PhD Thesis

Essays in transportation economics

by

Ismir Mulalic

April 2011

Department of Economics

University of Copenhagen

www.econ.ku.dk

Contents

Preface

This thesis was written from February 2008 to April 2011 while I was enrolled as a PhD student at the Department of Economics at the University of Copenhagen. The work in this thesis is part of a larger research project at the Technical University of Denmark (DTU): "Interaction between regulation, car choice, and energy consumption (ENS-33032-0207)" coordinated by Mogens Fosgerau. The research project was undertaken by researchers at the Department of Transport (DTU Transport) with financial support from the Danish Energy Agency (Energy Research Programme). Financial support from the Energy Research Programme is gratefully acknowledged. This thesis could not have been written without the help of a number of people. It would be impossible to name them all, but the following deserve special thanks.

First of all, I would like to thank my supervisors, Jørgen Birk Mortensen, Mogens Fosgerau, and Peter Birch Sørensen, who have been truly accessible and supportive throughout the entire program. I am very grateful for all their help, effort and inspiration. Thank you.

My most important sources of learning during my time as a PhD student have been my co-authors, without whom this thesis would be very different and most likely of poorer quality. They include Akli Berri, Jan Rouwendal, Jos N. van Ommeren, Ninette Pilegaard, Stéphanie Vincent Lyk-Jensen and Theodoros Zachariadis. It has been inspiring to work with all of you. I have learned a lot, and your co-authorship has greatly improved the quality of this thesis. Thank you.

During September 2009, I visited the University of Antwerp. This visit was highly rewarding and helped me improve parts of this thesis. I am grateful to the University of Antwerp for providing this opportunity and for their hospitality. I owe special thanks to Bruno de Borger who kindly devoted time to discuss my research with me. Thank you.

It has been rewarding as well as a pleasure to be part of the Department of Economics. Department of Economics has provided excellent facilities and an inspiring environment with lots of seminars and workshops. I would in particular like to thank Peter Norman Sørensen who I got to know both as a director of the PhD programme and during many seminars and workshops, both domestic and abroad. Furthermore, I thank DTU Transport for providing a very stimulating research environment with inspiring colleagues, excellent computational facilities, and access to data. Thank you to all my colleagues at the DTU Transport for contributing to this. I thank you for your comments and insightful discussions. I look forward to spending more time with you in the future. I also acknowledge the excellent assistance from the secretaries at the Department of Economics and DTU Transport. For the language check carried out not only in a professional but also in very supportive way I wish to thank Ulla Steen Salado-Jimena. I am grateful to Statistics Denmark (especially Karin Ørum Elwert and Søren Dalbro) for providing the register data used in this thesis. Special thanks go to all my office-buddies Katrine Hjorth, Simon Halphen Boserup (thanks for STATA support), Toke Fosgaard, Jori Veng Pinje, and Johannes Jacobsen, for the daily conversations that have helped me sharpen my thoughts and papers. Without you, the past three years would have been less fun and a lot more demanding. I would also like to thank my family and friends (especially Asim, Nadžija, Jasmin, Brynja, Nermin, Alma, Lejla, Edin, and Surface Below) for their unconditional support, and my parents Enver and Azra for their full understanding, support and encouragement during the course of my life and during this project. In addition, I have benefitted from a large number of helpful comments from seminars and conference participants, both in Denmark and abroad. Thank you for that.

Lastly, I am forever thankful to my wife, Selma Mustafic Mulalic, for being supportive and understanding during all the ups and downs of this project and our little son Mak for all his smiles and pleasant distractions. When I was absent because I worked late nights, because I was abroad for conferences, or because my thoughts were elsewhere, you have been amazingly patient with me. This is much more appreciated than you think. Words are inadequate to express my gratitude. I therefore dedicate this work to you.

If I have forgotten to name anyone, it is not because their help has not been appreciated. Needless to say, none of the above people bear the least of responsibility for any remaining errors or omissions in the thesis.

Ismir Mulalic *April 2011*

University of Copenhagen, Denmark

Introduction and summary

This thesis includes four self-contained chapters which cover a rather wide range of issues. All four chapters, however, share a common ground as they consider issues related to the broad field of transportation economics. The main aim of the thesis is not to produce a single message which is supported by all four chapters. Rather, each chapter was written to make a contribution of its own. Chapter one examines the causal effect of commuting distance on workers' wages. Chapter two shifts attention from labour market to the car taxation and household inequalities regarding consumption of transport goods and services. The chapter applies decomposition of the Gini index by expenditure component and estimates the redistributive effects of taxes on different transport commodity categories. Chapter three focuses on consumer choice behaviour for durables. The chapter develops a simple model to show that the marginal willingness to pay for a quality attribute of a durable has to be equal to the full marginal cost, which includes marginal fixed as well as variable costs, and apply the model to Danish data on car ownership and use. The last chapter analyses the determinants of the trucking firm fuel use. A summary of each of the four chapters is given below.

Chapter 1 is entitled "*Wages and commuting: quasi-natural experiments' evidence from firms that relocate*", and is joint work with Jos N. van Ommeren and Ninette Pilegaard. In this chapter, we study the causal effect of commuting distance on wages. The analysis is based on a wage bargaining. In this context workers get a fixed share of their commuting costs (equivalent to the bargaining power parameter) reimbursed through higher wages (see e.g. Marimom and Zilibotti, 1999; Van Ommeren and Rietveld, 2005). We estimate (reduced-form) panel data models using matched data from workers and firms for Denmark to determine the magnitude of this share. The data used in the empirical analysis are derived from annual register data from Statistics Denmark for the years 2003–2005.

Traditional assessments of the positive correlation between the commuting distance and wages have used cross-section data and it is possible that the endogeneity of the commuting distance is not appropriately controlled for, because of alternative explanations of the causal effect of commuting distance on wages (see e.g. Zax, 1991; White, 1977; Benito and Oswald, 1999; Manning, 2003). Using exogenous changes in commuting distance due to firm relocations, which fits in the literature of quasi-natural experiments, we are able to exclude many competing explanations of the wage-distance relationship, and interpret our estimates from a wage bargaining perspective.

We examine two types of effects of commuting distance on wages, i.e. the effect of commuting distance on wages for the range in commuting distance where the income tax reduction associated with commuting is not applicable (where the effect refers to overall (time and monetary) costs associated with commuting), and the effect where income tax reduction is applicable (where the effect refers principally to time costs losses associated with commuting). We show that commuting distance increases imply, on average, an overall hourly wage compensation of about 49% and 22%, for the commuting distance range where the income tax reductions do not apply and for the commuting distance range where income tax reduction is applicable, respectively. Moreover, the effect becomes zero at a commuting distance of 50 kilometres. The estimation results imply a wage bargaining parameter of about 0.5. The results appear robust with specification and accounting for selection effects.

There is a number of reasons why the effect of commuting distance on wages is of interest (for a review see Gibbons and Machin, 2006). Our findings put a price on commuting distance and points to the economic benefits from transport infrastructure improvements. We also show that wage bargaining with respect to commuting is an important characteristic of the (Danish) labour market.

Chapter 2 is entitled "*Household transport consumption inequalities and redistributive effects of taxes: a repeated cross-sectional evaluation for France, Denmark and Cyprus*", and is joint work with Akli Berri, Stéphanie Vincent Lyk-Jensen and Theodoros Zachariadis. In this chapter, we present an analysis of household inequalities regarding consumption of transport goods and services in three European countries (France, Denmark and Cyprus). A comparative analysis is carried out in the light of the differences between these countries in terms of car taxation systems and car ownership levels. More specifically, a decomposition of the Gini coefficient by expenditure component is applied to investigate temporal dynamics of household inequalities and to estimate the redistributive effects of taxes on different commodity categories. The analyses are carried out on data from repeated cross-sections of household expenditure surveys in France, Denmark and Cyprus spanning long time periods.

The analysis is motivated by the contention that apparently cars' social diffusion over time is likely to have lessened the progressivity of car taxes. In this study, we adopt a covariance-based formulation of the Gini coefficient that is used to obtain a decomposition of the Gini coefficient by the expenditure component (Lerman and Yitzhaki, 1994). The decomposition makes explicit the mechanisms by which each considered expenditure component (e.g. fuel expenditure) contributes to the global Gini and therefore enlightens the temporal patterns of inequalities. Each expenditure component appears through its proper Gini coefficient, its budget share and its degree of association with total expenditure (Lerman and Yitzhaki, 1985). Besides, decomposition of the Gini coefficient allows evaluating the redistributive effects of marginal changes in the different expenditure components considered which can be interpreted as the redistributive effect of a marginal change in taxes on a particular good or a service. By redistributive effect is meant the impact in terms of inequality increase or reduction.

We find that inequality regarding transport is mainly attributable to car purchases, while the relative contribution of public transport is very small. Moreover, the relative contribution of car use expenditures decreased over time. We also find evidence for overall progressivity of transport related taxes. However, this is principally due to the progressivity of taxes on car purchases. On the contrary, taxes on fuels are regressive except in Denmark, whereas the progressive character of taxes on the vehicle use goods and services has nevertheless become weaker over the years. In general, our findings reveal the effect of the gradual diffusion of cars in the last decades.

This study underlines the fact that equity issues should be considered when designing transport policies. In particular, the design of policy measures to reduce car use should take the inequality concerns into account.

Chapter 3 is entitled "*The willingness to pay for quality aspects of durables: theory and application to the car market*", and is joint work with Jan Rouwendal. In this chapter we study consumer choice behaviour for durables. We develop a simple model to show that the marginal willingness to pay for a quality attribute has to be equal to the full marginal cost, which includes marginal fixed as well as variable costs. Specifically, we allow the quality characteristics to appear in the utility function, so consumers also derive utility from quality characteristics of a durable in a direct way. We also show that the trade off between fixed and variable costs of energy-using durables studied by Hausman (1979) and Dubin and McFadden (1984) corresponds to the special case of our model in which the quality attributes do not affect preferences directly. We show that in this case the marginal variable and fixed costs associated with a quality attribute have to be of opposite sign implying the trade off. The conventional approach in housing economics where marginal willingness to pay is set equal to marginal fixed cost (see e.g. Sheppard, 1999) corresponds to another special case of our model in which the quality attributes have no impact on variable cost.

The model is applied to Danish data on car ownership and use derived from register data from Statistics Denmark and a car model database from the Danish car dealer association. The dataset includes, besides information on car attributes and household characteristics, also car prices, car costs and car use. This allows us to estimate hedonic price functions for fixed as well as variable costs. We follow Bajari and Kahn (2005) and use a highly recommended nonparametric estimation procedure to estimate a hedonic price functions (see e.g. Pace, 1995). More precisely, we use local linear regression to approach the hedonic price functions at each observation point (Fan and Gijbels, 1996; Härdle, 1993). We then recover each consumer's marginal willingness to pay, the marginal fixed costs, and the marginal variable costs for car attributes using first-order conditions for utility maximization. These estimates are household-specific in the sense that a unique set of the marginal fixed costs, the marginal variable costs and the total marginal willingness to pay are estimated for each household. So, we allow households to differ in their preferences for quality aspects of a car.

The empirical evidence strongly confirms that a model focusing on the trade-off between fixed and variable costs is not supported by the Danish data on car ownership and use. Further, our empirical analysis reveals that marginal variable car costs on average are about 20% of the total marginal cost. Moreover, for a broad range of car characteristics we found that the distribution of total willingness to pay is much smoother than that of marginal fixed cost and marginal variable cost. We also show that it is possible to relate the implicit prices for quality attribute that follows from our estimates of the fixed and variable cost functions to the marginal rate of substitution between the quality of the durable and the variable cost and that this variable can be considered as a structural preference parameter if one is willing to impose the necessary functional form assumptions. More precisely, we investigate how this marginal willingness to pay varies with household characteristics. Interesting correlations were found.

Chapter 4 is entitled "*The determinants of fuel use in the trucking industry - volume, size and the rebound effect*". The chapter is motivated by the continuing concern that surrounds the transportation economics literature on the transport sectors fuel efficiency and fuel use. This chapter contributes to the study of the determinants of the trucking firm's fuel use. We develop a simple, unified framework that embodies the trucking firm's behaviour as well as the technical relation relating to fuel use, traffic volume, and fuel efficiency. The trucking firm's fuel use is explained by highlighting the role of fuel efficiency. We build our model around the trucking firm profit maximization problem.

The model permits an analysis not only of the determinants of the fuel use but also the rebound effect. We decompose the standard definition of the rebound effect for motor vehicles (see e.g. Small and Van Dender, 2007), i.e. the elasticity of traffic volume with respect to fuel cost, into the elasticity of freight activity with respect to fuel cost per kilometre and the elasticity of traffic volume with respect to freight activity for the purpose of analysing the rebound effect for road freight transportation. Moreover, we analyse trucking firm fuel use by relaxing conventional assumption on exogenous change in fuel efficiency. We assume that the change in fuel efficiency is the result of changes in the average truck's attributes. We focus on the effect of changes in fuel price on the trucking firm's total fuel use. We show that disregarding dependence of fuel efficiency on fuel prices and road freight activity (the movement of cargo measured in ton-kilometres) may cause biased estimates of the effect of the change in the fuel price on the trucking firm's fuel use and in particular biased estimates of the rebound effect.

Using aggregate time-series data for Denmark we investigate how trucking firms react to changes in fuel prices. We estimate the simultaneous-equation model that simultaneously determines traffic volume, the size of the truck stock, average truck attributes (average truck capacity and average truck age), labour demand, freight activity, and fuel efficiency. The empirical results show that the average rebound effect for trucking in applied sample is 19% in the short run and 28% in the long run. Moreover, the empirical part of the chapter reveals that higher fuel prices decrease the average trucking firm's fuel use, but only by a small amount. The elasticity of fuel use with respect to fuel price has been estimated to -0.19 in the long run. Surprisingly, an increase in the fuel price has negative effect on the trucking firm fuel efficiency. The trucking firm responds to an increase in the fuel costs by rejuvenation of the truck stock, and the newer trucks use less fuel per kilometre. However, the increase in the fuel use caused by an increase in the average truck capacity offsets this effect. We show that the trucking firm also responses to an increase in the fuel price through expansion of the average truck capacity and, for a given traffic volume, trucks with higher capacity use more fuel. Thus, an increase in the fuel price decreases the trucking firm's fuel efficiency. However, less distance has to be driven for the same payload, so an increase in the fuel price results in the reduction in the trucking firm's overall fuel use.

The results have implications for the recent debate whether to adapt the rules on the optimal weights and dimensions of heavy trucks in EU. The estimation results show that, due to the rebound effect, strengthening fuel efficiency standards for heavy trucks in the EU can potentially result in undesirable effects.

References

Bajari, P. and Kahn, M. E., 2005. Estimating housing demand with an application to explaining racial segregation in cities. Journal of Business & Economic Statistics, 23, no. 1, 20-33.

Benito, A. and Oswald, A., 1999. Commuting in Great Britain in the 1990s Warwick. Department of Economics, University of Warwick.

Dubin, J. and McFadden, D., 1984. An econometric analysis of residential electric appliance holding and consumption. Econometrica, 52, 345-362.

Fan, J and Gijbels, I., 1996. Local polynomial modeling and its applications. Chapman & Hall, London.

Gibbons, S. and Machin, S., 2006. Transport and labour market linkages: empirical evidence, implications for policy and scope for further UK research. Commissioned for the Eddington Study.

Hausman, J., 1979. Individual discount rates and the purchase and utilization of energy-using durables. The Bell Journal of Economics, 10, 33-54.

Härdle, W., 1993. Applied nonparametric regression. Cambridge University Press.

Lerman, R.I. and Yitzhaki, S., 1994. Effect of marginal changes in income sources on U.S. income inequality. Public Finance Quarterly, 22 (4), 403-417.

Lerman, R.I. and Yitzhaki, S., 1985. Income inequality effects by income source: a new approach and applications to the United States. Review of Economics and Statistics, 67 (1), 151-56.

Manning, A., 2003. The real thin theory: monopsony in modern labour markets. Labour Economics 10 (2), 105–31.

Marimom, R. and Zilibotti, F., 1999. Unemployment versus mismatch of talents: Reconsidering unemployment benefits. Economic Journal, 109, 266–291.

Pace, R. K., 1995. Parametric, semiparametric, and nonparametric estimation of characteristics values within mass assessment and hedonic pricing models. Journal of Real Estate Finance and Economics, 11, 195-217.

Parry, I. W., 2008. How should heavy-duty trucks be taxed?. Journal of Urban Economics, 63, 651-668.

Sheppard, S., 1999. Hedonic analysis of housing markets. in Handbook of Regional and Urban Economics, eds. P. Cheshire and S. Mills, Volume 3, Chapter 41, pp. 1595-1635.

Small, K. A. and Van Dender, K., 2007. Fuel efficiency and motor vehicle travel: The declining rebound effect. The Energy Journal, Vol. 28, No. 1, 25-51.

Van Ommeren, J.N. and Rietveld, P., 2005. The commuting time paradox. Journal of Urban Economics 58 (3), 437–54.

White, M. J., 1977. A model of residential location choice and commuting by men and women workers. Journal of Regional Science, 17, 41–52.

Zax, J.S., 1991. The substitution between moves and quits. The Economic Journal, 101 (409), 1510–21.

Chapter 1

Wages and commuting: quasi-natural experiments' evidence from firms that relocate

Ismir Mulalic, Jos N. van Ommeren, and Ninette Pilegaard

Wages and commuting: quasi-natural experiments' evidence from firms that relocate

Ismir Mulalic

University of Copenhagen, Øster Farimagsgade 5, Building 26, 1353 Copenhagen K, Denmark, and
Technical University of Denmark, Bygningstorvet 116 Vest, 2800 Kgs. Lyngby, Denmark
imu@transport.dtu.dk

Jos N. Van Ommeren

VU University, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands
jommeren@feweb.vu.nl

Ninette Pilegaard

Technical University of Denmark, Bygningstorvet 116 Vest, 2800 Kgs. Lyngby, Denmark
np@transport.dtu.dk

Abstract

We examine the causal effect of commuting distance on workers' wages in a quasi-natural experiments setting using information on all workers in Denmark. We account for endogeneity of distance by using changes in distance that are due to firms' relocations. For the range of commuting distances where income tax reductions associated with commuting do not apply, one kilometre increase in commuting distance induces a wage increase of about 0.42%, suggesting an hourly compensation of about half of the hourly net wage. Our findings are consistent with wage bargaining theory and suggest a bargaining power parameter of about 0.50. Due to the experimental setup we are able to exclude many competing explanations of the wage-distance relationship.

Keywords: Commuting, wages, bargaining theory.

JEL codes: J22, R41.

1. Introduction

This paper examines the causal effect of commuting distance on wages from a wage bargaining perspective. One of the main issues we are concerned with is that distance may be endogenous with respect to wages. This is relevant because the literature emphasizes that it is not an easy task to find valid instruments for commuting distance, as it is related to labour and residence locations behaviour (Manning, 2003; Gubits, 2004). As emphasised by Manning (2003), but also in the literature study by Gibbons and Machin (2006), despite the large number of studies there is essentially no direct empirical evidence of the causal effect of commuting costs on wages. There are a number of reasons why the effect of the length of the commute on wages is of interest (for a review see Gibbons and Machin, 2006).

Evidence on equilibrium relationships between wages and commuting is informative about the frictions in the labour market that transport infrastructure may help to alleviate. In a competitive labour market without search frictions, firms do not determine wage levels based on the worker's commuting distance (as it is based on the worker's productivity level). If firms pay compensating wages for longer commutes, then firms must enjoy some monopsony power in the labour market which allows them to pay wages below marginal product. In a wage bargaining context with job search frictions, workers with long commuting distances are able to bargain for higher wages, because their opportunity costs of staying with the firm are less than those for other workers. To be more precise, a range of bargaining models imply that workers will get a fixed share of their commuting costs reimbursed through higher wages (e.g. Marimom and Zilibotti, 1999; Van Ommeren and Rietveld, 2005).[1] This share is determined by the degree of employer market power, which plays a major role in a wide range of bargaining models (Pissarides, 2000). In a competitive labour market, employers have no market power, so the share is zero. In a market where employers have full market power, the share is one (and worker receive a wage which makes them indifferent between working and being unemployed), so workers receive full compensation. There is virtually no direct evidence on what an appropriate value of this share (equivalent to the bargaining power parameter) should be (see e.g. Shimer, 2005; Mortensen and Nagypal, 2007; and Gertler et al., 2008). It is one of the objectives of the

[1] Wage compensation does not occur when workers with long commutes are fully compensated in the housing market through lower housing prices (see Zenou, 2009). We control for housing market compensation by keeping residential location constant.

current paper to determine the magnitude of this share for Denmark.[2] Evidence on a causal effect of the commuting distance on the wage implies that workers receive (a part) of their total commuting costs (inclusive time costs) reimbursed through higher wages.

The effect of the length of the commute on wages is also relevant in the context of income taxation. Income tax reductions for workers with a long commute can be found in many European countries (see Potter et al., 2006). In Denmark, workers with a one-way commute that exceeds 12.5 kilometres (about 50% of all workers) are entitled to a tax reduction. This tax reduction is based on a deduction that aims to compensate for monetary expenses associated with commuting per kilometre, i.e. fuel expenses, vehicle amortisation expenses, etc.[3] More precisely, in 2003 the workers were entitled to deduct 3.2 DKK per daily one-way kilometre for commutes between 12.5 and 50 kilometres and 1.6 DKK for longer commutes from taxable gross income.[4] On average, these reductions imply an increase in net wage of about 33% of the reduction, so 1.06 and 0.53 DKK per daily one-way kilometre respectively. This is substantial, as it implies an average net wage compensation of about 1.23% and 3.60% for commuting distances of respectively 12.5 and 50 one-way commuting kilometres (the average wage per working day was 1,114 DKK).

Despite the large theoretical and empirical debate around the relationship between wages and the length of the commute, it is maybe surprising that there is an absence of accurate empirical estimates of the *causal* effect of the length of the commute on wages (see e.g. Zax, 1991; Manning, 2003). Hence, there is a knowledge gap between the theoretical and empirical literature. We aim to fill this gap in the literature by estimating the *causal* effect of workers' commuting distance on wages. We estimate (reduced-form) panel data models using matched data from workers and firms for Denmark. We are interested in two types of effects of commuting distance on wages. So, we discuss (i) the effect of commuting distance on wages for

---

[2] Since the early 1980s, the Danish labour market has experienced a trend towards more decentralized bargaining regime based on flexible wage structures, inequality, and market-induced restraint (Iversen, 1996). Given the flexible wage structures in Denmark, individual unions and their employer counterparts determine the general wage level, while the workers bargain for additional bonuses at the level of firm, so the overall wage level is bargained at the individual level. In addition, the Danish labour market is relatively flexible, i.e. worker turnover is relatively high (Mortensen, 2001) and the high level of turnover applies to most categories of employees and is not caused by a minor share of (unskilled) workers being extremely mobile (Madsen, 2002).

[3] Employers seldomly reimburse commuting expenses explicitly (viz. through a fringe benefit), so we ignore this issue. Approximately 0.3% of workers have access to a company car. Including or excluding these workers does not affect estimation results.

[4] One DKK is approximately 0.13€. In 2005, per daily one-way kilometre the tax deductions were 3.36 DKK and 1.68 DKK respectively.

the range in commuting distance where the income tax reduction is not applicable (where the effect refers to overall (time and monetary) costs associated with commuting), and (ii) the effect where income tax reduction is applicable (where the effect refers principally to time costs losses associated with commuting).

Our study deals with a range of statistical difficulties that have not been properly addressed in the literature by making use of exogenous changes in commuting distance due to firm relocations, which fits in the literature of quasi-natural experiments. Therefore, our study strongly contrasts with previous studies. Our estimates can be interpreted from a wage bargaining perspective, whereas interpretation of previous studies, mainly based on cross-section data, is not straightforward, because of alternative explanations (see e.g. Zax, 1991; White, 1977; Benito and Oswald, 1999; Manning, 2003; Simonsohn, 2006). In principle, in addition to the wage bargaining explanation, there are at least four other explanations for a *positive correlation* between the length of the commute and wages.

First, according to urban economic theory, workers with a higher income choose a different residence location and therefore a different commute (Wheaton, 1974). This explanation relates to reversed causation. Second, unobserved variables, such as skills, may affect both commuting distance and wages, causing spurious correlation between the length of the commute and wages (Manning, 2003). A common method of dealing with these two explanations is the use of an instrumental variable estimation procedure. The problem with this approach in the current setting is finding suitable instruments for the length of the commute as argued by Manning (2003). We use employer-induced changes in distances rather than an instrument variable approach.[5] Third, given a competitive labour market, employers located at locations far from residences compensate workers with appropriately higher wages, which implies a spatial wage gradient (e.g. Fujita et al., 1997). This idea is confirmed by empirical findings (Timothy and Wheaton, 2001). We deal with this alternative explanation by using (year-specific) firm fixed effects. Fourth, Manning (2003) points out that, in a monopsonistic labour market with search frictions and a distribution of wages, workers receive many job offers but only those above a reservation wage are accepted; otherwise it pays to wait for further offers. The reservation wage rises with the commuting cost associated with the job offer. So, on average, wages rise with commuting

[5] Manning (2003) re-examines the results by Benito and Oswald (1999) and finds that the IV approach used by Benito and Oswald (1999) is sensitive to the choice of the instruments.

distance because workers only accept distant jobs that, at least partially, compensate for additional commuting costs. We control for this explanation by focusing on changes in wages of workers who remain with their employer.

The next section introduces the identification strategy to estimate the causal effect of commutes on wages in a wage bargaining framework; Section 3 provides information on the data employed; Section 4 presents the empirical results; Section 5 concludes.

2. Identification strategy

Wage bargaining theory predicts a positive relationship between wages and commutes for workers, ceteris paribus. To guarantee the ceteris paribus condition, it is useful to focus on workers who stay with their employer and do not change residence. The hypothesis is then investigated using Danish register data on *all* workers that are matched to *all* firms *that relocate*. The worker's commuting distance, defined by the residence and the workplace location, is usually self chosen by the worker. However, quite regularly, the workplace location is changed due to a relocation by the worker's employer. The commuting distance change is then employer-induced and therefore exogenous with respect to individual wages. In our approach, we only use these exogenous changes. The idea to use workplace relocation as a source of exogenous change in commuting distance is also exploited in Zax (1991) and Zax and Kain (1996), who analyse the effects of a relocation of a single firm on job and residential moving behaviour.[6] The analysis of the relationship between wages and commuting distances based on exogenous changes in the distance due to firm relocation fits in the literature of quasi-natural experiments.

More formally, our approach entails estimating a worker's first-differences wage model with controls for worker- and firm-specific time-invariant factors. Let $W_{i,f,t}$ denote the worker $i's$ wage in year t of firm f. We assume the following specification of wages:

$$log(W_{i,f,t}) = \alpha_0 + \alpha_1 d_{i,f,t} + \alpha_2 X_{i,f,t} + \eta_{f,t} + \varepsilon_i + u_{i,f,t} \quad (1)$$

where $d_{i,f,t}$ is the worker $i's$ commuting distance in period t employed by firm f. The matrix $X_{i,f,t}$ includes exogenous time-varying controls for workers' and firms' characteristics, ε_i is a worker fixed effect, and $u_{i,f,t}$ is the overall error. We emphasise that in (1) we have included year-specific firm fixed effects $\eta_{f,t}$, which control for all year-specific differences between firms

[6] See also Gutiérrez-i-Puigarnau and Van Ommeren (2010), who estimate distance effects on labour supply, but in their study firm relocations are not observed (but derived from job and residential mobility data).

(e.g., firms' size, firms' location, firms' sales, etc.). We estimate all models in terms of worker first-differences, that is, variables are formulated as changes from one time period to another, implying that:

$$log(W_{i,f,t}) - log(W_{i,f,t-1}) = \alpha_1(d_{i,f,t} - d_{i,f,t-1}) + \alpha_2(X_{i,f,t} - X_{i,f,t-1}) + \varphi_{f,t} + v_{i,f,t} \quad (2)$$

where $v_{i,f,t} = u_{i,f,t} - u_{i,f,t-1}$ and $\varphi_{f,t} = \eta_{f,t} - \eta_{f,t-1}$. Thus, we use within-workers' variation in commuting distance to explain within-worker's variation in wages and further control for year-specific changes in firm characteristics.[7] Consistent estimation of α_1 requires that the change in commuting distance, $d_{i,f,t} - d_{i,f,t-1}$, is exogenous and therefore not related to $v_{i,f,t}$.

In order to guarantee that the change in commuting distance is exogenous, we make two data selections. First, we select firms that changed location, so the change in commuting distance is the result of an employer-induced workplace relocation. This selection may create a selection bias as the set of firms that relocate may not be random. This bias is likely minimal however because we include (year-specific) firm fixed effects. Second, to control for *voluntary* worker changes in distance, we select workers (of firms that relocate) who did not change employer or residence (so we keep residence location constant). In this way, changes in distances are due to (usually unexpected) exogenous shocks in commuting distance.[8] Selecting a sample of workers who do not change employer and who do not change residence may create a selection bias. We will explicitly address the potential bias of this selection by comparing results of different samples and by estimating Heckman selection models.

We are mainly interested in the effect of changes in commuting distance on changes in wages, so we have experimented with functional form for commuting distance, and employed different samples and selection procedures. Our estimate imply that including observations of voluntary worker changes in distance through residential moves may bias estimation results, but including these changes through job moves does not bias the results. In the current paper, we discuss the results of a specification using commuting distance, and its square, and we explicitly allow for the nonlinear effect of income tax reduction associated with commuting. In addition,

[7] As we essentially have two periods in our data, this specification implies that we include only one $\varphi_{f,t}$ per firm.

[8] The shock is usually unexpected, because firms do not announce long in advance that they consider relocating. This phenomenon may have several explanations. For example, a long announcement period may increase uncertainty, which may immediately increase worker job quitting behaviour, absenteeism, etc., which firms prefer to avoid.

we also discuss the results which allow for the possibility that an increase in commuting distance induces a different wage effect than a decrease.

3. Empirical analyses

3.1. The data

The data used in the empirical analysis are derived from annual register data from Statistics Denmark for the years 2003–2005. Our period of observation is thus three years. For each year on 31 December, we have information on worker’s residence location and the workers’ establishment workplace location, annual *net* wages, and a range of explanatory variables (e.g. number of children).[9] Commuting distances have been calculated using information on exact residence and workplace addresses using the shortest route. For convenience, we will refer to establishments as firms.

3.2. Selection of sample and descriptive statistics

We observe the full population of 321,337 firms[10] and 2,710,462 workers. We select firms that changed address between January 2004 and December 2004 (11,314 firms; 64,643 workers). Records with missing information (4,209 firms; 5,857 workers); workers with more than 1 job (3,122 firms; 15,576 workers) and part-time workers (1,948 firms; 23,485 workers) were excluded. Furthermore, we excluded 337 observations referring to address changes that did not imply a change in commuting distance.[11] Moreover, observations for which commuting distance is greater than 100 km (179 firms; 878 workers), change in commuting distance is greater than 50 kilometers (19 firms; 434 workers), and the absolute change in log(wage) is greater than 0.5 (167 firms; 1,474 workers) were excluded as they were assumed to be outliers. Our econometric approach is based on a sample of (maximally) 1,333 firms and 8,601 workers. The full sample selection process can be found in Appendix A (Table A1).

Our focus is on a sample of workers who stayed with their firm and did not change residence from January 2003 to December 2005 (1,244 firms; 6,165 workers). We use wage data

[9] Wage data are derived from workers’ pay slips which are observed by Danish Tax Authorities.

[10] The statistical unit of firms is an administrative unit used by the tax authorities' register of enterprises liable to VAT. These units are identified by their so-called SE number. In most cases, the SE number unit is identical to the legal unit, i.e. the enterprise, but an enterprise might choose to split its registration up into several SE numbers (a divided registration). We assume that each SE number is a separate firm.

[11] One reason may be a change in street name, or building number, but it may also occur given a move within the same building (e.g. from one floor to another).

for the years 2003 and 2005, because within these years the commuting distance is constant (which is not the case for 2004). So, in the analysis, we focus on annual changes between 2003 and 2005. The data also contains information on worker's job function, so we are able to control for promotions.[12]

Table 1. *Summary statistics*

Variable	Mean	Std. Dev.	Minimum	Maximum
Change in commuting distance (km)	-0.4537	14.3492	-49.9630	49.8160
Abs. change in commuting distance (km)	9.2884	10.9461	0.0010	49.9630
log(wage2005)–log(wage2003)	0.0526	0.1841	-0.4994	0.4979
Change in workers function	0.4819	0.4997	0.0000	1.0000
Workers with commuting distance between 12.5 and 50 km in 2005 (share)	0.4381	0.4962	0.0000	1.0000
Workers with commuting distance > 50 km in 2005 (share)	0.0616	0.2405	0.0000	1.0000

Notes: Number of observations: 6,165.

Table 1 shows summary statistics of variables of interest. They show, as one may expect given random sampling, that the average change in commuting distance is close to zero. The average absolute change is 9.29 km, which is substantial compared to the average level of distance (17.50 km). So, we have a sufficient number of large exogenous changes in commuting distance. The share of workers entitled to a tax reduction (those with one way commute that exceeds 12.5 kilometres) is approximately 50%.

We have also calculated the correlation between changes in commuting distance and changes in wages, which appears to be 0.08. This is in line with a range of other studies (see e.g. Manning, 2003) although these studies include the change in distance and not only changes induced by firm relocations. The positive correlation suggests that variation in the commuting distance is important for determining variation in wages.

4. Empirical results

The econometric results of several specifications of first-differences models based on (2) are shown in Table 2. In these specifications, we initially do not correct for any sample selection effect. The first two columns show the results for a linear and a quadratic specification of commuting distance and where we also allow the distance effect to depend on whether the one-way commuting distance is between 12.5 and 50 kilometres or exceeds 50 kilometres, i.e. whether workers are entitled to an income tax reduction associated with commuting. The effect

[12] Variable 'change in workers function' is computed from labour market administrative register's variable RASDISCO, which is a 4-digit function code, including more than thousand different function descriptions. For some industries, particular for government sector, consulting etc., it is common practice that workers change function every 2nd or 4th year. This explains the high percentage of workers that change function from 2003 to 2005.

of commuting distance appears to be statistically significant and positive (within the relevant range) in all specifications. We also find evidence that the marginal effect is not linear, in line with the fact that income tax reduction apply to longer distances only and that the time cost of commuting is concave function of distance. Concavity makes sense. Speed of travel is strongly increasing in distance, implying that marginal costs of distance are smaller for longer distance. The quadratic specification [2] implies a marginal effect of 0.0049 for a minimal commuting distance. The marginal effect is only slightly lower at let's say 10 km, but substantially lower at the average commuting distance, where income tax reductions apply (see last two columns of Table 3). The marginal effect of commuting distance just above 12.5 where income tax reductions apply is 0.00190 (s.e. is 0.00041). It is positive up to 50 kilometres which applies to the large majority of observations (94%). For very long distances, the marginal time losses due to an increase in distance are too small to identify plausibly due to a high speed as well as income tax reductions. We do not reject the hypothesis that the marginal effect is zero at commuting distance just above 50 km suggesting that income tax reductions fully compensate for commuting time costs.

Both specifications imply that, for the commuting distance range where the income tax reduction does not apply, an increase in commuting distance by 1 kilometre induces on average a wage increase of 0.42%.[13] This is an economically significant effect. For example, if the commuting distance to a firm increases by 10 km, which is about the average change of a firm that relocates, wages increase by approximately 4.2%. This corresponds to 46.78 DKK per working day, or 2.34 DKK per additional kilometre travelled per day worked. Given a commuting speed of 35 km/h (this speed applies to distances of about 10 km), the compensation is about 81.86 DKK per hour, or 49.43% of the net hourly wage (which is 165.59 DKK on average). This estimate is likely an underestimate because it assumes that workers travel each day to their workplace, which is not the case due to business travel, teleworking and absenteeism.[14] Assuming that workers commute 90% of their workdays, the compensation will be closer to 54% of the hourly wage. Transport economists typically find that the value of time for commuters is about 50% of the wage (Small and Verhoef, 2007). This result seems to hold

[13] The effects are not different when we estimate the same models on sample including only workers with commuting distance below 12.5 kilometres for which income tax reductions are not applicable (see Appendix B).

[14] Absenteeism rates are about 3% and the sum of business travel and teleworking occur likely at similar rate.

for Denmark.[15] Monetary costs are typically of the same magnitude (Van Ommeren and Fosgerau, 2009). This result implies that workers bargain for about half a commuting costs. Our implicit estimate of the bargaining parameter is consistent with those reported in a number of papers in the labour market literature. Mortensen and Nagypal (2007) propose in their survey paper a value of 0.5 for this parameter.[16]

The marginal effect of commuting distance at the average commuting distance (17.5 km) is 0.0017. This corresponds to 1.9 DKK per working day, or 0.94 DKK per kilometre travelled per day worked. Given a commuting speed of 35 km/h, the compensation is now about 33.13 DKK per hour, or 22.01% of the net hourly wage, again about half of the commuting costs related to time losses.

Table 2. *First-difference wage model with firm fixed-effects*

	[1]	[2]	[3]	[4]	[5]
	all observations		abs. change in commuting distance>500m		
Change in commuting distance	0.00423*** (0.00066)	0.00494*** (0.00074)	0.00433*** (0.00070)	0.00502*** (0.00078)	
Change in commuting distance squared / 100		-0.00194** (0.00091)		-0.00189** (0.00095)	
Change in commuting distance (increase)					0.00443*** (0.00076)
Change in commuting distance (decrease)					0.00425*** (0.00074)
Change in commuting distance * $D_{12.5}$	-0.00263*** (0.00058)	-0.00256*** (0.00058)	-0.00264*** (0.00061)	-0.00258*** (0.00061)	-0.00265*** (0.00061)
Change in commuting distance * D_{50}	-0.00374*** (0.00064)	-0.00315*** (0.00070)	-0.00377*** (0.00068)	-0.00320*** (0.00073)	-0.00379*** (0.00068)
Change in worker's function	0.01912*** (0.00527)	0.01914*** (0.00526)	0.02418*** (0.00605)	0.02423*** (0.00605)	0.02425*** (0.00606)
Firm fixed effects (1,244)	yes	yes	yes	yes	yes
R-squared	0.3600	0.3606	0.3690	0.3696	0.3690
No. of observations	6165	6165	5085	5085	5085

Notes: Dependent variable is change in logarithm of wage; $D_{12.5}$ is dummy variable indicating if the worker one-way commuting distance is between 12.5 and 50 km; D_{50} is dummy variable indicating if the worker one-way commuting distance exceeds 50 km; ***,** indicate that estimates are significantly different from zero at the 0.01, and the 0.05 level, respectively; standard errors are in parentheses.

Our empirical results are consistent with those reported in a number of papers in the urban economics literature that examine the relationship between wages and commutes (but which ignore endogeneity issues as emphasized by Manning (2003)). For example, Madden (1985) investigates how wages vary with distance from the central business district (CBD). She regresses change in commuting distance on the change in wages for workers who changed job,

[15] Fosgerau et al. (2007) estimated value of time of 76 DKK per hour for Danish commuters for year 2005.
[16] Shimer (2005) proposes a value of 0.4 as a value for the worker's bargaining power parameter (based on the interpretation of this parameter as unemployment insurance), while Hall (2008) suggests 0.7 if one permits a broader interpretation of this variable. Cahuc et al. (2006) estimated the bargaining power of workers between 0 and 0.33.

changed residence, or both. For workers who changed job, she reports a positive relationship between wage and commuting distance changes.

The results are almost identical if one excludes observations referring to changes in commuting distance smaller than 500 meters (see the last three columns in Table 2).[17] We have also estimated models that distinguish between different effects of increases and decreases in commuting distance (see the last column in Table 2). A F-test (F=0.1403; p-value=0.7080) does not reject the null hypothesis that these effects are identical. As nominal wage decreases are extremely uncommon for workers who stay with the same firm, this indicates that workers with reduced distances receive smaller nominal wage increases than other workers in the same firm.

We have estimated the same models on other, less-selective, samples of data. So we have included data on (i) workers who change employer, (ii) workers who change residence, and (iii) workers who change both employer and residence. The effects of commuting distance, reported in Table 3, are very similar to the results reported in Table 2, except for a sample that includes residence change, most likely because distance changes of residential movers are compensated on the housing market (in line with theory, see Zenou, 2009).[18]

Table 3. *First-difference wage model. Effect of distance.*

Sample	N	Change in commuting distance	Change in commuting distance squared/100	Change in commuting distance * $D_{12.5}$	Wage effect at commuting distance of 10km	Wage effect at average commuting distance (17.5km)
Sample used for Table 2	6,165	0.00494*** (0.00074)	-0.00194** (0.00091)	-0.00256*** (0.00058)	0.00455*** (0.00074)	0.00170*** (0.00041)
Sample including employer changes	7,248	0.00431*** (0.00067)	-0.00222** (0.00082)	-0.00197*** (0.00052)	0.00387*** (0.00067)	0.00156*** (0.00037)
Sample including residence changes	7,338	0.00109*** (0.00026)	0.00056 (0.00070)	-0.00024 (0.00032)	0.00120*** (0.00026)	0.00105*** (0.00030)
Sample including employer and residence changes	8,601	0.00110*** (0.00024)	0.00029 (0.00064)	-0.00018 (0.00030)	0.00116*** (0.00024)	0.00102*** (0.00027)
Heckman selection model (selection regarding residence change)	8,601	0.00477*** (0.00067)	-0.00221** (0.00081)	-0.00239*** (0.00052)	0.00433*** (0.00067)	0.00161*** (0.00037)
Heckman selection model (selection regarding employer change)	8,601	0.00535*** (0.00074)	-0.00194** (0.00091)	-0.00294*** (0.00058)	0.00496*** (0.00073)	0.00173*** (0.00041)
Heckman selection model (selection regarding residence and employer changes)	8,601	0.00480*** (0.00068)	-0.00197** (0.00082)	-0.00247*** (0.00053)	0.00441*** (0.00067)	0.00164*** (0.00037)

Notes: ***,** indicate that estimates are significantly different from zero at the 0.01, and the 0.05 level, respectively; standard errors are in parentheses.

An alternative, and usually better, way to examine selection sample issues is to estimate Heckman selection models. The inclusion of instruments in the first step of the models is based on the presence of search frictions in labour and housing markets. Given these frictions, the

[17] These changes in distances are usually economically of no importance, but are relatively common in our data (18% of observations).

[18] The full results of the estimates are provided in Appendix C (Table C1- C3).

spatial configuration of jobs and residence affects job and residential mobility (Manning, 2003). It seems however reasonable to assume that the spatial configuration of jobs and residences does not directly affect annual changes in wages, so we use this configuration as an instrument. For single-wage earners, the spatial configuration is captured by the commuting distance. For two-earner households, it is captured by three variables: the commuting distances of both wage earners as well as the distance between the workplaces of the wage earners.[19] In Table 3, we report the results of Heckman selection models using the spatial configuration of jobs and residence as instrument. Accounting for sample selectivity in this way does not change our main result.[20] We have also estimated Heckman selection models applying another set of instruments. We control for number of children up to 12 years old, but use the presence of children in the age between 12 and 18 as an instrument. Children in this age group likely have no direct effect on changes of wages, but strongly affect residential, but also job, mobility. The results obtained from Heckman selection models applying these instruments are almost identical to the results presented above (see Appendix D).

5. Conclusion

This paper analyses the *causal* effect of commuting distance on wages using matched register data for firms and workers for Denmark. We deal with the endogeneity of commuting distance by means of an innovative approach using changes in commuting distance that are due to firm relocations and therefore exogenous. We take into account that above 12.5 kilometres income tax reductions apply. We show that, for the commuting distance range where the income tax reductions associated with commuting do not apply, commuting distance increases imply an overall hourly wage compensation of about 49% on average. The effect of commuting distance at the average commuting distance, where income tax reduction is applicable, is much lower, i.e. about 22%. The effect becomes zero at commuting distance of 50 kilometres. The estimated positive effect of change in commuting distance on wages is consistent with wage bargaining, and due to the quasi-natural experimental setup, excludes other competing explanations. Our

[19] Deding et al. (2009) hypothesize that residential mobility depends positively on the commuting distances of both spouses, but negatively on the distance between workplaces. Further, workers' job mobility depends positively on the worker's commuting distance, negatively on the spouse's commuting distance, and positively on the distance between workplaces. Using data for Denmark, Deding et al. (2009) show that these hypotheses hold, and that the effects of spatial configuration are rather large.

[20] The full results of the Heckman selection models are provided in Appendix D.

results imply wage bargaining parameter of about 0.5 for both the range in commuting distance where the income tax reduction associated with commuting is not applicable and for commuting distances where income tax reduction is applicable. The results appear robust with specification and accounting for selection effects.

Our findings have a number of implications. First it demonstrates that wage bargaining with respect to commuting is an important characteristic of the (Danish) labour market, in line with range of theoretical models (Marimom and Zilibotti, 1999). So it is able to demonstrate that employer have some labour market power and pay below workers productivity (Pissarides, 2000). Second, evidence of a wage-commute relationship puts a price on commuting distance and points to the economic benefits from transport infrastructure improvements.

Acknowledgements

The authors thank Mogens Fosgerau for comments. We are grateful to Statistics Denmark for providing the data. Research support from the Danish Energy Agency (Energy Research Programme) is acknowledged.

Appendix A

Table A1. *Sample selection procedure*

		Workers	Firms
1	Address changes, total	64,643	11,314
2	Workers with more than 1 job, part-time workers and missing information	44,918	9,279
2.1	*Missing information regarding number of jobs*	*1,713*	*638*
2.2	*Missing information regarding part-time / full-time job*	*3,404*	*3,178*
2.3	*Missing information regarding worker's wage*	*740*	*393*
2.4	*More than 1 job (in the last year)*	*15,576*	*3,122*
2.5	*Workers without full-time job (that last at least 1 year, continuously)*	*23,485*	*1,948*
3=1-2	Address changes	19,725	2,035
4	Change in commuting distance = 0	8,338	337
5=3-4		*11,387*	*1,698*
6	Commuting distance > 100 km	878	179
7=5-6		*10,509*	*1,519*
8	Change in log(wage)>0.5	1,474	167
9=7-8		*9,035*	*1,352*
10	Change in commuting distance > 50 km	434	19
11=9-10	**Full sample**	***8,601***	***1,333***
12	Employer change	1,263	66
13	Change in residence	1,353	132
14=11-12	***Sample 1 (exclude employer change)***	***7,338***	***1,267***
15=11-13	***Sample 2 (exclude change in residence)***	***7,248***	***1,201***
16	**Sample 3 (exclude residence and employer changes)**	**6,165**	**1,244**

Appendix B

Table B1. *First-difference wage model with firm fixed-effects (commuting distance < 12.5 km)*

	[1]	[2]	[3]
	all observations	abs. change in commuting distance>500m	
Change in commuting distance	0.00504*** (0.00108)	0.00507*** (0.00507)	
Change in commuting distance (increase)			0.00407 (0.00261)
Change in commuting distance (decrease)			0.00585*** (0.00220)
Change in worker's function	0.03034*** (0.00907)	0.03944*** (0.01228)	0.03870*** (0.01241)
Firm fixed effects (1,244)	yes	yes	yes
R-squared	0.4622	0.4932	0.4933
No. of observations	2097	1598	1598

Notes: as for Table 2.

Appendix C

Table C1. *First-difference wage model with firm fixed-effects. Sample includes employer and residence changes*

	[1]	[2]	[3]	[4]	[5]
	all observations		excl.|change in commuting distance|<500m		
Change in commuting distance	0.00111*** (0.00024)	0.00110*** (0.00024)	0.00115*** (0.00025)	0.00114*** (0.00025)	
Change in commuting distance squared / 100		0.00029 (0.00064)		0.00032 (0.00067)	
Change in commuting distance (increase)					0.00135*** (0.00030)
Change in commuting distance (decrease)					0.00081** (0.00033)
Change in commuting distance * $D_{12.5}$	-0.00009 (0.00022)	-0.00018 (0.00030)	-0.00008 (0.00023)	-0.00018 (0.00031)	-0.00002 (0.00023)
Change in commuting distance * D_{50}	-0.00080*** (0.00024)	-0.00097** (0.00046)	-0.00082*** (0.00025)	-0.00101** (0.00048)	-0.00077*** (0.00025)
Change in worker's function	0.02748*** (0.00388)	0.02748*** (0.00388)	0.03005*** (0.00445)	0.03005*** (0.00445)	0.02926*** (0.00449)
Firm fixed effects (1,333)	yes	yes	yes	yes	yes
R-squared	0.2836	0.2836	0.2871	0.2871	0.2872
No. observations	8601	8601	7234	7234	7234

Notes: as for Table 2.

Table C2. *First-difference wage model with firm fixed-effects. Sample includes employer changes*

	[1]	[2]	[3]	[4]	[5]
	all observations		excl.|change in commuting distance|<500m		
Change in commuting distance	0.00347*** (0.00060)	0.00431*** (0.00067)	0.00355*** (0.00062)	0.00437*** (0.00070)	
Change in commuting distance squared / 100		-0.00222** (0.00082)		-0.00216** (00085)	
Change in commuting distance (increase)					0.00386*** (0.00067)
Change in commuting distance (decrease)					0.00331*** (0.00065)
Change in commuting distance * $D_{12.5}$	-0.00203*** (0.00052)	-0.00197*** (0.00052)	-0.00205*** (0.00054)	-0.00200*** (0.00054)	-0.00207*** (0.00054)
Change in commuting distance * D_{50}	-0.00298*** (0.00057)	-0.00235*** (0.00061)	-0.00302*** (0.00060)	-0.00241*** (0.00064)	-0.00306*** (0.00060)
Change in worker's function	0.02563*** (0.00418)	0.02573*** (0.00417)	0.02789*** (0.00490)	0.02804*** (0.00490)	0.02708*** (0.00494)
Firm fixed effects (1,201)	yes	yes	yes	yes	yes
R-squared	0.3033	0.3041	0.3131	0.3140	0.3133
No. observations	7248	7248	6153	6153	6153

Notes: as for Table 2.

Table C3. *First-difference wage model with firm fixed-effects. Sample includes residence changes*

	[1]	[2]	[3]	[4]	[5]
	all observations		excl.|change in commuting distance|<500m		
Change in commuting distance	0.00111*** (0.00026)	0.00109*** (0.00026)	0.00117*** (0.00028)	0.00115*** (0.00028)	
Change in commuting distance squared / 100		0.00056 (0.00070)		0.00057 (0.00073)	
Change in commuting distance (increase)					0.00134*** (0.00035)
Change in commuting distance (decrease)					0.00096** (0.00038)
Change in commuting distance * $D_{12.5}$	-0.00007 (0.00024)	-0.00024 (0.00032)	-0.00007 (0.00025)	-0.00024 (0.00034)	-0.00004 (0.00025)
Change in commuting distance * D_{50}	-0.00078*** (0.00027)	-0.00112** (0.00050)	-0.00081*** (0.00028)	-0.00116** (0.00053)	-0.00079*** (0.00028)
Change in worker's function	0.02223*** (0.00465)	0.02221*** (0.00465)	0.02847*** (0.00524)	0.02844*** (0.00524)	0.02842*** (0.00524)
Firm fixed effects (1,267)	yes	yes	yes	yes	yes
R-squared	0.3190	0.3191	0.3219	0.3220	0.3220
No. observations	7338	7338	6221	6221	6221

Notes: as for Table 2.

Appendix D

Table D1. *Heckman selection models. Estimates of logarithm of changes in wage with changes in commuting distance, firm fixed-effects.*

		[1]	[2]	[3]	[4]	[5]	[6]
Logarithm of changes in wage	Change in commuting distance	0.00477***	0.00535***	0.00480***	0.00479***	0.00531***	0.00480***
		(0.00067)	(0.00074)	(0.00068)	(0.00067)	(0.00074)	(0.00068)
	Change in commuting distance squared / 100	-0.00221**	-0.00194**	-0.00197**	-0.00214**	-0.00192**	-0.00194**
		(0.00081)	(0.00091)	(0.00082)	(0.00081)	(0.00091)	(0.00082)
	Change in commuting distance * $D_{12.5}$	-0.00239***	-0.00294***	-0.00247***	-0.00238***	-0.00291***	-0.00246***
		(0.00052)	(0.00058)	(0.00053)	(0.00052)	(0.00058)	(0.00053)
	Change in commuting distance * D_{50}	-0.00288***	-0.00354***	-0.00304***	-0.00287***	-0.00351***	-0.00304***
		(0.00063)	(0.00069)	(0.00064)	(0.00063)	(0.00069)	(0.00064)
	Change in worker's function	0.01819***	0.01566***	0.01840***	0.01841***	0.01482***	0.01837***
		(0.00471)	(0.00513)	(0.00478)	(0.00472)	(0.00514)	(0.00478)
	Dummy indicating 1 child				0.01871***	0.01370**	0.01713**
					(0.00622)	(0.00656)	(0.00637)
	Dummy indicating 2 children				0.01457**	0.00811	0.01271**
					(0.00628)	(0.00609)	(0.00634)
	Dummy indicating 3 or more children				0.01909*	0.01715*	0.01712*
					(0.01004)	(0.01058)	(0.01007)
Dependent variable is dummy variable that defines the selection	Change in commuting distance	-0.01481***	-0.01265**	-0.01184***	-0.01228***	-0.01291***	-0.01085***
		(0.00179)	*(0.00503)*	(0.00163)	(0.00176)	(0.00501)	(0.00160)
	Change in commuting distance squared / 100	-0.00678	*0.03425*	-0.00226	-0.08096	0.00244	-0.00211**
		(0.00515)	*(0.00629)*	(0.00445)	(0.05338)	(0.06126)	(0.00449)
	Change in commuting distance * $D_{12.5}$	0.00649**	*0.00974***	0.00674***	0.00809***	0.00956**	0.00708***
		(0.00235)	*(0.00387)*	(0.00203)	(0.00239)	(0.00387)	(0.00204)
	Change in commuting distance * D_{50}	0.00825**	*0.01076***	0.00773***	0.01043**	0.01060**	0.00830***
		(0.00373)	*(0.00459)*	(0.00317)	(0.00385)	(0.00456)	(0.00319)
	Change in worker's function	-0.10028**	*-0.06952***	-0.05628**	-0.11304***	-0.06816**	-0.05311*
		(0.03578)	*(0.03290)*	(0.02867)	(0.03609)	(0.03278)	(0.02871)
	Commuting distance (inst.)	**-0.00704*** **	***0.00237*** **	**-0.00284*** **			
		(0.00115)	***(0.00098)***	**(0.00096)**			
	Commuting distance for spouse (inst.)	**0.00006**	***0.00095***	**0.00117**			
		(0.00102)	***(0.00089)***	**(0.00082)**			
	Distance between workplaces (inst.)	**0.00432*** **	***-0.00277*** **	**0.00002**			
		(0.00093)	***(0.00053)***	**(0.00058)**			
	Dummy indicating 1 child				0.26229***	0.05324	0.19066***
					(0.05211)	(0.04509)	(0.04101)
	Dummy indicating 2 children				0.47242***	0.02905	0.27109***
					(0.05468)	(0.04296)	(0.04086)
	Dummy indicating 3 or more children				0.44445***	0.01332	0.22748***
					(0.10289)	(0.07383)	(0.07219)
	Dummy indicating presence of children 12-18 years old **(inst.)**				***0.20456*** **	***-0.11123*** **	***0.04354***
					(0.05206)	***(0.03151)***	***(0.03554)***
	Intercept	1.13476***	0.86710***	0.60122***	0.88097***	0.87916***	0.45019***
		(0.03458)	(0.02972)	(0.02774)	(0.02889)	(0.02761)	(0.02364)
	Sigma	*0.15506****	*0.18177****	*0.15525****	*0.15378****	*0.18249****	*0.15441****
		(0.00183)	*(0.00244)*	*(0.00210)*	*(0.00159)*	*(0.00241)*	*(0.00179)*
	Rho	*0.23939****	*-0.83025****	*0.16996*	*0.13022*	*-0.83784****	*0.10310*
		(0.08353)	*(0.01636)*	*(0.11363)*	*(0.10630)*	*(0.01529)*	*(0.13848)*
	Log Likelihood	-355	-882	-2481	-283	-887	-2441
	No. Observations	8601	8601	8601	8601	8601	8601

Notes: (1) Columns [1] and [4] account for selection regarding residence change; columns [2] and [5] account for selection regarding employer change, columns [3] and [6] account for selection regarding both changes.

(2) Max.likelihood method is used to estimate the model. The coefficient associated with inverse Mills ratios is defined as σ multiplied with ρ.

(3) ***,**,* indicate that estimates are significantly different from zero at the 0.01, at the 0.05 and the 0.10 level, respectively.

(4) Standard errors are in parentheses.

(5) All explanatory variables in the selection equation are in 2003, so prior to the (possible) move.

(6) Instruments are indicated with bold type.

References

Benito, A. and Oswald, A., 1999. Commuting in Great Britain in the 1990s Warwick. Department of Economics, University of Warwick.

Cahuc, P., Postel-Vinay, F. and Robin, J.-M., 2006. Wage bargaining with on-the-job search: theory and evidence. Econometrica, Vol. 74, No. 2, 323-364.

Deding, M., Filges, T. and Van Ommeren, J.N., 2009. Spatial mobility and commuting: the case of two-earner households. Journal of Regional Science, Vol. 49, no. 1, pp. 113-147.

Fosgerau, M., Hjorth, K. and Lyk-Jensen, S. V., 2007. The Danish Value of Time Study. Report 5, Danish Transport Research Institute.

Fujita, M., Thisse, J-F. and Zenou, Y., 1997. On the endogeneous formation of secondary employment centers in a city. Journal of Urban Economics, 41, 337-357.

Gertler, M., Sala, L. and Traigari, A., 2008. An estimated monetary DSGE model with unemployment and staggered nominal wage bargaining. Journal of Money, Credit and Banking, Vol. 40, No. 8, 1713-1764.

Gibbons, S. and Machin, S., 2006. Transport and labour market linkages: empirical evidence, implications for policy and scope for further UK research. Commissioned for the Eddington Study.

Gubits, D.B., 2004. Commuting, work hours, and the metropolitan labor supply gradient. Mimeo.

Gutiérrez-i-Puigarnau, E. and Van Ommeren, J.N., 2010. Labour supply and commuting. Journal of Urban Economics, 68 (1), 82-89.

Hall, R., 2008. Sources and mechanisms of cyclical fluctuations in the labor market. Mimeo, Stanford University.

Iversen, T. 1996. Power, flexibility, and the breakdown of centralized wage bargaining: Denmark and Sweden in comparative perspective. Comparative Politics, Vol. 28, No. 4, 399-436.

Madden, J.F., 1985. Urban wage gradients: empirical evidence. Journal of Urban Economics, 18, 291–301.

Madsen, P. K., 2002. The Danish model of “flexicurity” – A paradise with some snakes. Interactions between Labour Market and Social Protection, Brussels May 16, 2002.

Manning, A., 2003. The real thin theory: monopsony in modern labour markets. Labour Economics 10 (2), 105–31.

Marimom, R. and Zilibotti, F., 1999. Unemployment versus mismatch of talents: Reconsidering unemployment benefits. Economic Journal, 109, 266–291.

Mortensen, D., 2001. How Monopsonistic is the (Danish) Labor Market?, Mimeo, Northwestern University.

Mortensen, D. and Nagypal, E., 2007. More on unemployment and vacancy fluctuations. Review of Economic Dynamics, 10:3, 327–47.

Pissarides, C. A., 2000. Equilibrium Unemployment Theory. second ed., MIT Press, Cambridge.

Potter , S., Enoch, T., Black, C. and Ubbels, B., 2006. Tax treatment of employer commuting support: an international review. Transport Reviews, 26, 2 , 221 – 23.

Simonsohn, U., 2006. New Yorkers commute more everywhere: contrast effects in the field. Review of Economics and Statistics, 88, 1-9.

Shimer, R., 2005. The cyclical behavior of equilibrium unemployment, vacancies, and wages: evidence and theory. American Economic Review, 95:1, 25–49.

Small, K. A. and Verhoef, E. T., 2007. The economics of urban transportation. Routledge.

Timothy, D. and Wheaton, W., 2001. Intra-urban wage variation, employment location and commuting times. Journal of Urban Economics, 20(2), 338-366.

Van Ommeren, J.N. and Fosgerau, M., 2009. Workers' marginal costs of commuting. Journal of Urban Economics, vol. 65, issue 1, 38-47.

Van Ommeren, J.N. and Rietveld, P., 2005. The commuting time paradox. Journal of Urban Economics 58 (3), 437–54.

Wheaton, W.C., 1974. A comparative static analysis of urban spatial structure. Journal of Economic Theory, 9, 223–237.

White, M. J., 1977. A model of residential location choice and commuting by men and women workers. Journal of Regional Science, 17, 41–52.

Zax, J.S., 1991. The substitution between moves and quits. The Economic Journal, 101 (409), 1510–21.

Zax, J.S. and Kain, J.F., 1996. Moving to the suburbs: do relocating companies leave their black employees behind?. Journal of Labor Economics, 14 (3), 472–504.

Zenou, Y., 2009. Urban Labor Economics. Cambridge University Press, New York.

Chapter 2

Household transport consumption inequalities and redistributive effects of taxes: a repeated cross-sectional evaluation for France, Denmark and Cyprus

Akli Berri, Stéphanie Vincent Lyk-Jensen, Ismir Mulalic, and Theodoros Zachariadis

Household transport consumption inequalities and redistributive effects of taxes: a repeated cross sectional evaluation for France, Denmark and Cyprus

Akli Berri
French Institute of Science and Technology for Transport, Development and Networks (IFSTTAR)
2, rue de la Butte Verte, FR-93166 Noisy-le-Grand cedex, France
akli.berri@ifsttar.fr

Stéphanie Vincent Lyk-Jensen
The Danish National Centre for Social Research (SFI)
Herluf Trolles Gade 11, DK-1052 Copenhagen K, Denmark
svj@sfi.dk

Ismir Mulalic
University of Copenhagen, Øster Farimagsgade 5, Building 26, 1353 Copenhagen K, Denmark, and
Technical University of Denmark, Bygningstorvet 116 Vest, 2800 Kgs. Lyngby, Denmark
imu@transport.dtu.dk

Theodoros Zachariadis
University of Cyprus, Economics Research Centre, P.O. Box 20537, CY-1678 Nicosia, Cyprus
t.zachariadis@cut.ac.cy

Abstract

We evaluate household transport consumption inequalities in France, Denmark and Cyprus, investigate their temporal dynamics, and estimate the redistributive effects of taxes on different commodity categories. Using household-level data from repeated cross-sections of expenditure surveys spanning long time periods, the paper applies a decomposition of the Gini index by expenditure component. The results highlight the effect of the social diffusion of the car. The relative contribution of vehicle use items to total expenditure inequality decreases over time, thus reflecting the increasingly widespread use of cars. Moreover, fuel taxes become regressive, while the progressive character of taxes on the remaining car use commodities weakens with time. Taxes on transport goods and services as a whole are progressive. However, this result is principally due to the progressivity of taxes on car purchases, a progressivity stronger by far in Denmark, where these taxes are so high that car purchase costs can be afforded only by those with high incomes. These findings demonstrate that policy makers must not overlook equity issues when designing policies to attenuate the environmental impact of cars. Increasing car use costs, notably fuel prices, through an increase of uniform taxes would be particularly inequitable.

Keywords: Inequality, transport consumption, household expenditure surveys, Gini index, decomposition by component, redistributive effects of taxes.

JEL codes: D12, H23, H24, R41.

1. Introduction

Car taxes are a source of public revenues, as well as a policy tool for reducing the adverse impacts of road traffic. Most of them were instituted at a time when the car was a luxury good (e.g. the 1924 Danish registration tax and the 1956 French vignette, an annual tax on vehicles owned). The large social diffusion of this good over the past six decades has doubtless lessened the progressivity of these taxes. The protests in several European countries against the rapid increase in fuel prices in the autumn of 2000 highlighted the households' sensitivity to the burden of fuel costs, and particularly the households who live in rural and suburban areas and who are more car-dependent.

This paper evaluates inequalities between households in the consumption of transport goods and services in France, Denmark and Cyprus, investigates their temporal dynamics, and estimates the redistributive effects of taxes on the different commodity categories considered. These three countries constitute an interesting field for comparison, as Cyprus has very poorly developed public transport, Denmark has highly developed public transport, with heavy car-related taxes, and France produces cars and therefore has an investment in the car as a means of transport. The paper carries out a comparative analysis in light of the differences between these countries, most notably those of car taxation systems and car ownership levels. Consumption is measured in terms of expenditures collected through budget surveys. As Deaton (1997) writes, by revealing who buys each good or service and the amounts spent, expenditure surveys show who bears most of the corresponding tax burden (especially by income level) and thus who are the potential losers and gainers from possible taxation changes.

Our analysis applies a decomposition of the Gini inequality indicator by expenditure component. Each component appears through its proper Gini coefficient, its budget share and its degree of association with total expenditure. This method provides a good understanding of the inequality mechanisms, in particular their temporal evolution. Moreover, it allows us to evaluate the redistributive effect of (a change in) a tax on a good or a service, i.e. its impact in terms of inequality increase or reduction. Finally, the analysis provides estimates of elasticities with respect to total expenditure (or income) without specifying a functional form for the Engel curves.

The data are from repeated cross-sections of household expenditure surveys. For each country, a few distant survey periods (about 10 years apart) are selected among the accessible data sets: two for Cyprus (1991 and 2003), two for Denmark (1997 and 2005), and three for France (1978-79, 1989 and 2000-01). The number of surveyed Cypriot households is 2,708 in 1991 and 2,990 households in 2003; the number of surveyed Danish households is 881 in 1997 and 725 in 2005; and the number of surveyed French households is 10,645 in 1978-79, 9,038 in 1989 and 10,305 in 2000-01.

After an exposition of the methodology of Gini decomposition by expenditure component in Section 2, Section 3 presents some of the characteristics of the car taxation systems in the three countries and examines the budget shares allocated to different expenditure groups according to a household's standard of living. Section 4 presents the results of the analyses of inequality and redistributive effects of taxes on the different categories of goods and services considered. Section 5 summarises the findings and concludes.

2. Decomposition of the Gini coefficient by component and redistributive effects of marginal changes in components

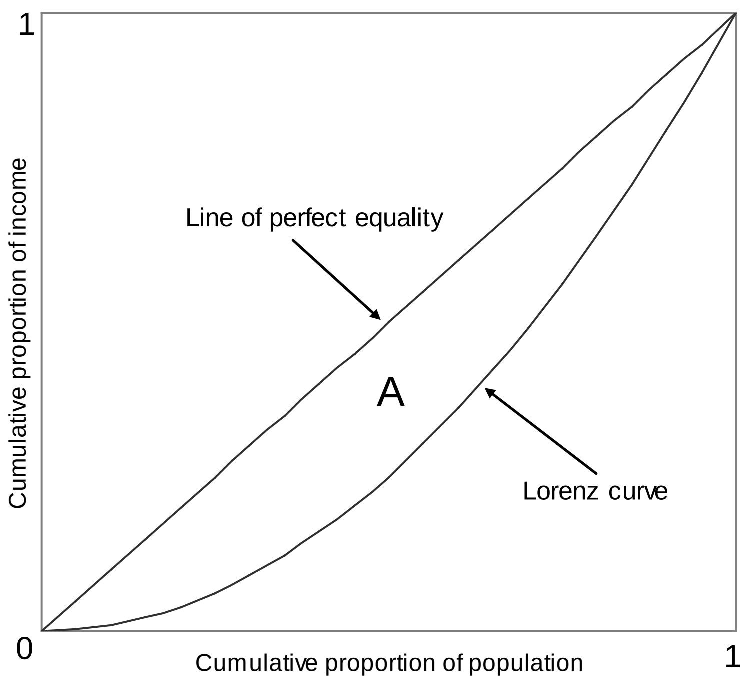
2.1. The Gini inequality index

The Gini coefficient is one of the more widely used indicators for evaluating inequalities (e.g. in income, wealth, consumption). A graphic visualisation of this index is based on the Lorenz curve, shown in Figure 1. The Lorenz curve of income, for instance, is constructed by arranging individuals from the poorest to the richest, and then representing their cumulative share of total income as a function of their cumulative proportion in the population.

If each individual had the same income, the curve would coincide with the main diagonal, the income share of a given group being equal to its weight in the total population. Apart from the case of perfect equality, the groups with the lowest incomes enjoy a share of total income lower than their weight in the population. Consequently, except in the case of perfect equality, every Lorenz curve lies below the main diagonal and its slope increases (in any case, it does not decrease) when approaching the highest incomes.

This graphic tool plays an important role in the characterisation of the robustness of inequality measures as to ranking distributions (Atkinson, 1970; Deaton, 1997). Thus, if the Lorenz curve of a distribution Y lies everywhere below that of another distribution X, Y is less egalitarian than X. Indeed, the distribution Y can be transformed by a series of transfers from the richer to the poorer in such a way to obtain the distribution X. Consequently, when two Lorenz curves do not cross, the upper one represents a more egalitarian distribution and will show a lower inequality level, provided that the inequality measure used satisfies the principle of transfers. The *principle of transfers* is stated as follows: if one transfers an amount d from a person having an income y_1 to another person having a lower income y_2 (with $y_2 \leq y_1 - d$, such that the transfer does not reverse their relative positions), then the new distribution should be preferred to the initial one. In the case where two Lorenz curves intersect, unambiguously ranking the corresponding distributions as to their degree of inequality is not possible without restraining choice to specific inequality measures.

Figure 1. *Illustration of a Lorenz curve*



The Gini coefficient is defined as the ratio of the area between the Lorenz curve and the main diagonal (designated by A in figure 1) to the area of the triangle below the diagonal (i.e. $1/2$), that is $G = 2A$. When the distribution is perfectly egalitarian, its Lorenz curve coincides with the diagonal, hence $A = 0$ and $G = 0$. Absolute inequality implies that A is the whole area of the triangle under the diagonal, so $A = 1/2$ and $G = 1$. Thus the Gini coefficient takes values between 0 and 1.

One of the advantages of the Gini coefficient as a measure of (income) inequality is that it is ‘a very direct measure of income difference, taking note of differences between every pair of incomes’ (Sen, 1997, p. 31). Indeed, one of the Gini expressions is based on the *average of absolute differences between pairs of observations*, called *Gini’s Mean Difference* (GMD):

$$\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}|y_i - y_j|,$$

where y_i is the income of individual i and N is the number of individuals in the population.

The Gini coefficient is defined as the GMD divided by twice the mean (m):

$$G = \frac{1}{2N^2 m}\sum_{i=1}^{N}\sum_{j=1}^{N}|y_i - y_j| \qquad (1)$$

Another advantageous feature of the Gini coefficient is that it handles negative values, a feature that is notably useful for its decomposition by income source, where taxes are considered as ‘negative incomes’ (Lerman and Yitzhaki, 1994).

Besides equation (1), the Gini coefficient has several expressions (Sen, 1997). In the following section, we adopt a formulation that is easy to apply directly to individual data. This formulation is used to obtain a decomposition by the constituents of the variable of interest. The decomposition makes explicit the mechanisms by which each component contributes to the global Gini and therefore sheds light on the temporal patterns of inequalities. Moreover, the decomposition allows us to evaluate the redistributive effects of taxes on the different components.

2.2. A practical formulation of the Gini coefficient

Lerman and Yitzhaki (1984) show that the Gini coefficient can be expressed as a function of the covariance between the variable of interest (X) and its cumulative distribution (F_X), and of its mean (m):

$$G(X) = \frac{2cov(X, F_X)}{m}. \qquad (2)$$

Estimation of the Gini coefficient using this formulation is easy to apply to individual survey data. Indeed, one only has to estimate the mean of X and the covariance between X and its empirical cumulative distribution, and to substitute for the corresponding terms in the expression above. With a random sample (same selection probability for all individuals) of size n, the cumulative distribution is estimated by ranking individuals according to increasing values of X

and by dividing their ranks i by the sample size, i.e. $\tilde{F}_X = I/n$, and the mean is estimated by $\tilde{m} = \sum_i x_i/n$. For a non-random sample (selection probability varying from one individual to another), the observations have to be weighted by the respective individual survey weights, w_i. The cumulative distribution and the mean of X are estimated as follows:

$$\hat{F}_i(x) = \sum_{j=0}^{i-1} \pi_j + \frac{\pi_i}{2}, with\ \pi_0 = 0, \text{ and}$$

$$\hat{m} = \sum_{j=1}^{n} \pi_i x_i,$$

where $\pi_j = w_j / \sum_{i=1}^{n} w_i$.[1]

By avoiding the usual practice of grouping data prior to estimation, this approach yields estimates that are more accurate and free of the (downward) bias from aggregation. Lerman and Yitzhaki (1989) show that this bias increases with the aggregation level and with the value of the Gini coefficient.

2.3. Decomposition of the Gini coefficient by component

Lerman and Yitzhaki (1985) use this covariance-based formulation to obtain a decomposition of the Gini coefficient by the constituents of X and apply it to the analysis of the effects of income sources on the global income inequality. Garner (1993) applies it to the analysis of inequalities in terms of expenditures.

Consider the case where X is household's total expenditure. Let $x_1, x_2, \ldots, x_k, \ldots, x_K$ be the amounts spent on the K budget components, such that:

$$X = \sum_{k=1}^{K} x_k\ . \tag{3}$$

Then, using the additivity property of covariance, the Gini coefficient of X can be written:

$$G(X) = 2\sum_{k=1}^{K} \frac{cov(x_k, F_X)}{m} \tag{4}$$

Let F_k and m_k be the cumulative distribution and the mean of x_k, respectively. Multiplying and dividing each term in k in equation (4) by $cov(x_k, F_k)$ and by m_k, one obtains the decomposition by component:

$$G(X) = \sum_{k=1}^{K} \left[\frac{cov(x_k, F_X)}{cov(x_k, F_k)}\right] \left[\frac{2cov(x_k, F_k)}{m_k}\right] \left[\frac{m_k}{m}\right] \tag{5}$$

[1] A derivation of $\hat{F}_X$ appears in Berri (2005), pp. 246-248.

Denoting the first term of the sum by R_k, the second by G_k, and the third by S_k, the Gini coefficient can be written:

$$G(X) = \sum_{k=1}^{K} R_k G_k S_k \qquad (6)$$

where R_k is the *Gini correlation coefficient* between expenditure k and total expenditure, G_k is the Gini coefficient of component k, and S_k is its budget share. A high Gini correlation for a category of goods and services means that expenditure devoted to this category is higher the higher the total budget. Gini correlation is a measure of association based on Gini's Mean Difference (Schechtman and Yitzhaki, 1987). The Gini correlation between two variables takes values between -1 and 1. It is equal to zero if the two variables are independent. For example, if one of the variables is a monotonously increasing function of the other, their Gini correlation will be equal to 1. Further details appear in the Appendix.

Thus the contribution of an expenditure category to total inequality is determined by three terms: its proper Gini coefficient, its average budget share, and the degree of its association with total expenditure (measured by their Gini correlation). The higher the value of each of the factors, the stronger the contribution of the category to total inequality. The expression of the contribution also means that a high Gini coefficient does not guarantee a large contribution to total inequality. As we will show, because of a very low budget share the contribution of the item 'two-wheeler purchases' is the lowest among the categories considered, despite its Gini coefficient being the highest.

This approach is advantageous in that it provides a decomposition of inequalities into elements easily interpretable and helps explain their temporal evolution by examining the evolution of the elements involved in the contribution of each component. Moreover, it avoids a major shortcoming of the usual method called 'before-after'. The latter consists of calculating an inequality index after excluding a particular component and comparing it with the value of the index when this component is included. The results of this method may depend on the order in which the components are considered. For instance, in the case of two income sources, Lerman (1999) shows that a component will appear to reduce or increase inequalities according to whether one accounts for that component before or after including the other component.

2.4. Redistributive effects of marginal changes in the components

Another advantage of this decomposition is that it allows us to evaluate the redistributive effects of marginal changes in the different expenditure categories. No explicit income transfer is considered here. The expression 'redistributive effect' refers to the impact in terms of increase or reduction of inequality.

Suppose that the expenditure on a particular item k undergoes a small percentage variation, e_k, identical for all households (e.g. a tax), such that $x_k(e_k) = (1 + e_k)x_k$, $e_k > 0$. In terms of variation of a tax t_k on expenditure k, one has $x_k(dt_k) = (1 + dt_k)x_k$. The initial rate t_k does not appear, its effect being incorporated in the observation on x_k. The tax change is imposed on the expenditure made, x_k, which is equivalent to a tax proportional to the price paid by the consumer. The effect on the global Gini is (Stark et al., 1986):

$$\frac{\partial G}{\partial e_k} = S_k(R_k G_k - G) , \tag{7}$$

the terms S_k, R_k, G_k and G being evaluated before the marginal variation in component k takes place. Dividing by G, one obtains:

$$\frac{\partial G / \partial e_k}{G} = \frac{R_k G_k S_k}{G} - S_k . \tag{8}$$

Equation (8) shows that the relative variation of the global Gini due to a small variation in expenditure on component k is equal to the relative contribution of the component to overall inequality minus its contribution to total expenditure. The sum of all relative marginal effects equals 0. Multiplication by $(1 + e)$ of all components leaves the overall Gini unchanged. It can also be seen that as long as the budget share S_k is not zero:

A. the relative marginal effect is negative if the Gini correlation between expenditure k and total expenditure is negative or null ($R_k \leq 0$);

B. if the Gini correlation is positive, the impact on inequality depends on the sign of $(R_k G_k - G)$. A necessary condition for this term to be positive is that the inequality of component k exceeds that of total expenditure: $G_k > G$ (since $R_k \leq 1$).

Equation (8) defines the concept of progressivity used here (Yitzhaki, 1997). A tax will be called progressive if an increase in this tax (or its imposition if it does not yet exist) reduces inequality of total expenditure (after taxes). A tax will be called regressive if it increases total inequality. This definition can also be justified as follows. Consider the compensation that is necessary to preserve the level of well-being enjoyed by each household before the modification in taxation. If the compensation is progressive (i.e. its share increases with total expenditure or income), the change in the tax will affect the rich more than the poor. The tax is then progressive and its increase (or its imposition) will yield a decrease in inequalities. Conversely, if the compensation is regressive (i.e. its share decreases when total expenditure increases), the change in the tax affects the poor more than the rich. The tax is therefore regressive and its increase (or its imposition) will induce an increase in inequalities.

If the expenditure component is a decreasing function of total expenditure (or income), as is the case of a regressive tax paid by all households, then its Gini correlation with total expenditure is -1 and the relative marginal effect is negative. Consequently, when the relative marginal effect is negative, the taxation should increase inequalities, as would a regressive tax do. If the component is an increasing function of total expenditure, as for a progressive or proportional tax, then its Gini correlation with total expenditure is $+1$. One is then in the configuration (B) previously mentioned, and the sign of the relative marginal effect depends on the quantities R_k, G_k and G.

Hence, the interpretation of equation (8) in terms of the impact on total inequality of (an increase of) a tax on an expenditure category k is as follows: when the relative marginal effect is positive (negative), the taxation should reduce (increase) global inequality. Such a tax would be progressive (regressive).

In addition, the decomposition as previously described provides estimates of elasticities (called *Gini elasticities*) with respect to total expenditure without specifying a functional form for the Engel curves. The term

$$\eta_k = \frac{R_k G_k}{G} = \frac{cov(x_k, F_X)}{cov(X, F_X)} \times \frac{m}{m_k} \quad (9)$$

can be interpreted as the elasticity of expenditure k with respect to total expenditure. Indeed,

$$\beta_k = \frac{cov(x_k, F_X)}{cov(X, F_X)} \quad (10)$$

can be seen as a non-parametric estimator of the marginal propensity to spend on the category of goods and services k (Olkin and Yitzhaki, 1992; Yitzhaki, 1994).

The relative marginal effect in equation (8) can also be written as:

$$\frac{\partial G / \partial e_k}{G} = S_k(\eta_k - 1). \quad (11)$$

Equation (11) makes even more immediately clear the interpretation of the relative marginal effect, in agreement with the usual classification of taxes according to elasticities with respect to income. A tax is progressive if it is imposed on a luxury commodity ($\eta_k > 1$), in which case the relative marginal effect is positive. It is regressive if it is imposed on a necessary or inferior good ($\eta_k < 1$), in which case the relative marginal effect is negative. However, the extent of the relative marginal effect depends on the magnitude of the component's budget share (S_k). Finally, the tax is neutral if the elasticity is equal to 1 (the relative marginal effect is zero).

3. Transport expenditures in the household budget

This section describes the budget shares of different transport items by standard of living and the temporal pattern of the shares. Households are grouped into quintiles of total expenditure, deflated by the number of consumption units (CU) to account for their composition. The number of CUs in a household is determined according to the Oxford scale: 1 for the reference person (or head), 0.7 for any other member aged 15 or older, and 0.5 for each child under 15 years of age. Total expenditure is chosen as a classifying variable because expenditure data are more reliable than income data in budget surveys. Moreover, a measure based on consumption (more precisely, expenditures) is more relevant than a measure based on income for providing an account of the level of material well-being, because households tend to smooth their consumption so as to maintain a stable standard of living over time (Rogers and Gray, 1994; Slesnick, 2001).

Private transport expenditures include purchases of cars and two-wheelers; insurance costs for cars and two-wheelers; purchases of fuels, lubricants, tyres and accessories; maintenance and repair costs; parking costs; lock-up garage or parking lot rental costs; car licence and annual registration taxes, and vehicle use-related fines.

Some background information about vehicle taxation systems in each country is necessary for understanding some of the differences between the countries in our sample as regards transport expenditures and their effect on inequality.[2]

In France, the VAT rate for car purchases was as high as 33.33% until 1987. It then gradually decreased to 28% in Sept. 1987, 25% in Sept. 1989, 22% in Sept. 1990, and 18.6% (the rate imposed on the majority of commodities) in April 1991. For most products, the rate increased to 20.6% in August 1995, then decreased to 19.6% as of April 2000. People pay a car

[2] A detailed account of taxation systems in France and Denmark appears, for example, in Bückman and Rienstra (1998).

registration tax when they buy a vehicle, whether new or second-hand; they also pay that tax when changing residential location from one administrative area to another. The tax is calculated on the basis of engine size expressed in fiscal horsepower. The tax rate per horsepower is fixed at the regional level. In 1995, this rate varied between about 14.5 € and about 19.7 €. In 2008, it ranged from 27 € to 46.15 €. In addition, there is an annual tax on individual ownership of a vehicle (but no such tax has been imposed on households since 2001) and taxes on insurance. For fuels, the domestic tax on petroleum products (TIPP) and VAT (19.6%) apply.

The structure of taxes and duties on cars in Denmark is exceptionally complex, partly because of many attempts at altering the taxation of private vehicles to obtain more energy-efficient car transport. The last significant reform was completed in 1997. Its most important objective was to reduce taxes on car ownership and simultaneously increase taxes on car use by increasing petrol duties. However, the price of new cars in Denmark is much higher than in other countries, due to the high vehicle registration fee. For example, the registration tax in 2003 (charged on the basis of the retail price) is 105% of the first 7,653 € of the value and 180% of the remainder. Moreover, all cars are subject to VAT at 25%. In addition to an annual road tax, a supplementary tax is also imposed on vehicles using fuel other than petrol. All motor fuels are subject to VAT (25%).

Car taxation in Cyprus has changed several times during the last two decades. Prior to 2003, registration taxes and annual vehicle taxes were calculated on the basis of vehicle weight and were higher for diesel-powered cars; from 2003 onwards taxes have been based on engine size regardless of the fuel they use, and are further differentiated according to carbon dioxide emissions (with low CO_2 cars enjoying tax reductions and high CO_2 cars bearing an additional tax penalty). Tax rates changed again in 2006. For second-hand cars, their current tax levels (also calculated on the basis of engine size) decrease with age. For fuel taxes, the major difference between the two survey years was the VAT. This tax was introduced in 1992 at the rate of 5%; its rate in 2003 was of 15%. In 1991, the excise tax was 0.18 eurocents per litre for both petrol and diesel fuel. This amount was 0.30 eurocents per litre in 2003.

Table 1 presents the expenditure shares for private transport in the three countries, while Table 2 shows the number of vehicles by household. The average expenditure devoted by French households to private transport (mainly car acquisition and use expenses) remained stable at about 14% of the total budget and then decreased slightly since the mid-1990s. However, this share differs greatly according to the standard of living and grows with income, i.e. the gap between the first and last quintiles is up to 9 percentage points.

Table 1. *Budget shares of private transport*

Quintile ∗	France			Cyprus		Denmark	
	1979	1989	2001	1991	2003	1997	2005
1	7.2	8.3	8.1	9.4	8.6	7.0	7.6
	[6.8; 7.6]	[7.8; 8.7]	[7.7; 8.5]	[8.5; 10.3]	[7.8; 9.4]	[6.0; 8.1]	[6.3; 8.9]
2	10.9	10.7	9.9	14.0	11.5	7.5	11.9
	[10.4; 11.5]	[10.2; 11.2]	[9.4; 10.3]	[12.7; 15.2]	[10.5; 12.5]	[6.4; 8.6]	[9.9; 13.9]
3	12.5	13.0	11.8	16.8	13.4	9.8	13.4
	[11.9; 13.1]	[12.3; 13.6]	[11.3; 12.4]	[15.3; 18.3]	[12.2; 14.5]	[8.4; 11.3]	[11.2; 15.6]
4	15.2	15.7	14.0	23.2	13.7	17.9	15.8
	[14.5; 15.8]	[14.9; 16.4]	[11.4; 14.6]	[21.2; 25.3]	[12.5; 14.8]	[15.3; 20.5]	[13.2; 18.3]
5	15.1	16.9	12.2	25.1	20.0	29.8	25.1
	[14.4; 15.7]	[16.0; 17.8]	[11.6; 12.8]	[22.7; 27.4]	[18.2; 21.7]	[25.5; 34.2]	[21.3; 29.0]
All hhs.	**13.5**	**14.4**	**11.9**	**21.3**	**15.3**	**17.6**	**17.0**
	[13.2; 13.8]	**[14.0; 14.7]**	**[11.6; 12.3]**	**[20.2; 22.3]**	**[14.6; 15.9]**	**[16.4; 18.9]**	**[15.7; 18.3]**

∗ Quintiles of total expenditure by consumption unit (Oxford scale).
Note: Confidence intervals at 95% are given in square brackets.

The same patterns can be observed in Denmark and Cyprus but with notably larger budget shares. The gap between the first and last quintiles is even higher in these two countries, particularly in Denmark, where it reached 23 percentage points in 1997. This difference reflects the structuring of household car ownership by income level, even though car diffusion progressed over the period: as Table 2 shows, the number of cars per household increased more strongly for the lowest incomes. Table 1 also shows a considerable decline of the budget share of private transport in Cyprus. This decline is due mainly to a decrease in car purchase costs after 1993, following a relaxation of restrictions on imports of second-hand cars, i.e. the maximum allowable age of an imported vehicle was raised from two to five years. Consequently, from 1993 onwards, second-hand cars constituted the majority of new car registrations in Cyprus. Moreover, new cars became cheaper. Clerides (2008) found evidence of a 5-10% drop in the real price of new cars as a result of this regulatory change.

Table 2 provides interesting insights into vehicle ownership levels according to standard of living. In Denmark, the large ownership differences among income groups are largely due to the effect of heavy taxes imposed on cars, making cars affordable only to households with relatively high incomes. These high car costs can also explain the very low overall levels of car ownership in Denmark, as compared to Cyprus and France.

Table 2. *Number of vehicles per household*

Quintile *	France			Cyprus		Denmark	
	1979	1989	2001	1991	2003	1997	2005
1	0.43	0.60	0.72	0.37	0.65	0.45	0.56
2	0.75	0.93	1.08	1.02	1.26	0.58	0.85
3	0.93	1.08	1.23	1.20	1.57	0.77	0.98
4	1.07	1.18	1.36	1.36	1.81	0.94	0.96
5	1.13	1.29	1.42	1.59	1.88	1.11	1.13
All hhs.	**0.86**	**1.02**	**1.16**	**1.11**	**1.43**	**0.77**	**0.90**

* Quintiles of total expenditure by consumption unit (Oxford scale).

The budget share of public transport, illustrated in Table 3, is very low, particularly in France and Cyprus. The decline in budget shares in Cyprus is due mainly to the near disappearance of public transport (buses) in the country between the two survey years (1991 and 2003). According to official statistics, the total number of passengers decreased by 50% during this period, and public transport now accounts for 1.8% of total trips and 2.7% of total passenger kilometres travelled (CYSTAT, 2008). Rising incomes, urban sprawl and – most importantly – lack of investments in public transport infrastructure are responsible for this decline in bus transport.

In general no regular pattern appears related to income level, probably because of a diversity of contexts in terms of urbanisation and population density, and hence the availability of local public transport means.

Table 3. *Budget shares of public transport*

Quintile *	France			Cyprus		Denmark	
	1979	1989	2001	1991	2003	1997	2005
1	0.78	0.98	0.85	3.09	0.84	3.96	3.32
	[0.74 ; 0.82]	[0.93 ; 1.03]	[0.81 ; 0.89]	[2.78 ; 3.40]	[0.76 ; 0.92]	[3.35 ; 4.57]	[2.76 ; 3.88]
2	0.91	0.81	0.84	2.24	1.05	3.76	2.26
	[0.87 ; 0.95]	[0.77 ; 0.85]	[0.80 ; 0.88]	[2.04 ; 2.44]	[0.96 ; 1.14]	[3.21 ; 4.31]	[1.88 ; 2.64]
3	0.99	0.93	0.94	1.98	1.07	3.40	2.59
	[0.99 ; 1.03]	[0.88 ; 0.98]	[0.90 ; 0.98]	[1.80 ; 2.16]	[0.98 ; 1.16]	[2.90 ; 3.90]	[2.17 ; 3.01]
4	1.10	1.07	1.13	1.85	1.25	3.23	3.64
	[1.05 ; 1.15]	[1.02 ; 1.12]	[1.08 ; 1.18]	[1.69 ; 2.01]	[1.14 ; 1.36]	[2.76 ; 3.70]	[3.05 ; 4.23]
5	1.39	1.35	1.18	1.69	1.32	3.32	3.22
	[1.33 ; 1.45]	[1.28 ; 1.42]	[1.12 ; 1.24]	[1.53 ; 1.85]	[1.20 ; 1.44]	[2.83 ; 3.81]	[2.72 ; 3.72]
All hhs.	**1.14**	**1.10**	**1.06**	**1.91**	**1.19**	**3.45**	**3.06**
	[1.11 ; 1.17]	**[1.07 ; 1.13]**	**[1.03 ; 1.09]**	**[1.82 ; 2.00]**	**[1.14 ; 1.24]**	**[3.20 ; 3.70]**	**[2.82 ; 3.30]**

* Quintiles of total expenditure by consumption unit (Oxford scale).
Note: Confidence intervals at 95% are given in square brackets.

4. Consumption inequalities and redistributive effects of taxes

Transport expenditures are grouped into sufficiently homogeneous categories: car purchases, two-wheeler purchases, fuels, other vehicle use items, and public transport. This section presents the results for these categories as well as for broader groups.

To account for household composition, we carried all estimations on the basis of expenditures per consumption unit (as defined in the Oxford scale). In the estimations, the data were weighted by the respective survey weights of the households. The estimators of all the parameters of the decomposition of the Gini coefficient are efficient (i.e. asymptotically unbiased), and their distributions converge to a normal distribution (Schechtman and Yitzhaki, 1987). Thus estimation of their standard errors allows us to construct confidence intervals according to values of a normal distribution. Standard errors are estimated with the *jackknife* method.[3]

Before turning to the estimation results, let us note certain characteristics of the data. First, the observed expenditures are the result of choices made under income and price constraints. Moreover, by their nature, some goods and services are not purchased frequently or regularly (e.g. durables). Likewise, some expenditures are conditional on others or on the existence of a stock of durables, as is the case with vehicle use expenditures. Finally, at the household level, certain expenditures may be insufficiently recorded because of the survey method and/or the observation period.

The effect of these aspects on the estimations appears, notably, through the more or less large frequency of zero expenditures in the sample. The level of a Gini coefficient indicates the degree of disparities between households in terms of expenditures on a category of goods and/or services. These disparities reflect differences in terms of amounts spent, as well as in terms of how widespread these expenditures are among households. In general, as Garner (1993, p. 137) points out, the greater the proportion of zero expenditures, whether as a result of consumer choice or of the method of observation, the higher the corresponding Gini index.

[3] The algorithms of estimation by jackknife of variances of the parameters of the Gini decomposition are described in Yitzhaki (1991).

4.1. Inequalities and redistributive effects

4.1.1. Inequalities by expenditure item and their contribution to overall inequality

Table 4 presents the estimated Gini coefficients for each one of the household surveys used in this study. The lowest Gini coefficients are observed in the categories of expenditures on vehicle use (fuels and other items). This finding is not surprising because vehicle use expenditures are increasingly widespread in the population, with rising car ownership over the years. Then come, in ascending order, the Gini indices of expenditures on public transport and those on car purchases. Finally, two-wheeler purchases show the highest concentration. This result is likely attributable to the relative scarcity of these purchases, e.g. they represent on average about 1% of the transport budget among French households and only 0.2% of their total budget.

Table 4. *Gini coefficients by expenditure item*

Expenditure item	France			Cyprus		Denmark	
	1979	1989	2001	1991	2003	1997	2005
Private transport	0.671	0.668	0.643	0.698	0.601	0.704	0.617
	[0.664; 0.679]	[0.662; 0.675]	[0.636; 0.650]	[0.685; 0.712]	[0.585; 0.617]	[0.682; 0.726]	[0.588; 0.646]
Veh. purchases	0.897	0.883	0.895	0.891	0.916	0.901	0.897
	[0.892; 0.902]	[0.878; 0.888]	[0.890; 0.900]	[0.881; 0.902]	[0.908; 0.924]	[0.886; 0.916]	[0.877; 0.917]
Cars	0.905	0.888	0.902	0.893	0.919	0.912	0.910
	[0.900; 0.910]	[0.883; 0.894]	[0.897; 0.907]	[0.882; 0.903]	[0.911; 0.927]	[0.896; 0.927]	[0.890; 0.931]
2-wheelers	0.956	0.977	0.972	0.987	0.993	0.976	0.931
	[0.949; 0.962]	[0.973; 0.980]	[0.968; 0.975]	[0.975; 0.999]	[0.991; 0.995]	[0.960; 0.992]	[0.907; 0.955]
Fuels	0.645	0.579	0.571	0.564	0.472	0.652	0.620
	[0.635; 0.654]	[0.571; 0.587]	[0.563; 0.579]	[0.547; 0.581]	[0.459; 0.486]	[0.624; 0.679]	[0.586; 0.655]
Other use exp.	0.690	0.648	0.644	0.530	0.450	0.594	0.508
	[0.676; 0.704]	[0.636; 0.660]	[0.628; 0.659]	[0.515; 0.546]	[0.437; 0.462]	[0.563; 0.624]	[0.471; 0.544]
Public transport	0.889	0.882	0.862	0.777	0.791	0.679	0.714
	[0.881; 0.897]	[0.870; 0.894]	[0.851; 0.873]	[0.764; 0.789]	[0.780; 0.802]	[0.646; 0.711]	[0.684; 0.745]
All transport	*0.644*	*0.638*	*0.617*	*0.658*	*0.578*	*0.611*	*0.541*
	[0.637; 0.652]	*[0.631; 0.645]*	*[0.610; 0.625]*	*[0.644; 0.672]*	*[0.562; 0.594]*	*[0.587; 0.636]*	*[0.512; 0.569]*
Total expend.	**0.338**	**0.336**	**0.356**	**0.397**	**0.306**	**0.224**	**0.222**
	[0.333; 0.344]	**[0.329; 0.343]**	**[0.348; 0.365]**	**[0.385; 0.409]**	**[0.298; 0.315]**	**[0.212; 0.236]**	**[0.206; 0.238]**

Note: Confidence intervals at 95% are given in square brackets.

Except for vehicle purchases for which there is no regular temporal pattern (probably because purchases of durables are not made frequently), Table 4 shows that inequalities have generally decreased in all countries over time. In particular, we see a steady decrease in expenditures on fuels and other vehicle use items.

The contribution of each transport expenditure category to the overall inequality in a country appears in Table 5. As equation (6) shows, the contribution of a component to overall inequality is determined by three factors: the proper inequality of the component (measured by its Gini coefficient), its degree of association with total expenditure (measured by their Gini correlation), and its weight in the total budget. Thus, despite a very high Gini coefficient (close to 1), the relative contribution of two-wheeler purchases to overall inequality is insignificant in all three countries, due to their small budget share and their weak association with total expenditure. By contrast, the relative contribution of car purchases is much more important in all countries (ranging from 8% to 27%) despite slightly lower Gini coefficients, due to greater budget shares and stronger correlation with total expenditure. In terms of significance for overall inequality, this component is followed by vehicle use expenditures other than fuels, then by fuels, and finally by expenditures on public transport.

For France, inequalities among households for transport as a whole are essentially attributable to car purchases (44% to 58% depending on the survey period), followed by vehicle use expenditures other than fuels (21% to 26%) and fuels (13% to 22%). The contribution of public transport is more modest (6% to 8%). That of two-wheelers expenditures is even lower (less than 2%). Over the entire observation period, the contribution to overall inequality declines for fuels and for the remaining vehicle use expenditures. It also declines slightly for public transport. The contribution of two-wheelers purchases remains stable at a negligible level. The contribution of transport as a whole decreases at the end of the period after a slight increase, thus following the trend of the most important of its components (car purchases).

Table 5. *Relative contribution to overall inequality (%)*

Expenditure item	France			Cyprus		Denmark	
	1979	1989	2001	1991	2003	1997	2005
Private transport	16.5	18.4	13.4	28.1	22.3	40.9	31.6
	(0.41)	(0.50)	(0.42)	(1.3)	(1.0)	(3.40)	(3.02)
Vehicle purchases	7.7	11.7	7.9	22.0	15.3	30.0	21.2
	(0.28)	(0.42)	(0.32)	(1.3)	(1.0)	(3.29)	(2.86)
Cars	7.6	11.4	7.8	21.8	15.1	27.3	20.8
	(0.28)	(0.42)	(0.32)	(1.3)	(1.0)	(3.13)	(2.87)
Two-wheelers	0.1	0.2	0.2	0.2	0.2	2.7	0.4
	(0.03)	(0.04)	(0.03)	(0.17)	(0.09)	(1.15)	(0.15)
Fuels	3.9	2.6	2.1	2.8	3.5	3.7	4.5
	(0.12)	(0.07)	(0.07)	(0.2)	(0.1)	(0.37)	(0.57)
Other use exp.	4.9	4.2	3.3	3.3	3.5	7.2	5.9
	(0.26)	(0.19)	(0.21)	(0.2)	(0.1)	(0.67)	(0.76)
Public transport	1.7	1.6	1.4	1.5	1.4	2.8	3.5
	(0.13)	(0.18)	(0.13)	(0.1)	(0.1)	(0.75)	(0.63)
All transport	*18.2*	*20.1*	*14.7*	*29.6*	*23.6*	*43.6*	*35.1*
	(0.42)	*(0.50)*	*(0.45)*	*(1.3)*	*(1.0)*	*(3.31)*	*(2.91)*

Note: Standard errors are given in brackets.

Trends are similar for Cyprus and Denmark. The notable differences lie in the greater contribution of car-related components to global inequality, especially car purchases. This contribution is highest for Denmark, as expected given the very high motoring costs.

4.1.2. Redistributive effects of taxes by expenditure item

Table 6 displays the estimated effect on overall inequality (i.e. inequality of total expenditure) of a marginal change in each transport expenditure item. For France, taxes on transport commodities as a whole remain progressive, though to a lesser extent in 2001 than earlier. A 1% proportional increase of transport expenditures would have reduced global inequality by 2% in the early 2000s (compared to 5% previously). The progressivity of taxes on transport as a whole is due mainly to the progressive character of taxes on car purchases, strongly linked to income and with a higher budget share than for the other expenditure items. However, with the social diffusion of the car and of its use, taxes on vehicle use items are decreasingly progressive and even become regressive for fuels. Although the extent of the induced variations is very small (the relative marginal effect decreased from 0.1% to −1%), the trend is important: it reflects a gradual transformation of the distributions of these expenditures as the penetration of cars in the population grew.

Table 6. *Change in overall inequality due to a marginal change in a component (%)*

Expenditure item	France			Cyprus		Denmark	
	1979	1989	2001	1991	2003	1997	2005
Private transport	4.0	4.9	1.9	7.4	7.5	23.2	14.6
Vehicle purchases	3.0	5.1	2.9	9.3	8.7	20.4	13.9
Cars	3.0	5.1	3.0	9.3	8.6	18.6	13.9
Two-wheelers	0.0	0.0	0.0	0.0	0.1	1.8	0.0
Fuels	0.1	-0.6	-1.0	-1.0	-0.6	1.0	0.9
Other use exp.	0.9	0.3	-0.1	-1.0	-0.5	1.9	-0.2
Public transport	0.4	0.4	0.2	-0.5	0.2	-0.7	0.5
All transport	*4.4*	*5.3*	*2.1*	*6.9*	*7.7*	*22.5*	*15.1*

The slightly progressive character of taxes on public transport services is attributable to long-distance trips. Indeed, as shown in Berri (2005), taxes on local public transport appear neutral at the national level (i.e. neither progressive nor regressive). However, this result conceals a diversity of local conditions in terms of supply of these transport means, according to the degree of urbanization and population density. Effectively, these taxes prove regressive when focusing on the Greater Paris region, a large urban area with a very well developed public transport infrastructure. As to purchases of two-wheelers, which are rare, there appears to be no effect whatsoever on global inequalities.

As Table 7 shows, the (Gini) elasticities with respect to total expenditure confirm these conclusions as to the regressive ($\eta_k < 1$) or progressive ($\eta_k > 1$) character of a tax on a category of expenditures.

Table 7. *Total expenditure (Gini) elasticities – France and Denmark*

Expenditure item	France			Denmark	
	1979	1989	2001	1997	2005
Private transport	1.318	1.358	1.165	2.318	1.861
	[1.280 ; 1.356]	[1.317 ; 1.399]	[1.122; 1.208]	[2.137; 2.449]	[1.652; 2.069]
Vehicle purchases	1.635	1.779	1.588	3.129	2.921
	[1.577 ; 1.692]	[1.720 ; 1.839]	[1.526; 1.650]	[2.895; 3.364]	[2.597; 3.245]
Cars	1.668	1.797	1.616	3.153	3.024
	[1.609 ; 1.727]	[1.737 ; 1.857]	[1.553 ; 1.680]	[2.903 ; 3.403]	[2.688 ; 3.361]
Two-wheelers	0.809	1.155	0.860	2.911	1.020
	[0.534 ; 1.084]	[0.920 ; 1.389]	[0.684 ; 1.036]	[1.728 ; 4.094]	[0.406 ; 1.634]
Fuels	1.020	0.817	0.688	1.325	1.239
	[0.977 ; 1.062]	[0.780 ; 0.854]	[0.651 ; 0.725]	[1.122 ; 1.581]	[0.986 ; 1.493]
Other use exp.	1.228	1.083	0.975	1.353	0.969
	[1.154 ; 1.301]	[1.023 ; 1.143]	[0.898 ; 1.055]	[1.155 ; 1.551]	[0.750 ; 1.188]
Public transport	1.364	1.346	1.215	0.789	1.147
	[1.243 ; 1.485]	[1.179 ; 1.513]	[1.088 ; 1.342]	[0.428 ; 1.168]	[0.827 ; 1.468]
All transport	*1.322*	1.357	1.169	2.069	1.752
	[1.287 ; 1.357]	[1.318 ; 1.395]	[1.127 ; 1.212]	[1.908 ; 2.230]	[1.574; 1.929]

Note: Confidence intervals at 95% are given in square brackets.

Thus, for France, the luxury character of transport commodities as a whole is obvious (with an elasticity of 1.32 in 1979, 1.36 in 1989 and 1.17 in 2001), because of the predominance of car purchases (the elasticity of which ranges between 1.6 and 1.8). Vehicle use expenditures show continuously decreasing elasticities (from 1 to 0.7 for fuels and from 1.2 to 1 for the remaining vehicle use items), thus confirming the increasing necessity of the car. The same decreasing tendency is observed for the elasticity of public transport (from 1.4 to 1.2). However, as previously mentioned, these elasticities give information on the sign of the relative marginal effect, not on its *magnitude*.

Although the results appear similar for Denmark, there are significant differences in magnitudes. The progressivity of taxes on car purchases is much stronger in Denmark where, as pointed out earlier, these taxes are so high that only those with high incomes can afford car purchase costs. Also, fuel taxes remain progressive in this country. Given the higher share of car purchases in the total budget as compared to other transport components, taxes on all transport are also much more progressive in Denmark.

5. Conclusions and implications for transport policy

The paper investigates, for three European countries, inequalities between households in the consumption of transport goods and services, as well as the redistributive effects of taxes on various expenditure categories. The analysis builds on a decomposition by expenditure component of the Gini inequality index, using individual-level data from repeated cross-sections of expenditure surveys spanning long periods. The results highlight the effect of the gradual penetration of cars into the European population in the last six decades.

Inequality regarding transport is mainly attributable to car purchases, followed by vehicle use items other than fuels, and fuels. The relative contribution of public transport is very small, due to a small budget share. The relative contribution of car use items decreased over time, thus reflecting the increasingly widespread use of the car.

Taxes on transport goods and services as a whole are progressive, i.e. they affect the rich more than the poor. However, this result is principally due to the progressivity of taxes on car purchases, strongly linked to income and with a high budget share as compared to the remaining types of expenditures. By contrast, taxes on fuels are regressive (i.e. they affect the poor more than the rich) except in Denmark, whereas the progressive character of taxes on the other vehicle use goods and services has become weaker over the years. This result, again, is evidence of the effect of the social diffusion of the car, evolving from a luxury to a necessity.

These findings, which are more or less typical of socio-economic conditions in all European countries, present some significant policy implications. So as not to worsen social inequalities, policy makers should not ignore equity issues when considering measures for attenuating the environmental impact of cars (pollutant emissions, congestion and noise). Increasing car use costs (notably fuel prices) through an increase of uniform taxes would be especially inequitable. In particular, the least wealthy of car-dependent households living in low population-density areas would face a heavy burden. Indeed, as shown by the example of the Greater Paris region (Berri, 2007), the peripheral location of modest income households, because of high property prices in the centre of the urban area, involves transport expenditures (mainly car purchase and running costs) that increase with distance from the centre. These expenditure levels are not necessarily chosen but rather induced by the absence of a credible alternative to the car. McCann et al. (2000) show similar patterns for American urban areas. The drift towards remote areas is in particular favoured by mortgage lenders' not taking transport expenditures into account when making home purchase loans (Bardy, 2001; Hare, 1995). By so doing, they assume that life in the outskirts (where land and property prices are cheaper but badly served by public transport) is more affordable than in the centre.

Area-specific measures may be more appropriate, such as urban tolls and restrictions of access into dense urban areas. In parallel, public transport supply needs improvement in terms of lines of service, speed, punctuality, comfort, etc. In addition, a global approach should include actions in the housing sector, to increase the density of the urban fabric and attenuate the tendency towards sprawl. In addition to the necessity of mortgage lenders' taking into account transport costs when evaluating solvency, other possible measures improving housing market conditions include stimulating construction and promoting low-cost accommodation in areas where public transport is easily accessible.

Acknowledgements

The authors are grateful to Prof. Shlomo Yitzhaki (Hebrew University of Jerusalem and Central Bureau of Statistics, Israel) for kind permission to use the programme Jackknife1612 (version of August 2001) for estimation of the Gini coefficient and other parameters resulting from its decomposition. They are thankful to Sotiroulla Hadjispyrou for excellent assistance in extracting and processing data from the Cypriot household surveys, and to Natalie Reid for thoroughly rereading the paper and providing a valuable language help. The authors also thank Dr. Jean-Loup Madre (INRETS) and Mogens Fosgerau (DTU Transport) for helpful comments and suggestions. Of course, responsibility for any remaining error or shortcoming is that of the authors.

Appendix: Gini correlation

The Gini correlation between two random variables X and Y is a measure of their degree of association, based on the Gini Mean Difference (Schechtman and Yitzhaki, 1987). The Gini correlation coefficient is intermediate between the (usual) Pearson correlation coefficient and the rank-based Spearman correlation coefficient, the expressions of which are respectively:

$\rho(X,Y) = \frac{cov(X,Y)}{\sqrt{var(X)var(Y)}}$, and

$r_s(X,Y) = cov(R_X, R_Y)/\sqrt{var(R_X)var(R_Y)}$.

R_X and R_Y represent the ranks according to the values of X and Y, respectively. Divided by the size of the population or sample, they give the (empirical) cumulative distributions of the corresponding variables. The Pearson correlation is based on the covariance of the two variables, whereas the Spearman correlation is based on the covariance of their cumulative distributions. The Gini correlation is a compromise between the two: it uses the covariance between one of the two variables and the cumulative distribution of the other. It is a non symmetric measure that can take the two following forms:

$$R(X,Y) = cov\big(X, G_Y(Y)\big)/cov(X, F_X(X)),$$

$$R(Y,X) = cov\big(Y, F_X(X)\big)/cov(Y, G_Y(Y)).$$

In general, the two correlations $R(X,Y)$ and $R(Y,X)$ are not equal.

The properties of the Gini correlation coefficient combine properties of the Pearson and Spearman coefficients (Schechtman and Yitzhaki, 1987). Among these properties:

- for every (X,Y), $-1 \leq R(X,Y) \leq 1$;
- if X and Y are independent, $R(X,Y) = R(Y,X) = 0$;
- if Y is an increasing (resp. decreasing) monotone function of X, not necessarily linear, $R(X,Y)$ and $R(Y,X)$ will be equal to $+1$ (resp., -1);
- if (X,Y) has a bivariate normal distribution with parameters μ_X, μ_Y, σ_X^2, σ_Y^2 and ρ, then $R(X,Y) = R(Y,X) = \rho$.

References

Atkinson, A.B., 1970. On the measurement of inequality. Journal of Economic Theory, 2, 244-263.

Bardy, J.-P., 2001. La prise en compte des dépenses de transports dans les projets d'accession : l'approche des organismes financiers. Report of the Association Départementale d'Information sur le Logement de Paris (ADIL 75) for the Direction Régionale de l'Equipement d'Ile-de-France (DREIF), 130 p.

Berri, A., 2007. Residential location and household expenditures for transport and housing: the example of the Greater Paris region. 11th World Conference on Transport Research, 24-28 June, Berkeley, CA (CD-ROM).

Berri, A., 2005. Dynamiques de la motorisation et des dépenses de transport des ménages. Analyses sur données individuelles et semi-agrégées. Ph.D. Thesis, Université de Paris 1 – Panthéon-Sorbonne, France.

Bückman, E.H. and Rienstra, S.A. (Eds.), 1998. EUROFISC: Comparison of fiscal measures in passenger transport – Country essays. Dutch Ministry of Transport, 103 p.

Clerides, S., 2008. Gains from trade in used goods: evidence from automobiles. Journal of International Economics, 76 (2), 322-336.

CYSTAT, Statistical Service of the Republic of Cyprus, 2008. Short Distance Passenger Mobility Survey in Cyprus 2007. Nicosia, Cyprus.

Deaton, A., 1997. The analysis of household surveys: a microeconometric approach to development policy. The Johns Hopkins University Press, Baltimore.

Garner, T.I., 1993. Consumer expenditures and inequality: an analysis based on decomposition of the Gini coefficient. Review of Economics and Statistics, 75 (1), 134-138.

Hare, P.H., 1995. Clunker mortgages and transportation redlining: how the mortgage banking industry unknowingly drains cities and spreads sprawl. Patrick H. Hare Planning and Design, Washington, DC, 21 p. Available online at: www.vtpi.org/clunker.pdf (accessed on June 27, 2010).

Lerman, R.I., 1999. How do income sources affect income inequality? In: Silber, J. (Ed.), Handbook on income inequality measurement. Kluwer Academic Publishers, Boston, pp. 341-362.

Lerman, R.I. and Yitzhaki, S., 1994. Effect of marginal changes in income sources on U.S. income inequality. Public Finance Quarterly, 22 (4), 403-417.

Lerman, R.I. and Yitzhaki, S., 1989. Improving the accuracy of estimates of Gini coefficients. Journal of Econometrics, 42, 43-47.

Lerman, R.I. and Yitzhaki, S., 1985. Income inequality effects by income source: a new approach and applications to the United States. Review of Economics and Statistics, 67 (1), 151-56.

Lerman, R.I. and Yitzhaki, S., 1984. A note on the calculation and interpretation of the Gini index. Economics Letters, 15, 363-68.

McCann, B. et al., 2000. Driven to spend: the impact of sprawl on household transportation expenses. Joint report of the Surface Transportation Policy Project (Washington, DC) and the Center for Neighborhood Technology (Chicago, IL), 44 p. Available online at: www.transact.org/report.asp?id=36 (accessed on June 27, 2010).

Olkin, I. and Yitzhaki, S., 1992. Gini regression analysis. International Statistical Review, 60 (2), 185-196.

Rogers, J. M. and Gray, M.B., 1994. CE data: quintiles of income versus quintiles of outlays. Monthly Labor Review, 117 (12), 32-37.

Schechtman, E. and Yitzhaki, S., 1987. A measure of association based on Gini's mean difference. Communications in Statistics – Theory and Methods, 16 (1), 207-231.

Sen, A., 1997. On economic inequality. Clarendon Press, Oxford [1973 edition, Oxford University Press, enlarged with an annex by J. Foster and A. Sen: "On economic inequality after a quarter century"].

Slesnick, D.T., 2001. Consumption and social welfare. Living standards and their distribution in the United States. Cambridge University Press, Cambridge.

Stark, O., Taylor, J.E. and Yitzhaki, S., 1986. Remittances and inequality. Economic Journal, 96 (383), 722-740.

Yitzhaki, S., 1997. The effect of marginal changes in prices on income inequality in Romania. Research on Economic Inequality, 7, 241-258.

Yitzhaki, S., 1994. On the progressivity of commodity taxation. In: Eichhorn, W. (Ed.), Models and measurement of welfare and inequality. Springer-Verlag, Heidelberg, pp. 448-465.

Yitzhaki, S., 1991. Calculating jackknife variance estimators for parameters of the Gini method. Journal of Business and Economic Statistics, 9 (2), 235-239.

Chapter 3

The willingness to pay for quality aspects of durables: theory and application to the car market

Ismir Mulalic and Jan Rouwendal

The willingness to pay for quality aspects of durables: theory and application to the car market

Ismir Mulalic

University of Copenhagen, Øster Farimagsgade 5, Building 26, 1353 Copenhagen K, Denmark, and
Technical University of Denmark, Bygningstorvet 116 Vest, 2800 Kgs. Lyngby, Denmark
imu@transport.dtu.dk

Jan Rouwendal

VU University, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands
jrouwendal@feweb.vu.nl

Abstract

Conventional hedonic analysis measures willingness to pay for attributes on the basis of marginal fixed costs. We argue that in many cases variable costs are also affected by these attributes and that this should be taken into account. We develop a simple model to show that the marginal willingness to pay for a quality attribute has to be equal to the full marginal cost, which includes marginal fixed as well as variable costs. The model is applied to Danish data on car ownership and use. We use a nonparametric estimation procedure to estimate hedonic price functions for fixed and variable costs. We recover each consumer's marginal willingness to pay, the marginal fixed costs, and the marginal variable costs for car attributes using first-order conditions for utility maximization. We show that the marginal fixed and variable costs have the same (positive) sign and that both contribute substantially to the marginal willingness to pay. Estimation results suggest that marginal variable costs are on average about 20% of the full marginal costs. Finally, we estimate the distribution of the marginal rate of substitution between quality attributes and variable costs, which can be interpreted as a structural parameter, and we investigate how this marginal rate of substitution varies with household characteristics.

Keywords: Durable goods, willingness to pay, hedonic analysis, nonparametrics, car market.

JEL codes: C14, D46, L62, L68.

1. Introduction

This paper examines consumer choice behavior for durables taking into account dependence of both the fixed and the variable costs associated with the durable on quality aspects of the durable. The costs associated with durable goods often depend on the use made of the good. That is, in addition to the fixed cost there are non-negligible variable costs. In many cases fixed as well as variable costs are functions of the characteristics of the good. For instance, the engine power of a car is related to its fuel use. Cars are certainly not the only example of this phenomenon. Many hedonic studies have shown that the volume of a house is related to its price or rent (the fixed cost), but volume also has consequences for the variable costs associated with living in the house as heating a larger house requires more fuel. The marginal willingness to pay for space should therefore be computed by taking into account the increase in fixed as well as variable costs implied by an additional cubic meter. Nevertheless, most studies on the willingness to pay for housing characteristics concentrate exclusively on house prices, thereby ignoring variable costs (see for instance, Sheppard, 1999). The example just given is not exceptional. Indeed, many characteristics of housing – the presence and size of a garden, the number of bathrooms, et cetera – should be expected to affect the costs of living in the house: maintenance of the garden, keeping the bathrooms clean, etc. Other examples of characteristics of durables that affect variable as well as fixed costs are easy to find.

Although most studies have ignored variable costs when studying the demand for durables there are notable exceptions to this rule. The trade-off between fixed and variable costs of energy using durables has been studied by Hausman (1979) and Dubin and McFadden (1984). Hausman (1979) computed the discount rate implied by the choices of the consumers, and Dubin and McFadden (1984) showed that, in reaction to increasing fuel prices, consumers switch to more fuel efficient varieties and decrease the use made of the durable to some extent. However, the examples mentioned above show that there is not always a trade-off between fixed and variable costs. Fuel efficiency as a cause of higher fixed and lower variable costs appears to be a special case. Many other characteristics affect fixed and variable costs in the same direction. More volume of a house implies higher costs of heating as well as a higher house price. More engine power of a car means higher fixed and variable costs. Making a car safer – for those inside it – often means that its weight increases, which implies (all else equal) that fuel costs will be higher. More cabin space implies that the car will be more voluminous and (again, all else equal) this will increase fuel costs. We will document the positive relationship between fixed and variable costs for car characteristics later in the paper.

This paper develops a simple model of consumer choice behavior for durables in which fixed as well as variable costs are functions of the characteristics of the durable. We allow for the possibility that both cost functions are increasing in particular characteristics. A special feature of the model is that the quality characteristics may appear in the utility function. Consumers derive utility from housing or car characteristics in a direct way, something that is less likely in the case of fuel efficiency. We then derive a simple relationship between the willingness to pay for these characteristics and the fixed and variable costs. The marginal willingness to pay for a quality characteristic has to be equal to the full marginal cost, which includes marginal fixed as well as variable costs. The conventional approach in housing economics, where marginal willingness to pay is set equal to marginal fixed costs corresponds to the special case in which the characteristic has no impact on variable cost. The situation studied by Hausman (1979) and Dubin and McFadden (1984) corresponds to another special case in which the characteristic under study does not affect preferences directly. We show that in the case studied by these authors we need a trade-off between both types of costs for an interior solution. Moreover, we demonstrate that there can be an interior solution with positive marginal fixed as well as variable costs in cases where the characteristic has a direct positive effect on the consumer's utility.

The model is applied to Danish data on car ownership and use. We analyze an unusually rich dataset that informs us about car prices, car costs and car use. This allows us to estimate hedonic price functions for fixed as well as variable costs. We implement our model and compare the full marginal willingness to pay with that implied by analyzing the fixed cost only and find important differences. Moreover, we show that a special version of the model motivates a structural interpretation of our results as the marginal rate of substitution between characteristics and variable costs and investigate how this marginal willingness to pay varies with household characteristics.

The paper proceeds as follows. The next section introduces the theoretical model of consumer choice behavior for durables. Section 3 provides information on the data employed; Section 4 presents the estimation strategy and the empirical results; Section 5 reports on a further investigation of preferences; and Section 6 concludes.

2. The model

This section discusses the model that underlies the empirical work that follows. We introduce quality in the standard (textbook, two good) microeconomic models of consumer behavior in a very general way: one of the two goods is quality differentiated and the consumer has preferences over quality. The differentiated good has a constant unit price, referred to as its variable cost and a fixed cost. Both depend on the quality level chosen by the consumer. In subsection 2.1 the model is introduced, 2.2 analyzes quality choice. In 2.3 we briefly discuss the hedonic fixed cost function that corresponds with market equilibrium and in 2.5 we consider the structural interpretation of the empirical results.

2.1 Preliminaries

We consider a household that derives utility u from car kilometers,[1] denoted as q, the quality of the car, k, and other goods x (which are treated as a single composite):

$$u = u(x, q, k). \qquad (2.1)$$

In what follows we usually treat k as a scalar for expositional simplicity, but most of the analysis remains unchanged when it is a vector of characteristics. Examples of relevant car quality attributes are engine power, transmission system and capacity. The utility function is increasing in its three arguments and its indifference curves are convex. Since our empirical application below is about cars, we will henceforth use this example for concreteness, but it should be kept in mind that the model has more general applicability. Car kilometers q and other goods x are conventional goods in the sense that they are available in continuous amounts at fixed unit prices. The price of car kilometers equals variable car costs p while the price of the composite good is normalized to 1. Car quality is different from the other goods: it is an intrinsic property of the car owned by the household and as such it affects fixed as well as variable costs. The budget constraint is:

$$x + p(k)q = y - f(k), \qquad (2.2)$$

where p denotes the variable cost (per kilometer) of car use, and f the (annual) fixed cost. Both depend on quality. The fixed cost should be interpreted as user cost, that is as the sum of fixed

[1] Although we realize that travel demand is in many cases derived from the demand for other goods, we follow the bulk of literature by treating car kilometers as a conventional good.

maintenance costs, taxes, and the difference between the value of the car at the beginning of the year and the present value of its price at the end of that year (i.e. depreciation).

Conditional upon the choice of k, the maximization of (2.1) subject to (2.2) is the textbook utility maximization problem that under standard conditions can be solved to derive the demand equations for q and x. The former can be expressed as:

$$q = q(y - f(k), p(k), k) \; . \tag{2.3}$$

Assumption 1. Demand for car kilometers is normal, that is, the demand function $q(y - f(k), p(k), k)$ is increasing in $y - f(k)$.

Economic theory (the Slutsky theorem) now implies that the demand for car kilometers will be decreasing in the variable car cost p.

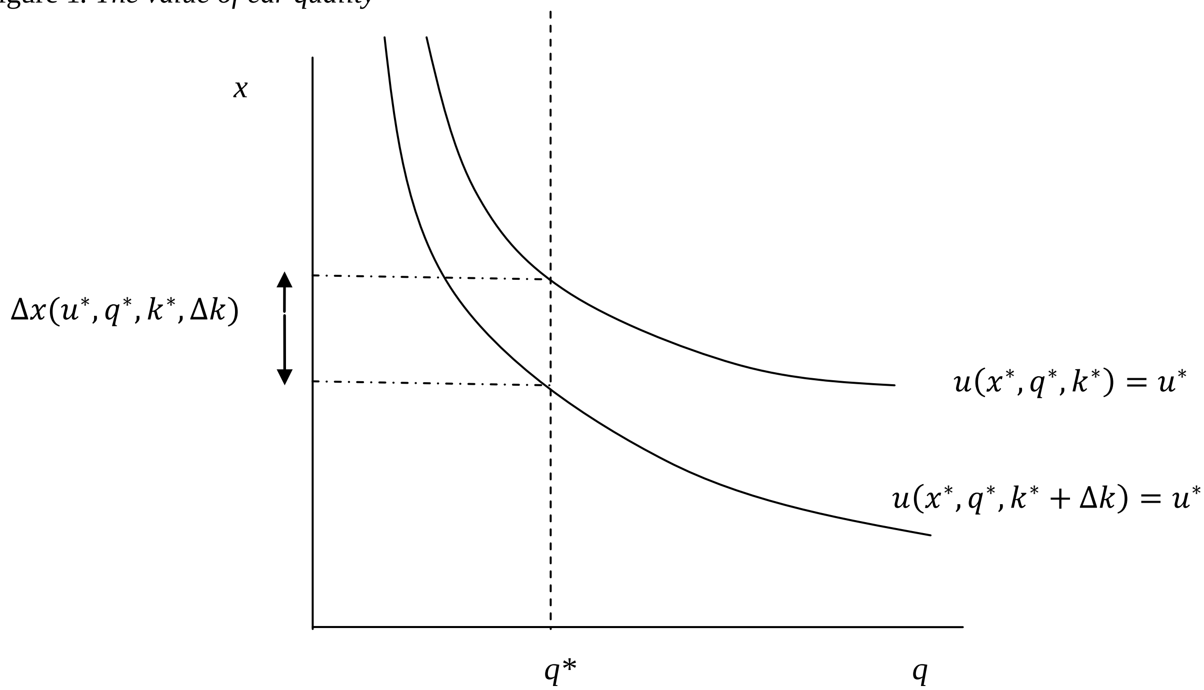
The sign of the effect of quality on the demand for car kilometers can be derived from a second assumption that refers to the relationship between a change in quality, Δk, and a change in the amount of the composite consumption, Δx, that compensates the consumer for the change in k, while keeping q constant. The value of Δx is determined by the initial values of x, q and k, and by the change in k, Δk. Since the first three variables determine initial utility, u, we can write Δx alternatively as a function of u, q, k and Δk. The variable Δx is implicitly defined by the following equation:

$$u(x - \Delta x(u, q, k, \Delta k), q, k) = u(x, q, k + \Delta k). \tag{2.4}$$

Assumption 2. For $k > 0$, $\Delta x(u, q, k, \Delta k)$ is an increasing function of q.

This assumption states that a consumer who drives more kilometers attaches a higher value to car quality in the sense that she is willing to give up more of the composite good in exchange for a higher quality of the car. The assumption is illustrated in Figure 1. This figure shows two indifference curves in q,x-space. Both indifference curves refer to the same level of utility, u^*, but to different level of car quality. Since car quality is valued positively by the consumer, the lower indifference curve in q,x-space refers to the higher level of car quality. For a given number of car kilometers, the value measure Δx defined above is the vertical difference between the two indifference curves.

Figure 1. *The value of car quality*



Assumption 2 implies that, for a given value of *q*, the indifference curve gets steeper when car quality increases. It follows that the demand for car kilometers is an increasing function of car quality. That is, if fixed and variable cost would both remain constant, and car quality is increased, the number of kilometers would also increase. In other words: the partial derivative of the demand function for kilometers with respect to car quality is positive.[2]

2.2 Quality choice

Conditional upon the choice of car quality, the indirect utility function can be written as:

$$u = v(y - f(k), p(k), k). \qquad (2.5)$$

Quality choice follows from the maximization of the conditional indirect utility function *v* through the choice of *k*. We assume quality is a continuous variable that has to be chosen from a closed interval $K = [k^{min}, k^{max}]$. Under conventional assumptions indirect utility is continuous and must therefore reach a maximum on *K* by Weierstrass' theorem.

[2] See Appendix A.

We maximize u in (2.5) by choice of k, taking into account the restrictions $k \le k^{\max}$, $k \ge k^{\min}$. Assuming differentiability, the first-order conditions of this maximization problem are:

$$-\frac{\partial v}{\partial(y-f)}\frac{\partial f}{\partial k}+\frac{\partial v}{\partial k}+\frac{\partial v}{\partial p}\frac{\partial p}{\partial k}+\lambda-\mu=0$$

$$\lambda\left(k-k^{min}\right)=0 \qquad (2.6)$$

$$\mu(k^{max}-k)=0$$

$$\lambda,\mu>0\ .$$

In this system λ and μ are the Lagrange multipliers associated with the two constraints. An interior solution exists when both Lagrange multipliers are equal to zero. In that case the first condition reads:

$$-\frac{\partial v}{\partial(y-f)}\frac{\partial f}{\partial k}+\frac{\partial v}{\partial k}+\frac{\partial v}{\partial p}\frac{\partial p}{\partial k}=0. \qquad (2.7)$$

This condition requires the marginal benefit $\frac{\partial v}{\partial k}$ (the higher utility caused by the better quality) to be equal to the total marginal cost $\frac{\partial v}{\partial(y-f)}\frac{\partial f}{\partial k}-\frac{\partial v}{\partial p}\frac{\partial p}{\partial k}$ (the lower utility caused by the higher costs). Note that this condition does not require that marginal fixed cost $\partial f/\partial k$ and marginal variable costs $\partial p/\partial k$ have opposite signs. Both can be positive and we have argued in the introduction that this is likely to be the case for at least some characteristics of housing, cars and other durables. However, condition (2.7) implies that with a positive marginal utility of quality the total marginal cost must be positive, which implies that at least one of the two marginal costs (fixed or variable) must be positive. If the marginal utility of quality is equal to zero, the two marginal costs have to be of opposite sign.

In what follows, we concentrate on interior solutions, since we do not see many households choosing the lowest or highest possible car qualities. That is, we interpret the market for automobiles as a continuum of submarkets for various qualities that are all supplied and demanded. We will later return to the conditions under which this is consistent with our theoretical model.

When marginal fixed and variable costs are both positive, (2.7) implies that an interior solution is only possible if the marginal utility of quality is positive. That is, it must be the case that the consumer derives utility from that quality characteristic immediately. If this condition is not fulfilled, as is likely to be the case, for instance, for fuel efficiency, the marginal utility of the characteristic is zero, and the two marginal costs must be of opposite sign for (2.7) to hold.

Condition (2.7) can be rewritten in two alternative ways. Application of Roy's identity and rearrangement of terms leads to the first of these equivalent statements:

$$wtp(k) = \frac{\partial f}{\partial k} + q\,\frac{\partial p}{\partial k} \quad , \tag{2.8}$$

where *wtp*(*k*) denotes (the absolute value of) the marginal willingness to pay for quality which is defined as:

$$wtp(k) = \frac{\partial v/\partial k}{\partial v/\partial (y-f)} \quad . \tag{2.9}$$

This equation states that the marginal willingness to pay for quality – its benefits – must be equal to the sum of the implied change in fixed and total variable costs. Alternatively, we can rewrite (2.7) as:

$$\frac{\partial f}{\partial k} = q\left(mrs(k) - \frac{\partial p}{\partial k}\right), \tag{2.10}$$

where *mrs*(*k*) denotes the marginal rate of substitution between quality and variable cost which is defined as:

$$mrs(k) = -\frac{\partial v/\partial k}{\partial v/\partial p}. \tag{2.11}$$

The right-hand side of (2.10) is the product of the quantity of the services provided by the durable and the difference between the marginal rate of substitution between the quality of the durable and the variable costs, and the marginal variable cost of quality. The former can be interpreted as the willingness to pay for extra quality, expressed as an increase in the variable cost. The expression between parentheses in (2.10) therefore gives an excess willingness to pay for quality over the increase in variable cost. This extra willingness to pay is needed to cover the increase in fixed costs associated with the higher quality.

Equation (2.10) is our preferred specification of the first order condition and we define a new variable $D(k)$ as:

$$D(k) = -\frac{\partial f}{\partial k} + q\left(mrs(k) - \frac{\partial p}{\partial k}\right) . \tag{2.12}$$

The necessary first order condition for an interior solution of the quality choice problem is: $D(k) = 0$. The sufficient condition for an interior solution to be a maximum is that the second derivative of indirect utility with respect to *k* is negative, or $\partial D(k)/\partial k < 0$:

$$\frac{\partial D(k)}{\partial k} = -\frac{\partial^2 f}{\partial k^2} + \frac{dq}{dk}\left(mrs(k) - \frac{\partial p}{\partial k}\right) + q\left(\frac{\partial mrs(k)}{\partial k} - \frac{\partial^2 p}{\partial k^2}\right). \tag{2.13}$$

In this equation dq/dk is the total derivative of the demand for kilometers with respect to *k*:

$$\frac{dq}{dk} = -\frac{\partial q}{\partial(y-f)}\frac{\partial f}{\partial k} + \frac{\partial q}{\partial p}\frac{\partial p}{\partial k} + \frac{\partial q}{\partial k}. \qquad (2.14)$$

The sign of $\partial D(k)/\partial k$ in (2.13) is indeterminate. To help determine its sign, we may assume that the functions *f* and p are both convex, and that the marginal rate of substitution between quality and variable cost is decreasing in *k*. However, this leaves us with the fundamental problem that the sign of dq/dk is indeterminate. The (reasonable) assumptions made earlier determine the signs of the three partial derivatives of q. Assuming that the fixed cost is increasing in *k*, as seems natural, implies that the first term on the right-hand side of (2.14) is negative. The third term is nonnegative, while the sign of the second term is ambiguous. In order to determine the sign of dq/dk in general, assumptions on preferences as well as the variable and fixed cost functions are necessary. It seems difficult to find a plausible combination of such assumptions at the level of the individual consumer.

We will therefore follow an alternative route to motivate the validity of the second order condition (2.13). We take the point of view that at the level of the market, car prices will adjust so as to ensure that every car quality in *K* will be demanded by some consumers. That is, we posit that the market for cars functions in such a way that the fixed cost function *f* adjusts to market circumstances to ensure that demand and supply for every quality in *K* are equal to each other. This will be discussed in detail in the next subsection.

2.3 Price equilibrium on the market for cars

A market equilibrium is a function $f(k)$ for which all consumers have a $k \in K$ for which the first and second order conditions for utility maximization are satisfied and all cars that are available in the market are owned by a consumer.

We start by limiting our attention to cases in which consumer preferences and variable cost are such that

Assumption 3. The marginal rate of substitution between quality and variable cost is larger than the marginal effect of quality on variable cost: $mrs(k) - \frac{\partial p}{\partial k} > 0$, is satisfied. This assumption says that consumers always prefer to drive a higher quality car if the marginal fixed costs of quality are equal to zero. Violation of this assumption implies that the first order condition (2.10) can only hold with marginal fixed cost $\frac{\partial f}{\partial k} < 0$, which does not seem to be a relevant case to analyze.

Now consider a market equilibrium in which optimal car quality is increasing in income under *ceteris paribus* conditions. That is, if all consumer characteristics, except income, are kept constant, a higher income will imply a higher optimal car quality. Taking the total derivative of (2.12) we can write this property as:

$$\frac{dk}{dy} = -\frac{\partial D(k)/\partial y}{\partial D(k)/\partial k} > 0. \qquad (2.15)$$

The numerator of the second expression can be elaborated as:

$$\frac{\partial D(k)}{\partial y} = \frac{\partial q}{\partial y}\left(mrs(k) - \frac{\partial p}{\partial k}\right) + q\,\frac{\partial mrs}{\partial y}\,. \qquad (2.16)$$

Since car kilometers are normal and the term between large brackets is positive by assumption 3, this expression is positive under:

Assumption 4. The marginal rate of substitution between car quality and variable cost, $mrs(k)$ is nondecreasing in income y: $\frac{\partial mrs(k)}{\partial y} \geq 0$.

This assumption seems plausible. Since we know, by assumption 2, that car kilometers are normal, assumptions 3 and 4 guarantee that $\partial D(k)/\partial y > 0$. We have therefore shown that the second order condition ($\partial D(k)/\partial k < 0$) is automatically satisfied for a function $f(k)$ for which all consumers have a $k \in K$ for which the first order condition for utility maximization is satisfied, if this *k* is increasing in consumer income (as indicated in 2.15).

The question that remains to be answered is whether there exists such a function $f(k)$ that also satisfies the other condition for market equilibrium, viz. that demand equals supply. Providing an answer to this question in the most general terms is outside the scope of this paper. In Appendix B we discuss the special case of a population of car-owning households with identical tastes that differ in incomes, where *k* is a scalar and show that a solution exists under general conditions and provide an affirmative answer.

2.4 Marginal prices and structural parameters

To investigate the possibility of giving a structural interpretation to our estimates we now discuss a simple example. Let the indirect utility function be given as:[3]

$$v = \left(\frac{\beta+\alpha\gamma+\beta\gamma p(k)+\gamma^2\big(y-f(k)\big)+\delta\gamma k}{\gamma^2}\right) e^{-\gamma(p(k)-(\delta/\beta)k)} \quad (2.17)$$

with $\alpha, \gamma, \delta > 0, \beta < 0$. The demand for kilometers follows by Roy's identity:

$$q = \alpha + \beta p(k) + \gamma\big(y - f(k)\big) + \delta k. \quad (2.18)$$

The marginal rate of substitution between car quality *k* and variable cost *p* can be determined as:

$$mrs(k) = -\frac{\delta}{\beta}, \quad (2.19)$$

which shows that Assumption 4 is satisfied. (2.19) shows that for indirect utility function (2.17) the marginal rate of substitution between car quality and variable cost is determined by the structural parameters of the utility function only.[4]

There exists other indirect utility functions for which this is also the case. Such indirect utility functions satisfy a strengthened version of assumption 4 in which $\frac{\partial mrs(k)}{\partial y} = 0$, which implies that indirect utility (2.5) can be written as: $u = v(y - f(k), v'(p(k), k))$. Moreover, the subutility function *v'* should be linear in its two arguments. For instance, for the indirect utility function (2.17) we have $v'(p(k), k) = \beta p(k) + \delta k.$ Apart from the linear demand function (2.18), some partial logarithmic demand functions can be derived from indirect utility functions that satisfy these requirements:

$$\ln(q) = \alpha + \beta p + \gamma(y - f(k)) + \delta k, \quad (2.20a)$$

$$\ln(q) = \alpha + \beta p + \gamma \ln(y - f(k)) + \delta k. \quad (2.20b)$$

If we can interpret the marginal rate of substitution between quality and variable cost as a parameter of the utility function, it reveals a structural aspect of consumer behavior. To see how it can be estimated, rewrite the first order condition (2.10) as:

$$mrs(k) = \frac{\partial p}{\partial k} + \frac{\partial f}{\partial k}/q. \quad (2.21)$$

Since we estimate the two marginal costs occurring in this equation for each consumer, and observe the number of kilometers driven in the data, this allows us to compute the *mrs*(*k*) at the

[3] This is a simple variant of the indirect utility function derived by Hausman (1981) for the case of a linear demand function.

[4] Eq. (2.19) relates the mrs to a ratio of coefficients of the utility function, but we can easily reparametrize in such a way that the marginal rate of substitution is a single parameter of the indirect utility function.

level of the individual consumer. To do so, we have to assume that each consumer's utility function satisfies the indirect separability conditions mentioned above, but the parameters of the utility function are individual-specific. This procedure is similar to the one Bajari and Benkard (2002) used to provide a structural interpretation to their investigation of the determinants of the willingness to pay for housing characteristics.[5]

To summarize, in this section we have developed a model for ownership, use and quality choice of a durable and studied the conditions for an interior solution in the context of market equilibrium. The validity of the second order condition for an interior solution is difficult to guarantee at the level of the individual agent for arbitrary fixed cost functions, but is automatically implied if we restrict attention to fixed cost functions that equilibrate the market.

3. The data

3.1 Introduction

We study car demand in Denmark. The data are derived from annual register data from Statistics Denmark and a car model database from the Danish car dealer association (DAF)[6] for the year 2004. The combination of these two databases results in a sample sufficiently detailed to explore the concepts proposed in this paper.

We have information from the car model database on car attributes of all new model variants supplied at the car market in Denmark in 2004 including the catalogue price of a new car, depreciation rate, car brand/model/type, vehicle cabin type (sedan, hatch, MPV, station car, and other), engine horsepower, and indicator for car transmission system. Other information on car attributes and household characteristics is derived from annual register data from Statistics Denmark, i.e. car tare, car total allowed weight incl. passengers and cargo, fuel efficiency, car vintage, and information on car owner's socio-economic characteristics (e.g. age, gender, income, etc.). Cabin capacity is most likely an important characteristic of a car. Since we do not observe cabin capacity measured in cubic meters, *car capacity* has been calculated as the difference between total allowed vehicle weight incl. passengers and cargo and the vehicle tare. The car model database includes more than thousand car brand/model/make combinations. We focus only on the most frequently purchased brand variants: Toyota, Suzuki, Hyundai, Peugeot,

---

[5] The utility function used by Bajari and Benkard (2002) is not suitable for the purposes of the present paper since it is quasilinear (implying that demand for car kilometers is insensitive to income).
[6] Danmarks Automobilforhandler Forening (DAF).

Fiat, Kia, Skoda, Mazda, Daewoo, Renault, Nissan, Volvo, Audi, Mitsubishi, other eastern car brands (Honda, Daihatsu, Subaru and others), and other car brands (Chrysler, Jeep, Land-Rover, Smart, Alfa Romeo, Mercedes-Benz, BMW, Seat, and others).

Our register data includes information about the so-called MOT tests. These are compulsory tests that take place before a car is sold, when it is four years old, and from then on every second year.[7] *Annual number of driven kilometers* has been calculated using information on exact odometer readings from the MOT tests. We use the odometer reading of the second MOT test (i.e. the first one after the car was sold) for the cars sold in 2004 that had not switched owner after being first sold. The times when the car was bought and when it passed its first MOT tests were both registered in months, which allows us to make a fairly accurate estimate of the annual number of kilometers driven during this time interval.

Annual fixed costs include annual depreciation, vehicle excise duty and insurance premium. The annual depreciation has been calculated, for example, for the first year of car use, by subtracting the provided average price for a one year old used car from the catalogue price.[8] The vehicle excise duty in Denmark is based on the vehicle fuel efficiency. Consequently, the vehicle excise duty has been calculated by using provided information on the vehicle fuel efficiency and the predefined statutory annual tax rates.[9] The estimates of the average annual insurance premium for ten car groups are provided by bilpriser.dk and FDM (2009). *Variable costs* have been calculated using information on fuel consumption and expenditures associated with car maintenance. Specifically, the fuel expenditure is compiled by dividing fuel price with specific car fuel efficiency in order to get a figure for fuel costs per kilometer. The estimates of average annual maintenance expenditures for ten car groups are provided by bilpriser.dk and FDM (2009). The units of all monetary amounts are in 2004 DKK.[10]

---

[7] The MOT test is a vehicle check that is compulsory for all vehicles registered in Denmark. The name derives from the Ministry of Transport. All Danish vehicles have to pass such MOT tests when first registered, and then at statutory time intervals. New automobiles owned by private persons (households) have to be approved by a second MOT test at least four years after being first registered, and then every second year. Each time a vehicle passes the MOT test, the inspection authority reads the odometer on the day of the MOT test, records date of the MOT test and several different identification data regarding the vehicle, such as vehicle id number, engine size, etc.

[8] The car model database includes information on expected prices for one year old cars for all brand/model/makes combinations calculated for the car associated with 20,000 driven kilometres. Moreover, the database provides brand/model/make specific price correction factors for deviation from the expected average annual kilometres driven. Thus, the expected depreciation has been corrected for the car wear. The depreciation includes also costs of delivery and all relevant financial expenditures (e.g. interests).

[9] The annual tax rates associated with the vehicle excise duty can be found in DAF (2005).

[10] One DKK is approximately 0.13€.

3.2. Selection of sample and descriptive statistics

We observe the full population of registered cars in Denmark. We select cars with petrol engines (77.8% of all cars) that were first registered in 2004 (87,798 cars). Records with missing information (8,469 observations) and households who sold the car between first registration and the second (MOT) test (38,792 observations) were excluded. Thus, our focus is on cases referring to new cars with petrol engines (we ignore imported second hand cars), purchased by households who did not sell the car between first registration and the first MOT test after being sold. We select households who did not sell the car between first registration and the second (MOT) test because we need annual kilometers driven for the computation of our willingness to pay and marginal rate of substitution measures. This left us with a sample of 40,537 observations. To ease the computational burden of estimating the flexible hedonic specifications, we drew a random sample of 10,000 observations.

Table 1. *Summary statistics*

Variable	Mean	Std. dev.	Minimum	Maximum
Fixed annual costs (DKK)	32,468	12,449	15,019	118,138
Variable costs (DKK/km)	2.1122	0.2521	1.6293	3.0537
Number of driven kilometers (km)	15,927	7,618	251	72,440
Car capacity (kg)	521	63	250	875
Engine horsepower	98.50	29.02	50.00	395.00
Automatic transmission (share)	0.0422	0.2011	0.0000	1.0000

Notes: Number of observations is 10,000.

Table 1 shows a summary of the main variables. They show that the mean annual fixed costs and mean annual expenditure associated with car variable costs are of more or less the same magnitude. The mean annual fixed costs are 32,468 DKK. The mean annual number of driven kilometers by one car is 15,927 kilometers, and the mean variable costs are 2.11 DKK per kilometer. The mean annual total expenditure associated with the variable costs of ownership and use of a car is then 33,641 DKK, about 3.6% higher than the mean fixed annual costs. The mean car capacity and the mean engine horsepower are 521 kg and 99 hp, respectively. The share of cars with automatic transmission is approximately 4.2%.

The correlation between variable and annual fixed costs equals 0.80,[11] which suggests that there does not exist a trade-off between both types of costs in the data. We have elaborated this issue by carrying out some regressions. Results are reported in Table 2. A basic regression of the fixed cost on the variable cost (column [1]) yields a positive and significant coefficient for the variable cost, while explaining almost two thirds of the variation in the fixed cost. Adding

[11] Pearson correlation; the correlation is significant at the 0.01 level.

controls for the car capacity, the engine horsepower and the automatic transmission (column [2]) decreases the coefficient for the variable cost, but it remains positive and highly significant. When dummies for the car brands and the cabin make are added the coefficient for the variable cost is again smaller, but still positive and significant at the 1% level (column [3]). Thus, a model focusing on the trade-off between fixed and variable costs would ignore much of the variation present in the data.

Table 2. *Regression of natural logarithm of fixed cost on natural logarithm of variable cost*

	[1]	[2]	[3]
Natural logarithm of variable cost	2.4868***	1.5311***	1.5001***
	(0.0087)	(0.0107)	(0.0193)
Natural logarithm of car capacity		0.2713***	0.1341***
		(0.0095)	(0.0082)
Natural logarithm of engine horsepower		0.4569***	0.3942***
		(0.0046)	(0.0052)
Dummy indicating automatic transmission		0.0733***	0.0399***
		(0.0042)	(0.0033)
Dummies indicating car brands	no	no	yes
Dummies indicating type of car cabin	no	no	yes
Intercept	8.4790***	5.3342***	6.5675***
	(0.0065)	(0.0503)	(0.0455)
R-squared	0.7163	0.8140	0.9009
No. observations	40,537	40,537	40,537

Notes: Dummies indicating car brands include: Opel, Ford, Toyota, Renault, Skoda, VW, Hyundai, Citroën, Peugeot, other eastern car brands (Honda, Nissan, Kia, Mazda, Suzuki and Mitsubishi), and other car brands (Chrysler, Jeep, Land-Rover, Smart, Alfa Romeo, Mercedes-Benz, BMW, Volvo, Audi, Fiat, Seat, and other). Dummies indicating type of car cabin include: sedan, hatch, MPV, station car, and other. ***,**,* indicate that estimates are significantly different from zero at the 0.01, at the 0.05 and the 0.10 level, respectively. Standard errors are in parenthesis.

4. Empirical strategy and results: the marginal willingness to pay for car attributes

Rosen (1974) pioneered the analysis of hedonic markets in a perfectly competitive setting and showed that the first derivative of the hedonic price function with respect to the individual attribute equals the marginal willingness to pay for this attribute. We will not provide a review of the subsequent literature that builds on his insights. Perhaps the most influential study referring to the car market is Berry et al. (1995) who study the market for new cars. They include fuel efficiency as one of the car characteristics in their model and focus on the price of new cars. The model we developed above suggests that the marginal fixed cost for quality characteristics is the difference between the marginal willingness to pay for that characteristic and the marginal variable cost. The conventional approach ignores the latter term and equates marginal willingness to pay with marginal fixed cost. This implies, for instance, that the willingness to pay for quality characteristics in terms of a higher price for new cars is independent of the fuel price, whereas our approach suggests that that consumer’s willingness to pay for quality characteristics of new cars varies inversely with the fuel price. Below, we estimate the full marginal willingness to pay for a car attribute as the change in the total annual costs of ownership and use of a durable that results from a small change in that attribute. We showed in section 2 that in our model it is proportional to the first derivative of the hedonic price (= fixed cost) function only if the variable cost of the durable is not affected by that attribute. If the variable cost is affected, there is a second term that is proportional to the hedonic variable cost function.[12] Application of our model therefore requires estimation of the hedonic price functions for both the variable and the fixed costs.

Our empirical analysis focuses on four basic car attributes that are able to explain much of the variance in both the fixed costs and the variable costs: engine horsepower, car capacity, type of the transmission system and the type of car cabin. Other, more subjective attributes, such as prestige of ownership, may affect both the fixed costs and the variable costs. We include the car brands in the empirical analysis as proxies for these difficult-to-quantify attributes.

4.1 Hedonic fixed and variable cost functions

Hedonic price functions for cars have been estimated ever since hedonic regressions have been run (see among others Court, 1939, and Grilliches, 1961). The recent literature stresses the importance of using flexible methods to estimate the hedonic functions (see e.g. Pace, 1995 and Anglin and Gencay, 1996). We follow Bajari and Kahn (2005) who use local linear regression to recover willingness to pay for housing attributes. Local linear methods have the same asymptotic variance and a lower asymptotic bias than the Nadaraya–Watson estimator, and the same asymptotic bias and a lower asymptotic variance than Gasser–Mueller estimator (Fan and Gijbels, 1996).

We assume that the fixed and variable costs of each car type $j = 1 \dots J$, satisfy:

$$\log f(k_j) = \log F(k_j) + \xi_j^f \qquad (4.1)$$

$$\log p(k_j) = \log P(k_j) + \xi_j^p \ , \qquad (4.2)$$

where *F* and *P* are unknown hedonic price functions, k_j is a vector of car characteristics and the ξs denote characteristics observed by the consumer but not by the researcher. The vector *k* of car characteristics consists of: hp, the engine horsepower; c, the car capacity; aut, a dichotomous

[12] For example, car cabin capacity affects automobile fuel efficiency (fuel consumption per kilometre) and thus variable costs.

variable that equals 1 if car's transmission system is automatic and 0 otherwise; $cabin$, a set of dummy variables indicating type of the car cabin; $brand$, a set of dummy variables indicating the car brands. Our local linear approach approximates the functions F and P locally for each observed car as:

$$\log F_j(k) = \alpha_{0,j}^f + \alpha_{1,j}^f\, log(hp) + \alpha_{2,j}^f\, log(c) + \alpha_{3,j}^f aut + \sum_{s=4}^{7} \alpha_{s,j}^f cabin_s + \sum_{r=8}^{22} \alpha_{r,j}^f brand_r + \xi_j^f, \quad (4.3)$$

$$\log P_j(k) = \alpha_{0,j}^p + \alpha_{1,j}^p\, log(hp) + \alpha_{2,j}^p\, log(c) + \alpha_{3,j}^p aut + \sum_{s=4}^{7} \alpha_{s,j}^p cabin_s + \sum_{r=8}^{22} \alpha_{r,j}^p brand_r + \xi_j^p. \quad (4.4)$$

We use local linear regression to approach the hedonic price functions at each observation point j.[13] In particular, we use weighted least squares to estimate the hedonic coefficients $\alpha_j = [\alpha_{0,j} \dots \alpha_{22,j}]'$. That is, for each j:

$$\alpha_j = arg\, min(\boldsymbol{r} - \boldsymbol{K\alpha})' \boldsymbol{W_j} (\boldsymbol{r} - \boldsymbol{K\alpha}) \quad (4.5)$$

where $\boldsymbol{r}$ is the $J \times 1$ vector of fixed or variable costs for all cars j ($\boldsymbol{r} = \boldsymbol{f} = [f_j]$ or $\boldsymbol{r} = \boldsymbol{p} = [p_j]$), $\boldsymbol{K}$ is a $J \times 23$ matrix of regressors (which for each product j includes an intercept and 22 attributes), and $\boldsymbol{W_j}$ is a $J \times J$ diagonal matrix of kernel weights. The kernel weights are a function of the distance between the characteristics of the car j' and car j. So, the local regression assigns greater importance to observations with attributes close to j.[14] We use normal kernel function:

$$W(z) = \prod_{m=1}^{22} N\left(\frac{z_m}{\hat{\sigma}_m^2}\right) \quad (4.6)$$

$$W_h(z) = W\left(\frac{z}{h}\right)/h. \quad (4.7)$$

In these equations N is the standard normal density function and h is the bandwidth. We evaluate the normal distribution for the *m-th* car characteristic at $z_m/\hat{\sigma}_m^2$, where $\hat{\sigma}_m^2$ is the sample standard deviation of attribute m. The choice of kernel bandwidth is central to the local regression (Altman, 1992) and the literature describes appropriate methods for choosing the bandwidth (see e.g. Fan and Gijbels, 1996, and Härdle, 1993) to approximate the hedonic price function. In the present study we focus on the first derivatives of the hedonic function and then a larger

[13] For description and discussion of local polynomial methods see e.g. Fan and Gijbels (1996) and Härdle (1993).

[14] Since we assume that data point that are close together have means that are more similar than data points that are far apart, it makes sense to use a weighted average, with smaller weight for data points farther from the centre of bandwidth.

bandwidth is recommended in the literature (see, for instance, McMillen, 2010). Based on visual inspection of the estimates, we choose the bandwidth (h) equal to 0.4. Moreover, the estimates of (4.5) allow us to recover estimates of the unobservable car characteristic associated with the fixed costs (ξ_j^f) and the variable cost (ξ_j^p) from (4.1) and (4.2). This unobservable car characteristic can be estimated as the residuals to the hedonic regression functions, i.e. $\xi_j^f = \log f(k_j) - \log F(k_j)$ and $\xi_j^p = logp(k_j) - \log P(k_j)$. We use the standard hedonic assumption that the unobserved car characteristics are independent of the observed car characteristics.[15]

4.2 Estimation results

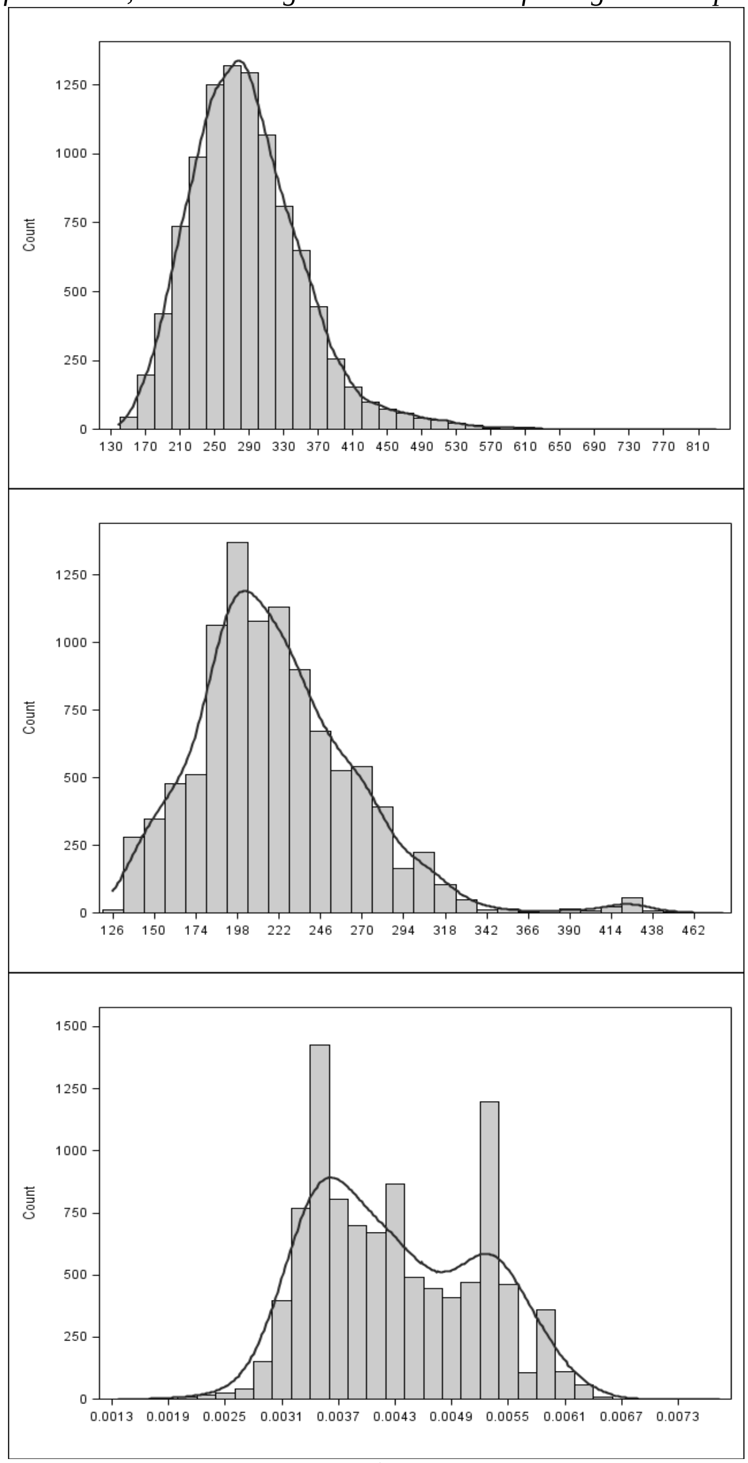
This section presents estimates of the hedonic price functions for the variable and the fixed cost functions (4.1) and (4.2). Estimated versions of the fixed and variable cost functions enable us to compute local (individual-specific) estimates of the marginal fixed costs and the marginal variable cost implied of the specific car attribute. As we noted above, estimation of marginal fixed costs is conventional in the hedonic price literature, whereas estimation of the marginal variable cost is much less common. We can also compute the (full) marginal willingness to pay for a car attribute from (2.8) using obtained estimates of the implicit prices faced by the household i from (4.1) and (4.2) together with information about the number of kilometers driven by the household.

For both the fixed cost and the variable cost functions the estimated implicit prices of the car attributes have intuitively plausible signs and magnitudes in almost all cases. Then for each car j we compute the marginal fixed costs of attribute m ($mfc_{j,m} = \partial F_j / \partial k_m$), and the marginal variable costs of that attribute ($mvc_{j,m} = \partial P_j / \partial k_m$). Moreover, using information about the number of kilometers driven, we compute on the basis on (2.8) the total marginal willingness to pay of attribute m ($wtp_{j,m}$). These estimates are household-specific in the sense that a unique set of the marginal fixed costs, the marginal variable costs and the total marginal willingness to pay are estimated for each value of $j = 1, \dots, J$.

Figure 2 presents histograms and estimated kernel distributions of the marginal willingness to pay, the marginal variable costs, and the marginal fixed costs for the engine horsepower and for the car capacity for the 10,000 Danish car owners.

[15] For discussion of the implication of this assumption for the car market equilibrium, see Berry et al. (1995).

Figure 2a. *Histogram and estimated kernel distribution of the marginal willingness to pay, the marginal fixed costs, and the marginal variable costs for engine horsepower (DKK)*



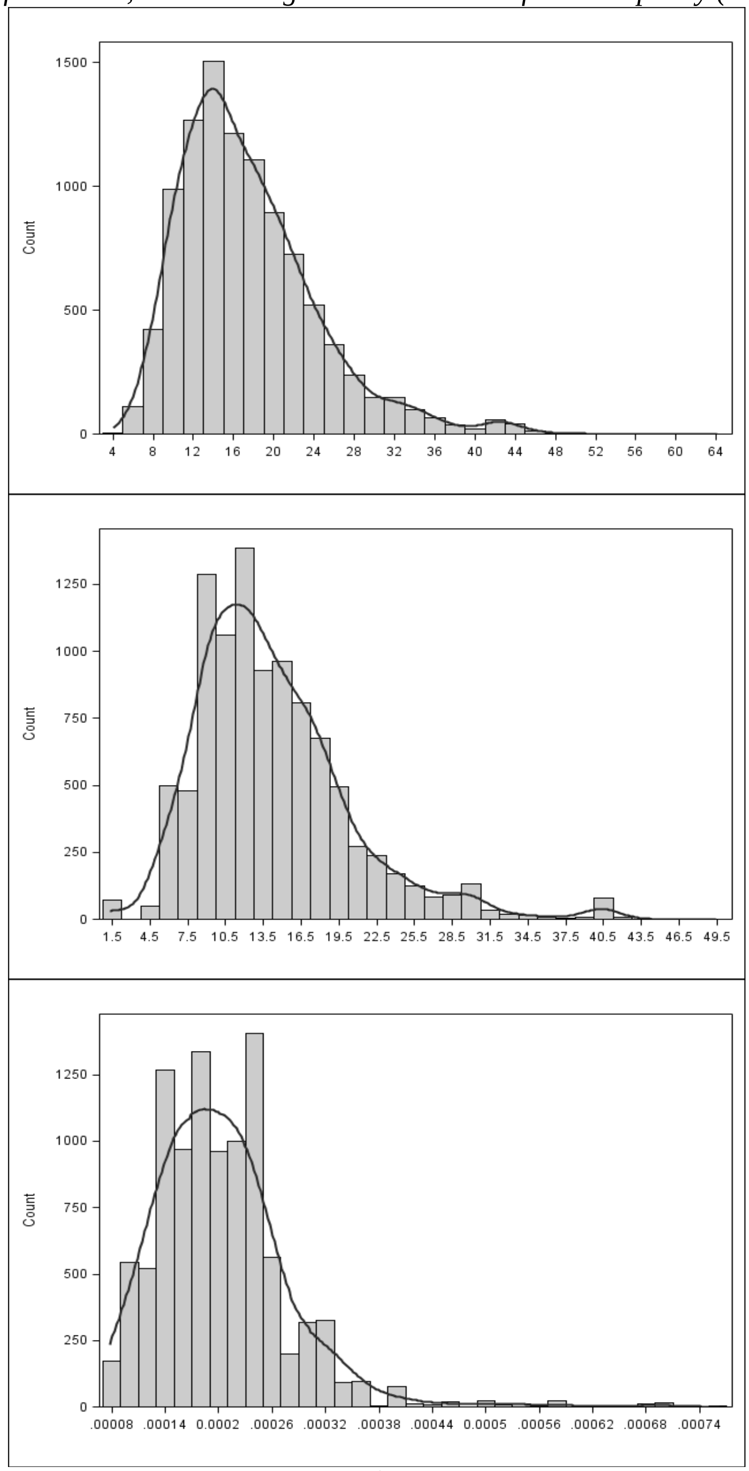
(a) *wtp*

(b) *mfc*

(c) *mvc*

Notes: The Kernel density estimation is performed here using SAS KDE procedure (SAS Institute Inc., 2009).

Figure 2b. *Histogram and estimated kernel distribution of the marginal willingness to pay, the marginal fixed costs, and the marginal variable costs for car capacity (DKK)*



(a) *wtp*

(b) *mfc*

(c) *mvc*

Notes: The Kernel density estimation is performed here using SAS KDE procedure (SAS Institute Inc., 2009).

This figure shows that the marginal willingness to pay is not symmetrically distributed, but skewed to the right. Moreover, the distribution of the full marginal willingness to pay is smoother than that of either the marginal fixed cost or the marginal variable cost as would be expected if a heterogeneous population of consumers whose tastes can be described by a smooth distribution function sorts over a large number of car makes by taking into account full marginal costs. It is also clear from these figures (in particular from a comparison of the distributions of marginal willingness to pay and the marginal fixed costs) that the marginal variable cost contributes substantially to our measurement of the total marginal willingness to pay. An average car's share of the marginal fixed costs of the total marginal willingness to pay is approximately 79% implying that one fifth of the full marginal cost refers to variable cost.[16]

Both the marginal fixed costs and the marginal variable costs are positively correlated with the number of kilometers driven, i.e. 0.36 and 0.19, respectively (see Appendix C, Table C.2). Thus, households with a relatively high demand for driving, demand more expensive cars.[17] This is consistent with our model even in the simplified case in which the marginal rate of substitution between quality and variable cost is a constant: a larger number of kilometers driven still increases the willingness to pay for quality in this situation (see (2.10)).

Table 3 shows that Danish households are on average willing to pay 287 DKK per year for an additional engine horsepower and 17 DKK per year for an additional kilogram of car capacity. Moreover, the total marginal willingness to pay for these two car attributes are dominated by the associated marginal fixed costs, i.e. 220 DKK and 14 DKK for engine horsepower and car capacity, respectively. The total marginal willingness to pay for the car automatic transmission system, compared to the car manual transmission system, is 4,529 DKK. The marginal fixed costs for this car attribute amount to approximately half of the total marginal willingness to pay of the car attribute. MPVs, station cars and other car cabin (including SUVs) have higher total marginal willingness to pay and only sedans have lower total marginal willingness to pay compared to hatchbacks. Unsurprisingly, car brands associated with high quality and prestige (e.g. Audi) have on average a higher total marginal willingness to pay than economical car brands (e.g. Hyundai).

---

[16] The average car in this context is a Toyota sedan with manual transmission system, 99 horsepower, and the capacity of 521 kilograms. For summary statistics for the computed willingness to pay measures, see Appendix C, Table C.1.

[17] For the histogram and the estimated kernel distribution of the number of kilometres driven, see the Appendix C, Figure C.1.

Table 3. *Means and standard deviations (DKK)*

	WTP		MFC		MVC	
	Mean	s.d.	Mean	s.d.	Mean	s.d.
Horse power	287	66.16	220	46.77	0.0043	0.0009
Capacity (kg)	17.44	6.91	14.25	6.18	0.0002	0.0001
Automatic transmission	4,529	1,994	2,065	1,057	0.1500	0.0170
MPV	11,207	4,103	6,123	2,115	0.3200	0.3400
Sedan	-4,480	1,822	-2,218	985	-0.1400	0.0180
Station car	11,110	4,171	6,847	2,694	0.2700	0.0290
Other cabin	12,054	4,531	6,647	2,435	0.3400	0.0390
Suzuki	-7,302	2,748	-6,101	2,548	-0.0750	0.0140
Hyundai	-8,071	2,856	-7,308	2,817	-0.0490	0.0130
Peugeot	-439	1,036	-887	1,066	0.0270	0.0150
Fiat	-8, 839	2,288	-5,418	2,193	-0.0220	0.0240
Kia	-6,352	2,253	-5,472	2,172	-0.0560	0.0180
Skoda	-7,299	2,682	-4,480	1,929	-0. 1800	0.0220
Mazda	1,008	901	980	946	0.0010	0.0120
Daewoo	-5,753	2,249	-6,055	2,432	0.0180	0.0140
Renault	-2,384	1,153	-1,764	1,100	-0.0390	0.0130
Nissan	581	1,106	324	1,049	0.0150	0.0160
Volvo	1,912	1,464	1,852	1,336	0.0024	0.0170
Audi	13,439	5,415	13,090	5,159	0.0200	0.0170
Mitsubishi	-5,010	1,575	-5,028	1,690	0.0001	0.0100
Other eastern car brands	-191	746	698	746	-0.0560	0.0110
Other car brands	1,049	1,159	1,241	1,127	-0.0110	0.0140

Notes: Wtp, mfc, and mvc are computed using the estimated coefficients from the hedonic price equations. The number of observations is 10,000.

5. Further investigation of preferences

The relationship between the marginal willingness to pay for the quality attributes and the structural parameters of the utility function (taste parameters) is of great interest. Bajari and Benkard (2002) have proposed a methodology for linking the marginal willingness to pay for the quality attributes to the (structural) parameters of individual specific utility functions. However, if only one choice per households is observed, as is the case in our data, severe restrictions have to be imposed on the utility function in order to recover household specific taste parameters, i.e. a log-linear specification for consumer preferences have to be assumed (Bajari and Benkard, 2002). The utility function assumed by Bajari and Benkard implies demand functions that do not depend on income and have price elasticity that equals -1, which is unattractive for the purposes of the present study. However, we showed in section 2 above that it is possible to relate the implicit prices for quality attribute k that follows from our estimates of the fixed and variable cost functions to the marginal rate of substitution between k and the variable cost ($mrs(k)$) and that this variable can be considered as a structural preference parameter if one is willing to impose the necessary functional form assumptions. Under these conditions, a structural investigation of consumer tastes can be performed on the basis of estimates of the individual marginal rates of substitution.

Our estimate of the *mrs* follows from (2.21), which we repeat here as:

$$mrs(k_m) = \frac{\partial p(k)}{\partial k_m} + \frac{\partial f(k)}{\partial k_m}\frac{1}{q}. \qquad (5.1)$$

The suffix *m* refers to the *m*-th quality attribute we consider in our empirical work, while *k* now denotes the vector of all car attributes considered. Since we have individual-specific estimates of the marginal fixed and variable costs and information about the number of kilometers driven, we are able to construct an individual specific estimate of the marginal rate of substitution $\widehat{mrs}_{i,m}$, where the index *i* refers to a household. An estimator of $mrs_{i,m}$ can be recovered from (4.1) and (4.2) as follows:

$$\widehat{mrs}_{i,m} = \frac{\hat{\alpha}^{p}_{j,m} p_j}{k_{j,m}} + \frac{\hat{\alpha}^{f}_{j,m} f_j}{k_{j,m}}\frac{1}{q_j}. \qquad (5.2)$$

Through (5.2) we recover household i's $\widehat{mrs}_{i,m}$ for characteristic *m* using available data and the estimate of the (local) implicit prices recovered from the hedonic price functions. Moreover, the marginal rate of substitution between car attribute m and variable cost p was shown in (2.19) to be equal to the negative of the ratio of two structural parameters of specific utility function. So, we are able to recover the ratio of two structural parameters of the utility function that is equal to the marginal rate of substitution.

After recovering household-level marginal rate of substitution between car attribute m and variable cost p, we can relate them to the household's socio economic characteristics d_i. We assume:

$$mrs_{i,m} = g_m(d_i) + \eta_{i,m} \qquad (5.3)$$

$$E(\eta_{i,m}|d_i) = 0. \qquad (5.4)$$

The $mrs_{i,m}$'s are modeled as functions, g_m, of household's socio economic characteristics, d_i, and an orthogonal household specific residual, $\eta_{i,m}$. We could easily do this estimation using flexible local linear methods. However, for presentation purposes, it is more convenient to model the joint distribution of mrs and demographic characteristics using a linear model. For continuous characteristics, we therefore simply estimate the following equation using robust regression:[18]

$$\widehat{mrs}_{i,m} = \theta_{0,m} + \sum_a \theta_{a,m} d_{i,a} + \eta_{i,m} \ . \qquad (5.5)$$

Given estimates $\theta_{a,m}$, the residuals ($\eta_{i,m}$) can be interpreted as household-specific taste shocks. Note that no parametric restrictions are imposed on the $\eta_{i,m}$.

For car attributes that take on the dichotomous values of 0 and 1, there is no first-order condition for utility maximization. Following Bajari and Kahn (2005) we apply a simple threshold decision making rule to estimate the $\widehat{mrs}_{i,m}$ for dichotomous attributes. Denote the value of the dichotomous car attribute as k_m. Utility maximization implies that:

$$[k_m = 1] \Rightarrow \left[mrs_{i,m} > \frac{\Delta p}{\Delta k_m} + \frac{\Delta f}{\Delta k_m}\frac{1}{q}\right] \qquad (5.6)$$

$$[k_m = 0] \Rightarrow \left[mrs_{i,m} < \frac{\Delta p}{\Delta k_m} + \frac{\Delta f}{\Delta k_m}\frac{1}{q}\right]. \qquad (5.7)$$

In these equations, the ratios of differences denote the changes in variable and fixed costs associated with a switch of k_m from 0 to 1. That is, if household i chooses $k_m = 1$, then we can infer that i's $mrs_{i,m}$ exceeds the implicit cost per kilometer for this attribute.

(5.6) and (5.7) show that the *mrs* for dichotomous characteristics is not identified. We can only infer that the preferences for a particular household are above or below the threshold value equal to the implicit cost per kilometer of the discrete characteristic. Following Bajari and Kahn (2005) we use a logit model to explain the choice of k_m.[19] Like these authors we normalize the coefficient on the implicit cost to −1 (see Bajari and Kahn, 2005).[20] For cabin types, we follow a similar approach and estimate a multinomial logit model.

The econometric results of two separate robust regressions where dependent variables are based on continuous quality attributes (i.e. engine horsepower and car capacity) are shown in table 4. In each regression we control for the age of the car owner, dummy variable for whether the car owner is male, dummy variable indicating presence of children in the household, and dummy variables indicating the population density of the car owner's municipality. We find that

[18] The main purpose of robust regression is to detect outliers and provide stable results in the presence of outliers. In order to achieve this stability, robust regression limits the influence of outliers. We are considering problems with outliers in the response direction. M estimation method, introduced by Huber (1973), has been applied for outlier detection and robust regression (SAS Institute Inc., 2009).

[19] These authors use a probit model which gives similar results in the case of dichotomous choice.

[20] An alternative approach, which does not require assuming that tastes lie in a parametric family, is to use the bounds approach described by Bajari and Benkard (2002).

the older car owners have higher mrs for car capacity and engine horsepower. For every year, the car owner's willingness to accept an increase in total annual variable costs rises by about 0.03 DKK for an additional kg of car capacity and 2.59 DKK for one additional engine horsepower.[21] Males are willing to accept increase in total annual variable costs of about 0.15 DKK and 34.38 DKK for a car with additional one kilogram of capacity and engine with one more horsepower. Households with children have slightly lower willingness to accept an increase in total annual variable costs for car capacity and engine horsepower, 0.09 DKK and 13.62 DKK, respectively. Finally, car owner's mrs associated with car capacity and engine horsepower decreases with population density. Compared to the car owners with residence in Copenhagen, car owners with residence in urban areas (with population density between 1,000 and 10,000 inhabitants) are willing to accept an additional increase in total annual variable costs of about 0.49 DKK for an additional kg of car capacity and 28.62 DKK for one additional engine horsepower.

Table 4. *Robust regressions of the marginal rate of substitution between quality attribute and variable cost on socio economic characteristics for continuous car attributes*

	[1] Engine horsepower	[2] Car capacity
Age	0.00016***	0.00001***
	(0.00001)	(0.00001)
Dummy indicating male	0.00212***	0.00005***
	(0.00014)	(0.00001)
Dummy indicating presence of children	0.00084***	0.00003***
	(0.00017)	(0.00001)
Dummy indicating population density 1-1,000	0.00177***	0.00016***
	(0.00019)	(0.00001)
Dummy indicating population density 1,000-10,000	0.00062***	0.00006***
	(0.00020)	(0.00002)
Dummy indicating population density 10,000- 100,000	0.00041**	0.00006**
	(0.00021)	(0.00002)
Constant	0.00890***	0.00050***
	(0.00030)	(0.00001)
R-square	0.0745	0.0645
No. observations	10,000	10,000

Notes: Dependent variables are the marginal rate of substitution between continuous quality attribute and variable cost. M estimation method, introduced by Huber (1973), has been applied for outlier detection and robust regression (SAS Institute Inc., 2009). Omitted variable associated with dummies representing population density is the Copenhagen area with the highest population density in Denmark. ***,** indicate that estimates are significantly different from zero at the 0.01, and the 0.05 level, respectively. Standard errors are in parentheses.

Table 5 reports logit estimation results for the car automatic transmission system. For every year, the car owner's willingness to accept an increase in total annual variable costs for the automatic transmission system rises by about 176 DKK. Males and households with children are

[21] The willingness to accept an increase in total annual variable costs has been calculated at the average number of kilometres driven (15,972) for a 1% increase in mean car capacity (521 kg) and mean number of engine horsepower (98.50 hp), i.e. 5.21 kg and 0.99 hp.

willing to accept an increase in annual variable costs of 3,562 DKK and 7,491 DKK, respectively, to own a car with the automatic transmission system.

Table 5. *Logit estimate for the marginal rate of substitution between automatic transmission system and variable cost*

	Automatic transmission
Age	0.011**
	(0.048)
Dummy indicating male	0.223**
	(1.030)
Dummy indicating presence of children	0.469***
	(2.150)
Dummy indicating population density 1-1,000	0.045
	(0.256)
Dummy indicating population density 1,000-10,000	-0.132
	(0.624)
Dummy indicating population density 10,000- 100,000	0.033
	(0.208)
Constant	-4.110***
	(17.700)
R-square	0.7351
No. observations	10,000

Notes: We have normalized the coefficient on implicit cost per kilometer equal to −1 instead of normalizing σ = 1. The estimation is performed with Biogeme (Bierlaire, 2005). Omitted variable associated with dummies representing population density is the Copenhagen area with the highest population density in Denmark. ***,** indicate that estimates are significantly different from zero at the 0.01 and the 0.05 level, respectively. Standard errors are in parentheses.

Table 6 reports MNL estimation results for the car cabin indices. Notice here that hatchback is the omitted variable in the hedonic price functions. The mrs associated with a station car decreases with the car owner's age (for every year by 527 DKK). Moreover, males and households with children prefer hatchbacks compared to other car cabin types, while the average Danish household prefers a station car.

Table 6. *MNL for the marginal rate of substitution between car cabin types and variable cost*

	MPV	Sedan	Station car	Other
Age	-0.002	0.003	-0.033***	-0.017
	(0.006)	(0.006)	(0.046)	(0.027)
Dummy indicating male	-0.490***	-1.680***	-1.350***	-0.826***
	(0.689)	(2.340)	(1.880)	(1.180)
Dummy indicating presence of children	-1.430***	-0.639***	-0.814***	-1.600***
	(1.990)	(0.897)	(1.140)	(2.240)
Dummy indicating population density 1-1,000	-0.511***	-0.317**	0.096	0.118
	(0.725)	(0.466)	(0.194)	(0.378)
Dummy indicating population density 1,000-10,000	-0.391**	-0.137	-0.268*	-0.709*
	(0.566)	(0.246)	(0.403)	(1.060)
Dummy indicating population density 10,000- 100,000	0.001	0.032	0.056	-0.383
	(0.163)	(0.169)	(0.179)	(0.677)
Constant	-0.999***	-1.430***	0.574**	-3.470***
	(1.410)	(1.990)	(0.795)	(4.850)
R-square	0.2570	0.2570	0.2570	0.2570
No. observations	10,000	10,000	10,000	10,000

Notes: We have normalized the coefficient on implicit cost per kilometre equal to −1 instead of normalizing σ = 1. Hatchback is the omitted variable in the hedonic price function. The estimation is performed with Biogeme (Bierlaire, 2005). Omitted variable associated with dummies representing population density is the Copenhagen area with the highest population density in Denmark. ***,**,* indicate that estimates are significantly different from zero at the 0.01, at the 0.05 and the 0.10 level, respectively. Standard errors are in parentheses

6. Conclusion

In this paper we have developed a model for choice of durable goods when variable costs are affected by quality attributes. This issue is ignored in the conventional hedonic analysis of, for instance, housing choice which restricts attention to the impact of quality attributes on fixed cost. Existing literature that considers fixed as well as variable costs concentrates on energy efficiency and the associated trade-off between fixed and variable costs. We concentrate on cases in which the quality attributes have a direct impact on utility and are positively related to fixed as well as variable costs. We developed a simple model that reduces to the standard two-good textbook model when quality is given. In this model quality attributes can be an argument of the utility function, while they also affect variable and fixed cost. The model covers situations in which variable costs are independent of quality attributes, or in which quality attributes affect variable and fixed costs in opposite ways as special cases. We showed that under plausible assumptions the second order condition for an interior solution is satisfied in market equilibrium where fixed costs (prices) have adjusted so as to equilibrate supply and demand for all quality levels.

We applied the model to Danish data. In these data fixed costs were positively related to variable costs even after controlling for car characteristics. Variable car costs were shown to be a substantial part of the total marginal cost of engine power and cabin capacity. We computed the total marginal willingness to pay for car characteristics. The distribution of this total willingness to pay was much smoother than that of its two cost components: marginal fixed cost and marginal variable cost. Marginal variable costs are on average about 20% of the full variable cost.

Finally, we related the marginal willingness to pay to household characteristics. Interesting correlations were found, and a structural interpretation of these results is possible if one is willing to make some additional assumption on the utility function.

Acknowledgements

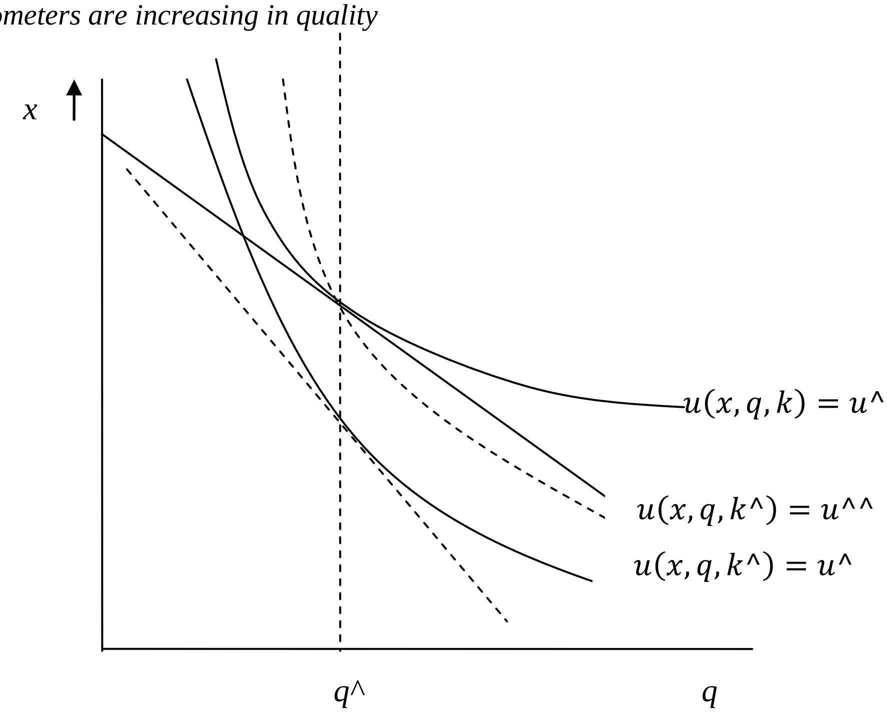
The authors thank Mogens Fosgerau and Bruno de Borger for helpful discussions on earlier versions of this paper and Alexandros Dimitropoulos, Piet Rietveld and Jos van Ommeren for comments. We are grateful to Statistics Denmark and the Danish car dealer association (Danmarks Automobilforhandler Forening) for providing the data. Research support from the Danish Energy Agency (Energy Research Programme) is acknowledged.

Appendix

Appendix A. Demand for car kilometers is increasing in car quality

To elaborate this, consider a consumer whose optimal *q* is *q*^ at the initial quality level *k* and reaches a utility level *u*^ in that situation. That is, the solid budget line is tangent to the indifference curve corresponding to *u*^ at *q*^ (see Figure A1). If quality increases, say to *k*^, the indifference curve corresponding to *u*^ shifts downwards. The slope of the shifted indifference curve at *q*=*q*^ is now steeper, as is indicated by the dashed budget line. For *q*=*q*^ there is now another indifference curve, corresponding to a higher level of utility than *u*^ , say *u*^^ that crosses the budget line (see the dashed indifference curve in the graph). Since demand for *q* is normal, the slope of this indifference curve must be steeper than the slope of the indifference curve corresponding to *u*^ at the higher quality level, and hence steeper than that of the indifference curve corresponding to *u*^ at the original quality level. In the graph, this means that the dashed indifference curve is steeper at *q*=*q*^ than the dashed budget line. And since the dashed budget line is steeper than the solid budget line, this implies that the dashed indifference curve crosses the solid budget line at *q*^ from above. In other words: the slope of the indifference curve that crosses the budget line at *q*=*q*^ gets steeper when quality increases and all else remains equal. The optimal *q* will therefore be higher than *q*^ after the increase in quality.

Figure A1. *Car kilometers are increasing in quality*



Appendix B. Fixed costs in market equilibrium

We adopt a short run perspective in which a given stock of – new and second hand – cars has to be distributed over a given number of households. The cars differ only in quality, which is here treated as a scalar variable, and the distribution of quality in the stock is a continuous function $G(k)$ that has positive support on K. Households all have the same tastes and differ only in income y. The income distribution is $H(y)$, and has positive support on a closed interval [y^{min}, y^{max}]. The total number of households is $H(y^{max})$. The total number of cars is $G(k^{max})$, which is assumed to be smaller than the number of households ($G(k^{max}) < H(y^{max})$). An equilibrium in the market is a fixed cost function $f(k)$ that allows all consumers to realize their utility maximizing quality choice and allocates all cars to households.

We conjecture that in equilibrium consumers with higher incomes drive higher quality cars while the households with the lowest incomes do not own a car. Denoting the critical income level, at which a consumer is just indifferent between owning and not owning a car as y^*, we must then have:

$$H\big(y(k)\big) - H(y^*) = G(k). \qquad \text{(B.1)}$$

This equation implicitly defines $y(k)$ as the income level to which car quality k must be allocated in equilibrium $\Big(y(k) = y^* + H^{-1}\big(G(k)\big)\Big)$. In other words, $y(k)$ gives us the allocation of cars over households that must be realized in equilibrium.

This allocation rule should be supported by utility maximization of all consumers faced with a fixed cost function $f(k)$ that is as yet unknown. For all these consumers the first order condition (2.10) must hold at the allocation described by $y(k)$:

$$\frac{\partial f}{\partial k} = q(y(k) - f(k), p(k), k)\left(mrs(y - f(k), p(k), k) - \frac{\partial p}{\partial k}\right) \qquad \text{(B.2)}$$

This equation repeats (2.10), after substitution of $y(k)$ for y, and makes all arguments of the demand and mrs function explicit.

(B.2) is a differential equation in $f(k)$. Its solution gives us the fixed cost function associated with market equilibrium, provided the second order condition is satisfied for all consumers. We can find a solution to (B.2) by interpreting it as an initial value problem. Start by observing that we must have:

$$y\big(k^{min}\big) = y^*. \qquad \text{(B.3)}$$

That is, the consumer with the critical income level chooses the car with the lowest quality. This consumer must – by the definition of the critical income – be indifferent between owning and not owning a car. This provides us with the value of $f(k^{min})$. For instance, if consumers who do not own a car use public transport, which has zero fixed cost, variable cost p^{pt} and quality k^{pt}, the utility of a such a consumer with income y is $v(y, p^{pt}, k^{pt})$. If consumer with income y^* is indifferent between car ownership and the use of public transport we must have:

$$v(y^*, p^{pt}, k^{pt}) = v\left(y^* - f\left(k^{min}\right), p\left(k^{min}\right), k^{min}\right), \quad \text{(B.4)}$$

and the value of $f(k^{min})$ is determined implicitly by this equation.

Once we know $f(k^{min})$, we can compute $\partial f\left(k^{min}\right)/\partial k$ by substituting $f(k^{min})$ into the right-hand-side of (B.2). The next step is to approximate $f(k^{min} + \Delta)$ for a small value of Δ as: $f\left(k^{min} + \Delta\right) = f\left(k^{min}\right) + \Delta \frac{\partial f(k^{min})}{\partial k}$. We can then compute $\partial f\left(k^{min} + \Delta\right)/\partial k$ by substituting $f(k^{min} + \Delta)$ into the right-hand side of (B.2) (also using $y(k^{min} + \Delta)$ instead of $y(k^{min})$ and continue the procedure until we reach k^{max}.

This procedure is an application of the Euler method for solving differential equations, which is known to converge to the true solution of the differential equation for $\Delta \rightarrow 0$. General conditions for existence and uniqueness of a solution $f(k)$ are provided by the Picard-Lindelöf theorem. Essentially what is needed is that the function f satisfies a Lipschitz condition.

An illustration

A closed form solution for the fixed cost function can be reached if we assume that all consumers have an indirect utility function (2.17). Here we assume that the parameters of this function are identical for all consumers (only incomes are different). Moreover, we assume that the variable cost is a linear function of quality:

$$p(k) = \pi_0 + \pi_1 k, \quad \text{(B.5)}$$

and that $\pi_1 < -\frac{\delta}{\beta}$ to satisfy Assumption 3. Substituting the appropriate expressions into (B.2) gives the following differential equation:

$$\frac{\partial f}{\partial k} = \left(\alpha + \beta(\pi_0 + \pi_1 k) + \gamma\left(y(k) - f(k)\right) + \delta k\right)\left(-\frac{\delta}{\beta} - \pi_1\right). \quad \text{(B.6)}$$

If the distributions of income and car quality are both uniform, then:

$$G(k) = \frac{k - k^{min}}{k^{max} - k^{min}}, \quad \text{(B.7)}$$

$$H(y) - H(y^*) = \frac{y - y^*}{y^{max} - y^*}. \quad \text{(B.8)}$$

Using these distributions, we can solve for $y(k)$ as:[22]

$$y(k) = y^* + \left(k - k^{min}\right)\left(\frac{y^{max} - y^*}{k^{max} - k^{min}}\right). \quad \text{(B.9)}$$

After substitution of this result we can solve (B.6) for $f(k)$ as:

$$f(k) = \left(\left(k - k^{min}\right)C + y^* + \frac{\alpha + \beta(\pi_0 + k\pi_1) + \delta k}{\gamma} + \frac{\beta}{\gamma^2} + \frac{\beta C}{\gamma(\delta + \beta\pi_1)}\right) \quad \text{(B.10)}$$

$$+ e^{\left(k - k^{min}\right)\gamma(\delta + \beta\pi_1)/\beta}\left(f\left(k^{min}\right) - y^* - \frac{\alpha + \beta\left(\pi_0 + k^{min}\pi_1\right) + \delta k^{min}}{\gamma} - \frac{\beta}{\gamma^2} - \frac{\beta C}{\gamma(\delta + \beta\pi_1)}\right)$$

In this equation C denotes $(y^{max} - y^*)/\left(k^{max} - k^{min}\right)$. Note that $\delta + \beta\pi_1 > 0$, because of assumption 3. This says that the positive effect of increased quality on car kilometers is larger than the negative price effect that occurs through the increase in variable costs associated with the higher quality.

The first expression in large parentheses in (B.10) is linear in k. Its slope is $C + \frac{\delta + \beta\pi_1}{\gamma}$, which is positive. The second expression in large parentheses – that appears after the exponent – can be rewritten as:

$$-\frac{1}{\gamma^2}\left[\beta + \left(\alpha + \beta\left(\pi_0 + k^{min}\pi_1\right) + \gamma\left(y^* - f\left(k^{min}\right)\right) + \delta k^{min}\right)\gamma\right] - \frac{\beta C}{\gamma(\delta + \beta\pi_1)}. \quad \text{(B.11)}$$

The expression in square brackets is the Slutsky term, which must be negative. The term (B.11) as a whole is therefore positive.

We conclude that the function $f(k)$ is the sum of an upward sloping linear curve and a nonlinear term that decreases asymptotically to zero. It can be shown that $f(k) > 0$ whenever $k \geq k^{min}$.

Appendix C. Additional information about the data

Table C.1. *Summary statistics for the willingness to pay measures for the average car*

Variable	Mean	Std. Dev.	Minimum	Maximum
Marginal willingness to pay (DKK)	37,578	14,761	14,563	135,586
Marginal fixed costs (DKK)	29,409	12,023	13,283	110,134
Marginal variable costs (DKK/km)	0.5064	0.0712	0.3578	0.8719
Marginal variable costs multiplied with the number of km driven (DKK)	8,169	4,293	108	41,556

Notes: The average car in this context is a Toyota sedan with manual transmission system, 99 horsepower, and the capacity of 521 kilograms. Number of observations is 10,000.

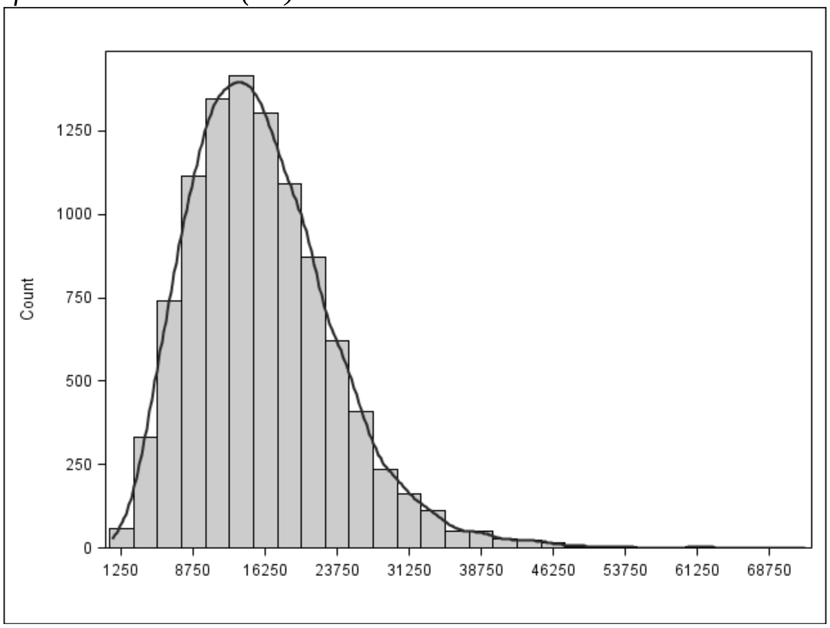
[22] Recall from the discussion following (B.1) that $y(k) = y^* + H^{-1}\left(G(k)\right)$.

Table C.2. *Correlation matrix of wtp, mfc, mvc and q*

	q	wtp	mfc
wtp	0.56724		
mfc	0.35577	0.96919	
mvc	0.19161	0.72125	0.72597

Notes: Pearson correlation. All the correlations are significant at the 0.01 level. Number of observations is 10,000.

Figure C1. *Histogram and estimated kernel distribution of the number of kilometers driven (km)*



Notes: The Kernel density estimation is performed here using SAS KDE procedure (SAS Institute Inc., 2009). Mean value is 15,927 and standard deviation is 7,618. Bandwidth is 1,105. Number of observations is 10,000.

References

Altman, N. S., 1992. An introduction to kernel and nearest-neighbor nonparametric regression. The American Statistician, 46, no. 3, 175-185.

Anglin, P. M. and Gencay, R., 1996. Semiparametric estimation of a hedonic price function. Journal of Applied Econometrics, 11, 633-648.

Bajari, P. and Kahn, M. E., 2005. Estimating housing demand with an application to explaining racial segregation in cities. Journal of Business & Economic Statistics, 23, no. 1, 20-33.

Bajari, P. and Benkard, C. L., 2002. Demand estimation with heterogeneous consumers and unobserved product characteristics: a hedonic approach. working paper, Stanford University.

Berry, S., Levinsohn, J. and Pakes A., 1995. Automobile prices in market equilibrium. Econometrica, 63, 841-890.

Bierlaire, M., 2005. An introduction to biogeme (version 1.4). biogeme.epfl.ch.

Court, A. T., 1939. Hedonic price indexes with automotive examples. in the Dynamics of automobile demand. General Motors Corporation, New York, 99-117.

DAF, 2005. Brugtbilkatalog nr. 184. Danmarks Automobilforhandler Forening, Tåstrup.

Dubin, J. and McFadden, D., 1984. An econometric analysis of residential electric appliance holding and consumption. Econometrica, 52, 345-362.

Fan, J and Gijbels, I., 1996, Local polynomial modeling and its applications. Chapman & Hall, London.

FDM, 2009. Beregningseksempel fra Motor nr. 1 2009 . http://www.fdm.dk/filer/ Beregningseksempel.pdf, (accesed 24/10 2010).

Griliches, Z., 1961. Hedonic price indexes for automobiles: an econometric analysis of quality change. mimeo, NBER.

Hausman, J., 1979. Individual discount rates and the purchase and utilization of energy-using durables. The Bell Journal of Economics, 10, 33-54.

Hausman, J., 1981. Exact consumer's surplus and deadweight loss. American Economic Review, 71, 662-676.

Huber, P.J., 1973. Robust regression: asymptotics, conjectures and monte carlo. The Annals of Statistics, 1, no. 5, 799–821.

Härdle, W., 1993. Applied nonparametric regression. Cambridge University Press.

McMillen, D., 2010. Issues in spatial data analysis. forthcoming in the Journal of Regional Science.

Pace, R. K., 1995. Parametric, semiparametric, and nonparametric estimation of characteristics values within mass assessment and hedonic pricing models. Journal of Real Estate Finance and Economics, 11, 195-217.

Rosen, S., 1974. Hedonic prices and implicit markets: product differentiation in perfect competition. Journal of Political Economy, 82, 34-55.

SAS Institute Inc., 2009. SAS/STAT® 9.2 User's guide second edition. Cary, NC

www.bilpriser.dk, Mikrobilsanalysen 2009. http://www.bilpriser.dk/analyselist.do (accesed 24/10 2010).

— Småbilsanalysen 2009, http://www.bilpriser.dk/analyselist.do (accesed 24/10 2010).

— Mellemklasseanalysen 2009, http://www.bilpriser.dk/analyselist.do (accesed 24/10 2010).

— Stationcarsanalysen 2009, http://www.bilpriser.dk/analyselist.do (accesed 24/10 2010).

— MPV-analysen 2009, http://www.bilpriser.dk/analyselist.do (accesed 24/10 2010).

— Cabrioletanalysen 2009, http://www.bilpriser.dk/analyselist.do (accesed 24/10 2010).

Sheppard, S., 1999. Hedonic analysis of housing markets. in Handbook of Regional and Urban Economics, eds. P. Cheshire and S. Mills, Volume 3, Chapter 41, pp. 1595-1635.

Chapter 4

The determinants of fuel use in the trucking industry – volume, size and the rebound effect

Ismir Mulalic

The determinants of fuel use in the trucking industry – volume, size and the rebound effect

Ismir Mulalic

University of Copenhagen, Øster Farimagsgade 5, Building 26, 1353 Copenhagen K, Denmark, and
Technical University of Denmark, Bygningstorvet 116 Vest, 2800 Kgs. Lyngby, Denmark
imu@transport.dtu.dk

Abstract

We analyse the determinants of trucking firm fuel use. We develop a simple model to show that trucking firm fuel use depends, in addition to the fuel price and the traffic volume, also on the output of the trucking firm's production process (the movement of cargo) measured in ton-kilometres, characteristics of the truck stock, and congestion. We also analyse the rebound effect for road freight transportation, i.e. the percentage of increased energy efficiency that does not result in the reduction of fuel used. For the purpose of analysing the rebound effect for road freight transportation, we decompose the standard definition of the rebound effect for motor vehicles, i.e. the elasticity of traffic volume with respect to fuel cost, into the elasticity by which changes in fuel costs affect freight activity and the elasticity by which changes in freight activity affect traffic volume. We estimate these elasticities using a simultaneous-equation model based on aggregate time-series data for Denmark for 1980-2007. Our best estimates of the short run and the long run rebound effects for road freight transportation are 19% and 28%, respectively. We also find that an increase in the fuel price surprisingly has a small but significant negative effect on the fuel efficiency (measured here as vehicle kilometres travelled (VKT) per litre of consumed fuel), i.e. a 1% increase in the fuel price decreases the fuel efficiency by 0.13% in the long run. However, less distance has to be driven for the same payload. An 1% increase in the fuel price decreases the VKT by 0.19% in the short run and 0.28% in the long run. Finally, a 1% increase in the fuel price results in a 0.19% reduction in the trucking firms' overall fuel use.

Keywords: Road freight transportation, fuel use, energy efficiency, rebound effect.

JEL codes: L91, Q41, R41.

1. Introduction

This paper examines the determinants of the road freight transportation fuel use. In 2004, the transportation sector was responsible for more than a quarter of the total world energy use, and roughly a third of this energy use was dedicated to road freight transportation (IEA, 2006; WBCSD, 2004). The analysis of the determinants of the road freight transportation fuel use is relevant because the road freight transportation's energy use is expected to grow in both the EU and the US (IEA, 2010; Léonardi and Baumgartner, 2004).

In Denmark, freight transportation accounts for a rising share of the total energy use as well. The road freight transportation activity (measured in ton-kilometres) increased by 59% from 1980 to 2007. Energy use of the road freight transportation increased by 105% in the same period. The main reason for the evident energy efficiency decline is presumably the 'just-in-time' behaviour of trucking firms, which resulted in lower utilization of the vehicles' capacity (Sathaye et al., 2010).[1]

Large reductions in road freight transportation energy use can be achieved by structural changes in the trucking industry towards improved matching of truck capacity to load (Kamakaté and Schipper, 2008). However, an often observed effect of policies directed at higher utilization of the vehicles' capacity is that better-utilized trucks are regularly heavier and use more fuel per kilometre, but, in theory, less distance has to be driven for the same payload (Léonardi and Baumgartner, 2004; Sathaye et al., 2010).[2] In Denmark, improvements in fuel efficiency of individual trucks were offset by growth in production and the overall change in the structure of the truck-stock (Kveiborg and Fosgerau, 2007).

As with all changes that improve energy efficiency, there may be some rebound effect that to some extent offsets the original energy saving.[3] As the energy efficiency of road freight transportation improves, freight road transportation becomes cheaper, thereby providing an incentive to increase its use. Thus total fuel use responds less than proportionally to changes in

[1] Just-in-time is an inventory strategy that strives to improve a business's return on investment by reducing in-process inventory and associated carrying costs (see e.g. Bonney, 1994).

[2] The reduction of freight truck trips with the general purpose to reduce congestion and environmental impacts has been a common policy goal for many governments around the world in recent years (Sathaye et al., 2010). For example, freight centres for facilitating cargo transfer have been constructed in several European countries implying significant savings for trucking firms using these centres through reduced fuel consumption (McKinnon, 2003).

[3] The rebound effect has been studied in different contexts (for survey see Greening, et al., 2000), including transportation (see e.g. Small and van Dender, 2007; Hymel, et al., 2010).

fuel efficiency. The rebound effect is typically quantified as the extent of the deviation from this proportionality.

Substitution between freight modes also has a large impact on freight transportation energy use, mostly because the energy intensity (measured in energy use per ton-kilometre) of trucks, ships and trains is considerably different (Forkenbrock, 1999). This paper focuses solely on road freight transportation because substitution between freight modes in Denmark is relatively limited and more than three quarters of all goods in Denmark are transported by trucks. According to Rich et al. (2010) a large proportion of the road freight transport services between OD pairs in Denmark cannot be substituted since there is only one option available, i.e. trucks. Moreover, Bjørner and Jensen (1997) calculated a cross-price elasticity of about 0.2 between road freight transportation versus train and ships (for a given transport demand).[4] In general, the share of road freight transportation compared to other modes is large in small countries (Kamakaté and Schipper, 2008).

Considering the debate about the road freight transportation fuel use, the absence of empirical estimates about it may be surprising. We aim to fill this gap in the literature. The aim of the current paper is to analyse the determinants of the trucking firm fuel use. We estimate a simultaneous-equation model based on aggregate time-series data for Denmark for 1980-2007. Our study deals with a range of the statistical difficulties by accounting for the endogeneity of fuel efficiency, and by distinguishing between autocorrelation and lagged effects. The paper adds to the transportation literature, contributing with two main improvements. First, we explicitly analyze the determinants of the fuel use in road freight transportation. To our knowledge, such an analysis has not been undertaken before. Matos and Silva (2011) analysed the effect of increasing energy efficiency based on the estimation of a direct rebound effect for road freight transportation in Portugal for the period between 1987 and 2006 using aggregate time series data. They estimated the demand for road freight transportation focusing on the effect of a change in energy cost of transportation on a change in demand for road freight transportation taking into account detected endogeneity of the price variable. Parry (2008) presented an analytical framework for estimating optimal taxes on the fuel use and mileage of heavy duty trucks in the United States that indirectly includes measures of the rebound effect. Bjørner (1999) carried out

[4] Surveys of the studies on price elasticities for freight transportation are given by Oum, Waters and Yong (1992) and Zlatoper and Austrian (1989).

an empirical analysis of the environmental benefits from better road freight transportation management in a Danish context, in a VAR model based on aggregate time series. Second, we show that the rebound effect for the road freight transportation can be decomposed from the standard definition of the rebound effect for motor vehicles, i.e. the elasticity of traffic volume with respect to fuel cost, into the elasticity of freight activity with respect to fuel cost per kilometre and the elasticity of traffic volume with respect to freight activity. The next section introduces the analytical model; Section 3 provides the empirical specification of the model; Section 4 presents the empirical results; and Section 5 concludes.

2. Trucking firm behaviour

We consider a small open economy where a representative competitive trucking firm ships goods at a given price denoted by $\bar{P}^Y$. Mode choice is not considered, thus there is only one means of transportation (road freight transportation). The output of the production process in trucking is the movement of cargo, or freight activity (Hubbard, 2003). A fundamental difficulty associated with studying trucking firm behaviour is finding an appropriate measure of output. Since trucking activity can be characterized by point of origin and destination, commodity type, and shipment size, the ideal measure of output would include all of these dimensions. In this study, freight activity is measured in ton-kilometres (tkm), which is the product of the mass of freight (measured in tonnes) and the distance it is carried (measured in kilometres).

Certain fundamental relationships exist between average load, aggregate ton-kilometres, vehicle kilometres travelled (VKT), and tons (Smith, 1957). The technical relation relating to freight activity (Y), traffic volume (V), and the average load (W) per shipment can be approximated as $V = Y/W$.[5] The capacity utilization depends largely on how well trucking firms can identify and agglomerate complementary demands onto individual trucks (Hubbard, 2003; Baker and Hubbard, 2003). We assume that the trucking firm can to some extent reduce traffic volume for a given freight activity through investments in logistics. These reductions will mainly be the result of better matching of the trucks capacity to shipment, i.e. change in the average load.

[5] Exact definition of the relationship between Y, V and W includes an adjustment factor to take into account the effect of the nonlinear statistical relationship between length of haul and size of load (see Smith, 1957). Due to data unavailability, we use the approximation of this relationship in an unadjusted form.

The firm employs labour (L), purchases fuel, and purchases and uses trucks to produce freight service. The total number of trucks (M) and the average truck's attributes affect the firm's costs. The average truck attributes S and H could be anything that affects the trucking firm's decision making process, i.e. the trucking firm's total revenue and total costs. For concreteness, we define S as the average truck capacity (measured here as axle weight). The firm's decision making process is also affected by the truck vintage since newer trucks depreciate more than older trucks and the truck vintage is presumably correlated with truck technology (for example fuel injection), so we define H as the average truck age.[6] Moreover, the firm's choice set includes also consideration of use of fuel, and consequently traffic volume, in producing freight service. Fuel consumption (F) and traffic volume (V) are related through the identity:

$$F = \frac{V}{E(S,H,D)}\,, \tag{2.1}$$

where $E = E(S,H,D)$ is the fuel efficiency measured in VKT per litre of consumed fuel. Fuel efficiency is a function of the average truck capacity (S), the average truck age (H), and the level of congestion (D), where trucks with larger capacity are assumed to have higher fuel consumption and where newer trucks through improved technology increase fuel efficiency, i.e. $E_S \leq 0$, $E_H \leq 0$ where subscripts stand for partial derivates. Increasing congestion is assumed to reduce fuel efficiency, i.e. $E_D \leq 0$.

2.1 Trucking firm profit maximization problem (PMP)

We consider a competitive market consisting of identical trucking firms producing a homogeneous service. When determining its optimal policy, the trucking firm faces the market constraint existing in any competitive market, i.e. the prices are assumed to be independent of the production plans of the firm (for discussion see Mas-Colell et al., 1995, chapter 10).[7]

The representative trucking firm attempts to maximize its profit (Π); that is the trucking firm chooses actions so as to maximize the total revenue minus total costs. It faces at least three types of the production costs: fuel costs, wages, and capital costs (Schipper and Price, 1997). Moreover, external factors (such as time of vehicle use, weather conditions, and traffic congestion) have proven to be relevant for the road freight transportation fuel efficiency, and consequently for the firm's costs (Samuelsson and Tilanus, 2002; Calthrop and Proost, 2003).

[6] For detailed discussion on vehicle vintage and fuel efficiency see Fullerton and West (2001).
[7] In a perfect competitive industry entry and exit costs are zero and firms are endowed with perfect foresight.

However, it is only the choice of the vehicle type and the implementation of the IT scheduling systems that can be influenced by the managers of trucking firms (Léonardi and Baumgartner, 2004).

The representative trucking firm is concerned only with determining the profit-maximizing levels of freight activity (Y) and inputs in production. The profit maximization problem (PMP) facing the firm can be written as simply a choice over its input levels (L, M, S, H, and V) for a given price vector and given $\overline{\mathrm{D}}$:

$$\max_{L,M,S,H,V\geq 0} \Pi = \overline{\mathrm{P}}^{\mathrm{Y}} f(L,M,S,H,V\,) - \left[wL + \frac{P^F}{E(S,H,\overline{D})}V + g(S,H,M;\boldsymbol{\alpha})\right], \qquad (2.2)$$

where $f(\cdot)$ is a quasi-concave differentiable production function with substitution possibilities between production inputs and $g(\cdot)$ is a differentiable cost function of S, M and H, where $\boldsymbol{\alpha}$ is the corresponding price vector. The vehicle capital costs (g) are equal to the costs of maintaining a truck fleet, i.e. costs related to S, M and H.[8] The price of labour and the price per litre of fuel are w and P^F, respectively.[9] The fuel price is divided by E to get a figure for fuel costs per kilometre (P^V). The functions that give the optimal choices of inputs and output as a function of the prices are known as the factor demand functions $Z^* = Z(\overline{D}, \overline{Y}, w, P^F, \boldsymbol{\alpha})$, $Z = L, M, S, H, V$ and output supply function, correspondingly.[10] The necessary first-order condition for Y^* to be profit maximizing is:

$$\overline{\mathrm{P}}^{\mathrm{Y}} - \frac{\partial C(Y,\overline{\mathrm{D}},\mathrm{w},\mathrm{P^F},\boldsymbol{\alpha})}{\partial Y} \leq 0\ , \qquad (2.3)$$

with equality if $Y^* > 0$, where $C(\cdot)$ is the cost function. Thus, at an interior optimum (i.e., if $Y^* > 0$), price equals marginal costs. This result will be useful in the analysis of the trucking firm fuel use.[11]

[8] We can specify the vehicle capital costs function as $g = (P^M + P^S S + P^H H)M$, where P^M is average truck capital/maintenance costs, P^S is the price for adding one additional unit of capacity to average truck capacity (S), and P^H is the price for adding one additional unit of age (e.g. year) to average truck age (H). We consider for simplicity of notation only the more general form of the vehicle capital costs function, i.e. $g(\cdot)$.

[9] A study of Denmark showed that fuel choice is almost exclusively diesel for trucks due to the very low diesel fuel cost (Lee Schipper and Price, 1997).

[10] It is easy to show that the first-order conditions for S, H and V, and the economic rate of substitutions between inputs are adjusted for the changes in the fuel efficiency, indicating the likely existence of the rebound effect.

[11] Analysis of the effects of changes of exogenous variables on the choice set variable (for example the effect of change of P^F on L, S, H, M, and V) requires determination of the signs of the bordered Hessian matrix of the Lagrangian principal minors. Determination of the signs of these principal minors, and consequently the analysis of the effects of changes of exogenous variables on the choice set variables, is not considered in this study because of the dimensionality of the bordered Hessian.

2.2 Trucking firm fuel use

When changes in the fuel efficiency are assumed to be exogenous, it is easy to show that fuel use responds to exogenous changes in E according to the elasticity equation (see Appendix A):

$$\varepsilon_{F,E} = -1 - \varepsilon_{V,Y}\varepsilon_{Y,P^V} \ , \tag{2.4}$$

where P^V is the per-kilometre fuel cost, $\varepsilon_{F,E}$ is elasticity of F with respect to E, $\varepsilon_{V,Y}$ is elasticity of V with respect to Y, and ε_{Y,P^V} is elasticity of Y with respect to P^V. Consequently a non-zero value of $\varepsilon_{V,Y}\varepsilon_{Y,P^V}$ implies that change in F is not proportional to change in E. Thus, $\varepsilon_{V,Y}\varepsilon_{Y,P^V}$ is taken as the measure of the rebound effect for road freight transportation. The rebound effect arises because traffic volume depends (among other things) on the freight activity, and the freight activity depends (among other things) on the variable cost per kilometre driven, a part of which is the per-kilometre fuel cost. Therefore, improved fuel efficiency reduces fuel cost per kilometre and consequently increases Y and V. The rebound effect refers to this response in Y and V which tends to reduce the beneficial effects of the improved fuel efficiency.

We analyse trucking firm fuel use by accounting for the endogeneity of fuel efficiency. We define fuel efficiency as a function of the average truck attributes and congestion. Consequently, the change in fuel efficiency will be the result of changes in the average truck capacity and the average truck age which again are determined by the level of freight activity, congestion, wages, capital costs, and fuel price. We will in this study focus on the effect of changes in fuel price on the trucking firms' total fuel use. A simple calculation using the definition of elasticity and the solution to PMP shows that (see Appendix A):

$$\varepsilon_{F,P^F} = \varepsilon_{V,P^F} + \varepsilon_{V,Y}\varepsilon_{Y,P^V}\left(1 - \varepsilon_{E,P^F}\right) - \left(\varepsilon_{E,P^F} + \varepsilon_{E,P^V}\right), \tag{2.5}$$

where $\varepsilon_{E,P^F} = \varepsilon_{E,S}\varepsilon_{S,P^F} + \varepsilon_{E,H}\varepsilon_{H,P^F}$ and $\varepsilon_{E,P^V} = \varepsilon_{Y,P^V}\left(\varepsilon_{E,S}\varepsilon_{S,Y} + \varepsilon_{E,H}\varepsilon_{H,Y}\right)$. ε_{E,P^F} measures the effect of fuel price on fuel efficiency and ε_{E,P^V} measures the effect of fuel cost per kilometre on fuel efficiency. The potential difference between ε_{F,P^F} and ε_{V,P^F} therefore requires that the last two terms in (2.5) be considerably different from zero. Disregarding this dependence of E on P^F may cause biased estimates of the effect of the change in the fuel price on the trucking firm fuel use and in particular biased estimates of the rebound effect.

3. Empirical analyses

3.1. System of simultaneous equations

The empirical specification is based on trucking firm PMP that simultaneously determines traffic volume (V), number of trucks (M), average truck capacity (S), average truck age (H), labour demand (L), and freight activity (Y). The factor demand functions are determined by the level of output and the factor input prices (see section 2.1). Thus, the trucking firm chooses traffic volume, size of the truck stock, average truck attributes, and demand labour based on freight activity, fuel price, input prices for capital (capital/maintenance costs) and labour (wages), and the level of congestion. The freight activity is determined (among other things) by the output price. Since we do not observe the output price we specify the freight activity equation based on the PMP first-order condition (see (2.3)), i.e. at the optimum output price equals marginal costs. Therefore, the freight activity is assumed to be the function of the level of congestion and the factor input prices, i.e. the fuel cost per kilometre ($P^V = P^F/E$), the input prices for capital (capital/maintenance costs), and wages.[12] The empirical specification of the freight activity also includes GDP that is used here as proxy for general economic development. We also include fuel efficiency (E) in the estimation with the purpose of explicitly analyzing the determinants of the fuel efficiency. The fuel efficiency is determined by accounting for the average truck attributes and the level of congestion. Moreover, the empirical specification of the fuel efficiency includes time trend to proxy for unmeasured changes (technology).[13] These assumptions lead to the following structural model:

---

[12] Notice here, $\frac{\partial C}{\partial Y} = w\frac{\partial L^*}{\partial Y} + \frac{P^F}{E}\frac{\partial V^*}{\partial Y} - V^*\frac{P^F}{E^2}\left(\frac{\partial E}{\partial S^*}\frac{\partial S^*}{\partial P^F} + \frac{\partial E}{\partial H^*}\frac{\partial H^*}{\partial P^F}\right) + \left(\frac{\partial g}{\partial S^*}\frac{\partial S^*}{\partial Y} + \frac{\partial g}{\partial H^*}\frac{\partial H^*}{\partial Y} + \frac{\partial g}{\partial M^*}\frac{\partial M^*}{\partial Y}\right)$ where $C = wL^* + \frac{P^F}{E(S^*,H^*,\bar{D})}V^* + g(S^*,H^*,M^*;\alpha)$ is the cost function, and $Z^* = Z(\bar{D},\bar{Y},w,P^F,\boldsymbol{\alpha})$, $Z = L,M,S,H,V$ are the conditional factor demand functions.

[13] We have also experimented with a producer provided indicator for the expected average fuel use per kilometre for a 40-tonne truck as proxy for the average truck technology. European Automobile Manufacturers' Association reports that average fuel consumption for a 40-tonne truck decreased from 50 litres per 100 km in 1967 to 32 litres per 100 km in 2004 (see ACEA, 2007). Improved fuel efficiency is the result of improvements in engine technologies (e.g. fuel injection), Selective Catalytic Reduction (SCR), telematics technologies (e.g. satellite navigation systems), tires, aerodynamics, etc., i.e. improved technology. So, the producer provided average expected fuel use per kilometre for a new 40-tonne truck has been used as proxy for the average truck technology. Moreover, this indicator has been divided by the average truck age with the purpose of accounting for speed of implementation of new truck technology in the existing truck stock. This experiment was not successful.

$$
\begin{aligned}
E &= E(S, H, \overline{D}, X_E) \\
L &= L(Y, P^F, \overline{D}, w, \boldsymbol{\alpha}, X_L) \\
S &= S(Y, P^F, \overline{D}, w, \boldsymbol{\alpha}, X_S) \\
H &= H(Y, P^F, \overline{D}, w, \boldsymbol{\alpha}, X_H) \\
M &= M(Y, P^F, \overline{D}, w, \boldsymbol{\alpha}, X_M) \\
V &= V(Y, P^F, \overline{D}, w, \boldsymbol{\alpha}, X_V) \\
Y &= Y(P^F/E, \overline{D}, w, \boldsymbol{\alpha}, X_Y),
\end{aligned}
\tag{3.1}
$$

where X_E, X_L, X_S, X_H, X_M, X_V, and X_Y are additional exogenous variables including constants.

We analyse the trucking firm fuel use based on the system in (3.1). Following Small and van Dender (2007) we generalize estimation in two ways to handle dynamics. First, we allow the error terms to be autoregressive of order 1. It means that unobserved factors influencing decisions in a given state will be similar from one year to the next. This could be caused by unobserved factors that persist over time, such as for instance business organizational styles. Second, we include the one-year lagged value of the dependent variable among the explanatory variables. The coefficient of this variable determines the difference between short run and long run effects on the independent variables. The inertia of such movement can arise due to lack of knowledge or slow turnover of the truck stock, or simply because trucking firms respond only slowly to changed circumstances. Consistent estimates of variables in a time series data may depend on autoregression and autocorrelation. Both autoregression and autocorrelation are important in determining the short run and long run effects, because the measurements of the lagged values of the dependent variables are sensitive to whether or not autocorrelation is controlled for. However, it is difficult to separate the presence of a lagged dependent variable from the presence of autocorrelation, especially when aggregate time-series data are used. In the current paper, we discuss the results of a specification incorporating both autoregression and autocorrelation.[14] We will explicitly address the potential bias of this specification by comparing results of this specification with results of a specification incorporating only autoregression.

We specify the equations as linear in parameters and with most variables in logarithms, leading to the following system:

[14] Survey of the empirical studies on rebound effect for motor vehicles is given by Small and van Dender (2005).

$$e_t = \alpha^e e_{t-1} + \alpha^{es} s_t + \alpha^{eh} h_t + \beta^e X_t^e + u_t^e$$
$$l_t = \alpha^l l_{t-1} + \alpha^{ly} y_t + \beta_1^l p_t^f + \beta_2^l X_t^l + u_t^l$$
$$s_t = \alpha^s s_{t-1} + \alpha^{sy} y_t + \beta_1^s p_t^f + \beta_2^s X_t^s + u_t^s$$
$$h_t = \alpha^h h_{t-1} + \alpha^{hy} y_t + \beta_1^h p_t^f + \beta_2^h X_t^h + u_t^h \qquad (3.2)$$
$$m_t = \alpha^m m_{t-1} + \alpha^{my} y_t + \beta_1^m p_t^f + \beta_2^m X_t^m + u_t^m$$
$$v_t = \alpha^v v_{t-1} + \alpha^{vy} y_t + \beta_1^v p_t^f + \beta_2^v X_t^v + u_t^v$$
$$y_t = \alpha^y y_{t-1} + \alpha^{ye} e_t + \beta_1^y p_t^f + \beta_2^y X_t^y + u_t^y$$

with autoregressive errors:

$$u_t^i = \rho^i u_{t-1}^i + \varepsilon_t^i \qquad i = e, l, s, h, m, v, y\ , \qquad (3.3)$$

where lower case notation indicates that the variable is in logarithm.[15] The individual variables in each vector X may be in either levels or logarithms. Subscript t designates a year, and u and ε are error terms assumed to have zero expected value, with ε assumed to be "white noise". The following section provides an overview of the variables used in the system (3.2).

3.2. Data and variables

The data used in the empirical analysis are aggregate time-series data for Denmark covering the years 1980-2007. Our period of observation is thus 28 years. For each year, we have information on aggregate freight activity measured in tkm, aggregate VKT of all trucks registered in Denmark, aggregate fuel consumption (of all trucks registered in Denmark), total actual hours worked in road freight transportation, average truck capacity (measured as average truck total axle weight), average truck age, number of trucks in the truck stock, fuel price, compensation of employees in road freight transportation, price index for vehicles and spare parts, and a range of explanatory variables (GDP, total annual VKT for all motor vehicles registered in Denmark, and the infrastructure measure (kilometre road in Denmark)). *Energy efficiency* (E) has been approximated as the VKT per litre of consumed fuel calculated as the ratio between the total annual VKT and the total annual fuel use. Our measure of *congestion* (D) has been compiled as the ratio of the total annual VKT for all motor vehicles registered in Denmark to the total kilometres of road in Denmark.

---

[15] Notice here: $\log(P^V) = \log\left(\frac{P^f}{E}\right) = \log(P^f) - \log(E)$.

We identify each variable using both the generic notation in (3.1) and the variable name used in our empirical specification (3.2). We express all the dependent variables and (most of the) independent variables in natural logarithms because this seems a more plausible relationship and because it is easy to interpret estimation results as elasticities. All monetary variables are real. Table 1 shows summary statistics for the data used in our specification. Data sources are given in Appendix B.

Table 1. *Summary statistics for selected variables*

Variable	Mean	Std. Dev.	Minimum	Maximum
Freight activity Y, ($millions\ tkm$)	9,528	1,370	6,941	11,738
Vehicle kilometre travelled V, ($millions\ VKT$)	2,041	154	1,798	2,364
Actual hours worked in road freight transportation L, (1,000 $hours$)	64,458	4,508	56,150	71,972
Number of trucks M, ($trucks$)	47,220	1,715	44,014	50,764
Average truck capacity S, ($tonnes$)	10.59	1.47	7.60	14.10
Average truck age H, ($years$)	7.01	1.11	5.02	8.40
VKT per litre of consumed fuel E, (km/l)	2.74	0.48	2.18	4.00
Fuel price P^F, (DKK/l)	5.92	0.88	3.71	7.53
Price index for vehicles and spare parts PIT, ($index$)	0.82	0.17	0.43	1.03
Compensation of employees in freight transportation w, ($DKK/hour$)	95.974	32.441	37.065	150.530

Notes: Number of observations: 28. One DKK is approximately 0.13€ in 2005.

The dependent variables are:

Y: Freight activity (logarithm: y).

V: Vehicle kilometre travelled (VKT) (logarithm: v).

L: Actual hours worked in road freight transportation (logarithm: l).

M: Truck stock (logarithm: m).

S: Average truck capacity (logarithm: s).

H: Average truck age (logarithm: h).

E: Number of driven kilometres per litre of consumed fuel (logarithm: e).

The independent variables are:

P^F : Fuel price (logarithm: p^f).

X_E includes index of congestion D (logarithm: d) and time trend to proxy for unmeasured changes (for example technological improvements).

X_L, X_S, X_H, X_M and X_V include the price index for vehicles and spare parts (PIT) (logarithm: pit), and the average compensation of employees in road freight transportation per hour (logarithm: w).

X_Y includes the price index for vehicles and spare parts (PIT) (logarithm: pit), the average compensation of employees in road freight transportation per hour (logarithm: w), and the gross national income (GDP) in constant 2000 prices (logarithm: gdp).

4. Empirical results

Two procedures are available for estimating systems of simultaneous equations containing several endogenous variables, i.e. two-stage least squares (2SLS) and three-stage least squares (3SLS). 2SLS first estimates a reduced form of the system in which each equation contains as variables only the exogenous contemporary variables and (for technical reasons) one lagged value of all the exogenous variables and two lagged values of all endogenous variables (Wooldridge, 2002, ch. 8). Then it estimates each equation by replacing the endogenous variables on its right-hand side by their predicted values from the first stage. 3SLS in addition estimates also correlations in the error terms among equations, and then re-estimates the system taking these correlations into account.[16] This is likely in our system because, for example, unobserved factors like economic expectations might influence both the truck usage and the truck stock. Moreover, there is only little difference between 3SLS and 2SLS estimates. The 3SLS provides slightly better precision of estimates. Thus, there is no indication for problems that might arise from misspecification. We therefore consider the 3SLS results as our best estimates. The ordinary least squares (OLS) results are shown for comparison.[17] Consequently, we present results from two estimation methods: OLS and 3SLS.

We reduce each equation to the simplest form including only the significant variables, due to the small number of observations and high correlation between the factor input prices. So the final model specification was obtained by a systematic process of eliminating the insignificant variables. The results of estimating the final specification of the structural system (3.2) are presented in tables 2-8.

[16] The advantage of 3SLS is that it makes more efficient use of the data, by taking advantage of the information in the correlations among the endogenous variables, and therefore permits a more precise measurement of parameters. The disadvantage is that if there are errors in the specification of one equation, then this error affects the other equations more directly than with 2SLS.

[17] Recall here that the OLS procedure ignores the reverse causation.

4.1. Structural equations

The VKT per litre of consumed fuel equation (Table 2) explains the approximated fuel efficiency for constant average truck capacity and constant average truck age. Most coefficients are measured with good precision and demonstrate strong and plausible effects. Unsurprisingly, increase in average truck capacity and average truck age decreases average VKT per litre of consumed fuel. The effect from an increase in the average truck capacity is -1.12 in the short run and $-1.12/(1-0.31)=-1.62$ in the long run. A one percent increase in the average truck capacity therefore implies a 1.12% decrease in the average VKT per litre of consumed fuel in the short run and 1.62% in the long run. The corresponding effects from an increase in the average truck age are -0.91 in the short run and -1.31 in the long run. Moreover, our measure of congestion has statistically significant negative effect on the average VKT per litre of consumed fuel (negative coefficient on d). The negative effect of congestion can be seen as a confirmation that increasing congestion implies environmental externality in the form of higher fuel use and consequently higher traffic related emissions, a result found by many other researchers. The long run effect of an increase in our measure of congestion is -2.15.

Table 2. *VKT per litre of consumed fuel equation*

	[1]	[2]
	Estimated using OLS	Estimated using 3SLS
Lagged natural logarithm of VKT per litre of consumed fuel (e_{t-1})	0.5410**	0.3096
	(0.1916)	(0.2027)
Natural logarithm of average truck capacity (s)	-0.7658*	-1.1200**
	(0.4092)	(0.4441)
Natural logarithm of average truck age (h)	-0.5829**	-0.9054***
	(0.2619)	(0.2789)
Natural logarithm of index of congestion (d)	-0.8538	-1.4845**
	(0.6149)	(0.5994)
Trend	0.0342**	0.0546***
	(0.0162)	(0.0165)
Constant	6.6992*	10.9149***
	(3.4695)	(3.4668)
Rho	0.0017	0.0366
	(0.2620)	(0.2578)
Adjusted R-squared	0.8620	0.8057
SSE	0.0623	0.0654
No. of observations	27	26

Notes: Dependent variable is the natural logarithm of VKT per litre of consumed fuel (e); ***,**,* indicate that estimates are significantly different from zero at the 0.01, at the 0.05 and the 0.10 level, respectively; standard errors are in parentheses.

The positive significant coefficient associated with the time trend shows a tendency toward a more energy efficient truck stock for a constant average truck capacity and constant average truck age, i.e. presumably due to the improvements in the available technology. The coefficient on the lagged dependent variable implies that VKT per litre of consumed fuel demonstrates considerable inertia in trucking firm behaviour, with the adjustment in VKT per litre of consumed fuel in a given year by approximately 69% percent of the ultimate adjustment. The equation does not exhibit autocorrelation.

Table 3. *Labour demand equation*

	[1] Estimated using OLS	[2] Estimated using 3SLS
Lagged natural logarithm of labour demand (l_{t-1})	0.2759	0.3729
	(0.3437)	(0.2343)
Natural logarithm of fuel price (p^f)	0.1046	0.0801
	(0.0875)	(0.0855)
Natural logarithm of freight activity (y)	0.1721	0.4939**
	(0.1706)	(0.1827)
Natural logarithm of wages (w)	0.0272	-0.1526**
	(0.0730)	(0.0630)
Constant	6.3156	1.9220
	(3.8081)	(1.9680)
Rho	0.6544**	0.2176
	(0.3057)	(0.2409)
Adjusted R-squared	0.8084	0.8085
SSE	0.0194	0.0161
No. of observations	27	26

Notes: Dependent variable is the natural logarithm of labour demand (l); ***,**,* indicate that estimates are significantly different from zero at the 0.01, at the 0.05 and the 0.10 level, respectively; standard errors are in parentheses.

The labour demand (Table 3) is explained, unsurprisingly, by the freight activity and the wages. Increase in the freight activity has positive effect on labour demand (0.49 and 0.79 in the short run and the long run, respectively), while an increase in wages has negative effect on labour demand (-0.15 and -0.24 in the short run and the long run, respectively). The relatively small wage effect on labour demand is possibly due to the fact that a truck has to be operated by a driver regardless of the wage level. The labour demand equation does not exhibit autocorrelation (insignificant coefficient associated with rho).

The average truck capacity equation (Table 4) shows a significant effect of fuel price; but the effect is small (0.12). This effect is however more than four times higher in the long run (0.50). Thus, the trucking firm responses to increase in the fuel costs through expansion of the average truck capacity. The expansion of the average truck capacity increases the fuel use per kilometre (see table 2) but less distance has to be driven for the same payload. We will see that the latter effect offsets the effect of the fuel price on the average truck capacity and that an increase in the fuel price results in the reduction in the overall annual fuel use. The price index for vehicles and spare parts has as expected negative impact on the average truck capacity. Moreover, wages have positive significant effect, possibly because increase in wages decreases the labour demand (see table 3), so for a given freight activity, the trucking firm must extend the average truck capacity in order to be able to ship the same amount of cargo using less labour. The truck capacity demonstrates considerable inertia in trucking firm behaviour. The equation does not exhibit autocorrelation.

Table 4. *Average truck capacity equation*

	[1]	[2]
	Estimated using OLS	Estimated using 3SLS
Lagged natural logarithm of average truck capacity (s_{t-1})	0.8039***	0.7548***
	(0.1670)	(0.1660)
Natural logarithm of fuel price (p^f)	0.1282*	0.1234*
	(0.0656)	(0.0664)
Natural logarithm of price index for vehicles and spare parts (pit)	-0.2843	-0.2175
	(0.1728)	(0.1692)
Natural logarithm of wages (w)	0.1808	0.1697
	(0.1223)	(0.1187)
Constant	1.9333	1.7357
	(1.2015)	(1.1720)
Rho	-0.4108*	-0.2494
	(0.2366)	(0.2201)
Adjusted R-squared	0.9533	0.9439
SSE	0.0159	0.0158
No. of observations	27	26

Notes: Dependent variable is the natural logarithm of average truck capacity (s); ***,**,* indicate that estimates are significantly different from zero at the 0.01, at the 0.05 and the 0.10 level, respectively; standard errors are in parentheses.

The results for the average truck age equation (Table 5) show a small but significant effect of fuel price, indicating that trucking firms response to increases in the fuel cost through rejuvenation of the truck stock, i.e. improvements in the truck technology. The fuel price effect on the average truck age is -0.15 in the short run and -0.53 in the long run. Predictably, the price index for vehicles and spare parts has positive effect on average truck age.

Table 5. *Average truck age equation*

	[1]	[2]
	Estimated using OLS	Estimated using 3SLS
Lagged natural logarithm of average truck age (h_{t-1})	0.5144**	0.7113***
	(0.1912)	(0.2299)
Natural logarithm of fuel price (p^f)	-0.1540	-0.1517*
	(0.0914)	(0.0856)
Natural logarithm of price index for vehicles and spare parts (pit)	0.0856	0.2030
	(0.2122)	(0.2188)
Natural logarithm of wages (w)	0.0294	0.1093
	(0.1493)	(0.1880)
Constant	0.9339	0.1914
	(1.3116)	(1.4101)
Rho	0.9202***	0.6053**
	(0.1153)	(0.2293)
Adjusted R-squared	0.9549	0.9639
SSE	0.0230	0.0171
No. of observations	27	26

Notes: Dependent variable is the natural logarithm of average truck age (h); ***,**,* indicate that estimates are significantly different from zero at the 0.01, at the 0.05 and the 0.10 level, respectively; standard errors are in parentheses.

Wages do not have significant effect on average truck age. The average truck age equation exhibits, as expected, considerable autocorrelation and demonstrates substantial inertia in trucking firm behaviour. The long run effect of the estimated coefficients is approximately 3.5 times higher than the short run effect.

In the truck stock equation (Table 6) most of the coefficients have strong and plausible effects. As expected, fuel price has negative significant effect, but this effect is relatively small (-0.07 in the short run and -0.14 in the long run). Moreover, freight activity does not have significant effect. Since a truck is an ordinary good, the effect of the price index for vehicles and spare parts is, as expected, negative. This effect is however relatively small and significantly different from zero only at 16%. Wages have positive significant effect on truck age (an increase in wages increases the size of the truck stock) possibly for the same reason as for the average truck capacity. Unsurprisingly, there is a considerable inertia in expanding or contracting the truck stock. This most likely reflects the transaction costs of buying and selling trucks. The equation exhibits considerable autocorrelation (significant coefficient associated with rho).

Table 6. *Truck stock equation*

	[1]	[2]
	Estimated using OLS	Estimated using 3SLS
Lagged natural logarithm of average number of trucks (m_{t-1})	0.6233***	0.5252*
	(0.1626)	(0.2592)
Natural logarithm of fuel price (p^f)	-0.0794**	-0.0660*
	(0.0353)	(0.0340)
Natural logarithm of freight activity (y)	-0.0205	0.0343
	(0.0699)	(0.0660)
Natural logarithm of price index for vehicles and spare parts (pit)	-0.0658	-0.1274
	(0.0866)	(0.0874)
Natural logarithm of wages (w)	0.1234**	0.1280*
	(0.0536)	(0.0663)
Constant	4.9708**	5.7835*
	(1.7556)	(2.8322)
Rho	0.8216***	0.7454**
	(0.1583)	(0.2895)
Adjusted R-squared	0.8870	0.8727
SSE	0.0031	0.0033
No. of observations	27	26

Notes: Dependent variable is the natural logarithm of average number of trucks (m); ***,**,* indicate that estimates are significantly different from zero at the 0.01, at the 0.05 and the 0.10 level, respectively; standard errors are in parentheses.

The VKT equation (Table 7) explains the amount of driving performed by the average trucking firm for constant freight activity. The fuel price does not have significant effect on VKT. So, the direct effect of a change in fuel price on traffic volume is unsurprisingly limited, because the trucking firm can only to some extent reduce the traffic volume for a given freight activity through better matching of the trucks' capacity to shipment (see section 2). However, the trucking firm's decision regarding the traffic volume is highly dependent of the freight activity for given factor input prices as indicated by the estimated coefficient associated with the freight activity. The elasticity of VKT with respect to freight activity is 0.40 in the short run and 0.49 in the long run. The price index for vehicles and spare parts does not have a significant effect on VKT. Wages have positive effect on traffic volume, probably because the wage indicator does not adequately measure the truck drivers' wages in this equation but instead the general economic development, since wages rise in periods of economic prosperity. VKT demonstrates mild inertia in trucking firm behaviour, reflecting the time needed to adjust planned travel behaviour. The VKT equation exhibits substantial autocorrelation.

Table 7. *VKT equation*

	[1]	[2]
	Estimated using OLS	Estimated using 3SLS
Lagged natural logarithm of VKT (v_{t-1})	0.3737*	0.1791
	(0.2021)	(0.1744)
Natural logarithm of fuel price (p^f)	-0.0943	-0.0005
	(0.0865)	(0.0872)
Natural logarithm of freight activity (y)	0.1568	0.4029**
	(0.1600)	(0.1629)
Natural logarithm of price index for vehicles and spare parts (pit)	-0.0085	0.0799
	(0.1918)	(0.1865)
Natural logarithm of wages (w)	0.1051	0.3790*
	(0.1251)	(0.2143)
Constant	3.8049*	2.9213
	(1.9057)	(1.7153)
Rho	0.8352***	0.9004***
	(0.1552)	(0.0351)
Adjusted R-squared	0.8616	0.8487
SSE	0.0161	0.0164
No. of observations	27	26

Notes: Dependent variable is the natural logarithm of VKT (v); ***,**,* indicate that estimates are significantly different from zero at the 0.01, at the 0.05 and the 0.10 level, respectively; standard errors are in parentheses.

Table 8 shows the estimation results for freight activity. The fuel price has a significant negative effect and the VKT per litre of consumed fuel (the approximated fuel efficiency) has a positive effect, confirming that an increase in fuel cost will raise the marginal costs of production and consequently decrease the demand for freight activity.[18] The elasticity of freight activity with respect to fuel cost per kilometre is $-0.20 - 0.26 = -0.46$ in the short run and -0.57 in the long run. An increase in GDP has, as expected, a positive and significant effect on freight activity (0.55 and 0.67 in the short run and in the long run, respectively). The dynamic effects are small and insignificant, suggesting that, in the short run, the trucking firms adapt the freight activity to changes in the economic environment.

[18] Recall here that, at the market equilibrium, the output price equals marginal costs.

Table 8. *Freight activity equation*

	[1]	[2]
	Estimated using OLS	Estimated using 3SLS
Lagged natural logarithm of freight activity (y_{t-1})	0.1200	0.1803
	(0.1648)	(0.1522)
Natural logarithm of fuel price (p^f)	-0.1756*	-0.2048**
	(0.0915)	(0.0905)
Natural logarithm of VKT per litre of consumed fuel (e)	0.3654***	0.2614**
	(0.0946)	(0.1000)
Natural logarithm of GDP (gdp)	0.3916*	0.5517**
	(0.2002)	(0.2079)
Natural logarithm of wages (w)	0.1218	0.1066
	(0.2191)	(0.2237)
Natural logarithm of price index for vehicles and spare parts (pit)	0.3871	0.3683
	(0.2442)	(0.2308)
Constant	7.1955***	7.5793***
	(2.0766)	(2.1548)
Rho	-0.0078	0.2373
	(0.0299)	(0.1818)
Adjusted R-squared	0.9655	0.9612
SSE	0.0136	0.0132
No. of observations	27	26

Notes: Dependent variable is the natural logarithm of freight activity (y); ***,**,* indicate that estimates are significantly different from zero at the 0.01, at the 0.05 and the 0.10 level, respectively; standard errors are in parentheses.

4.2 Rebound effect and other elasticities

We consider the 3SLS results our best estimates and use them for the analysis of the determinants of trucking firm fuel use and the rebound effect. Table 9 shows selected elasticities implied by the structural model, the effect of fuel price on trucking firm fuel use, and the rebound effect.

We estimate the rebound effect based on (2.4). In system (3.2), the formula for rebound effect becomes:

$$\varepsilon_{V,Y}\varepsilon_{Y,P^V} = \alpha^{vy}\left(\beta_1^y - \alpha^{ye}\right). \tag{4.1}$$

The long run rebound effect has been calculated using the same formula, and in addition by accounting for lagged values. Our best estimate of the average rebound effect in the applied sample is 18.8% in the short run and 27.9% in the long run (see Table 9). This is in line with a range of other studies (see e.g. Matos and Silva, 2011).[19] The elasticity of VKT with respect to freight activity $(\varepsilon_{V,Y})$ and the elasticity of freight activity with respect to fuel cost per kilometre (ε_{Y,P^V}) are of more or less same magnitude. The elasticity of VKT with respect to freight activity has a slightly smaller effect. This appears to confirm the theoretical expectation that higher fuel

[19] Matos and Silva (2011) estimated the long run rebound effect for the road freight transportation in Portugal to be about 24.1%. Moreover, estimates of personal motor-vehicle rebound effect for the motor vehicles lie within a range of 10-30% (Small and van Dender, 2007; Hymel et al., 2010).

prices first and foremost imply a decrease in freight activity which again has a considerable effect on the traffic volume. Use of OLS underestimates the short run and long run rebound effects by 54.9% and 44.9%, respectively.[20] This is possibly the case because OLS ignores reverse causation.

Table 9. *Rebound effect and other elasticities*

	Short run	**Long run**
Elasticity of freight activity with respect to fuel price (ε_{Y,P^F})	-0.2048	-0.2498
Elasticity of freight activity with respect to fuel efficiency $(\varepsilon_{Y,E})$	0.2614	0.3190
Elasticity of VKT with respect to freight activity $(\varepsilon_{V,Y})$	0.4029	0.4908
Elasticity of VKT with respect to fuel price (ε_{V,P^F})	-0.0005	-0.0006
Elasticity of fuel efficiency with respect to average truck capacity $(\varepsilon_{E,S})$	-1.1200	-1.6222
Elasticity of fuel efficiency with respect to average truck age $(\varepsilon_{E,H})$	- 0.9054	-1.3114
Elasticity of average truck capacity with respect to fuel price (ε_{S,P^F})	0.1234	0.5031
Elasticity of average truck age with respect to fuel price (ε_{H,P^F})	-0.1517	-0.5255
Rebound effect $(-\varepsilon_{V,Y}\varepsilon_{Y,P^V})$	0.1878	0.2792
Elasticity of fuel use with respect to fuel price (ε_{F,P^F})	-0.1877	-0.1883

Notes: All elasticities are estimated using 3SLS; $\varepsilon_{Y,P^V} = \varepsilon_{Y,P^F} - \varepsilon_{Y,E}$.

Table 9 also shows the total effect of a change in fuel price on the average trucking firm fuel use. We estimate this effect based on (2.5). Since the effects of the freight activity on the average truck age and the average truck capacity are not significant, (2.5) reduces to:

$$\varepsilon_{F,P^F} = \varepsilon_{V,P^F} + \varepsilon_{V,Y}\varepsilon_{Y,P^V}\left(1 - \varepsilon_{E,P^F}\right) - \varepsilon_{E,P^F}\ , \tag{4.2}$$

where $\varepsilon_{E,P^F} = \varepsilon_{E,S}\varepsilon_{S,P^F} + \varepsilon_{E,H}\varepsilon_{H,P^F}$. In system (3.2), the formula for the elasticity of fuel use with respect to fuel price becomes:

$$\varepsilon_{F,P^F} = \beta_1^v + \alpha^{vy}\left(\beta_1^y - \alpha^{ye}\right)\left(1 - \varepsilon_{e,p^F}\right) - \varepsilon_{e,p^F}\ , \tag{4.3}$$

where $\varepsilon_{e,p^F} = \alpha^{es}\beta_1^s + \alpha^{eh}\beta_1^h$. The long run effect has been calculated using the same formula, and in addition by accounting for lagged values.

Table 9 shows that higher fuel prices decrease the average trucking firm fuel use, but only by a small amount. The estimation results suggest that the response to a fuel price increase is dominated by changes in the freight activity and the traffic volume rather than changes in the VKT per litre of consumed fuel (approximated fuel efficiency). An 1% increase in the fuel price decreases the VKT through the freight activity ($\varepsilon_{V,Y}\varepsilon_{Y,P^V}$) by 0.19% in the short run and 0.28% in the long run. As discussed in the previous section, an increase in the fuel price has more or less no effect on the VKT. Finally, changes in the VKT per litre of consumed fuel of a change in

[20] Notice the insignificant elasticity of VKT with respect to freight activity in Table 7. So, OLS fails to estimate rebound effect.

the fuel price $(\varepsilon_{E,S}\varepsilon_{S,P^F} + \varepsilon_{E,H}\varepsilon_{H,P^F})$ also affect the trucking industry fuel use. The trucking firm responds to increase in the fuel price through expansion of the average truck capacity and an increase of the average truck capacity decreases the VKT per litre of consumed fuel, presumably because for given VKT trucks with higher capacity use more fuel. The total effect on the VKT per litre of consumed fuel of a change in the fuel price through average truck capacity is -0.14 in the short run and -0.82 in the long run. The trucking firm also responds to an increase in the fuel costs through rejuvenation of the truck stock, and the newer trucks use less fuel per kilometre. The total effect on the VKT per litre of consumed fuel of a change in the fuel price through the average truck age is 0.14 in the short run and 0.69 in the long run. Thus, an increase in the fuel price has negative effect on the average VKT per litre of consumed fuel (approximated fuel efficiency), i.e. a 1% increase in the fuel price decreases average VKT per litre of consumed fuel by 0.001% and 0.13% in the short run and in the long run, respectively. However, less distance has to be driven for the same payload, so the total effect on the average trucking firm fuel use is negative. Thus, an increase in the fuel price results in the reduction in the trucking firm's overall fuel use. The elasticity of fuel use with respect to fuel price is -0.19 in the short run and in the long run.

4.3 Robustness checks

In this section, we discuss the sensitivity of the estimation results to assumptions regarding the model dynamics (autoregression and autocorrelation) and to known problems with the aggregate freight activity data.

First, our estimates (especially long term estimates) rely on assumptions regarding the model dynamics, i.e. the one-year lagged value of the dependent variable (autoregression) and the autoregressive error terms (autocorrelation). Moreover, the role of the one-year lagged value of the dependent variable in determining the long run effect is sensitive to whether or not autocorrelation is controlled for. In order to check the dependence of the estimated effects on the autocorrelation, we estimate a model shown in Appendix C where the autoregression of the error term is omitted. The exclusion of the autoregressive error terms increases the estimates of the rebound effect from 19% to 26% in the short run and from 28% to 69% in the long run. Furthermore, in the unrestricted model (model specification incorporating both autoregression and autocorrelation), the total effect of changes in fuel price on the average trucking firm fuel use (ε_{F,P^F}) in the short run is more or less identical to this in the restricted model (model specification incorporating autoregression but not autocorrelation). The effect of changes in fuel prices on the trucking firm fuel use in the long run is almost two times higher in the restricted model than in the unrestricted model. However, the overall performance of the restricted model is unsatisfactory, especially its dynamic properties. The Godfrey Lagrange multiplier test for serially correlated residuals indicates strong autocorrelation in almost all equations (see Appendix C).[21] So, we use the model specification incorporating both autoregression and autocorrelation, since the Godfrey Lagrange multiplier test for serially correlated residuals rejects the null hypothesis that the errors are serially uncorrelated in the model specification incorporating autoregression but not autocorrelation. Therefore we have some confidence that the resulting estimates of the coefficients of the lagged endogenous variables in the preferable empirical specification are accurate and give a valid indication of the extent of long-run effects. Furthermore, including both autoregression and autocorrelation does not seem to affect the precision of the other estimates.[22]

The second robustness check concerns the aggregate freight activity data collected by the Statistics Denmark. Data are collected in quarterly sample surveys. The response rate at the closing of the survey is relatively high (98%).[23] However, about 40% of the questionnaires do not contain journey data. In these questionnaires selected vehicles were inactive in the reference period because of lacking orders, holiday closure, or vehicle technical service. Consequently, we have reason to think that this exceptionally high share of inactive vehicles in the reference period biases the estimation results. However, if the aggregate freight activity data are underestimated every year by the roughly same percent, then the impact of this high share of inactive vehicles in the reference period on the estimation results will be minimal. We have no reason to think that the sources of measurement error are persistent over time and unrelated to the independent variables, and because we do not have information of the share of inactive vehicles for every

[21] The null hypothesis of Godfrey's tests is that the equation residuals are white noise. However, if the equation includes autoregressive error model of order t ($AR(t)$) the test is for the null hypothesis that the structural errors are from an $AR(t+1)$ process versus the alternative hypothesis that the errors are from an $AR(t)$ process.
[22] The estimation of a specification including the two-year lagged value of the dependent variable could not be performed due to the limited number of observations.
[23]http://www.dst.dk/HomeUK/Guide/documentation/Varedeklarationer/emnegruppe/emne.aspx?sysrid=992 (accessed 25/12 2010).

year, we can only conclude that better data on the aggregate freight activity would add considerably to the confidence in estimation results.

5. Conclusion

This paper analyses the determinants of road freight transportation fuel use. We develop a simple model to show that the trucking firm fuel use depends on traffic volume, freight activity, characteristics of the truck stock, factor input prices, and congestion. We show that the rebound effect for road freight transportation can be decomposed into the negative of the product of the elasticity by which changes in fuel costs affect the freight activity and the elasticity by which changes in freight activity affects traffic volume. The model is applied to Danish aggregate time series data covering the years 1980-2007. The empirical results provide some insights into the determinants of the road freight transportation fuel use.

We find that higher fuel prices decrease the trucking firm fuel use, but only by a small amount. Surprisingly, an increase in the fuel price has negative effect on the fuel efficiency, i.e. a 1% increase in the fuel price decreases the fuel efficiency by 0.13% in the long run. However, less distance has to be driven for the same payload, so an increase in the fuel price results, as expected, in the reduction in the trucking firm fuel use. Moreover, we find that the short run and the long run rebound effects for road freight transportation are 19% and 28%, respectively.

Analyses of the determinants of the trucking firm fuel use and estimates of the rebound effect are highly relevant for policy. For example, measurements of the rebound effect for road freight transportation can contribute to the ongoing debate whether to adapt the rules on the optimal weights and dimensions of heavy trucks in EU. Arki (2009) shows that introducing longer and heavier vehicles (up to 20.75 meters, 44 tonnes) Europe-wide will be overall beneficial for society. Moreover, Arki (2009) argues that the introduction of longer and heavier vehicles could lower fuel consumption of road freight transportation by 3.6%. We showed that an increase in the weight of heavy trucks will reduce the fuel efficiency and consequently affect the fuel cost per kilometre implying the rebound effect that to some extent will offset the original energy saving. So, the introduction of longer and heavier vehicles will most likely not result in a 3.6% fuel saving, but only in a 2.6% reduction due to the rebound effect. This stresses the importance of including rebound effects in assessments of new policies. Moreover, strengthening fuel efficiency standards for heavy trucks in the EU can potentially result in undesirable effects on traffic congestion, because strategies that increase fuel efficiency, and therefore reduce the per-kilometre cost of driving, tend to increase total truck use. It is therefore important to account for the rebound effect to more accurately evaluate energy policy changes.

Acknowledgements
The author thanks Bruno de Borger for helpful discussions on earlier versions of this paper and Mogens Fosgerau and Jørgen Birk Mortensen for comments. The author is grateful to Statistics Denmark and Kaj Jørgensen (RISØ DTU) for providing the data. Research support from the Danish Energy Agency (Energy Research Programme) is acknowledged.

Appendix A. Rebound effect

Assume now that E exogenously changes. We know that traffic volume V depends (among other things) on the fuel price and freight activity (see section 2.1). Moreover, we know that the freight activity depends (among other things) on marginal costs, a part of which is the per-kilometre fuel cost (P^V).[24] Fuel consumption and VKT are related through fuel efficiency (see (2.1)):

$$F = \frac{V\left(Y(P^V),\bar{D},w,P^F,\boldsymbol{\alpha}\right)}{\bar{\mathrm{E}}} \quad , \tag{A.1}$$

where $\mathrm{P^V} = \frac{\mathrm{P^F}}{\mathrm{E}}$. Differentiating (A.1) with respect to E, we have:

$$\frac{\partial F}{\partial E} = \frac{-P^F}{E^3}\frac{\partial V}{\partial Y}\frac{\partial Y}{\partial P^V} - \frac{V}{E^2}. \tag{A.2}$$

Now multiplying both sides with E/F and rearranging we get:

$$\varepsilon_{F,E} = -1 - \varepsilon_{V,Y}\varepsilon_{Y,P^V}. \tag{A.3}$$

Notice now that, using the solution to PMP, fuel use can be shown to be:

$$\mathrm{F} = \frac{\mathrm{V}\left(\mathrm{Y}\left(\frac{\mathrm{P^F}}{\mathrm{E}\left(\mathrm{S}(\mathrm{P^F},\bar{\mathrm{Y}},\bar{\mathrm{D}},\mathrm{w},\boldsymbol{\alpha}),\mathrm{H}(\mathrm{P^F},\bar{\mathrm{Y}},\bar{\mathrm{D}},\mathrm{w},\boldsymbol{\alpha})\right)}\right),\bar{\mathrm{D}},\mathrm{w},\mathrm{P^F},\boldsymbol{\alpha}\right)}{\mathrm{E}\left(\mathrm{S}\left(\mathrm{Y}\left(\frac{\mathrm{P^F}}{\bar{\mathrm{E}}}\right),\bar{\mathrm{D}},\mathrm{w},\mathrm{P^F},\boldsymbol{\alpha}\right),\mathrm{H}\left(\mathrm{Y}\left(\frac{\mathrm{P^F}}{\bar{\mathrm{E}}}\right),\bar{\mathrm{D}},\mathrm{w},\mathrm{P^F},\boldsymbol{\alpha}\right)\right)}. \tag{A.4}$$

Moreover, a simple calculation using the definition of elasticity shows that:

$$\frac{\partial F}{\partial \mathrm{P^F}} = \frac{1}{E}\left[\frac{\partial V}{\partial \mathrm{P^F}} + \frac{\partial V}{\partial \mathrm{Y}}\frac{\partial Y}{\partial \mathrm{P^V}}\left(\frac{1}{E} - \frac{\mathrm{P^F}}{E^2}\left(\frac{\partial E}{\partial \mathrm{S}}\frac{\partial S}{\partial \mathrm{P^F}} + \frac{\partial E}{\partial \mathrm{H}}\frac{\partial H}{\partial \mathrm{P^F}}\right)\right)\right]$$

$$-\frac{V}{E^2}\left(\frac{\partial E}{\partial \mathrm{S}}\left(\frac{\partial S}{\partial \mathrm{P^F}} + \frac{\partial S}{\partial \mathrm{Y}}\frac{\partial Y}{\partial \mathrm{P^V}}\frac{1}{E}\right) + \frac{\partial E}{\partial \mathrm{H}}\left(\frac{\partial H}{\partial \mathrm{P^F}} + \frac{\partial H}{\partial \mathrm{Y}}\frac{\partial Y}{\partial \mathrm{P^V}}\frac{1}{E}\right)\right)$$

$\Leftrightarrow$

[24] Smith (1957) showed that the trucking firm fuel use is a function of the VKT and the vehicle gross weight, while the total aggregate fuel use by the trucking industry is a function of both VKT and freight activity.

$$\varepsilon_{F,P^F} = \varepsilon_{V,P^F} + \varepsilon_{V,Y}\varepsilon_{Y,P^V}\left(1 - \varepsilon_{E,P^F}\right) - \varepsilon_{E,P^F} - \varepsilon_{E,P^V}\ , \quad \text{(A.5)}$$

where $\varepsilon_{E,P^F} = \varepsilon_{E,S}\varepsilon_{S,P^F} + \varepsilon_{E,H}\varepsilon_{H,P^F}$ and $\varepsilon_{E,P^V} = \varepsilon_{Y,P^V}\left(\varepsilon_{E,S}\varepsilon_{S,Y} + \varepsilon_{E,H}\varepsilon_{H,Y}\right)$.

If $\varepsilon_{S,Y} = 0$ and $\varepsilon_{H,Y} = 0$, then:

$$\varepsilon_{F,P^F} = \varepsilon_{V,P^F} + \varepsilon_{V,Y}\varepsilon_{Y,P^V}\left(1 - \varepsilon_{E,P^F}\right) - \varepsilon_{E,P^F}. \quad \text{(A.6)}$$

Appendix B. Data sources

Aggregate freight activity has been compiled by the National Environmental Research Institute – Aarhus University from several different reports (Statistics Denmark, 2000; The Danish Car Importers Association, 2001-2008; The Danish Road Directorate, 1998; Winther, 2007), which in turn are based on data submitted by enterprises performing transport for their own account or for hire or reward. The data are collected by Statistics Denmark in quarterly sample surveys including trucks over 6 tonnes of maximum permissible weight. The survey is described in Statistics Denmark's online documentation.[25] *Aggregate VKT* of all trucks registered in Denmark has been compiled by Statistics Denmark based on exact odometer readings from the so-called MOT tests, a more accurate basis than asking respondents to remember VKT.[26] Data on the *size of the truck stock* are published regularly by Statistics Denmark in "News from Statistics Denmark" ("Nyt fra Danmarks Statistik"), in the series "Statistical News" ("Statistiske Efterretninger"), and in Statistics Denmark's online-database www.statbank.dk (accessed 25/12 2010). *Average truck capacity* (measured as axle load in kilograms) and *average truck age* are computed from administrative register data. Data on *fuel consumption* are taken from Danish Environmental Accounts (see www.statbank.dk); the statistics on fuel consumption are reprinted in many sources, such as Winther (2007). *Fuel prices* are from The Danish Petroleum Association web page (http://oliebranchen.dk/da-DK/Service/English.aspx (accessed 25/12 2010)). Applied infrastructure measure (*kilometre road in Denmark*) is easily taken from the Danish Road Directorate's online database (http://www.vejdirektoratet.dk (accessed 25/12 2010)). Data on *total actual hours worked in road freight transportation*, *compensation of road freight transportation employees, price index for vehicles and spare parts*, and *GDP* are taken

[25] For detailed description of the survey see http://www.dst.dk/HomeUK/Guide/documentation/Varedeklarationer/ (accessed 25/12 2010).

[26] The MOT test is a vehicle check that is compulsory for all vehicles registered in Denmark. The name derives from the Ministry of Transport. All Danish trucks have to pass such MOT tests when first registered, and then at statutory time intervals, i.e. every year. Each time a truck passes the MOT test, the inspection authority reads the odometer on the day of the MOT test, records date of the MOT test and several different identification data regarding the vehicle.

from Statistics Denmark's online database www.statbank.dk. The data is available from the author on request.

Appendix C. Estimation results for specifications without control for autocorrelation

Table C1. *VKT per litre of consumed fuel equation*

	[1]	[2]
	Estimated using OLS	Estimated using 3SLS
Lagged natural logarithm of VKT per litre of consumed fuel (e_{t-1})	0.5414***	0.3843**
	(0.1668)	(0.1547)
Natural logarithm of average truck capacity (s)	-0.7664*	-0.9942**
	(0.3991)	(0.4149)
Natural logarithm of average truck age (h)	-0.5826**	-0.8434***
	(0.2460)	(0.2283)
Natural logarithm of index of congestion (d)	-0.8527	-1.1583**
	(0.5695)	(0.5164)
Trend	0.0341**	0.0465***
	(0.0155)	(0.0146)
Constant	6.6952*	9.0816***
	(3.2353)	(2.9981)
Adjusted R-squared	0.8686	0.8226
SSE	0.0623	0.0629
Godfrey LM test statistics	0.00	1.48
No. of observations	27	26

Notes: Dependent variable is the natural logarithm of VKT per litre of consumed fuel (e); ***,**,* indicate that estimates are significantly different from zero at the 0.01, at the 0.05 and the 0.10 level, respectively; standard errors are in parentheses.

Table C2. *Labour demand equation*

	[1]	[2]
	Estimated using OLS	Estimated using 3SLS
Lagged natural logarithm of labour demand (l_{t-1})	0.5339**	0.2654
	(0.2115)	(0.1602)
Natural logarithm of fuel price (p^f)	0.0509	0.0666
	(0.0851)	(0.0730)
Natural logarithm of freight activity (y)	0.2901	0.5918***
	(0.2107)	(0.1756)
Natural logarithm of wages (w)	-0.0628	-0.1683***
	(0.0634)	(0.0568)
Constant	2.2669	2.2018 *
	(1.5754)	(1.2640)
Adjusted R-squared	0.7778	0.7900
SSE	0.0235	0.0186
Godfrey LM test statistics	4.52	3.64
No. of observations	27	26

Notes: Dependent variable is the natural logarithm of labour demand (l); ***,**,* indicate that estimates are significantly different from zero at the 0.01, at the 0.05 and the 0.10 level, respectively; standard errors are in parentheses.

Table C3. *Average truck capacity equation*

	[1]	[2]
	Estimated using OLS	Estimated using 3SLS
Lagged natural logarithm of average truck capacity (s_{t-1})	0.7382***	0.7781***
	(0.2072)	(0.1763)
Natural logarithm of fuel price (p^f)	0.1448*	0.1336*
	(0.0783)	(0.0742)
Natural logarithm of price index for vehicles and spare parts (pit)	-0.1460	-0.1185
	(0.2079)	(0.2075)
Natural logarithm of wages (w)	0.1213	0.1042
	(0.1535)	(0.1463)
Constant	1.3083	1.0716
	(1.4981)	(1.4513)
Adjusted R-squared	0.9498	0.9426
SSE	0.0179	0.0169
Godfrey LM test statistics	3.09	2.18
No. of observations	27	26

Notes: Dependent variable is the natural logarithm of average truck capacity (s); ***,**,* indicate that estimates are significantly different from zero at the 0.01, at the 0.05 and the 0.10 level, respectively; standard errors are in parentheses.

Table C4. *Average truck age equation*

	[1]	[2]
	Estimated using OLS	Estimated using 3SLS
Lagged natural logarithm of average truck age (h_{t-1})	0.8504***	0.8273***
	(0.1121)	(0.0698)
Natural logarithm of fuel price (p^f)	-0.0277	-0.0219
	(0.1086)	(0.0891)
Natural logarithm of price index for vehicles and spare parts (pit)	-0.0836	0.0405
	(0.3187)	(0.2449)
Natural logarithm of wages (w)	0.0875	0.0583
	(0.2099)	(0.1559)
Constant	0.9324	0.3459
	(2.0443)	(1.5417)
Adjusted R-squared	0.9220	0.9444
SSE	0.0416	0.0277
Godfrey LM test statistics	6.81	5.71
No. of observations	27	26

Notes: Dependent variable is the natural logarithm of average truck age (h); ***,**,* indicate that estimates are significantly different from zero at the 0.01, at the 0.05 and the 0.10 level, respectively; standard errors are in parentheses.

Table C5. *Truck stock equation*

	[1] Estimated using OLS	[2] Estimated using 3SLS
Lagged natural logarithm of average number of trucks (m_{t-1})	0.6390***	0.5516***
	(0.1683)	(0.0968)
Natural logarithm of fuel price (p^f)	-0.0725	-0.0659
	(0.0475)	(0.0423)
Natural logarithm of freight activity (y)	0.0390	0.0689
	(0.0934)	(0.0613)
Natural logarithm of price index for vehicles and spare parts (pit)	-0.2200	-0.2135*
	(0.1333)	(0.1047)
Natural logarithm of wages (w)	0.1669*	0.1484**
	(0.0823)	(0.0646)
Constant	5.0243**	5.6073***
	(1.8823)	(1.1877)
Adjusted R-squared	0.7728	0.7733
SSE	0.0066	0.0061
Godfrey LM test statistics	15.73	10.93
No. of observations	27	26

Notes: Dependent variable is the natural logarithm of average number of trucks (m); ***,**,* indicate that estimates are significantly different from zero at the 0.01, at the 0.05 and the 0.10 level, respectively; standard errors are in parentheses.

Table C6. *VKT equation*

	[1] Estimated using OLS	[2] Estimated using 3SLS
Lagged natural logarithm of VKT (v_{t-1})	0.4993**	0.5508***
	(0.2152)	(0.1684)
Natural logarithm of fuel price (p^f)	0.0316	-0.0284
	(0.1063)	(0.0946)
Natural logarithm of freight activity (y)	0.4036*	0.5434***
	(0.2069)	(0.1660)
Natural logarithm of price index for vehicles and spare parts (pit)	-0.1992	-0.2940
	(0.2317)	(0.2120)
Natural logarithm of wages (w)	0.0392	0.0298
	(0.1528)	(0.1362)
Constant	1.0346	-0.1330
	(1.9393)	(1.6669)
Adjusted R-squared	0.8188	0.8381
SSE	0.0221	0.0185
Godfrey LM test statistics	4.17	1.32
No. of observations	27	26

Notes: Dependent variable is the natural logarithm of VKT (v); ***,**,* indicate that estimates are significantly different from zero at the 0.01, at the 0.05 and the 0.10 level, respectively; standard errors are in parentheses.

Table C7. *Freight activity equation*

	[1]	[2]
	Estimated using OLS	Estimated using 3SLS
Lagged natural logarithm of freight activity (y_{t-1})	0.1006	0.1351
	(0.1413)	(0.1410)
Natural logarithm of fuel price (p^f)	-0.1670*	-0.1714*
	(0.0837)	(0.0884)
Natural logarithm of VKT per litre of consumed fuel (e)	0.3667***	0.3196***
	(0.0920)	(0.0870)
Natural logarithm of GDP (gdp)	0.3829*	0.4731**
	(0.1928)	(0.1732)
Natural logarithm of wages (w)	0.1321	0.0464
	(0.2091)	(0.1969)
Natural logarithm of price index for vehicles and spare parts (pit)	0.3919	0.4304*
	(0.2377)	(0.2317)
Constant	7.3584***	6.8632***
	(1.9136)	(1.8520)
Adjusted R-squared	0.9671	0.9618
SSE	0.0136	0.0137
Godfrey LM test statistics	0.78	2.08
No. of observations	27	26

Notes: Dependent variable is the natural logarithm of freight activity (y); ***,**,* indicate that estimates are significantly different from zero at the 0.01, at the 0.05 and the 0.10 level, respectively; standard errors are in parentheses.

References

ACEA, 2007. Commercial vehicles - key facts and figures. Technical report, European Automobile Manufacturers Association.

Arki, H., 2009. European study on heavy goods vehicles' weights and dimensions. Paper Presented at European Transport Conference, Leeuwenhorst Conference Centre, The Netherlands, 05/10/2009 - 07/10/2009.

Baker, G. P. and Hubbard, T. N., 2003. Make versus buy in trucking: asset ownership, job design, and information. The American Economic Review, vol. 93, no. 3, 551-572.

Bjørner, T. B., 1999. Environmental benefits from better freight transport management: Freight traffic in a VAR model. Transportation Research Part D, 4, 45-64.

Bjørner, T. B. and Jensen, T. C., 1997. Gods på vej eller bane?. Nationaløkonomisk Tidsskrift, 135, 144-159.

Bonney, M. C., 1994. Trends in inventory management. International Journal of Production Economics, Volume 35, 107-114.

Calthrop, E. and Proost, S., 2003. Environmental pricing in transport. in D. Hensher and K. Button (eds.), Handbook of transport and environment, Elsevier, Oxford.

Forkenbrock, D. J., 1999. External costs of intercity truck freight transportation. Transportation Research Part A, 33, 505-526.

Fullerton, D. and West, S., 2001. Tax and subsidy combinations for the control of car pollution. Workshop on environmental economics, Venice International University, San Servolo, 18-19 July 2001.

Greening, L. A., Greene, D. L. and Difiglio, C., 2000. Energy efficiency and consumption - the rebound effect - a survey. Energy Policy, 28, 389-401.

Hubbard, T. N., 2003. Information, Decisions, and Productivity: On-Board Computers and Capacity Utilization in Trucking. The American Economic Review, vol. 93, no. 4, 1328-1353.

Hymel, K. M., Small, K. A. and Van Dender, K., 2010. Induced demand and rebound effects in road transport. Transportation research part B, 44, 1220-1241.

IEA, 2006. World energy outlook 2006. Technical report, International Energy Agency.

IEA, 2010. Transport energy efficiency – Implementation of IEA recommendations since 2009 and next steps. Technical report, International Energy Agency.

Kamakaté, F. and Schipper, L., 2008. Trends in truck freight energy use and carbon emissions in selected OECD countries from 1973 to 2003. TRB 2008 Annual Meeting CD-ROM.

Kveiborg, O. and Fosgerau, M., 2007. Decomposing the decoupling of Danish road freight traffic growth and economic growth. Transport Policy, Volume 14, 39-48.

Léonardi, J. and Baumgartner, M., 2004. CO2 efficiency in road freight transportation: Status quo, measures and potential. Transportation Research Part D, 9, 451.464.

Mas-Colell, A., Whinston, M. D. and Green, J. R., 1995. Microeconomic theory. Oxford University Press.

Matos, F. J. S. and Silva, F. J. F., 2011. The rebound effect on road freight transport: Empirical evidence from Portugal. Energy Policy, doi:10.1016/j.enpol.2011.02.056.

McKinnon, A., 2003. Logistic and the environment. In: Hensher, D. A. and Button, K. J. (eds.), Handbook of transport and the Environment. Elsevier, 665-685.

Oum, T. H., Waters, W. G. and Yong, J.-S., 1992. Concepts of price elasticities of transport demand and recent empirical estimates: An interpretative survey. Journal of Transport Economics and Policy, Vol. 26, No. 2, 139-154.

Parry, I. W., 2008. How should heavy-duty trucks be taxed?. Journal of Urban Economics, 63, 651-668.

Rich, J., Kveiborg, O. and Hansen, C. O., 2010. On structural inelasticity of modal substitution in freight transport. Journal of Transport Geography, article in press.

Samuelsson, A. and Tilanus, B., 2002. A framework efficiency model for goods transportation, with an application to regional less-than-truckload distribution. in K. B. A. McKinnon and P. Nijkamp (eds.), Transport Logistics Classics in Transport Analysis, Elgar, Cheltenham.

Sathaye, N., Horvath, A. and Madanat, S., 2010. Unintended impacts of increased truck loads on pavement supply-chain emissions. Transportation Research Part A, 44, 1-15.

Schipper, L. S. and Price, L., 1997. Energy use and carbon emissions from freight in 10 industrialized countries: An analysis of trends from 1973 to 1992. Transportation Research Part D, 2, 57-76.

Small K.A. and Van Dender, V., 2005. A study to evaluate the effect of reduced greenhouse gas emissions on vehicle miles travelled. Final report for the California Air Resources Board (ARB), number 02-336.

Small, K. A. and Van Dender, K., 2007. Fuel efficiency and motor vehicle travel: The declining rebound effect. The Energy Journal, Vol. 28, No. 1, 25-51.

Smith, V. L., 1957. Engineering data and statistical techniques in the analysis of production and technological change: fuel requirements of the trucking industry. Econometrica, vol. 25, no. 2, 281-301.

Statistics Denmark, 2000. Transport 2000. Statistics Denmark, Copenhagen 2000.

The Danish Car Importers Association, 2001-2008, Vejtransporten i tal og tekst. The annual report "Road transport – in text and figures", The Danish Car Importers Association.

The Danish Road Directorate, 1998. Tal om vejtraffik. Report 156.

WBCSD, 2004. Mobility 2030: Meeting the challenges to sustainability. Technical report, World Business Council for Sustainable Development.

Winther, M., 2007. Danish emission inventories for road transport and other mobile sources. Technical Report Technical Report NERI Research Report 236, DMU.

Wooldridge, J. M., 2002. Econometric analysis of cross section and panel data. The MIT Press.

Zlatoper, T. J. and Austrian, Z., 1989. Freight transportation demand: A survey of recent econometric studies. Transportation, 16, 24-46.