UDC: 336.226(437.3); 336.226.3(437.3)
Keywords: consumer demand – excise taxes – tax simulation – nonparametric regression

Demand for Fuel and its Consequences for Indirect Taxation: A Microeconomic View

Michal SLAVÍK*

1. Introduction

The development of public finance is probably the main mid term economic risk in the Czech Republic. Government deficits are growing steadily and fiscal consolidation is necessary not only to avoid serious macroeconomic instability, but also to fulfil the liabilities that stem from EU membership. The government recently introduced a concept of public finance reform that is based on some hypothetical expenditure cuts and proposals of revenue increases. The increase of revenues is represented among others by higher indirect taxes – value added tax and excise taxes. The government concept assumes that the budget will be improved by almost 90 milliard CZK of additional indirect tax revenues in the next three years.

Any government would try to take advantage of the fact that demand for some goods is less price elastic than others and additional taxation of these commodities should not considerably affect the quantity consumed. Excise taxes – e.g. on petrol, tobacco products or alcohol – are among the first candidates for an increase in taxation in any country, because it is general knowledge that the demand for these commodities are not very price elastic. A tax change will influence different people differently, because the elasticities may differ across the population depending on many factors. Families with higher revenues will not pay as much attention to minor changes in prices compared to low income families. A worker who needs to use a car to get to work every day will respond to a price change differently than somebody who can easily substitute the car for another mean of transportation. An alcoholic will respond to a price change of his favourite bottle diversely to a person who only drinks occasionally. Looking at the aggregate data does not reveal much about a behavioural response of particular individuals or social groups. However it could be interesting for policymakers to consider who will be mainly affected by a tax change, before the tax rate or the tax base is changed. Using micro-data could give answers about a probable impact of a tax change on different groups of people.

* Czech National Bank and PhD student at Faculty of Social Sciences, Charles University, Prague (michal.slavik@seznam.cz)
I wish to thank James Banks, Karel Janda, Bill Kerr, Jaroslav Kochaníček, Jaroslav Kotlík, Ivan Matalík, Robert Murárik, Milan Sojka and Pavel Soukup for helpful comments, data and support. All possible mistakes are however my own.

An existence of household expenditures micro-data on fuel from the *Czech Family Expenditure* Survey, which is carried out by the Czech Statistical Office, creates the opportunity to investigate effects that would not be depicted using the (macro) aggregated data. By using the micro-data, we can estimate the effects of a tax change on different social groups, because the *Family Expenditure Survey* includes information about surveyed households: size of the family/household, composition of a household, number of children, number of economically active members, age, education and social group of the head of the household, size of the town where the household lives and many other demographic characteristics. The distinction whether to use macro or micro data depends on the purpose of the analysis. If we are only interested in how much extra money the government gets, it is preferable to use macro data and carry out a time series analysis. If we are curious about who will actually pay the tax burden, we should take into account the micro data. It would be theoretically possible to use the micro data to estimate the change of the overall tax revenue as well, but the characteristics of the population surveyed in the family expenditure surveys is not identical to the whole population of a country. Some people are not surveyed (e.g. people who do not have a permanent address, live abroad but pay taxes here, etc). Also there may be substantial differences between what people report to the survey and what they report to the tax collection authorities. These facts would cause potentially significant estimation errors. Both approaches are therefore complements and it is necessary to combine them to answer different questions about a possible tax reform. It is surprising that the Czech government does not use any micro oriented study to support the suggested tax changes.

The excise duty on fuel plays an important role in the Czech tax system. The revenue from the excise tax on fuel products generates around 60–65 % of government's excise income and four other excise taxes (on wine, tobacco, beer and spirits) constitute just 35–40 % of total excise revenues. The taxation of fuel is also a taxation of transportation, since the majority of fuel is consumed in the form of petrol for cars. A change of this tax may therefore have significant impacts among others on the labour market flexibility, unemployment, environment issues and quality of life in rural areas.

In the centre of the following tax analysis is an investigation of the demand function. The economic theory offers several possible demand specifications. Here, we use the Stone's linear expenditure model and a single equation from the Deaton and Muellbauer's "Almost ideal demand model" enlarged by a quadratic income term. The nature of the problem permits to use two econometric estimation approaches, the instrumental variables that deal with an endogeneity issue, and the Heckman 2-step selection model, that deals with the selection issue connected to the ownership of a car, since a change of petrol taxation will only directly influence people who buy petrol products (because they have/use a car).

For the empirical part of the study we used data from the Czech Statistical Office (CSU) – from the *Czech Family Expenditure Surveys* 1996–2000. The other data was obtained from the Ministry of Finance web pages.

The contents of this paper is as follows: we briefly summarise the importance of excise taxes in the current tax system in the first part and give

TABLE 1

Commodity	Minimal tax rate (EUR)	Minimal tax rate (CZK)*
Leaded petrol	337	10 110
Unleaded petrol	287	10 110
Motor diesel	245	7 350
Light gas oil	18	7 350
Heavy gas oil	13	390
Liquefied petroleum gas	100	3 000
Compressed gas for transportation	100	3 000

Note: * Assuming for simplicity the exchange rate approximately 1 EUR = 30 CZK.
Source: Ministry of Finance of the CR

the basic intuition about the micro-data from the Family expenditure survey in the second section. We give a short overview of the demand analysis in the third part and present the empirical results of estimated models in the following section. A part that is concentrated on the simulation aspects follows. Particular attention is given to effects of a tax rate change on households with different incomes, education, age and households living in differently sized towns. The last section concludes the paper.

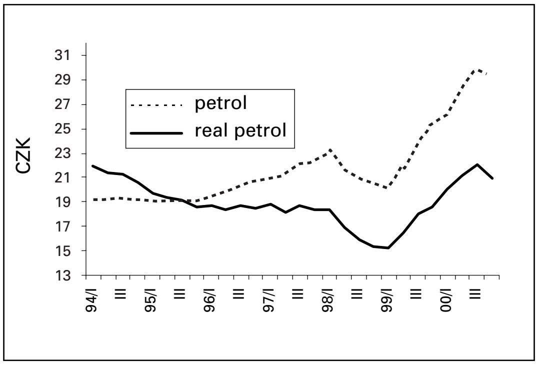
2. Excise Tax on Fuel

The importance of fuel products excise tax creates a natural interest to investigate this tax more closely. Fuel products are divided into 7 categories: unleaded petrol, leaded petrol, motor diesel, light gas oil, heavy gas oil, liquefied petroleum gas (used as a car fuel) and compressed gas (also used as a car fuel). These groups are similar to the EU classification and their taxation already fulfils the EU requirements for minimal tax rates. The current tax law imposes the same rates for leaded and unleaded petrol and also the same rates for diesel and light gas oil. This is mainly because these fuels are close substitutes and the temptation to interchange them because of tax evasion may be too high. The other two categories, the heavy gas oil and compressed gas for transportation, have a zero tax rate. This may reflect some social considerations, because lower income families might use this as a source of heating. The *Table 1* shows the minimal tax rates according to the EU requirements.

Fuel excise duty has an *in rem* form. Each physical unit is firstly taxed by the excise tax that corresponds to the physical volume. In the second step, it is taxed by the VAT, which is computed from the sale-price.

In the following analysis we concentrate only on the first three items from table 1 that are recorded in the Czech family expenditure survey together under the item “household expenditure for fuel, oil and similar preparation for personal transportation vehicles”. The other items from the table are registered in a different expenditure category, fuel used for heating purposes. These items are not a subject of this study. From this point hereinafter the term “fuel” or “petrol” will refer just to the fuel used for the transportation purposes.

FIGURE 1 Price of Petrol



Source: Czech Statistical Office

3. Data Description

3.1 Price of Fuel

Before we proceed to the demand analysis, it would be useful to mention some issues connected to the price of fuel. Because all three kinds of fuel are recorded in one category in the Czech Family Expenditure survey, it was necessary to construct an (weighted) average price that would appropriately reflect the consumption composition of different types of fuels. The replacement of old cars, which used leaded petrol, by cars, which use unleaded petrol or diesel, take place in the period covered by the available family expenditure data. This shift is fully reflected in a change of weights in the consumer price index (CPI). The CPI weights from 1994 were changed by the CSU in 2001 to reflect different structures of household consumption. The new weights of leaded petrol dropped significantly. In the opposite direction, the weight of unleaded petrol used by new cars increased to compensate at least a part of the leaded fuel consumption decline. Using linear interpolation we constructed hypothetical consumption weights for each type of fuel in every year in the sample period (years 1996–2000) using the official 1994 and 2001 CPI weights. However, since leaded and unleaded petrol are nowadays taxed by a same excise tax rate, different weights do not play such a crucial role.

Demand analysis is usually carried out on real values, since nominal values may be highly misleading in the presence of fluctuating inflation. We used the GDP deflator to convert the nominal price of fuel into the real price. *Figure 1* represents the development of the nominal and real prices per litre of the "composite" fuel (computed using the weight discussed above) expressed in the domestic currency. The dashed line depicts the nominal price of fuel, the continuous one the real price between years 1994 and 2000. A gradual decrease of the real price can be identified starting from the 1st quarter of 1994 until the 4th quarter of 1995, followed by a stagnation between the 1st quarter of 1996 and the 1st quarter of 1998, then a rapid drop during the whole of 1998 and finally a relatively fast increase beginning in the 1st quarter of 1999 and ending in the 3rd quarter of 2000. This price development has a significant consequence for a further demand analysis – it helps to identify the demand responses to different price changes. If the development of the price of petrol was monotonous, the price elasticity would be hardly distinguishable from a time trend and the estimates would be probably seriously biased.

Unfortunately, we could not obtain valid data for petrol prices in different regions. We could not use therefore the regional differences approach to estimate the demand function and was left to rely solely on price variation over time. According to the CSU, the reason why prices of petrol in individual regions are not tracked separately is the fact that the price differences of fuels is not significantly different between various regions. Generally, the regional differences could be also almost insignificant due to the relatively small size of the country. It seems that the fuel price is more dependent on an actual location of a particular petrol station – the price of petrol will differ more between a petrol station which is on the motorway as compared to one in the middle of a village.

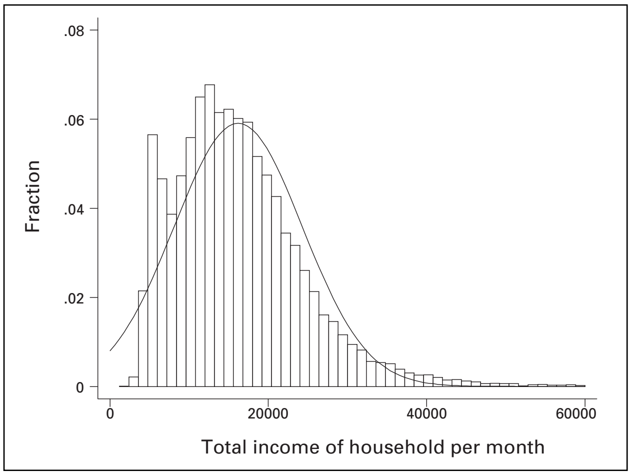
3.2 Family Expenditure Survey Data

The data for this analysis was taken from *Czech Family Expenditure Survey* that is conducted by the CSU. The whole sample includes around 3,000 households questioned in every quarter. We used for this analysis, mainly just because of the budget constrains, a random sub-sample with just 1500 household observations in each quarter in five consecutive years starting in 1996, which gave in total 30 000 various data points. Such a rich source of information allows apart from other things the application of modern econometrics methods such as non-parametric regressions. The data of each household/family includes information about: region and district where the family lives, size of the town, number of household members, number of its economically active members, number of dependent children, number of dependent children younger than 5 years of age, social group of the head of the household, his/her age and education, total monthly income, total monthly expenditures, monthly expenditures on fuel and a dummy variable for whether the family has a car or not. A list of the variables is included in the *Appendix*.

The family expenditure surveys are compiled in every country on slightly different principles and assumptions. The *Czech Family Expenditure Survey* is not a random survey as it is in many other countries. The CSU has a stable set of households, whose main characteristics approximately correspond to the distribution of the characteristics in the whole population.[1] The selected households are in a long-term relationship with the CSU and a kind of mutual confidence has developed over time, which helps to obtain

[1] There may be however deviations in same characteristics; i.e. in our sample the *social group* of the head of the household doesn't precisely correspond to the real distribution in the population.

FIGURE 2



more accurate (frank) data. In addition this approach allows the CSU to also get information from a group of people who would probably never respond to randomly sent questionnaires, i.e. households with extremely high incomes.

For having an intuition for the data, let's briefly describe their main features. Only one suspicious observation was dropped out from the total number of 30 000 observations, because of having a negative value of total expenditures. Approximately 57 % of the households in the sample have a car and 43 % do not. The median of total income in the sample was 15,145 CZK and the mean was 16,418 CZK (several households with extraordinary high total income caused the difference between these two statistics). The distribution of total income is shown in the histogram in *Figure 2* and it is not far from the normal distribution.[2] However, we have 85 observations of households with a total monthly income of higher than 60,000 CZK that are not plotted in this chart.

The dataset covers all regions of the Czech Republic. The number of respondents from each region however varies. The dataset is made up of the following regions (the frequencies of occurrence in the dataset are in the brackets): Prague (14,0 %), Central (11,2 %), South (7,0 %), West (8,7 %) and North Bohemia (11,4 %), South (18,7 %) and North Moravia (17,3 %). The head of the household, which can be used as one of the proxy variables indicating the social group of the whole family, is divided between these 5 categories: worker (27,5 %), self-employee (13,5 %), employee (26,2 %), farmer (12,2 %) and retired/unemployed (20,4 %). The distribution of the so-

[2] One can also notice the common phenomenon of expenditure surveys – households with lower (than average) incomes are, probably due to opportunity costs or because the respondents are slightly remunerated for their effort, more willing to take a part in a survey than the households with higher incomes. This fact is documented in the data histogram and the normal distribution – frequencies of occurrences of households below the average total income are slightly above the ideal normal distribution, the opposite is true for households above the average.

TABLE 2

Education	Frequency
Uncompleted primary school	0.0 %
Primary school completed	6.5 %
Primary school completed and secondary training institution	37.4 %
Apprenticeship secondary school without A-levels	5.7 %
Complete apprenticeship secondary school (with A-levels)	5.0 %
Completed grammar secondary school (with A-levels)	4.6 %
Completed practical secondary school or conservatory (with A-levels)	27.4 %
Undergraduate university education (bachelor level)	0.6 %
Graduate university education (master level)	8.2 %
Postgraduate, doctoral	4.5 %

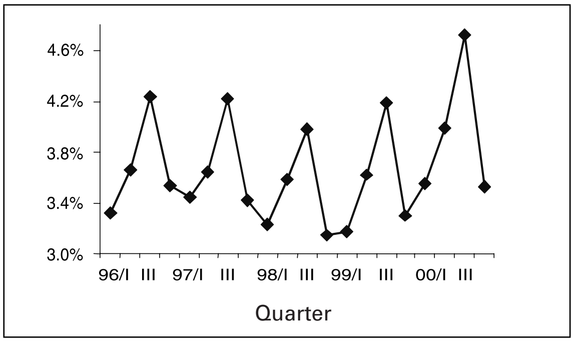
cial group in our sample does not correspond (due to the way how the data was collected) to the distribution of population. We cannot therefore construct any "average household" according to the social group.

A larger part of the sampled households comes from bigger towns or cities. Towns are divided, according to their size, into 9 groups (the figures in brackets are the corresponding frequencies of the households in the dataset): towns with less than 500 citizens (7,0 %), 500 to 999 (7,7 %), 1000 to 1999 (7,3 %), 2000 to 4999 (8,6 %), 5000 to 9999 (4,8 %), 10 000 to 19 999 (9,1 %), 20 000 to 49 999 (17,9 %), 50 000 to 99 999 (14,8 %) and cities with more than 100 000 people (22,8 %).

Frequencies of the head of household education in the dataset are depicted in the *Table 2*. A couple of interesting things can be read from this table. First, there are no uneducated people in the dataset (illiterate people may be self un-selected from the data collection process if they were not able to fill in a questionnaire). There is also a very small number of undergraduate people; (hoping that they did not face the same selection problem as the previously mentioned group) this is probably caused by the fact, that the undergraduate level of education was introduced just a couple of years ago and people who studied at a university before 1991 could study only master courses, because no undergraduate courses were available. A rather suspicious feature of the dataset is a high number of postgraduates compared to masters. This is more an interpretation problem: the previous university degrees were not fully compatible with the current segmentation. Some of students (e.g. in medicine, law, philosophy, theology) got the title "doctor" for a master-like level of studies. The last row of the table is therefore probably a mixture of real PhD's. (or CSc.) holders and previous "doctoral" masters. BTW even more peculiar things may happen in other countries – e.g. in Italy students are allowed to use the title doctor after obtaining a bachelor degree.

Among the other variables that portray a household in the dataset were: the number of household members, number of children, number of children younger than 5 years old, number of economically active members in a household and household type that describes composition of a household (if there are both parents or if it is a one parent family, etc.) All categorical variables were converted into sets of dummies, i.e. variables that are equal

FIGURE 3 Average Fuel Budget Share



pramen: ČSÚ

to 1 if a particular household e.g. belongs to a particular education group, town size population etc. and equal to 0 otherwise.

A closer look at the average budget fuel share (i.e. the average fuel expenditure over all households in each quarter divided by the total expenditures) in *Figure 3* reveals a strong seasonal pattern over the year. The first three quarters in every year have increasing fuel shares that peak in the third quarter. This is probably caused by the summer holidays where a lot of people use a car for travelling. In the forth quarter, there is a sharp drop – this may be a reflection of the fact that a lot of drivers will stop using their cars for safety reasons during winter (1st and 4th quarter). Overall fuel expenditures fluctuated between 3 % and 4.6 % of the total expenditures. There is no observable trend, but the visible periodicity is calling for inclusion of dummy variables for each quarter into the empirical model.

4. Demand Analysis

The relationship between the price and the quantity consumed is known as a demand function and it is extensively studied in microeconomics. We can apply this analysis to answer questions of how a specific tax might influence consumption of a particular good. The tax change from the consumers' point of view is nothing else than some alteration of the price. The law of demand says that an decrease in the price of a product should not lead to a decrease in quantity demanded[3], in other words the own price elasticities, which measure a proportionate rate of change of demanded quantity to a unit proportionate price change, should be non-positive.

There exists in the literature many studies that investigate a demand generally or just a demand for fuel, which suggest several possible specifications of the demand function. E. Engel pioneered work in this field in the nineteen century, and was followed by H. Working in 1940s and C. Leser in 1960s. Nonetheless their significant contribution to the demand analysis was not absolutely consistent with the utility maximization theory and another im-

[3] Let's assume that there are only normal and not Giffen goods that we are talking about.

provement had to be reached by the works of J. Muellbauer, A. Deaton and D. Jorgenson in 1970s and 1980s.

One approach how to tackling the problem is to estimate an Engel curve – which is the relationship between expenditures on a particular commodity and the total income or expenditures. In theory, Engel curves are demand equations in which prices are supposed to be constant. The Engle curves are also known as income expansion paths. When we think only about a "two commodity economy", we can identify three possible cases: in the first case the income expansion path is straight and is going through the origin[4], the demand curves have then unit income elasticity and consumers will buy the same proportion of each commodity at every level of his/her income. In the second case the curve is bends towards one of the axes – commodities depicted there are luxury goods, the other case is necessary goods. In the last case the curve bends backwards and after a certain point of income increase a consumer starts to consume less of one good. Such situations are known as inferior goods – these are goods for which more income results in less quantity demanded – see (Varian, 1992, p. 117).

The literature about specification form does not give a simple answer as to which form should be used. For more details about various functional forms see (Blundell, 1988, pp. 22–24) or (Brown – Deaton, 1972, pp. 1207–1211). Probably the most frequently used in the empirical studies were either the linear expenditure specification suggested by Stone or the almost ideal model suggested by Deaton and Muellbauer (1980). The next section gives a brief summary of different demand model specifications that are often employed in empirical studies.

4.1 Brief Overview of Standard Models

The first modelling effort concentrated on individual demand equations, but the testing of these models whether they correspond to the economic theory was rather limited. This created an impulse to develop demand systems – systems of simultaneous equations that describe demands for several commodities. These models are more convenient for testing the theory,[5] because they include more testable parameters. The functional specification of a demand function is not important just for being able to test the theory, but it is also crucial for tax simulations, estimation and optimal tax reform design and analysis of its sensitivity – see e.g. (Madden, 1996).

The Working-Leser model. It is one of the most frequently used demand specifications (also because it is a starting point for the almost ideal model). It relates the value shares w_i to the natural logarithm of total expenditures x:

$$w_i = \alpha_i + \beta_i \log x$$

where both parameters α and β could be functions of prices made in many

[4] of the graph where there are depicted on axes quantities of goods consumed

[5] By testing the theory we mean testing of the demand function properties: adding up, homogeneity, symmetry and negativity – for details see (Deaton – Muellbauer, 1980, p. 43).

different ways. If β_i is greater than 0, the commodity i is a luxury; if it is lower than 0, it is a necessity. Usual assumption is that the sum of α_i over i is equal to 1 and the sum of β_i is equal to 0.

The Stone's model. The model from 1954, as presented by Deaton and Muellbauer (1980, pp. 61–63), starts with the logarithmic demand function:

$$\ln q_i = \alpha_i + e_i . \ln x + \sum_{k=1}^{n} e_{ik} \ln p_k$$

where e_i is the total expenditure elasticity and e_{ik} is the cross-price elasticity of k^{th} price on the i^{th} demand. Using the Slutsky equation and assuming a homogeneity restriction ($\sum_k e^*_{ik} = 0$) this function can be adjusted to the form:

$$\ln q_i = \alpha_i + e_i \cdot \ln \left(\frac{x}{P}\right) + \sum_{k \in K} e^*_{ik} \ln \left(\frac{p_k}{P}\right)$$

where P is a general price index and K is the set of "close" complements and substitutes. The equation above is – often appended with a time trend term – a point of departure for most of Stone's consumption analysis.

The Linear expenditure model. The famous linear expenditure model is also connected to Stone's work. He imposed adding up, homogeneity and symmetry restrictions to get a linear expenditure system:

$$p_i q_i = p_i \gamma_i + \beta_i \left(x - \sum p_k \gamma_k\right)$$

where $\Sigma \beta_k = 1$. Parameters γ_i are sometimes interpreted as minimal quantities required (or subsistence quantities). The first term on the right side denotes the expenditures, which are made first. The second term describes the residual after realisation of those first-round expenditures, which is then allocated between commodities in the fixed proportion β_i.

This demand form was derived from the utility function: $u(q) = \sum_i \beta_i \ln(q_i - \gamma_i)$ and hence fully satisfies the theoretical requirements. Its popularity comes from the easy way of obtaining own- and cross-price responses.

The Almost ideal model. This model, developed by Deaton and Muellbauer (1980), can be written in the budget share form as:

$$w_i = \alpha_i + \sum_j \gamma_{ij} \ln p_j + \beta_i \ln (x/P)$$

where w_i is the budget share, P is the price index, x is total expenditure and p_j is price of the goods j. Common restrictions suggest that the sum of α_i over all i should be 1, the sum of γ_{ij} over all i should be 0, and also the sum of β_i over all i should be 0. The symmetry also implies that $\gamma_{ij} = \gamma_{ji}$. Working with a complete set of goods will result in the sum of w_i equal to 1. Changes in real expenditures affect the budget share through β_i parameters, which are positive for luxuries and negative for necessities. Changes in relative prices are captured by the γ_{ij} coefficients. Deaton and Muellbauer (1980, p. 314) defined the price index P as:

$$\ln P = \alpha_0 + \sum_k \alpha_k \ln p_k + \frac{1}{2} \sum_j \sum_k \gamma_{kj} \ln p_k \ln p_j$$

The cost function corresponding to the almost ideal demand system can be expressed as:[6]

$$\ln c(u, p) = \alpha_0 + \sum_k \alpha_k \ln p_k + \frac{1}{2} \sum_k \sum_j \gamma_{kj} \ln p_k p_j + u\beta_0 \prod_k p_k^{\beta_k}$$

and satisfies by multiplying both sides by $p_i / c(u, p)$:

$$\frac{\partial \ln c(u, p)}{\partial \ln p_i} = \frac{p_i q_i}{c(u, p)} = w_i$$

4.2 Augmentation of the Almost Ideal and Linear Expenditure Models

We applied in the empirical part of this study the following two specifications that differ in some ways from the theoretical models presented above. The first model is the specification of the almost ideal demand model augmented by household demographic characteristics $\overline{z}$ and the second model is the log linear specification with added demographic variables.

Generally, we can estimate – using micro data for households – a regression model like:

$$w_h = \alpha_1 + \beta_1 \cdot \ln x_h + \gamma_1 \cdot \ln p_v + \delta_1 \cdot \overline{z}_h + \varepsilon_{1h} \tag{1}$$

where w_h is the share of fuel expenditure in the total expenditures of a household, x_h is a household's total expenditure (or income), p_v is the price of fuel, z_h is a vector of demographic and other household characteristics and ε_{1h} is an error term. Another model we can estimate is:

$$\ln v_h = \alpha_2 + \beta_2 \cdot \ln x_h + \gamma_2 \cdot \ln p_v + \delta_2 \cdot \overline{z}_h + \varepsilon_{2h} \tag{2}$$

where $\ln v_h$ is the natural logarithm of expenditure on a fuel and the variables on the right side have the same meaning as in (1). The only difference is the dependent variable – in the first case we work with the expenditure share, in the second case with the logarithm of expenditures. These two specifications include also a price term, because the dataset used for this estimation is a cross-section repeated over several years. Therefore, we cannot rely on the theoretical assumption of the Engel curves that the prices are constant as would be appropriate in a cross-section sample taken at one moment in time.

The equation (1) is connected to the work of Deaton and Muellbauer (1980) and could form a system (in the case of more commodities) called the "almost ideal demand system" (AIDS). Its advantage is that it was derived from the utility maximization and can be used under certain circumstances for testing homogeneity and symmetry – see (Phlips, 1983, p. 136).

The second model – the constant elasticity or logarithmic curve – is more restrictive, since we are not able to calculate a logarithm of zero expendi-

[6] the demand function can be derived using the Shephard's lemma directly from the cost function since: $\partial c(u,p) / \partial p_i = q_i$.

tures on fuel. Therefore only households with a positive consumption of fuel will be captured. We must also keep in mind that the interpretation of the estimated parameters differs for both specifications. In model (2) we calculate just one overall price elasticity γ_2, but in model (1) the individual household uncompensated price elasticity depends on budget share and can be evaluated using:

$$e = \frac{\gamma_1}{w_h} - 1 \tag{3}$$

4.3 Linear or Quadratic Engel Curves?

Is the previously mentioned linear specification of the Engel curve sufficient? Some studies – see e.g. (Banks et. al., 1997) – suggest to include a quadratic income term into the Engel curve specification to better display the true underlying relationship, which need not be necessarily linear. They revealed, based on the *U.K. Family Expenditure Survey* analysis, that some groups of goods – in particular expenditures on alcohol and closing – display distinct non-linear behaviour. It is therefore legitimate to investigate if such a non-linearity exists also in the demand for the fuel in the Czech Republic.

We can use nonparametric regressions as a diagnostic tool, which might help to decide whether including a quadratic term into the specification is appropriate or not. Nonparametric methods do not relay on a particular function form, and allow us to also capture non-linear relationships between variables.

We used a kernel regression with a Gaussian kernel and different bandwidths to investigate the legitimacy of inclusion the quadratic term into the Engel curve linear specification. We estimate in the kernel regression instead of eq. (1) the budget share as:

$$w_h = \alpha + \beta \ln p_v + m(x_h) + \varepsilon_h \tag{1a}$$

where the $m(.)$ is any – not necessarily linear – function of household expenditures x_h.

First, we estimated using ordinary least squares (OLS) method a regression similar to eq. (1), but without the expenditure term. Secondly, we regressed using kernel regression the residuals from the previous OLS on the logarithm of real total expenditures. Probably the most suitable result (displayed on *Figure 4*) was produced by the bandwidth 0.25. This result indicates non-linear behaviour and justifies the usage of the quadratic income term in the Engle curve specification.

Applying the Nadaraya-Watson kernel estimator in Stata on (1a) results in the function that is drawn in *Figure 5*. The observations here were restricted only to car owners; non-owners were excluded. This graph with the logarithm of real expenditures on x-axis and the share of fuel in a household budget on the y-axis supports the idea that the Engel curve in this case is very likely not linear.

We can defend, based on the previous kernel regression results, the en-

FIGURE 4

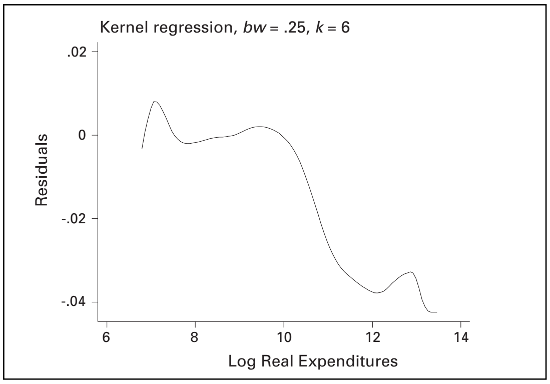
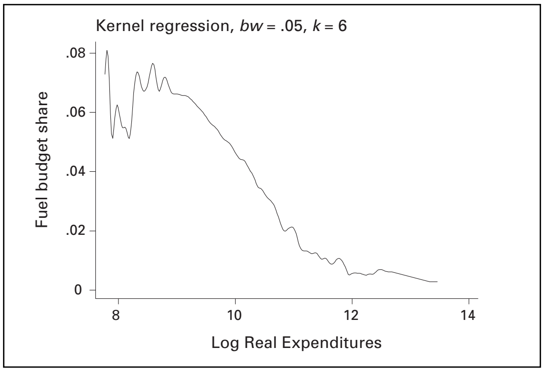


FIGURE 5



largement of a demand curve by a quadratic income term inclusion. For that reason, our linear specification will instead of eq. (1) be:

$$w_h = \alpha_1 + \beta_1 \cdot \ln x_h + \lambda_1 \cdot (\ln x_h)^2 + \gamma_1 \cdot \ln p_v + \delta_1 \, \overline{z}_h + \varepsilon_{1h} \tag{4}$$

where λ_1 represents the quadratic income parameter. Similar adjustment of the log linear model can be made to eq. (2) resulting in this quadratic Engle curve specification:

$$\ln v_h = \alpha_2 + \beta_2 \cdot \ln x_h + \lambda_2 \cdot (\ln x_h)^2 + \gamma_2 \cdot \ln p_v + \delta_2 \cdot \overline{z}_h + \varepsilon_{2h} \tag{5}$$

4.4 Empirical Model

The OLS method may be used for estimation of the models (4) and (5) using some cross-section data. Nonetheless some researchers have noticed several potential drawbacks of this approach – see e.g. (Blow – Crawford, 1997, p. 35). Two main problems are connected to endogeneity and selectivity. The selectivity comes from the fact that we observe fuel expenditure only from households that have at least one car. We do not observe the fuel expenditures for those who do not have a car, so we see just part of the overall distribution. The endogeneity problem may be connected to the unobserved tastes for the usage of a car. This usage may be correlated with the cost per kilometre. We observe mainly the fuel expenditures of households that have higher preferences for car usage than households that are not car-owners. These groups are very likely not similar and this may result in a bias in the regression.

The *endogeneity* problem, meaning in fact that there is a correlation between regressors and the error term, may be removed by using the instrumental variables (IV) approach instead of the OLS. The endogeneity occurs when the expectation of the error term in a regression function conditional on explanatory variables does not equal to zero ($E(\varepsilon | X = x) <> 0$). If this is the case, applying the OLS method leads toward biased estimates of regression parameters. To overcome this problem, we try by using the IV approach to find variables[7] with observed realisations Z – called *instrumental variables* – that have the property: $E(\varepsilon | Z = z) = 0$.

When applying instrumental variables we cannot use for model selection purposes one of standard diagnostic tools – the R^2 coefficient, which measures goodness of fit. In instrumental variable regressions it often happens that the R^2 is not positive. Pesaran and Smith (1994) suggested a generalization of R^2 that is suitable also for the IV approach. Their "*generalized* R^2" provides an asymptotically valid model selection criterion appropriate for making a choice between different nested or non-nested IV models. The idea of GR^2 is based on a two step OLS regression. In the first step, the models are estimated by OLS, where the dependent variables are the instruments (from the IV model) and the dependent variables are the same as in the original IV model. One can calculate using these estimates the fitted values of dependent variables. The second step consists of regressing (also by OLS) the original (IV) dependent variable on those fitted values of instruments and the other explanatory variables from the IV specification. The R^2 (called GR^2) from this estimate could be used as a criterion for model selection in the IV context.

The *selectivity* issue may be overcome by application of the Heckman 2-stage selectivity model (Heckman, 1979). This approach separates the car ownership question and Engel curve modeling. First, a household decides whether it will own a car and then in the second step, how much fuel they will consume driving their car. The ownership of a car may be dependent on certain demographic and social characteristics of the family, e.g. social status, distance to work, size of town where the household lives, number of adult household members and their economic activity. Generally, what we want to estimate is the *population* regression function:

$$E(Y_{1i} | X_{1i}) = X_{1i}\beta_1 \tag{6}$$

[7] The capital letter Z was used just because it is a convention in the literature that describes the IV approach and does not have any link to the previously used letter z when we were talking about demographic characteristics.

where i denotes an individual household. But unfortunately due to the selection problem we observe just a part of the Y_1 distribution, which is included in the sample:

$$E(Y_{1i} | X_{1i}, \textit{sample selection rule}) = X_{1i}\beta_1 + E(U_{1i} | \textit{sample selection rule}) \quad (7)$$

Equation (7) may equal to (6) only if the conditional expectation of the error term U_1 is zero (this may happen when observations Y_{1i} are missing randomly). James Heckman (1979, pp. 156–157) suggests a procedure, of how to cope with such problems. He calculates the conditional expectation of the error term in (7) using the inverse of the Mill's ratio, which is the monotone function of the probability that an observation is selected into the sample.

Except those two problematic issues, there is also a question regarding the data *aggregation*. It is surely a much smaller problem in cross-section micro-data than in the aggregate time series, but even this relatively small aggregation from individuals to households may cause trouble, because the microeconomic theory models the behaviour of an individual, but *Family Expenditure Surveys* use a household (and not an individual) as the basic unit. Another type of aggregation problem is incorporated in the broad categories in which commodities are recorded – the item "fuel" is a composite basket of several commodities. These deviations from the common microeconomic assumptions may create some difficulties in an estimation of a demand system; fortunately in an estimation of one single demand function, they should not play a significant role.

5. Empirical Results

Having absorbed the theory that underlies the demand analysis and some connected econometrics issues, we may advance to the empirical part, where we restricted our focus only on the two estimation techniques: the instrumental variables approach and Heckman's two-step selection model. These methods are often applied to overcome estimation problems – namely endogeneity and selectivity – mentioned in the previous section. During the search for the best empirical specification, we have estimated model (4) and (5) using instrumental variables (we instrumented total expenditures by total income) with different sets of explanatory variables.

5.1 Instrumental Variables Results

We have investigated several possible regression models – each with a different set of explanatory variables. Trying to keep the presentation of results tractable, we chose 3 different models (see *Table 3*): *model 1* includes only explanatory variables that were significantly different to zero (some of the dummies were removed) and insignificant variables whose exclusion would substantially worsen the model. *Model 2* is a minimalist version, because it excludes the majority of demographic dummies. *Model* 3 is the op-

TABLE 3 Almost Ideal Model [equation (4)]

Dependent variable	Fuel share		
	Model 1	Model 2 (mini)	Model 3 (maxi)
Log (fuel price)	0.0291 (0.0029)	0.0301 (0.0029)	0.0293 (0.0029)
Log (total expenditures)	0.0738 (0.0253)	0.0359 (0.0203)	0.0757 (0.0254)
Log^2 (total expenditures)	–0.0046 (0.0013)	–0.0027 (0.0011)	–0.0047 (0.0013)
Children	–0.0047 (0.0004)	–0.0049 (0.0004)	–0.0050 (0.0004)
Children < 5 years old	–	0.0017 (0.0007)	0.001 (0.0007)
Age	–0.0005 (0.00003)	–0.0005 (0.00003)	–0.0005 (0.00004)
Seasonal dummies	✓	✓	✓
Town size dummies	✓	–	✓
Region dummies	✓	–	✓
Education dummies	✓	–	✓
Social group dummies	✓	–	✓
Household type dummies	✓	–	✓
GR^2	10.4 %	7.3 %	10.5 %

Note: Price and expenditure explanatory variables are in real terms; corresponding standard errors are in brackets.

posite – all possible demographic characteristics were included, even when they were not statistically different to zero. The difference between the first and the third model is rather small. The third one includes slightly more explanatory variables (more precisely, there are all explanatory variables in the 3rd model specification, but some of the statistically insignificant dummies are after stepwise re-estimations eliminated in the 1st specification). Each of these three models produces similar estimates of parameters of our interest. The parameter signs generally fulfil our expectations – e.g. one may expect, that the fact that there is a child in a family will result in a lower proportion of fuel share; the family expenditures will be oriented very likely more toward this child. On the other hand, if there is a young child (younger than 5 years) the petrol expenditures may be slightly higher, because of the necessary of transportation. Increasing the age of the head of household seems to be negatively related to the fuel share – older people use cars less often than young families (or may have higher incomes keeping their fuel consumption unchanged). Surprising was the conclusion, that the variable characterizing the number of household members was in the IV approach not statistically significantly different from zero. It suggests that there may be no connection between the size of a household and fuel expenditure.

The empirical evidence shows that the quadratic expenditure term is significantly different from zero at the 10% level. Our initial assumption that the relationship in the Engel curve may not be strictly linear seems to be therefore supported.

The same parameters signs were obtained in the estimates of a model given by (5), which has a different dependent variable. The results are depicted in the *Table 4*.

The equation (5) results outperform the results of equation (4) in terms of the *generalized* R^2 that are reported in the last rows of both previous tables. However, a model given by the equation (4) is more serviceable for a de-

TABLE 4 Log-linear Specification [equation(5)]

Dependent variable	Log (fuel expenditures)		
	Model 1	Model 2 (mini)	Model 3 (maxi)
Log (fuel price)	0.5337 (0.052)	0.5502 (0.0520)	0.5375 (0.0521)
Log (total expenditures)	3.8072 (0.3893)	3.221 (0.3618)	3.9797 (0.4531)
Log^2 (total expenditures)	–0.1644 (0.0201)	–0.1354 (0.0188)	–0.1728 (0.0231)
Children	–0.0673 (0.0064)	–0.0651 (0.0064)	–0.0718 (0.0068)
Children < 5 years old	–	0.0185 (0.0123)	0.0117 (0.0131)
Age	–0.0078 (0.0005)	–0.0082 (0.0005)	–0.0083 (0.0006)
Seasonal dummies	✓	✓	✓
Town size dummies	✓	–	✓
Region dummies	✓	–	✓
Education dummies	✓	–	✓
Social group dummies	✓	–	✓
Household type dummies	✓	–	✓
GR^2	17.8 %	15.1 %	17.9 %

Note: Price and expenditure explanatory variables are in real terms; corresponding standard errors are in brackets.

scription of the price elasticity differences among various households, because it allows us to compute – using the relationship (3) – the individual price elasticity for every household in the dataset.

The estimated parameters of dummy variables cannot be compared among these three models, simply because we use in each specification a different set of dummies. The third (maxi) specification that fully represents the demographic characteristics of a family may be the most relevant according to the dummy parameters. It is appropriate to mention for completeness that we have to omit one dummy for each category to avoid a multicolinearity between regressors, e.g. for the town size (*townX*) the smallest settlements that have up to 499 inhabitants (*town* = 1), for education (*edX*) people who finished their basic education[8] (*edu* = 1) etc. were omitted. A parameter of a particular dummy (e.g. *town*2) then describes how much people classified in this category (e.g. *town* = 2) differ from people who live in the first category (in *town* = 1).

5.2 The Heckman Two-step Selection Model Results

The estimates in the previous section are based only on households that have at least one car. The Heckman two-step selection model is designed to solve this selection problem. Some analysis, e.g. Wheaton (1982), suggest that income elasticity of petrol consumption is more elastic than had been previously thought. This comes partly from the income effect of car ownership. Wheaton (1982, p. 440) states that the econometric results are in some sense ambiguous.

The Heckman 2-step selection model estimates two equations. The first one is trying to explain the fuel expenditure or the budget share according

[8] There were no observations for people with *edu* = 0 (those who didn't complete the basic school education).

TABLE 5

Dependent variable:	Fuel share log(fuel expenditures)	
	Model 1	Model 2
Log (fuel price)	0.0348 (0.0031)	0.6017 (0.054)
Log (total expenditures)	0.0043 (0.0145)	3.8167 (0.2595)
Log^2 (total expenditures)	–0.0015 (0.0007)	–0.1747 (0.0129)
Children	–0.0066 (0.0014)	–0.1149 (0.0245)
Children < 5 years old	0.0014 (0.0008)	0.0101 (0.0132)
Members	0.0021 (0.0014)	0.0551 (0.0246)
Trend	✓	✓
Seasonal dummies	✓	✓
Town size dummies	✓	✓
Region dummies	✓	✓
Education dummies	✓	✓
Social group dummies	✓	✓
Household type dummies	✓	✓

Note: Price and expenditure explanatory variables are in real terms; standard errors are in brackets.

to a chosen demand specification. The second – the selection equation – is modelling the probability of car ownership.

The following *Table 5* summarizes the results of the Heckman two-step selection model.

The left column in Table 5 corresponds to the equation (4), the right to the equation (5). I had included/excluded different dummy variables into/from both equations in the Heckman's model (the main and the selection one for a car ownership) and the price elasticity displayed quite a stable pattern – its value was usually between 0.034 and 0.036 for the almost ideal (fuel share) specification and 0.59 and 0.62 for the logarithmic specification. The number of household members seems to have in both these models statistical significance at the 10% level, which was not the case in the IV regression. The Wald test shows that the logarithm of total expenditure in the almost ideal specification is not statistically significant, how-ever the joint test with the quadratic term indicates that these parameters are significant. These parameters are in the logarithmic specification both separately and jointly significant.

The Heckman's two-step model requires finding a variable or variables that effect whether a household has a car or not and at the same time does not effect how much one drives. These variables are then applied to the selection equation that specifies the factors of car ownership. One can think of e.g. the price of cars, lagged price of petrol or a variables that combine different demographic characteristics. In the specification equation for cars we used as additional explanatory variables for the car ownership: the real prices of fuel lagged by 1 quarter, 1 year and 2 years, combination of the education and size of town and combination of the number of working members and size of the town (assuming that people from bigger towns, achieved higher levels of education and households with more working people are more likely to own a car). All of these variables seem to be statistically significant in both demand specifications.

5.3 Distribution of Elasticities

One interesting question, we may ask, is how price elasticities differ between different groups of people. This tells us something about the effect of possible tax changes and its impacts on different groups of population. Some groups of people will increase their consumption after a fuel price change, which may be induced by a change in taxation or by other reasons, while others will consume less. The effect of price variation on different groups will probably be different. Using the micro-data allows us to learn something about the behaviour of certain groups and to estimate their demand reactions in advance before a price change occurs.

We may investigate the connection between the age of the household head, the size of the town where the household lives and the price elasticity. The majority of people have an elasticity of lower than 1, so an increase of petrol price would lower their budget share of petrol expenditures. The data show that the people in bigger towns are not changing their budget share so much as compared to people in smaller municipalities.

6. Tax Simulation

After estimating the demand equations we are able to answer questions about the effect of tax changes on fuel products and their distributional impact on the population. Depending on people/household characteristics, different groups will be affected by the taxation change in different ways. Some of them may respond to the tax change by decreasing their consumption of petrol by driving less, some of them, who could not find an easy substitute for driving their car, will continue to consume the same amount as before the tax change, but will pay more (assuming that the tax was increased). A tax change can be carried out in two forms, in the first one the price of the commodity will be altered by a change of indirect taxation, in the second one the direct taxes could modify the household's budget constrain. Depending on income, we sometimes speak also about progressive or regressive taxes. A progressive tax is a tax that takes an increased proportion of an income as the income increases, for a regressive tax the opposite applies. Taxes that keep the same proportion of income are called proportional – for more details see e.g. (James – Nobes, 1996, pp. 14–15). The differentiation in a context of indirect taxes is not as straightforward as in direct taxes, since a particular commodity is usually taxed by a tax rate that is not directly dependent on the buyers' income. Nevertheless, if a commodity that tends to be purchased by people with higher income is taxed by a higher tax rate, the indirect tax may be progressive.

Using aggregate data might be helpful when one is interested in the predictions of tax revenue after a tax rate change. To answer questions about behavioural responses of a population one needs to carry out the simulation on the micro-data. A reason, why a government might want to increase taxes, could be an intention to obtain extra revenues or to change the behaviour of citizens. In the case of the excise tax on fuel products, this second reason may represent an effort to change driving behaviour and could be

TABLE 6

Period	Nominal price of petrol after taxation	Nominal price of petrol before tax	Deflator GDP	Real price of petrol before tax
2000q1	26.33	10.90	1.31467	8.29
2000q2	28.42	12.62	1.34121	9.41
2000q3	29.90	13.83	1.35442	10.21
2000q4	29.29	13.33	1.39844	9.53

TABLE 7

	0 %	25 %	50 %
Average excise tax rate on fuel	10,674	13,342	16,010
Tax rate change	0	2,668	5,337
Average price after taxation	28.5	31.7	35.0

possibly justified by environmental reasons. From the micro-data analysis we might be able to answer questions such as who is going to pay the tax and who is more likely to change their behaviour and substitute the taxed commodity for something else.

Before we give some answers about the response of consumers to the tax change, we have to decompose the observed price (after taxation) into the price before taxation and both taxes that are imposed on fuel – the value added tax (VAT) and the excise tax. The observed price may be expressed as the price before taxation increased by the excise and then multiplied by the VAT tax rate:

$$\textit{price after tax} = (\textit{price before tax} + \textit{excise on 1 unit}) \cdot \textit{VAT taxation} \quad (8)$$

Knowing the VAT (22 %) rate and the excise tax rate, we can easily compute by eq. (8) using observed prices the price before taxation. The same steps in the reversed direction starting with the previously computed price before taxation allows us to know the price after taxation under different tax rates. Different fuels are taxed with different excise taxes (see Table 1). It was necessary for the following analysis to compute (using weights from the consumer price index) an average excise tax on fuel.[9] Because the demand equations estimated in the previous section were formulated in real terms, we need to adjust the nominal values for the development of the price level. *Table 6* gives us the real price of petrol before taxes in 4 quarters of year 2000.

Let's assume that the VAT rate is kept constant and consider changes in only the excise tax. We might assume for example that the fuel excise tax is increased by 25 % or 50 % as indicated in *Table 7*. The first column shows

[9] The majority of fuel consumption (94 %), according to the 2001 CPI weights, is made in categories of leaded and unleaded petrol. These fuels are presently taxed by same tax rates. Household consumption of diesel in comparison with it is still very low and reaches just approx. 6 %. These numbers are hard to verify using overall consumptions of fuels, because the overall consumption of fuel consists also of business consumption, where usage of diesel engine cars is much higher.

no change in the excise taxation; the others depict how the tax rate would look like after the increase (measured in CZK/1000 litres). The second row shows the (absolute) difference between the new increased tax rate and the current tax rate (in CZK). The last row shows the average nominal price of fuel for different excise tax rates. We have to keep in mind that any change in an excise tax rate is further magnified by a VAT rate multiplication.

We chose for the simulation of the tax impact on different groups of population the third (maxi) model presented in the IV approach (see Table 3) with the *fuel share* as the depending variable and the second Heckman 2-step selection model from Table 5 with the log (*fuel expenditures*) as the dependent variable. The selection of the appropriate model was intuitive rather than theoretically based – a lack of reliable diagnostic tools (compared to the time-series econometrics where we have several model selection criteria as Akaike's or Schwarz's ones) is a common feature of microeconometrics. Having estimated these models we can easily using the price of fuel under different excise tax rates[10] (from Table 7) obtain predictions of the effect of taxation alteration on specific demographic groups.

6.1 Distributional Impact of Fuel Excise Tax

We may be interested which income groups will be affected the most, which education or age groups will bear the tax burden or how the tax encumbrance is distributed according to the size of the town or where a household lives. These questions are closely related to the normative side of taxation and also to a political economy dimension of any tax reform. The next section presents some of the main findings. Its purpose is more to demonstrate a way in which we can conduct the analysis, rather than drawing precise conclusions for policy makers. We should be aware that the econometric estimation was done with a dataset of only 1,500 households (although the resulting total number of observations was close to 30,000). A larger dataset could result in more robust estimates. It is also possible to dispute the way in which the nominal prices were transformed into real values – maybe using the CPI instead of GDP deflator would be more appropriate.

6.2 The Effect on Different Income Groups

Among the most frequently asked questions are queries about the response of different income groups to indirect taxation changes. We divided the households for the sake of answering this question into deciles according to their monthly total income. Just to remind the basics statistics, the people in the first decile represent the group of 10 % households with the lowest total income, etc. *Table 8* shows a household average monthly income in the year 2000 for each decile of the income distribution (measured in CZK per month). Its first row represents only the households with

[10] after converting them into real terms by dividing with an appropriate GDP deflator

TABLE 8

Percentile	10	20	30	40	50	60	70	80	90	100
Average total income for car-owners	9,592	12,921	15,106	17,239	19,240	21,114	23,224	25,749	29,890	46,053
Average total income for all	5,954	8,833	11,626	13,376	15,492	17,720	20,063	22,635	26,300	39,502

FIGURE 6 Average Change of Quality Consumed after Duty Rate Increase

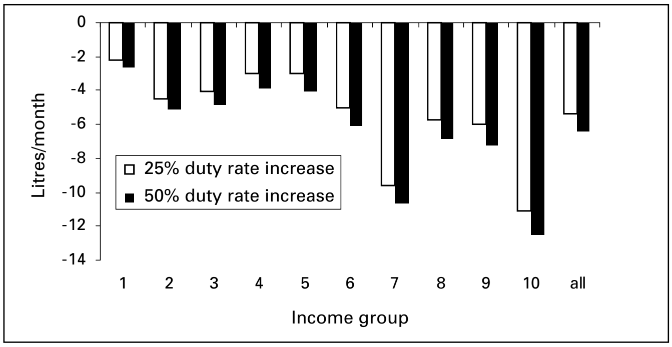
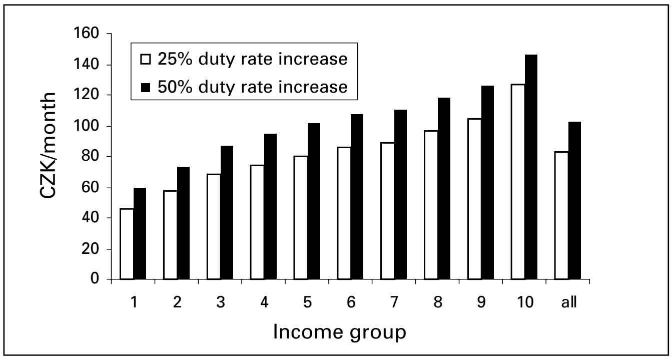


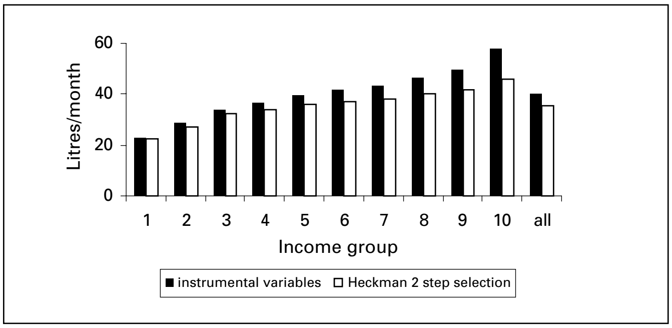
FIGURE 7 Increase of Average Duty Tax Paid per Month for Different Income Groups



a car, the second row both types of households – with and without a car. The average income of car owners in each decile is higher than the average of the all households in the corresponding deciles. It may support our intuition that car ownership is very likely related to family income.

Figures 6 and 7 depict the reaction of different income groups to a fuel excise tax rate increase of 25 % and 50 % – the difference between both figures is in the measurement unit. Figure 6 verifies the usual assumption that after a price increase, the demanded quantity should decline. The size of bars on the graph reflects the difference between the average change of quantity demanded each month after a new excise tax rate is introduced and before such change. The decline of quantity is indeed higher with

FIGURE 8 Quantities Consumed after 25% Tax Rate Increase



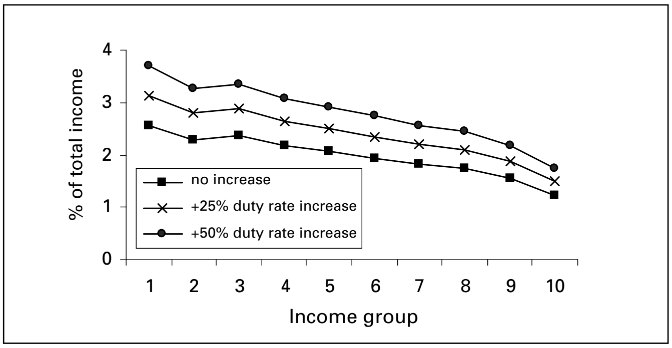
the 50% increase of the excise tax rate than with the 25 % increase; higher income households tend to react to such increases more rapidly than lower income families.

Figure 7 expresses the response to an excise rate change, but instead of measuring it in the physical units as in Figure 6, it is represented by the average tax paid per month over all families who belong to a certain decile of income distribution. This number was calculated using the quantities before and after the tax change multiplied by the corresponding tax rates (taken from Table 7). It seems that despite the drop in quantities of petrol consumed, the tax revenue due to higher excise tax rate increases as indicated in Figure 7. The groups who are contributing the most to the tax revenue are people with higher incomes.

Patterns that are visible on previous charts were also visible in both models outputs: the instrumental variables and the Heckman selection model. Comparing the IV results and the Heckman 2-step selection results shows that the effects of a tax increase were generally higher in the instrumental variables estimates than in the Heckman selection model (see Figure 8). The reason may be the selectivity issue; if this is the case the IV method very probably overestimates the true underlying parameters. Figure 8 depicts the average quantity of petrol that each income group is likely to consume per month after a 2% tax increase. The bars on the left correspond to the estimates obtained by instrumental variables, the right bars to the Heckman 2-step selection procedure. Although the absolute numbers in every category differ, the overall trend in both methods is analogous.

What is the economic interpretation of *Figures 8* and *9*? Based on Figure 8 we can confirm our expectations, that the consumption of fuel is positively related to income, in microeconomic words: fuel is a normal good; its demand rises with increasing income. Figure 7 tells the same story, since the excise tax rate is the same for every income group. If fuel consumption is a positive function of the total income, the households who earn more are paying a higher tax in absolute terms. If the government introduces a tax rise, the higher income households will pay higher taxes (in absolute terms) than lower income households. Nonetheless this does not hold in relative terms. We can clarify this from Figure 9. This graph shows the average tax paid by different income groups under 3 possible scenarios: no policy change

FIGURE 9 Duty Tax on Fuel as Percentage of Total Income



(the bottom line), 25% (the middle line) and 50% (the upper line) increase of the fuel excise rate. Relative to income, the fuel excise tax is higher for lower income households. Shifting the tax rate above has a bigger impact on these groups than on the higher earning families. The vertical distance between the no policy change curve and the 25% (or 50%) excise tax rate increase line is bigger for lower income families than for higher income families. Using indirect taxes as a tool for wealth redistribution is therefore rather controversial. It is a bit surprising that some left-wing oriented governments are trying to increase the indirect taxation to balance public finances – it will hurt their potential voters (assuming that the majority of their voters comes from that part of the income distribution).

6.3 The Effect on Different Age Groups

We can also try to answer questions about the possible effects of a change in the excise tax rate on different age groups. Our dataset includes only information about the age of the head of the household, but we might assume that it will be somehow correlated to the age of other household members.[11] Another assumption we may wish to make is that the head of the household will probably be also a car driver or he/she will have significant influence on the frequency in which a car is used. For simplification of the results presentation we divided the households into 6 age groups (with cut-points: 29, 39, 49, 59, and 69 years). We computed, using again the Heckman selection model, the average change of fuel quantity after an increase of the duty rate consumed by different age groups. The results are presented in *Figure 10*, which is methodologically similar to Figure 6 presented in the previous section.

Again, we can see on this graph that a larger increase of duty rate tax will result in a larger drop in consumption. The groups who react the most intensely to the change in taxation are families where the head of the house-

[11] In other words we take the age of the household head as a proxy for the age of the whole household.

FIGURE 10 Average Change of Quantity Consumed after Duty Rate Increase

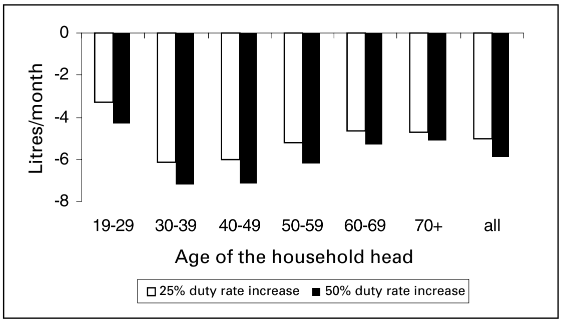
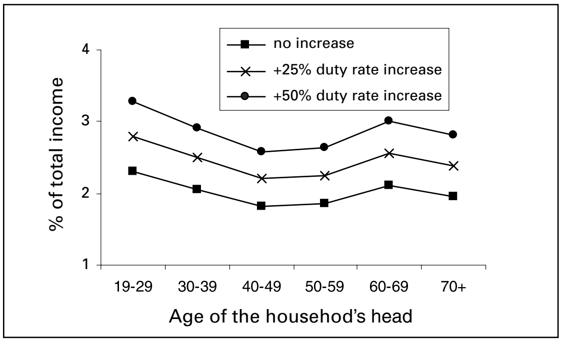


FIGURE 11 Duty Tax on Fuel as Percentage of Total Income



hold is aged between 30–49 years. Young people, aged up to 29 years tend to keep their consumption pattern unchanged; the seniors seem to adjust their habits fairly close to the overall average.

We may also look at the distribution of the proportion of fuel excise tax paid on total income across households with different ages (assuming again that the age of the head of the household is significantly correlated to the age of the other household members). *Figure 11* exhibits *U*-shape functions for 3 different excise rate levels. Young and old people are spending more of their income on fuel than the middle-aged households. A decline for households with the head older than 70 years can be explained by the fact, that a majority of responsible car drivers stops driving car before reaching this age, because of safety reasons.

Combining the information from Figures 10 and 11 may lead us to a conclusion that younger and older household are spending relatively more of their budgets on fuel. At the same time, they are less willing to change their demand if a tax rate change is introduced. Increasing excise taxation on fuel has consequently asymmetric impacts on differently aged households; unfortunately socially weaker groups (as people in the beginning or after their productive carrer) probably carry a relatively higher tax burden. Increasing taxes on petrol and redistributing them to the "socially weaker" is from this perspective an inefficient activity, since they get back what they

TABLE 9

Education group of the household's head	Avg. tax paid before any change	Avg. tax paid when there is demand response	Avg. tax paid when no change in quantity demanded
Basic (no A levels)	2.20 %	2.43 %	2.63 %
Secondary	2.07 %	2.30 %	2.50 %
University	2.17 %	2.37 %	2.57 %

have paid and the redistribution of wealth is carried out from a substantial part within this group itself, meaning losses from the superfluous transaction costs and market distortions.

6.4 The Effect on Different Education Groups

The tax impact can be decomposed also according to the other demographic variables. We may think of decomposition according to the education groups.[12] Not to repeat the same steps as in the previous sections we may enlarge our analysis by an investigation of the assumption that the tax rate would not have any response to a consumer demand (in other words a change of an excise tax and consequently of the price will not result in any change of quantity demanded). This is rather a silly assumption, since it may only hold true with absolutely non-elastic demands.

Table 9 shows the effect of a 20% increase in the fuel excise rate on the tax paid (expressed as a % of a household's total income) for different education levels of the household head. Table 9 shows for three education groups in the first column the average fuel excise paid before any changes in taxation (as a percentage of total income). The second column describes the same if we assume the demand will respond to the change of price that is induced by an increase of excise by 20 %. Finally, the last column shows us how much tax on average a household would pay (again as a percentage of its total income) if we assume that there are no adjustments of quantity demanded after a tax (and price) change.

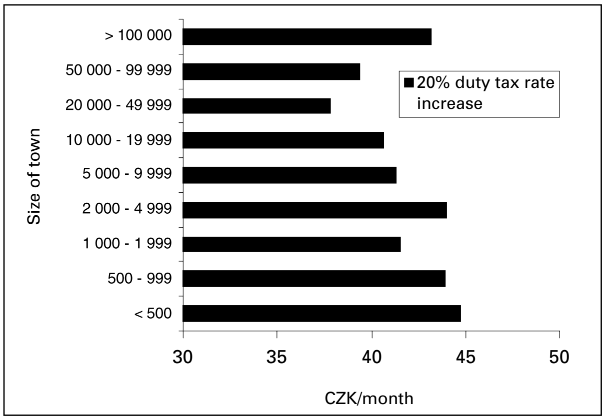
The values in Table 9 confirm our thoughts, that an increase of a tax will result in the decrease of quantity demanded. The finally obtained tax revenue will be somewhere in the region between the previous tax revenue and the hypothetical revenue calculated as previous quantity consumed multiplied by a new tax rate.

6.5 Effects of Fuel Taxation on People Living in Towns of a Different Size

Another interesting topic, we could investigate using the demand analysis based on micro data, is how a tax rate change may affect people living

[12] Instead of investigating all 10 possible groups of a different education type, we aggregated the education into a variable with just three values: *Basic* education for those who don't have A-levels, *Secondary* for those who have A-levels and *University* for any graduates.

FIGURE 12 Aditional Average Tax Paid after Duty Rate Increase



in different places. Tax on fuel may have a different impact on households living in different regions, but also on households living in different sized towns. We may assume that people from smaller towns are more likely to use a car as a mean of daily transportation to their jobs outside of where they live. It is probable that people living in bigger towns have a job in the same towns and do not need a car for travelling to work as the people from smaller town do. An increase of taxation should more heavily hit the people from smaller places, because their demand should be for this reason less elastic.

Investigating the impact of an excise tax change on people living in different sized towns reveals that the effect of a tax rate increase has indeed a bigger impact on people in smaller towns. The importance of the size of a town has been clearly indicated already in the estimates of the demand equation where town size dummies were highly significant. *Figure 12* shows how much would an average household in each town category pay after the excise tax rate was increased by 20 %. We can clearly see that the collected tax has a declining path as the size of the town grows from the smallest up to towns with 49,000 inhabitants. However, for the 2 largest groups of localities (with more than 49,000 people) the absolute amount of tax paid every month increases.

This finding may have important consequences for tax policy design. An increase of fuel taxation is relatively felt more by people in smaller towns and its increase may have negative impacts both on the labour market (i.e. flexibility of labour force to move) and on the living conditions in countryside.

7. Conclusion

In this study we were investigating the demand for fuel products in the Czech Republic. Firstly, we described the dataset to provide better intuition about the variables used in the followed analysis. Secondly, we reviewed the minimum necessary underlying microeconomic theory that is related to the applied demand analysis. In the empirical part, we estimated several model specifications using the instrumental variables econometric method and the Heckmen 2-step selection model. Finally, based on these demand function estimates, we showed several tax simulation consequences.

The demand modelling on the micro-data level allows us to estimate impacts of a tax rate change on different specified groups of people. We investigated the effect of a fuel excise tax rate increase of 25 % and 50 % on different income groups, households with differently olded and educated head and population of variously large towns. Because the *Family Expenditure Survey* considers a household as a unit of a measurement, the assumption about significant correlation between a head of a household and other household members had to be made to allow generalization of results.

The increase of the excise tax rate on fuel will result in a higher drop in consumption by higher income households. The reaction of lower income groups is milder and their adjustment in demanded quantity is lower. It suggests that the price elasticity of lower and higher income households is different and a tax rate change could have more negative consequences to the first group. Nevertheless, the view oriented on the tax revenues shows that the average increase of tax paid is positively dependent on income and hence higher income households will pay after a tax reform, on average more than households with lower incomes. These seemingly contradicting results are caused by the fact that investigated tax rate changes, more than offsets the drop in demanded quantity.

Younger and older households have relatively higher price elasticity compared to middle age households. Their reaction in demanded quantity to a tax rate increase is therefore smaller. Knowing that the share of fuel expenditures is also higher for younger and older households, they are the groups who are more likely to take the additional tax burden.

The households with a lower educated head pays a larger proportion of their total income on the fuel tax for every tax rate considered. The second largest proportion is paid by households, in which the head has a university degree. The households where the head completed secondary school are between these two groups.

Significant economic policy implications could be maybe derived from the fact that a fuel excise increase will harder hit people who live in smaller towns and villages. Their need for personal transportation might be higher for objective reasons. An increase in tax rates could have negative repercussion to the quality of life outside urban centres.

The microeconomic analysis of various economic policy programs based on observable data about individual households or firms could provide richer information of awakened policy responses. One of main advantages of the empirically based analyses, that fully reflect the principles of the economic theory and are not just the data mining, is a close relation to the reality. Their results often do not appear to be so *"scientifically looking and precise"* as the outcomes of deductively built models, but they could be used for solving practical economic policy problems.

APPENDIX

Data Description

Variable Description

Variable	Description
hsn	identification number of household in a specific region and district
region	region where the household lives *(there are 8 regions in the Czech Republic)*
district	smaller part of a region *every region (except Prague) is divided into 8–13 smaller administrative units*
soc_grp	social group of the head of household *(1: worker, 2: self-employee, 3: employee, 5: farmer, 7: retired/unemployed)*
members	number of household members
work	number of household members who are economically active
children	number of dependent children in a household
children5	number of dependent children under 5 years age
town_size	size of town or city 1: up to 499 2: 500 to 999 3: 1000 to 1999 4: 2000 to 4999 5: 5000 to 9999 6: 10 000 to 19 999 7: 20 000 to 49 999 8: 50 000 to 99 999 9: more than 100 000
hshl_type	type of household 1: complete pure family, head of household is economically active a) 2 members without children (husband & wife, 2 partners) b) family with juvenile children (no grandparents or other relatives living with the family) 2: complete pure family, head of household is not economically active 3: complete pure family, head of household and his wife are economically active 4: complete composite family a) 2 members without children with another relative (mother-in-law, grandson, grandmother) b) family with children where some children are economically active c) family with children with another relative in a household 5: pure incomplete family with children (only one parent) other similar to 1b) 6: composite incomplete family (only one parent) other similar to 2b, c) 7: non-family household (2 sisters, grandmother with a grandson) 8: individual person – male 9: individual person – female

age	age of the head of the household
edu	education of the head of the household 0: uncompleted primary school 1: primary school completed 2: primary school completed and secondary training institution 3: apprenticeship secondary school without A-levels 4: complete apprenticeship secondary school (with A-levels) 5: completed grammar secondary school (with A-levels) 6: completed practical secondary school or conservatory (with A-levels) 7: undergraduate university education (bachelor level) 8: graduate university education (master level) 9: postgraduate (PhD level)
total_inc	net money income of household per month in Czech Koruna (CZK) (gross income minus income taxes, mandatory, social insurance and selected savings and loans)
total_exp	net money expenditures of household in CZK per month
fuel_exp	net expenditures for car petrol in CZK per month
car	ownership of a car 0: household does not have any car 1: household has car

REFERENCES

ATKINSON, A. B. – STIGLITZ, J. E. (1980): *Lectures on Public Economics*. New York, McGraw Hill, 1980.

BAKER, P. – McKAY, S. – SYMONS, E. (1990): The Simulation of Indirect Tax Reforms: the IFS Simulation Program for Indirect Taxation (SPIT). *Institute for Fiscal Studies,* London, *Working paper,* no. W90/11.

BAKER, P. – BLUNDELL, R. – MICKLEWRIGHT, J. (1989): Modelling Household Energy Expenditures Using Micro-Data. *Economic Journal*, vol. 99 (September 1989), pp. 720–738.

BANKS, J. – BLUNDELL, R. – LEWBEL, A. (1996): Tax Reform and Welfare Measurement: Do We Need Demand System Estimation? *Economic Journal*, vol. 106 (September 1996), pp. 1227–1241.

BANKS, J. – BLUNDELL, R. – LEWBEL, A. (1997): Quadratic Engel Curves and Consumer Demand. *Review of Economics and Statistics*, vol. LXXIX, no. 4 (Nov. 1997), pp. 527–539.

BLOW, L. – CRAWFORD, I. (1997): The Distributional Effects of Taxes on Private Motoring. *Institute for Fiscal Studies,* London, *Commentary*, no. 65, December 1997.

BLUNDELL, R. (1988): Consumer Behaviour: Theory and Empirical Evidence – A Survey. *Economic Journal,* vol. 98, issue 389 (March 1988), pp. 16–65.

BROWN, A. – DEATON, A. (1972): Surveys in Applied Economics: Models of Consumer Behaviour. *Economic Journal*, vol. 82, no. 328 (Dec. 1972), pp. 1145–1236.

CRAWFORD, I. – LAISNEY, F. – PRESTON, I. (1998): Estimation of Household Demand Systems Using Unit Value Data. *Institute for Fiscal Studies, Working paper*, no. 02/17 (London, October 1998).

CRAWFORD, I. – SMITH, S. – WEBB, S. (1993): VAT on Domestic Energy. *Institute for Fiscal Studies*, London, *Commentary*, no. 39.

DEATON, A. (1997): The *Analysis of Household Surveys: A Microeconometric Approach to Development Policy*. The Johns Hopkins University Press, Baltimore and London, 1997, pp. 271–334.

DEATON, A. – MUELLBAUER, J. (1980): An Almost Ideal Demand System. *American Economic Review*, vol. 70, June 1980, no. 3, pp. 321–326.

DEATON, A. – MUELLBAUER, J. (1980): *Economics and Consumer Behaviour*. Cambridge University Press, reprinted in 1993.

DUBIN, J. A. – McFADDEN, D. L. (1984): An Econometric Analysis of Residential Electric Appliance Holdings and Consumption. *Econometrica*, vol. 52, March 1984, no. 2.

GALLINI, N. (1983): Demand for gasoline in Canada. *Canadian Journal of Economics*, vol. XVI, 1983, no. 2.

HÄRDLE, W. (1990): *Applied Nonparametric Regression*. Cambridge University Press, 1990.

HECKMAN, J. (1979): Sample Selection Bias as a Specification Error. *Econometrica*, vol. 47, issue 1 (Jan 1979), pp. 153–162.

ICHIMURA, H. (2002): *Lecture Notes on Microeconometrics*. London, UCL – mimeo.

JAMES, S. – NOBES, CH. (1996): The *Economics of Taxation, Principles, Policy and Practice*. Prentice Hall Europe, 1997/1998 edition.

MADDEN, D. (1996): Marginal Tax Reform and the Specification of Consumer Demand Systems. *Oxford Economic Papers,* vol. 48, 1996, pp. 556–567.

MAYSHAR, J. – YITZHAKI, S. (1995): Dalton-Improving Indirect Tax Reform. *American Economic Review*, vol. 85, 1995, no. 4 (September 1995).

PESARAN, M. H. – SMITH, R. J. (1994): A Generalized R^2 Criterion for Regression Models Estimated by the Instrumental Variables Method. *Econometrica*, vol. 62, 1994, no. 3, pp. 705–710.

PHLIPS, L. (1983): *Applied Consumption Analysis*. North-Holland Publishing Company, 1983.

SALTER, M. (2001): A *Study of Optimal Fuel Taxation*. Dissertation at the University College London.

SLAVÍK, M. (2002): *Demand for petrol in the Czech Republic and its consequences for indirect taxation.* Dissertation at the University College London.

SMITH, S. – HEADY, CH. (1995): Tax and Benefit Reform in the Czech and Slovak Republic. In: Newbery, D. (ed): *Taxation and Benefit Reform in Central and Eastern Europe.* London, CEPR, 1995.

VARIAN, H. R. (1992): *Microeconomic Analysis*. 3rd ed. W. W. Norton & Company, 1992.

WHEATON, W. C. (1982): The long-run structure of transformation and gasoline demand*. Bell Journal of Economics*, 1982, pp. 439–454.

SUMMARY

JEL classification: C21, D12, H31
Keywords: consumer demand – excise taxes – tax simulation – nonparametric regression

Fuel Demand and Indirect Taxation: A Microeconomic View

Michal SLAVÍK – PhD candidate, Faculty of Social Sciences, Charles University, Prague and Czech National Bank (michal.slavik@seznam.cz)

The author uses microeconomic theory, particularly demand theory, and econometrics to investigate the impact of proposed tax changes on the consumers. The effect of taxation can be investigated in two ways: the first possibility is to construct a multi-equation model able to analyze all taxes simultaneously. The second is to create a partial model for a specific tax, which would register subtler effects. The second approach is adopted here, with an exclusive focus on fuel-product excise taxes and their microeconomic impact.

Two alternative approaches, instrumental variables and the Heckman two-step selection model, are used to estimate the demand function for fuel. The estimated demand function is then used toward a tax simulation and an investigation of the tax-change effects on different segments of the population.