Sveriges lantbruksuniversitet
Swedish University of Agricultural Sciences

Department of Economics

Elasticity of demand for gasoline in Sweden

Emma Dahlkvist

Master's thesis · 30 hec · Advanced level
European Erasmus Mundus Master Program: Agricultural Food and Environmental Policy Analysis (AFEPA)
Degree thesis No 1047 · ISSN 1401-4084
Uppsala 2016

Elasticisty of demand for gasoline in Sweden

Emma Dahlkvist

Supervisor: Rob Hart, Swedish University of Agricultural Sciences, Department of Economics

Assistant supervisor: Thomas Heckelei, Rheinische Friedrich-Wilhelms-Universität Bonn, Institute for Food and Resource Economics

Examiner: Sebastian Hess, Swedish University of Agricultural Sciences, Department of Economics

Credits: 30 hec
Level: A2E
Course title: Independent project/degree in Economics
Course code: EX0537
Programme/Education: European Erasmus Mundus Master Program: Agricultural Food and Environmental Policy Analysis (AFEPA)
Faculty: Faculty of Natural Resources and Agricultural Sciences

Place of publication: Uppsala
Year of publication: 2016
Name of Series: Degree project/SLU, Department of Economics
No: 1047
ISSN 1401-4084
Online publication: http://stud.epsilon.slu.se

Keywords: Distributional effects, Fuel tax, Gasoline price elasticity, Income elasticity

Sveriges lantbruksuniversitet
Swedish University of Agricultural Sciences

Department of Economics

Declaration

I hereby affirm that I have prepared the present paper self-dependently, and without the use of any other tools than the ones indicated. All parts of the text, which were taken over verbatim or analogously from published or unpublished works have been identified accordingly. The thesis has not yet been submitted in the same or a similar form, within the context of another examination.

__________________________ __________________________

Date, place of submission Emma Dahlkvist

Abstract

Policy measures in the transport sector have been widely debated during the recent decades, specifically in terms of increasing carbon emissions from passenger transport. Fuel taxes is receiving most receptive consideration by governments, although households tend to respond little to these measures, especially in rural regions. The aim of this paper, which focuses at the gasoline consumption among households, is to develop a model to estimate price and income elasticities in rural and urban regions in Sweden. While obtaining overall price elasticity to be within the range of previous studies, I find that there is a significant variation in price elasticities across regions. As a consequence, these differences might result in unwanted distributive effects.

CONTENTS

1. Introduction

In January of 2016 the Swedish government initiated an increase of the fuel tax with the aim to reduce carbon emissions and energy dependence in the transport sector. Whether this aim is reachable will depend on how households response to this increased tax. This response can be measured with elasticities, which indicate how much more or less a household demand as the price of the gasoline change. Gasoline elasticities have been found to vary within countries depending on household income as well as location. These differences might cause unwanted distributive effects where some households are more punished that others.

In urban areas it is reasonable to suppose that the long run price elasticity of demand for fuel is rather high, since there are many alternatives to car transport. On the other hand, in rural areas with fewer options for public transport and where work and home may be far apart, the fuel demand may be less price sensitive. In the worst case, the high fuel tax may be seen as a punitive charge applied to those living in the countryside, who are unable to adapt their behaviour to avoid its effects.

For reasons both of equity and policy effectiveness, it is relevant to recognise how an increase in fuel tax will affect different types of households. It is also useful to know from which groups the demand response will be most distinct. In this setting, whether rural households in Sweden are likely to be more affected and more responsive to fuel taxes compared to urban households, is contingent on an analysis of price and income elasticities of gasoline demand.

The research question of this study is whether households have different price and income elasticities of gasoline demand in Sweden, depending on in which region they are located. The aim is thus to test the potential existence of a difference in the elasticity of demand of gasoline fuel by Swedish households, given their place of residence. To do so an econometric model is assessed with annual disaggregate data from the Swedish Household Expenditure Survey of the period between 2003-2009 and 2012. The price and income elasticities of the demand of gasoline is compared for households located in the major towns in Sweden with households living in rural and other regions. Conclusions are drawn on the effectiveness and distributional impacts of the tax increase.

Many studies have used disaggregate data to find the regional effect of a price shock in car fuel. This have been done by including location of the household as a variable in the model (e.g., Poterba, 1991; Kayser, 2000; Bento et al., 2009; West and Williams, 2004). Fewer studies intend to more specifically measure the actual price elasticity of each region and its implications.

Bureau (2011) studies the distributive effects of an increased fuel tax in France with panel data from 2003-2006, and distinguishes between urban and peri-urban/rural households that own a car. The estimated elasticities in the former region are between -0.30 for low income households to -0.19 for high income households and for the peri-urban/rural -0.25 to -0.17. In the UK Blow and Crawford (1997) and Santos and Catchesides (2005) use the same data source in different time periods. Blow and Crawford (1997) examine an earlier time period of 1988-1993 and find the greatest price elasticity to be -0.54 of poor households in urban areas and the smallest to be rich households in rural areas of -0.2. The income elasticity shows little variation with respect to population density and is on average 0.26. Santos and Catchesides (2005) study the time period 1999-2000, and find the greatest price sensitivity to be among the poor households living in urban areas of -0.93. The smallest price response is for households living in rural areas, where middle-income households has the smallest price response with an elasticity of -0.75. Income elasticity is found to be 0.63 for households in the urban areas and 0.6 for households in the rural areas.

These studies show that households in the more densely populated regions, in general, tend to be more price sensitive to gasoline than households in the rural areas, and that urban households might have greater income response than the rural ones.

In Sweden however the pattern seems to be different, based on the very limited number of studies. Brännlund and Nordström (2004) study the period 1985, 1988 and 1992 and predict price elasticities to be consistent around -0.99, regardless of the location of the household. Although Brännlund and Nordström use a more sophisticated modelling approach in relation to the studies listed above, these results are quite contradictory even to studies using the same modelling approach. Nicol (2003) applies a similar methodology as Brännlund and Nordström (2004) and find Canadian households to vary in price elasticity estimates between -0.47 to -0.83 depending on their location. Nicol indicate that while elasticities might change due to certain characteristics of a household, this might not be the same in another country. Thus, there is a need for them to be country specific.

The small number of elasticity studies of gasoline in the different regions in Sweden is the main motivation of this study. While previous studies of increased fuel taxes tend to focus on the income equity issue, this study will mainly focus on the effect on the different regions. In the methodology part, an alternative approach applicable in a single demand function is used to estimate elasticities, by including interaction terms between several regions and gasoline price as well as income. Interaction terms of this kind have previously been used by Wadud et al. (2010), although not in the Swedish case.

In this thesis gasoline demand is modelled as a function of price, total expenditure and other relevant household characteristics. The goal of the model is to assess the influence of prices on fuel consumption and the distributive effects on households in different regions. The model is calibrated on five different regions in Sweden, with rural and non-rural characteristics. Price and income elasticity are estimated, by using the static single equation model. In order to derive the elasticity estimates directly from the estimated coefficients the static log-log linear model is used with the ordinary least square regression (OLS) in line with previous gasoline demand studies (e.g Hughes et al., 2008). In Hughes's model specification of the log-log linear OLS, no household characteristics are included. In my model on the other hand various aspects of the model is considered based on the data at hand and findings in the literature, specifically; time, regional-dimension, demographic characteristics. The time trend is included as fixed time effects using dummy variables representing each year. In addition also the regional effect will be studied using dummy variables. Further since the interest is to see how elasticities varies with the regions interaction terms region-price and region-expenditure is integrated, following Wadud et al. (2010) approach as well as other household demand studies (e.g. Archibald and Gillingham, 1980).

Households characteristics included in the model are based on empirical findings in the literature. Kayser (2000) suggests that demographic characteristics such as income, age and occupation status affect gasoline demand, while education level have no significant impact on demand.

With considerations to the data, where about 20 percent of households report zero consumption for gasoline, the OLS regression is not appropriate to use for the whole sample since it does not take special account of zero gasoline expenditure and consequently yield inconsistent estimates of the parameters. Therefore, subsample OLS including only households with positive expenditure on gasoline will be tested. Wales and Woodland (1983) indicate however that this will reduce the sample size and the standard estimators may be biased and inconsistent. In order to correct for these potential sample selections issues Heckman's two stage approach is used.

The results indicate a significant variation of the price elasticities in the different regions in Sweden. Urban households tend to be more responsive to gasoline prices, particularly in Stockholm, Gothenburg and Malmö. While rural households in the northern regions seem to be almost price insensitive. With regards to the income elasticity, the regional location of the household does not seem to affect the level. These are measures in the intermediate run and effects in the long run will partly depend on how the revenue of the tax increase is recycled.

Past decades of urbanization and depopulation of rural areas in Sweden has lead to extensive engagement into rural development strategies by the Swedish government. Substantial investments are directed towards this policy area in order to maintain the livelihood of remote regions. Within this context, transport is a main component. Thus if strict transport policies hurt households living in rural areas, these measures might contradict with investments supporting these regions. The region specific elasticity estimates of gasoline demand resulting from this study might therefore be desirable for policy-makers in order to design overall policy measures more efficiently.

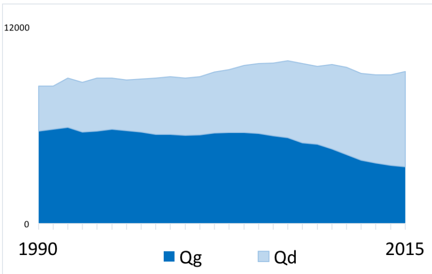
Next section will review and discuss the transport sector in Sweden. Following the theoretical framework of this study is presented with focus on the neoclassical framework. Section 4 discuss the data applied in this thesis, followed by a review of my model and alternative approaches that could have been chosen. Section 6 specifies the final model, followed by the results and discussion. Lastly section 8 concludes the paper.

2. The transport sector

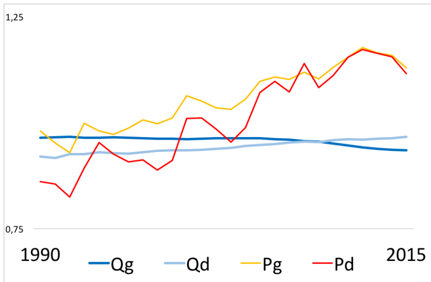
In this section, the general transport policy in Sweden is discussed and more specifically the recent increase in the fuel tax and the distributive effects it might result in.

Our modern society is highly dependent on a functioning transport system. While mobility is a crucial feature of our lifestyle and passenger transport is relevant for economic development together with individual and social welfare, it also causes a number of serious environmental and health problems. The level and increasing trend of automobile fuel consumption as well as the dependence of fossil fuel as energy source is a growing concern among countries, particularly in terms of increasing carbon emissions and security of energy supply. Although policy measures that discourage fossil fuel use in transport are in place and alternative fuels are available, passenger transport is still the fastest growing emitter of carbon emissions worldwide of all energy source sectors (IPCC, 2015).

2.1. **Fuel consumption and gasoline price in Sweden.** Overall fuel consumption specifically in Sweden, has on the other hand been relatively constant since the beginning of the 90s. Although, there's been a slight decrease of gasoline consumption in favour for diesel consumption as illustrated in figure 1(A), where Q_g is the quantity consumed of gasoline and Q_d is the quantity consumed of diesel. The real price of fuel has on the other hand been fluctuating in a upward-sloping motion, see figure 1(B), where P_g is the real price of gasoline and P_d is the real price of diesel.



(A) Fuel consumption



(B) Consumption and prices

FIGURE 1. Consumption and prices of automobile fuel in Sweden
Source: Data from Swedish Petroleum and Biofuels Institute (2016)

The pattern in figure 1(B) is similar in most EU-countries and is explained by that drivers seem to respond to higher fuel prices by investing in fuel economy and less so by reducing driving (Van Dender, 2009).

2.2. **Transport policy.** Transport policy decisions have far-reaching and long-term consequences for the structure of societal transportation. In order to enable a more environmental and socially sustainable use of transport and to provide guidance for policy makers in designing efficient policies, prediction of how car fuel demand react to changes in price, income and other explanatory variables is preeminent.

Transport policy should in theory correct the external effects from traffic as direct as possible. Thus either a tax or a quantity restriction will generally be favoured to alternative direct controls such as limiting traffic levels. That is simply because taxes and quantity restriction are more efficient in encouraging the use of less-polluting technology and achieve pollution reduction in a more economic efficient way. This reasoning is supported by findings in the literature (e.g De Jong and Gunn, 2001) where price elasticity of gasoline is generally higher in absolute terms, than the price elasticity of automobile travel demand.

When governments initiate passenger transport policies, they are usually driven by four types of concern: the tax base, climate change, the security of oil supply, profits and employment in the domestic car industry (Proost and Van Dender, 2011). Since Sweden is part of the European Union, the outline of the transport policies are partly affected by regulations settled in the Union. In 2009, the European Commission passed a legislation requiring car producers to reduce the average per-kilometre carbon emissions of newly manufactured automobiles to 130g/km by 2015 (Council of the European Union, 2009). In addition to this fuel efficient standards, minimum rates of fuel taxation are set for all member states (Council of the European Union, 2003).

There are divided opinions in the literature regarding the effectiveness of combining fuel efficiency standards with fuel taxes. While Van Dender (2009) suggest that fuel efficiency standards and fuel taxes are complements rather than substitutes, more recent studies by Frondel and Vance (2013) and Liu (2015) demonstrate that tightening efficiency standards will partially offset the effectiveness of taxes on reducing fuel consumption, since the improvements of vehicle fuel efficiency leads to lower price elasticity and weakens the consumer response to gasoline price changes. Although, according to Sperling and Nichols (2012) a collection of different policy instruments are needed in order to have large emission reductions from the transport sector.

Despite policy efforts within the EU, technical progress and potential for cost-effective energy efficiency improvements, the transport system has not fundamentally changed and is not sustainable (Akerman et al., 2000). Recent political activities suggests that transport charges and taxes must be reconstructed in the direction of application of the "polluter-pays" and "user-pays" principle (European Commission et al., 2011).

2.3. **Fuel tax.** Among the policies used in the transport sector, fuel taxes seem to be receiving most receptive consideration by governments as they raise revenue but also through their effect on price that can affect demand and consumption (Nicol, 2003). That is also true in Sweden, where the government support the economists' idea of fuel taxes being of central importance to meet the climate and energy objectives in a cost effective manner (The Swedish Government, 2015).

The use of the price mechanism is particularly important due to its contribution to revenue of the public fund, but also to assist markets to operate more efficiently by ensuring that the the external cost of using a vehicle are met by the users.

Figure 2 show the four components of the gasoline price met by households in Sweden: the gross margin, production cost, tax and Value Added Tax.

Taxation of automobile fuel in Sweden consist of excise duties (tax) and value added tax. Excise duties are divided into energy tax and carbon tax and are derived on the basis of the national Energy Tax Act(1994:1776). The carbon tax was introduced to reduce carbon emissions from fossil fuels,

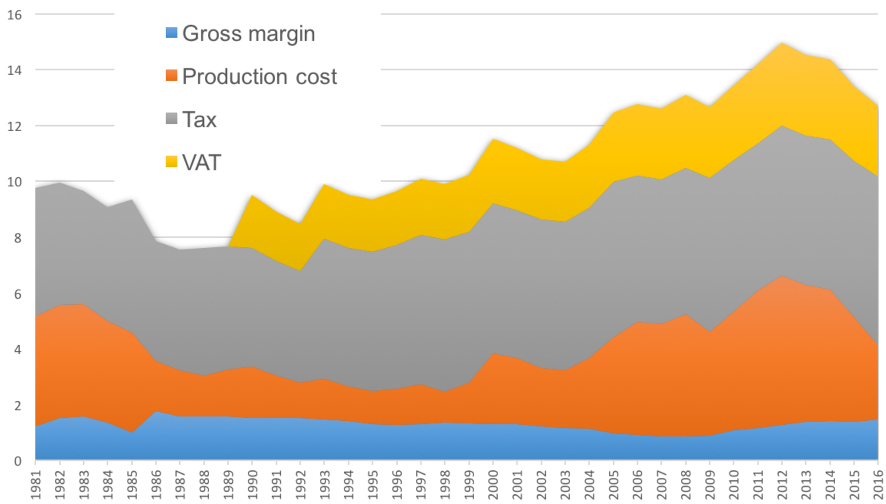


FIGURE 2. Components of the gasoline price in Sweden
Source: Data from Swedish Petroleum and Biofuels Institute (2016)

whereas the energy tax first was motivated by fiscal reasons but later have become the most important measure to reach the Swedish energy intensity goal.

Nevertheless, since the literature generally suggest a low price elasticity and high income elasticity of automobile fuel demand, the effectiveness of fuel taxes might therefore be limited in reducing fuel consumption. Espey (1998) among others imply that in order for fuel taxes to be effective, fuel prices rise must faster than income and Van Dender (2009) indicate that fuel consumption has become price elastic over time, due to rising incomes and falling real fuel prices.

Empirically, however, it seems like fuel taxes had a great influence on carbon emissions generated from the transport sector. Stern (2007) studies the effects if all OECD countries have had as low taxes as the US, and finds that total fuel consumption within the OECD counties could have been 30 percent higher. These results indicate that even though increasing fuel taxes might be insufficient in influencing demand, they clearly have provided an incentive for the development of alternative and more fuel efficient technologies.

2.4. **Motivations of increasing the fuel tax in Sweden.** In the budget proposal of 2015 the Swedish Government established that the environmental impact from the transport sector must be reduced at a faster pace. That is to be in line with the national goal of efficient energy use as well as the fossil fuel independent vehicle fleet by 2030. This should be done by increasing the controlling effect of taxes on energy and fossil fuels on automobile fuel. Specifically, the Government initiates, by the first of January 2016, increased energy taxes on fossil by 0.48 SEK on gasoline and 0.53 SEK on diesel per litre of fuel. The tax increase is meant to incentives automobile owners to reduce their car use which is estimated to reduce carbon emissions slightly. In addition, the Government believe that higher energy taxes may also result in automobile owners to choose energy-efficient vehicles, which could lead to further carbon reductions. In order for this aim to be fulfilled it's important to examine the consumer response to the price change. The theoretical framework for doing so will be explained further in the next chapter.

2.5. **The distributional effects of environmental taxes.** With an increased taxation on fuel, households will be affected differently depending on their income and their reliance to travel by car. The change in fuel taxes will therefore generate distributional effects depending on household characteristics. There are many studies focusing on the income equity issue of gasoline taxes. Bureau (2011), conclude that fuel tax is regressive in France, which also is true in the UK according to Santos and Catchesides (2005), who further concludes that middle-income households suffer the most.

Although, with an environmental tax, or a tax correcting for external effects, Eliasson et al. (2016) states that it might be unclear whether these distributional effects, when it comes to income equity, are relevant to consider. If the price of car travel is lower than the full social cost, those who drive the most should also bear the burden of the cost of external effects. The economic welfare effect on households due to a tax increase, is theoretically the same for all regardless of income or wealth, if they face the same price. The strive for income equity among households is rather handled by taxation and welfare systems.

However, when discussing distributional effects, there are other aspects to consider such as the regional dimension of the household. Within this area there a much fewer studies in the literature. Bureau (2011) finds that welfare losses are significantly higher in rural areas when distinguishing between urban and rural residents in France. In Sweden Eliasson et al. (2016) finds the same pattern.

The regional dimension of the distributional impacts of an increased fuel tax may be of high interest in Sweden. That is since the government invests a substantial amount into rural development strategies due to decades of increased urbanisation causing depopulation of some rural areas. Within this policy area transport and infrastructure are two significant components. Thus if the transport policy cause negative distributional effects in rural areas, these effects will counteract with the large investments in rural development.

Before getting into the theoretical framework for doing so, figure 3 illustrates how the budget share put on automobile fuel differ between regions. Region 0 and 1 are the most populated in Sweden and region 3-4 the most rural. This clearly illustrate that households in rural areas allocate a larger amount of their total expenditure towards gasoline consumption than other households.

A complete analysis of the net welfare effect needs to consider the use of revenues from the tax are recycled. This could be read further in an extensive study by Bento et al. (2009) within the US context and Eliasson et al. (2016) who examine the Swedish prospective.

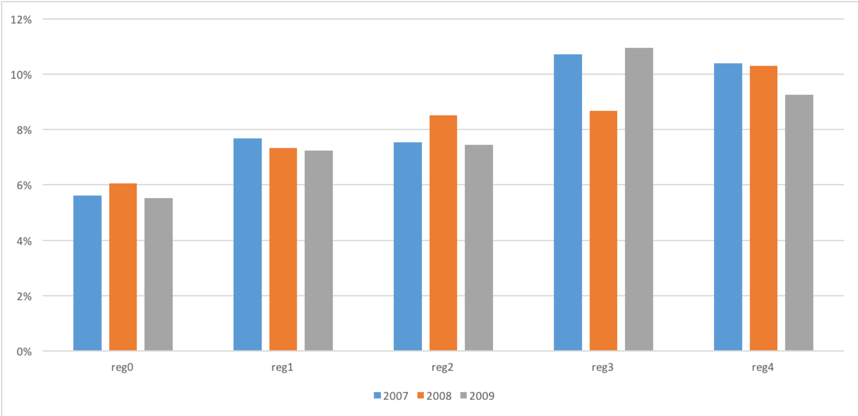


FIGURE 3. Share of gasoline expenditure in each region: reg0= Stockholm, Gothenburg, Malmö, reg1=Urban areas, reg3= Major towns, reg4=Rural areas
Source: Data from Statistics of Sweden (2016)

3. Theoretical framework

Demand of gasoline can be affected by many different factors such as price shocks, policy measures, but also price changes of substitute or complement goods. Consumer demand theory is fundamentally concerned with how a rational consumer make consumption decisions. Apart from the general problem of choice theory, this problem is worth studying do to its particular structure that allows to derive economically meaningful results.The structure arises due to the fact that a consumer's choice set is assumed to be defined by prices and that consumer's income. In this section the consumer behavioural process will be explained on the basis of neoclassical consumer demand theory.

3.1. **Neoclassical consumer demand theory.** The neoclassical approach is set to explain and predict the behaviour of individual actors. In the framework of neoclassical economics consumers attempt to maximize their gain of obtaining goods by expanding the number of goods in their purchasing basket until what is gained from and extra unit of good is balanced by what they must give up to get it. Generally, there are three central assumptions of neoclassical economics; (i) individuals make rational choices in consumption based on their preferences, (ii) individuals are maximizing utility and (iii) individuals have full and all relevant information to make a decision. These assumption relates to consumer behaviour and will be explained in following.

The consumer is said to have preferences between the total set of bundles of goods. Considering two bundles in the total set x and y, if $x \succeq y$ and $y \succeq x$, the consumer is said to be indifferent or have no preferences between y and x. If only $y \succ x$, then the consumer is said to strictly prefer y over x. The consumer is assumed to base her actions according to her preferences. However, if $x \succeq y$ the consumer believe that bundle x is at least as good as bundle y. This belief is based on the individual consumer's feelings that

determine her choice. The belief and action processes are independent of each other, although put together they form the behaviour of the consumer (Deaton and Muellbauer, 1980).
Mas-Colell et al. (1995) defines it (emphasis added),

> The theory [of consumer behaviour] is developed by first imposing rationality axioms on the decision maker's preferences and then analysing the consequences of these preferences for her choice behaviour (i.e. on decisions made)

The rationality axioms Mas-Colell et al. (1995) is referring to must be met by the consumer preference relations in order for her to be assumed to act rational. First the preferences must be complete, indicating that for any two consumption bundles x and y, $x \succeq y$, $y \succeq x$, or both. The second is transitivity which imply that for any three bundles x, y and q, if $x \succeq y$ and $y \succeq q$, then $x \succeq q$. The preferences must also be reflexive, where any bundle is at least as good as itself, and continuous, indicating no big jumps in consumer preferences. Further that preferences are monotonicity, meaning that more is always preferred to less and lastly that they are convex, where any combination of two equally preferable bundles are more desirable than these bundles by themselves.

Given the completion of these axioms, the consumers indifference sets between bundles can be illustrated with indifference curves, in which each curve is the set of all bundles generating the same utility for the consumer. The slope of the line tangent to a bundle on the indifference curve is called the marginal rate of substitution, MRS, which implies the rate at which the consumer is willing to exchange x for y . This exchange also depends on the prices of the goods as well as the constrained budget of the consumer. The most preferred bundle of goods the consumer can afford is found by choosing the bundle on the budget line where MRS equals the price ratio.

In order for the consumer to be utility maximizing, stated by assumption (ii), the consumer will choose the bundle of goods that maximizes her utility the most constrained by her income. Consider an individual with an utility function $u(q, z)$ where q is the vector of a number of goods on which the consumer must make consumption decisions and z represent the individual's characteristics. The total amount of income to spend is y, and the budget constraint is $y = p'x$ where p' is a vector of prices of the goods. The maximizing utility problem facing the consumer is,

$$\begin{aligned} \textbf{max} \quad & u(q, z) \\ \textbf{s.t} \quad & y = p'x \end{aligned}$$

The solution to the maximization problem is a set of demand equations for the goods i in vector q,

$$q_i = f(p, y, z)$$

From this demand equation we find that besides consumer choice strategies, consumption patterns are also affected by price changes. A consumer with a fixed budget in the short term has three possible responses due to a price change: (i) The consumer buy another good as a substitute; (ii) the consumer purchase less of the good and no substitute goods; (iii) the consumer continue to buy the same amount of the good and decrease the expenditure on other goods in her consumption basket.

3.2. **Substitutes and complement goods.** For many goods demand is dependent on price and consumption of other goods. If the price increase for a certain good leads to an increased demand of another good these goods are substitutes. Considering gasoline and diesel, these are substitutes if, $(\partial Q_d/\partial P_g) > 0$ where a price increase of gasoline (the denominator) would lead to an increased quantity demand of diesel (the nominator). If the price of one good on the other hand leads to a decrease in quantity demanded of another the goods they are complements. Automobiles and gasoline are an example of the latter types of goods, which require the following setting $(\partial Q_v/\partial P_g) < 0$. This implies that if the price of gasoline increases, the quantity demand of vehicles driven on gasoline will decrease. Although we can establish that gasoline and automobiles are compliments, what consumers really demand is transportation. Consumers demand transportation where gasoline serves as an input and where gasoline consumption will depend on the efficiency and price of vehicles.

3.3. **Consumer production theory.** Some researchers have implicitly argued that commodities purchased by consumers are inputs into the the production of goods within the household (e.g Lancaster, 1966; Becker, 1965). This leads into Lancaster's (1966) assumption that consumption is an activity in which goods are inputs and in which output is a collection of characteristics. The framework of this view on consumer theory is summarized in three assumptions: (i) the good will not give utility to the consumers, rather it generates characteristics which give rise to utility, (ii) generally a good will possess more than one characteristic and most characteristics will be shared by more than one good, and (iii) many goods in combination can generate characteristics different from those from an individual good.

3.4. **Elasticity.** While the price change effect of a certain good depends on its price sensitivity or elasticity, the elasticity indicate how consumers react to a price increase. Price elasticities indicate how willingly consumers are to purchase substitutes of a good that has gone up in price and how much consumers value a particular good. These can be used in this way because the underlying theory of consumer response to a price change, as explained above. With reference to the demand equation in (1), the price and income elasticity we can be specified as,

$$\eta_p = \frac{\partial q_i}{\partial p}\frac{p}{q_i}$$
$$\eta_y = \frac{\partial q_i}{\partial y}\frac{y}{q_i}$$

According to price elasticities, $\eta_p < -1$ indicate an elastic response and $\eta_p > -1$ an inelastic. When a good is price elastic, the consumption for that good will decrease in favour for substitutable goods as the price increase. If the good is price inelastic, the consumption will not change at all or very little, due to the lack of substitutable goods.

With respect to the income elasticity, a normal good can be classified in terms of its importance to the consumer where, $0 < \eta_y < 1$ are necessity goods and $\eta_y > 1$ luxury.

The degree of elasticity depends on a number of factors, primarily the availability of alternative (substitute) goods but also the consumer's individual tastes and other infrastructural factors particularly in the context of fuel. Given a consumer owns a gasoline car, gasoline will have limited degree of substitutability especially in the short term. Over time, however, the consumer has the option to purchase a vehicle that is more fuel efficient and/or use alternative fuels and still enjoy the same level of utility using less fuel (if the infrastructure allows). Typically, switching vehicle requires replacing an expensive good and for many consumers it is considered a long-run adjustment to high fuel prices. Therefore, empirical estimates on price elasticity of gasoline tend to be elastic in the long term and inelastic in the short. There are several methods to estimate the elasticity of fuel demand that will be further explained in section 5.

4. Data and method selection

In this section the underlying data of this study is discussed.

4.1. **Data.** The annual gasoline price used are the yearly average for petrol provided by the Swedish Petroleum and Biofules Institute. In order to estimate consistent elasticities and to ensure identification of actual demand responses, the prices are adjusted for inflation. To do so the annual Consumer Price Index (CPI) series, published by Statistics of Sweden on a monthly basis, is used. The reference is the annual average of 2015.

The household data is collected from Statistics of Sweden between 2003-2009 and 2012. The data is based on the Swedish Household budget survey (HBS), including mainly household expenditure on consumption goods but also household and individual characteristics. The data is repeated cross sectional with different households reporting each year but answering the same type of questions. Due to the repeated cross sectional structure, the yearly samples are assumed to be independent. In order to efficiently combine the two sets of household characteristics and individual characteristics, the individual characteristics are based on the head of the household only. The total expenditure level will be used as a proxy for lifetime income, which previous studies suggest to be a better predictor of consumption than annual income (see Friedman, 1957; Poterba, 1991; West and Williams, 2004). The annual quantity of gasoline consumption will be derived by dividing the annual household expenditure share spent on gasoline with the annual average price of gasoline.

In order to account for household features that may affect consumer behaviour with respect to gasoline demand, we need to include these household characteristics in the model estimation. A summary statistics for the variables extracted from the HBS survey and included in my model to represent household characteristics are presented in 1. All of them have been shown to affect gasoline consumption (e.g Puller and Greening, 1999; West and Williams, 2004; Kayser, 2000).

Table 1. Summary statistics

Statistic	N	Mean	St. Dev.	Min	Max
Total annual expenditure	17 911	345 428	198 751	7520	3 082 046
Total quantity of gasoline demand	17 911	1255	1640	0.000	111 660
Real gasoline price	17 911	12.641	1.272	10.692	14.982
Expenditure on gasoline	17 911	15 672	20 426	0.000	1 459 579
Child	17 911	0.955	1.160	0	9
Adult	17 911	1.842	0.595	1	7
Male	17 911	0.318	0.466	0	1
Age (Head of household)	17 911	48.254	14.538	13	92

Source: Data from Statistics of Sweden (2016)

The different regions studied are specified in Figure 4. This figure also include the composition of different car fuel consumed by the households, including gasoline, diesel, other types such as electricity or natural gas and non which indicate that the household consume zero car fuel.

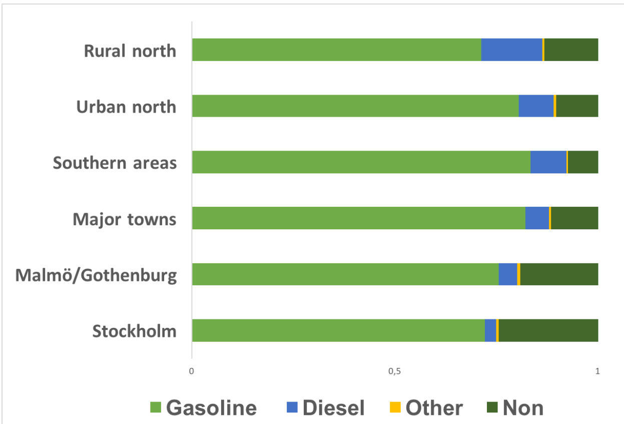


Figure 4. Type of car fuel consumed in each region
Source: Data from Statistics of Sweden (2016)

5. Methodology

With the data presented, this section will begin to explain the model I have chosen to use to answer my research question and then what alternative approach could have been chosen instead. Recall that I estimate price and income elasticities of the different regions using a static single equation model with OLS.

5.1. **My model.** To fulfil the aim of the study, two key structural parameters need to be estimated. These parameters are price and income elasticities of gasoline demand of different rural and non-rural regions in Sweden.

The single equation approach is used in this thesis since it is simple and very flexible in its specification. Apart from different data types that can be adopted within this approach, also static or dynamic framework can be used as well as different functional forms.

The econometric approach to the single equation model rely on a regression function that is set to determine to what extent one or a set of independent variables, denoted x, explains the variation of a dependent variables, denoted y. The starting point is usually the linear regression model, which must be linear in its parameters,

$$
\begin{aligned}
y &= \beta' x + u \\
E(u|x) &= 0
\end{aligned}
$$

This regression function can be estimated by ordinary least squares (OLS), by defining one of the elements of x to be a constant (an intercept term). In this setting, the slope β are the coefficients or parameters of the regression line, where the slope is the change in y associated with a unit change in x. The term u is a residual, disturbance or error term illustrating omitted determinants of y, including measurement error. It contains all of the other factors besides x that determine the value of y (Stock and Watson, 2012).

The OLS method is simple, although very flexible and therefore a good choice in this thesis. However, in order to derive and use the OLS estimator there are five assumptions to consider. First, the dependent variable y must be calculated as a linear function of the specific set of the independent variables, x, and the error term, u. Thus, the equation must be linear in parameters $\beta's$, but not in the $x's$. The second assumption is that x and y are independently and identically distributed (i.i.d.) across observations. This assumption refers to how the sample is drawn and holds if the sample is randomly drawn from the population. The third assumption is that the conditional distribution of u given x has a mean of zero, $E(u|x) = 0$. This assumption indicate that independent variables must be exogenous, the x variables are not allowed to include any information on the error term u. The fourth assumption states no perfect collinearity, which imply no linear relationship among the independent variables. The last assumption refers to homoskedasticity, indicating that all the error terms have the same variance and are not correlated with each other (Verbeek, 2008).

Further in this thesis I use the static approach, indicating that demand is in equilibrium with observed prices, and therefore also time-invariant since it do not consider short term adjustment. Formally the static approach of quantity of gasoline demand i can be stated as following,

$$\ln Q_i = \beta_i + \eta_y \ln Y + \eta_p \ln P_i + \sum \ln Z_k$$

where Q_i denotes quantity of gasoline, Y real income and P_i real price of car fuel. The η represent the elasticity estimates of Y and P. Z_k denote household characteristics, other exogenous variables and time, to account for steady changes in tastes.

The observations in my data are from different years, thus, the time-invariant model needs some adjustments in order to correct for the different time periods. With the flexible properties of the OLS model, I'll then include time fixed effects by dummy variables for each year. In addition to the time specific effects there are also a wide range of regional factors that could affect gasoline consumption, availability of public transportation, infrastructure or cultural differences. Since many of these factors are unmeasurable, dummies representing each region is included to capture the regional fixed effects. More detailed description on these model specifications are discussed in the next section.

The alternative approach to the static model could have been to use the dynamic approach. The dynamic approach accounts for time-dependent changes in the demand for fuel, by including a lagged endogenous variable of previous period level of demand. Many fuel demand researchers are following this approach and argue that fuel consumption should be a function of not only present price and income, but also of previous periods demand,

$$\frac{Q_t}{Q_{t-1}} = \left(\frac{Q_t^*}{Q_{t-1}}\right)^{1-\lambda}$$

That is since households are inflexible in their stock of durable goods such as car or location, thus adaptation to a change in the fuel price or their income is expected to be done partially in each period. Houthakker et al. (1974) provides further description of this approach in the context of gasoline demand. The dynamic model can be specified as,

$$\ln Q_t = \beta + (1-\lambda)\eta_y \ln Y_t + (1-\lambda)\eta_p \ln P_t + \lambda Q_{it-1} + \sum \ln Z_k$$

This model is a partial adjustment model or more commonly called the lagged endogenous model (Dahl and Sterner, 1991). The estimated regression coefficients of Y_t and P_t are the short run income- and price elasticity estimates. Solving for η_y and η_p, by dividing them with $(1-\lambda)$ then yields the long run elasticity estimates.

In relation to the static approach this model can explicitly estimate long and short run elasticities. Although, with reference to the length of the data used in this study with not much variation in the price of gasoline, the long run estimates might be difficult to get consistent.

Since regional specific elasticities is the main interest of my study, I'll include interaction terms between region dummies and the price variable and the expenditure variable. With this method I can both facilitate the utilization of the entire sample, allowing a more efficient estimation, and derive the elasticity estimates directly from the estimated coefficients. Interaction terms have been used in several previous household demand studies (e.g West and Williams, 2004;Wadud et al., 2010). An alternative could be to divide into regional subsamples (Pollak et al., 1995), but since previous studies using this approach have had problems with insignificant estimates, possibly explained by the reduced sample size (e.g. Archibald and Gillingham, 1980), it will not be further studied in this analysis.

There are one main statistical problem to estimate the type of model proposed based on the data. The data set includes about 20 percent households with zero consumption of gasoline. In table 2 the specific share of non-using gasoline households are shown in percentage form.

TABLE 2. Number of gasoline/non-gasoline consumers in each region

Region	0	1	Non users
Rural north	269	780	25.5 %
Urban north	201	830	19.5 %
Southern areas	519	2598	16.7 %
Major towns	1213	5578	17.9 %
Malmö and Gothenburg	643	1981	24.5 %
Stockholm	921	2378	27.9 %
Overall	3766	14145	21 %

Source: Data from Statistics of Sweden (2016)

The issue with zero observations is a well-known problem in household demand analysis. Basically since it complicates the estimation procedure as the dependent variable, demand of gasoline, will be zero in some cases and thus censored. Using the OLS model for the total sample is not efficient, since it cannot take special account for zeros in the dependent variables. Therefore is a subsample OLS used as the base model in this study, only including the positive observations of gasoline consumption. Although, this type of model might cause problems with sample-selection bias since the sample is not random any more. In order to correct for this, I'll include an additional model to correct for this potential selection bias. Before deciding on what econometric model to use in this sense, I need to identify why there might be zero observations.

From a theoretical point of view, demand is constrained to be nonnegative, i.e. no zero observations. In household demand studies on the other hand it is often treated as a special case of rationing, where the rationing out (zero consumption) of some commodity goods can typically be due to three possible sources according to Newman et al. (2001), (i) the household does not purchase the good for economic reasons, e.g. price or income; (ii) the survey period is shorter than the good's purchasing cycle thus fail to include positive positive purchase by a consuming household; (iii) the household does not participate in the market for non-economic reasons e.g habitual, norms or environmental considerations.

Econometric models where the dependent variable include zero observations generally use a latent variable to represent it, indicating that each household has an latent expenditure (unobserved) which is known for some and unknown for others. These models are called censored models, where information on the dependent variable is lost but not data on the independent variables. The problem with censored dependent variables was initially recognized by Tobin (1958), who found that the use of OLS for such models results in biased and inconsistent estimates. The Tobit model initiated by Tobin, assumes that zero expenditure is exclusively due to economic factors such as lack of income or high prices, known as a standard corner solution. This assumption is rather restrictive in the settings of a household demand of gasoline, since zero observations can also be attributed to other factors such as habitual non-usage, infrequency of travelling or better alternative transport modes.

To better address this potential bias in the gasoline demand context, the Heckman's two stage sample selection estimator (Heckman, 1976) has been widely used (e.g Blow and Crawford, 1997; Kayser, 2000; West and Williams, 2004) and will therefore also be followed in this study.

Heckman's two stage procedure is specified by a selection equation and an outcome equation. Statistically, the outcome equation is estimated upon the choice of consuming gasoline (selection equation). This allows to make statements about how the choice to consume may structurally affect its consumption level. Assuming the initial demand equation of gasoline to be summarized in the following way,

$$\ln Q_i = \beta_i + \eta_y \ln Y + \eta_p \ln P_i + \sum \ln Z_k \equiv x_1\beta_1 + \upsilon_1$$

Then to include the zero observations consistently by the Heckman's approach, the first stage is to use the selective equation that is estimated as a probit regression,

$$g_i^* = x\delta_2 + \nu_2$$

where the dependent variable g_i^* is a latent variable which takes values zero or one, where one represent if the household consumes gasoline or alternatively (e.g Kayser, 2000 and West and Williams, 2004) if the household owns a car. This regression express the choice of a household to consume and from the equation the independent variables x are different from those in the initial demand equation, x_1 due to the exclusive restriction. Variables influencing a household to own or not own a car should be different from the variables influencing how much quantity of gasoline to purchase. From the selective equation the inverse Mills ratio can be calculated based on the parameter estimates $\hat{\delta_2}$. Then by including the inverse Mills ratio in the initial regression model we get to the second step known as the outcome equation,

$$\ln Q_i = x_1\beta_1 + \gamma_1\lambda(x\hat{\delta_2}) + \epsilon_i$$

this equation will give consistent estimates of the parameter vector β_1, by the inclusion of λ. A well written description about this method can be studied further in Heien and Wesseils (1990).

5.2. **Potential drawbacks with single equation model.** Overall the single equation approach is attractive in its simplicity, nonetheless Sadoulet and De Janvry (1995) points out some important drawbacks. Primarily the choice of functional forms of the demand equation and variables to include is arbitrary. The structure employed are not based on economic theory and usually rely on computational convenience, common sense and interest in specific elasticity estimates. As a result there are uncertainties on what is actually measured and if the estimates are sufficiently derived from household behaviour. Further concerning the inflexibility of the functional forms where the elasticities are constant over all values of the exogenous variables. Specifically this might cause inconsistent results when studying gasoline, since consumers are found to have different income elasticities related to their income but also that this elasticity change as income increase. All of these shortcomings are considered in the model I intend to use in this study.

Many gasoline household studies use alternative approaches where gasoline demand is incorporated in the context of consumer production theory. In these studies households' need for transportation is fulfilled through the utilization of a household's vehicle stock in which gasoline is used. Thus gasoline demand is employed as derived demand and thus more in line with economic theory (Lancaster, 1966). This can be done using the Heckman's two stage approach as I intend to do also in this study.

Considering the functional form, a standard linear model will be used but including interaction terms to explore impacts from regional features of demand of gasoline. Interaction terms have previously been used in household demand studies (e.g. Archibald and Gillingham, 1980; Rouwendal, 1996; Kayser, 2000; Nicol, 2003), although most of them include a price and income interaction term and report variation of price elasticity with respect to income or income groups. This however fails to study differences in price elasticity between households based on other characteristics than income level. Wadud et al. (2010) acknowledge this shortcomings and use additional interaction terms between price and rural location as well as income and rural location. This method will be followed in this thesis.

The main disadvantage of using a single equation model, however, is that it does not allow for analysing possible complementarity or substitutability between the various goods comprising the household consumption basket. In order to model this, a complete demand system is required, which is explained in the next sub section. Although, this method will not be used in this thesis, primarily for reasons of data limitations. The data at hand simply better suits a standard linear OLS method.

5.3. **Demand systems.** Demand systems can efficiently include the whole consumption basket and therefore require a large amount of detailed data on consumption as well as prices. In the context of household gasoline consumption this method has been used by researchers worldwide (e.g Brännlund and Nordström, 2004; Nicol, 2003; West and Williams, 2004)

These system of demand equations applied with consumer demand theory was originally introduced by Stone (1954). Since then, a number of different specifications and functional forms have been proposed, where one of the most examined is the almost ideal demand system (AIDS) which can be studied further in Deaton and Muellbauer (1980). The benefits of these systems are that they are able to take into account the interdependence of large numbers of goods in the choices made by consumers. In addition, efficiently incorporate the neoclassical assumption of maximizing utility. For a utility maximizing consumer, the total expenditure or income y is equal to a cost function representing the minimum expenditure necessary to attain the maximum utility level at given prices, $c(u, p)$. This equality can be inverted to yield the indirect utility function, $v(p, y)$ which is the maximum utility the consumer can reach for a given income y at given prices p. There are many ways such a system can be built and they can not easily be compared since the interpretation of the elasticities is model specific. Nicol, 2003 indicate that while elasticities might change in the same direction as family size in one country, this is not observed in another country. In order for gasoline elasticity estimates to provide guidance for policy makers, this statement imply the importance of actually studying individual countries more precise as my study intend to do.

6. Model specification

The gasoline demand model examined in this study is first estimated by using subsample OLS in the log-log linear and static framework. By subsample, it means only including households with positive spending on gasoline.

$$\ln Q_{it} = \beta + \beta_e \ln E_{it} + \beta_p \ln P_t + H_i + \sigma_t + \gamma_j + \varepsilon_{it} \tag{1}$$

where Q_{it} is gasoline consumption of household i in year t, P_t is the aggregate real price of gasoline in year t, E_{it} denote total expenditure of household i in year t and η the corresponding elasticities. H_i denotes a vector of household-specific characteristics. σ_t represent the fixed time effects to capture the seasonality of gasoline demand and γ_j represent the regional fixed effects.

This constant elasticity model, is adopted also in this study since it provide a good fit of the data and further allows for direct comparison with previous results from the literature. Since the household data is repeated cross-section we could in principle treat is as a large pooled cross-section. However, this ignores the time dimension and therefore year specific dummies will be included to represent the fixed time effects (ε_t). In addition there are regional factors that could affect gasoline consumption, therefore dummies representing each region is included to capture the regional fixed effects (ε_j).

6.1. **Interaction Parameter Model.** To accommodate the possibility that households in different regions have different price or income responses, interaction between the price and total expenditure with dummies representing each region will be used. The final model estimated with OLS is then formally,

$$\ln Q_{it} = \beta_0 + \sum_{j=5}^{J} \beta_j D_{ij} + \sum_{t=7}^{T} \beta_t D_{it} + \left(\beta_e + \sum_{j=5}^{J} \beta_{ej} D_{ij} \right) \ln E_{it} + \left(\beta_p + \sum_{j=5}^{J} \beta_{pj} D_{ij} \right) \ln P_t + \sum_{k=7}^{K} H_i + \varepsilon_{it} \equiv X_i^{'} \beta + \varepsilon_{it} \tag{2}$$

D_{it} are the dummy variables of the the different years, D_{ij} are the dummy variables of household i in region j. Since we include $k-1$ number of dummy variables, one of them will be excluded among the regressors and work as the reference. The parameters β represent the corresponding elasticities for the variables, besides β_0 which is a constant and ε the error term.

The interaction term $\ln P_t D_{ij}$ captures the extent to which the responsiveness of households to price changes increase or decrease as place of resident changes. In this specification, the price elasticity of gasoline demand for each region is equal to $\beta_{pj} = \beta_p + \beta_{pj}$, where region one has the price elasticity estimate $= \beta_p + \beta_{p1}$. Given that the price elasticity is less than zero, a positive coefficient estimate of β_{pj} indicate a decrease in the price response for households in the given region. The same interpretation goes for the interaction $\ln E_t D_{ij}$.

With the region-year repeated cross-section data at hand, Bertrand et al. (2002) and Kezdi (2003) emphasize that clustering can be present even after including region and year effects in the regression and valid statistical inference requires controlling for clustering within regions. Failure to do so might lead to under-estimated standard errors and low p-values. A solution is to use cluster-robust standard errors which allows for independence across clusters but correlation within clusters. This is convenient for estimators to retain their consistency when statistical inference since the usual cross-section assumption of independent observation is no longer appropriate (Cameron and Miller, 2015). Even though I also have different time periods, the clustering should not be year-region since the error for urban areas in 2006 is likely to be correlated with the error for urban areas in 2007.

6.2. **Heckman selection specification.** The second estimation is the Heckman selection model. This is done in order to correct for possible sample selection bias from the subsample OLS. First, a probit regression is computed that determines the probability that a given household will consume gasoline, where the decision to consume is modelled as a dichotomous choice problem. From this regression the inverse Mills ratio (λ_i) is estimated, which is included as an instrument in the second stage. The second stage is known as the outcome equation and is the initial demand model (5), including the λ_i which incorporates the censoring latent variables to control for the bias caused by non-random sampling.
First stage probit regression:

$$S_i^* = \alpha_o + \alpha_1 \ln p_i + \alpha_2 \ln y_i + \alpha_3 TEN_i + \sum_k \alpha_k H_{ik} + \nu_i \equiv Z_i^{'} \alpha + \nu_i \tag{3}$$

where S^* is the unobserved spending on gasoline which is equal to one if the household has positive spending on gasoline and equal to zero if not. H_{ik} is a vector of household characteristics that determine different preferences in the spending decision on gasoline. TEN is a dummy variable of whether the household owns their place of residence of not. Previous theoretical work regarding the specification of (6) is limited, Heien and Wesseils (1990) suggest that prices, total expenditure as well as demographic characteristics should be of equal importance in the probit model to those expected in traditional demand analysis. However, in order to correctly interpret the parameters in the two equations either the error terms must be uncorrelated, or if not there must be at least one variable in the probit equation that is not included in the outcome equation (Maddala, 1983).

West (2004), suggest home ownership to be a good alternative since it acts as a proxy for wealth and access to credits and therefore increase the likelihood of owning a car and thus the likelihood of consuming gasoline, but is not expected to affect quantity of gasoline consumed. In the data

I have, there's no information regarding car ownership. Although, it can be assumed that households with positive expenditure own or at least have access to a car and therefore it can be used as a proxy for car ownership. Thus, variables that explain car ownership can also be used in this settings and therefore is home ownership used as the additional variable in the probit regression.

Furthermore, we assume that $(\nu_i, \varepsilon_{it})$ has a bivariate normal distribution with correlation ρ and zero means. Following the specification by Kayser (2000), the expected gasoline demand of (5) become,

$$E[\ln(Q_{it})|S_i > 0] = E[X_i^{'}\beta + \varepsilon_{it}|H_{ik}^{'}\alpha > \upsilon_i] = X_i^{'}\beta + \rho\sigma_\varepsilon \frac{\phi(Z_i^{'}\alpha)}{\Phi(Z_i^{'}\alpha)}$$

where,

$$\frac{\phi(Z_i^{'}\alpha)}{\Phi(Z_i^{'}\alpha)}$$

is the inverse Mills ration that denotes the non-selection hazard. Lastly the second stage outcome equation is expressed,

$$\ln Q_{it} = X_i^{'}\beta + \beta_\lambda\lambda_i + \varepsilon_{it}\lambda = \frac{\phi(Z_i^{'}\alpha)}{\Phi(Z_i^{'}\alpha)} \tag{4}$$

Leung and Yu (1996) points out that the degree of collinearity between the regressors used in the outcome equation and λ_i is the decisive criteria to judge the appropriateness of the Heckman's approach in relation to subsample OLS. The lack of exclusion restrictions, as in this study, is likely to cause collinearity issues and to test this we follow
's example and estimate λ_i on all the regressors in the outcome equation.

Following (Kayser, 2000), the short run elasticities can be estimated by,

$$\eta_p = \frac{\partial E[\ln(Q_i)|X_i]}{\partial p} = \beta_1 + \beta_2$$

$$\eta_y = \frac{\partial E[\ln(Q_i)|X_i]}{\partial y} = \beta_2 + \beta_2$$

7. Results and discussion

Table 3 shows the empirical results from the models estimated in this thesis. The first column (1) is the initial subsample OLS estimates, the second (2) denote the results from the outcome equation of Heckman's two stage model and (3) the probit regression of Heckman's model. The numbers in the parenthesis under the coefficient estimate are the standard errors. In the initial subsample OLS (1) these are adjusting for clustering. The clustering procedure resulted in overall lower standard errors of the estimates of the parameters.

Since dummy variables are used in the model to represent the different regions and other demographic characteristics of the household, a reference household is defined. The reference household is headed by a female, with no children and living alone. The household is located in the region defined as "Major towns". Interpretation of the variables in table 3 is then done with respect to this reference household.

The purpose of using the Heckman's model, was to eliminate problems with selection bias by including the inverse Mills ratio in the gasoline demand (outcome) equation. Although, since the Mills ratio,λ, is significant it means we have high correlation between the error terms of the two models and the problem with selection bias is not solved. Regressing λ on all the variables in the outcome equation result in highly significant parameters and adjusted R^2 of 0.89, which imply problems with collinearity. The problem with collinearity in this study is most likely related to that only one additional variable, home ownership, is included in the selection model and/or that home ownership also explains the quantity of gasoline consumed.

In order to correct for collinearity, the standard procedure is to find more appropriate exclusive restrictions. That is, find variables that determine the probability to consume gasoline, but do not determine the quantity consumed directly. Two examples could be distance to work or if holding a driver license. Puhani (2000) suggest if there are problems with collinearity in the Heckman's procedure that can not be solved, subsample OLS may be the most robust estimator. Due to problem with collinearity connected to the Heckman's approach, and the lack of possibilities to reduce collinearity with the data at hand, from now on I'll only focus on the estimates from initial subsample OLS (1) since this model is assumed to be the most robust alternative.

First looking at the household characteristics, the male variable is significant and positive, which indicate that households where male is the head of household have a positive influence on demand of gasoline. This finding is well-supported in the literature. Carlsson-Kanyama and Linden (1999) studies 45 000 individuals in Sweden and confirm that men with high income are the most intensive users of car fuel. Since a dummy variable is not a continuous variable, we can not interpret the parameter estimate as the marginal effect of log of demands. Instead Halvorsen et al. (1980) show that the percentage change of a dummy variable is given by $e^{\beta} - 1$ where β is the coefficient estimate of the dummy. In this case, households with male head consume 14.8 percent more gasoline than those with female head.

Age of household head is significant and have expected signs as the age of household head has a positive influence on gasoline demand, while the negative sign of age squared means that as the head of household gets older, the effect of age is lessoned. The more adults and children there are in the household have a significant positive effect on gasoline consumption, which is consistent to results by West and Williams (2004).

Turning over to the research question of this study, which questioned if there are potential differences between the gasoline elasticities of demand in rural and urban regions in Sweden. In table 3 we first find that the positive statistically significant coefficient of the interaction term between price and rural north establish that gasoline demand for households in the rural areas are less price elastic than households in major towns. This finding

Table 3. Results

	Dependent variable:		
	ln gasoline demand		gasoline expenditure
	OLS		*probit*
	(1) subsample	(2) outcome	(3) selection
ln p(price)	-0.725^{***}	0.163	-1.897^{***}
	(0.123)	(0.144)	(0.206)
ln exp(total expenditure)	0.447^{***}	0.266^{***}	0.427^{***}
	(0.023)	(0.027)	(0.038)
male	0.141^{***}	0.059^{***}	0.174^{***}
	(0.015)	(0.017)	(0.024)
ln age	7.610^{***}	5.745^{***}	2.554^{***}
	(0.565)	(0.582)	(0.829)
ln age^2	-0.971^{***}	-0.747^{***}	-0.306^{***}
	(0.075)	(0.077)	(0.111)
adult	0.280^{***}	0.105^{***}	0.256^{***}
	(0.020)	(0.024)	(0.029)
adults	0.364^{***}	0.274^{***}	0.070^{**}
	(0.017)	(0.030)	(0.049)
child	0.070^{***}	0.040^{*}	0.054
	(0.025)	(0.021)	(0.035)
children	0.096^{***}	0.094^{***}	-0.025
	(0.029)	(0.019)	(0.032)
2004	0.007	-0.038	0.083^{*}
	(0.001)	(0.024)	(0.044)
2005	0.073^{***}	-0.040	0.219^{***}
	(0.001)	(0.025)	(0.041)
2006	0.049^{***}	-0.097^{***}	0.284^{***}
	(0.002)	(0.026)	(0.041)
2007	0.082^{***}	-0.010	0.153^{***}
	(0.002)	(0.024)	(0.039)
2008	0.055^{***}	-0.003	-0.078^{**}
	(0.002)	(0.024)	(0.038)
2009	0.013^{***}	-0.016	0.010
	(0.024)	(0.024)	(0.046)
Stockholm	-1.185^{***}	1.358^{*}	-4.850^{***}
	(0.771)	(0.786)	(1.133)
Southern areas	0.751^{***}	-0.786	3.648^{***}
	(0.798)	(0.734)	(1.180)
Urban north	0.029	-1.279	2.766
	(1.100)	(1.098)	(1.764)
Rural north	-1.616	-3.320^{***}	2.951^{*}
	(1.158)	(1.161)	(1.645)
Malmö/Gotheburg	-0.722	1.556^{*}	-2.789^{**}
	(0.803)	(0.812)	(1.208)
Interaction terms			
ln exp:Stockholm	0.138^{***}	-0.100^{**}	0.375^{***}
	(0.038)	(0.043)	(0.060)
ln exp:Southern areas	-0.072^{**}	0.016	-0.195^{***}
	(0.038)	(0.038)	(0.062)
ln exp:Urban north	-0.068	0.008	-0.175^{**}
	(0.057)	(0.057)	(0.088)
ln exp:Rural north	-0.008	0.064	-0.222^{***}
	(0.062)	(0.062)	(0.086)
ln exp:Malmö/Gotheburg	0.107^{**}	-0.036	0.144^{**}
	(0.041)	(0.042)	(0.063)
ln p:Stockholm	-0.329	-0.078	-0.085
	(0.215)	(0.215)	(0.331)
ln p:Southern areas	0.095	0.258	-0.461
	(0.209)	(0.209)	(0.355)
ln p:Urban north	0.349	0.494	-0.244
	(0.331)	(0.330)	(0.541)
ln p:Rural north	0.734^{**}	1.093^{***}	-0.185
	(0.336)	(0.335)	(0.498)
ln p:Malmö/Gotheburg	-0.288	-0.430^{*}	0.291
	(0.227)	(0.226)	(0.356)
λ(Inverse Mill's ratio)		-1.169^{***}	
		(0.099)	
TEN			0.131^{***}
			(0.010)
Constant	-11.812^{***}	-7.577^{***}	-5.624^{***}
	(1.111)	(1.161)	(1.625)
Observations	14,145	14,145	17,911
R^2	0.209	0.216	
Adjusted R^2	0.207	0.215	
F Statistic	124.068^{***} (df = 30; 14114)	122.463^{***} (df = 31; 14113)	

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$
In brackets= Standard errors

is consistent with similar studies in the UK by Blow and Crawford (1997) and Santos and Catchesides (2005).

In order to answer this research question we have to look at the estimated elasticities in this study of the different regions illustrated in table 4. These elasticity estimates will be interpreted to be in the intermediate run, since Dahl and Sterner (1991) among others indicate that static models tend to produce price elasticities between the short and long run. The elasticity estimates in table 4 are only based on results from the initial OLS, since we assume that this model generated the most robust estimates, and they are found by adding the coefficient estimate of the price and expenditure (income) variable with the coefficient estimate of the interaction between these and the regions. The reference is major towns, indicating that the price elasticity of major towns is equal to the coefficient of the price variable, β_p=-0.72, and the same for income elasticity, which then is equal to the expenditure coefficient, β_i=0.45. For the other regions, recall that the price and income elasticities are estimated by, $\beta_{pj} = \beta_p + \beta_{pj}$ and $\beta_{ej} = \beta_e + \beta_{ej}$ where the subscript j denote the different regions, thus, Stockholm will have price and income elasticity estimate of,

$\beta_p + \beta_{pStockholm} = -0.72 + (-0.329) = -1.05$

$\beta_e + \beta_{eStockholm} = 0.45 + 0.138 = 0.58$

where $\beta_{p_{Stokcholm}}$ is the coefficient of the interaction term of the price variable and the regional dummy for Stockholm (= *p:Stockholm*) and $\beta_{e_{Stokcholm}}$ is the coefficient of the interaction term *exp:Stockholm.*

TABLE 4. Elasticity estimates of the different regions

	Initial subsample OLS	
Region	β_{pj}	β_{ij}
Major towns	-0.72***	0.45***
Stockholm	-1.05***	0.58***
Southern areas	-0.63***	0.38***
Urban north	-0.38***	0.38 ***
Rural north	0.01***	0.44*
Malmö/Gothenburg	-1.01	0.55
Average	-0.63	0.46

Note: *p<0.1; **p<0.05; ***p<0.01

The average price elasticity in Sweden is estimated to be -0.63, which means that as the price of gasoline increase with 10 percent, gasoline demand will decrease with 6.3 percent. In relation to previous findings in the literature, the average price elasticity I've found in this study is slightly higher. Dahl (2012) finds that the average price elasticity from previous static models worldwide is in the range of -0.41 - 0.00. The explanation for my higher results could be first that European price elasticities of gasoline generally tend to be higher than average, due to the better availability to other transport modes such as public transportation. Second if we look at the sample, 70 percent of the households studied are located in major towns, Stockholm and Malmö/Gothenburg, all areas in which alternative transport to car transport for sure is an alternative.

With respect to the different regional elasticity estimates, all are significant but Malmö/Gothenburg. The overall pattern of the price elasticities illustrate that the rural region is inelastic and Stockholm is very elastic. A similar pattern is found by Bureau (2011) in France and Santos and Catchesides (2005) in the UK, where the rural region seem to be less responsive to price changes and urban areas on the other hand are more responsive. Also Nicol (2003), find regional differences of price elasticities to vary between -0.103 and -0.894 in Canada.

The reason why the regional price elasticities are different could be that households in rural regions for instance tend to keep their car longer and thus their car stock turns over to efficient vehicles more slowly and therefore lowering their price responsiveness. This is also connected to infrastructural issues, where it in rural areas might not be possible to switch to vehicles driven by alternative fuels since there might, for instance, be limited places to charge an electric car. Also, infrastructural issues in terms of public transportation, which is less developed in rural areas and thus decrease the opportunity to substitute between different modes of transport.

In table 4 it's shown that the urban north price elasticity is -0.38 compared to southern areas of -0.68 , which is quite interesting if the reasoning that access public transportation might have influence on the price elasticity. In this case the urban north should in principal be more price responsive than southern areas since it is not defined as an urban area, but this is not the case. One explanation for this could be that work and home may be further apart in the northern regions, thus, even if there are better public transportation these transits are not suitable to fit the travelling behaviour of the households. Another reason could be related to cultural differences, where it's more accepted to use other types of transport than a car in the southern areas compared to in the northern region.

Even though the general pattern of the price elasticity estimates found in this study are similar to those reported by others, it is possible that they are misspecified which may lead to inconsistent estimators whose the properties are unknown (Blundell et al., 2012). In the rural north for instance, the price elasticity is estimated to be almost perfectly inelastic at 0.01, which indicate that as the price of gasoline increase with 10 percent, the demand of gasoline will be almost unchanged or slightly go up. This might be due to that households in these areas face higher prices in the real world, compared to the annual average gasoline price in Sweden used as the price variable in this thesis. Using a constant price variable is problematic when analysing regional effects, and is used in this study mainly because of data limitations of regional gasoline prices. The best solution for this would be to find more detailed price data and then apply a non parametric estimator to be sure to find consistent estimates, i.e. use a system of demand model instead of the linear single equation. An alternative solution may have been to drop the price variable from the estimation and use an expenditure function instead of a demand equation. Thus, in this case I would not have found any price elasticity estimates of the different regions but only the income elasticities.

Turning to the income elasticity estimates in table 4, the average is 0.46 which is in line with previous studies. Dahl (2012) finds that the average income elasticity is between 0.00-0.53.The income elasticity estimates are more reliable compared to the price elasticity, since these are based on the households individual total expenditure and therefore more precise compared to the average price variable. However, the various income elasticity estimates of the regions are positive and below one, thus, gasoline seem to be a normal good in all regions. In addition the elasticity estimates do not change much between the regions, which is also found by Blow and Crawford (1997) and Wadud et al. (2010). Therefore we can conclude that the regional dimension in terms of income elasticity is not very strong. More specifically, however, Stockholm has an income price elasticity of 0.58, which indicate that as income increase with 10 percent, gasoline demand increase with 5.8 percent. Where the rural north on the other hand has an income elasticity of 0.44, which imply that as income increase with 10 percent, gasoline demand increase with 4.4 percent. Although the difference is quite small these results stress that the income effect is slightly stronger in Stockholm, which means that as income for households in Stockholm increase they will demand more gasoline compared to if income would rise with the same percentage share for households in the rural region.

Finally, in order to see the overall how well the subsample OLS preform and if its results is trustful, first we can turn to the standard adjusted R^2. In this model it's 20.9 percent, which means that the independent variables explain 20.9 percent of the variation in the dependent variable. This might be a low result, however the adjusted R^2 is not the final answer whether the model preform well or not since it does not provide a formal hypothesis test for the relationship between the independent and dependent variables. For this we need to do an F-test. The F-test provides a test of the overall significance. It compares an intercept only model with the model specified (subsample OLS). The null hypothesis for the F-test is that there's no relationship between quantity of gasoline demanded and the independent variables. The overall F-test is significant and therefore the null hypothesis is rejected and the subsample OLS model is expected to provide a better fit than the intercept only model.

8. Conclusion

In this thesis the aim has been to investigate the potential differences of gasoline demand elasticities in different regions in Sweden, using the static log-log linear OLS method and Swedish Household budget data between 2003-2006 and 2012. The results show that the regional differences in terms of price elasticity of gasoline seems to vary more distinct in comparison to the income elasticity.

Thus, the increased fuel tax initiated by the Swedish government will trigger different responses in households living in different regions. Given the price elasticity estimates, northern households will carry a greater share of the financial burden of the tax, especially the rural ones. In Stockholm, Gothenburg and Malmö on the other hand the tax increase might be very efficient in reducing demand of gasoline since the price elasticity is elastic.

This means that if there is a political will to equalize the distributional effects, households in rural areas should be compensated. Lower fuel taxes in these areas might not necessarily be favoured though, since differentiation would kill the incentive to restrict gasoline consumption. The distributional impacts from a tax increase should rather be handled by recycling the revenues efficiently.

In addition, it should be remembered that the model presented here is based on estimates in the intermediate-run. In the long-run, households can adapt their travelling behaviour in response to higher gasoline taxes by switching to a more fuel efficient car or using alternative transport modes. In this case the distributive effect of a tax reform will be reduced. Although, this is very hard to test, since it will depend on future infrastructural changes specifically relevant in the rural areas.

8.1. **Ethical considerations.** Two ethical dimensions can be related to this study. First, with respect to the regional location of the household with the ethical perspective of the direct welfare effects connected to the regions. Second, as the goal of the environmental tax to reduce carbon emissions from private car travelling, thus take responsibility for the welfare of future generations.

The sometimes troublesome dilemma of an environmental tax, where it's in theory is justified to tax those who bear the burden of negative external effects as long as the social cost exceeds the cost of the tax. However, if those most economically punished by the tax, such as rural households in this study, are so for reasons out of their reach the environmental tax might not be justified in reality. The environmental fuel tax with the goal to change behaviour to decrease carbon emissions, is not ethical in the sense that rural households don't have the same ability to change their behaviour as the urban households. This study aim to highlight the importance of evaluating behavioural changes in order to actually reach the goal of an environmental fuel tax policy.

In the long run, on the other hand, there's an ethical responsibility for future generations. The transport sector has an overall increasing trend of carbon emissions. What is really needed is to clarify how we should transport ourselves in the future since the current situation is not sustainable. If a fuel tax can assist financially to develop a future strategy then the unequal welfare changes today might not be as relevant.

8.2. **Future studies.** Future research with more detailed price data on regional level using a more sophisticated system of demand approach would be interesting to see with respect to specific elasticity estimates. In addition a more in depth study on why rural regions seem to respond to less to price changes and what is actually needed to change behaviour in terms of driving less or more efficient. More relevant research is needed on a sub-national level of the transport sector.

References

Akerman, J., Banister, D., Dreborg, K., Nijkamp, P., Schleicher-Tappeser, R., Stead, D., Steen, P., 2000. European transport policy and sustainable mobility. Routledge.

Archibald, R., Gillingham, R., 1980. An analysis of the short-run consumer demand for gasoline using household survey data. The Review of Economics and Statistics, 622–628.

Becker, G. S., 1965. A theory of the allocation of time. The economic journal 75 (299), 493–517.

Bento, A. M., Goulder, L. H., Jacobsen, M. R., Von Haefen, R. H., 2009. Distributional and efficiency impacts of increased us gasoline taxes. The American Economic Review, 667–699.

Bertrand, M., Duflo, E., Mullainathan, S., 2002. How much should we trust differences-in-differences estimates? Tech. rep., National Bureau of Economic Research.

Blow, L., Crawford, I., 1997. The distributional effects of taxes on private motoring.

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

Brännlund, R., Nordström, J., 2004. Carbon tax simulations using a household demand model. European Economic Review 48 (1), 211–233.

Bureau, B., 2011. Distributional effects of a carbon tax on car fuels in france. Energy Economics 33 (1), 121–130.

Cameron, A. C., Miller, D. L., 2015. A practitioner's guide to cluster-robust inference. Journal of Human Resources 50 (2), 317–372.

Carlsson-Kanyama, A., Linden, A.-L., 1999. Travel patterns and environmental effects now and in the future:: implications of differences in energy consumption among socio-economic groups. Ecological Economics 30 (3), 405–417.

Council of the European Union, 2003. Council directive 2003/96/ec of 27 october 2003 restructuring the community framework for the taxation of energy products and electricity. Official Journal of the European Union L283.

Council of the European Union, 2009. Regulation (ec) no 443/2009/ec of the european parliament and of the council of 23 april 2009 setting emission performance standards for new passenger cars as part of the community's integrated approach to reduce co 2 emissions from light-duty vehicles. Official Journal of the European Union L140.

Dahl, C., Sterner, T., 1991. Analysing gasoline demand elasticities: a survey. Energy economics 13 (3), 203–210.

Dahl, C. A., 2012. Measuring global gasoline and diesel price and income elasticities. Energy Policy 41, 2–13.

De Jong, G., Gunn, H., 2001. Recent evidence on car cost and time elasticities of travel demand in europe. Journal of Transport Economics and Policy, 137–160.

Deaton, A., Muellbauer, J., 1980. An almost ideal demand system. The American economic review 70 (3), 312–326.

Eliasson, J., Pyddoke, R., Swärdh, J.-E., et al., 2016. Distributional effects of taxes on car fuel, use, ownership and purchases. Tech. rep., CTS-Centre for Transport Studies Stockholm (KTH and VTI).

Espey, M., 1998. Gasoline demand revisited: an international meta-analysis of elasticities. Energy Economics 20 (3), 273–295.

European Commission, et al., 2011. Roadmap to a single european transport area-towards a competitive and resource efficient transport system. White Paper, Communication 144.

Friedman, M., 1957. A theory of the consumption function. Vol. 63. JSTOR.

Frondel, M., Vance, C., 2013. Fuel taxes versus efficiency standards–an instrumental variable approach. Ruhr Economic Paper (445).

Halvorsen, R., Palmquist, R., et al., 1980. The interpretation of dummy variables in semilogarithmic equations. American economic review 70 (3), 474–75.

Heckman, J., 1976. The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models.

Heien, D., Wesseils, C. R., 1990. Demand systems estimation with microdata: a censored regression approach. Journal of Business & Economic Statistics 8 (3), 365–371.

Houthakker, H. S., Verleger, P. K., Sheehan, D. P., 1974. Dynamic demand analyses for gasoline and residential electricity. American Journal of Agricultural Economics 56 (2), 412–418.

Hughes, J. E., Knittel, C. R., Sperling, D., 2008. Evidence of a shift in the short-run price elasticity of gasoline demand. The Energy Journal, 113–134.

IPCC, 2015. Climate Change 2014: Mitigation of Climate Change. Vol. 3. Cambridge University Press.

Kayser, H. A., 2000. Gasoline demand and car choice: estimating gasoline demand using household information. Energy economics 22 (3), 331–348.

Kezdi, G., 2003. Robust standard error estimation in fixed-effects panel models. Available at SSRN 596988.

Lancaster, K. J., 1966. A new approach to consumer theory. The journal of political economy, 132–157.

Leung, S. F., Yu, S., 1996. On the choice between sample selection and two-part models. Journal of econometrics 72 (1), 197–229.

Liu, W., 2015. Gasoline taxes or efficiency standards? a heterogeneous household demand analysis. Energy Policy 80, 54–64.

Maddala, G., 1983. Qualitative and limited dependent variable models in econometrics. Cambridge: Cambridge University Press.

Mas-Colell, A., Whinston, M. D., Green, J. R., et al., 1995. Microeconomic theory. Vol. 1. Oxford university press New York.

Newman, C., Henchion, M., Matthews, A., 2001. Infrequency of purchase and double-hurdle models of irish households' meat expenditure. European Review of Agricultural Economics 28 (4), 393–419.

Nicol, C. J., 2003. Elasticities of demand for gasoline in canada and the united states. Energy economics 25 (2), 201–214.

Pollak, R. A., Wales, T. J., et al., 1995. Demand system specification and estimation. OUP Catalogue.

Poterba, J. M., 1991. Is the gasoline tax regressive? Tax Policy and the Economy, 145–164.

Proost, S., Van Dender, K., 2011. What long-term road transport future? trends and policy options. Review of Environmental Economics and Policy 5 (1), 44–65.

Puhani, P., 2000. The heckman correction for sample selection and its critique. Journal of economic surveys 14 (1), 53–68.

Puller, S. L., Greening, L. A., 1999. Household adjustment to gasoline price change: an analysis using 9 years of us survey data. Energy Economics 21 (1), 37–52.

Rouwendal, J., 1996. An economic analysis of fuel use per kilometre by private cars. Journal of Transport Economics and Policy, 3–14.

Sadoulet, E., De Janvry, A., 1995. Quantitative development policy analysis. No. 338.9011 S3. Johns Hopkins University Press Baltimore.

Santos, G., Catchesides, T., 2005. Distributional consequences of gasoline taxation in the united kingdom. Transportation Research Record: Journal of the Transportation Research Board (1924), 103–111.

Sperling, D., Nichols, M., 2012. California's pioneering transportation strategy. Issues in Science and Technology 28 (2), 59.

Statistics of Sweden, 2016. Household expenditure. Accessed on 2016-06-01. URL http : / / www . scb . se / sv _ /Hitta-statistik / Statistik-efter-amne/Hushallens-ekonomi/Hushallens-utgifter/Hushallens-utgifter-HUT/

Stern, N., 2007. The economics of climate change: the Stern review. cambridge University press.

Stock, J. H., Watson, M. W., 2012. Introduction to econometrics. Pearson Education.

Stone, R., 1954. Linear expenditure systems and demand analysis: an application to the pattern of british demand. The Economic Journal 64 (255), 511–527.

Swedish Petroleum and Biofuels Institute, 2016. Gasoline price and consumption. Accessed on 2016-06-01.
URL ttp://spbi.se/statistik/priser/bensin/

The Swedish Government, 2015. The government budget propsition 2015/16:1.

Tobin, J., 1958. Estimation of relationships for limited dependent variables. Econometrica: journal of the Econometric Society, 24–36.

Van Dender, K., 2009. Energy policy in transport and transport policy. Energy Policy 37 (10), 3854–3862.

Verbeek, M., 2008. A guide to modern econometrics. John Wiley & Sons.

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

Wales, T. J., Woodland, A. D., 1983. Estimation of consumer demand systems with binding non-negativity constraints. Journal of Econometrics 21 (3), 263–285.

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

West, S. E., Williams, R. C., 2004. Estimates from a consumer demand system: implications for the incidence of environmental taxes. Journal of Environmental Economics and management 47 (3), 535–558.

Appendix A.
Variable definition

Table 5. Variable definition

Variable name	Definition	Proportions of households
$\ln exp$	Logarithm of total annual expenditure (in 2015 prices)	
$\ln p$	Logarithm of annual gasoline price (in 2015 prices)	
$\ln d$	Logarithm of annual gasoline consumption (litre)	
$\ln age$	Logarithm of age of household head	
adult	Number of adults 20+ years	
child	Number of children 0-19 years	
Discrete characteristics		**Proportions of households**
male	=1 if head of household is male =0 otherwise	71.7
Major towns	=1 if located in Major towns =0 otherwise	37.8
Stockholm	=1 if located in Stockholm =0 otherwise	18.4
Southern areas	=1 if located in southern areas =0 otherwise	17.4
Urban north	=1 if located in urban north =0 otherwise	5.8
Rural north	=1 if located in rural north =0 otherwise	5.9
Malmö and Gothenburg	=1 if located in Malmö or Gothenburg =0 otherwise	14.7
$\ln p : R1...5$	interaction term between price and each of the regions	
$\ln exp : R1..5$	interaction term between expenditure and each of the regions	
λ	Inverse Mills ration	