Energy Economics 67 (2017) 346–354

Contents lists available at ScienceDirect

Energy Economics

journal homepage: www.elsevier.com/locate/eneeco

A Bayesian sampling approach to measuring the price responsiveness of gasoline demand using a constrained partially linear model

Haotian Chen [a], Russell Smyth [b,*], Xibin Zhang [a]

[a] Department of Econometrics and Business Statistics, Monash University, Australia
[b] Department of Economics, Monash University, Australia

ARTICLE INFO

Article history:
Received 26 August 2016
Received in revised form 24 August 2017
Accepted 29 August 2017
Available online 8 September 2017

JEL classification:
C14
C11
D12
H31

Keywords:
Kernel estimator
Markov chain Monte Carlo
Price elasticity
Slutsky condition
Smoothness

ABSTRACT

Partial linear models provide an intuitively appealing way to examine gasoline demand because one can examine how response to price varies according to the price level and people's income. However, despite their intuitive appeal, partial linear models have tended to produce implausible and/or erratic price effects. Blundell et al. (2012) propose a solution to this problem that involves using Slutsky shape restrictions to improve the precision of the nonparametric estimate of the demand function. They propose estimating a constrained partially linear model through three steps, where the weights are optimized by minimizing an objective function under the Slutsky constraint, bandwidths are selected through least squares cross-validation, and linear coefficients are estimated using profile least squares. A limitation of their three-step estimation method is that bandwidths are selected based on pre-estimated parameters. We improve on the Blundell et al. (2012) solution in that we derive a posterior and develop a posterior simulation algorithm to simultaneously estimate the linear coefficients, bandwidths in the kernel estimator and the weights imposed by the Slutsky condition. With our proposed sampling algorithm, we estimate a constrained partially linear model of household gasoline demand employing household survey data for the United States for 1991 and 2001 and for Canada for 2006–2009 and find plausible price effects.

1. Introduction

Estimating the demand for gasoline is an important research topic given the growth in motor vehicle ownership and resultant environmental implications of carbon dioxide emissions from this source. Of particular concern to policy makers is the effect of price changes on gasoline demand. The responsiveness of motor vehicle users to rises in the price of gasoline dictates the extent to which governments can impose taxes in order to raise revenue, encourage conservation through discouraging demand for gasoline and encouraging switching in favour of cleaner fuel sources, as well as attain certain national security objectives in terms of energy independence (see Dahl, 1979; Dahl and Sterner, 1991; Goel, 1994; Hausman and Newey, 1995; Deaton and Paxson, 1998; Manzan and Zerom, 2010; Chang and Serletis, 2014). As a consequence, a large literature has evolved that seeks to estimate the price elasticity of gasoline demand in order to inform policymakers. Much of this uses aggregate data, although some uses household-level data (see for example, Dahl and Sterner, 1991; Graham and Glaister, 2002; Brons et al., 2008; Espey, 1998; Havranek et al., 2012, for surveys).

Economic theory provides no specific guidance on the appropriate functional form for the gasoline demand function. Most analyses of gasoline demand employ some form of parametric specification, such as linear, log-linear or translog and assume the distribution of the error term to be normal with zero mean and fixed variance. The results from such models are easy to interpret because parametric modelling gives constant price and income elasticities. For 'representative consumers', parametric modelling of gasoline demand mostly suggests that gasoline demand is price inelastic. Meta-analyses suggest that the long-run price elasticity is in the range −0.31 (Havranek et al., 2012) to −0.84 (Espey, 1998). However, the problem is that it will often be inappropriate to assume that demand elasticities are constant across groups of consumers with different incomes and facing different prices. As Hausman and Newey (2016, p.1125) have recently emphasised: "Demand functions could vary across individuals in general ways". Parametric specifications are not well suited to study potential variation in the elasticity of demand, although understanding how people's response to price varies according to the price level, and across the income distribution, is of importance to policymakers (Blundell et al., 2012).

In such circumstances, using a partial linear model is an attractive solution to model household gasoline demand if we believe that the relationship between some independent variables and gasoline is linear, while the relationship between other independent variables and

* Corresponding author at: 20 Chancellors Walk, Wellington Road, Clayton, VIC 3800, Australia
E-mail address: russell.smyth@monash.edu (R. Smyth).

http://dx.doi.org/10.1016/j.eneco.2017.08.029
0140-9883/

gasoline is non-linear. In response, beginning with Hausman and Newey (1995), there has been growing interest in modelling gasoline demand through applying semi-parametric models to household survey data (see for example, Schmalensee and Stoker, 1999; Yatchew and No, 2001; Coppejans, 2003; Blundell et al., 2012; Manzan and Zerom, 2010; Wadud et al., 2010b; Liu, 2014).

Yet, while estimating partial linear models with household data seems intuitively attractive, as Blundell et al. (2012) have noted, non-parametric and semi-parametric regression can yield implausible and erratic estimates. A feature of much of the earlier literature to estimate gasoline demand with semi-parametric regression is that it found erratic price effects. For example, using household survey data for the United States (US), Hausman and Newey (1995) and Schmalensee and Stoker (1999) found that as price increases, the fitted demand curve exhibits an upward slope when the price is within a middle range and a downward slope elsewhere. Moreover, Schmalensee and Stoker (1999) found that the average derivative of the estimated demand with respect to price is positive. These findings contradict conventional consumer theory, which implies that gasoline demand should decrease as gasoline price increases. There continues to be substantial differences in the more recent semi-parametric regression literature regarding estimates of the price elasticity of gasoline in the US (see for example, Blundell et al., 2012; Manzan and Zerom, 2010; Hausman and Newey, 2016).

Blundell et al. (2012) suggested that erratic and/or implausible price effects with semi-parametric regression reflect imprecision of the unconstrained semi-nonparametric estimates. Their solution to this problem was to impose the Slutsky condition on the kernel estimator of the nonlinear component, through which gasoline price and household income effect gasoline demand. Fitting a constrained partially linear model to US household survey data for 2001, they found an inverse relationship between gasoline price and demand, consistent with consumer theory. A practical limitation of their solution, though, is that estimation of this constrained model is complicated because the vector of weights imposed by the Slutsky condition has a dimension equal to the sample size. Hence, the total number of unknown identities including coefficients, bandwidths and weights is larger than the sample size.

Blundell et al. (2012) suggested estimating this constrained partially linear model through three steps, where the weights are optimized by minimizing an objective function under the Slutsky constraint, bandwidths are selected through least squares cross-validation, and linear coefficients are estimated using the profile least squares method discussed by Speckman (1988) and Robinson (1988). To the best of our knowledge, no method is available to simultaneously estimate the linear coefficients, bandwidths and weights imposed by the Slutsky condition. We fill this gap in the literature through proposing such a method and present a sampling approach to the estimation of a partially linear model subject to the Slutsky condition.

We propose a methodological improvement to the manner in which partial linear models are used to estimate household gasoline demand and illustrate our contribution through estimating household gasoline demand using US survey data. Our starting point is that price elasticity, which is computed through the first derivative of the fitted demand with respect to price, depends on the smoothness of the kernel estimator of the nonlinear function, through which price and income effect demand. The performance of this kernel estimator is mainly determined by bandwidths (see, for example, Härdle, 1990). Schmalensee and Stoker's (1999) finding of an implausible price effect is likely to be due to the lack of Slutsky constraint on the kernel estimator and a subjective choice of bandwidths. Blundell et al. (2012) introduced such a constraint, but a limitation of their three-step estimation method is that bandwidths are selected based on pre-estimated parameters. We improve on the approach adopted by Blundell et al. (2012) in that we derive a posterior and develop a posterior simulation algorithm to simultaneously estimate the linear coefficients, bandwidths in the kernel estimator and the weights imposed by the Slutsky condition.

Specifically, we contribute to the literature on estimating partial linear models of household gasoline demand with survey data in the following important ways:

- We derive a posterior for the partially linear model with its kernel estimator of the nonlinear component being subject to the Slutsky condition. A Markov chain Monte Carlo (MCMC) sampling algorithm is developed to simultaneously sample linear coefficients and bandwidths, as well as the weights imposed by the Slutsky condition.
- With the proposed sampling algorithm, we estimate the constrained partially linear model of household gasoline demand employing the same 1991 US survey data as originally employed by Schmalensee and Stoker (1999) and find, contrary to their result, that the price elasticity for household gasoline demand is negative.
- With the proposed sampling algorithm, we also estimate the constrained partially linear model of household gasoline demand employing the 2001 US survey data used by Blundell et al. (2012). We find not only a negative price effect, which is largely consistent with Blundell et al.'s (2012) finding, but, contrary to their findings, that middle-and high-income earners are less sensitive to gasoline price increases, while low-income households are more sensitive to gasoline price changes.
- We also apply our algorithm to estimate household demand for gasoline in Canada over the period 2006 to 2009. We find a negative price effect, that gasoline is, on average, price inelastic and that as gasoline prices increase, the price elasticity decreases.

The rest of this paper is organized as follows. In the next section, we present the posterior for a partially linear model with its kernel estimator of the nonlinear component being subject to the Slutsky constraint. We also present a sampling algorithm. In Section 3, we apply this sampling procedure to estimate constrained partially linear models for household gasoline demand and examine the relevant price effects. To do so, we first use data from the 1991 and 2001 US household transportation surveys, which facilitate comparison between our results and those in Schmalensee and Stoker (1999) and Blundell et al. (2012). We then apply our approach to more recent data; namely, Canadian household data for the period 2006–2009. Section 4 concludes the paper and draws out the broader advantages of our approach for modelling gasoline demand.

2. A partially linear model subject to the Slutsky condition

2.1. Model and posterior

Suppose that a response variable y can be explained by $\boldsymbol{x}$ and $\boldsymbol{z}$ in the form of

$$y = \boldsymbol{x}'\beta + g(\boldsymbol{z}) + \varepsilon, \tag{1}$$

where β is a $d \times 1$ vector of unknown parameters, and $g(\cdot)$ is an unknown nonlinear function with its argument being either a scalar or a vector. Let $(y_i, \boldsymbol{x}_i', \boldsymbol{z}_i)'$, for $i = 1,2,\cdots,n$, denote observations of $(\boldsymbol{y},\boldsymbol{x}',\boldsymbol{z})'$. It is also assumed that the error terms ε_i, for $i = 1,2,\cdots,n$, are independent and identically distributed (iid) with an unknown density $f(\varepsilon)$.

Let $\boldsymbol{z} = (z_1,z_2)'$ and $\boldsymbol{z}_i = (z_{1i},z_{2i})'$, where z_1 and z_2 are the price and income variables in the household data of gasoline demand. In the partially linear model (1), $g(\boldsymbol{z})$ is estimated by a weighted Nadaraya-Watson (NW) estimator. It is a function of price and household income and is expressed as (Hall and Huang, 2001; Blundell et al., 2012)

$$\hat{g}(\boldsymbol{z}_i|\boldsymbol{h}) = \frac{\sum_{j=1}^{n} w_j K_{\boldsymbol{h}}(\boldsymbol{z}_i - \boldsymbol{z}_j)\left(y_j - \boldsymbol{x}_j'\beta\right)}{n^{-1}\sum_{j=1}^{n} K_{\boldsymbol{h}}(\boldsymbol{z}_i - \boldsymbol{z}_j)}, \tag{2}$$