that households with lower incomes do not respond as much to higher gasoline prices as wealthier households. As discussed earlier, it is not unreasonable to assume that households whose burden from gasoline expenditures is relatively high because of their low incomes are likely to have reduced their gasoline consumption to the amount that is necessary for them to get to work or to town, leaving little room to respond noticeably to higher prices. Though wealthier households may have less of an incentive to respond to higher gasoline prices, they are probably more able to do so because some of their consumption is less of a necessity.

Results from the regression of miles driven serve to decompose the effects of the

---

15 To evaluate price and income elasticity at the mean, I use unweighted average income instead of the weighted average presented in Table A1.
explanatory variables on gasoline demand into effects on the quantity of driving and on the efficiency with which the miles were driven, as well as to make sure that the estimation results are reasonable and not an artifact from the imputed fuel efficiency measure. Regression results using reported miles driven as the dependent variable are fairly similar to those for gasoline demand. Households headed by a woman drive significantly fewer miles than households headed by a man, marital status and the number of children has no significant effect on miles driven, and education plays at best a marginal role in determining the amount of driving. Households who live in a rural setting tend to drive more miles, while access to public transportation lowers the amount of miles driven. Age affects miles driven positively but at a decreasing rate, and households with a non-white head or without work tend to drive fewer miles. Income has a positive effect on miles driven but at a decreasing rate. The income elasticity is significant and with a magnitude of 0.46 similar to the one estimated for gasoline demand.

A number of factors are different across the two regressions, however. Most importantly, the price coefficients and the price elasticity are not statistically significant at all. The \( \chi^2 \)-test for joint significance of the coefficients involving the price of gasoline shows a significance level of 0.40. At least in the sample at hand, the price of gasoline has no significant effect on the number of miles households drive though it had a marginally significant effect on gasoline consumption. Households do not appear to respond to higher prices by reducing the amount of driving but rather by traveling the same miles more efficiently by shifting transportation to the most fuel-efficient car in their fleet, or — in the longer run — by replacing older and less fuel-efficient cars with newer ones.

Other noteworthy results from this comparison are that rural households not only drive longer distances but also do so in less fuel-efficient cars, and that households with a non-white head drive fewer miles while consuming less gasoline, suggesting that white households drive more but do so in more fuel-efficient cars. A very similar result holds for unemployed households when compared to households where the head of household or the spouse are employed. Though the former travel approximately 30% fewer miles, they consume only 19% less gasoline.

In another attempt to see whether using the imputed fuel-efficiency measure may lead to bias because of errors in measurement, a separate regression was run using as the dependent variable the natural logarithm of gasoline consumption where gasoline consumption was calculated by deflating the miles driven by the national average fuel efficiency of 15.94 mpg. The results obtained are very similar to the results using miles traveled, with a significant income elasticity of 0.47. Again the price coefficients fail to be statistically significant with a significance level of 0.38.\(^{16}\)

\(^{16}\)The log likelihood for the regression using average miles per gallon to calculate gasoline demand is rather similar to that using the imputed fuel efficiency — 6899.
5. Conclusions

The results from the empirical work in this study confirm that the price elasticity of gasoline demand is low in the short-run. The price elasticity found in this study is lower than the corresponding price elasticities found by Archibald and Gillingham, the only other study using household-level data to estimate a gasoline demand equation. While the short-run price elasticity estimate is $-0.23$ when evaluated at mean prices and mean income, Archibald and Gillingham estimate a price elasticity of $-0.254$ for households with one car and $-0.376$ for households with several cars. Both studies indicate, however, that higher gasoline prices will not lead to a substantial reduction in the amount of gasoline consumed by households in the short-run.

Gasoline demand is calculated by deflating the reported miles traveled by the imputed household-specific fuel efficiency. Estimating the demand for miles traveled separately reveals a positive and highly insignificant price elasticity. If we believe that the imputed household-specific fuel efficiencies capture some of the true variation in fuel efficiency across the population, then this result suggests that gasoline demand responds to changes in the price of gasoline mainly through adjustments in the fuel efficiency of the car fleet and not through an adjustment of miles traveled.

The estimated gasoline demand equation provides short-run income elasticity estimates of 0.49, and the miles traveled equation provides short-run income elasticity estimates of 0.48. In both cases the income elasticity is highly significant, and they are quite similar in magnitude. The magnitude is also in line with other estimates of the short-run income elasticity. The fuel efficiency regression shows that income is a significant explanatory variable for the fuel efficiency of a household's car fleet, and Appendix C also shows a consistently positive relationship between income and fuel efficiency for each of the chosen subgroups. It appears that higher income allows households to purchase newer cars that will on average be more fuel-efficient because cars in 1981 are subject to the corporate fuel efficiency standards.

Finally, the empirical results show that there are clear differences in gasoline demand across the population. Households living in a rural setting and households with no public transportation available for travel to work will on average be affected more strongly by higher gasoline prices than similar households in an urban setting and with access to public transportation. Households whose head and/or spouse are working also consume significantly more gasoline. The working poor who have to commute to work by car and have no access to public transportation, for example, will be hurt more by raising prices than non-working poor who live off transfer programs. Information of this kind can be very valuable to policymakers who are interested in raising gasoline taxes for a variety of perfectly

17Dahl and Sterner (1991) show short-run income elasticities varying from 0.39 to 0.52.
legitimate reasons but recognize the need for adequately compensating those who are economically disadvantaged and may be disproportionately affected by higher gasoline costs.

Future research in the direction of making the car choice equation richer appears desirable. I would assume that most responses to higher gasoline prices in terms of a household’s car fleet will not occur at the extensive margin but rather at the intensive margin. For example, households are likely to purchase more fuel-efficient cars when gasoline prices are high and they are in the market for replacing their car, and it is likely that households with several cars will shift some of the uses from the less fuel-efficient car to the more fuel-efficient one. Capturing effects on gasoline demand of these kinds of adjustments to price changes requires significantly more detailed information about a household’s car fleet than is available in the data used for this study.

Appendix A: Measurement error

The dependent variable in this paper is the gallons of gasoline a household consumed in the previous year. This variable is the ratio of the reported miles driven by the household divided by the imputed fuel efficiency measure. The fuel efficiency variable is measured with error and can thus be expressed as the true, unobserved fuel efficiency value, mpg*, and the equally unobserved measurement error, v. The effect of the measurement error on the estimated coefficients can be demonstrated using for simplicity a simple regression of gasoline consumption on the price of gasoline. Let \( Y \) be the logarithm of gasoline consumption. Then \( Y \) can be expressed as:

\[
Y = \ln\left(\frac{\text{miles}}{\text{mpg}}\right) = \ln\left(\frac{\text{miles}}{\text{mpg}^* + v}\right) \equiv \ln(\text{gallons})^* - \eta = Y^* - \eta \quad (A1)
\]

The regression equation becomes:

\[
Y^* - \eta = X^T \beta + \epsilon \quad (A2)
\]

and the regression coefficients can be calculated as:

\[
b = \frac{\text{cov}(x, y)}{\text{var}(x)} = \frac{E[xy] - E[x]E[y]}{\text{var}(x)} = \frac{E[(y^* - \eta)x] - E[(y^* - \eta)E[x]]}{\text{var}(x)}
\]
\[
\begin{align*}
    &= \frac{E[y^* x] - E[(y^* - \eta) x]}{\text{var}(x)} \\
    &= \beta - \frac{E[\eta x]}{\text{var}(x)}
\end{align*}
\]

The direction of the bias from a measurement error thus depends on the covariance between the unobserved error and the explanatory variable, i.e. the coefficient will be biased negatively if \(E[\eta x] > 0\) and negatively if \(E[\eta x] < 0\). The magnitude of the bias will also be influenced by the variance of the regressor.

**Appendix B:**

**Table A1:** Descriptions and weighted means for the 1981 variables used in the regression model

<table>
<thead>
<tr>
<th>Variable</th>
<th>All households in the sample</th>
<th>Car-owning households only</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average mean income in 1982 dollars</td>
<td>27,952</td>
<td>30,738</td>
</tr>
<tr>
<td>Average gasoline consumption</td>
<td>869</td>
<td>1015</td>
</tr>
<tr>
<td>Average car cost by state</td>
<td>2977</td>
<td>3090</td>
</tr>
<tr>
<td>Proportion of non-white household heads</td>
<td>0.12</td>
<td>0.09</td>
</tr>
<tr>
<td>Proportion of unmarried household heads</td>
<td>0.41</td>
<td>0.33</td>
</tr>
<tr>
<td>Proportion of female household heads</td>
<td>0.27</td>
<td>0.20</td>
</tr>
<tr>
<td>Proportion with rural residency</td>
<td>0.25</td>
<td>0.27</td>
</tr>
<tr>
<td>Proportion living in the south</td>
<td>0.30</td>
<td>0.30</td>
</tr>
<tr>
<td>Proportion living in the west</td>
<td>0.19</td>
<td>0.20</td>
</tr>
<tr>
<td>Proportion living in north-central US</td>
<td>0.29</td>
<td>0.29</td>
</tr>
<tr>
<td>Proportion of households with one adult</td>
<td>0.33</td>
<td>0.26</td>
</tr>
<tr>
<td>Proportion of households with two adults</td>
<td>0.54</td>
<td>0.60</td>
</tr>
<tr>
<td>Proportion of households with more than two adults</td>
<td>0.13</td>
<td>0.14</td>
</tr>
<tr>
<td>Proportion of households with no children</td>
<td>0.60</td>
<td>0.57</td>
</tr>
<tr>
<td>Proportion of households with one child</td>
<td>0.16</td>
<td>0.17</td>
</tr>
<tr>
<td>Proportion of households with several children</td>
<td>0.24</td>
<td>0.25</td>
</tr>
<tr>
<td>Proportion of households with head not working</td>
<td>0.27</td>
<td>0.21</td>
</tr>
<tr>
<td>Proportion with head with no high school degree</td>
<td>0.31</td>
<td>0.26</td>
</tr>
<tr>
<td>Proportion with head with high school degree</td>
<td>0.18</td>
<td>0.19</td>
</tr>
<tr>
<td>Proportion with head with additional education, no degree</td>
<td>0.32</td>
<td>0.34</td>
</tr>
<tr>
<td>Proportion with high quality public transportation</td>
<td>0.40</td>
<td>0.36</td>
</tr>
<tr>
<td>Head's age</td>
<td>48</td>
<td>46</td>
</tr>
<tr>
<td>Number of observations</td>
<td>4,944</td>
<td>3,979</td>
</tr>
</tbody>
</table>

**Table A2:** Ordinary least squares regression of fuel efficiency using data from the 1983 survey of consumer finances (dependent variable: miles per gallon)

---

The descriptive statistics presented here are weighted means. The PSID oversamples poor, non-elderly households and provides family weights to adjust to make the sample more representative of the US population.
Variable | $\beta$ | Standard error
---|---|---
Intercept | 16.68* | (0.462)
Household income | 0.0225* | (0.006)
Household income² | − 0.0005* | (0.00002)
Value of the house | − 0.0046* | (0.002)
Number of cars | − 0.368* | (0.104)
Two adults employed | 0.665* | (0.208)
No adult in household employed | − 0.889* | (0.298)
Living in west | − 0.120 | (0.268)
Living in south | − 0.203 | (0.228)
Living north-central | − 0.753* | (0.232)
More than three children | − 0.966* | (0.514)
Three children | − 0.231 | (0.356)
On child | 0.411 | (0.264)
No child | 0.443* | (0.230)
More than three adults | 0.277 | (0.365)
Three adults | − 0.084 | (0.253)
One adult | − 0.099 | (0.353)
Head has a college degree | 1.637* | (0.223)
Head went to college, no degree | 0.906* | (0.234)
Head did not complete high school | − 0.802* | (0.216)
Head is Caucasian | 1.621* | (0.492)
Head is not Caucasian and not black | 1.440* | (0.280)
Head is male | − 0.342* | (0.165)
Head has never been married | 0.392 | (0.429)
Head is widowed | − 1.158* | (0.417)
Head is divorced | 0.642 | (0.394)
Head is separated | 0.239 | (0.570)
Head own a house | − 0.357 | (0.215)

**Appendix C**

*Table A1: Distribution of the estimated fuel efficiency across gender, age and ethnicity in the 1983 PSID*

<table>
<thead>
<tr>
<th>Number</th>
<th>Average income 10 000</th>
<th>Average income 10 000–15 000</th>
<th>Average income 15 000–20 000</th>
<th>Average income 20 000–30 000</th>
<th>Average income 30 000–40 000</th>
<th>Average income &gt; 40 000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male head</td>
<td>3189</td>
<td>15.38</td>
<td>15.58</td>
<td>15.61</td>
<td>15.92</td>
<td>16.21</td>
</tr>
<tr>
<td>Male head</td>
<td>815</td>
<td>15.90</td>
<td>15.93</td>
<td>16.23</td>
<td>16.35</td>
<td>16.55</td>
</tr>
<tr>
<td>Male head</td>
<td>3640</td>
<td>15.91</td>
<td>16.14</td>
<td>16.21</td>
<td>16.25</td>
<td>16.41</td>
</tr>
<tr>
<td>Male head</td>
<td>364</td>
<td>15.56</td>
<td>15.20</td>
<td>15.09</td>
<td>15.23</td>
<td>15.48</td>
</tr>
<tr>
<td>Head age 25–44</td>
<td>2041</td>
<td>15.89</td>
<td>15.83</td>
<td>15.96</td>
<td>16.09</td>
<td>16.2</td>
</tr>
<tr>
<td>Head age 45–64</td>
<td>1156</td>
<td>15.68</td>
<td>15.40</td>
<td>15.40</td>
<td>15.63</td>
<td>16.14</td>
</tr>
<tr>
<td>Head age &gt; 64</td>
<td>470</td>
<td>15.54</td>
<td>15.43</td>
<td>15.30</td>
<td>15.60</td>
<td>15.87</td>
</tr>
<tr>
<td>All</td>
<td>4004</td>
<td>15.71</td>
<td>15.71</td>
<td>15.79</td>
<td>15.99</td>
<td>16.24</td>
</tr>
</tbody>
</table>
Table A1 shows the distribution of miles per gallon (mpg) across a few different demographic groups in the 1983 SCF. In all but the highest income bracket, households with a female household head drive more fuel-efficient cars than households with a male head, and households with a white head of household drive more fuel-efficient cars than households with a non-white head. Fuel efficiency of the car fleet tends to decrease with the age of the head while it increases with average income. These results and further results regarding the distribution across regions, education of the head of household, and rural or urban location coincide with the results found by Greenlees (1980).\textsuperscript{16} Since the distribution of the imputed fuel efficiency in the 1981 PSID is similar to the one Greenlees finds in his study, it appears that at least some of the distributional differences in the fuel efficiency across different population groups have been captured correctly through the imputation.

References


\textsuperscript{16}Greenlees uses the 1973/1974 Consumer Expenditure Survey to model the choices of individual households between four, six, and eight cylinder cars. The number of cylinders is highly correlated with the fuel efficiency of the car.


