The Determinants of Energy Demand of the Swiss Private Transportation Sector

Inaugural dissertation

in fulfillment of the requirements for the degree of

Doctor rerum oeconomicarum

at the Faculty of Business, Economics

and Social Sciences of the University of Bern

Submitted by

RETO TANNER

from Eriswil, Bern

2012

The faculty accepted this work as dissertation on 23 February 2012 at the request of the two advisors Prof. Dr. Klaus Neusser [advisor] and Prof. Dr. Denis Bolduc [second advisor], without wishing to take a position on the view presented therein.

Table of Contents

Appendix

List of Figures

Appendix

Mathematical Appendix

List of Tables

Appendix

Introduction

The transportation sector plays an important role with respect to government expenditure and carbon dioxide emissions. Since traffic demand is gradually increasing, additional and increasingly expensive infrastructure needs to be built. In addition, the rise in traffic demand and the trend towards larger cars have led to a gradual increase in carbon dioxide emissions in recent decades. Since driving demand and car ownership levels play a dominant role with respect to the use of traffic infrastructure and carbon dioxide emissions, these aspects of the private transportation sector shall be examined in this dissertation. The principle aim of this dissertation is to explain the impact of fuel and car ownership taxes on driving demand and on car ownership by private households. The differences in the effects arising if these tax revenues are reimbursed to households or deployed for additional government expenditures will also be shown. With these results, policy-makers can choose from a mix of such taxes to achieve a particular level of car ownership and a certain aggregate level of driving distance. Since household income rises over the car's lifetime, I also show how income affects these two measures. Since the household's car choice also plays an important role with respect to the aggregate fleet consumption, I will also present a model that explains households' willingness to pay for more fuel-efficient cars.

The study is based on two cross-sectional datasets and two datasets with a panel structure based on a stated preference survey. Since fuel prices vary only slightly in these cross-sectional datasets and the fixed costs of cars do not vary between households for a certain car type, the values of interest cannot be explained using traditional models. To overcome this problem, I use the so-called Multiple Discrete-Continuous Extreme Value Model (MDCEV), first introduced by Bhat (2005). The model is based on a microeconomic framework using a certain type of direct utility with a random component. In contrast to a study by Bhat (2008) on car ownership and use, the model I develop captures the fixed costs of cars and explains the household's demand for total annual driving distance.

The principle findings determined using this MDCEV model are that the elasticities of fuel demand I generated (-0.28 .. -0.25) are larger than those reported by traditional models using stated preference data (-0.17). These values (-0.28 .. -0.25) are also larger than those determined by Baranzini et al. (2009) for fuel demand using Swiss time series data (-0.20), but smaller than the average values found in international studies (-0.31), such as Graham and Glaister (2004). A further result is that a tax on fuel is much more effective at reducing the aggregate driving distance than a tax on car ownership. According to the model, the impact of a car ownership tax on driving distance is approximately four times smaller per unit of tax revenue than that of a fuel tax. Since, based on a stated preference dataset, the impact of fuel prices on the household's car choice with respect to fuel efficiency is rather low compared to other countries, there is no significant difference between the fuel price elasticity of driving demand (-0.25) and the fuel price elasticity of fuel demand (-0.31). An additional finding is that households' driving demand increases by about 47% when they move from urban to rural areas.

The willingness to pay for fuel-efficient cars is computed using a Multinomial Logit model. The results show that there is a rather high heterogeneity between specific household segments. On average, however, the households' willingness to pay for fuel efficiency of a car was identical to the amount they expected to save on fuel over the period during which they own that car. This is a strong result since some of the measures taken by the Swiss government to reduce the average consumption level of car fleets rely intrinsically on households' motivated behaviour.

The dissertation is structured as follows: In Chapter 1, I provide an overview of the results of other studies and different types of models. I shall also describe the datasets I use and explain the models used to compute the results. In Chapter 2, I present the results yielded by traditional models, namely the OLS and the Tobit model. In the main chapter of this thesis, Chapter 3, I present the MDCEV model. After presenting the unmodified MDCEV model, I derive the modified MDCEV model that captures the fixed costs of car ownership and show the corresponding results. At the end of Chapter 3, I compare the results yielded by the Tobit and the MDCEV model using different datasets. The results of the research concerning household behaviour with respect to car choice are presented in Chapter 4. This thesis also has an extensive appendix, containing proofs and the detailed results of the different models. Finally, a mathematical appendix contains a list of mathematical rules of the extreme value distribution on which the MDCEV and the Logit model used in Chapter 4 are based. The mathematical appendix also contains a chapter in which I derive the error correction term of the discrete-continuous model of Dubin and McFadden (1984). This derivation of the widely used error correction term can not yet be found in the literature. However, I do not present any results computed using the discrete-continuous model of Dubin and McFadden, since the available data does not allow the computation of satisfactory results.

Acknowledgements

I offer my thanks and regards to all those who supported me in any respect during the completion of this project, namely:

Klaus Neusser (first referee)
Denis Bolduc (second referee)
Stefan Boes
Michael Gerfin
Ralph Winkler
Marc Blatter
Michael Schmutz
Michael Mayer
Lutz Dümbgen
Guido Baldi
Dolores Messer
Samuel Mühlemann
Christoph Leuenberger
Teresa Gehrs
Kai Axhausen (for the dataset)
Rolf Wüstenhagen (for the dataset)

1. An overview of the models and data

1.1 Introduction

In this first chapter, I will provide an overview of the field of transportation research. I will explain how the current traffic volume evolved, and will present the results of several Swiss and international studies. Furthermore, I will present models that can be used to address certain questions in the field of transportation research. Finally, I will describe the datasets which are used and provide a brief introduction to the chosen models.

More specifically, in Subchapter 1.2, I will show a set of diagrams that illustrate how several key variables – such as the average driving distance and the ratio of carless households – evolved over time on an aggregated level. These variables will be presented for three countries: the USA, Germany and Switzerland. I chose these countries because they are all highly developed and have a similar level of average household income. Since the annual driving distances per capita and the level of car ownership in the USA differ dramatically from those in Germany and Switzerland, I will present data that may help to explain these differences. In contrast, the data on car use and car ownership in Switzerland and Germany differ only slightly due to these countries' similar geographical structure, service level of public transportation and income per capita.

In Subchapter 1.3, I provide an overview of previous studies undertaken at Swiss and international level. First, I will present the results of a couple of meta-analyses of studies from various countries. These results will then be grouped by type of elasticity, e.g. the income and fuel price elasticity of aggregate vehicle kilometres, whether they are short- or long-run elasticities, and which type of model was used to compute them. I will distinguish between time-series models and models based on cross-sectional or panel data. Only a few studies exist on car use and travel demand in Switzerland. For this reason, I consider the corresponding results to be benchmarks to which I will refer when discussing the results computed using my models.

In Subchapter 1.4, I will briefly describe all of the datasets at the Swiss household level used for my analysis. I used three sets of revealed preference micro-census data on the travelling behaviour of Swiss private households in 2000 and 2005. This data contains the annual distances driven in cars and the data of consumer surveys from 2000 to 2005 containing information on household fuel consumption. In addition, I used two stated preference datasets. The first involves car ownership and use, given different fictitious price levels of fuel. The second dataset was created from a survey that focused exclusively on the choice of car type. Households had to select a car from a set of cars with fictitious fuel efficiencies and prices. These stated preference datasets are very valuable because the explanatory variables of interest, such as fuel price and fuel efficiency, vary considerably, enabling the corresponding elasticities to be computed using standard models.

In the final subchapter of Chapter 1, I will present the different types of model commonly used in transportation research. I will then discuss the properties and limitations of each model to determine whether they suit the purpose of my thesis, given the available data. I will then proceed to state which models I decided to use, substantiating my choice. In particular, I shall elaborate why I chose to extend the multiple Discrete-Continuous Extreme Value Model (MDCEV) for my application.

1.2 Some stylized facts

Before introducing the different models to explain the aggregate driving demand, car stock and average fuel efficiency, I shall present some stylized facts. The purpose of this presentation is to identify the key variables that influence both car ownership and driving demand of private households. In the following, I always compare data from the USA, Germany and Switzerland covering the last decades. The reason I chose these countries is that Germany is a very similar country to Switzerland with respect to income and the traffic system, particularly regarding the service level of the public transportation system. In contrast, the spacial structure, price level of fuel and the service level of public transport in the USA are completely different to those in Germany and Switzerland. First, I will present the evolution of driving demand and the ratio of carless households. I will then present trajectories of variables, such as the Gross Domestic Product (GDP) per capita and fuel prices, which I believe influence the two measures. In the process, I shall put forward hypotheses on relations between variables. After stating my hypotheses, I will identify interesting information that could be provided by a model, and will derive basic requirements on models at the end of this subchapter.

Two of the principal measures for describing driving demand generated by private households on an aggregate level are driving distance per capita and the share of households that own at least one car. Here, I first present the trajectories of these measures for Switzerland, Germany and the USA. I will then show the trajectories of other variables to identify the forces behind these two measures.

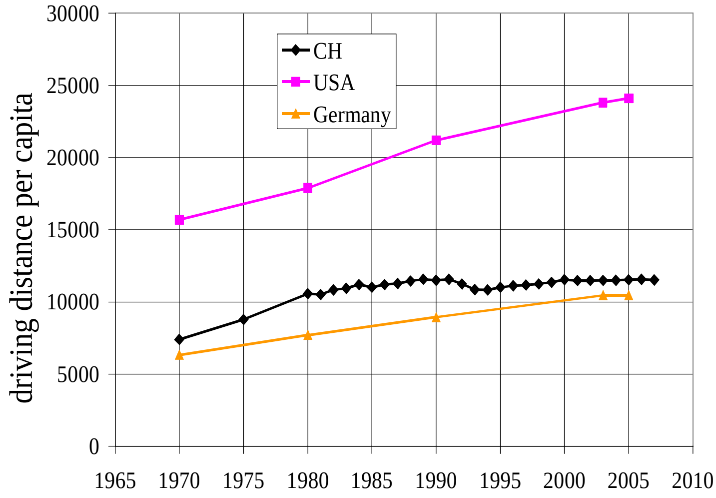


Figure 1.2.1: Annual distances driven per capita in Switzerland, Germany and the USA.[21]

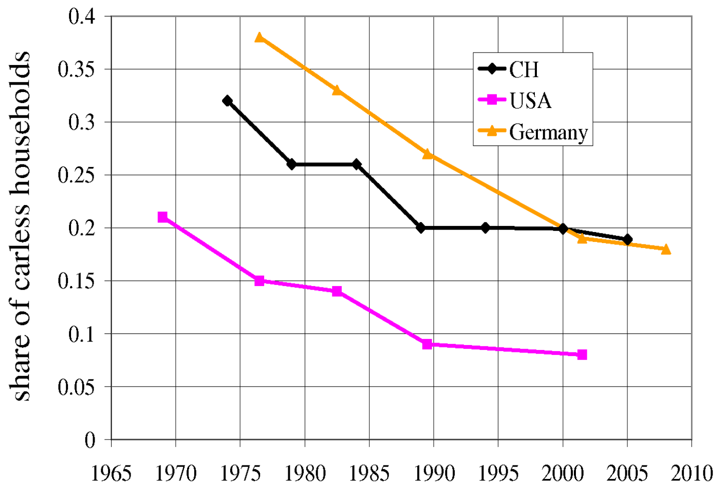


Figure 1.2.2: Ratio of carless households in Switzerland, Germany and the USA.[22]

Figures 1.2.1 and 1.2.2 show that there is a co-movement between driving demand and the ratio of households that own at least one car: both measures increase over time, but the growth rate is decreasing. The figures also reveal that both car use and car ownership are similar in Germany and Switzerland, whereas these figures are much higher in the USA. It seems natural to endeavour to explain these differences and the trajectories by differences in fuel prices and income per capita. First, I present the trajectories of fuel prices.

[21] This data was collected by Deutsches Institut für Wirtschaftsforschung (2002), Deutsches Institut für Wirtschafts-forschung (2010), Federal Highway Administration (2010) for data on vehicle distance travelled and is further contained in Heston et al. (2009).

[22] This data is based on Simma (2004) and Bühler and Kunert (2008).

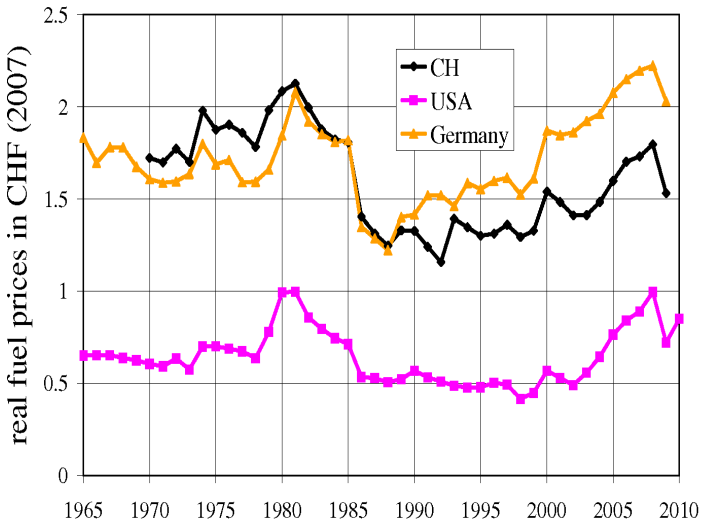


Figure 1.2.3: Real fuel price in Switzerland, Germany and the USA.[23]

Figure 1.2.3 shows that fuel prices in the USA are almost three times lower than in Germany and Switzerland. This difference in fuel price, the differences in the service level of public transport and spacial structure[24] may be a principal explanation for the huge differences in driving distance and the ratio of carless households between the USA and the other two countries. In contrast, there is no visible impact of the short-term fluctuations of fuel price on driving demand per capita, with the exception of the sharp increase in fuel prices in the period from 2004 to 2008. It seems that driving demand and car ownership are the result of a medium- or long-term decision, independent of minor changes in fuel price. In the period from 2004 to 2008, however, the increase in fuel prices was dramatic, even leading to a decrease in driving demand in the USA. The reason why the aggregate driving distance per capita declined after 2005 in the USA but not in Germany and Switzerland may be that fuel prices there have increased more considerably than in Germany and Switzerland since 2002, namely by approximately 100% in the USA compared to only 30% in the other two countries. Additionally, the impact was greater because the fuel efficiency of the average US car is much lower than in Germany and Switzerland, namely 11.7 l/100 km versus 8.3 l/100 km,[25] and therefore the rise in fuel price had a more significant impact on the marginal costs of driving. Finally, the consumption

[23] This data was collected by Bundesamt für Statistik (2009a), Allgemeiner Deutscher Automobil-Club (2010) and Energy Information Administration (2010). To convert the currencies, I took the average exchange rate from 2007 based on Schweizerische Nationalbank (SNB) (2010). Data for US fuel prices are already given in real values in the original dataset. To deflate the German prices, I used data from Statistisches Bundesamt (2010). For Switzerland I used data from Bundesamt für Statistik (2010b).

[24] Even though the level of urbanization is higher in the USA than in Switzerland and Germany (see Figure 1.2.5), the spatial structure of the USA leads to more demand in traffic. This is due to the huge suburban areas consisting primarily of detached houses. In this areas, the population density is rather low and the distances to facilities are long.

rate in the USA remained virtually unchanged,[26] meaning that this reduction cannot be the result of a general reduction in consumption.

Before discussing the medium- and long-term impact of fuel prices on driving demand, I shall present diagrams of the real income per capita, since this variable is also assumed to have a major impact on driving demand. The development of the income per capita may also explain the decrease in the ratio of carless households.

Next, I aim to find out why the growth rate of annual distances driven per capita decreased after 1990. Below, I will discuss possible reasons for this change for Switzerland and Germany. The first potential cause, besides fuel price, could be the GDP per capita. Also with respect to income per capita, no short-term impacts on driving demand and car ownership are visible.[27] As in the case of fuel prices, the reason for this could be that decisions on car ownership are medium- to long-term decisions that do not depend on minor- and short-term changes in income. Nonetheless, the trajectories of income may give us an indication of its medium- and long-term impact on car ownership and driving demand.

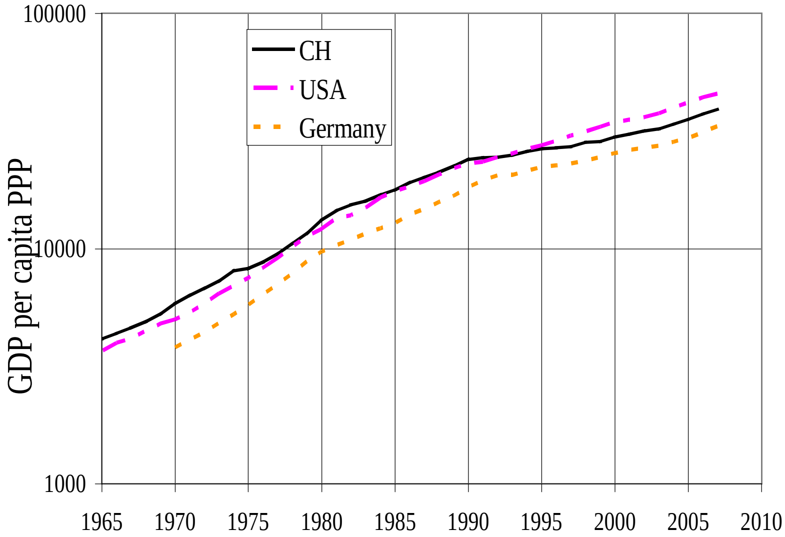


Figure 1.2.4: Gross domestic product, measured in purchasing power parity (PPP).[28]

---

[25] Note that fuel costs as a percentage of all marginal driving costs is approximately 50% in Switzerland, whereas it is less than 30% when fuel prices are around CHF 0.5/litre. The value of 8.3 l/100 km applies to Germany. It was computed by dividing the total fuel demand by the total distance driven by car on the aggregate level.

[26] According to the data contained in Heston (2009), the consumption rate remained virtually unchanged at a level of 70% in the period from 2004 to 2008, see data series of the variable “Consumption Share of Real GDP per capita (RGDPL)”.

[27] One reason for this is also that the short-term fluctuations of the gross domestic product (GDP) per capita were very small in the period considered here.

[28] This data is based on Heston et al. (2009).

Considering the trajectories of income per capita and fuel prices, it is not clear whether the declining growth of driving demand per capita from 1990 is due to the gradual increase in fuel prices in Switzerland and Germany or to the declining growth rate of their GDP. It is also interesting to note that the difference between fuel prices in Germany and Switzerland due to a gradual increase in fuel taxes in Germany did not affect the different driving demand per capita in these two countries. In fact, the very opposite was the case: despite the fact that the growth rate of GDP in both countries was virtually identical, the difference in driving distance per capita between Germany and Switzerland has increased slightly since 1992.

Country	Year	GDP	Avg. growth rate	Total growth rate
GER	1992	20'492		
	2007	33'181	3.27%	61.92%
CH	1992	24'510		
	2007	39'161	3.17%	59.78%
USA	1992	24'472		
	2007	45'597	4.24%	86.32%

Table 1.2.1: Growth in gross domestic product between 1992 and 2007.[29]

This indicates that factors other than fuel prices and GDP per capita may play an important role in the decline in the growth rate of driving demand from 1992 and the slight increase in the difference between the values for Germany and Switzerland. One explanation for this slight increase could be the difference in the development of spatial structure between Germany and Switzerland. Since households in rural areas have a higher preference for using cars, trends towards living in rural areas could boost driving demand. One way to gauge the ratio of households living in rural areas is to explore the rate of urbanization.[30] The hypothesis I therefore put forward is that the differences in the growth rate of driving demand per capita between the two countries are driven by changes in the differences of the rate of urbanization.[31]

[29] These are the results of my own computations, based on the data contained in Heston et al. (2009).

[30] The rate of urbanization measures the ratio of the population living in urban areas.

[31] Note also that the increase in the rate of urbanization in the USA did not seem to reduce the growth in driving demand.

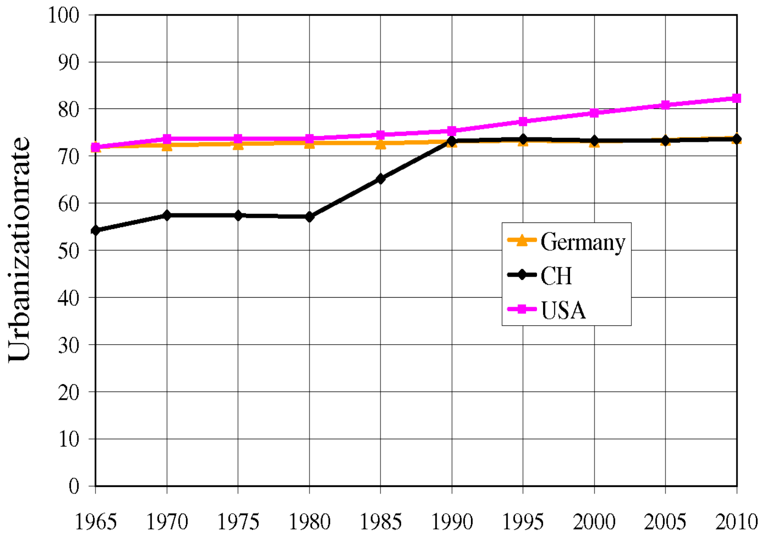


Figure 1.2.5: Ratio of people living in urban areas.[32]

Figure 1.2.5 shows that the difference in the rate of urbanization between Switzerland and Germany remained unchanged in the period from 1990 to 2010. It can further be seen that both rates of urbanization virtually remained at the same level, 73%. There must therefore be another reason for the increase in the difference in driving demand between Switzerland and Germany from 1990.

I now wish to test a final explanation for the difference in the growth in driving demand per capita between Switzerland and Germany. I will test whether the preference for car driving evolved differently due to changes in price level or the service level of the public transportation system. To test this, I present the trajectories of passenger kilometers generated by the public transportation system as a percentage of the total of passenger kilometers generated by all modes.

The following Figure 1.2.6 confirms that the preference for the public transportation system must have increased more considerably in Switzerland than in Germany. Below, I will present various diagrams to identify the potential reasons for this change. The first reason could be that in 1985, the Swiss Federal Railways (SBB) cut the price of the half-fare travelcard[33] from CHF 350 to CHF 100. This strategy succeeded in encouraging many new customers to use the railways. The effect can be seen in Figure 1.2.6, which shows a leap in the ratio of passenger kilometers in the public transport sector in Switzerland in 1985/86.

---

[32] This data is based on Heston et al. (2009).

[33] Passengers with a half-fare travelcard are given a 50% discount on all train tickets and on most other means of public transport. The discount is lower only for public transport within cities. The price of CHF 100 in 1986 relates to a one-year travelcard.

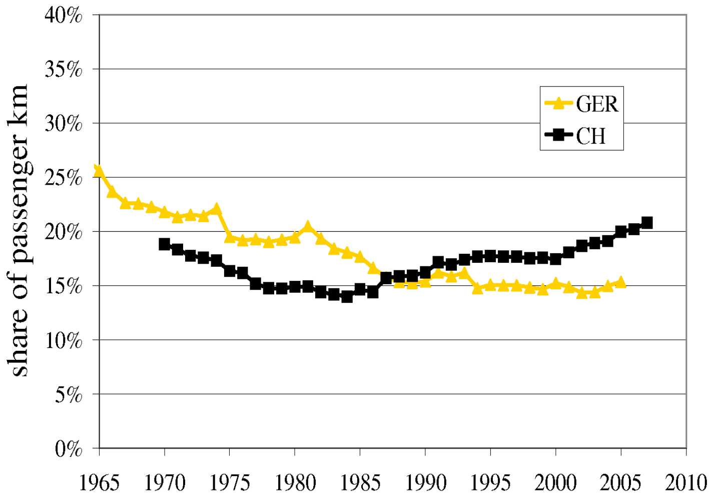


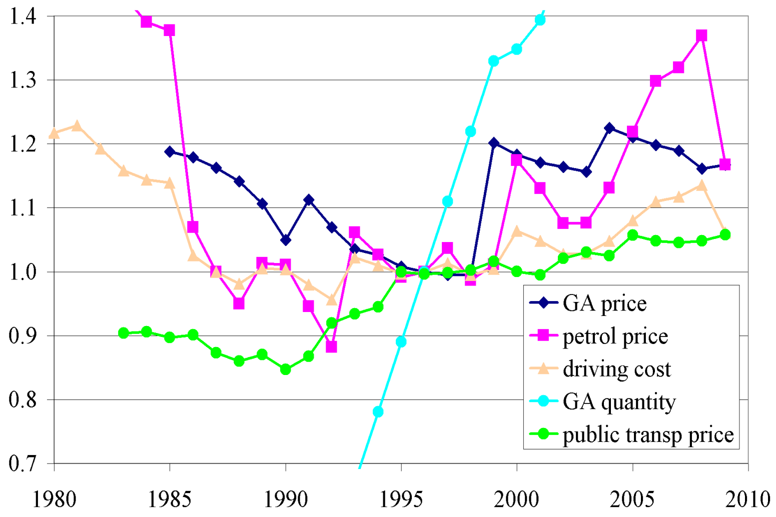
Figure 1.2.6: Passenger kilometres as a percentage of the total road and rail distances covered using public transport.[34]

One potential factor that could lead to a rise in the ratio of public transport is therefore a drop in the level of the price index of public transport. However, Figure 1.2.7 shows that this was not the case: after a short period of decline from 1985 to 1990, the real price index of public transport rose again whilst the real price of fuel decreased, at least up to 1998.

It was therefore probably the change in the service level of public transport that led to the increase of the share of public transportation: from 1990, the service level has been greatly improved by upgrading some of the major railway lines and by opening new lines between major cities. The public transportation system therefore became more attractive. An indicator of this increase in attractiveness is the dramatic increase in sales of the flat rate "General Abonnement" (GA) travelcard, which more than doubled between 1996 and 2009, namely from 165,500 to 400,000.[35] Note that the GA travelcard allows use of all public transportation services for a fixed annual cost. Even though the price of this travelcard has increased more significantly than fuel prices and driving costs since 1995, the number of tickets sold has increased dramatically. This is a sign that the service level of the public transportation system must have improved greatly.

[34] The data for passenger kilometres in Switzerland was collected by Bundesamt für Statistik (2010c) for public transport and Bundesamt für Statistik (2008) for individual transport. The data for Germany was collected by Bundesministerium für Verkehr, Bau und Stadtentwicklung (2006).

[35] This data was provided by Swiss Federal Railways (SBB) for exclusive use in this thesis.



Note: All values are normalized to one (1996)

Figure 1.2.7: Real prices of public transport and demand for flat rate tickets (GA).[36]

The indicators shown in the following Figure 1.2.8 confirm this. Figure 1.2.8 reflects that the service of the railway network was improved by increasing both train frequency and train speed. Moreover, travel comfort was enhanced by running new trains and constructing new railway lines with fewer bends and switches than the old lines. It follows that the decrease in demand for car driving, and particularly the stabilisation of the ratio of carless households at around 20%, is also due to a better service level of the public transportation system.

[36] The "GA price" denotes the annual price of the flat rate ticket, the "General Abonnement (GA)" (English: "GA travelcard") which allows passengers full use of public transport. The "GA quantity" represents the number of GA travelcards sold in any specific year; the "driving costs" are the costs of driving one additional kilometre by car. The latter figure is based on the computation presented in Subchapter 3.2, section entitled "Data". These costs consist of a component that is not reflected in the fuel costs, such as mechanical wear. I assume that the first cost component evolves in the same way as the Swiss Consumer Price Index (LIK) ("*Landesindex der Konsumentenpreise*") and that the fuel price enters the marginal cost by a constant factor, which means that I assume that the fuel efficiency did not change within the period considered. The price index of the GA travelcard is based on the price of an adult GA travelcard for 2nd class travel. In Figure 1.2.7, all values are deflated by the price index (LIK). The data sources are as follows: The "GA price" is based on data found in Verband öffentlicher Verkehr (2007), the "GA quantity" is based on Verband öffentlicher Verkehr (2007) for data up to 2002; data after 2002 is based on Schweizerische Bundesbahnen (2004-2009). The fuel prices were collected by Bundesamt für Statistik (2009a) and the price index by Bundesamt für Statistik (2010b). The data on the flat rate GA travelcard was supplied by the Swiss Federal Railways (SBB) for exclusive use in this thesis. The data on the price index of entire public transportation is from Bundesamt für Statistik (2010b).

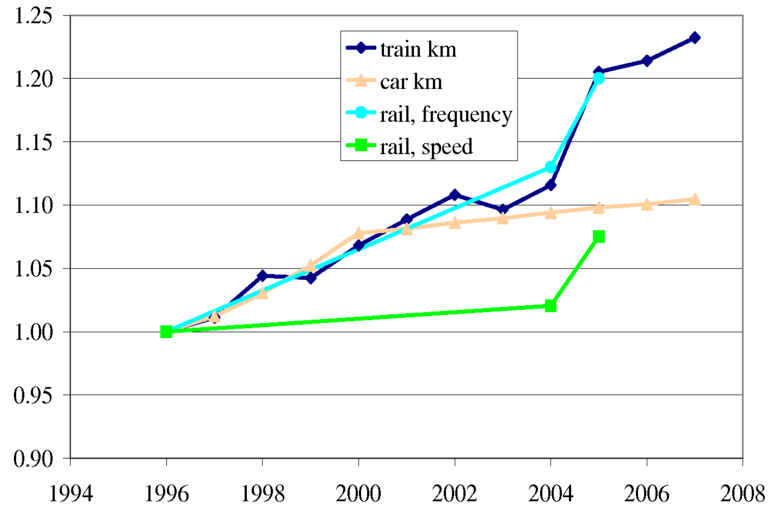


Figure 1.2.8: Increase in frequency and speed of Swiss Federal Railways trains.[37]

The reason for the difference in the growth rate of driving demand from 1990 therefore seems to be that the increase in the service level of public transport was more important in Switzerland than in Germany. The fact that fuel prices in Germany increased more considerably than in Switzerland from 1990 supports this hypothesis. Hence it follows that the drop in the growth rate of driving demand cannot be explained solely by the increase in fuel prices or by a possible satiation mechanism and the increase in fuel prices.[38]

To conclude this section, I shall summarize the results derived from the data and state which variables should be included in a model. In addition to presenting past changes and differences in travel demand between three countries, I also showed which factors influenced travel demand. The data also shows that the level of car ownership and the aggregate driving demand exhibit a co-movement over time. Since there is also a causal effect between car ownership and driving demand, my model should capture both of these values.[39] Since the variables income per capita, fuel price and service level of the transportation system presented in the diagrams seem to influence both of these values, my model

[37] The label "train km" denotes the total distance travelled by train, "car km" the sum of all distances travelled by car. "Rail, frequency" denotes the average frequency of train connections for a distance exceeding 30 km; "rail, speed" denotes the average speed of trips for a distance exceeding 30 km. To compute the last two values, the data on individual railway lines was aggregated using weights based on the total number of annual trips. The data sources used to produce this figure were: Bundesamt für Statistik (2008) for data on road traffic, Bundesamt für Statistik (2010c) for data on total train kilometres, and Bundesamt für Verkehr (2006) for data on train speed and frequency.

[38] Note that if the drop was caused solely by the increase in fuel prices, the drop in the growth rate of driving demand should have been greater in Switzerland than in Germany. This was obviously not the case.

[39] It is also intuitive that the level of car ownership strongly influences the aggregate driving demand. My model should therefore contain both these values.

should at least capture these as explanatory variables. Furthermore, the variables price level of the public transportation system, the rate of urbanization and the fixed costs of cars should be covered by the model to test whether they have a significant effect on the level of car ownership and the aggregate driving demand. If more detailed data at the household level is used, additional dependent variables can of course be considered. As expected, the figures show that the lower the fuel price, the higher the ratio of households that own a car and the higher driving demand is. The same holds for a higher income and a lower service level of the public transportation system. This implies that a model structure imposing that all explanatory variables influence the variables car ownership and aggregate driving demand either in the positive or in the negative direction, may be feasible.[40] The option of choosing a model structure with such a restriction could be useful for the models used in Chapters 2 and 3.

1.3 Results of other studies

There is a large number of studies that use different methods to estimate the elasticities of fuel demand. I will present only the results of a few major meta-studies that attempt to aggregate the results of the numerous studies found in the literature. I will also present the results of a German study, since I assume that the mechanisms in traffic economy in Switzerland are very similar to those in Germany. Finally, I will present the results of four Swiss studies.

International studies

I will consider three major meta-studies that aggregate past studies on elasticities of traffic demand. All of these studies aggregate the results for the elasticities of fuel demand, aggregate vehicle kilometres, average fuel efficiency and the car fleet with respect to fuel prices and income. In the following, I shall focus on the two most recent of these three studies, namely the studies by Goodwin et al. (2004) and Brons et al. (2006).

I shall start by presenting the study by Goodwin et al. (2004). Goodwin et al. added 69 studies to the set of studies already used by Goodwin (1992) for their meta-study. They only added results that were not based on repeat publications in a different form or publications based on updates of the same base material. The countries addressed in the studies included the USA (number of studies = 63), the UK (29), Canada (12), France (7), Germany (7), Belgium (6), 12 OECD countries (6) as well as Denmark, Italy, the Netherlands, Austria, Sweden, Norway, Spain, Australia, Japan, specific US states and various multi-country groupings (1-4 each). These studies contain a total of 175 equations and produced 491 elasticities.

[40] Note that, in general, a model structure should not restrict the sign of the impact of explanatory variables.

To analyse these studies, the results were divided into two groups. The first group contains results based on dynamic estimation routines, namely time series analysis. The second group comprises results based on static econometric methods, such as OLS, Tobit, etc. The values they present are simply summary statistics – mean values, standard deviations and populations size – where the population size reflects the number of results used to compute the mean and standard deviation. The results are listed in Table 1.3.1, which also includes the results given in Graham and Glaister (2005). These results are very similar to those presented in Goodwin et al. (2004).

Study	Type	Fuel price			Income		
		sr[c]	lr[c]	static[c]	sr	lr	static
Goodwin et al. (2004)[21]	Fuel demand	-0.25 (0.15)[a] {46}[b]	-0.64 (0.44) {51}	-0.43 (0.23) {24}	0.39 (0.25) {45}	1.08 (0.35) {50}	0.49 (0.40) {20}
Goodwin et al. (2004)	Fuel efficiency	-0.15 (--) {--}	-0.35 (--) {--}	-0.30 (0.22) {22}	0.09 (--) {--}	0.35 (--) {--}	0.55 (0.35) {19}
Goodwin et al. (2004)	Vehicle-km (total)	-0.10 (0.06) {3}	-0.29 (0.29) {3}	-0.51 (0.25) {2}	0.30 (0.21) {7}	0.73 (0.48) {7}	0.49 (0.42) {15}
Goodwin et al. (2004)	Vehicle stock	-0.08 (0.06) {8}	-0.25 (0.17) {8}	-0.06 (0.08) {3}	0.32 (0.21) {15}	0.81 (0.43) {15}	1.09 (0.56) {5}
Graham and Glaister (2005)[22]	Fuel demand	-0.25 {377}	-0.77 {213}	--	0.47 {333}	0.93 {150}	--
Graham and Glaister (2005)[23]	Fuel efficiency	-0.10	-0.46	--	0.17	0.20	--
Graham and Glaister (2004)[24]	Vehicle-km (total)	-0.15 {}	-0.31 {}	--	0.30 {}	0.73 {}	--

[a]: The values in parentheses “( )” denote the standard deviation of the values of the different studies.
[b]: The values in parentheses “{ }” denote the number of results on which the values are based.
[c]: Short-run (sr) and long-run (lr) elasticities are based on results of dynamic estimation methods based on time series data. The results of static econometric methods, such as OLS and Tobit, correspond to the values in the column “static”.
Note: Since only one study referred to explaining the impact of fuel prices on fuel efficiency, I computed this value by subtracting the value of the elasticity of total vehicle-km from the elasticity of fuel demand. The same approach was used by Goodwin et al. (2004) to derive their summary of results on page 278.

Table 1.3.1: Elasticities of a meta-analysis based on computing averages of a set of results of surveys.

[21] See Goodwin et al. (2004: 282-285).

[22] See Graham and Glaister (2005: 26), Table 2.

[23] The values for elasticities with respect to fuel price were taken from Table 3 on page 26 in Graham and Glaister (2005). I computed the values for elasticities with respect to income by subtracting 0.30 from 0.47 and 0.73 from 0.93.

[24] Graham and Glaister (2005) refer on page 26, Table 2 to Graham and Glaister (2004: 271), Table 8.

Note that the elasticity of fuel demand was computed based on the following identity:

Fuel consumption (C) = fuel consumption per km (FE) · vehicle-km (VKM) (1.3.1)

The following relation is then yielded between elasticities $\varepsilon_{C,p_{fuel}}$, $\varepsilon_{FE,p_{fuel}}$ and $\varepsilon_{VKM,p_{fuel}}$:[25]

$$\varepsilon_{C,p_{fuel}} = \varepsilon_{FE,p_{fuel}} + \varepsilon_{VKM,p_{fuel}} . \qquad (1.3.2)$$

The reason why I determined elasticities ε_{FE} by computing $\varepsilon_{FE} = \varepsilon_C - \varepsilon_{VKM}$ is that only in one survey ε_{FE} is directly determined and, to make matters worse, the corresponding result does not appear to be very plausible.

The results above show that the elasticity of fuel efficiency with respect to fuel price makes up about one half of the elasticity of fuel demand. This is quite plausible in the case of long-term elasticities since households may switch to more fuel-efficient cars. Surprisingly, this contribution of the elasticity of fuel efficiency on short-run elasticity is even more than half in the case of Goodwin et al. (2004). This could be because households with more than one car may switch to a more fuel-efficient car or may drive more slowly and efficiently to save fuel. It seems that households prefer to choose this method of reducing fuel costs when fuel prices rise, rather than driving less. In contrast, the results of Graham and Glaister (2004) and Graham and Glaister (2005) imply that ε_{FE} contributes more significantly to fuel elasticity ε_C in long-term rather than in short-term elasticities. This is true not only for elasticity with respect to income, but also to fuel price.[26]

To gain more precise information about the extent to which the elasticity of fuel efficiency contributes to fuel demand, Brons et al. (2006) conducted a survey that considers the relationship between the elasticities (1.3.2). They defined a linear system of equations on which they based their estimation procedure.[27] They added a number of equations based on further splitting the elasticities. For instance, the product of kilometres per car and size of car stock is used to explain the total annual driving distance. The elasticities are computed using the Maximum Likelihood method. In their first step, Brons et al. (2006) did not distinguish between short- and long-term elasticities. The results are presented in Table 1.3.2.

---

[25] Relation (1.3.2) results from Equation (1.3.1) by taking the logs, $\ln(C) = \ln(FE) + \ln(VKM)$, differentiating $\frac{\partial C/\partial p}{C} = \frac{\partial FE/\partial p}{FE} + \frac{\partial VKM/\partial p}{VKM}$ and multiplying by *p*, $\frac{\partial C}{\partial p} \cdot \frac{p}{C} = \frac{\partial FE}{\partial p} \cdot \frac{p}{FE} + \frac{\partial VKM}{\partial p} \cdot \frac{p}{VKM}$.

[26] It seems that in the event of an increase in household income, a household will not immediately replace its car by a larger, less fuel-efficient one. Fuel demand therefore reacts more strongly to changes in income in the long run rather than in the short run.

[27] See Brons (2006: 8).

	Fuel demand	Fuel efficiency	Vehicle-km (total)	Vehicle stock
Brons et al. (2006)	-0.53 (0.003)[a] {158}[b]	-0.22 (0.02) {15}	-0.32 (0.13) {3}	-0.22 (0.15) {14}
Graham and Glaister (2002)	-0.698	-0.373	--	-0.312
Hanly et al. (2002)	-0.451	--	-0.257	-0.148
Espey (1998)	-0.442	--	--	--

[a]: The values in parentheses "( )" denote the standard deviation of the values of the different studies.
[b]: The values in parentheses "{ }" denote the number of results of different studies that actually directly explain the elasticity.
Note 1: The elasticities are "average" elasticities. No distinction was made between short- and long-run elasticities or between type of method.
Note 2: The results of the studies by Graham and Glaister (2002), Hanly et al. (2002) and Espey (1998) were computed by summing up the short- and long-run elasticities weighted by the corresponding number of results of different studies.

Table 1.3.2: Fuel price elasticities of a meta-analysis based on computing averages of a set of results of surveys.

The results show that the elasticity of fuel efficiency with respect to fuel price contributes about 40% to the elasticity of fuel demand. This is less than the proportions in Table 1.3.1; a proportion of 40% seems more realistic.

Further, Brons et al. (2006) studied whether elasticities could differ with respect to the estimation method, between short- and long-run elasticities and the type of country. This survey was conducted at a decomposed level.[28] The results split by the different types of elasticities are shown in Table 1.3.3.

	Fuel demand	Fuel efficiency	km per car	Vehicle stock
Constant	-0.107 (0.067)[a]	--	--	--
USA, Canada, Australia	0.148** (0.038)	ns[b]	ns	0.204** (0.069)
Cross section	-0.226** (0.0724)	ns	ns	-0.344* (0.173)
Long run	-0.366** (0.047)	-0.130 (0.047)	-0.327* (0.130)	0.090 (0.083)
Dynamic	-0.197** (0.041)	-0.295** (0.100)	ns	ns

[a]: The values in parentheses "( )" denote the standard deviation based on the Maximum Likelihood estimation method.
[b]: I only list statistically significant coefficients. "ns" stands for "not statistically significant".
Significance levels: * : 5% ** : 1%
Note 1: Some explanatory variables were not included in this table. The complete table can be found in Brons et al. (2006: 16).
Note 2: The constant relates to panel data time series models and their estimates of the short-run elasticities.
Note 3: relates to panel data time series models and their estimates of the short-run elasticities.

Table 1.3.3: Differences in the values of the elasticities due to the type of the model and type of country.

[28]This means that they ignored identities such as those stated in (1.3.2).

The results presented in Table 1.3.3 can be interpreted as follows: long-run models have elasticities that are on average -0.366 below the short-run elasticities. There are two reasons for this difference. First, households may adapt their mode of transport or even their location in the long run, which they tend not to do in the short run. Second, in the long run households may switch to a more fuel-efficient car when fuel prices increase, for instance when replacing their old car. In total, the long-run elasticities of fuel demand are approximately 0.366 higher in absolute terms than short-run elasticities. This value corresponds roughly to those gained by computing the differences between the short- and long-run elasticities in Table 1.3.1. In other words, the value found by Brons et al. (2006) is plausible.

Another interesting result given in Table 1.3.3 is that the elasticities with respect to fuel prices in the USA, Canada and Australia are smaller in absolute terms than those in other countries. This is not very surprising because the price of fuel is low in these three countries, due to low taxes on fuel. Fuel costs therefore contribute to a lesser extent to driving costs. To give an example: a one percent increase in fuel prices increases driving costs in these three countries by a lower percentage than in other countries.[29] This is because the costs of mechanical wear comprises over half of all the marginal costs of driving, given that the fuel price is around CHF 1.50/litre.[30] Another explanation is that the income level is high in these three countries, meaning that households there are not very sensitive to fuel price changes because such expenses make up only a small proportion of the household budget. In addition, the pubic transportation infrastructure in all three countries is comparatively poor, rendering it difficult for households to switch from private to public transportation.[31] The results further show that the computed elasticities depend on the estimation method applied. Price elasticities of fuel demand that resulted from cross section data and are processed by traditional static estimation methods, such as OLS and Tobit, are on average approximately a factor of 0.497 smaller in magnitude than the long-

[29] Note that taxes made up only approximately 13% of the retail fuel price in 2008 in the USA, whilst average taxes in Europe constituted around 66% of retail prices there in 2003. These figures are based on American Petroleum Institute (2010) or U.S. Department of Energy (2010) for the USA. For the European countries I carried out my own calculations based on Diagram 1 on page 66 of Bundesministerium der Finanzen (2005).

[30] Even in the case of a standard car in Switzerland, fuel expenses at a rather high fuel price of CHF 1.79/litre contribute only 45.7% to the marginal driving costs, see TCS (2007). Since fuel prices are much lower in the USA, namely about USD 2.8/gal, which corresponds to approximately CHF 0.74/litre versus CHF 1.69/litre in Switzerland in 2010, I conclude that fuel as a percentage of all marginal driving costs is much lower in the USA than in Switzerland, even though the fuel economy of cars in the USA is lower than in Switzerland. Note that in the USA, the average fuel economy of a passenger car is about 9.6 litres/100 km, see Research and Innovative Technology Administration, Bureau of Transportation Statistics (2010), while in Switzerland it is about 8 litre/100 km.

[31] The argumentation resembles that in Brons et al. (2006: 14): "One possible explanation for this is that the combination of high income and low gasoline prices renders consumers less price sensitive. Another explanation might be that, due to the combination of sparse population and relatively underdeveloped public transport infrastructure, car dependence is higher in these countries."

term elasticities computed using dynamic time series analysis methods, namely -0.333 versus -0.670.[32] Note that the long-term elasticities computed using dynamic time series analysis methods can be considered to be the most realistic elasticities because they also map lagged effects of changes in fuel prices, such as selling the car when fuel prices increase.

Graham and Glaister (2004) presented a table of results of studies conducted to explain the effects of income, fuel prices and fixed costs on car ownership.[33] These studies were based on the model introduced by De Jong (1990). Since in these studies the elasticities of car ownership with respect to income and fixed costs are more than four times higher than the values found in international studies, I consider the results yielding from the model of De Jong (1990) as not trustworthy.[34]

I consider the works by Johansson and Schipper (1997) and Dargay (2001) to be the most relevant studies on car ownership and aggregate car stock. The results provided by Johansson and Schipper (1997) are based on the restriction that total fuel consumption per capita can be explained by the product of driving distance per capita. Johansson and Shipper applied several statistical methods on time-series data of 12 OECD countries[35] for the years 1973 to 1992. The authors computed the effect of fuel price, taxation of cars and population density on car stock, mean fuel efficiency, mean driving distance of cars, fuel demand and households' total car travel demand. In contrast, Dargay (2001) explains only car ownership. His study is based on cohort data concerning the UK from 1970 to 1995.[36] He computed his results using an econometric normal fixed effects model. Table 1.3.4 shows the results of these two studies. For completeness, I have also listed the results of Johansson and Schipper (1997) for the other elasticities.

---

[32] This factor results from the following calculation:
$(-0.107-0.226)/(-0.107-0.366-0.197)=(-0.333)/(-0.670)=0.497$, for values, see Table 1.3.3.

[33] See Graham and Glaister (2004: 264).

[34] In Subchapter 1.5, I will explain why the accuracy of the results produced by De Jong's model seems to be doubtful.

[35] These countries were the USA, the UK, Japan, Australia, Germany, France, Italy, the Netherlands, Sweden, Denmark, Norway and Finland. See Johansson and Shipper (1997: 278).

[36] "... based on cohort data constructed from the UK Family Expenditure Survey for the years 1970 to 1995. This survey has been carried out continuously on an annual basis since 1960, each survey providing a random sample of around 7,000 households, ..." Dargay (2002: 811).

Study	Type	Fuel price	Income	Fixed costs
Johansson and Schipper (1997)	Car fuel demand	-1.0 to -0.40 [-0.7][a]	0.05 to 1.6 [1.2]	-0.254 to -0.032 [-0.175]
Johansson and Schipper (1997)	Car travel demand	-0.55 to -0.05 [-0.3]	0.65 to 1.25 [1.2]	-0.0635 to 0.127 [0.0]
Johansson and Schipper (1997)	Mean fuel intensity (fuel efficiency)	-0.45 to -0.35 [-0.4]	-0.6 to 0.0 [0.0]	-0.190 to -0.159 [-0.175]
Johansson and Schipper (1997)	Car stock	-0.20 to 0.0 [-0.1]	0.75 to 1.25 [1.0]	-0.127 to -0.063 [-0.095]
Dargay (2001)	Car stock	--	0.48 (sr)[b] 0.74 (lr)	0.16 (sr) 0.26 (lr)

[a]: The values in parentheses "[..]" denote the "best guess" values according to Johansson and Schipper (1997), 290.
[b]: The values provided by Johansson and Schipper (1997) are long-run (lr) elasticities. Dargay (2001) computed both long- and short-run (sr) elasticities.
Note: The original values of the elasticities with respect to fixed costs related to the sum of amortisation and taxes in the case of the tables in Johansson and Schipper (1997) and to car purchase costs, and therefore amortisation, in the case of Dargay (2001). Since, according to Touring Club der Schweiz (2007), these costs constitute only 63% and 50%, respectively, of the fixed cost, I have corrected the values by factors of 0.63^{-1} and 0.50^{-1}, respectively.

Table 1.3.4: Long-run elasticities of the car stock with respect to fuel prices, income and fixed costs.

So far the elasticities that were presented were average values found by different models using data from different countries. It is quite intuitive to assume that elasticities could systematically differ between certain types of countries. A common presumption is that the higher the income level, the lower all these elasticities – as listed in Table 1.3.4 – are expected to be in magnitude, due to the satiation effect.[37] In the study by Dargay et al. (2007), this satiation effect is used as a premise. Dargay et al. (2007) also present some empirical evidence that this satiation effect occurs with respect to income:[38] in the period from 1960 to 2002, the ratio of growth of vehicle per capita versus growth of income per capita of low-income non-OECD countries was almost twice that of high-income OECD countries, namely 1.30 versus 2.39.[39] Furthermore, Dargay et al. (2007) projected elasticities of car ownership with respect to income of 0.42 for OECD countries and 1.61 for non-OECD countries using a non-linear model containing a number of control variables.[40] Figures 1.2.1 and 1.2.2 in the previous subchapter support this finding: since 1990, the growth rates of driving distance have decreased more

[37] This satiation effect is based on the assumption that the marginal demand of car driving approaches zero at a certain income level because households simply no longer have time to drive. At this point, the impact of fuel price on driving demand is also lower than at lower levels of income.

[38] Dargay et al. (2007: 163-164), Figures 10 and 11. Figures 2.3 and 2.4 on page 6 show even more clearly that the elasticities are lower for high-income countries such as the USA, Canada, Australia and all Western European countries.

[39] See Dargay et al. (2007: 147), Table 1.

[40] See Table 4 on page 266 of Dargay et al. (2007). This table also shows that the results of other studies are in a similar range.

considerably than the GDP growth rate decreased in the USA, Germany and Switzerland. In addition, the fact that the ratio of carless households in Switzerland and Germany has not decreased further in recent years also lends support to the hypothesis that the economies are approaching a satiation level of driving demand and car ownership. Since Switzerland is one of the countries with the highest income per capita, I expect that the income elasticity of 0.42 for the car stock for the OECD countries for 2030 marks an upper limit for the case of Switzerland. I also expect that the elasticities of car stock and fuel efficiency are lower in Switzerland. The reason for this assumption is that, since income levels are high and driving distances per capita are much lower than in the USA, for instance, the proportion of expenditures for car driving is low.[41] I therefore assume that households will not readily sell or replace their cars when fuel prices rise. In contrast, I expect the elasticity of driving distance with respect to fuel prices to be similar to the average of other countries, since the previously mentioned effect of a high income level is balanced out by a high-standard network of public transport. The latter implies that most individuals can easily use public transport for at least some of the journeys they used to undertake by car.

Swiss studies

There are only very few studies of the Swiss transportation sector, none of which have been published in a scientific journal. Nonetheless, I wish to summarize the results of a number of studies on travel or fuel demand in the Swiss transport sector. I will also present the results of a German study because I expect the mechanism in the German transport sector to be very similar to the one in Switzerland, as can be concluded from the figures in the previous Subchapter 1.2. Table 1.3.5 summarizes the results of these studies.

---

[41] This argumentation is used by Brons et al. (2006) to explain the smaller fuel elasticities of the USA, Canada and Australia by Brons et al. (2006). They argue "... price sensitivity is lower in the US, Canada and Australia. Espey (1998) and Hanly et al. (2002) have found similar results. One possible explanation for this is that the combination of high income and low gasoline prices renders consumers less price sensitive, ..." see Brons et al. (2006: 14).

		Fuel price elasticities		Income elasticities	
Author	Data	Short run	Long run	Short run	Long run
Schleiniger (1995)	Time-series 1967-1994	-0.24 (0.01)[a]	--	1.24 (0.258)	1.65 (0.077)
Baranzini et al. (2009)[42]	Time-series 1970-2008 fuel demand	**-0.097 (0.055)**	**-0.202 (0.029)**	**0.453 (0.188)**	**0.829 (0.138)**
Baranzini et al. (2009)	Time-series 1970-2008 only petrol	**-0.084 (0.061)**	**-0.333 (0.035)**	**0.033 (0.193)**	**0.627 (0.172)**
Axhausen and Erath (2010)[43]	Stated preference, panel		-0.15 (--)	--	--
Peter et al. (2002)	Time-series	-0.3 .. -0.4		0.65	
Progtrans (2004)	Time-series 1970-2002 (Germany)	-0.18 (0.05)	--	0.59 (0.20)	--
Progtrans (2004)	Rolling panel, 1995-2005 driving demand (Germany)	0.3 .. 0.5 (0.15)		--	--

[a] The values brackets "(..)" denote standard deviations of the corresponding values.

Note: The results generated by Baranzini et al. (2009) are given in bold because I consider them to be the most reliable results.

Table 1.3.5: Results of Swiss studies on fuel demand.

Schleiniger (1995) and Baranzini et al. (2009) apply time-series error correction models using the Engle-Granger methodology. Baranzini et al. added a structural break in 1993 accounting for the tax level increase in 1993 in the error correction term. Integration tests were conducted in both studies. The income elasticity of fuel demand in Schleiniger (1995) seems to be rather high. One reason for this could be that Baranzini et al. added a trend variable into the equation of the long-term equation as well as the car stock.[44] It seems problematic to add the car stock in the long-run equation as an explanatory variable as both the car stock and the demand for driving are driven by income, the state of the road infrastructure and other factors that influence preference for car ownership and use. I therefore assume that an endogeneity problem could lead to biased estimators. The results in Axhausen and Erath (2010) are based on a stated preference dataset. I consider their value of -0.19 to be a lower bound in absolute value, since I suspect that the respondents in a stated preference framework do not seriously evaluate solutions to alter their driving habits when fuel prices increase. This is because they

[42] These figures correspond to the results based on annual data, see Tables 6.1 and 6.2 on pages 38 and 39.

[43] See Axhausen and Erath (2010: 57).

[44] See Equation (1) in Baranzini et al. (2009: 5).

do not have to pay for the actual increase in fuel cost that arise when fuel fictively rise. Since the authors gave me access to the dataset, I will present the results of different models based on this dataset in the following chapters, along with my comments. The values of the study by Peter et al. (2002) are computed by use of time-series models. These models were based on growth rates between subsequent quarterly periods. Their results might therefore be determined by the intra-year co-movement between economic activities and the GDP rather than by the short- or long-term effects of non-seasonal fluctuations. Therefore, the results from Peter et al. (2002) are likely to be biased. Finally, I list the results presented in Progtrans (2002) that refer to German data. The first row lists the results generated by Progtrans based on aggregate time series data. They regressed the growth rates of fuel prices and GDP on fuel demand. They also processed long-term elasticities. Since they did not test for co-integration, however, I have not listed the results. Progtrans also processed results based on household consumer survey data with a rolling panel structure. All households responded for three subsequent years. The elasticities were computed using a random effects estimation model. Since the variable income was not available to them, these results may suffer from an omitted variable bias.

Unfortunately, all of these studies suffer from severe imperfections, most of which are related to mis-specification. I was therefore unable to include values for the elasticities I regarded as benchmarks in my study. In view of the preceding discussion, I conclude that the most reliable results are those presented by Baranzini et al. (2009), also because they control for fuel tourism by incorporating a fuel price index for neighbouring countries.

1.4 Swiss data

In this subchapter, I shall present the Swiss data used in my empirical analysis, all of which are at the household level. I will start by presenting the micro-census survey data on travel behaviour, conducted in 2000 and 2005, see Bundesamt für Statistik (2001) and Bundesamt für Statistik (2006a). I will then present the data generated by the Swiss consumer survey conducted between 2000 and 2005, see Bundesamt für Statistik (2007b). In addition, I shall describe how the stated preference panel data on car ownership and use was collected by Axhausen and Erath (2010). Finally, I will present the stated preference panel dataset on car ownership compiled by Wüstenhagen and Sammer (2007).

Micro-census data on travel behaviour

Every five years, an extensive survey on travel behaviour is conducted at the household level in Switzerland. I will consider only data from 2000 to 2005, see Bundesamt für Statistik (2001) and Bundesamt für Statistik (2006a). This data was compiled by a telephone survey. In total, 30,000 (2000) and 33,000 (2005) Swiss residential households were interviewed. The interview dates were more or less evenly distributed over the year in question. This dataset contains detailed information on travel behaviour, ownership of cars, motorbikes and bicycles, and information about each household. The survey covers all aspects of travel, namely the use of cars, public transport and air travel. Data on total annual distances travelled is available only for each household's car. Data on the use of public transport and air travel is available only for one person on a certain day. For that day, the modes of transport and the distances of each leg of the journey were collected. The same information is available for the main legs of the last journey involving at least one overnight stay. Since the purpose of this study is to investigate fuel demand, I will use the information provided on total kilometres driven by cars. Since the focus is on the effect of fuel prices on distances travelled and whether or not to own one or more cars, I will use only the household variables that appeared to have the greatest impact on travel distance or fuel demand as control variables.[45] Since the models I will use can capture only one type of car, this type is considered to be an “average car”. The fixed costs of maintaining a car and the marginal costs of driving are assumed to be equal to those of an average car owned by a Swiss household. The fixed and marginal costs of such a car were published by the Swiss touring club, see TCS (2007). According to this publication, the fixed annual costs incurred in 2000 and 2005 were CHF 7,935 and CHF 7,033, respectively. Taking the data from TCS (2007) and TCS (2001), I derive formulas determining the marginal costs per kilometre p_2 :[46]

$$p_2 = 0.163+0.088 \cdot p_{fuel}\text{, (2000)} \quad (1.4.1a)$$

$$p_2 = 0.1601+0.0778 \cdot p_{fuel}\text{, (2005)} \quad (1.4.1b)$$

[45] This choice is based on comparing the results of different specifications based on a Tobit model.

[46] According to TCS (2007), the total annual costs of an average car amounted to CHF 11,600 when the annual distance driven was 15,000 kilometres (km). 17.4% of these costs, namely CHF 2,018.4, were fuel costs. Based on the average fuel price paid for petrol 98 octane of CHF 1.729/litre in 2007 (Bundesamt für Statistik 2009), it can be computed that the TCS (2007) based this fuel cost on a fuel consumption of 7.7825 litres/100 km: (CHF 2,018.4/15,000 km) / (CHF 1.729/litre) = 7.7825 litres/100 km. The fuel costs of an average car per kilometre is therefore 7.7825 litres/100 km/100 multiplied by the fuel price per litre paid by households. Non-fuel-related marginal costs of a car were accounted as 20.7% of the total costs, 0.207 · CHF 11,600 = CHF 3,312, amounting to CHF 3,312/15,000 km = CHF 0.1601/km, see TCS (2007). For the dataset of the 2000 survey, all these costs were computed based on the data of TCS (2001).

Fuel price p_{fuel} is the average fuel price from the last twelve months prior to interviewing the household. I assume that households react to the actual fuel price in their decisions on how far to drive and that, therefore, the average fuel price will influence the total distance driven during this period.[47] I also assume that households base their decisions regarding whether or not to own a car on this price.[48] Note that the constant component of the formula accounts for marginal costs unrelated to fuel consumption, such as the wear and tear of tyres and mechanical components. The computation of p_{fuel} is based on the monthly average price of petrol 98 octane, as published by the Bundesamt für Statistik (2009a). Due to the variation in fuel price in 2004 and 2005, the average fuel price of 12 months ranges from CHF 1.404 to CHF 1.574, and the marginal cost p_2 from CHF 0.2694/km to CHF 0.2826/km. In relative terms, the minimal value of p_2 is 4.7% lower than the highest value.[49] In other words, the variation is very small. This is a key feature and key disadvantage of cross-sectional datasets on travel behaviour: The fuel price is roughly the same for every household at all locations within a country. Due to the lack of variation in the observed fuel prices, the coefficients estimated by traditional econometric models accounting for the effect of the fuel price are not statistically significant.

Before carrying out my estimation, I eliminated all households that drive more than 60,000 kilometres a year or that spend more than one third of their income on the costs of driving. There are two reasons for eliminating these observations. First, I consider these observations to be wrong or inaccurate information provided by the households, so-called "outliers". Second, I assume these households use their cars professionally. I have excluded these households from the dataset because I am interested in how households react to fuel price changes concerning private travel requirements.

I also eliminated households that drive less than 1,500 kilometres per year. I consider it to be irrational to own a car and bear the fixed costs of at least CHF 5,000 per year if only less than 1,500 km are driven per year. I therefore assume that these households provided incorrect information on their driving distances. Further, I eliminated data on households that failed to state their income. Note that relatively more households that possess a car were eliminated than carless households. To keep the proportion of carless households identical to that of the original dataset, I eliminated

[47] Note that the driving distance refers also to this period.

[48] Of course, it is reasonable to assume that households base their decisions on whether or not to own a car on the expected fuel price throughout the car's life instead. However, this price is not observable. The average fuel price of the last twelve months seems to be a good approximation, since there is no reason to believe that fuel prices increase more or less rapidly than the consumer price index, which is more or less identical to price p_1; thus the ratio p_2/p_1 will remain stable over time.

[49] The variation in the 2000 dataset is also very low.

randomly drawn data of carless households from the dataset. Since households' incomes are available only as an interval value, I computed the actual annual income by allocating the mean of the corresponding intervals.[50] The summary statistics of the resulting dataset is reported in the following table:

	Driving distance	Ratio of carless households	Driving costs p_2	Annual income	Income, urban households	Income, rural households	Ratio of rural households
mean	13,890	18.90%	0.2745	80,187	81,555	75,561	22.8%
stdev	12,195	-	0.0036	43,373	44,340	39,589	-
min	0	-	0.2692	18,000	18,000	18,000	-
max	59,731	-	0.2838	228,000	228,000	228,000	-

Note: The annual income and all prices are measured in Swiss Francs (CHF). The figures from the 2000 dataset are similar.

Table 1.4.1: Summary statistics of data used (2005).

Swiss consumer survey

The Swiss consumer survey "Einkommens- und Verbrauchserhebung (EVE)", Bundesamt für Statistik (2007) is based on a written report provided by households. Each household reports all its expenditures in a very detailed diary.[51] Finally, the households' expenditures are grouped into 458 categories, such as "bananas", "apples", "pears". For each household, the total number of units consumed and the total expenditures in each category is available. The households also reported all sources of income (63 categories) and ownership of durable goods (27 categories), such as possessing a car. Additionally, the dataset contains a number of household-specific variables, such as number of children and type of residence. Although this dataset does not contain any information on the number of kilometres households drove by car, it does provide information about diesel and petrol consumption. To aggregate these two measures, I multiplied the amount of diesel consumed by a factor of 1.1228 to obtain a value that accounts for carbon dioxide in petrol litre equivalents.[52] Since the fuel efficiency of diesel cars is better on average than that of a petrol-driven cars, this value – up to

---

[50] The first interval was "less than CHF 2,000 per month". For this category, I allocated 12 · CHF 1,500 = CHF 18,000. Category 2-8 were CHF 2,000 - 3,999, CHF 4,000 - 5,999, etc. For these categories, I allocated 12 · CHF 3,000, 12 · CHF 5,000, 12 · CHF 7,000, ..., 12 · CHF 15,000. For the category "more than CHF 16,000 per month, I allocated 12 · CHF 19,000.

[51] The households list the quantities and prices of all products they purchase. The diary of the actual survey in 2010 can be downloaded: "Erhebungen, Quellen – Haushaltsbudgeterhebung (HABE)" http://www.bfs.admin.ch/bfs/portal/de/index/infothek/erhebungen__quellen/blank/blank/habe/Resultate.html. The diary of the surveys from 2002 to 2005 can be ordered from Bundesamt für Statistik der Schweiz.

a certain factor – is also a proxy for driving distance. To achieve comparability with the results I computed using the micro-census datasets described above, I multiplied the carbon dioxide equivalents by a factor, such that the same average driving distance is yielded as in these datasets.[53] For my data, I eliminated observations according to the same rules as those applied in the case of micro-census data on travel behaviour.[54] The Swiss consumer survey is conducted every year. A random sample of 3,500 households submits a report every year.[55] For my survey, I used data compiled for 2002 to 2005.[56]

Stated preference dataset by Axhausen and Erath (2010)

This dataset is based on interviews conducted with 409 households. These households were asked about their actual behaviour concerning private transport. All households owned at least one car at the time of the interview (June and July, 2009) and were chosen representatively.[57] The data was collected by thirteen interviewers. Households reported about car ownership and the kilometres they drive. They were then asked how they would react to an increase in fuel prices, changes to the price of public transport and to tax rebates for certain car types. They could choose from a set of nine car types. Information on fixed and marginal driving costs was listed for each car type. The marginal costs were split into fuel-related costs and non-fuel related costs. In addition, they were asked to choose from a set of engine sizes and engine types (gasoline, diesel, natural gas, hybrid and electric). Information on energy costs for a fuel-based engine were given for each engine type.[58] The questioning was organised as follows: First, the levels of fuel price (1.5, 2, 3, 4, 5 CHF/litre), the price level of public transport[59] and the tax rebates for certain car types relevant to the following choice situation were announced to

[52] Note that the carbon dioxide emission of burning one litre of petrol is 2.36 kilogrammes;for diesel it is 2.65 kilograms, i.e. 12.28% more.

[53] The average driving distances given in the two micro-census datasets were 14,128 kilometres (Mz00) and 13,890 kilometres (Mz05). I chose the factor such that the average driving distance of the EVE dataset is 13,853 kilometres. All values correspond to the datasets from which outliers were eliminated. Note that I ignore the fact that part of the fuel could have been bought for use in motorbikes and gardening equipment, etc.

[54] To do this, I computed the annual distance by assuming that households own cars with an average fuel efficiency.

[55] In fact, it would be a rolling sample. For data protection reasons, the data does not contain a key that would enable the corresponding data from the different years to be linked.

[56] The data after this period is not yet available.

[57] The micro-census dataset collected by Bundesamt für Statistik (2006a) was taken as the benchmark for being representative, see Axhausen and Erath (2010: 35 ff.).

[58] The factors were: 1.00 for gasoline engines (standard), 0.75 for diesel engines, 0.9 for natural gas engines, 0.7 for hybrid drive systems and 0.6 for electric cars.

[59] The following price levels were possible: 10% cheaper, or 20% and 50% more expensive, respectively, than the present price level.

the household representative. The household representative then reported which type of car and engine he would choose, and how many kilometres per year would be driven in each of these cars. It was also possible to choose the option of no longer possessing a car. The interview was supported by a computer program, which helped respondents by displaying the fixed and marginal driving costs for each possible car and engine type. All respondents had to report their choice for six different price schemes. For some of these price schemes, there was a bonus for cars with an energy label[60] “A” or a surcharge for cars with an energy label “F/G”.[61] Since this dataset does not contain carless households, I added a number of randomly sampled observations from the micro-census dataset Bundesamt für Statistik (2006a). The number of samples I chose ensured that the proportion of carless households resembled that in the micro-census dataset. Since all models assume that households can only choose a standard car, I will ignore the fact that these costs could vary due to the fictitious taxes or rebates mentioned in the questionnaire by Axhausen and Erath (2010).

Stated preference dataset by Wüstenhagen and Sammer (2007)

Wüstenhagen and Sammer (2007) conducted a survey to examine consumers' behaviour when purchasing cars with respect to fuel economy. The survey was based on stated preferences, and was conducted by telephone using a prepared document. This document comprised all questions, including 21 sheets of car-related information. Each of these 21 sheets contained three car types that differ in car brand and model, engine size, fuel type, fuel economy, energy label[62] and price.[63] For each of these choice sets, households had to choose the car they would most probably select if they were to purchase a car on the day of the interview and only had these three cars to choose from. Since households had to choose 21 times, the dataset has a panel structure. To simulate a situation close to reality, only households that purchased a car in the previous twelve months were interviewed. In total, 156 households that purchased a small car costing under CHF 25,000 and 159 households that bought a medium-class car costing between CHF 25,000 and CHF 45,000 were interviewed. Consequently, households that purchased a small car had to choose only from choice sets containing only small cars, and households that bought a medium-class car only had to select from choice sets containing medium-class cars. In addition, consumers also provided information on their income, occupation,

[60] The energy label is a standardised label used widely for labelling energy-consuming products, such as light bulbs, dish washers, etc. in the European Union and Switzerland. The purpose of this label is to provide consumers with simple information about the energy efficiency of a product. Energy label “A” stands for high energy efficiency; label “G” means the energy efficiency of the corresponding product is very low.

[61] For more detailed information, see Axhausen and Erath (2010: 26).

[62] Energy label “A” stands for high energy efficiency; label “G” means the energy efficiency of the corresponding product is very low.

[63] A sample sheet with table containing all car types and the possible values of the attributes can be found in Appendix A1.1.

household size, etc. as well as on a number of political opinions, the car they purchased recently and their annual mileage. In addition, they were also asked to respond to questions concerning their decision-making process when they recently bought a car, such as whether they paid attention to fuel economy and where they obtained their information about the various car models.

1.5 Model types

In this subchapter, I will present the models commonly used to simulate traffic demand. First, I will discuss models used to analyse time-series data. I will then present traditional models used to evaluate cross-sectional and panel data. I call these models traditional because they are often used to estimate the demand for goods. I will then focus on models that are adapted for particular contexts in the field of transportation economics, based on cross-section data. I will present a model that focuses on car choice and use, which may be able to forecast the effect of taxes on car ownership and use. Since the models will be presented in detail in later chapters, I will merely give brief descriptions of them at this point. The aim of this subchapter is to create an overview of existing models, to discuss their basic properties and limitations. Finally, I will outline the models I decided to use for my particular applications.

Time-series models

Time-series models can be used to evaluate aggregated data at the country level, if the data series are available for a longer period. In almost all cases, an error correction model by Engle and Granger (1987) is applied. One principal problem is how to treat the size of the car stock. On the one hand, the size of the car stock strongly influences driving demand. However, the size of the car stock is also driven by the same factors as those that drive demand itself, e.g. the unobserved preference for travelling or the GDP. Since there is no macroeconomic variable that affects only the demand for car ownership without influencing the aggregate driving distance, an instrumental variable approach is not possible. If, in contrast, the car stock is omitted from the regression, there is an omitted variable problem because the omitted variable is correlated with the GDP, which is used as an explanatory variable. In each case, therefore, the estimated parameters will be biased. A further problem is finding an integrated relation between variables. The existence of such an integrated process would, for instance, mean that there is a fixed ratio between the GDP and the total driving demand or that there is a long-run relation between GDP, fuel price and total driving demand that is constant. Further, given the data for a specific country, coefficients denoting the elasticities of interest should, of course, be statistically significant. Unfortunately, neither Graham and Glaister (2002) nor Schleiniger (1995) present a significant coefficient corresponding to the long-term fuel price elasticity of fuel demand using Swiss data.[64] In contrast, in the model applied in Baranzini et al. (2009), the coefficient relating to the long-term fuel price elasticity was significant. This is probably because they added a trend variable and a dummy variable that relates to the increase in fuel taxes in 1993. Since I am more interested in comparing the effect of fuel taxes versus taxes on car ownership, I did not analyse time-series models.[65] This is because the effect of taxes on car ownership cannot be tested using the standard error-correction model. I also do not consider extending the model to enable this as promising.[66]

Traditional models for evaluating cross-sectional data

In many cases, survey data on household level is available as cross-sectional data. Such datasets usually consist of a vast number of household-specific variables, allowing for the use of numerous explanatory variables. As mentioned in Subchapter 1.4, one basic problem concerning this data is that fuel prices do not vary sufficiently across households for two reasons. First, the regional differences are minor or even imperceptible. Second, in most cases, fuel prices did not vary much during the survey when the individual households reported their data. However, cross-sectional data is very useful when attempting to identify individual factors that influence driving demand.

The most basic model to explain driving distance or fuel demand is the OLS model, which is defined as follows:

$$y = X\beta + \varepsilon \ , \qquad (1.5.1)$$

where X is a matrix defined as $X = \left(x_1,..,x_n,..,x_N\right)'$, where the vectors x_n contain the observations of explanatory variables, y is the vector of observations $y = \left(y_1,..,y_n,..,y_N\right)'$, β is a parameter vector with the dimension of x_n and ε is a vector of random variables $\varepsilon = \left(\varepsilon_1,..,\varepsilon_n,..,\varepsilon_N\right)'$, which are independently and normally distributed, with $\varepsilon \sim N\left(0,\sigma^2 I_N\right)$.

---

[64] See table in Graham and Glaister (2002) on page 8 and comments on the results for Switzerland on page 7. For Schleiniger (1995), see Table 2 on page 6.

[65] I consider the price index on car ownership to be inadequate for computing a coefficient corresponding to the impact of the cost of car ownership on driving demand. This is because this index is not corrected for quality and comfort features of the "average" car it reflects.

[66] One possibility would be to use a vector error correction model (VECM), introduced in Johansen (1991).

The key problem with this approach in the context of traffic demand studies is that some assumptions of the OLS model are violated.[67] Figure 1.5.1 illustrates these violations by showing the distribution of driving distance of urban households with an annual income of CHF 60,000.

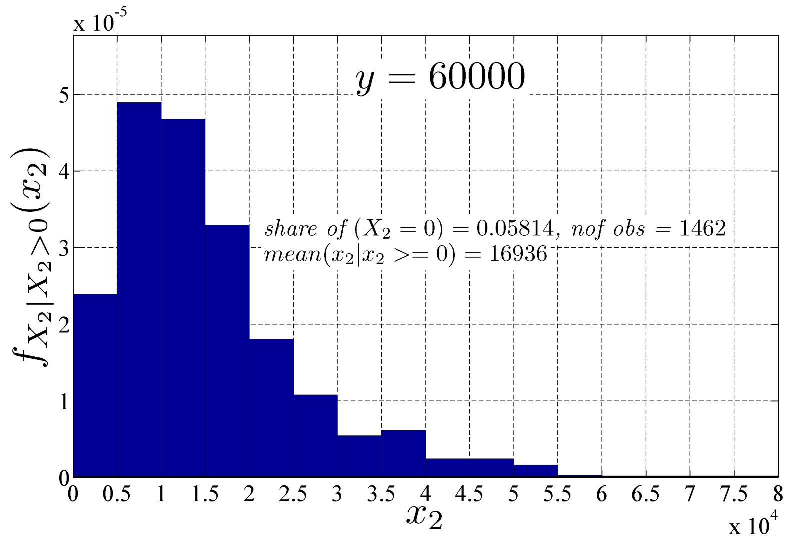


Figure 1.5.1: Distribution of driving distance of urban households with an income of CHF 60,000.

The shape of the histogram clearly does not correspond to a normal distribution.[68] First, the distribution is asymmetric and there is a heavy tail at the right. Second, the distribution is cut at the left due to the non-negativity constraint. Third, the variance of the error term increases with income.[69] The assumptions on the error term are therefore violated. Moreover, due to this, the assumption that the error term and the explanatory variable are uncorrelated is violated.[70] The higher the income, the more distinct the latter problem becomes. This is due to the fact that, when a household's income decreases, the mean of the normal distribution also decreases. There is therefore more density on the lower limit of the distribution, which should have been set to zero. In other words, the actual error term is negatively correlated with income. Since the error term and the explanatory variable are correlated, it follows that the estimated parameters are biased, as are any elasticities computed using these parameters.

[67] For the complete list of all assumptions, see Chapter 2.1.2.

[68] Note that, firstly, the shape of the histogram may also be non-normally distributed due to the omission of a number of explanatory variables. Secondly, in its general form, the OLS model is not based on normally distributed error terms, but on iid error terms. Thus the large sample properties of the OLS model would not be violated due to this.

[69] See histograms in Appendix A3.4. Since the variance of the error term depends on the income, the iid assumption of the OLS model in its general definition is violated.

[70] See Assumption (4) in Chapter 2.1.2.

The problem caused by the non-negativity limitation of the dependent variable can be overcome by using a limited dependent variable model. The most frequently used model in this context is the Tobit model. Although this model is also based on the assumption that the error term is normally distributed, unlike the OLS model, it takes into account that the dependent variable may not be negative. To this end, the Tobit model treats all negative outcomes as zero. In other words, the model predicts a probability that the outcome is zero. In the specific application of travel demand, this means that the model predicts a certain probability that a household does not own a car. Similar to the OLS model, the Tobit model is based on the assumption that the error term is normally distributed.[71] Formally, the Tobit model can be written as follows:

$$y_i = y_i^*,\ if\ y_i^* \geq 0,$$
$$y_i = 0,\ \ if\ y_i^* < 0,$$

where $y_i^* = x_i\beta + \varepsilon_i$. (1.5.2)

Matrix X, vector y and the random vector ε are defined as in the OLS model. Vector y^* consists of so-called latent variables y_i^*. This vector is distributed as $y^* \sim N\left(X\beta, \sigma^2 I_J\right)$. The probability that the observed variable y_i - in this case the driving demand - is equal to zero is therefore $P\left(y_i^* < 0\right)$. The conditional distribution of $y_i \mid y_i^* \geq 0$ still does not adapt the shape of empirical distribution presented in Figure 1.2.1 very well,[72] but this should not cause a greater error than in the case of the OLS model since the problem of the limited dependent variable no longer exists. One key feature of this model is that both the expected value of driving demand $E\left(y_i^*\right)$ and the probability of being carless $P\left(y_i^* < 0\right)$ depend only on $x_i\beta$ and σ^2. This implies that the relative impact on the expectation value of driving demand[73] versus the impact on the probability of being carless is identical for each explanatory variable.[74] Since it could be the case that some variables have a stronger impact on the decision

---

[71] Note that it is also often assumed that ε is logistically distributed, "Logit" model. This modification does not change the results considerably, but this assumption is often made because the computation effort involved in estimating parameters $X\beta$ and σ is less demanding.

[72] The shape of the empirical distribution is very different to the shape of a normal distribution.

[73] Note that $\partial E(Y)/\partial x_i$, with $x_i = p_{fuel}$ is used to compute the marginal effect of driving demand with respect to the fuel price p_{fuel} elasticity of driving demand.

[74] This fact can be illustrated as follows: Imagine that the coefficients corresponding to the household's income and the population density at the place of residence are 1.2 and -2.4, respectively. Let us assume further that an increase in income by one unit increases the expectation value of driving demand by 1% and decreases the probability of being carless by 2% at the mean values of the explanatory variables. The change of the latter induced by an increase in income or any other explanatory variable is therefore double the change of the expectation value of driving demand. If, in contrast, the population density at the place of residence decreases by one unit, the expectation value of driving demand increases by 2% and the probability of being carless decreases by 4%, the ratio is again 1:2. This is because these effects are only driven by a change in $x\beta$ and, since

whether or not to own a car than on driving demand than other variables, the restriction of the Tobit model, where this is not possible, is its major disadvantage.[75]

One model that can overcome this disadvantage is the sample selection model. This type of model consists of a choice and a use equation. Formally, the sample selection model can be written as follows:

$$z_i^* = w_i\gamma + u_i, \tag{1.5.3a}$$

$$y_i = x_i\beta + \varepsilon_i\,,\ y_i > 0 \text{ only if } z_i^* > 0\text{, and } y_i = 0 \text{ otherwise,} \tag{1.5.3b}$$

where u_i and ε_i are bivariate normally distributed $(u_i, \varepsilon_i) \sim N\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}1 & \rho\\ \rho & \sigma_\varepsilon\end{bmatrix}\right)$. (1.5.3c)

The choice equation (1.5.3a) describes whether a household owns a car, which implies that the observed driving distance is positive. Vector w_i contains variables that influence the probability of a household owning a car, e.g. income, household size or type of residence. It becomes apparent from condition $z_i^* > 0$ that the higher value $w_i\gamma$, the higher the latent variable of the choice equation z_i^* is, and therefore the greater the probability that a household owns a car. Equation (1.5.3b) describes the driving distance in the event that a household decides to own a car; vector x_i contains variables that influence driving distance y_i. The random terms (u_i, ε_i) contain unobserved variables that influence the choice of owning a car and driving demand, respectively. Since unobserved variables exist that influence both the choice of car ownership and driving demand,[76] the error terms ε_i and u_i are correlated. The correlation coefficient between ε_i and u_i is denoted by ρ. Note that the vector corresponding to choice decision w_i may contain variables that are also contained in the vector corresponding to demand decision x_i. It follows from the model structure that part of the driving demand, mapped by Equation (1.5.2b), can be explained by information gained from the choice model mapped by Equation (1.5.2a), due to the correlation between error terms ε_i and u_i. For this reason, when applying a two-step procedure where first the choice model (1.5.3b) and then the demand model (1.5.2a) is estimated, a correction term $\rho/\sigma_\varepsilon \cdot \lambda\left(w_i\gamma/\sigma_\varepsilon\right)$ is added to the demand equation (1.5.3b). The term $\lambda\left(w_i\gamma/\sigma_\varepsilon\right)$ is the so-called inverse Mills ratio.[77] Theoretically, vectors w_i and x_i can contain the same variables. In practice, however, there should be at least one variable in w_i that

---

the parameter accounting for the impact of the population density is twice that corresponding to the household's income, the impact of the first variable on $x\beta$ is also twice that of the income per unit increase.

[75] Note that there is also the problem in the case of the Tobit model that the variance of the error term increases with income, which leads to problems that become evident in Figures A3.15.3 and A3.15.17.

[76] For example, the unobserved preference for driving a car.

influences only the choice whether to own a car at all.[78] Unfortunately, there is no explanatory variable that explains only the decision whether to own a car at all: I expect all available explanatory variables to influence both the probability of owning a car and the use of it.[79] Due to this, the sample selection model is unfortunately not applicable for mapping both decisions regarding car ownership and use.

Some datasets also contain information about the types of cars households own or the type of car they would choose in a fictitious situation. In these two cases, choices can be mapped by a discrete-choice model. A discrete-choice model is based on the assumption that a household *i* chooses the option I_i out of a possible set of options C_i that provides the highest utility z_i^*:[80]

$$I_i = \arg\max_j z_j^* = x_{ij}\beta_j + \varepsilon_{ij}, \ \forall j \in C_i \tag{1.5.4}$$

Vector x_{ij} contains household-specific data and attributes of the corresponding options, such as engine size or fuel efficiency, in the case of a survey on car choice. Note that the coefficients of β_j relating to household-specific attributes differ for each option *j*. As an example, this reflects that the preference for utilising a VW Golf could decrease when income increases, whereas preference for a Mercedes C-Class increases. Note that the parameters corresponding to car attributes can also principally vary between car types. For instance, it could be assumed that a 400 cm^3 increase in engine size would

---

[77]The inverse Mills ratio is defined as $\lambda\left(w_i\gamma/\sigma_\varepsilon\right) = \frac{\phi\left(-w_i\gamma/\sigma_\varepsilon\right)}{1-\Phi\left(-w_i\gamma/\sigma_\varepsilon\right)}$, see Puhani (2002: 55), where ϕ and Φ denote the pdf and cdf, respectively, of the standard normal distribution.

[78] Puhani argues that when all explanatory variables w_i and x_i are identical, there will be an identification problem of the parameters, due to a collinearity problem. This collinearity problem occurs in this case because the parameters can only be identified by the non-linearity of the inverse Mills ratio. In practice, however, most observations will be in the almost linear range of function $\lambda(\bullet)$, see Figure 1 in Puhani (2002: 57). Referring to Little and Rubin (1987: 230), Puhani implies that there must be at least one variable that predicts latent variable z_i^* well but does not predict variable y_i.

[79] One of the only variables which I believe influence only the choice of whether to own a car is if an individual is unable to walk and uses a wheelchair. In this case, use of a car is a very good option for travelling, and the fact an individual has a walking disability does not imply that he has a higher driving distance demand. Since, however, there is no variable indicating whether a person has a walking disability, there is no solution for the collinearity problem. Furthermore, even if this variable were available, I doubt whether its use would provide a solution. Since only a very small proportion of people have a walking disability, this variable would not vary much. Note that the collinearity problem occurs in both the two-step method using the error-correction term $\rho/\sigma_\varepsilon \cdot \lambda\left(w_i\gamma/\sigma_\varepsilon\right)$ and the Maximum Likelihood estimation method.

[80] Note that it is often the case that not all options are considered to be contained in choice set C_i. For instance, an expensive luxury car is not considered to be contained in the choice set of a low-income household. Also, in the case of a stated preference survey where people are asked about their choice conditional on fictitious attributes such as a certain price level, only limited sets of choices are presented to households. There are two reasons for this. First, the number of choices should be limited because it would otherwise imply that households require too much effort to decide, involving the risk that it would simply choose one of the options randomly. Second, only relevant alternatives should be presented to households. For instance, a luxury car is not a relevant alternative for a low-income household.

increase utility of a VW Golf more than that of a Mercedes C-Class. In most cases in practice, however, it is assumed that the impact of car attributes on z_j^* is the same for each option *j*. The error term ε_{ij} stands for unobserved household attributes, impacts of explanatory variables not captured by the model, or unobserved attributes of the different options *j* that can be chosen.

Traditional models for evaluating panel data

In some cases, survey data at the household level is available for several years. This allows us to examine the impact of changes of exogenous variables on the dependent variable at the household level. Another type of panel data stems from stated preference surveys. In these datasets, households are repeatedly asked about their planned consumption – in the context of transport, about their planned driving behaviour and choice of mode of transport – conditional on various price regimes or changes in other exogenous variables. In both cases, panel data models are use to evaluate the data. These models differ from the standard models presented in the previous paragraph in that they either add a household-specific constant, in the case of a fixed effects model, or they add a household-specific error term, in the case of a random effect model. Both the constant and the random variable should account for unobserved household-specific attributes. Note that adding a constant for each household is only possible in the case where all relevant explanatory variables vary between the observations of each household. Note that even if all relevant explanatory variables do vary, the random effects model can still be the "better"[81] model compared to the fixed effects model. If the panel consist of data stemming from different points in time, it often makes sense to capture a variable that accounts for a trend in time or for variables that change over time, but are identical for all households. In the case of traffic demand models, such variables could be the number of sunny days in one year or, of course, fuel price, since these two variables influence the total number of kilometres driven. The advantage of such data is that fuel prices vary since the observations stem from different years, and therefore parameters referring to fuel price can be estimated.

Microeconometric models that simultaneously capture choice of car type and use

It is interesting to use models that capture both driving demand and the car type chosen for three reasons: first, the consumers' choice of car type strongly influences fuel demand since every car has a specific fuel efficiency. Second, a more fuel-efficient car has lower marginal costs of driving, therefore driving demand will be higher. Third, a more comfortable car will increase the frequency of driving and will therefore increase demand for driving. Since the studies presented in Subchapter 1.3 showed that the effect of fuel efficiency on fuel demand accounts for up to half of the total effect on fuel demand when fuel prices increase, it should be highly advantageous for a model to include households' choice of car.

[81] Statistical test methods can be used to decide whether a fixed or a random effects model is likely to fit the data better.

The first model in the literature that is able to capture both car choice and use is the discrete-continuous choice model introduced by Dubin and McFadden (1984). They used this model framework to examine households' choice of heating system type, namely gas versus electricity. They explained households' choice of a type of heating system[82] and the intensity of use[83] by electricity and gas prices, the average ambient temperature in winter at the households' place of residence, various characteristics of the house, such as floorspace, and a number of household characteristics. Since this approach can also be used for a set of more than two types, these models are also applicable when households choose between different types of car and the intensity of use, i.e. the annual distance they will drive this car. The drawback of this model is that it can only capture the behaviour of households with just one car. This model also offers the option of a household deciding against owning a car if there is sufficient data on the total distance travelled using public transport. It is not possible to capture the behaviour of households that decide to maintain more than one car using this model. Formally, the model consists of a Marshallian function that stands for driving demand and its corresponding indirect utility function. This framework maps the following microeconomic calculus: the household takes into account the choice of a certain car and then chooses the number of kilometres it would drive in this car. After paying the fixed costs of the car and the driving costs, the remaining budget is used for all other goods. Given the car choice and the corresponding consumption level, the household calculates its overall utility. The household applies this procedure to all car models available and then ranks the cars according to the level of overall utility. It then chooses the car at the top of the ranking. The outcome of this decision process is what is assumed to be observed in the data. The household's behaviour is mapped by this model as follows:

$$\max_i v_i(p_{in}, y_n - r_i, b_i, s_n, \varepsilon_{in}, \xi_{in}) = e^{-\beta p_{in}} \left(\frac{\alpha}{\beta} + \beta (y_n - r_i) + \alpha \beta p_{in} + \gamma s_n + \delta b_i \right) + \xi_{in}, \qquad (1.5.5a)$$

$$x_{in} = x_i(p_{in}, y_n - r_i, s_n, \varepsilon_{in}) = \alpha p_{in} + \beta (y_n - r_i) + \gamma s_n + \delta b_i + \varepsilon_{in}, \qquad (1.5.5b)$$

where y_n is the income of household *n*, r_i are the fix costs of car type *i*, and p_{in} are the costs per kilometre when driving car type *i,* which depends strongly on the fuel price and fuel efficiency of car type. The socio-demographic variables are denoted by s_n and the car attributes by b_i. The function $v_i(p_{in}, y_n - r_i, b_i, s_n, \varepsilon_{in}, \xi_{in})$ is an indirect utility function, indicating the level of utility household *n* can achieve, given its income y_n and its choice of car type *i.* The Marshallian function

[82] The survey was conducted in a region where households use either gas or electric heating systems. The model therefore captures only these two types of heating system.

[83] The intensity of use was measured by annual energy consumption.

$x(p_{in}, y_n - r_i, s_n, \varepsilon_{in})$ describes the number of kilometres per year the household would drive with a car of type *i*. In contrast to all previously presented models, Equations (1.5.5) show that the model takes into account not only driving costs, but also the fixed costs of car ownership. The random terms ξ_{in} and ε_{in} represent unobserved socio-demographic variables, unobserved car attributes and measurement errors. There are some assumptions on the error terms, such that an easily computable error-correction term is yielded. This error term can be computed from results obtained by solving a multinomial discrete choice model, based on the indirect utility function (1.5.5a). This correction term captures information that can be gained from the choice model, which captures car choice and partly explains driving demand. The following example may help give an insight into the purpose of this correction term. Imagine a household has an unobserved preference for large, comfortable cars and hence also for driving an over-average distance. A simple solution would be to add a dummy variable accounting for luxury cars in the demand equation. But since both driving demand and the probability of buying a luxury car are driven by the same unobserved factor, there would be an endogeneity problem that would lead to biased estimators. The purpose of the error-correction term is now to include the information that can be obtained from the choice model in the demand equation such that it does not cause biased estimators.

The second model in the literature that can capture both car choice and use, the so-called multiple discrete-continuous extreme value model (MDCEV), which was introduced by Bhat (2006).[84] In addition to the model by Dubin and McFadden (1984), this model also maps a possible decision by households to own none or several cars. It is assumed that households may choose to own one or several cars from a set of car types; they can also decide for each car how far they will drive. The only restriction is that households may not choose more than one car of the same type. The model is specified as follows:

$$U = \sum_{i=1}^{J} u_i(X_i) = \sum_{i=1}^{J} \exp(m_i + \xi_i) \cdot (X_i + a_i)^{d_i}, \tag{1.5.6a}$$

where $m_i = \gamma_i \cdot s + \delta \cdot b_i$,

subject to $y = \sum_{i=1}^{J} p_i X_i$. (1.5.6b)

[84] Initially, Bhat (2005) developed this model to explain the activities of tourists. This model maps a situation in which tourists may choose from a set of leisure activities, for each of which they choose a certain duration. The model's restriction is a time budget of 24 hours. Hence the marginal “price” of an additional minute spent doing one activity is that one minute less is available for the other activities. I extended this model to the car choice decision, as mentioned in this paragraph. Bhat followed the same approach in his publication Bhat (2006). Additionally, however, he also used slightly different specifications of the utility function, which is a positive transformation of the function I use here.

Variables X_i denote the driving distances of the corresponding car types $i = 2,...,J;$ X_1 denotes level of consumption of the remaining goods and U denotes the overall utility of a household. Note that if a household does not hold a certain car type i then the corresponding driving distance is $X_i = 0$.[85] As in the discrete-continuous model by Dubin and McFadden, which is described in the previous paragraph, r_i denotes the fix costs of car type i, p_i the costs per kilometer when driving car type i, s socio-demographic variables and b_i car attributes like the engine size or the number of doors. The assumptions on error terms are such that an easily computable Maximum Likelihood function of a closed form is yielded, which is a great advantage of this model. Its major drawback is that it does not capture the fixed costs of car ownership. For this reason, MDCEV does not exclude that it is irrational to hold a large number of cars, since such a choice does not affect the available budget by way of the fixed costs this causes. Note that one key advantage of this model over the previously presented models is that it contains a budget restriction. Thus, in contrast to the other model, the density function corresponding to the driving demand is strictly zero for any distance above which households would spend their entire income for car driving.

A microeconometric model that captures the option of being carless and includes fixed costs of car ownership

To overcome the drawbacks of the two models presented in the previous paragraph, De Jong (1990) developed a model that captures household's choice between being carless or owning one standard car and the distance driven in it. In contrast to the traditional microeconometric models I presented in a previous section, De Jong (1990) included the fixed cost of car ownership in his model. De Jong's model is based on the same principle as the discrete-continuous model by Dubin and McFadden (1984). However, he uses a different Marshallian demand function, which means that the indirect utility function is different, too.[86] In the framework of this model, the indirect utility function v of household n and the natural logarithm of the Marshallian demand function x are given by

$$v_n = v(p_n, y_n - r, b, s_n, \varepsilon_n) = \frac{1}{\alpha} \cdot e^{-\alpha p_n + \gamma s_n + \delta b + \varepsilon_n} + \frac{1}{1-\beta} \cdot (y_n - r)^{1-\beta}, \tag{1.5.7a}$$

$$\ln(x_n) = \ln(x(p_n, y_n - r, s_n, \varepsilon_n)) = \alpha\, p_n + \beta \ln(y_n - r) + \gamma s_n + \delta b + \varepsilon_n, \qquad \text{(1.5.7b)[87]}$$

where, again, y_n denotes the income of household n, r the fixed costs of car ownership, p_n the driving

[85] Note that the parameters must be chosen so that $\partial u_i(X_i)/\partial X_i$ is finite if $X_i = 0$.

[86] This Marshallian demand function and its corresponding indirect utility function can be found in Hausman (1981: 669). It can also be seen there how the indirect utility function can be derived from the Marshallian demand function.

[87] These formulas are identical to those in De Jong (1990) and De Jong (1997).

costs, s_n socio-demographic variables of household n and b are car attributes. Note that driving costs p_n and fixed costs r correspond to an average car. Following from $\lim_{\varepsilon_n \to -\infty} \ln(x_n) = -\infty \Rightarrow x_n \downarrow 0$, De Jong (1990) argues that

$$v_0 = v_0(p_n, y_n, b, s_n, \varepsilon_n) = \lim_{\varepsilon_n \to \infty} v(p_n, y_n, b, s_n, \varepsilon_n) = (1-\beta)^{-1} \cdot y_n^{1-\beta} = \frac{1}{1-\beta} \cdot y_n^{1-\beta} \tag{1.5.8}$$

corresponds to the utility a carless household yields. The household then decides whether $v_0(p_n, y_n, b, s_n, \varepsilon_n) > v(p_n, y_n - r, b, s_n, \varepsilon_n)$. This inequality is satisfied for any ε_n below a certain household-specific threshold $\varepsilon_{n,0}$, $\varepsilon_n < \varepsilon_{n,0}$.

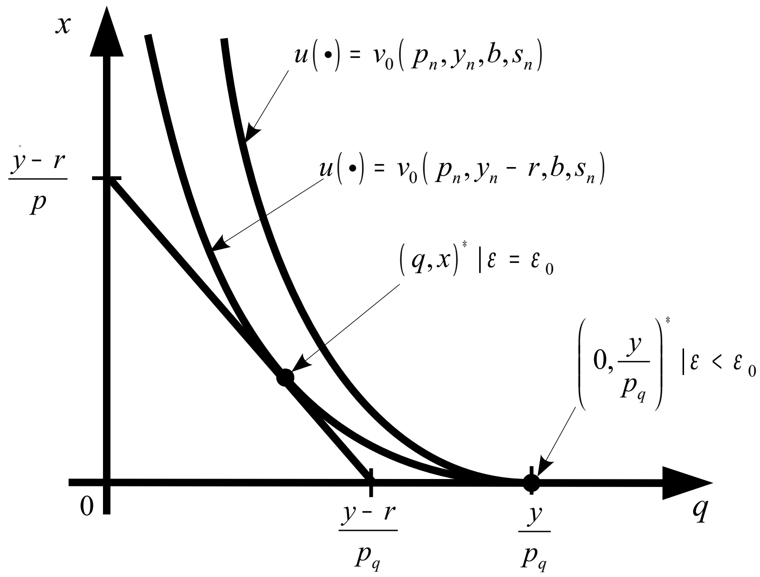


Figure 1.5.2: Minimum driving distance according to the model of De Jong (1990).

Note that, from the researcher's perspective, random variable ε_n is unobserved, although it is known to households. Note that the optimum driving distance x is given when $\varepsilon_n = \varepsilon_{n,0}$ denotes the minimum driving distance given parameters $\alpha, \beta, \gamma, \delta$ and variables p_n, y_n, r, b, s_n. According to the model, therefore, there should be no observations below this value. Since in reality there are often observations below this value, it would not be possible to estimate the model using the Maximum Likelihood method because these observations have a theoretical probability of zero. To overcome this problem, De Jong (1990) adds an additional error term to the Marshallian demand function:

$$\ln(x_n) = \ln(x(p_n, y_n - r, s_n, \varepsilon_n)) = \alpha\, p_n + \beta \ln(y_n - r) + \gamma s_n + \delta b + \varepsilon_n + \kappa_n . \tag{1.5.9}$$

De Jong justifies this by claiming that this error term accounts for planning errors. He claims that households plan according to the indirect utility function whether or not they wish to purchase a car and how much to use it. Once they actually have the car, however, they will use it more or less than planned because of changes in preferences or altered traffic conditions, e.g. an increase in traffic jams. The problem of this approach is that the Marshallian demand function and the indirect utility function are linked by Roy's identity. In other words, adding error term κ_n to the Marshallian function would also add this term to the indirect utility function. Since De Jong (1990) neglected this, the choice and demand model no longer correspond to one another. In my eyes, De Jong's extension is only feasible if the variance of κ_n is much smaller than that of ε_n, since then De Jong's model would be almost identical to that without error term κ_n and therefore Roy's identity would be approximately satisfied. Unfortunately, the resulting variance of κ_n is even larger than that of ε_n. I doubt, therefore, whether the results are realistic. An initial indicator of why De Jong's results are unrealistic is that the estimated standard deviation of κ_n is 0.5, see De Jong (1990: 980),[88] This would mean that the +/- 1 standard deviation interval of planning error is on average -40% .. 65% in relation to the planned distance, which seems very unrealistic to me. A second indicator of why De Jong's results are unrealistic is that the results yielded by his model are very different when using different data, as the following table shows.

		Bjorner (1999)	Average of three studies[89]	Ramjerdi and Rand (1992)	De Jong (1990)
Country		Denmark	Denmark	Norway	Holland
Model type		"De Jong"	time-series	"De Jong"	"De Jong"
Income:	Car ownership	0.33	0.6 .. 0.9	0.41	0.15
	km driven	0.42	0.6 .. 0.9	0.63	0.26
Driving costs:	Car ownership	-1.33	-0.50[a] (0.34)[b]	-0.78	-0.41
	km driven	-1.63	-0.58[a] (0.58)[b]	-1.11	-0.80
Fixed costs:	Car ownership	-2.65	-0.4 .. -0.6	-1.29	-0.80
	km driven	-2.48		-0.88	-0.48

a These values do not refer to the three studies using Danish data. This data is based on the average of international studies.
b The values in parentheses "( )" denote the standard deviation of the values given in the different studies. The values are based on the study by Goodwin et al. (2004), and can be found in Table 1.3.2. I multiplied the values given in Table 1.3.2 by a factor of two because fuel costs account for only approximately half of all variable costs.

Table 1.5.1: Optimum decision of a household that chooses not to own a car.

[88] Bjorner (1999) using the same model even gets a value of 0.92, see Bjorner (1999: 385).

[89] I found these results in Bjorner (1999: 389). I computed the elasticity of car ownership with respect to fixed costs by multiplying the values of the elasticity of car ownership with respect to car purchase costs by a factor of two because the amortisation of the car makes up only around half of the fixed costs.

In addition, the results differ considerably to the results yielded by other studies: whilst elasticities with respect to income are smaller than those in other studies, elasticities with respect to driving costs are higher on average than those in the other studies. In particular, I believe that the elasticities of driving distance with respect to variable and fixed costs yield in De Jong's model are too high in absolute values because the model yields effects of the two variables on car ownership that are too high, see Table 1.3.4.

My further criticism is that the marginal utility tends to infinity when driving distance approaches zero. Economically, this implies that if there were no fixed costs of car ownership, everybody would own and drive a car irrespective of the fuel price level. I do not expect this to be the case because people living close to a highly frequented node of public transport may still own a car, even if the fixed costs were zero. I believe that if the model could be parametrized such that driving demand can be zero in the case of the absence of fixed costs, the decision regarding car ownership and car use could be treated more independently and the model would yield more realistic results. Another drawback of this model is that it captures only one standard car and does not allow for swapping for more fuel-efficient cars when fuel prices increase. Finally, the results of these studies should also be viewed with caution because the authors do not use marginal changes of the independent variables when computing elasticities. Instead, they use changes of 10%. Since the functional dependency of car ownership on the dependent variables of interest is strongly non-linear in the range of interest, this leads to rather considerable inaccuracies.[90]

I conclude from the above that use of De Jong's model is not very promising. I therefore choose to base my main model on the framework of the multiple discrete-continuous extreme value (MDCEV) model. The main reason for this choice is that the impact of taxes on car ownership and use can only be computed using the MDCEV model. In addition, even when using cross-sectional data in which the fuel price variation is very low, the impact of changes in fuel price on car ownership and use can be computed. I will present this model in Chapter 3. To obtain results comparable to those of the MDCEV model, I will also compute results using the OLS and the Tobit model, see Chapter 2. Further, using a Multinomial Logit model, I will estimate the willingness of households to pay for improvements in fuel efficiency using a stated preference data set, see Chapter 4. A more detailed discussion on the models chosen can be found in the subsequent Subchapter 1.6.

[90] See Figure 2 in De Jong (1990: 981).

1.6 Model choice

In this subchapter, I will discuss the models presented and used in the following chapters. With the exception of the car choice model, which uses panel data, all other models use cross-sectional data.[91] Since use of the time-series data available to me led to unsatisfactory results, I did not compute elasticities based on time-series data. In the following, I will first summarize the major aim of my thesis. I will then decide which models to use to capture the decision whether or not to own a car and how much to use it. Then I will justify my choice in model for explaining car choice with respect to car fuel efficiency. Finally, I will discuss the possibility of using discrete-continuous choice models, given the available data.

The aim of the thesis

The aim of this thesis is to determine the effect of tax increases in fuel and in car ownership on fuel demand and car ownership. My main objective is to determine the effect of such taxes on total driving distance, total fuel demand and car ownership. Policy-makers can use my model to forecast the effect of a given tax scheme. The model will also capture household income, which strongly influences the level of car ownership and driving demand. This will also allow us to forecast the aggregate driving demand when GDP is growing. Furthermore, incorporating income enables us to examine the effect of the use of tax yield, for instance the effect of reimbursing tax in proportion to income can be tested. I shall also examine the effect of rebate systems for fuel-efficient cars and fuel prices on the car fuel efficiency chosen when households purchase a car. In the process, I will determine their willingness to pay for fuel efficiency. In the following I will summarize the models and describe which information they can perform. Finally, I will show by which models the results I aim for can be computed, given the data I have.

Models that capture choice of car ownership and use based on cross-sectional data

When choosing the optimum model that should explain the decision of car ownership and its use for the data given, a "good" model should be able to capture some peculiarities concerning Switzerland. The major peculiarity is that – despite the high income level – the ratio of car-less households in Switzerland is rather high, namely 20% versus approximately 8% in the USA. Further, the level of public transport is rather high, making it relatively easy to swap driving demand for use of public transport. Moreover, the regional differences in car use and car ownership are relatively high, due to topographic differences and differences in levels of public transport. As in all other countries, fuel demand is also influenced by a couple of household-specific variables, such as the number of

[91] Two exceptions are the OLS and the Tobit model when I compute results based on the stated preference dataset by Axhausen and Erath (2010).

household members. A model that maps fuel demand should therefore cover both car choice and car use, as well as a number of household attributes. As mentioned in the previous subsection, the fuel price in the data available does not vary sufficiently across households because regional differences in fuel price are small or imperceptible and, in most cases, fuel prices do not vary much in the course of the survey when individual households make their responses. The same holds for the fixed costs for cars: taxes on car ownership do not vary sufficiently such that an effect on car ownership and use could be measured directly. To capture both driving and fixed costs, a microeconomic model framework incorporating both these variables has to be used. Unlike De Jong's (1990) model I presented in the previous chapter, which is based on an indirect utility function, my model will be based on a direct utility function. The utility function I will use is based on the multiple discrete-continuous extreme value (MDCEV) model framework introduced by Bhat (2005). The utility function will also contain a number of parameters that correspond to marginal utilities that are finite when driving demand approaches zero, which is more realistic than in De Jong's model (1990). This renders it possible for the choice and the use model to be more independent from one another. In contrast to De Jong (1990), the model framework I use remains completely within the microeconomic framework. I solve the problem of observations below the minimum driving distance according to the model by eliminating them.[92] Since I use a utility function, I regard the driving distance to be the entity that provides utility. I therefore use household driving distance rather than fuel demand as the explained variable. As in the case of De Jong's model (1990), I assume that households may choose only a standard car. I treat the aspect of household behaviour with respect to car fuel efficiency when purchasing a car separately. To this end, I use a choice model with a different dataset, namely the stated preference dataset by Wüstenhagen and Sammer (2003).

Note that apart from the fact that this model captures driving and marginal costs, one incentive for choosing this was that, unlike De Jong's model (1990), it can be extended to the case where households can choose from several car types. The implementation of that extension is much more complex, and there is a risk that the available computation power is insufficient. For this reason, I only implement the simplest case where households only have one car type to choose from.

As mentioned in the previous section, the traditional model that is able to capture both car choice and use, namely the Tobit model, cannot perform a significant coefficient related to the impact of driving and fixed costs. Nonetheless, I will use this model to compute elasticities with respect to income because I believe they are close to the true values, since the income varies a lot between the households and the corresponding estimated parameter will be highly significant. I will then compare

[92] Since the minimum driving demand according to the model depends on model parameters. I therefore introduced a criterion to determine the optimum level of minimum driving demand.

these results with those obtained using the extended MDCEV model I developed. In addition, using the stated preference dataset by Axhausen and Erath (2010) in which they asked households about their fictitious behaviour in the event of fuel prices varying dramatically, I will also compute elasticities with respect to fuel prices and driving costs using a Tobit model. Since the dataset by Axhausen and Erath (2010) has a panel structure, I will also test whether the results change if this panel structure is captured by the model. I will also compute all results using an OLS model to determine the elasticity of driving demand. The reason for doing this is that I wish to examine whether the values of the elasticity of driving demand depend strongly on the type of model used to compute it. All of the models mentioned in this section will be applied to the micro-census data collected by Bundesamt für Statistik (2001) and (2006a) and to the data provided by Axhausen and Erath (2010). The data contained in the Swiss consumer expenditure survey collected by Bundesamt für Statistik (2007) will only be applied to the Tobit and OLS models, since it does not contain any information on driving distances.

A model that maps choice of car type and its fuel efficiency

The aspect of car choice is relevant with respect to fuel demand since car fuel efficiency plays an important role. Based on the stated preference dataset of Wüstenhagen and Sammer (2007), I will examine household behaviour with respect to car choice. In particular, I will test the hypothesis that suggests that households' willingness to pay for fuel efficiency is equal to the fuel costs they can save during the use of the car. In the survey by Wüstenhagen and Sammer (2007), households were repeatedly asked about their choice in a setting where different car types are presented in the choice set. In particular, the car types differed with regard to fuel efficiency. Since each household responded to several choice sets, the dataset has a panel structure. I will therefore use a discrete-choice model that accounts for this panel structure. Use of this model will enable me to determine household behaviour when purchasing a car, given that fuel prices rise.

Models that map both choice of car type and driving distance

The standard approach for capturing both car choice and its use is the discrete-continuous model framework of Dubin and McFadden (1984), as presented in the previous subsection.[93] Given the cross-sectional data that also includes information on car type, namely the micro-census data of Bundesamt für Statistik (2001) and (2006a), again, the key problem is that fuel prices do not vary across regions. The variation of driving costs is therefore solely determined by the fuel efficiency of the specific car type, and the household's reaction to fuel prices cannot actually be measured directly. Measuring its

[93] Note that capturing the car choice enables to compute the fuel demand of the households more precisely since the fuel efficiency of each car is known.

impact indirectly by means of differences in fuel efficiency is also problematic for the following two reasons. First, the model can only be performed by defining a restricted number of vehicle types. Typically, the number of vehicle types is about four to nine, depending on the size of the dataset and the available data. These car types therefore actually include a large group of similar car types – a car category – according to the model. Since I do not know which car type within such a car category would attract a specific household most, I assume that households can only choose from car types that have the average attributes of the cars in the corresponding category. This, of course, is a rather restrictive assumption. An additional problem is that the attributes of these car types within a car category do not vary between households. It is therefore not clear, for instance, whether the parameter representing driving costs really expresses the impact of these costs or whether it represents also some unobserved attributes of the car. When distinguishing car types by their engine size and omitting all attributes of cars, for instance, the outcome of a simple regression would be the lower the fuel efficiency due to larger engine size, the higher the driving demand.

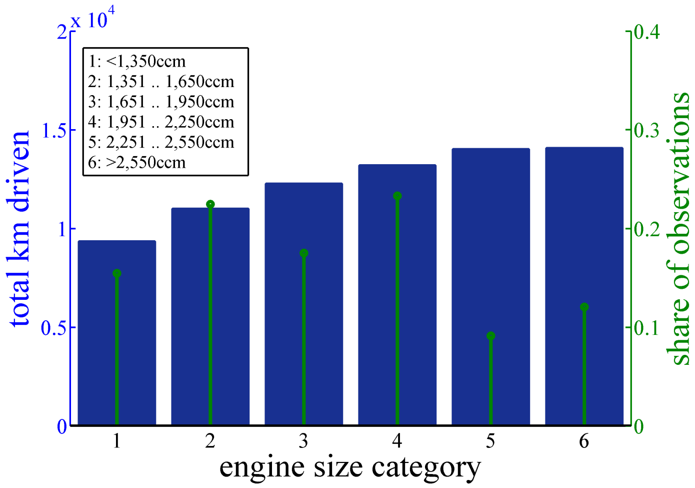


Figure 1.6.1: Average driving distance and car's engine size.[94]

Even if part of the effect of unobserved car attributes can be captured by the model chosen, I doubt whether the model will be able to map the true mechanisms. I particularly doubt that the model will reflect the impact of driving costs, and therefore fuel prices, correctly. Note that in their application on heating systems, Dubin and McFadden (1984) did not have the problem of a lack of variation in fixed and operating costs. Nor did they have the problem of unobserved attributes or unobserved household preferences, since the output – namely heat – is always the homogeneous good. Hence, their model mapped their decision problem correctly, and the results were satisfactory because the parameters were significant.

[94] This figure is based on the micro-census data of the Bundesamt für Statistik (2001).

I conclude from this that the discrete-continuous model by Dubin and McFadden is not suitable for computing results based on the available Swiss data. In the context of my extensive research into this model, I derived the error correction term presented by Dubin and McFadden (1984). The aim was to understand the exact mechanisms of this model and to establish whether modifications are possible such that given the data the aggregate own price elasticity of fuel demand could be computed.[95] Since this derivation cannot yet be found in the literature, I shall present it in Appendix A4.

[95] Note that I could not find a possibility for an extension that led to satisfactory results.

2. Analysis based on traditional models

Introduction

In this chapter, I shall analyse the data using "traditional" models. By traditional models I mean the Ordinary Least Squares (OLS) method and the Tobit model. As mentioned in the first chapter, these models may not provide all of the information I aim to compute. Neither may give information on the impact of fuel prices[96] and the fixed costs of maintaining a car on car ownership and driving demand. Further assumptions on which these models are based may be violated. Despite these drawbacks, the results of these models shall serve as benchmarks when testing the validity of the model I will present in Chapter 3. Furthermore, using these models, I will test which variables have a significant impact on driving distance. To obtain reasonable computation times, I will then only capture the most relevant of these variables for the MDCEV model.[97]

In this chapter, all model approaches consider a simplified situation in which households have only the choice between no car and a standard car. This standard car is considered to be an average medium-class car with corresponding fixed and variable costs. If a household owns several cars, it is considered as having one standard car.

In the following, I will only very briefly present the two models, namely the OLS and the Tobit model, since they are documented in detail in numerous standard textbooks. I will then show and discuss the results obtained using four different datasets based on surveys on Swiss private households.

[96] Except for the case of the stated preference dataset with variant fuel prices presented in Erath and Axhausen (2009).

[97] Note that the inclusion of many explanatory variables in the MDCEV model implies very high computation time effort when estimating the parameters. Since I found that the inclusion of additional variables in the MDCEV model does not change the values of interest, namely the elasticities with respect to income and fuel price, I include only variables I consider to be most relevant in the MDCEV model.

2.1 The Ordinary Least Squares model

The Ordinary Least Squares (OLS) approach is the very basic, simplest approach that can be used to explain the fuel demand of Swiss households. Of course, this approach has many drawbacks because it is unable to capture all relevant economic decisions taken by households, such as whether or not to own a car. Despite this deficiency, I shall start with this approach. I will show that certain assumptions upon which the OLS model is based are not met. This will lead to the need for a more sophisticated Tobit model, presented later in this chapter.

The OLS model is defined by the statistical model

$$y_n = \beta' x_n + \varepsilon_n,\ n = 1,...,N, \tag{2.1.1}$$

where x_n is a column vector of explanatory[98] variables, y_n is the explained[99] variable and β is a parameter column vector.Index n indicates the observation within a dataset; N is the total number of observations. The random term ε_n reflects the random component of y_n that cannot be explained by the explanatory variables x_n. It basically consists of two components: first, variables that explain y_n but are not available to the researcher. Second, the random term ε_n may contain measurement errors with respect to y_n or x_n.

The model above can be rewritten as:

$$y = X\beta + \varepsilon, \tag{2.1.2}$$

where X is a matrix defined as $X = \left(x_1,..,x_n,..,x_N\right)'$, y is the column vector of observations $y = \left(y_1,..,y_n,..,y_N\right)'$ and ε is a column vector of random variables $\varepsilon = \left(\varepsilon_1,..,\varepsilon_n,..,\varepsilon_N\right)'$.

[98] Alternative notations are “independent”, “control” and “predictor variable” or “regressor”.

[99] Alternative notations are “dependent”, “response” and “predicted variable” or “regressand”.

The OLS model is based on the following assumptions:[100]

(1) There is a linear relationship between the explanatory variable x_n and the explained variable y_n, as described by the function (2.1.2).

(2) The expected value of the random variable ε_n is zero for all observations *n*, i.e. $E[\varepsilon_n] = 0$.

(3) The variance of the random term ε_n is identical for all observations *n,* $\text{var}[\varepsilon_n] = \sigma \;\; \forall n$, and there is no correlation between the random terms of the different observations ε_n: $\text{cov}[\varepsilon_n, \varepsilon_m] = 0, \forall n \neq m$. This condition is equivalent to $E[\varepsilon, \varepsilon'] = \sigma I_N$, where I_N is an identity matrix of dimension *N.* This property is denoted as "homoscedasticity".

(4) The random terms are uncorrelated across the observations. There is no correlation between explanatory variables $X = (x_1,..,x_n,..,x_N)'$ and random terms ε, $E[x_n \cdot \varepsilon_m] = 0, \forall n \neq m$.

(5) Explanatory variables x_n are deterministic; there is no multicollinearity. This means that each component of $X = (x_1,..,x_n,..,x_N)'$, e.g. row *j* of *X,* is linearly independent of all other rows of *X,* and *N* is the total number of observations.

(6) The random term ε_n is normally distributed with variance σ^2: $\varepsilon_n \sim N(0,\sigma^2) \;\; \forall n$.

Note that all these conditions are conditional on the observed explanatory variables x_n.

The parameters are estimated using the OLS method. This means that the estimated parameters minimize the sum of squared deviations between the values predicted by the model and the observed values y_n:

$$\hat{\beta}_{OLS} = \arg\min_{\beta} \sum_{n=1}^{N} (y_n - \beta' x_n)^2 . \tag{2.1.3}$$

[100] See, for instance, in Greene (2003: 7 ff.).

(7) Solving this minimization problem yields[101]

$$\hat{\beta}_{OLS} = \left(X'X \right)^{-1} X'y \, . \tag{2.1.4}$$

Given Assumptions 1 - 7 above, the OLS estimator is equivalent to the Maximum Likelihood and the General Method of Moments (GMM) estimator, see, e.g. Ruud (2000: 539).

The problem with the data I wish to use is that the explained variable *y*, namely the driving distance, is non-negative. Assumption (2) is therefore violated if *y* is zero, since for this case, the error term ε_n is equal to the deterministic value $\beta' x_n$, $\varepsilon_i = -\beta' x_n$. It also follows from the non-negativity limitation of *y* that assumption (5) is violated.[102] Since the fraction of the total dataset of *y* being zero is about 20%, this may lead to biased estimators β. On the other hand, for high values of *y*, this effect does not occur, such that the bias of the estimated parameters should not be too strong.

The following tables show the results yielded for the different preference datasets I presented in Chapter 1. I shall start by presenting a table containing the results of a "large" model with a considerable number of household attributes. I will then provide a table containing the results of the model including only the most essential variables. I apply the second model for two reasons: first, I also intend to use a small number of household attributes for the MDCEV model to be discussed in

[101] Formula (2.1.2) can be rewritten as: $\hat{\beta}_{OLS} = \arg\min_{\beta} \left(y - X'\beta \right)' \left(y - X'\beta \right)$. The minimum is where

$$\frac{\partial \left(y - X'\beta \right)' \left(y - X'\beta \right)}{\partial \beta} = \frac{\partial \left(y'y - 2y'X\beta + \beta' XX'\beta \right)}{\partial \beta} = -2X'y + \left(XX' + \left(XX' \right)' \right) = -2X'y + 2XX'\beta = 0 \Rightarrow \beta = \left(XX' \right)^{-1} X'y.$$

To prove that this is a minimum, the second derivative must also be computed:

$$\frac{\partial^2 \left(y - X'\beta \right)' \left(y - X'\beta \right)}{\partial \beta^2} = \frac{\partial \left(-2X'y + 2XX'\beta \right)}{\partial \beta} = \left(2XX' \right)' = 2XX'.$$

Since the second derivative is positive definite for any value of β, the minimum value computed above is a global and unique minimum. The proof for positive definiteness is as follows: matrix *A* is positive definite if, for any vector values *s*, $s'As > 0$. Testing this definition for $2XX'$ yields $s'\left(2XX' \right) s = 2\left(s'X \right)\left(X's \right) = 2\left(s'X \right)\left(s'X \right)' > 0$.
See also Greene (2003: 21).

[102] This fact can be illustrated by the following example: imagine that variable x_1 denotes household income and that the corresponding coefficient β_1 is positive. Assume that $y = 0$ and therefore $\varepsilon = -\beta' x$. This implies that if income x_1 decreases, $-\beta' x$ increases, as does ε. For this reason, ε is negatively correlated with x_1 and Assumption (5) of the OLS model is therefore violated: $E\left[x_n \cdot \varepsilon_n \right] \neq 0, \forall\, n$ and the OLS estimation routine will yield biased estimators.

Chapter 3.[103] To compare the results of OLS to those of the MDCEV, I will use the same control variables for the "small" model as in the MDCEV model. Second, I would like to examine whether the results change dramatically when fewer household attributes are used. If this were the case, I would have to suspect that the results of the MDCEV might suffer from a significant bias. The following results are based on the standard OLS estimation procedure. Note that all datasets would contain weights for each observation, which I ignored for the following reasons: first, for computational reasons, I ignored the weights in the MDCEV model. Since I focus on determining whether the results differ if other models are used, I need to exclude the differences that can occur from capturing the weights in the model. Second, some empirical studies question whether the weights of these datasets are correct. Third, the datasets are very large. It is therefore unnecessary to improve the efficiency of the estimators as a result of using the weights. Fourth, the coefficients do not differ much when weights are used.[104]

In the following, I will first show the results of the elasticities of driving demand with respect to income and change of type of residence. The income elasticity of driving demand is defined as follows:

$$\varepsilon_{E(Y),inc} = \sum_{n=1}^{N} \frac{\partial E(Y_n)}{\partial inc_n} \cdot inc_n \cdot \left(\sum_{n=1}^{N} E(Y_n) \right)^{-1}, \qquad (2.1.5)$$

where y denotes driving demand. In the specific case of the OLS model, the elasticity can be computed by[105]

$$\varepsilon_{E(Y),inc} = \hat{\beta}_{inc} \cdot \overline{inc} / \overline{y} \,. \qquad (2.1.6)$$

[103] I incorporate only a small number of variables in the MDCEV model to save computational power.

[104] In the case of the dataset Mz00 when using weights, the coefficient relating to the income differs less than 5% from the value that yields by the OLS model when ignoring the weights.

[105] Note that in the case of the OLS model, $E(Y_n) = \hat{\beta}_{OLS} \cdot x_n$ and therefore $\partial E(Y_n) / \partial inc_n = \hat{\beta}_{OLS,\, income}$.

The “large” model that includes a vast number of control variables yields the following results:

	Mz00	Mz05	EVE
	OLS	OLS	OLS
	[1]	[2]	[3]
$\varepsilon_{E(Y),income} \mid \hat{\beta}_{OLS}, \bar{x}$	0.5612 (0.0135)	0.5727 (0.01108)	0.4631 (0.01711)
$\left(\frac{E(Y)_{rural}}{E(Y)_{city}} - 1 \right) \mid \hat{\beta}_{OLS}, \bar{x}$	0.4239 (0.02028)	0.5586 (0.0192)	0.3557 (0.0237)

Note: The notation “$\mid \bar{x}$” means that the values were computed conditional on mean values $\bar{x}$ of explanatory variables *x;* the values in brackets denote standard deviations.

Table 2.1.1: Effects of income and type of residence based on the large OLS model.[106]

The table above shows that all results are in a similar range. Nonetheless, I shall discuss a number of differences in the following. First, I do not find any obvious reason why the income elasticity of fuel demand is greater than the income elasticity of driving demand. It could be explained partially by the fact that the fuel demand captured by the dataset EVE contains also fuel used for motorbikes, which might tend to be used by households with lower incomes. In contrast, however, households with higher incomes own cars with lower fuel efficiency, such that fuel demand should depend more strongly on income than driving distance.[107] Second, the reason why the difference in fuel consumption between urban and rural households is smaller when considering fuel consumption (EVE) instead of driving distance (Mz00 and Mz05) could be that urban households drive more in urban traffic, where fuel consumption per kilometre is higher than when driving in rural areas. Further, rural households pay more attention to fuel economy when purchasing a car, since their mileage is higher.[108] When comparing the results of the Mz00 and the Mz05, the difference in driving distance

[106] Note that “Mz00” and “Mz05” stand for the micro-census datasets of 2000 and 2005, respectively, Bundesamt für Statistik (2001) and (2006a). The label “EVE” stands for the “Einkommens- und Verbrauchserhebung (EVE)”, Bundesamt für Statistik (2007).

[107] Using the dataset in Axhausen and Erath (2010), a simple regression of the income, fuel price and type of residence on car fuel economy yielded that households purchase cars that consume 0.139 litre more fuel per 100 kilometres for every CHF 1,000 more monthly income. The regression accounted for the panel structure of the dataset. The Institute for Environmental Decisions at ETH in Zürich, see Mueller (2010), provided me with a dataset that includes information such as floorspace and fuel efficiency for around 10% of all cars in the dataset. I was therefore also able to run a regression of household income on fuel efficiency of each household's car or fleet of cars. In this case, only 0.0652 litre more fuel per 100 kilometres resulted for every CHF 1,000 more monthly income per household. Note that if households own more than one car, I used the average fuel economy of the cars computed using the driving distance as the weight.

[108] The regressions mentioned in the previous footnote yield that households in rural areas own cars that consume 0.453 litres or 0.050 litres – Mueller (2010) – less per 100 kilometres than the cars owned by urban households. In this case, the difference between these results is considerable. Unfortunately, the last result is based on a non-significant parameter, and the first is based on a stated preference dataset. But at least the negative signs correspond with our intuition.

between rural and urban areas has increased. This could be because the service level of public transport in urban areas has increased, leading to a switch to this mode of transport by a significant proportion of urban households. Unfortunately, no data is available to support this hypothesis, since no data exists on kilometres travelled using public transport. Examining the rural effect separately for each year based on the EVE dataset yielded positive differences for 2003 to 2005 compared to 2002. None of them, however, were statistically significant. The coefficients of all explanatory variables are listed in Table 2.1.2. In Table 2.1.2 the coefficients of all three models show a similar pattern. As to be expected, due to the smaller income elasticity of driving demand, the coefficient corresponding to income is considerably smaller when using the data of the consumer survey (EVE). Interesting results are revealed by the dummies that stand for household attributes. Although the focus of this thesis is on the effect of income and fuel prices on driving demand and car ownership, I would like to briefly address the impact of household properties. When considering the family structure, it is interesting to note that one-parent family households do not drive more on average than single households. I suspect that one-parent family households have less disposable income after having paid for their accommodation. In contrast, couples with children drive much larger distances than single households. Surprisingly, for two-parent family households an additional child does not increase driving demand. This may be because the time available and the monetary budget are more restricted, and parents seem to drive all of their children to the same place at the same time. When considering fuel consumption, however, an additional child leads to increased demand. This may be because these households require larger cars that consume more fuel. Also, this data contains fuel expenditures for motorbikes, which may have been driven by children. Note that when considering driving demand per person, single households rank top, with a driving demand of 12,324 km.[109]

[109] This is the predicted value, given that the values of all other explanatory variables are at their mean. The mean value of the driving distance of all single households is lower, namely 8,008 km. This difference mainly arises due to the lower income of non-single households, namely CHF 55,549 versus CHF 90,890 and because single households are more likely than non-single households to be located in city areas (43% versus 28%). These values are all based on the micro-census dataset Mz05.

	Mz00	Mz05	EVE
	[1]	[2]	[3]
Constant	357.1 [1.28]	-789.2* [-2.52]	1006* [2.00]
Income	95.42*** [42.95]	99.20*** [53.68]	61.06*** [27.56]
Agglomeration	2677*** [13.61]	3908*** [23.06]	2733*** [11.77]
Town	2276** [3.07]	3650*** [4.14]	1599 [1.08]
Rural	5013*** [23.55]	6885*** [33.32]	4514*** [16.44]
Regional variables:			
Leman	1265.2*** [5.07]	717.0** [3.03]	3232*** [9.51]
Midland	417.0 [1.62]	820.0*** [3.80]	856.8** [2.73]
Northwest	374.1 [1.23]	37.72 [0.11]	-131.9 [-0.38]
East	370.7 [1.02]	900.4*** [3.34]	828.9* [2.27]
Central	674.8 [1.70]	430.1 [1.63]	132.2 [0.32]
Ticino	2639*** [5.46]	3470*** [8.40]	5554*** [13.70]
Couple	3073*** [13.61]	3192*** [16.46]	3270*** [12.05]
One-parent family	3091*** [6.58]	-470.3 [-0.42]	193.8 [0.11]
Two-parent family		5738*** [8.93]	2353* [2.52]
Flat share	3349 [1.36]	3017* [2.43]	-3990* [-2.27]
Subtenancy	-1716 [-0.22]	2508 [0.48]	--
Effect of additional person:			
One-parent family	219.1 [1.83]	967.6* [2.33]	507.9 [0.82]
Two-parent family		-21.63 [-0.14]	751.6** [3.25]
Flat share	-433.4 [-0.44]	186.5 [0.39]	2591*** [4.69]
Subtenancy	-415.9 [-0.11]	--	--
σ	10846	10431	10874
R^2 / pseudo R^2	0.1821	0.2690	0.1827

Note 1: Notation " $|\,\bar{x}$ " means that the values were computed conditional on the mean values of explanatory variables x.
Note 2: The benchmark household is a single household living in a city in the region of Zürich.
Note 3: For the EVE dataset, "flat share" corresponds to "other household types". This could also be the reason for the different result.
Note 4: The values in square brackets "[..]" denote t-values.
Note 5: The levels of significance are denoted by: *: p<0.05, **: p<0.01 and ***: p<0.001.
Note 6: For the Mz00 dataset, only information on the number of children was given. It was not clear whether the children lived with both parents or just one.

Table 2.1.2: Estimated parameters based on the OLS model.

With respect to spacial structure, first of all the coefficients reveal the obvious differences between the driving demand of city centre households and others. It comes as no surprise that households in rural areas drive more than urban households in agglomerations. I believe the reasons for the increased driving demand of households living outside city centres are their longer distances to work or to recreational facilities, and the lower service level of public transport in their area. This is also why the driving demand of households in the most densely populated area of Zurich – the benchmark region in the three models – is the lowest of all regions.[110] The areas that differ most to the benchmark region are the regions of "Leman" in the west and the "Ticino" in southern Switzerland. This difference arises because these areas consist of a high proportion of remote mountainous areas where the distances to facilities are long and the service level of public transport is low. Further, the results computed by EVE dataset, based on fuel consumption, show that fuel demand in these two regions is much higher. Again, this is due to the mountainous area of large parts of these regions, causing higher fuel consumption per kilometre.

Note that for the model based on the dataset of the EVE consumer survey, some dummy variables were added for particular months because all household diaries refer to a certain month. The dummies revealed the following interesting seasonal pattern: dummies for January and February show significantly negative values (-667 km, -1,330 km), presumably due to poor road conditions. For July, August and October they show significantly positive values (1,363 km, 985 km, 1,038 km) that can most likely be attributed to holidays and outdoor leisure activities. Note that December is the benchmark month. Also, dummy variables were used to control for years. None of these year-dummies were significant.

Next, I will present the results based on a limited number of explanatory variables. As mentioned in Subsection 1.5, the reason for doing this is that I wish to compare the results to those obtained using the data in Axhausen and Erath (2010), denoted by "Erath" in the following, as well as to those obtained using the MDCEV models. For this reason, I wish to use the same explanatory variables to make the corresponding parameter values comparable.

Since the dataset "Erath" in Axhausen and Erath (2010) has a panel structure, I will estimate not only a model that neglects this panel structure, but also one that takes it into account.[111] The OLS model that considers the panel structure is defined as follows:

[110] At the highest level of geographical distinction, the Swiss Federal Statistical Office has divided Switzerland into seven regions. For an exact definition of these regions, see Bundesamt für Statistik (2010e).

[111] The reason for this is that the Multiple Discrete-Continuous Extreme Value Model (MDCEV) ignores the panel structure. I therefore want to find out whether including the panel structure in the model would change the results. This would be the case if inclusion of the panel structure in the Tobit model led to differences in the results.

$$y_{jn} = \alpha + \beta \cdot x_{jn} + \eta_n + \varepsilon_{jn}, \tag{2.1.7}$$

where

$$\eta_n \sim iid\, N\left(0, \sigma_\eta\right), \tag{2.1.7a}$$

$$\varepsilon_{jn} \sim iid\, N\left(0, \sigma_\varepsilon\right) \text{ and independent of } \eta_n \text{, and} \tag{2.1.7b}$$

$$j = 1, \ldots, J \text{ and } n = 1, \ldots, N. \tag{2.1.7c}$$

Index n represents the household, with N being the total number of households. Index j denotes the fuel price level presented to households and J the total number of fuel levels presented to them. Parameter vector β captures weights relating to x_n, which stands for variables such as fuel prices, the marginal costs of driving and household properties. Since only the fuel price within variables x_i changes within the data of an individual household, it is not possible to estimate a model with fixed effects where a constant α_n would be estimated for each household. For this reason, only a random effect model could be estimated. Note that the panel models are discussed in further detail in Appendix A2.1.

Table 2.1.3 lists the whole results of the "small" model for all available data. The results of the regression model that includes the panel structure corresponding to (2.1.7) show that variance σ_ε^2 is much lower than variance σ_η^2. This implies that the hypothesis that all error terms η_n are zero can be rejected[112]. Although this implies that the model captures the panel structure, the values of the estimated parameters do not differ much. Note that by using the stated preference dataset of Erath, it was now possible to compute statistically significant parameters that account for the impact of fuel prices on driving demand. The elasticities of driving demand with respect to driving cost and fuel prices could then be computed from these parameters. These values are slightly lower than those determined by Baranzini et al. (2009), who conducted a survey based on time-series data for Switzerland.[113]

[112] The Breusch and Pagan Lagrangian multiplier test for random effects yielded a χ^2 value of $\chi^2(1) = 5813$. The null hypothesis $\eta_n = 0 \ \forall\, n = 1..N$ could therefore be rejected at every level, namely $p < 0.0001$.

[113] See Table 1.3.5.

Dataset	Mz00	Mz05	EVE	Erath	Erath
	OLS	OLS	OLS	OLS (pooled)	Panel (rand. eff.)
Number of models				(2)	(3)
(Constant) b_1	1790*** (8.24)	562** (3.04)	2779*** (6.15)	-2225* (752)	-2893* (949)
(Income) b_2	101.2*** (46.5)	0.106*** (58.6)	63.23*** (29.3)	0.1011 (0.00375)	0.1016* (0.00887)
(Driving costs) b_3	--	--	--	-6396* (1548)	-7768* (487.2)
(Fuel price) b_3	--	--	--	-497.8* (120)	-605.5* (38.45)
(Rural place of residence) b_4	3811*** (20.3)	4673*** (26.3)	2640*** (11.6)	529.2* (347)	347 (819)
(Number of people in household) b_5	1053*** (15.4)	1613*** (25.9)	1662*** (19.2)	1354.5* (155)	1360.2* (366)
σ, σ_ε	10960	10641	11094	8577	2673
σ_η	--	--	--	--	8185
R^2	0.164	0.239	0.148	0.279	0.279
$\varepsilon_{E(X_2),p_2}$	--	--	--	-0.245 (0.0593)	-0.298 (0.0187)
$\varepsilon_{E(X_2),p_{fuel}}$	--	--	--	-0.150 (0.0362)	-0.182 (0.0114)
$\varepsilon_{E(X_2),y}$	0.5950 (0.0133)	0.6139 (0.0110)	0.4795 (0.0167)	0.774 (0.0287)	0.778 (0.0679)
$\varepsilon_{E(Y),rural}$ at mean values $\overline{x}$	0.2891 (0.0144)	0.3364 (0.0129)	0.1908 (0.0165)	--	--

Table 2.1.3: Estimated parameters based on the OLS model,[114] "small" model.

In contrast, the income elasticity is virtually identical to the values generated by Baranzini et. al. (2009). The parameter accounting for the type of place of residence in a rural area is not significant. For this reason, I did not compute the corresponding elasticity. It is unclear to me why this parameter is insignificant. Although Axhausen and Erath (2010) claim that households were randomly drawn,[115] those sampled in rural areas seem to drive systematically less than those in the other datasets.[116] At

[114] Note that the panel model is estimated using a Maximum Likelihood estimation method.

[115] "The quotas are representative for the Swiss population over 18 years of age and living in a household with at least one car. With the exception of the car type market shares, ... , all values are derived from the Mikrozensus Verkehr Swiss Federal Statistical Office (2006a).", Axhausen and Erath (2010, 33). Further, Axhausen and Erath (2010) state on page 34: "Overall, the quotas were fulfilled satisfactorily and the sample can be considered representative."

[116] The average driving distances in the Erath dataset are 12,444, 15,344 and 14,328 km (rural, urban, total). Note that although the average driving distance of households resident in rural areas is lower than that in urban areas, the coefficient accounting for the impact of living in a rural area on driving demand is positive. This is mainly because the income of households living in rural areas is lower, namely CHF 85,997 versus CHF 90,513 and CHF 88,929 (rural, urban, total).

least the average driving distance of households in the Erath dataset is almost identically to the one of the other datasets.[117]

I have now commented all of the results yielded from OLS regression, with the exception of the model that captures the panel structure. As mentioned above, all of these coefficients, and therefore the results for elasticities, are biased because the OLS model is inadequate. Nevertheless, the result can still be used as a rough guide. In the next subsection, I shall present the results based on Tobit models.

2.2 The Tobit model

In this subchapter, I present the results obtained using a Tobit model. I will use the same datasets as in the previous subchapter. The Tobit model can cope with the fact that driving demand is non-negative and that there is a discrete probability that the driving distance is zero. I therefore consider the results generated using the Tobit model to be more realistic than those obtained by the OLS model. Since I have already addressed the Tobit model in Subchapter 1.5 and it is also covered by numerous textbooks, I shall not present it here again . The ML function corresponding to both the OLS and the Tobit model that includes the panel structure can be found in Appendix A2.1. This appendix also provides an exact definition of the elasticities $\text{mean}\left(\varepsilon_{E(Y),income}\right)$ and $\text{mean}\left(\varepsilon_{P(Y=0),income}\right)$ which are shown in Table 2.2.1 and derivations of the formulas I use to compute them for both the OLS and the Tobit model (see Formulas (A2.1.11) and (A2.1.12)).

Tables 2.2.1 to 2.2.3 show the elasticities and parameter values resulting from the Tobit model. First, I will show two tables containing the results of the “large” model that has a large number of control variables. I will then present a table with the results of the “small” model, which contains only the absolutely essential variables. Since the Erath dataset contains only a small number of explanatory variables, the results based on this dataset are only contained in the tables corresponding to the small model.

The results of the income elasticity of driving demand $\text{mean}\left(\varepsilon_{E(Y),income}\right)$ are very similar to those generated by the OLS model. For this reason, although the Tobit model captures the fact that the explained variable cannot be negative, this has little impact on the results. This could be because the proportion of carless households is rather low, namely 20%.

The results for the elasticity of being carless with respect to income are in a rather narrow bandwidth, namely between -0.77 and -0.67. This means that the proportion of carless households reacts rather

[117] The average driving distance of households in the Erath dataset is 14,328 km. The average driving distance given in dataset Mz05 is 13,890 km, see Table 3.2.3.

sensitively to income. Note that in both cases, elasticities $\text{mean}\left(\varepsilon_{E(Y),income}\right)$ and $\text{mean}\left(\varepsilon_{P(Y=0),income}\right)$ differ from the value of those which are computed at the mean $\bar{x}$. This difference is due to the fact that the marginal effects on which these elasticities are based depend non-linearly on the explanatory variable.[118]

	Mz00	Mz00	Mz05	Mz05	EVE	EVE
	OLS	Tobit	OLS	Tobit	OLS	Tobit
	[1]	[1']	[2]	[2']	[3]	[3']
$\varepsilon_{E(Y),income} \mid \bar{x}$	0.5612 (0.0135)	0.5687 (0.0135)	0.5727 (0.01108)	0.5904 (0.0115)	0.4631 (0.01711)	0.4659 (0.0171)
$\text{mean}\left(\varepsilon_{E(Y),income}\right)$	--	0.5692 (0.0135)		0.5737 (0.0115)	--	0.4636 (0.01708)
$\varepsilon_{P(Y=0),income} \mid \bar{x}$	--	-1.0122 (0.0252)	--	-1.1719 (0.0244)	--	-0.9215 (0.03487)
$\text{mean}\left(\varepsilon_{P(Y=0),income}\right)$	--	-0.7725 (0.0252)	--	-0.7593 (0.0244)	--	-0.6663 (0.03487)
$\left(\frac{E(Y)_{rural}}{E(Y)_{urban}} - 1\right) \mid \bar{x}$	0.4626 (0.01632)	0.5133 (0.0187)	0.7482 (0.0192)	0.8243 (0.0188)	0.4272 (0.0237)	0.4673 (0.0224)
$\left(\frac{P(Y=0)_{rural}}{P(Y=0)_{urban}} - 1\right) \mid \bar{x}$	--	-0.4748 (0.0258)	--	-0.6353 (0.0273)	--	-0.4891 (0.03549)
$P(Y=0)$ $\text{mean}(y=0)$ $\ln\left(P(Y=0)/\text{mean}(y=0)\right)$	--	0.19531 0.21108 0.0808	--	0.18898 0.20036 0.0602	--	0.18211 0.18561 0.01921
$E(Y)$ $\text{mean}(y)$ $\ln\left(E(Y)/\text{mean}(y)\right)$	--	13183 13270 0.00665	--	13890 13908 0.00128	--	13835 13962 0.00921
R^2	0.18	0.0123	0.2690	0.0191	0.1827	0.0122

Note 1: The notation " $\mid \bar{x}$ " means that the values were computed conditional on the mean values of explanatory variables x.
Note 2: The standard deviations of the relative effects when households move from urban to rural areas refer to the changes in log values.

Table 2.2.1: Effects of income and type of place of residence based on the large OLS and Tobit model.

The predicted relative changes in driving distance when households move from urban to rural areas according to the Tobit model are all almost exactly by a factor of 1.10 greater than those predicted by the OLS model. I believe that the values of the Tobit model are more realistic because this model also captures the effect of not owning a car when the preference for driving is low. Table 2.2.1 also shows that the Tobit model replicates the average driving distance very well, whereas the replication error of the proportion of car-less households is greater, but still below 10%. All parameters are listed in Table 2.2.2.

---

[118] For a discussion on this difference, see Appendix A2.2.

	Mz00	Mz00	Mz05	Mz05	EVE	EVE
	OLS	Tobit	OLS	Tobit	OLS	Tobit
	[1]	[1']	[2]	[2']	[3]	[3']
Constant	357.1 [1.28)	-4962*** (-14.25)	-789.2* (-2.52)	-6589*** (-17.19)	1006* (2.00)	-4017 (605.9)
Income	95.42*** (42.95)	115.3*** (43.01)	0.0992*** (53.68)	0.116*** (52.77)	61.06*** (27.56)	71.44*** (27.57)
Type of place of residence:						
Agglomeration	2677*** (13.61)	3549*** (14.88)	3908*** (23.06)	5435*** (26.44)	2733*** (11.77)	3910*** (14.18)
Town	2276** (3.07)	3074*** (3.43)	3650*** (4.14)	4680*** (4.42)	1599 (1.08)	2756 (1.59)
Rural	5013*** (23.55)	6111*** (23.72)	6885*** (33.32)	8747*** (35.21)	4514*** (16.44)	5822*** (18.00)
Regional variables:						
Lemand	1265.2*** (249.5)	1561.3*** (303.7)	717.0*** (236.6)	1025*** (286.3)	3232*** (339.9)	4051*** (400.9)
Midland	417.0 (1.62)	570.0 (1.82)	820.0*** (3.80)	1117*** (4.29)	856.8** (2.73)	1301*** (3.50)
Northwest	374.1 (1.23)	594.4 (1.61)	37.72 (0.11)	16.24 (0.04)	-131.9 (-0.38)	75.50 (0.18)
East	370.7 (1.02)	639.9 (1.45)	900.4*** (3.34)	1364*** (4.22)	828.9* (2.27)	1458*** (3.38)
Central	674.8 (1.70)	866.1 (1.80)	430.1 (1.63)	582.4 (1.83)	132.2 (0.32)	541.2 (1.12)
Ticino	2639*** (5.46)	3261*** (5.58)	3470*** (8.40)	4509*** (9.14)	5554*** (13.70)	6801*** (14.33)
Type of household:						
Couple	3073*** (13.61)	5065*** (18.33)	3192*** (16.46)	5451*** (23.20)	3270*** (12.05)	4865*** (15.09)
One-parent family	3091*** (6.58)	4676*** (8.26)	-470.3 (-0.42)	167.9 (0.12)	193.8 (0.11)	-314.4 (-0.16)
Two-parent family	--	--	5738*** (8.93)	8256*** (10.92)	2353* (2.52)	3866*** (3.57)
Flat share	3349 (1.36)	4921 (1.58)	3017* (2.43)	4074** (2.75)	-3990* (-2.27)	-3397 (-1.65)
Subtenancy	-1716 (-0.22)	-4376 (-0.45)	2508 (0.48)	2747 (0.43)	--	--
Effect of additional person :						
One-parent family	219.1 (1.83)	298.8* (2.08)	967.6* (2.33)	1263* (2.55)	507.9 (0.82)	1027
Two-parent family	--	--	-21.63 (-0.14)	-69.30 (-0.38)	751.6** (3.25)	809.6** (3.04)
Flat share	-433.4 (-0.44)	-746.2 (-0.60)	186.5 (0.39)	232.5 (0.41)	2591*** (4.69)	2832*** (4.41)
Subtenancy	-415.9 (-0.11)	922.4 (0.19)	--	--	--	--
σ	10846	12792*** (162.9)	10431	12177*** (178.0)	10874	12454*** (140.1)
R^2 / pseudo R^2	0.1821	0.0123	0.2690	0.0191	0.1827	0.01

Note 1: The notation "$\mid \bar{x}$" means that the values were computed conditional on the mean values of explanatory variables x.
Note 2: The benchmark household is a single household living in a city in the region of Zürich.
Note 3: For the EVE dataset, "flat share" corresponds to "other household types". This could also be the reason for the different result.
Note 4: The levels of significance are denoted by: *: $p<0.05$, **: $p<0.01$ and ***: $p<0.001$.

Table 2.2.2: Estimated parameters based on the OLS model.

The estimated parameters of all models show a similar pattern. The parameters of the Tobit model are usually approximately 1.3 times greater than those of the OLS model. The relative marginal effects of the parameters on the models' outcomes are therefore roughly the same in the OLS and Tobit models.

I would now like to present and discuss the results of the small model. The following table will also contain the results for the "Erath" dataset in Axhausen and Erath (2010). Since Axhausen and Erath varied fuel prices sufficiently, significant parameter values corresponding to fuel prices can be computed. Again, as in the case of the OLS model in Subchapter 2.1, I computed the Tobit model once including the panel structure and once neglecting it ("pooled model").

The Tobit model that includes the panel structure is defined as follows:

$$y^*_{jn} = \alpha + \beta \cdot x_{jn} + \eta_n + \varepsilon_{jn}, \tag{2.2.1}$$

$$y_{jn} = \left\{ \begin{array}{l} y^*_{jn} < 0 : y_{jn} = 0 \\ y^*_{jn} \geq 0 : y_{jn} = y^*_{jn} \end{array} \right\} \tag{2.2.1a}$$

where

$$\eta_n \sim iid\, N\left(0, \sigma_\eta\right), \tag{2.2.1b}$$

$\varepsilon_{jn} \sim iid\, N\left(0, \sigma_\varepsilon\right)$ and independent of η_n and (2.2.1c)

$j = 1,\ldots,J$ and $n = 1,\ldots,N$. (2.2.1d)

The parameters and indices are defined as in the case of the OLS model presented in Subchapter 2.1, namely: index n indicates the household, N is the total number of households, j denotes the fuel price level presented to the household and J the total number of fuel levels presented to them. Parameter vector β captures weights relating to x_n, which stands for variables such as fuel prices, the marginal costs of driving and household properties. As with the OLS, it is not possible to estimate a fixed effects model where a constant α_n would be estimated for each household, since of variables x_i only fuel prices change within the data of an individual household. For this reason, only a random effect model could be estimated. Note that the panel models are discussed in further detail in Appendix A2.1. This appendix also explains how I computed the marginal effects used to calculate the elasticities. Note in particular how the elasticities were determined in the case of the panel model used to compute the Erath data.

Table 2.2.3 lists all of the results generated by the "small" model for all available data.

Dataset	Mz00	Mz00	Mz05	Mz05	EVE	EVE	Erath	Erath	Erath	Erath
Model type	OLS	Tobit	OLS	Tobit	OLS	Tobit	OLS	OLS panel	Tobit	Tobit panel
_cons	1790*** (217)	-2776*** (272)	562** (185)	-4098*** (230)	2779*** (452)	-1060* (538)	-2225* (752)	-2893* (949)	-1693* (945)	-1856* (519)
inc	101.2*** (2.18)	124.3*** (2.636)	0.106*** (0.0018)	0.127*** (0.0022)	63.23*** (2.158)	75.57*** (2.535)	0.1011 (0.0038)	0.1016* (0.0089)	0.1228* (0.0046)	0.1250* (0.0089)
driving costs							-6396* (1548)	-7768* (487.2)	-7835* (1931)	-9917* (604)
fuel price							-497.8* (120)	-605.5* (38.45)	-609.8* (150.4)	-771.8 (47.01)
rural	3811*** (188)	4563*** (226)	4673*** (177)	5653*** (212)	2640*** (227.6)	3230*** (266.9)	529.2* (347)	347 (819)	1566* (430)	2112* (750)
nof_pers	1053*** (68.2)	1462*** (82.2)	1613*** (62.3)	2153*** (74)	1662*** (86.6)	2089*** (100.9)	1354.5* (155)	1360.2* (366)	1796* (190)	2192* (345)
σ	10960	12960*** (79.7)	10641	12478*** (70.3)	11094	12742*** (139.8)	8577	2673	10347 (157.8)	2995 (48.18)
σ_η	--	--	--	--	--	--	--	8185	--	10876 (349.2)
$\varepsilon_{E(Y),income} \mid \bar{x}$	--	0.6080 (0.0132)	--	0.5904 (0.0115)	--	0.4864 (0.0165)	--	--	--	0.802 (0.0581)
$\text{mean}\left(\varepsilon_{E(Y),income}\right)$	0.5950 (0.0133)	0.6136 (0.0132)	0.6139 (0.011)	0.5737 (0.0115)	0.4795 (0.0167)	0.4905 (0.0165)	0.774 (0.0287)	0.778 (0.0679)	0.790 (0.0374)	0.786 (0.0581)
$\varepsilon_{P(Y=0),income} \mid \bar{x}$	--	-1.0717 (0.0247)	--	-1.1719 (0.0244)	--	-0.9427 (0.0331)	--	--	--	--
$\text{mean}\left(\varepsilon_{P(y=0),income}\right)$	--	-0.7725 (0.0252)	--	-0.7593 (0.0244)	--	-0.7118 (0.0331)	--	--	-0.913 (0.0435)	-0.950 (0.1109)
$\left(\frac{E(Y)_{rural}}{E(Y)_{urban}} - 1\right) \mid \bar{x}$	0.3136 (0.0169)	0.3016 (0.0157)	0.3643 (0.0151)	0.4501 (0.0140)	0.2220 (0.0183)	0.2268 (0.01749)	--	--	--	0.183 (0.0617)
$\varepsilon_{E(X_2),p_2}$	--	--	--	--	--	--	-0.245 (0.0593)	-0.298 (0.0187)	-0.231 (0.0765)	-0.379 (0.0204)
$\varepsilon_{E(X_2),p_{fuel}}$	--	--	--	--	--	--	-0.150 (0.0362)	-0.182 (0.0114)	-0.140 (0.0467)	-0.171 (0.0124)
$\varepsilon_{P(X_2=0),p_2}$	--	--	--	--	--	--	--	--	0.378 (0.8175)	0.472 (0.0391)
$\varepsilon_{P(X_2=0),p_{fuel}}$	--	--	--	--	--	--	--	--	0.233 (0.0499)	0.290 (0.023)
$\left(\frac{P(Y=0)_{rural}}{P(Y=0)_{urban}} - 1\right) \mid \bar{x}$	--	-0.4747 (0.0258)	--	-0.4783 (0.02362)	--	-0.310 (0.0298)	--	--	--	-0.215 (0.0826)
$P(Y=0)$		0.1953		0.1890		0.1821			0.1953	
$\text{mean}(y=0)$	--	0.2098	--	0.1969	--	0.1837	--	--	0.2111	--
$\ln\left(\frac{P(Y=0)}{\text{mean}(y=0)}\right)$		0.0740		0.0421		0.0090			0.0808	
$E(Y)$		13183		13890		13835			13183	
$\text{mean}(y)$	--	13261	--	13889	--	13957	--	--	13270	--
$\ln\left(\frac{E(Y)}{\text{mean}(y)}\right)$		0.0059		0.0014		0.0088			0.0067	
R^2 / pseudo R^2	0.164	0.011	0.239	0.016	0.148	0.009	0.279	0.280	0.020	

Note 1: The notation "$\mid \bar{x}$" means that the values were computed conditional on the mean values of explanatory variables x.
Note 2: The benchmark household is a single household living in a city in the region of Zürich.
Note 3: For the EVE dataset, "flat share" corresponds to "other household types". This could also be the reason for the different result.
Note 4: The values in parentheses "(..)" denote standard deviations.
Note 5: The levels of significance are denoted by: *: p<0.05, **: p<0.01 and ***: p<0.001.

Table 2.2.3: Estimated parameters based on the OLS model.

The results in Table 2.2.3 show that the income elasticity of driving demand does not vary considerably between the different model types OLS and Tobit. The values based on the Erath dataset are approximately 25% higher than the results based on the other datasets. They are therefore higher than the long-term elasticity determined by Baranzini et al. (2009) of 0.629, but in the range established by Goodwin et al. (2004) as the average of international studies, namely 1.08. The results for the elasticity of the probability of being car-less with respect to income $\varepsilon_{P(Y=0),income}$ are also very similar to each other and, again, the value yielded using the Erath data is around 25% higher in absolute terms. If the elasticity of the probability of owning a car is computed, the value would be around four times[119] smaller, namely about 0.21. This value is considerably lower than the income effect on car stock determined in international studies, e.g. the average value of 0.81 yielded by Goodwin et al. (2004) (see Table 1.3.1). Nonetheless, the Tobit model does not capture the fact that households may own more than one car. Moreover, it could be the case that households considering whether or not to purchase an additional car may react more sensibly to income, since ownership of a second or third car suggests possession of a luxury good. For this reason, the values generated by this model – in absolute terms – are rather a lower bound of the true values. Further, in the case of Switzerland, there may be a kind of satiation effect because there is only a small proportion of households in Switzerland with an income that does not enable them to own a car.[120] Despite this fact, a value of 0.21 seems too low. The same holds for the elasticity of the probability of being car-less with respect to fuel price $\varepsilon_{P(Y=0),p_2}$. Dividing the corresponding value by four to obtain the elasticity with respect to the probability of owning at least one car yields a value of approximately 0.06. In absolute terms, this is also well below the value of -0.25 established by Goodwin et al. (2004), listed in Table 1.3.1.

I will now comment on the estimated fuel price elasticity of driving demand. The values I computed are very similar for both the OLS and the Tobit model that does or does not capture the panel structure of the data. The values are very close to the long-term elasticity (-0.202) yielded by Baranzini et al. (2009). Compared to the average of international studies (-0.29) presented in Goodwin et al. (2004), however, the values are rather low.[121]

---

[119] From the definition of $\varepsilon_{P(Y>0),p_2}$ it follows that

$$\varepsilon_{P(Y>0),p_2} = \frac{\partial P(Y>0)}{\partial p_2} \cdot \frac{p_2}{P(Y>0)} = -\frac{\partial P(Y=0)}{\partial p_2} \cdot \frac{p_2}{P(Y>0)} = -\frac{\partial P(Y=0)}{\partial p_2} \cdot \frac{p_2}{P(Y=0)} \cdot \frac{P(Y=0)}{P(Y>0)} = -\frac{P(Y=0)}{P(Y>0)} \cdot \varepsilon_{P(Y>0),p_2} .$$

Since $P(Y=0) \approx 0.2$, it follows that $\varepsilon_{P(Y>0),p_2} \approx -0.25 \cdot \varepsilon_{P(Y>0),p_2}$.

[120] For a model that captures a satiation effect, see Dargay et al. (2007).

[121] For an overview of the results of Swiss and international studies and a discussion on differences in elasticities, see Chapter 1.

I would now like to comment on the effects yielded when households move from urban to rural areas. With respect to the effect on driving demand, the values computed by the OLS and the Tobit model do not differ considerably. Similar to the results of the EVE dataset, those based on the Erath dataset are much smaller than those yielded from datasets Mz00 and Mz05. I have already discussed possible reasons for this in Subchapter 2.1. Further, the results show that the standard errors of random variables η_n are much larger than those of the random term ε_{jn}. For this reason, the household-specific random term dominates the error structure, and the model that includes the panel structure is presumably the model that better reflects reality. Finally, all of the small models, including that based on the Erath dataset, replicate the mean aggregate driving distance and the proportion of car-less households quite well. Note that a more in-depth discussion on how well the Tobit model reflects the data can be found in Appendix A3.15.

3. The Multiple Discrete-Continuous Extreme Value Model (MDCEV)

3.1 Introduction

As discussed in Chapter 1, the OLS and Tobit models used to estimate household demand for annual travel distance cannot provide significant coefficients for fuel prices when using cross- sectional data. The model presented here is able to overcome this problem. The key to obtaining elasticities for fuel price is that this model is based on a micro-economic model framework. In its general form, this model can capture several car types and their annual driven distances. Each distance driven by a certain car type would then be treated as a separate good that provides a utility. If the household does not own a certain car type, the corresponding distance is simply zero.

The model framework I will use is based on Bhat (2005) and Bhat (2006).[122] Bhat (2006) developed the Multiple Discrete-Continuous Extreme Value Model (MDCEV) to study car type choice and use. He assumed that the total driving distance is given for each household, which is equal to the sum of kilometers driven by the vehicles declared in a survey by households. Bhat further assumed that households are not restricted by their budget when deciding to own one or several cars. In other words, this model captures only the households' preference for car types, but not their economic behavior. For instance, Bhat's model does not capture the fact that it is economically irrational for households to own a vast number of cars. This is because the model neglects the fact that car ownership involves fixed costs that reduce available income and therefore the achievable utility level. Further, Bhat (2006) failed to explain the total driving demand of households. The purpose of extending Bhat's model is therefore to include the effect of fixed costs on the economic behavior of households and to explain the total annual driving distance. Since the complexity of the model increases dramatically if fixed costs are included in the decision-making process, I merely developed the case where households can choose only whether or not to own and use one car type. Thus, the model I shall present is an extension of the Tobit model, as presented in Chapter 2. In contrast to the Tobit model, the extended MDCEV model enables us to compute the impact of taxes on car ownership and the impact of fuel on the proportion of carless households and driving demand.

In the following, the foundations of the models are presented in Subchapter 3.1. In Subchapter 3.2, the simplest model, according to which households can choose between owning a car and choosing the driving distance or not owning a car, is derived for the case where fixed costs of car ownership are neglected. In Subchapter 3.3, the model is extended to the case in which owning a car implies fixed costs.

[122] Bhat (2005) analysed the time spent on tourist activities. Using the same model framework, Bhat (2006) analysed car type choice and car use. I will therefore present an outline of Bhat (2006).

This subchapter is structured as follows. First, the basic principle of the model framework is presented. Second, the micro-economic optimization problem, where households may choose between several cars, is stated in a general form. Third, an illustration of the two-good case is presented. Fourth, I explore the problem that researchers are unable to observe household preferences, and describe how this problem is captured by the model. Finally, I shall present and discuss the distinct utility function used in this framework.

The basic principle of the modelling framework

The model presented in the following describes the micro-economic decision of a household with respect to car ownership and use. In its general form, it is assumed that households can choose one or several cars from a set of cars. The choice is restricted only to the extent that households may not choose two cars of the same type. It is assumed that the choice between one and several car types and the choice of distance driven in those particular cars are simultaneous. Households can also decide against owning a car. Households are assumed to maximize utility, subject to a budget constraint. The utility function values the utility households yield by driving cars of different types and from a composite good. The composite good includes all goods apart from driving cars: housing, health, food, insurance, etc. Households can choose zero, one or several cars from each set of different car types. Each car type can only be chosen once. The framework considers the decision for one period. Decision-makers can optimize each single period and the outcome. It is assumed that decisions are independent of decisions made in any other period. That means that decision-makers have no switching costs when deciding whether or not to own one or more cars: the model treats car ownership as though the decision-maker would be able to rent a car type period by period. Further, it is assumed that there is no habit persistence. The budget constraint contains expenditures for driving one or several cars, namely the number of kilometers multiplied by the cost per kilometer driven by a specific car type. With the exception of simplified cases, where fixed costs are neglected, the budget constraint also contains the fixed costs of owning one or more cars. The remaining expenditure is spent on the composite good. Including fixed costs of car ownership enables a realistic description to be given of household behavior with respect to deciding whether or not to own a car. This decision is particularly relevant for low-income households. It is assumed that households are in possession of perfect information. This means that household know their precise preferences and are informed about the features of all of the car types from which they can choose. In contrast, from the researcher's perspective, households' utility functions are not known exactly, which is why they are stochastic functions. For empirical research, parametrized utility functions will be used. Some parameters are stochastic, accounting for the fact that the utility functions are stochastic. Some of the parameters depend on household and car characteristics. These parameters will be estimated by the Maximum

Likelihood estimation (MLE). In order to obtain a simple formula for the Maximum Likelihood (ML) function, the utility function must be of a certain type, and certain assumptions on the distribution of the stochastic term are necessary. The concrete utility function used and the assumptions on the parameters are presented in the final part of this section.

The household optimization problem

I shall now describe the problem solved by a household. The household is considered to behave as though it had maximized a utility function

$$\max_x u(x), \tag{3.1.1}$$

with $\lim_{x_1 \to 0} \partial u(x) / \partial x_1 = \infty$, (3.1.1a)

subject to budget restriction

$$y \geq p_1 x_1 + \sum^J p_i x_i + \sum^J I_{x_i > 0} \cdot k_i \text{ and } x_i \geq 0 \ \forall i = 1..J. \tag{3.1.2}$$

The composite good is denoted by x_1 and *y* denotes the households budget.[123] Condition (3.1.1a)[124] ensures that x_1 always has to be positive, since it contains also essential goods such as food and housing. Index *I* = 2..*J* is an index for car types. The annual distance in kilometers driven using car type *i* is denoted by x_1. Vector *x* contains all x_1. Price p_i denotes the per-kilometer costs of driving car *i*. The costs per kilometer consist of the fuel costs and the depreciation caused by driving the car, e.g. the wear and tear of the mechanical components and tyres. Fixed costs k_i refer to the annual fixed costs of owning car type *i*. These fixed costs consist of parking costs, insurance costs, taxes and depreciation. In this context, depreciation captures only the loss of value caused by factors unrelated to the use of cars, such as rusting, and loss in value due to technical obsolescence. It is assumed that if a household owns a car, the household will also use this car and therefore the annual distance x_i is assumed to be strictly positive. When a household decides not to own a certain car type *i*, the corresponding distance x_1 is zero. Ownership of car type *i* is therefore equivalent to a positive value x_i. Since fixed costs only arise if a car is owned, an indicator is needed when totalling the fixed costs of each car type. Indicator $I_{x_i > 0}$ is one if x_1 is greater than zero, and zero otherwise.

The optimization problem stated above differs from the standard problem, as described in many textbooks, where the budget restriction is linear in all x_i. The difference arises because the budget

---

[123] Note that in Chapter 3 the notation is different from that in Chapter 2 where *y* denotes the driving demand and the x_i denotes explanatory variables.

[124] This is called the INADA condition.

restriction is now non-linear in $x_i \ \forall i = 2..N$, due to indicator functions $I_{x_i>0}$. Since indicator functions $I_{x_i>0}$ are not differentiable at $x_i = 0$, the Kuhn-Tucker optimization method is not directly applicable.[125]

The problem therefore has to be restated as follows: assume that the household first chooses zero, one or several cars out of set $S_c = \{2,3,..,J\}$ of cars plus always the composite good with index one. Note that the numbers in set S_c indicate the different car types. Each car in set S_c may only be chosen once. Let the set of cars chosen by the household be denoted by $S \in \mathrm{Pot}(\mathrm{S_c})$. Note that the power set of S_C contains all possible combinations of car types from which a household can choose. The total number of such combinations is 2^{J-1}. The household then maximizes the utility conditional on choice set *S*, yielding the optimum consumption bundle x_S^* and the corresponding utility level $u_S = u(x_S^*)$. This means households choose all distances driven by the cars in set *S* and the amount of the composite good optimally. The budget available is reduced by the fixed costs of the cars in set *S*. This maximization problem can be stated as follows:

$$\max_{x_i,\, i\in S} u(x) \text{, with } x_j = 0,\ j \in S_c \setminus S, \tag{3.1.3}$$

subject to the budget constraint:

$$y \ge h(x) \text{, with } h(x) = p_1 x_1 + \sum_{i\in S} p_i x_i + \sum_{i\in S} k_i - y \text{ and } x_i \ge 0 \ \forall i \in S. \tag{3.1.4}$$

This maximization problem can now be solved by setting up the following Lagrangian for each choice set *S*:[126]

$$L = u(x) - \sum_{i\in S} \mu_i x_i - \lambda h(x) \text{, with } x_j = 0,\ j \in S_c \setminus S. \tag{3.1.5}$$

The corresponding conditions of the Kuhn-Tucker theorem are now

$$\frac{\partial u(x_S^*)}{\partial x_i} = \mu_i + \lambda p_i,\ i \in S \text{, and } \frac{\partial u(x_S^*)}{\partial x_1} = \lambda p_1,\ i \in S \tag{3.1.6}$$

with the corresponding complementary slackness conditions:

$$\mu_j > 0 \text{, if } x_j = 0, \tag{3.1.7a}$$

$$\mu_j = 0 \text{, if } x_j > 0, \tag{3.1.7b}$$

where $j \in S$.

---

[125] See Varian (1992: 503).

[126] Note that due to the INADA condition (3.1.1a) the term $-\mu_1 \cdot x_1$ does not need to be added to the Lagrangian (3.1.5).

Note that $h(x)$ depends linearly on $x_i, i \in S$, and in this case it is a convex function. The same holds for the non-negativity constraints $-x_j \le 0,\ j \in S \cup \{1\}$. The is therefore a unique solution x_S^* such that the Kuhn-Tucker conditions hold. This means that both the Kuhn-Tucker theorem and the Kuhn-Tucker sufficiency theorem hold.[127] Households then choose the set that yields the highest utility:

$$x_{\hat{S}}^*, \text{ with } \hat{S} = \arg \max_{S \in Pot(S_c)} u_S \tag{3.1.8}$$

Note that corner solutions are principally possible, and for this case one or several distances driven would be zero. But if some values *x* are zero, the solution would correspond to another set $S' \subset S$. Note that set S' implies that the fixed costs corresponding to boundary solutions $x_i = 0$ do not arise. The total fixed costs of S' are therefore lower than those of choice set *S*, $\sum_{i \in S'} k_i < \sum_{i \in S} k_i$, and the utility that can be yielded by S' is always greater than the utility yielded by *S*, $u_S > u_{S'}$, meaning that households would never choose *S*.

The fact that in the case of boundary solutions *S* would always be dominated by S' can be illustrated in the simplest way in the two-good case: assume that a household chooses to own a car and to bear the fixed costs of ownership, $S = \{2\}$. Let the household's preferences be such that utility maximization would cause the household not to drive the car, $x_2 = 0$. The whole budget remaining, after subtracting the fixed costs, would therefore be spent on good one $x_1 = (y - k_2)/p_1$. Since $x_2 = 0$, the corresponding choice set S' is $S' = \{\ \}$. In choice set S' only good one is consumed, and thus there are no fixed costs. The entire budget is therefore spent on good one: $x_1 = y/p_1$. Since marginal utility is positive in all goods, maximal utility of choice set S' is greater than that of choice set *S*, $u(y/p_1, 0) > u((y - k_2)/p_1, 0)$, and solution $x_1 = y/p_1,\ x_2 = 0$ is optimal.

Illustration of the maximization principle in the case of two goods

In the following, I present the formal calculation of the optimum consumption level of the model in the simplest case, where households can only choose between having one car and having no car. Further, this optimization is illustrated in several diagrams. Later, it will be shown how changes in prices, income and preferences affect optimum consumption levels. In this case, the possible sets *S* are: $S \in \{\{\ \}, \{2\}\}$. For the two-good case, therefore, the utilities of two choice sets, choice set one, $S_1 = \{\ \}$, and choice set two, $S_2 = \{2\}$, must be computed:

[127] See Varian (1992: 503).

Choice set 1: $S_1 = \{ \ \}$

In this case, the households will spend all income for good one: $x_1 = y/p_1$. The maximal utility level for this choice set is therefore:

$$u_{S_1} = u\left(y/p_1, 0\right). \tag{3.1.9}$$

Choice set 2: $S_2 = \{2\}$

In this case the Lagrangian (3.1.5) reduces to

$$L = u(x) + \mu_2 x_2 - \lambda\left(p_1 x_1 + p_2 x_2 - y\right). \tag{3.1.10}$$

This yields maximization conditions

$$\frac{\partial u\left(x_{S_2}^*\right)}{\partial x_1} = p_1 \lambda \text{ and} \tag{3.1.11a}$$

$$\frac{\partial u\left(x_{S_2}^*\right)}{\partial x_2} = -\mu_2 + p_2 \lambda, \tag{3.1.11b}$$

and two possible combinations of complementary slackness conditions:

$$\mu_2 = 0, \ x_2 > 0, \ \lambda > 0, \ p_1 x_1 + p_2 x_2 + k_2 - y = 0 \text{ or} \tag{3.1.12}$$

$$\mu_2 > 0, \ x_2 = 0, \ \lambda > 0, \ p_1 x_1 + p_2 x_2 + k_2 - y = 0. \tag{3.1.13}$$

Recall that setting $\lambda = 0$ would violate both maximization condition (3.1.12) and (3.1.13), as shown above. Of these two combinations, complementary slackness condition (3.1.13) shall first be examined. This condition implies that all income net fixed costs are spent for good one $x_1 = \left(y - k_2\right)/p_1$. This implies that, since μ_2 is greater than zero, the following relation must be met:

$$\left[\frac{\partial u\left(x_1, x_2\right)}{\partial x_2}\right]_{\substack{x_1=(y-k_2)/p_1, \\ x_2=0}} < \frac{p_2}{p_1} \cdot \left[\frac{\partial u\left(x_1, x_2\right)}{\partial x_1}\right]_{\substack{x_1=(y-k_2)/p_1, \\ x_2=0}}. \tag{3.1.14}$$

Whether or not this condition is fulfilled depends on the functional form of the utility function. A utility function with the following isoquant would satisfy this condition (3.1.14):

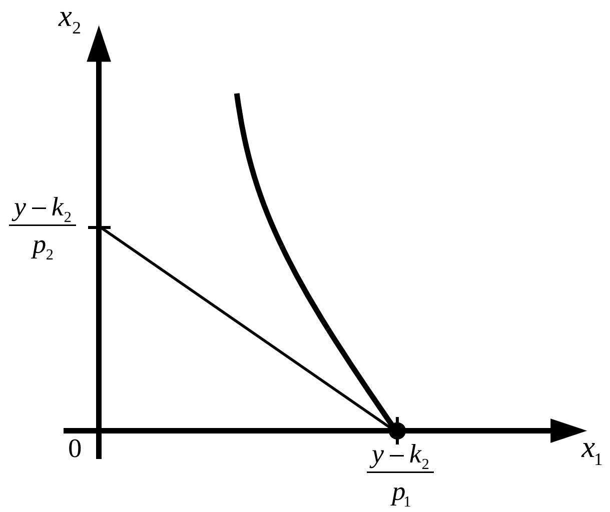


Figure 3.1.1: Isoquant of utility function when owning a car but not driving it is optimal.

Since in this case the household would not drive the car, the choice set corresponds to $S_1 = \{\ \}$. Since there are no fixed costs for case $S_1 = \{\ \}$, however, case $S_1 = \{\ \}$ would provide higher utility because the entire budget could be spent on good x_1.

If, in contrast, the utility function is such that it does not meet condition (3.1.14), it will then meet condition (3.1.12), from which it follows that

$$\left[\frac{\partial u(x_1, x_2)}{\partial x_2}\right]_{\substack{x_1 = x^*_{1,S_2} \\ x_2 = x^*_{2,S_2}}} = \frac{p_2}{p_1} \cdot \left[\frac{\partial u(x_1, x_2)}{\partial x_1}\right]_{\substack{x_1 = x^*_{1,S_2} \\ x_2 = x^*_{2,S_2}}} \text{, where } 0 < x^*_{1,S_2} < \frac{y-k_2}{p_1} \text{ and } x^*_{2,S_2} = \frac{y-k_2-p_1 \cdot x^*_1}{p_1}.$$

This is illustrated in the following figure:

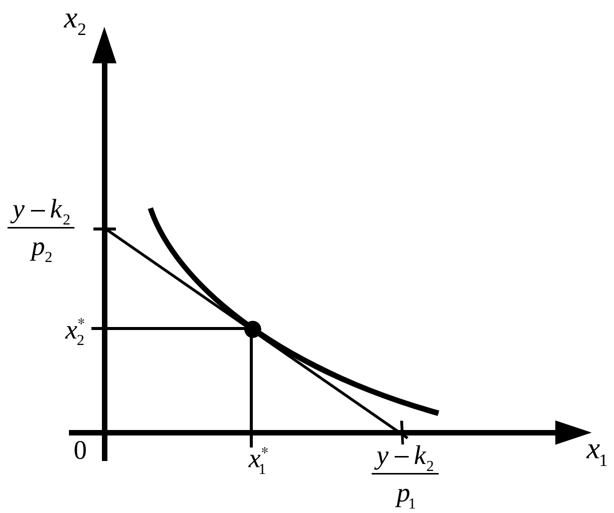


Figure 3.1.2: Isoquant of utility function when owning a car and driving it is optimal.

For this case, the utility is

$$u_{S_2} = u\left(x^*_{1,S_2}, x^*_{2,S_2}\right). \tag{3.1.15}$$

The household will now compare the utility levels of the two cases and choose to consume the following amounts of the composite good and car driving:

$$\left(x_1, x_2\right) = \begin{Bmatrix} \left(y_1/p_1, 0\right), if\ u_{S_1} > u_{S_2} \\ \left(x^*_{1,S_2}, x^*_{2,S_2}\right), if\ u_{S_1} \leq u_{S_2} \end{Bmatrix}. \tag{3.1.16}$$

Recall that if a boundary solution were optimal for choice set $S_2 = \{2\}$, the household would always choose S_1, which yields optimal consumption levels $\left(x_1, x_2\right) = \left(y/p_1, 0\right)$, since $u\left(y/p_1, 0\right) > u\left(\left(y-k_2\right)/p_1, 0\right)$.

The decision described above can also be illustrated graphically. The following figure shows an example where the utility function is such that the household would choose to own a car.

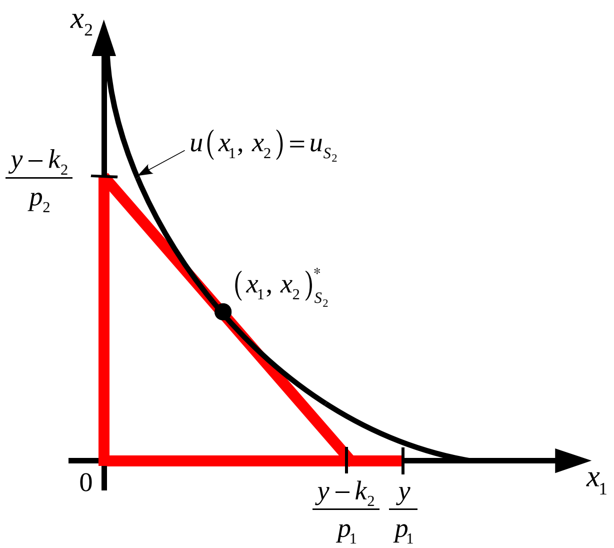


Figure 3.1.3: Optimum decision of a household that chooses to own a car.

The thick solid lines show the boundaries of the budget set that contains all feasible consumption bundles $\left(x_1, x_2\right)$. The triangular set represents the budget set corresponding to choice set S_2; the line between the origin and point $\left(x_1, x_2\right) = \left(y_1/p_1, 0\right)$ depicts the budget set corresponding to choice set S_1. The solution $\left(x_1, x_2\right) = \left(y_1/p_1, 0\right)$ of choice set S_1 is below the isoquant of the utility function, which means that it yields a lower utility. This is why the consumer chooses the optimum of set $S_2 = \{2\}$, $\left(x_1, x_2\right)^*_{S_2}$.

The following figure illustrates the case in which the utility function is such that solution $(x_1, x_2) = (y_1/p_1, 0)$ yields the highest utility:

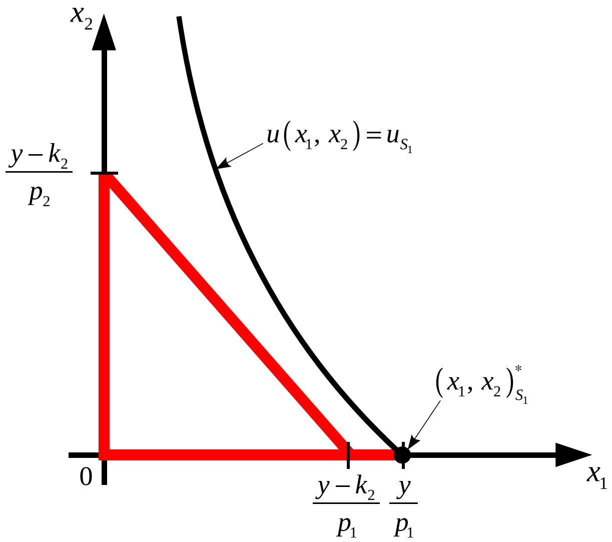


Figure 3.1.4: Optimum decision of a household that chooses not to own a car.

Note that the slope of the indifference curve at $x_1 = 0$ is infinite,[128] whilst the slope of the indifference curve at $x_2 = 0$ is less than zero.[129]

Note also that the budget set indicated by the solid red lines is not convex, and therefore several local optima may exist,[130] as the following figure shows:

---

[128] Proof: $\lim_{x_1\to0} \frac{dx_2}{dx_1} = \lim_{x_1\to0} -\frac{\partial u(x_1,x_2)/\partial x_1}{\partial u(x_1,x_2)/\partial x_2} = -\infty$, since $\lim_{x_1\to0}\left[\partial u(x_1,x_2)/\partial x_1\right] = \infty$ and $\lim_{x_1\to0} \partial u(x_1,x_2)/\partial x_2$ is finite and positive.

[129] $\lim_{x_2\to0} \frac{dx_2}{dx_1} = \lim_{x_2\to0} -\frac{\partial u(x_1,x_2)/\partial x_1}{\partial u(x_1,x_2)/\partial x_2} < 0$, since $\lim_{x_2\to0}\left[\partial u(x_1,x_2)/\partial x_1\right]$ and $\lim_{x_2\to0}\left[\partial u(x_1,x_2)/\partial x_2\right]$ are both strictly positive and finite.
Note that model frameworks exist so that $\lim_{x_2\to0} dx_2/dx_1 = 0$ and where for some positive prices p_1, p_2 and k_2 solutions $0 < x^*_{1,S_2} < (y-k_2)/p_1$ exist, so that $u\left(x^*_{1,S_2}, x^*_{2,S_2}\right) > u(y/p_1, 0)$ for some positive prices p_1, p_2 and k_2. One such example is the framework used by De Jong (1990). As mentioned in Subchapter 1.4, however, the property $\lim_{x_2\to0}\left[\partial u(x_1,x_2)/\partial x_2\right]$ does not correspond to intuition.

[130] A (global) unique solution only exists if both the Kuhn-Tucker theorem and the Kuhn-Tucker sufficiency theorem hold. Both these theorems are only satisfied if all restrictions and the utility functions are concave; see Varian (1992: 503).

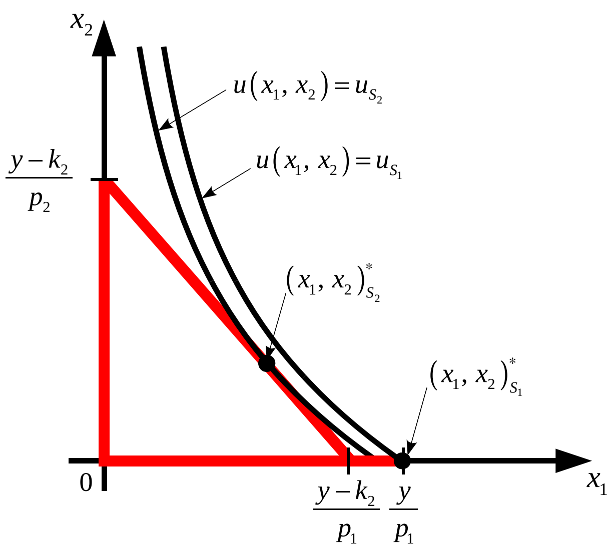


Figure 3.1.5: Optimum decision of a household that chooses not to own a car.

Note that these local maxima – $(x_1, x_2)^*_{S_2}$ and $(x_1, x_2)^*_{S_1}$ – are equal to the maxima discussed above. The local maximum on the budget line corresponds to the optimum solution of choice set S_2; the local maximum on the x_1 axis corresponds to the optimum solution of choice set S_1. In this case, the household would decide to choose the optimum of choice set S_1.

The following figure shows the local optima for the case where the household chooses the optimum of choice set S_2:

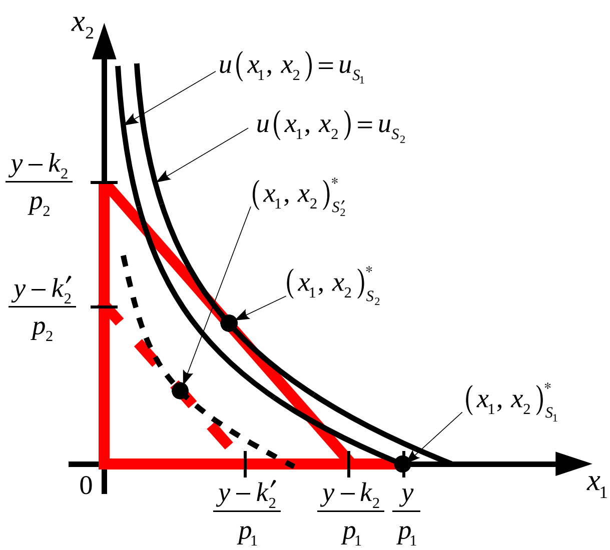


Figure 3.1.6: Optimum decision of a household that chooses not to own a car.

The diagram above also illustrates that if the fixed costs increase to k_2', the budget line will shift parallel towards the origin. At some level of fixed costs, the household will choose $(x_1, x_2)_{S_1}^*$ instead of $(x_1, x_2)_{S_2'}^*$.

No further discussion of the impact of economic variables on household choice shall be undertaken at this point because most results will depend on the choice of utility function.

3.2. Model with two goods and no fixed cost

In this subchapter, I present the Multiple Discrete-Continuous Extreme Value Model (MDCEV), according to which a household can choose between one car and no car. In contrast to the previous subchapter, where this model was presented for a general utility function, here I consider a specific type of a utility function. This utility function is parametrized. Some of these parameters are household-specific, accounting for the households' preferences for driving, which may vary between households. Since not all variables that influence the households' preferences for car driving are observed, it follows that individual preferences cannot be exactly described from the researcher's perspective. To account for this, the utility function contains an error term. Hence, whether or not a household chooses to own a car – given observed household characteristics and economic variables – can be predicted by the researcher only with a certain probability. Also, in the event that a household chooses to own a car, the amount consumed of composite good X_1 and the distance driven cannot be exactly determined. It is only possible to compute the probability that this distance is within a certain interval. One of the key issues of this subchapter is to compute these probabilities for each household. The function that describes these probabilities is a so-called Maximum Likelihood (ML) function. The ML function describes the probability of observing the consumption levels of households in a certain dataset, given the model parameters, market prices and household incomes. The ML function will be used to determine the model parameters by applying the Maximum Likelihood Estimation (MLE) procedure. This procedure chooses the parameters such that the ML function is maximized.

This subchapter is structured as follows: first, I will state the basic assumptions concerning the utility function and its random term. I will then derive the Marshallian demand function for driving a car for the case where households decide to own a car.[131] The density function of the distance driven is computed from this Marshalling demand function for the case where the household owns a car. Further, the probability that the household chooses not to own a car is computed. I will subsequently discuss how changes in parameter values influence both the density function and the probability that households decide not to own a car, illustrating this in diagrams. I will then proceed to compute the ML function, which incorporates the density, and the probability function. After this, I will show how to compute the effect of changes in variables, such as an increase in driving costs due to higher fuel prices, on driving demand and on the proportion of carless households. Finally, after presenting the results yielded by this model based on micro-census data[132], I will discuss how successfully this model describes the data.

---

[131] Note that since the utility function is of an additive separable type, this Marshallian demand function will depend negatively on the price p_2 for any values of the parameters.

[132] Micro-census on the travel behaviour of Swiss households in 2005, Bundesamt für Statistik (2006a).

The utility function and the assumptions on its random parameter

The model is based on a parametrized additive utility function. One parameter is household-specific, and accounts for the difference in relative preference for driving across households. This parameter is the only one that will be explained by household-specific properties. A random term is added to this parameter. This random term stands for the fact that researchers cannot observe all household properties relevant to the individual preference. The other parameters are identical for all households. The assumptions on the type of utility function and the random term are identical to the assumptions made by Bhat (2005). The assumptions on the type of utility function and the random term are crucial to acquire an easily computable ML function of an explicit form. In this subchapter, I will present only the model that captures two goods. Good one X_1 is a composite good that contains all goods apart from car driving. Good two X_2 denotes the kilometres driven by car. To start with, I present the utility function for the general case. I will then show how it can be transformed to enable one parameter to be interpreted as the preference for car driving relative to the preference of composite good X_1.[133] The utility function is defined as follows:[134]

$$U = u_1(X_1) + u_2(X_2) = \exp(m_1 + \xi_1) \cdot (X_1 + a_1)^{d_1} + \exp(m_2 + \xi_2) \cdot (X_2 + a_2)^{d_2}, \qquad (3.2.1)$$

where: $m_j = \gamma_j \cdot s + \delta \cdot b_j,\ j = 1, 2.$

Note that since optimal values X_1 and X_2 depend on the value of the random terms ξ_j, optimal values X_1 and X_2 are also random variables. I therefore use the capital letters X_1 and X_2. I assume a positive marginal utility and a decreasing marginal utility in all arguments, thus d_1 and d_2 are bound

---

[133] In the following, I will simply call *m* preference for car driving.

[134] This utility function is the same as in Bhat (2005: 684) and Bhat (2006: 39). In Bhat (2005), vector *x* was the time used for activities on holiday on a certain day, and the restriction was that the sum of duration activities equals 24 hours. The “price” was therefore equal to one for all activities, since the “price” of prolonging an activity by an extra hour is having to reduce the time left for other activities by one hour. In the context of the model as discussed in this chapter, the *x*'s are regarded as “ordinary” goods with individual prices. I therefore extended the model of Bhat (2005) to ensure that the restriction is now an ordinary micro-economic budget restriction where the sum of prices multiplied by the amounts of each good consumed equals the budget available to the household. The budget is considered to be the household’s annual income. Later, Bhat (2008) also extended the model framework to the case where the *x*'s represent ordinary goods and the model’s restriction is a budget constraint. In this later publication, he used a slightly different functional type of utility function, which is a positive transformation of the function I have used here. The partial utility of the different goods *k* can be rewritten as follows: $\exp(m_k + \xi_k) \cdot (X_k + a_k)^{d_k} = \exp(m_k + \xi_k) \cdot a_k^{d_k} \cdot (X_k / a_k + 1)^{d_k}$. This implies that parameter a_k corresponds to parameter γ_k in Bhat (2008), and d_k corresponds to a_k in Bhat (2008). Parameter ψ_k, which Bhat (2008) uses, is equal to $\psi_k = d_k \cdot \exp(m_k + \xi_k) \cdot a_k^{d_k - 1}$.

to be between zero and one:[135] $0 < d_j < 1,\ j = 1,2$. The smaller d_j is, the more rapid the marginal utility of good *j* decreases when X_2 increases. Parameters a_1 and a_2 can be considered as shifting parameters, since they shift the indifference curves of the utility function along the *x*- and in the *y*-axis, respectively. Note that the marginal utility of X_1 is infinite if X_1 approaches $-a_1$, and the marginal utility of X_2 is infinite if X_2 approaches $-a_2$. Values $-a_1$ and $-a_2$ therefore define the lower limits of optimal solutions for X_1 and X_2. Since I assume that good one X_1 is essential, a_1 is non-positive in order to ensure that the solution for X_1 is always positive. Like Bhat (2008)[136], I chose $a_1 = 0$. The minimal value of X_1 must therefore be greater than zero in the optimum, since $\lim_{x_1 \to 0} \partial u(x_1, x_2)/\partial x_1 = \infty$. Since good two X_2 is not essential, a_2 must be positive, so that bounded solutions $X_2 = 0$ are possible. Parameter a_2 will be determined by the estimation routine. Expression $\exp(m_j + \xi_j)$ is a weight on $(X_j + a_j)^{d_j}$. The higher $\exp(m_j + \xi_j)$ is, the stronger the preference for good *j*. This weight is determined by socio-demographic variable *s* and characteristic b_j of the corresponding good *j*, $m_j = \gamma_j \cdot s + \delta \cdot b_j,\ j = 1, 2$. This means, for instance, that households in rural areas usually have a greater preference for driving than households in urban areas. If a household moves from an urban to a rural area, therefore, m_2 is expected to increase. The utility of good X_2, therefore, would be weighted relatively higher compared to good X_1 for such households. The random terms ξ_j represent socio-demographic variable $\tilde{s}$ and the goods' characteristic $\tilde{b}$ that cannot be observed by the researcher. These random terms are assumed to be iid Gumbel distributed[137]:

$$\xi_j \sim iid\ gu(0,1),\ f_\xi(x) = e^{-x} \cdot \exp(-e^{-x}). \qquad (3.2.2)$$

---

[135] $\frac{\partial^2 U}{\partial X_j^{\ 2}} = d_j \cdot (d_j - 1) \cdot \exp(m_j + \xi_j) \cdot (X_j + a_j)^{d_j - 2} < 0$, if and only if $0 < d_j < 1$. This also implies that the utility function is concave and that the Hessian matrix is therefore negative (semi-)definite:

$$\begin{vmatrix} \frac{\partial^2 U}{\partial X_1^{\ 2}} & \frac{\partial^2 U}{\partial X_1 \partial X_2} \\ \frac{\partial^2 U}{\partial X_1 \partial X_2} & \frac{\partial^2 U}{\partial X_2^{\ 2}} \end{vmatrix} = \begin{vmatrix} \frac{\partial^2 U}{\partial X_1^{\ 2}} & 0 \\ 0 & \frac{\partial^2 U}{\partial X_2^{\ 2}} \end{vmatrix} = \frac{\partial^2 U}{\partial X_1^{\ 2}} \cdot \frac{\partial^2 U}{\partial X_2^{\ 2}} > 0 \text{ and} \frac{\partial^2 U}{\partial X_j^{\ 2}} < 0$$

, if and only if $0 < d_j < 1,\ j = 1,2$ and $X_1 > -a_1$ and $X_2 > -a_2$.

The term $\frac{\partial^2 U}{\partial X_1 \partial X_2}$ is equal to zero because the utility function is of the additive separable type.

[136] "Note that there is no translation parameter γ_k for the first good, because the first good is always consumed" (Bhat 2008: 290). Note that γ_k, which Bhat uses, corresponds to α_k, which I use.

[137] The Gumbel distribution is a non-symmetric distribution that has a similar shape to the normal distribution, see figure MA1.1 in Appendix MA1. The Gumbel distribution also has a number of useful properties that are necessary to obtain an ML function that is an explicit function of the parameters. Please refer to Mathematical Appendix MA1. The Gumbel distribution is also often called the extreme value distribution of type I.

This special form of the utility function and the assumptions on the error terms renders a number of formal simplifications possible when deriving the ML function. These simplifications will yield an ML function in closed form. Further, the cumulative density function that will appear in the ML function is of a simple form, hence permitting a short computation time. It is assumed that the household maximizes its utility by choosing optimal values for X_1 and X_2, subject to its budget constraint:

$$y = p_1 \cdot X_1 + p_2 \cdot X_2 \,. \tag{3.2.3}$$

Before deriving the conditions of optimality, I set $d = d_1 = d_2$, since the Marshallian demand function will only be of a closed form for this special case.[138] This property will be necessary when I extend the model to the case where I capture the fixed costs of maintaining a car, too. Further, it will transpire that even if this restriction is imposed, not all model parameters can be identified, meaning that an additional restriction has to be imposed.[139] Another reason for imposing such a restriction at this point is that I wish to ensure the comparability of the results of the model I am presenting in this subchapter and the model with fixed costs to be presented in Subchapter 3.3. Further, by assuming $d = d_1 = d_2$ the following useful transformation of the utility function can be made. Utility function (3.2.1) can be divided by $\exp(m_1 + \xi_1)$:[140]

$$U = \frac{u_1(X_1) + u_2(X_2)}{\exp(m_1 + \xi_1)} = (X_1 + a_1)^d + \exp(m + \varsigma) \cdot (X_2 + a_2)^d \,, \tag{3.2.4}$$

where $\varsigma = \xi_2 - \xi_1$, $m = m_2 - m_1$, and the random variable ς is logistically distributed[141], i.e. ς has the following density:

$$F_\varsigma(x) = \frac{1}{1 + e^{-x}} \tag{3.2.5}$$

---

[138] If this was not the case, it would also be impossible to make simplifications that would enable the probability of a households being carless to be computed. Further, the density function of driving demand will also be of a more simple form, and therefore quicker to compute.

Note that since the utility function is of an additive separable type, this Marshallian demand function will depend negatively on price p_2 and positively on household income y for any feasible values of the parameters. Since it is very plausible that household driving demand will increase when income increases or fuel prices decrease, this does not impose any infeasible restriction on the model.

[139] Also Bhat (2008) proposes to set $d = d_1 = d_2$ as one of three possible parameter restrictions that enable the model to be identifiable; see Bhat (2008: 290), Formula 32.

[140] Note that this is feasible since it is a positive transformation.

[141] For proof, see Rule 3 of MA1.

Expression $\exp(m+\varsigma)$ is the relative weight of the partial utility $u_2(X_2)$ with respect to partial utility $u_1(X_1)$, and hence the relative utility weight of driving. This transformation also shows that it is impossible to identify both parameters m_1 and m_2, since it is only the difference between them, $m = m_2 - m_1$, that matters. The random variable ς contains unobserved preferences for car driving. Since only one car model is available in this model with two goods, it is assumed that this car is an "average" car. It is therefore impossible to include any car feature *b*, thus only one coefficient δ can be estimated. This coefficient amounts to the preference for driving an average car without taking socio-demographic variables into account. Since the first component of vector *s* contains a one, the first component of γ, which I denote as γ_1, will correspond to this parameter δ. In the context of this two-good model, therefore, there are only socio-demographic variables plus a constant that defines the deterministic part *m* of the relative preference for car driving: $m = \gamma \cdot s$.

Note that, in this case, no fixed costs are assumed for good two. This implies that both the utility function and the restrictions are concave and differentiable functions. The following Lagrangian can therefore be solved using the Kuhn-Tucker theorem.

$$L = (X_1 + a_1)^d + \exp(m+\varsigma)\cdot(X_2 + a_2)^d - \lambda(p_1X_1 + p_2X_2 - y) + \mu_2X_2 \text{, with } a_1 = 0 . \qquad (3.2.6)$$

where λ is the Lagrangian multiplier corresponding to the budget constraint and Lagrangian multiplier μ_2 corresponds to the non-negativity constraint of X_2.[142] The random term ς represents the characteristics that are unobserved by the researcher. I assume that households have perfect information and that they exactly know their preferences. Further, I assume that their preferences can be exactly described by the utility function, as stated in the Lagrangian function (3.2.6). Households therefore know ς, which implies that they can be assumed to behave as though they would solve the optimization problem as stated in (3.2.6). I now present the solution of this optimization problem (3.2.6) by computing the maximization conditions:

$$d\cdot\frac{1}{(X_1+a_1)^{1-d}} - \lambda\cdot p_1 + \mu_1 = 0 , \qquad (3.2.7)$$

$$d\cdot\exp(m+\varsigma)\cdot\frac{1}{(X_2+a_2)^{1-d}} - \lambda\cdot p_2 + \mu_2 = 0 \text{, with } m = \gamma\cdot s \qquad (3.2.8)$$

and $-(p_1X_1 + p_2X_2 - y)\cdot\lambda \geq 0$ and $X_2\cdot\mu_2 \geq 0$ (3.2.9a-b)

[142] Note that since $\lim_{x_1\to 0}\partial u(x_1, x_2)/\partial x_1 = \infty$, it is not necessary to add μ_1X_1 to (3.2.6).

and two possible combinations of complementary slackness conditions:

$\mu_2 = 0$, $X_2 > 0$, $\lambda > 0$, $p_1 X_1 + p_2 X_2 - y = 0$ or (3.2.10a),

$\mu_2 > 0$, $X_2 = 0$, $\lambda > 0$, $p_1 X_1 + p_2 X_2 - y = 0$. (3.2.10b)

Derivation of the Maximum Likelihood function

Since the ML function is required at a later stage to estimate all parameters of this model by MLE, the probability of observing case $X_2 = 0$ and the density function of X_2 when $X_2 > 0$ have to be computed.

I start with the case $X_2 > 0$, which corresponds to the complementary slackness condition (3.2.10a). In this case, the Lagrangian multiplier μ_2 is zero, $\mu_2 = 0$, and therefore the first-order conditions (3.2.7) and (3.2.8) can be written as:

$$\frac{d}{p_1} \cdot \frac{1}{\left(X_1 + a_1\right)^{1-d}} = \lambda, \tag{3.2.11}$$

$$\frac{d}{p_2} \cdot \exp\left(m + \varsigma\right) \cdot \frac{1}{\left(X_2 + a_2\right)^{1-d}} = \lambda, \text{ with } m = \gamma \cdot s. \tag{3.2.12}$$

The aim is now to compute the density of X_2. Since the distribution of random variable ς is known, I solve[143] conditions (3.2.11) and (3.2.12) for ς:

$$\varsigma = V_1 - V_2, \tag{3.2.13}$$

with:

$$V_1 = \ln\left(d\right) - \ln\left(p_1\right) - \left(1-d\right) \cdot \ln\left(X_1 + a_1\right), \tag{3.2.14a}$$

$$V_2 = \ln\left(d\right) - \ln\left(p_2\right) + m - \left(1-d\right) \cdot \ln\left(X_2 + a_2\right), \text{ with } m = \gamma \cdot s. \tag{3.2.14b}$$

[143] First I take the logarithm of (3.2.11) and (3.2.12), which yields $\log(d) - \log(p_1) - (1-d)\cdot\log(X_1 + a_1) = \log(\lambda)$ and $\log(d) + m - \log(p_2) - (1-d)\cdot\log(X_2 + a_2) + \varsigma = \log(\lambda)$. When denoting $V_1 = \log(d) - \log(p_1) - (1-d)\cdot\log(X_1 + a_1)$ and $V_2 = \log(d) + m - \log(p_2) - (1-d)\cdot\log(X_2 + a_2)$, the two conditions can be written as $V_1 = \log(\lambda)$ and $V_2 + \varsigma = \log(\lambda)$. Replacing $\log(\lambda)$ and solving for ς yields (3.2.13). Note that Lagrangian multiplier λ is always strictly positive, since $\lambda = 0$ would mean that not all of the budget is exhausted. Since the utility function strictly increases in both arguments, however, not spending the whole budget would violate the assumption of utility maximization and would therefore not be feasible.

Further, using the budget restriction defined in (3.2.10a), X_1 can be expressed as a function of X_2:

$$X_1 = \frac{y - p_2 X_2}{p_1}. \tag{3.2.15}$$

It follows from Equations (3.2.13), (3.2.14) and (3.2.15) that random variable ς can be expressed as a function of X_2. The density of X_2 can be derived by using the theorem of densities of transformed variables[144]:

$$f_{X_2 \wedge (X_2>0)}(z \mid \theta, p_1, p_2, y, s) = f_\varsigma(V_1 - V_2) \cdot \left(\frac{1-d}{\frac{y - p_2 z}{p_1} + a_1} \cdot \frac{p_2}{p_1} + \frac{1-d}{z + a_2} \right), \tag{3.2.16}$$

where $f_\varsigma(x)$ is the probability density function (pdf) of the logistic distribution, $f_\varsigma(x) = \frac{\exp(e^{-x})}{\left(1 + \exp(e^{-x})\right)^2}$, V_1 and V_2 are given by (3.2.14a) and (3.2.14b), and $\theta = \{d, a_2, m, \beta\}$.

Note that $f_{X_2 \wedge (X_2>0)}(z \mid \theta, p_1, p_2, y, s)$ is continuously differentiable in all parameters θ within their feasible range.

I would now like to compute the probability that $X_2 = 0$. The driving demand is zero if condition (3.2.10b) is fulfilled and if the Lagrangian multiplier μ_2 takes a value greater than zero, $\mu_2 > 0$. It follows from (3.2.7) and (3.2.8), therefore, that[145]

---

[144] See Theorem 13 in MA1. For this special case, the result can also be computed as follows:
From $P(\varsigma \le V_1 - V_2) = F_\varsigma(V_1 - V_2)$ it follows that:

$$f_{X_2 \wedge (X_2>0)}(z) = \frac{\partial F_\varsigma(V_1 - V_2)}{\partial (V_1 - V_2)} \cdot \frac{\partial (V_1 - V_2)}{\partial z} = f_\varsigma(V_1 - V_2) \cdot \left(\frac{\partial V_1}{\partial X_2} \cdot \frac{\partial X_2}{\partial z} - \frac{\partial V_2}{\partial z} \right),$$

where $\frac{\partial V_1}{\partial X_1} \cdot \frac{\partial X_1}{\partial z} = -\frac{1-d}{x_1 + a_1} \cdot \frac{-p_1}{p_2} > 0$, $\frac{\partial V_2}{\partial z} = -\frac{1-d}{z + a_2} < 0$ and $x_1 = \frac{y - p_2 \cdot z}{p_1}$. Note that the expression $\frac{\partial V_1}{\partial X_2} \cdot \frac{\partial X_2}{\partial z} - \frac{\partial V_2}{\partial z}$ is positive for any value z that is in the feasible range $0 \le z < y/p_2$. Note that this is a necessary condition for the validity of the theorem of densities of transformed variables.

[145] Dividing (3.2.7) by p_1 and (3.2.8) by p_2, putting Lagrangian multipliers on one side and taking logarithms, yields $\log(d) - \log(p_1) - (1-d) \cdot \log(X_1 + a_1) = \log(\lambda)$ and $\log(d) + m - \log(p_2) - (1-d) \cdot \log(X_2 + a_2) + \varsigma = \log(\lambda - \mu_2/p_2)$. When denoting $V_1 = \log(d) - \log(p_1) - (1-d) \cdot \log(X_1 + a_1)$ and $V_2 = \log(d) + m - \log(p_2) - (1-d) \cdot \log(X_2 + a_2)$, the two conditions can be written as $V_1 = \log(\lambda)$ and $V_2 + \varsigma = \log(\lambda - \mu_2/p_2)$. Note that here both Lagrangian multipliers λ and μ_2 are positive. This implies that $\log(\lambda) > \log(\lambda - \mu_2/p_2)$, and therefore $V_1 > V_2 + \varsigma \Leftrightarrow \varsigma < V_1 - V_2$.
It is important to note that $\lambda - \mu_2/p_2$ is always positive, because the expression in (3.2.8) $d \cdot \exp(m + \varsigma) \cdot \frac{1}{(X_2 + a_2)^{1-d}}$ is always positive. Taking the logarithm of (3.2.8) therefore always yields real values for any finite value X_1.

$$\varsigma < V_1 - V_2 . \tag{3.2.17}$$

Hence the probability of observing $X_2 = 0$ is equal to the probability of ς being smaller than $V_1 - V_2$:

$$P(X_2 = 0 \mid \theta, p_1, p_2, y, s) = P(\varsigma < V_1 - V_2) = F_\varsigma(V_1 - V_2), \qquad (3.2.18)^{146}$$

with

$$V_1 = \ln(d) - \ln(p_1) - (1-d) \cdot \ln\left(\frac{y}{p_1} + a_1\right), \tag{3.2.19a}$$

$$V_2 = \ln(d) - \ln(p_2) + m - (1-d) \cdot \ln(a_2), \text{ with } m = \gamma \cdot s \tag{3.2.19b}$$

and where $F_\varsigma(x)$ is the cumulative distribution function (cdf) of the logistic distribution, $F_\varsigma(x) = \frac{1}{1+\exp(e^{-x})}$.

I have now computed both the probability X_2 being zero, $P(X_2 = 0 \mid \theta, p_1, p_2, y, s)$, and the density of X_2 for positive values. The following diagram shows the concrete shape of the probability distribution of X_2 if X_2 is positive, given the values of the parameters and the economic variables:

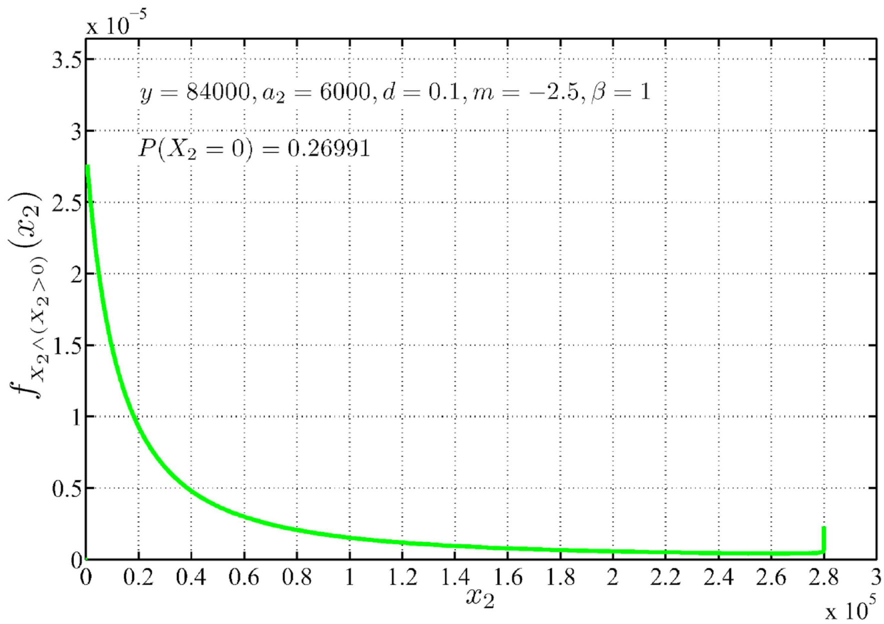


Figure 3.2.1: Density function of distances driven.

[146] Note again that the probability function $P(X_2 = 0)$ is conditional on parameters d, a_2, m, β and economic variables p_1, p_2, y. The probability should then, in fact, be written as $P(X_2 = 0 \mid \theta, p_1, p_2, y)$ with $\theta = \{d, a_2, m, \beta\}$. To keep the notation short and simple, however, I used the notation as in (3.2.18).

This diagram shows that the density at very low annual kilometres is rather high. Further, the density at very high annual kilometres, where households almost spend the entire budget on car driving, increases. Both are unrealistic, and changing parameter values do not change this shape significantly either. When looking at real data[147] of households with an annual income of CHF 84,000, it becomes obvious that this probability distribution is unrealistic:

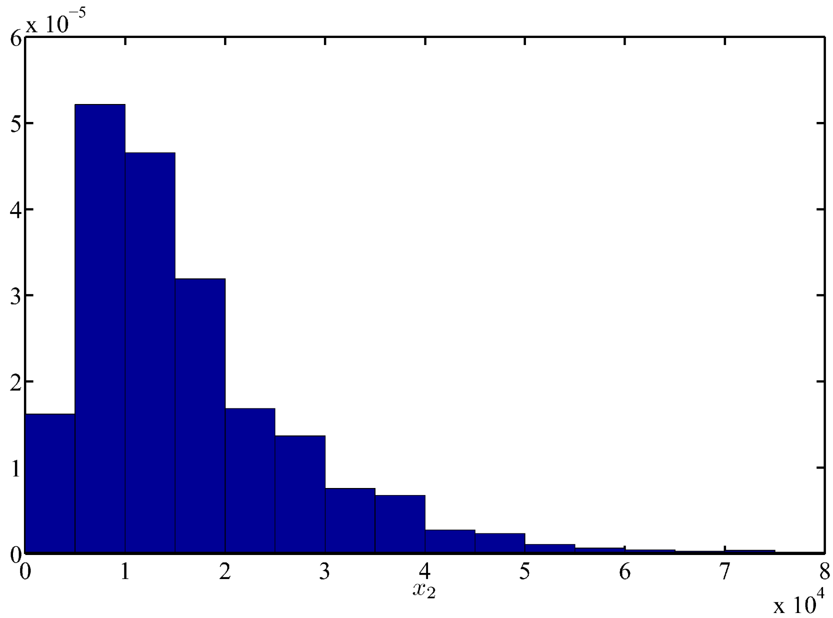


Figure 3.2.2: Histogram of distances driven by households living in non-rural areas with an annual income of CHF 84,000.[148]

Comparing the histogram and the density function reveals that the density function has too large densities at very low values and high values. This leads to the conjecture that the variance of the random term ς is too high. I therefore replace random term ς by $\beta \cdot \varsigma$ in the following. Parameter β is a strictly positive scaling factor that I expect to be smaller than one, such that it reduces the variance of the error term.[149]

---

[147] The data is taken from a survey on Swiss households concerning travel behaviour; Bundesamt für Statistik (2006a). For more detailed information on this dataset, see section “Data” in this subchapter.

[148] Note that the surface of this histogram is normalized so that it is equal to probability $P(X_2 > 0)$ conditional on an income of CHF 84,000.

[149] Note that Bhat (2008) denotes this scaling factor as σ and specifies $\sigma \cdot (m_k + \xi_k)$; Bhat (2008: 285), Formulas 14 and 15.

Since the error term has changed, it is necessary to derive the probability of X_2 being zero and the density function of X_2 again. To this end, I start with condition

$$\beta \cdot \varsigma = V_1 - V_2, \tag{3.2.20}$$

which corresponds to condition (3.2.13) of the previous case without parameter β and follows from condition (3.2.10a), corresponding to the case when the consumption of X_2 is positive. Again, from condition (3.2.20) the density of X_2 can be derived using the theorem of densities of transformed variables[150],

$$f_{X_2 \wedge (X_2>0)}(z \mid \theta,\, p_1,\, p_2,\, y, s) = \frac{1}{\beta} \cdot f_\varsigma\left(\frac{V_1 - V_2}{\beta}\right) \cdot \left(\frac{1-d}{\dfrac{y - p_2 z}{p_1} + a_1} \cdot \frac{p_2}{p_1} + \frac{1-d}{z + a_2} \right), \tag{3.2.21}$$

where V_1 and V_2 are defined in (3.2.19a) and (3.2.19b), and θ contains all parameters, $\theta = \{d, a_2, \gamma, \beta\}$.

Analogously to the case where the error term was not multiplied by a scaling parameter β, probability $X_2 = 0$ can be computed using condition (3.2.10b). Again, the only modification of condition (3.2.17) is to replace the error term ς by $\beta \cdot \varsigma$:

$$\beta \cdot \varsigma < V_1 - V_2 \,. \tag{3.2.22}$$

Reformulating this expression by dividing both sides by β yields:

$$\varsigma < \frac{V_1 - V_2}{\beta} \,. \tag{3.2.23}$$

It follows from this that probability $P(X_2 = 0)$ yields:

$$P(X_2 = 0 \mid \theta, p_1, p_2, y, s) = F_\varsigma\left(\frac{V_1 - V_2}{\beta}\right). \tag{3.2.24}$$

---

[150] See Appendix MA1, Theorem 13. In this special case, the result can also be computed as follows:
It follows from $P(\varsigma \le V_1 - V_2) = F_\varsigma(V_1 - V_2)$ that:

$$f_{X_2 \wedge (X_2>0)}(z) = \frac{\partial F_\varsigma\left(s^{-1} \cdot (V_1 - V_2)\right)}{\partial s^{-1} \cdot (V_1 - V_2)} \cdot \frac{\partial \left(s^{-1} \cdot (V_1 - V_2)\right)}{\partial (V_1 - V_2)} \cdot \frac{\partial (V_1 - V_2)}{\partial z} = f_\varsigma\left(s^{-1} \cdot (V_1 - V_2)\right) \cdot s^{-1} \cdot \left(\frac{\partial V_1}{\partial X_2} \cdot \frac{\partial X_2}{\partial z} - \frac{\partial V_2}{\partial z} \right),$$

where, again, $\frac{\partial V_1}{\partial X_1} \cdot \frac{\partial X_1}{\partial z} = -\frac{1-d}{x_1 + a_1} \cdot \frac{-p_1}{p_2} > 0$, $\frac{\partial V_2}{\partial z} = -\frac{1-d}{z + a_2} < 0$ and $x_1 = \frac{y - p_2 \cdot z}{p_1}$. Note that the expression $\frac{\partial V_1}{\partial X_2} \cdot \frac{\partial X_2}{\partial z} - \frac{\partial V_2}{\partial z}$ is positive for any value *z*, which is in the feasible range $0 \le z < y/p_2$. Note that this a necessary condition for the validity of the theorem of densities of transformed variables.

The density function of X_2 (3.2.21) yields a shape that adapts histogram 3.2.2 better, as illustrated in Figure 3.2.3.

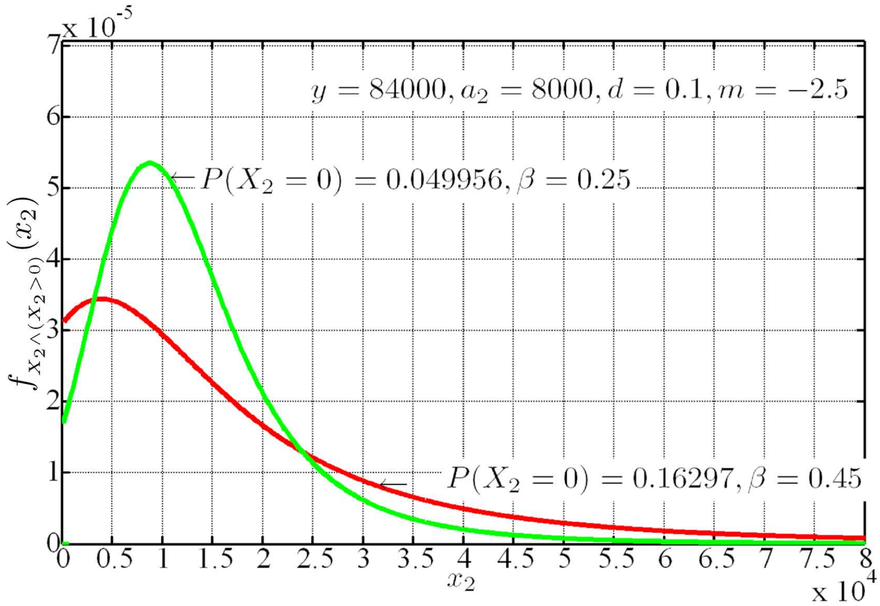


Figure 3.2.3: Density function for different values of scaling factor of random term.

This diagram shows that the smaller scaling factor β, the more the density function of X_2 is concentrated around a value defined by the other parameters. By choosing an appropriate value for scaling factor β, which reduces the variance of the error term, density function (3.2.21) adapts the empirical distribution illustrated in Histogram 3.2.2 much better than density function (3.2.16). Thus, the density function can be adapted much more effectively to the data by introducing scaling factor β

It now remains to define the ML function that yields the probability of observing data of a complete dataset. First, I derive the probability of observing the demand of a single household, which can be computed using (3.2.21) and (3.2.24):

$$L(z = X_2 \mid \theta, p_1, p_{2n}, y_n, s_n) = f_{X_2 \wedge (X_2>0)}(z \mid \theta, p_1, p_{2n}, y_n, s_n) \cdot I(z>0) + \\ + P(X_2 = 0 \mid \theta, p_1, p_{2n}, y_n, s_n) \cdot I(z=0), \qquad (3.2.25)$$

where $I(z>0)$ and $I(z=0)$ are indicator functions, being one, when the argument is true, and zero otherwise. Parameter-vector θ contains all parameters, $\theta = \{d, a_2, m, \beta\}$. The probability $P(X_2 = 0 \mid \theta, p_1, p_2, y, s)$ is defined in (3.2.24); the density $f_{X_2 \wedge (X_2>0)}(z \mid \theta, p_1, p_2, y, s)$ is defined in (3.2.21). Using probability (3.2.25), the probability of observing the whole dataset $(x_{2n}, y_n, p_{2n}, s_n)_{n=1,2,..,N}$ can be computed by:

$$L_{MLE}\left(\left(X_2 = x_{2n}\right)_{n=1,2,..,N} \mid \theta,\ p_1, p_{2n=1,2,..,N}, y_{n=1,2,..,N}, s_{n=1,2,..,N}\right) = \prod_{n=1}^{N} P\left(x_{2n} = 0 \mid \theta,\ p_1, p_{2n}, y_n, s_n\right)^{I\left(x_{2n}=0\right)} \cdot$$

$$\cdot \prod_{i=n}^{N} f_{X_2 \wedge (X_2 > 0)}\left(x_{2n} \mid \theta,\ p_1, p_{2n}, y_n, s_n\right)^{I\left(x_{2n}>0\right)}. \qquad (3.2.26)$$

Note that it is necessary to assume independence between observations. Otherwise the observation probability of the whole dataset L_{MLE} cannot be written as a product of the probabilities of the single observations $L\left(z = X_2 \mid \theta, p_1, p_{2n}, y_n, s_n\right)$. Since the households in the dataset were chosen randomly from the telephone directory list, the assumption of independence is reasonable.

For a numerical computation, it is more efficient to compute the log likelihood function:

$$\log\left(L_{MLE}\left(\left(X_2 = x_{2n}\right)_{n=1,2,..,N} \mid \theta,\ p_1,\ p_2,\ y_{n=1,2,..,N},\ s_{n=1,2,..,N}\right)\right) =$$

$$\sum_{i=1}^{N} I\left(x_{2n} > 0\right) \cdot \log\left(f_{X_2 \wedge (X_2>0)}\left(x_{2n} \mid \theta,\ p_1,\ p_{2n}, y_n, s_n\right)\right) +$$

$$+ \sum_{i=1}^{N} I\left(x_{2n} = 0\right) \cdot \log\left(L\left(x_{2n} = 0 \mid \theta,\ p_1,\ p_{2n}, y_n, s_n\right)\right). \qquad (3.2.27)$$

I compute parameters $\theta = \{d, a_2, m, \beta\}$ using the Maximum Likelihood Estimation (MLE) method. The MLE routine chooses parameters such that the (log-) ML function is maximized. In this case, the MLE method chooses parameters $\theta = \{d, a_2, m, \beta\}$ that maximize the ML function (3.2.27). Intuitively, the MLE procedure can be regarded as though the parameters were chosen so that the density function adapts histogram 3.2.2 as closely as possible and that the difference between probability (3.2.25) and the proportion of carless households is as small as possible. Since the logarithm is a monotone function, the parameter values that maximize (3.2.26) also maximize (3.2.27). The properties of the estimators have already been described in numerous textbooks. For a good description, see Chapter 5 in Cameron et al. (2005).

Simulation

The principal aim of this study is to simulate the impact of a change in fuel price and households' income on car ownership and travelling behaviour. To do this, I will consider only data provided by the households, with the exception of the annual kilometres driven. Given this information and the parameters, I will compute the probabilities that households do not own a car using the estimated parameters $\tilde{\theta}$[151] and the corresponding economic variables p_1, p_2, y, s, and the expected annual kilometres, that households would drive a car, for each individual household. Summing up these

[151] I use the notation $\tilde{\theta}$ instead of $\hat{\theta}$ to emphasize that the estimated parameter vector is in fact a random vector, see (3.2.39).

probabilities and expected annual kilometres driven and dividing them by the number of observations yields the average simulated proportion of households with a car and the average simulated driving distance. I will then change the marginal costs of driving a car (p_2) and income y of each household by one percent and rerun the simulation as described above. I then consider the difference in results to be the actual effect of fuel price and income on car ownership and average driving demand.[152] Recall the underlying assumption that households can always switch between owning and not owning a car without any costs. In reality, this switch does incur costs. Also, it can be assumed that households will show a considerable habit persistence. For this reason, I expect that the simulated number of households switching from car ownership to non-car ownership when the fuel price increases will be larger than in reality. I assume that this simulated change in the proportion of carless households applies rather in the long run. For the same reasons, the simulated change of driving distance when fuel prices or incomes change will also yield an upper bound when considering the short-term effect. On the other hand, reducing the total driving distance is not connected to any switching costs. I therefore expect the simulated effect of an increase in fuel price on driving distance to be closer to reality than the simulated change of the proportion of carless households. In the long run, habit persistence and switching costs do not play a role; the simulated change in driving demand therefore applies to this case.

In the following, I present the formulas used for the simulation, starting with the formula for the simulated probabilities. The function I use is exactly the same as that used for the ML estimation. The simulated probability of a single household is therefore:

$$P_{sim,n} \mid \tilde{\theta}, p_1,\ p_{2n}, y_n, s_n = P\left(x_{2n} = 0 \mid \tilde{\theta}, p_1,\ p_{2n}, y_n, s_n\right), \tag{3.2.28}$$

where $\tilde{\theta}$ is a random vector with the distribution of the estimated parameter vector $\tilde{\theta}$. The distribution of $\tilde{\theta}$ is defined further below, see (3.2.39).

The expected distance driven by a household is defined as:

$$E_{sim,n}\left(X_2 \mid \tilde{\theta}, p_1,\ p_{2n}, y_n, s_n\right) = \int_{z=0}^{z=y/p_2} z \cdot f_{X_2 \wedge (X_2>0)}\left(z \mid \tilde{\theta},\ p_1,\ p_{2n}, y_n, s_n\right) dz\ . \tag{3.2.29}$$

[152] This procedure corresponds to the note in Cameron et al. (2005: 122).

Since this integral is calculated numerically and the density function yields very small values for large values of z, this integral should be transformed via integration by parts[153]:

$$E_{sim,n}\left(X_2 \mid \tilde{\theta}, p_1, p_{2n}, y_n, s_n\right) = \frac{y}{p_2} - \int_{z=0}^{z=y/p_2} F_{X_2 \wedge (X_2>0)}\left(z \mid \tilde{\theta}, p_1, p_{2n}, y_n, s_n\right) dz \, , \tag{3.2.30}$$

where $F_{X_2 \wedge (X_2>0)}$ is the cumulative distribution function of car driving, given that when the household owns a car, it drives a positive annual distance. This function is defined as:

$$F_{X_2 \wedge (X_2>0)}\left(z \mid \tilde{\theta}, p_1, p_{2n}, y_n, s_n\right) = F_{\varsigma}\left(\frac{V_{1n} - V_{2n}}{\beta}\right), \tag{3.2.31}$$

where V_{1n} and V_{2n} are defined as

$$V_{1n} = \ln\left(\tilde{d}\right) - \ln\left(p_1\right) - \left(1-\tilde{d}\right) \cdot \ln\left(\frac{y - p_{2n} z}{p_1}\right), \tag{3.2.32a}$$

$$V_{2n} = \ln\left(\tilde{d}\right) - \ln\left(p_{2n}\right) + \tilde{m}_n - \left(1-\tilde{d}\right) \cdot \ln\left(z + \tilde{a}_2\right), \text{ with } \tilde{m}_n = \tilde{\gamma} \cdot s_n \, . \tag{3.2.32b}$$

The population average marginal effects of changes of economic variables – here the case of the population average marginal effect of an increase in fuel price – are computed as follows:

$$\begin{aligned} &\frac{\Delta P_{sim} \mid \tilde{\theta}, p_1, p_{2n}, y_n, s_n}{\Delta p_2} = \frac{1}{N} \cdot \frac{1}{\Delta p_2} \cdot \sum_{n=1}^{N} \frac{\Delta P_{sim,n} \mid \tilde{\theta}, p_1, p_{2n}, y_n, s_n}{\Delta p_2} = \\ &= \frac{1}{N} \cdot \frac{1}{\Delta p_2} \cdot \sum_{n=1}^{N} P\left(x_{2n} = 0 \mid \tilde{\theta}, p_1, p_{2n} + \Delta p_2, y_n, s_n\right) - \frac{1}{N} \cdot \frac{1}{\Delta p_2} \cdot \sum_{n=1}^{N} P\left(x_{2n} = 0 \mid \tilde{\theta}, p_1, p_{2n}, y_n, s_n\right), \end{aligned} \tag{3.2.33}$$

$$\begin{aligned} &\frac{\Delta E_{sim}\left(X_2 \mid \tilde{\theta}, p_1, p_{2n}, y_n, s_n\right)}{\Delta p_2} = \frac{1}{N} \cdot \frac{1}{\Delta p_2} \cdot \sum_{n=1}^{N} \frac{\Delta E_{sim,n}\left(X_2 \mid \tilde{\theta}, p_1, p_{2n}, y_n, s_n\right)}{\Delta p_{2n}} = \\ &= \frac{1}{N} \cdot \frac{1}{\Delta p_2} \cdot \sum_{n=1}^{N} E_{sim,n}\left(X_2 \mid \tilde{\theta}, p_1, p_{2n}\right) + \Delta p_2, y_n, s_n - \frac{1}{N} \cdot \frac{1}{\Delta p_2} \cdot \sum_{n=1}^{N} E_{sim,n}\left(X_2 \mid \tilde{\theta}, p_1, p_{2n}, y_n, s_n\right). \end{aligned} \tag{3.2.34}$$

The population average marginal effects (3.2.33) and (3.2.34) are conditional on the estimated parameter vector $\tilde{\theta}$ These marginal effects therefore have certain distributions, which I intendo determine to see how accurate the model's predictions are. As an example, if the distribution of the

---

153 $\int_a^b x \cdot f(x) dx = b \cdot F(b) - a \cdot F(a) - \int_a^b F(x) dx$. Plugging in the limits $a = 0$ and $b = y/p_2$ and using $F(x) = F_{X_2 \wedge (X_2>0)}(x)$ yields: $\int_0^{y/p_2} x \cdot f_{X_2 \wedge (X_2>0)}(x) dx = b \cdot F_{X_2 \wedge (X_2>0)}(y/p_2) - \int_0^{y/p_2} F_{X_2 \wedge (X_2>0)}(x) dx$.

own price elasticity (3.2.33) is such that the 90% confidence interval is $[-1.2\% \ldots +0.8\%]$, it is not even possible to conclude which sign applies to the elasticity. To determine the distribution of the marginal effects, I will apply the delta method. Since the functions that describe the elasticities of driving demand contain an integral that can only be solved numerically, the derivatives can also only be computed numerically with respect to the components of θ_1. Since the computation of these derivatives may be rather inaccurate, I prefer to compute the distribution of the marginal effects as follows: I draw a sequence of draws $\{\theta_k\}_{k=1..K}$ from the distribution of $\tilde{\theta}$. For each draw, I then compute the simulated values (3.2.33) and (3.2.34):

$$\frac{\Delta P_{sim} \mid \theta_k, p_1, \; p_{2n}, y_n, s_n}{\Delta p_{2n}}, \tag{3.2.35}$$

$$\frac{\Delta E_{sim}\left(X_2 \mid \theta_k, p_1, p_{2n}, y_n, s_n\right)}{\Delta p_{2n}}. \tag{3.2.36}$$

I generated sequence $\{\theta_k\}_{k=1..K}$ using the Gibbs sampling algorithm. The computed values (3.2.35) and (3.2.36) were then illustrated in a histogram; see Figure 3.2.7. I was able to compute the 95% interval of the marginal effects using a Kernel function. A further advantage of this procedure is that it yields the distribution based on the actual non-linear elasticity function. This approach may shed light on whether it is inappropriate to use linearised functions, on which the delta method is based.[154] If it was inappropriate, the method I present here or a bootstrapping method should be used to compute the expectation value and the distribution of the marginal effects.

Note that marginal effects (3.2.33) and (3.2.34) are conditional on estimated parameters $\tilde{\theta}$. Since these estimated parameters are in fact a random vector, I wish to compute the unconditional simulated marginal effects that reflect the expected value of the actual marginal effects, which are therefore "the

---

[154] Let us define the marginal effect of fuel price on driving demand to be a non-linear function of a θ as $f(\theta)$. Further, I define function $g(\theta)$ as the linear approximation of $f(\theta)$ at $\hat{\theta}$: $g(\theta) = a_0 + \sum_{j=1}^{J} a_j \cdot \theta_j$, with $a_j = \partial f(\theta)/\partial\theta_j \mid \theta = \hat{\theta}$ and $a_0 = h(\hat{\theta}) - \sum_{j=1}^{J} a_j \cdot \hat{\theta}_j$. If $f(\theta)$ was perfectly linear at $\hat{\theta}$ and if $E(\theta) = \hat{\theta}$, it follows that:

$$E\left(f(\tilde{\theta})\right) = E\left(g(\tilde{\theta})\right) = a_0 + \sum_{j=1}^{J} a_j \cdot E\left(\tilde{\theta}_j\right) = a_0 + \sum_{j=1}^{J} a_j \cdot \hat{\theta}_j = f\left(\hat{\theta}\right).$$

This implies that if the difference $\left|E\left(g(\tilde{\theta})\right) - f\left(\hat{\theta}\right)\right|$ exceeds a certain level, the linear function $g(\theta_1)$ may be an inappropriate approximation of function $f(\theta)$ at $\theta = \hat{\theta}$, and determining the distribution of the marginal effects using the delta method would be inappropriate. I do not define a certain level of $\left|E\left(g(\theta)\right) - f\left(\hat{\theta}\right)\right|$ above which I consider the linear approximation to be infeasible, but I will compute this difference to see whether there may be any doubt that the linearization could be infeasible from a qualitative point of view. Note that I will approximate $E\left(f(\theta)\right)$ by $E\left(f(\theta)\right) = \frac{1}{K} \cdot \sum_{k=1}^{K} f(\theta_k)$, in which case θ_k is a sequence of random draws from the distribution of θ.

best" measure. Since the simulated marginal effects depend non-linearly on $\tilde{\theta}$, the unconditional marginal effects differ to the marginal effects at $\theta=\hat{\theta}$.[155] To obtain the unconditional simulated marginal effects, the random variable $\tilde{\theta}$ has to be eliminated: $P_{sim,n}=E_{\tilde{\theta}}\left(P_{sim,n}\,|\,\tilde{\theta}\right)$, $E_{sim,n}\left(X_2\right)=E_{\tilde{\theta}}\left(E_{sim,n}\left(X_2\right)|\,\tilde{\theta}\right)$. Since $\tilde{\theta}$ is multivariate, normally distributed, as shown below, it is virtually impossible to compute such an integral. To circumvent computing this integral, I compute an approximation of the unconditional simulated probabilities $P_{sim,n}$ and expected values $E_{sim,n}\left(X_2\right)$ as follows:

$$\frac{\Delta P_{sim}\,|\left\{p_1,\ p_{2n},y_n,s_n\right\}_{n=1..N}}{\Delta p_2}\cong\frac{1}{K}\cdot\sum_{k=1}^{K}\frac{\Delta P_{sim}\,|\,\theta_k,p_1,\ p_{2n},y_n,s_n}{\Delta p_2}, \tag{3.2.37}$$

$$\frac{\Delta E_{sim}\left(X_2\right)|\left\{p_1,\ p_{2n},y_n,s_n\right\}_{n=1..N}}{\Delta p_2}\cong\frac{1}{K}\cdot\sum_{k=1}^{K}\frac{\Delta E_{sim}\left(X_2\right)|\,\theta_k,p_1,\ p_{2n},y_n,s_n}{\Delta p_2}. \tag{3.2.38}$$

I now show how the estimated parameters $\tilde{\theta}$ are actually distributed. Since the number of observations of the dataset is rather high, the large sample theory can be applied. I therefore assume that $\tilde{\theta}$ is normally distributed as follows:

$$\tilde{\theta}\sim\Phi\left(\hat{\theta},\operatorname{var}\left(\tilde{\theta}\right)\right), \tag{3.2.39}$$

where $\hat{\theta}$ is the estimated value by MLE and $\operatorname{var}\left(\hat{\theta}\right)$ can be approximated by the inverse Hessian matrix of the log-likelihood function (3.2.27) at $\hat{\theta}$:[156]

$$\operatorname{var}\left(\tilde{\theta}\right)\cong\left[\frac{\partial^2\log\left(L_{MLE}\left(\left(X_2=x_{2n}\right)_{n=1,2,..,N}\,|\,\theta,\ p_1,\ p_2,\ y_{n=1,2,..,N},\ s_{n=1,2,..,N}\right)\right)}{\partial\theta\partial\theta'}\right]_{\theta=\hat{\theta}}^{-1} \tag{3.2.40}$$

Data

The data I used to estimate the parameters is the micro-census data on the travel behaviour of Swiss households, Bundesamt für Statistik (2006a). 33,000 households were interviewed. The dates of the interviews were more than less evenly distributed over the year 2005. This dataset contains a vast number of information on traveling behavior, ownership of cars, motorbikes and bicycles, and

---

[155] Note that the expected value of $\tilde{\theta}$ is given by $E\left(\tilde{\theta}\right)=\hat{\theta}$. Let $f\left(x\right)$ be a non-linear function, yielding a scalar. Then, usually, $E\left(f\left(\tilde{\theta}\right)\right)\neq f\left(\hat{\theta}\right)$. Since marginal effects conditional on $\tilde{\theta}$ are also non-linear in $\tilde{\theta}$, (3.2.37) will be unequal to (3.2.35), and (3.2.38) will be unequal to (3.2.36).

[156] See Wooldridge (2002: 395) or Cameron et al. (2005: 143-146), where we implicitly assume that the Hessian matrix is non-singular.

information on the households. Since the purpose of this study is to investigate fuel demand, I will use the information on total kilometers driven by cars. Since in the present model I do not consider the choice of different car types, I will use the total annual kilometers driven by all households as a proxy for fuel demand. Since I am basically interested in the effect of fuel prices on the distance traveled and the decision of whether or not to own one or several cars, I will only use the household variables that appeared to have the most important impact on travel distance or fuel demand in other models.[157] In this case, I will only use only the income and the place of living residence as explanatory variables, namely whether the households live in a rural area or in a non-rural area, which I denote "urban areas". As in Bhat (2008)[158], I choose the price of the composite good X_1 to be numeraire.[159] Since p_1 is one, amount X_1 is nothing but the remaining income minus the amount spent of driving, $X_1 = y - p_2 \cdot X_2$, since I assume that households spend all of their income and do not save anything. Of course, it is a simplification to assume that households will spend all of their income on consumption, but no data on savings is available in the dataset.

Results

By using the ML function (3.2.27), the parameters $\theta = \{d, a_2, \gamma, \beta\}$ can be estimated by MLE. Note that the parameter *m* is explained by the sum of constant γ_1 and the effect of the household location expressed by parameter γ_2 multiplied by dummy s_2 "rural" (rural = 1, if a household is located in a rural area). Therefore, the parameter *m* that represents the deterministic part of the household's relative preference for car driving is household-specific: $m_n = \gamma_1 \cdot s_1 + \gamma_2 \cdot s_{2n}$, where *n* is the index for household and $s_1 = 1$. From this also follows that the utility function (3.2.4) is household-specific, which is an important feature of this model.

When estimating all parameters $\theta = \{d, a_2, \gamma_1, \gamma_2, \beta\}$, it turned out that not all of these parameters can be identified. There is one degree of freedom in the optimum. The following diagram shows the iteration process of the optimization routine for the parameters $\theta = \{d, a_2, \gamma_1, \gamma_2, \beta\}$. In the diagram, the trajectories of all parameters are plotted as ratio to their own value in the last iteration step. If a trajectory is closed to one at some iteration step, it implies that it is closed to the value it reached in the last iteration step. The number of iterations was limited to 1,000 steps. Therefore, the last step cannot be considered as if the optimization process would have found the optimum values of the parameters.

---

[157] Note that when including more explanatory variables, the resulting elasticities do not change much, see Appendix A3.12.

[158] "If an outside good is present, label it as the first good which now has a unit price of one," Bhat (2008: 290). Note that Bhat denotes an "outside good" as a good that is always chosen.

[159] This is reasonable since the price of a composite good is a price index, and a price index is scale-free.

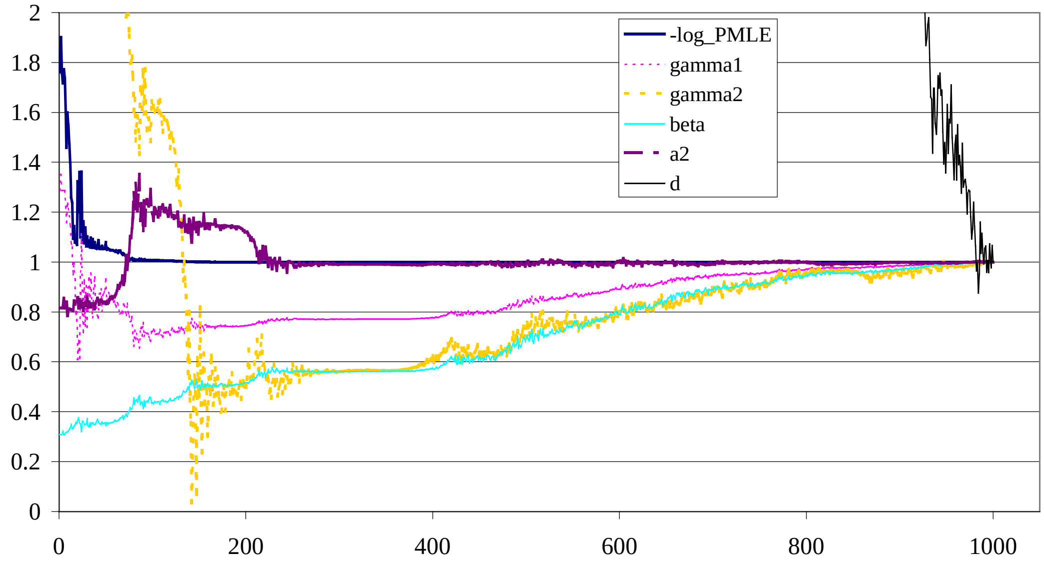


Figure 3.2.4: Parameter values during iteration process.[160]

The diagram shows that the value of the ML function (3.2.27) reaches a value closed to its final value after about 150 iterations. After that, there is no significant improvement in the log likelihood value "log_PMLE" that corresponds to the value of (3.2.27), but all parameter values change quite dramatically. This leads to the conclusion that not all parameters are identifiable. Further, the Hessian matrix, necessary for computing the covariance matrix of estimated parameters, could numerically not be computed, since the variation of the log likelihood function "log_PMLE" was too small with respect to some parameter values as computed in the last iteration step. Since in the vicinity of the last iterations steps parameter d variates quite a lot, I decided to estimate all parameters conditional on d. This procedure is also proposed in Bhat (2008).[161] Bhat also provides some intuition as to why not all parameters are identifiable: He shows that for some combinations of parameters d and a_2 the shape of the partial utility function of good two, $\exp(m+\varsigma)\cdot(X_2+a_2)^d$, is almost identical,[162] which may be the reason for the identification problem.

[160] Note that the results of Figure 3.2.4 and the following Figures 3.2.5 and 3.2.6 and Table 3.2.1 below are based on a dataset that only contains 20% of the complete data for the reason of saving computation time. This size of the dataset is sufficient to compute these results.

[161] "Alternatively, the analyst can stick with one functional form a priori, but experiment with various fixed values of a_k for the γ_k-profile [...]"; Bhat (2008: 282), footnote 9. The term "functional form" refers to the three utility functions (32) in Bhat (2008: 290). The so-called "γ_k-profile" corresponds to the model based on the third utility function of (32) in Bhat (2008: 290). The utility function (3.2.4) I use is a positively transformed function of that third utility function, and I fix its parameter value $d = d_1 = d_2$ and estimate all the other parameters.

[162] See Bhat (2008: 281f.) . Particularly, the figures 4a and 4b show interesting results for the case $a_k \to 0$ that corresponds to the case $d \to 0$.

In the following, the parameters $\{a_2, \gamma_1, \gamma_2, \beta\}$ are estimated by the maximum likelihood method conditional on some fixed parameters values of *d*, namely $d = \{0.8, 0.5, 0.3, .., 0.0001, 0.00001\}$, using a dataset of size 4,174. This data was generated by randomly choosing 20% of the households of the original dataset, as described in the section “data”. The size of this dataset is sufficient to estimate parameters with large *t*-values. Using these data and the estimated parameters, I simulated also the elasticity of demand with respect to income and marginal costs of driving. The results are listed in the following table:

d	$\hat{\gamma}_1$	$\hat{\gamma}_2$	$\hat{\beta}$	$\hat{a}_2$	$E_{sim}(X_2)$	P_{sim}	$\log(P_{MLE})$
0.8	-1.577	0.0703	0.0766	7,754.72	16,879.23	0.1636887	37,762.1553
0.5	-2.000	0.1754	0.1894	7,877.31	16,812.34	0.1632946	37,744.7949
0.3	-2.283	0.2456	0.2648	7,892.84	16,804.72	0.1632504	37,743.0759
0.2	-2.424	0.2806	0.3026	7,897.16	16,802.68	0.1632386	37,742.6451
0.15	-2.494	0.2982	0.3215	7,898.86	16,801.89	0.1632339	37,742.4825
0.1	-2.565	0.3157	0.3403	7,900.34	16,801.21	0.1632300	37,742.3452
0.07	-2.607	0.3263	0.3516	7,901.14	16,800.84	0.1632279	37,742.2728
0.05	-2.636	0.3333	0.3592	7,901.63	16,800.61	0.1632265	37,742.2282
0.03	-2.664	0.3403	0.3667	7,902.10	16,800.40	0.1632253	37,742.1861
0.02	-2.678	0.3438	0.3705	7,902.33	16,800.30	0.1632247	37,742.1660
0.01	-2.692	0.3473	0.3743	7,902.56	16,800.19	0.1632242	37,742.1465
0.001	-2.705	0.3505	0.3777	7,902.75	16,800.11	0.1632236	37,742.1294
0.0001	-2.706	0.3508	0.3780	7,902.77	16,800.10	0.1632235	37,742.1277
0.00001	-2.706	0.3508	0.3781	7,902.77	16,800.10	0.1632236	37,742.1275
0	-2.706	0.3508	0.3781	7,902.77	16,800.10	0.1632236	37,742.1275

Table 3.2.1: Estimated parameters conditional on a fixed parameter *d*.

The table shows that below the value *d* = 0.001 neither the estimated parameters $\{a_2, \gamma_1, \gamma_2, \beta\}$ vary much, nor does the log ML function change considerably any more. Further, I computed the population average of the simulated probabilities that households do not hold a car, P_{sim}, and the population average of the simulated expectation value for driving car, $E_{sim}(X_2)$.[163] Both values P_{sim} and $E_{sim}(X_2)$ are almost the same for all values of *d*.[164] This is also a first indication that the choice of *d* does not critically influence the model’s outcomes. But most importantly, the maximum value of the

[163] Note that for notational reasons, in the table above both these functions were not noted as conditional on parameters and data, $P_{sim} \mid d, \hat{a}_2, \hat{\gamma}_1, \hat{\gamma}_2, \hat{\beta}, p_1, p_2, y, s$ and $E_{sim}\left(X_2 \mid d, \hat{a}_2, \hat{\gamma}_1, \hat{\gamma}_2, \hat{\beta}, p_1, p_2, y, s\right)$, what would have been actually correct. The exact functional form of both formulas is described in the section “simulation” that can be found previously in this Subchapter 3.2.

log-ML function $\log(P_{MLE})$ that yields the estimate does decrease rather dramatically in the range $0.1 \le d \le 0.8$, but in the range $d \le 0.1$ the decrease is rather small if *d* decreases. The results of Table 3.2.1 can also be illustrated in a diagram:

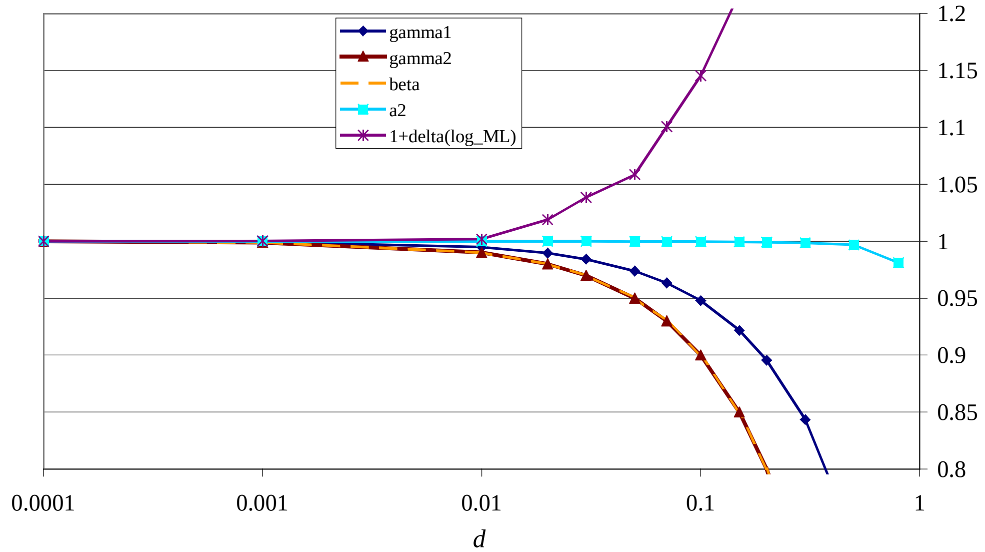


Figure 3.2.5: Estimated parameters conditional on a fixed parameter *d*.

Note that for this diagram the trajectories of all parameters are plotted as ratio to their values at $d = 0.000001$. Only the trajectory of the log-ML function is plotted as difference to its value at $d = 0.000001$ plus one. These trajectories show that all estimated parameter values are already very close to their value at $d = 0.000001$, when *d* is below the value $d = 0.01$. More importantly, the trajectory of the log-ML values shows that decreasing *d* further at $d = 0.001$ does not lower the value of the log-ML function any more significantly.

Of course, the arguments stated until now are not sufficient to justify that any choice of *d* below the value $d = 0.1$ is feasible or does not affect any result that can be computed by the model. Therefore, I examine if the results I am interested in, namely the simulated own price elasticity[165] and the income elasticity of the total demand of the economy for driving distance and the relative change of the share of carless households when driving costs or income changes, are affected by the choice of parameter *d*. Therefore, I compute all these changes for different values of *d* and estimate its corresponding parameters.

---

[164] Note, that the actual average kilometres of the population is 13,901 km per year and the share of carless households is 18.9%.

[165] In contrast to the section "Simulation" of this subchapter, I will compute the *relative* changes of driving demand and the share of households that do not own a car in order to insure comparability with other studies done in this field.

The simulated own price elasticity of driving demand and the relative change of the share of carless households when driving costs change are defined as follows:

$$\varepsilon_{P,p_{2n},sim,n} \mid \theta_1, d = \frac{\partial P_{sim,n} \mid \{\theta_1, d, p_1, p_{2n}, y_n, s_n\}}{\partial p_{2n}} \cdot \frac{p_{2n}}{P_{sim,n} \mid \{\theta_1, d, p_1, p_{2n}, y_n, s_n\}}, \tag{3.2.41}$$

$$\varepsilon_{E(X_2),p_{2n},sim,n} \mid \theta_1, d = \frac{\partial E_{sim,n}\left(X_2 \mid \{\theta_1, d, p_1, p_{2n}, y_n, s_n\}\right)}{\partial p_{2n}} \cdot \frac{p_{2n}}{E_{sim,n}\left(X_2 \mid \{\theta_1, d, p_1, p_{2n}, y_n, s_n\}\right)}, \tag{3.2.42}$$

with $\theta_1 = \{a_2, \gamma_1, \gamma_2, \beta\}$.[166]

Since I am interested in the relative changes on the level of the total population, I compute the following expressions:

$$\varepsilon_{P,p_2,sim} \mid \theta_1, d = \sum_{n=1}^{N} \frac{\partial P_{sim,n} \mid \{\theta_1, d, p_1, p_{2n}, y_n, s_n\}}{\partial p_{2n}} \cdot \frac{p_{2n}}{\sum_{n=1}^{N} P_{sim,n} \mid \{\theta_1, d, p_1, p_{2n}, y_n, s_n\}}, \tag{3.2.43}$$

$$\varepsilon_{E(X_2),p_{2n},sim} \mid \theta_1, d = \sum_{n=1}^{N} \frac{\partial E_{sim,n}\left(X_2 \mid \{\theta_1, d, p_1, p_{2n}, y_n, s_n\}\right)}{\partial p_{2n}} \cdot \frac{p_{2n}}{\sum_{n=1}^{N} E_{sim,n}\left(X_2 \mid \{\theta_1, d, p_1, p_{2n}, y_n, s_n\}\right)}, \tag{3.2.44}$$

with $\theta_1 = \{a_2, \gamma_1, \gamma_2, \beta\}$.

Note that these expressions, (3.2.43) and (3.2.44), are defined correspondingly with respect to the income and to the effect of a simulated change in household location.[167] For the following diagrams, I chose a finite change of marginal cost per kilometre of $\Delta p_{2n} = 0.01 \cdot p_{2n}$ and a finite change of income of $\Delta y_n = 0.01 \cdot y_n$.[168] All simulated changes were computed at $\theta_1 = \hat{\theta}_1$. As will be discussed further

[166] To simplify the notation, I omitted to state that $\varepsilon_{P,p_{2n},sim,n} \mid \theta_1, d$ and $\varepsilon_{E(X_2),p_{2n},sim,n} \mid \theta_1, d$ are conditional on the households' economic and sociodemographic variables $\{p_1, p_{2n}, y_n, s_n\}$. The elasticities are defined analogously to (3.2.41) and (3.2.42) with respect to the income y_n, $\varepsilon_{E(X_2),y_n,sim,n} \mid \theta_1, d$ and $\varepsilon_{E(X_2),y_n,sim,n} \mid \theta_1, d$.

[167] The exact definition of the income elasticity of driving demand and the relative change of the share of carless households when the income changes can be found in Appendix A3.2.

[168] The own price elasticity of the total population is computed as follows:

$$\varepsilon_{E(X_2),p_{2n},sim} \mid \hat{\theta}_1, d = \frac{\sum_{n=1}^{N} E_{sim,n}\left(X_2 \left\{\hat{\theta}_1, d, p_1, 1.01 \cdot p_{2n}, y_n, s_n\right\}\right)}{\sum_{n=1}^{N} E_{sim,n}\left(X_2 \mid \left\{\hat{\theta}_1, d, p_1, p_{2n}, y_n, s_n\right\}\right)} - 1.$$

below, this is a simplification, but for the purpose of this diagram the results computed this way are assumed to be sufficiently accurate.

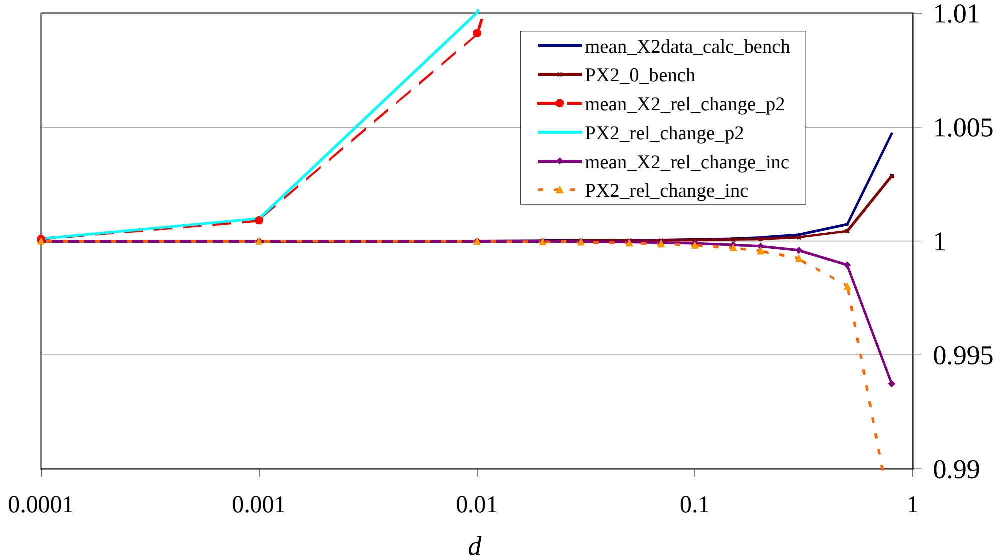


Figure 3.2.6: Simulated relative changes conditional on a fixed parameter *d*.

The results show that both the average simulated demand for distance and the average share of carless households[169] do not change below any value of $d < 0.1$. The same holds for the relative changes with respect to a relative change of income, whereas relative changes with respect to a relative change of the marginal costs of driving p_2 do not change below any value of $d < 0.001$, which is much lower. From this follows that to choose any value $d < 0.001$ is feasible, since in this range both the value of the ML function in the optimum and the marginal effects I am interested in do not vary with *d*.

Note that the average simulated demand for distance is about 16,800 km per year, and the share of carless households is about 16.3% for any $d < 0.001$, while as the average distance driven of the households is 13,901 km per year, and the share of carless households accounts for 18.9%; see Table 3.2.3. This difference may appear to be rather large, but since I am only interested in the change of these values when the income or the driving costs change, this "bias" does not matter. Further, this bias may be also due to the ignorance of fixed costs of car ownership. In Subchapter 3.3, I will present the same model including the fixed costs, and it turns out that due to this extension these two values - the average distance driven and the share of carless households - can be better approximated.

[169] These values were computed as follows:

$$P_{sim} \mid \{\theta_1, d, p_1, p_{2n}, y_n, s_n\} = \frac{1}{N} \cdot \sum_{n=1}^{N} P_{sim,n} \mid \{\theta_1, d, p_1, p_{2n}, y_n, s_n\} \text{ and}$$

$$E_{sim}\left(X_2 \mid \{\theta_1, d, p_1, p_{2n}, y_n, s_n\}\right) = \frac{1}{N}\sum_{n=1}^{N} E_{sim,n}\left(X_2 \mid \{\theta_1, d, p_1, p_{2n}, y_n, s_n\}\right).$$

It is important to note that so far the simulated elasticities were computed at the point estimates $\hat{\theta}_1 = \left\{\hat{a}_2, \hat{\gamma}_1, \hat{\gamma}_2, \hat{\beta}\right\}$. Note that here I use the notation $\hat{\theta}_1$ instead of $\hat{\theta}$, since $\hat{\theta}_1$ does not contain the parameter *d*. For reasons of completeness, I now restate the formulas (3.2.43) and (3.2.44) for the case where parameter value *d* is pre-chosen.

$$\tilde{\theta}_1 \sim \Phi\left(\hat{\theta}_1, \text{var}\left(\tilde{\theta}_1\right)\right), \qquad (3.2.45)$$

where $\hat{\theta}_1$ is the estimated value by MLE conditional on a fixed parameter value *d*, $\theta_1 = \{a_2, \gamma_1, \gamma_2, \beta\}$, and $\text{var}\left(\tilde{\theta}_1\right)$ can be approximated by the inverse Hessian matrix (assumed to be non-singular) of the log-likelihood function (3.2.27) at $\hat{\theta}_1$:

$$\text{var}\left(\tilde{\theta}_1\right) \cong \left[\frac{\partial^2 \log\left(L_{MLE}\left(\left(X_2 = x_{2n}\right)_{n=1,2,..,N} \mid \hat{\theta}_1, d, p_1, \ p_2, \ y_{n=1,2,..,N}, \ s_{n=1,2,..,N}\right)\right)}{\partial\theta_1 \partial\theta_1'}\right]^{-1}_{\theta=\hat{\theta}_1} \qquad (3.2.46).$$

The fact that the estimated parameter vector $\hat{\theta}_1 = \left\{\hat{a}_2, \hat{\gamma}_1, \hat{\gamma}_2, \hat{\beta}\right\}$ is a random vector has the same implications as already mentioned in the section "Simulation" of this subchapter, namely that the simulated marginal effects are also random variables too.

For a sample of 20,870 that corresponds to 100% of all households of the Bundesamt für Statistik micro-census dataset (2006a) and the parameter *d* = 0.0001, the following estimates result by use of MLE based on the ML function (3.2.27):

$$\hat{a}_2 = \underset{\substack{(118.6)\\ [65.25]}}{7736.5}, \quad \hat{\gamma}_1 = \underset{\substack{(0.00923)\\ [-294.4]}}{-2.7165}, \ \hat{\gamma}_2 = \underset{\substack{(0.0111)\\ [31.49]}}{0.3494}, \ \hat{\beta} = \underset{\substack{(0.00309)\\ [126.3]}}{0.3899}, \qquad (3.2.47a)$$

$$\text{var}\left(\hat{\theta}_1\right) = \text{var}\left(\left\{\hat{a}_2, \hat{\gamma}_1, \hat{\gamma}_2, \hat{\beta}\right\}\right) = \begin{pmatrix} 14058204 & 878.134 & -177.29 & -215.40 \\ 878.134 & 0.08516 & -0.0413 & -0.0146 \\ -177.29 & -0.0413 & 0.1231 & 0.00340 \\ -215.40 & -0.0146 & 0.00340 & 0.009528 \end{pmatrix} \cdot 10^{-3}, \qquad (3.2.47b)$$

where the values in brackets "(...)" in (3.2.47a) denote the standard deviations and the values in square brackets "[...]" the t-values of the corresponding estimates. The results show that all coefficients are highly significant. Further, the sign of the variable "rural" is positive, as was expected: Households located in a rural area have a stronger preference for driving cars.

In order to compute the distribution of the simulated relative changes as described in the paragraph "Simulation" of this subsection, I now take 80 draws[170] $\{\theta_{1,k}\}_{k=1..80}$ of the distribution (3.2.45) for which the parameters are defined by (3.2.47). For each of these draws, I computed the simulated relative changes of driving demand and share of carless households according to (3.2.43) and (3.2.44). For example, the distribution of the elasticity of driving demand is as follows:

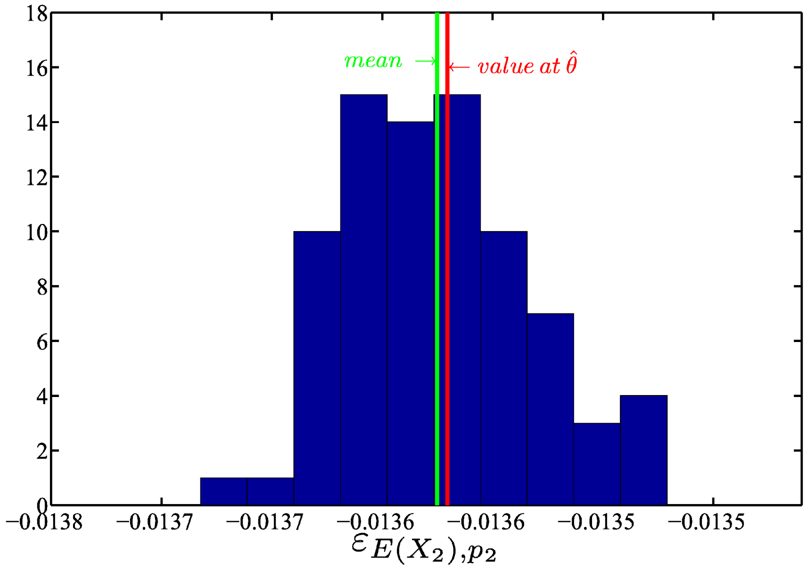


Figure 3.2.7: Elasticities of the mean of driving demand with respect to the driving costs.

This result shows that the variation of the simulated elasticities is rather small. All realizations are within an interval of the width of 0.0004, which is less than +/- 2% of the magnitude of the elasticity. This means that the simulated elasticities do not vary much when being computed for random draws of the distribution of the estimated parameters. Further, the mean value of the simulated values is very close to the value conditional on the point estimate $\varepsilon_{E(X_2),p_2,sim} \mid \hat{\theta}_{1,k}, d$, denoted by "at value $\hat{\theta}_1$" in the diagram above. Therefore, there seems no doubt that the delta method is feasible.[171] The mean of the simulated value, denoted by "mean" in the diagram above, is an approximation of the value I am actually interested in, namely the expectation value of $\varepsilon_{E(X_2),\,p_2} \mid d = E_{\tilde{\theta}_1}\left(\varepsilon_{E(X_2),\,p_2} \mid \tilde{\theta}_1, d\right)$. This value can approximately be calculated by the mean of the simulated values:

$$\varepsilon_{E(X_2),p_2} \mid d = E_{\tilde{\theta}_1}\left(\varepsilon_{E(X_2),p_2} \mid \tilde{\theta}_1, d\right) \approx \frac{1}{K} \cdot \sum_{k=1}^{K} E_{\tilde{\theta}_{1,k}}\left(\varepsilon_{E(X_2),p_2} \mid \tilde{\theta}_{1,k}, d\right), \tag{3.2.48}$$

---

[170] A Gibbs sampler was used to generate the sample. The first 50 out of the 130 samples were deleted, since these usually deviate from the true distribution when using the Gibbs sampling routine. This effect is called "burn-in" effect.

[171] For more detailed information see the last paragraph in the section "Simulation" of this subchapter.

where $\{\theta_{1,k}\}_{k=1..K}$ is a sequence of the realizations of the random distribution defined by (3.2.45).

The following diagram shows the simulation results for the relative change of the share of carless households with respect to a relative change of driving costs as defined by (3.2.43). Note that all subsequent diagrams are always based on the same draws $\{\theta_{1,k}\}_{k=1..80}$.

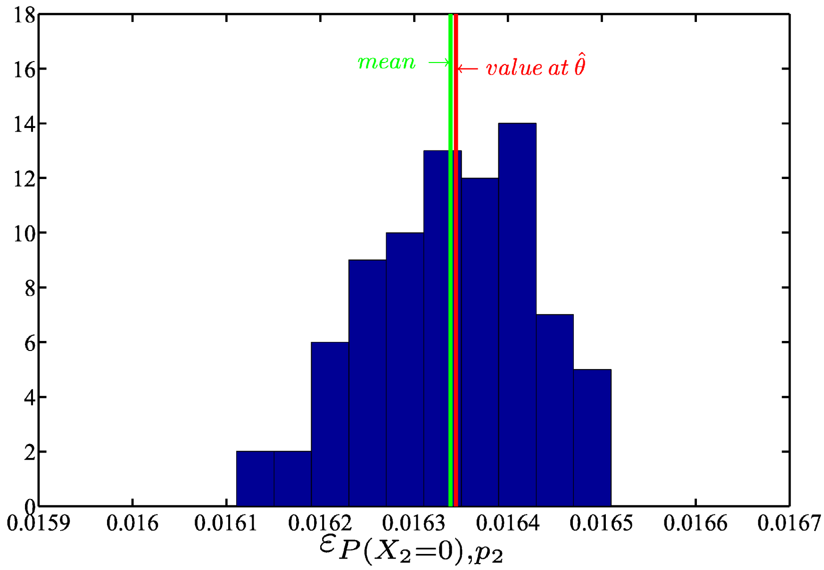


Figure 3.2.8: Distribution of the relative change of probability of car ownership with respect to the driving costs.

In this case, the same result holds: The variance is relatively small with respect to the level of the change of the share of carless households. Also, the value of the point estimate $\hat{\theta}_1$ "at value $\hat{\theta}_1$" is similar to the mean of the simulated values denoted by "mean" in the diagram above. Similar results yield for the elasticities with respect to income and for the relative changes of driving demand and the simulated change of share of carless households when a change of household location is simulated; see Appendix A3.3. Therefore, I conclude that the delta method is feasible and that it is not necessary to use a bootstrapping method in order to compute the distribution of the relative changes of driving demand and the share of carless households.[172] Further, it is sufficient to consider the point estimates if one is only interested in the expectation value of relative changes when prices and income change or a change of the households' location is simulated. I now summarize the results:

[172] To test whether it really does not make any difference, I also applied a parametric bootstrapping method. The resulting standard deviations were similar to those computed by the method presented in this subchapter, see Appendix A3.6.

$\varepsilon_{E(X_2),p_2} \mid d=0.0001=\underset{(0.0042703)}{\underset{\langle -1.362\rangle}{-1.3625}}$	$\varepsilon_{P(X_2=0),p_2} \mid d=0.0001=\underset{(0.00884)}{\underset{\langle 1.6345\rangle}{1.6339}}$
$\varepsilon_{E(X_2),p_{fuel}} \mid d=0.0001=\underset{(0.00195)}{\underset{\langle -0.62243\rangle}{-0.62266}}$	$\varepsilon_{P(X_2=0),p_{fuel}} \mid d=0.0001=\underset{(0.00404)}{\underset{\langle 0.74697\rangle}{0.74669}}$
$\varepsilon_{E(X_2),y} \mid d=0.0001=\underset{(0.004227)}{\underset{\langle 1.3483\rangle}{1.3488}}$	$\varepsilon_{P(X_2=0),y} \mid d=0.0001=\underset{(0.0087525)}{\underset{\langle -1.6181\rangle}{-1.6176}}$
$\Delta_{\ln(E(X_2)),urban\to rural} \mid d=0.0001=\underset{(0.017023)}{\underset{\langle 0.50485\rangle}{0.50276}}$	$\Delta_{\ln(P(X_2=0)),urban\to rural} \mid d=0.0001=\underset{(0.010019)}{\underset{\langle -0.45479\rangle}{-0.45479}}$
$\Delta_{\ln(E(X_2)),rural\to urban} \mid d=0.0001=\underset{(0.0076374)}{\underset{\langle -0.34175\rangle}{-0.34065}}$	$\Delta_{\ln(P(X_2=0)),rural\to urban} \mid d=0.0001=\underset{(0.032088)}{\underset{\langle 0.80436\rangle}{0.80022}}$

Table 3.2.2: Simulated parameters conditional on a fixed parameter *d*.[173]

The results show that the income elasticity is about 1.35, and the fuel price elasticity is about -0.62. In this micro-census data of the Bundesamt für Statistik (2006a), which I am using, the observed data can be considered as the outcome of a long-term equilibrium, since the households' decision between holding or not holding a car is the result of a long-term decision. Further, the fuel and car prices did not change much in the years before the survey was conducted, so the households' decision was based on fuel and car prices that could actually be observed in the year 2005. Further, it can be shown that the effect of the change in the share of carless households does not have a strong effect on the expected driving demand accounts for about one third of the total effect of the change in fuel prices on driving demand.[174] Therefore, the effect of selling the car when fuel prices rise has quite a high impact on the fuel demand according the results of this model.

---

[173] Households' individual driving costs per kilometre were computed as 0.1601 + 0.077825 * (average fuel price during the period the household was driving). The units of these values are: kilometres and CHF. Since the average fuel price in 2005 was about CHF 1,729, this yields fuel costs of 0.077825 litres/km * 1.729 CHF/litre = 0.1346 CHF/km. Thus, the share of fuel costs on the total driving costs are about 0.1346 CHF/km / (0.1346 CHF/km + 0.1601 CHF/km) = 45.7%. Therefore, an increase in fuel price by 1% increases the total cost of an additional kilometre only by 0.457. From this follows that the fuel price elasticity is equal to 0.457 times the elasticity of the driving demand with respect to the driving costs. To compute these values in this way is an approximation. In fact, a simulation with $p_{2n}'=0.01601+0.077825\cdot p_{fuel,n}\cdot(1+0.01)$ should be processed. But since the fuel prices vary only very little, I expect only a very small difference in the results, such that a computation by use of this more accurate procedure is not worth doing it.
The values in brackets "$\langle ... \rangle$" are the values at the point estimate of $\hat{\theta}_1$, and the values in brackets "$(...)$" are the standard deviations of the forecasted elasticities.

[174] The elasticity with respect to the marginal driving costs can be decomposed in two parts: The first part consist only of the effect that households that drive cars drive less if the marginal driving costs increase. The second part consists of the effect that the expected driving demand decreases because the households' probability not to own a car increases. I now show how the share of the latter effect on the total effect can be computed. I start with computing the marginal effect of a change of the marginal driving costs on the expected driving demand:

$$\frac{\partial E(X_2)}{\partial p_2}=\frac{\partial\left(1-P(X_2=0)\right)\cdot E(X_2\mid X_2>0)}{\partial p_2}=\left(1-P(X_2=0)\right)\cdot\frac{\partial E(X_2\mid X_2>0)}{\partial p_2}-\frac{\partial P(X_2=0)}{\partial p_2}\cdot E(X_2\mid X_2>0).$$

Comparison with results of other studies

Comparing these elasticities to the results of international studies as presented in Table 1.3.4[175] the elasticities with respect to driving demand found by this model are much too high in magnitude, namely by about factor two for both the fuel price and the income elasticities. Note, that difference to the values found by Baranzini et al. (2009) is even larger. To compare the effects on car ownerships the results above have to be transformed by multiplying by the factor -0.197,[176] so that $\varepsilon_{P(X_2>0),p_2}=0.147$ and $\varepsilon_{P(X_2>0),y}=-0.318$. Note, that the measures $\varepsilon_{P(X_2>0),p_2}$ and $\varepsilon_{P(X_2>0),y}$ represent lower bounds of the elasticity with respect to car ownership. These two values are about by factor two smaller in magnitude that the ones presented in Table 1.3.4. Since the measures $\varepsilon_{P(X_2>0),p_2}$ and $\varepsilon_{P(X_2>0),y}$ represent lower bounds it cannot be concluded whether values of these measures are unrealistic. But, in the case of the values of $\varepsilon_{E(X_2),y}$ and $\varepsilon_{E(X_2),p_2}$ the conclusion is clear: The values this model yields are not realistic.

One explanation that the income elasticity of driving demand is far too high is could be that in my model the difference between the expectation value of driving distance X_2 and the average of the observed data is increasing with the income of the corresponding household categories. As shown in Appendix A3.3, this effect could account for an overestimate of the income elasticity of about 0.58. So, if I corrected for that effect, then the income elasticity would be 0.77 instead of 1.35. In Appendix A3.4 I also argue that this difference is caused by the too heavy tail of the density function of X_2 for household segments with incomes higher than CHF 108,000. A possibility to overcome this problem is also sketched in Appendix A3.4. But it has to be noted that it is not sure if correcting the value of the

---

Multiplying this expression by $\frac{p_2}{E(X_2)}$ and using $E(X_2)=\left(1-P(X_2=0)\right)\cdot E(X_2\mid X_2>0)$ yields:

$$\varepsilon_{E(X_2),p_2}=\frac{\partial E(X_2)}{\partial p_2}\cdot\frac{p_2}{E(X_2)}=\frac{\partial E(X_2\mid X_2>0)}{\partial p_2}\cdot\frac{p_2}{E(X_2\mid X_2>0)}-\frac{\partial P(X_2=0)}{\partial p_2}\cdot\frac{p_2}{1-P(X_2=0)}=\varepsilon_{E(X_2\mid X_2>0),p_2}-\frac{\varepsilon_{P(X_2=0),p_2}}{P(X_2=0)^{-1}-1}.$$

Plugging in the average simulated values of the whole population yields:

$$\varepsilon_{P(X_2=0),p_2}\cdot\left(P(X_2=0)^{-1}-1\right)^{-1}=0.016339\cdot\left(0.163^{-1}-1\right)^{-1}=0.0032.$$

Decomposing these effects for each household and taking the average of these effect - which would actually be the correct method - yields $\frac{1}{N}\cdot\sum_{n=1}^{N}\varepsilon_{P(X_2=0),p_2,n}\cdot\left(P(X_2=0)_n^{-1}-1\right)^{-1}=0.004221$. Therefore, the contribution of the change of share of carless households to the change of the total driving demand is $0.0032/0.01362=0.235$ or $0.0042/0.01362=0.31$, respectively, which is less than one third of the total effect in both cases. Considering the income elasticity of the driving demand, the effect of households switching from car owning to not owning a car contributes 0.42% of 1.35%, which is 0.31 of the the total effect.

175 See the results of Johansson and Schipper (1997).

176 Note that $\varepsilon_{P(X_2>0),p_2}=\frac{\partial P(X_2>0)}{\partial p_2}\cdot\frac{p_2}{P(X_2>0)}=-\frac{\partial P(X_2=0)}{\partial p_2}\cdot\frac{p_2}{P(X_2=0)}\cdot\frac{P(X_2=0)}{P(X_2>0)}=-\varepsilon_{P(X_2>0),p_2}\cdot\frac{P(X_2=0)}{1-P(X_2=0)}$.

For $P(X_2=0)$ I plug in the average simulated values of the dataset, 0.164, see Table 3.2.3.

income elasticity by 0.58 would be correct.[177] But at least, the observation that the difference between the expectation value of driving distance X_2 and the average of the observed data is increasing with the income of the corresponding household categories leads to the presumption that the simulated income elasticity is rather too high than too low.

It will be interesting whether these results change if the model that includes fixed costs of car ownership is applied, see Subchapter 3.2.

Model quality

So far I have shown that if I consider the variance of the estimated parameters the variance of the distribution of simulated elasticities will be very small. Further, the results for income and fuel price elasticities are within the range found by other surveys using data from other countries. I now want to give more reasons why the model I specified reflects the micro-census data of the Bundesamt für Statistik (2006a) quite well. First, I show that the average simulated values are rather close to the mean values of the data. The average simulated values are defined as follows:

$$P_{sim,avg} = \frac{1}{N} \cdot \sum_{n=1}^{N} P_{sim,n} \mid \hat{\theta}_1, d, p_1, \ p_{2n}, y_n, s_n \ , \tag{3.2.49}$$

$$X_{2,sim,avg} = \frac{1}{N} \cdot \sum_{n=1}^{N} E_{sim,n}\left(X_2 \mid \hat{\theta}_1, d, p_1, \ p_{2n}, y_n, s_n \right), \tag{3.2.50}$$

where $\hat{\theta}_1$ is the estimated parameter vector $\theta_1 = \{a_2, \gamma, \beta\}$. The formulas $P_{sim,n} \mid \hat{\theta}_1, d, p_1, \ p_{2n}, y_n, s_n$ and $E_{sim,n}\left(X_2 \mid \hat{\theta}_1, d, p_1, \ p_{2n}, y_n, s_n \right)$ are defined in (3.2.28) and (3.2.29).

Plugging in $\hat{\theta}_1$ that yielded at d = 0.0001 and using the complete micro-census dataset yields the following values:

Empirical values	Simulated values
$mean\left(x_{2,n} = 0\right) = 18.90\%$	$P_{sim,avg} = 16.43\%$
$mean\left(x_{2,n}\right) = 13{,}890$	$X_{2,sim,avg} = 16{,}915$

Table 3.2.3: Average simulated and average empirical driving distance and probability of not owing a car.

[177] In fact, the value of 0.77 for the income effect would result, if one simply computed the average of the driving distance for each income category and then computed the increase of this average driving distance with respect to the increase of income.

The comparison of these values shows that the simulated value of the average probability of households is by 2.47% too low. This is 13% less than it actually is. The average simulated expected value of car driving is 3,025 km higher than the actual average value, which is 22% more than it actually is.[178] This seems rather a lot, but since I am only interested in the changes of these simulated values when I compute elasticities and changes in probabilities, the absolute difference with respect to the empirical values should not matter, as I already mentioned in the paragraph "Results". But I want to examine the cause of these differences by looking at a diagram where both the histogram of the actual driving demand and the density that is determined by the model are illustrated. I do this for the households with an annual income of CHF 60,000 located in urban areas.

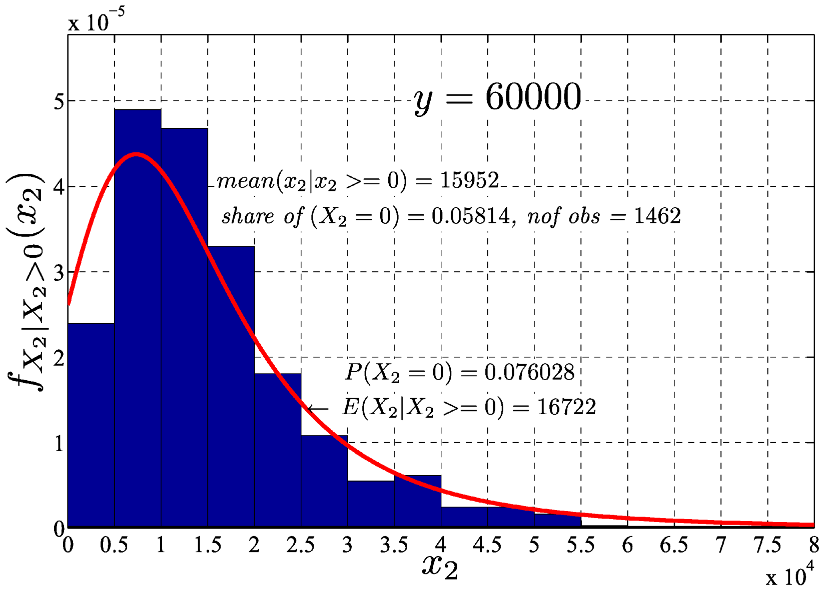


Figure 3.2.9: Histogram of households living rural areas with an income of CHF 60,000.

This diagram shows that the conditional probability density function (pdf) $f_{X_2|(X_2>0)}$ determined by the model is well adapting the actual distribution of the driving demand. The difference occurs mainly, because the conditional pdf $f_{X_2|(X_2>0)}$ has too much mass at the upper tail. This implies that the simulated expected value of driving demand $E(X_2)$ is higher than the observed data. A second difference is that the conditional pdf has too much density at values below 5,000 km. This might be due to the following reason: Consider the non-conditional density, which would be strictly positive for some negative x_2. Then, the area below this density for negative x_2 would correspond to $P(X_2 = 0)$.[179] From this follows that if the non-conditional density has heavier tails, the probability $P(X_2 = 0)$

[178] (16.43-18.9)/18.9 = -0.1307, (16,915-13,890)/13,890 = 0.2178.

[179] For an illustration of this argumentation see Figures A3.5.3 and A3.5.4 in Appendix A3.5.

increases. Compared to the resulting density as illustrated above, the density would better adapt the distribution of positive values x_2 if the the tails were less heavy. But then the probability $P(X_2 = 0)$ that is already too small would become even smaller. Therefore, there is a trade off between adapting the shape of the distribution of positive values x_2 and to approximate the probability $P(X_2 = 0)$. Since the ML function values both these criteria, the resulting probability functions $P(X_2 = 0)$ does neither approximate the share of carless households $P(X_2 = 0)$, nor does the density adapt the distribution of positive values x_2 very closely. But the ML estimation procedure yields a density function that are a good compromise when regarding these two conflicting criteria.

Visualization of simulated changes

Before I conclude this subchapter, I wish to visualize a simulated rise in income on the conditional pdf $f_{X_2|(X_2>0)}(x_2)$ in order to provide some intuition about the simulation procedure.

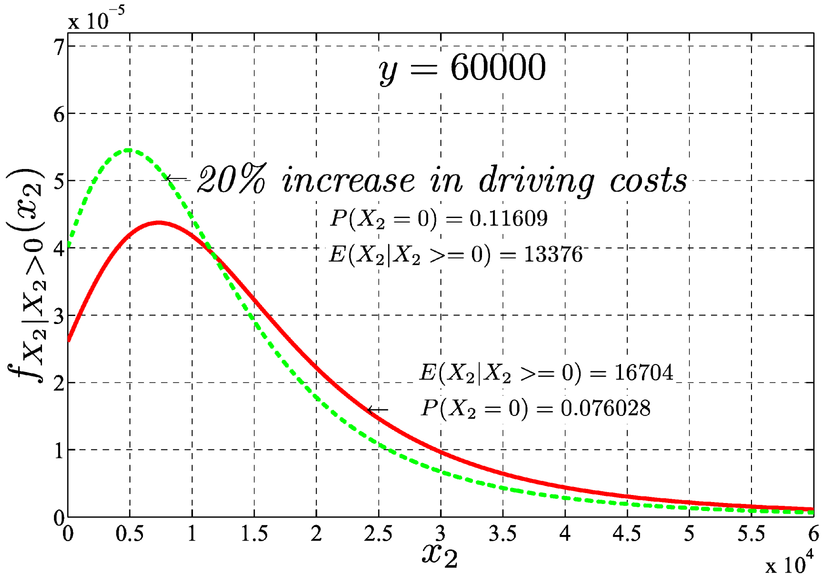


Figure 3.2.10: Change of conditional probability density function if driving costs increase.[180]

This diagram shows that if the driving costs increase, the density at lower annual driving distances increases, and consequently the density at high annual driving distances decreases. Therefore, the expected driving distance deceases, given a household owns a car. At the same time, the probability of not owning a car increases. All these effects are quite intuitive: If the driving costs increase, we expect the household to drive shorter distances, and for some households we expect that they will decide to

[180] Note that here the value $E(X_2 | X_2 > 0)$ was computed at the mean value of p_2, $mean(p_2)$, while the value corresponding to Figure 3.2.9 is the mean value of $E(X_2 | X_2 > 0)$ that was computed for each individual household. The difference in the results is due to the fact that not all households have the same marginal costs of driving p_2.

get rid of their car. The illustration of an increase in income and a change of household location from a rural to an urban area can be found in Appendix A3.5.

In the following subchapter I will show that when incorporating the fixed costs of holding a car in the model, the distribution can be better adapted to both the observed distribution of positive values x_2 and the probability $P(X_2 = 0)$. When we take a look at the density function, the consequence of the inclusion of the fixed cost is that the distribution is restricted to zero at a positive value $x_{2,crit.}$, the so-called "critical" level of x_2. This means that there is less mass at the tails of the distribution necessary to yield a certain probability $P(X_2 = 0)$. Therefore the simulated expectation of driving demand $E(X_2)$ will be closer to the average values x_2 of observed data. Therefore, the incorporation of the fixed costs of car ownership should help to overcome the problems connected to the model presented in this subchapter.

3.3 Model with two goods (good two with fixed costs)

In this subchapter, I shall present an extended Multiple Discrete-Continuous Extreme Value Model (MDCEV). To this end, the common MDCEV presented in the previous subchapter was extended by capturing the fixed costs when a household decides to own and drive a car. This extension allows us to map the economic decision of a household more realistically. So far, I have only found a model by De Jong (1990) that incorporates fixed costs and that is based on a micro-economic model framework mapping the economic behaviour of a single household.[181] So far, studies have only captured both car stock and car use based on time series on aggregate data of car stock and fuel consumption.[182] In contrast, the model I present here is based on the behaviour of individual households; cross-sectional data is used to estimate the model parameters. This enables us to map household properties to explain individual households' decisions on car ownership and use. This model enables me to compute the effects of fuel tax on the proportion of carless households and on the aggregate driving distance. An important feature of this model that included fixed costs is that I can simulate the effect of car taxes on the proportion of carless households and the aggregate driving distance. I expect this tax will mainly impact the proportion of carless households. Since the proportion of carless households is rather high in the case of Switzerland – 18.90% in 2005 versus less than 9% in the USA – given the average income of a Swiss household and the comparatively low degree of urbanization, a change in the proportion of carless households may have a significant impact on the total driving demand.[183] Using this model, I will compare the effects of a tax on fuel and a tax on car ownership. It will transpire that the impact of these two taxes on car ownership are approximately the same per unit of tax revenue, whereas the effect of the fuel tax on the aggregate driving distance is much larger than that of a tax on car ownership. It also emerges that this model yields density functions that successfully adapt the actual data of the household segments as presented in Figures A3.4.1-3.4.10. Further, I will show that both models with and without fixed costs can explain the co-movement of the proportion of households owning a car and the aggregate driving demand, as observed in the data from different countries in recent decades.[184]

---

[181] In Chapter 1, I justify why I decide against using the model by De Jong (1990).

[182] See, for instance, Johansson and Schipper (1997).

[183] For instance, in the USA the proportion of carless households was only about 8% in 2000/01. In contrast, Germany has the same level of carless households, namely 19%; see Bühler and Kunert (2008: 10).

[184] The co-movement of driving demand and the proportion of households that own a car is presented in Subchapter 1.2.

The model

The solution is based on the procedure described in Subchapter 3.1. The microeconomic decision corresponds to the illustration in Figure 3.1.3. Recall that the key decision taken by households is whether it should own and use a car and bear the fixed costs or whether it should save the fixed costs and spend all its income on good one, which contains all goods apart from car driving. Formally, this corresponds to comparing the maximum utility of choice set $S_1 = \{\ \}$, where only good one can be consumed to the maximum utility of choice set $S_2 = \{2\}$, where both goods can be consumed as presented in the paragraph "Illustration of the maximization principle in the case of two goods" in Subchapter 3.1. I shall now compute the maximum utility levels for these two cases; the concrete parametrized utility function is given. Recall that the parametrized utility function is

$$U = \frac{u_1(X_1) + u_2(X_2)}{\exp(m_1 + \xi_1)} = (X_1 + a_1)^d + \exp(m + \beta \cdot \varsigma) \cdot (X_2 + a_2)^d, \qquad (3.3.1)$$

as already stated in (3.2.4).

I start computing the utility level for case 1, u_{S_1} corresponding to choice set $S_1 = \{\ \}$. Since in this case only good one is consumed and the household does not have to bear the fixed costs, all income is spent on good one:

$$X_1 = \frac{y}{p_1}. \qquad (3.3.2)$$

Utility u_{S_1} is therefore

$$u_{S_1} = \left(\frac{y}{p_1} + a_1\right)^d + \exp(m + \beta \cdot \varsigma) \cdot (a_2)^d \cdot \qquad (3.3.3)$$

I shall now compute the utility for the case where both goods can be consumed. This problem is equivalent to the problem in Subchapter 3.2, which was solved by solving Lagrangian (3.2.6). In this case, I need to compute the Marshallian demand functions, because these have to be plugged into utility function (3.3.1) to compute the indirect utility function. To this end, I set the first-order conditions (3.2.11) and (3.2.12) as being equal:

$$\frac{d}{p_1} \cdot \frac{1}{(X_1 + a_1)^{1-d}} = \frac{d}{p_2} \cdot \exp(m + \beta \cdot \varsigma) \cdot \frac{1}{(X_2 + a_2)^{1-d}}, \text{ with } m = \gamma \cdot s. \qquad (3.3.4)$$

For budget restriction $y = p_1 \cdot X_1 + p_2 \cdot X_2 + k_2$, it follows that

$$X_1 = \frac{y - k_2 - p_2 \cdot X_2}{p_1}. \tag{3.3.5}$$

Plugging (3.3.5) into (3.3.4) and solving for X_2 yields the following Marshallian demand function:[185]

$$X_2 = \max\left(x_2\left(y-k_2,\ p_1,\ p_2,\ A,\ a_1,\ a_2\right),0\right), \tag{3.3.6}$$

with $x_2\left(y-k_2,\ p_1,\ p_2,\ A,\ a_1,\ a_2\right) = \dfrac{A \cdot \dfrac{y-k_2}{p_1} - a_2 + A \cdot a_1}{1 + A \cdot \dfrac{p_2}{p_1}}$ and $A = \left(\dfrac{p_1}{p_2} \cdot \exp\left(m + \beta \cdot \varsigma\right)\right)^{\frac{1}{1-d}}$.

Note that $x_2\left(y-k_2,\ p_1,\ p_2,\ A,\ a_1,\ a_2\right)$ depends on the random term ς and that a value $\varsigma = \varsigma_0$ exists such that

$$x_2\left(y-k_2,\ p_1,\ p_2,\ A,\ a_1,\ a_2\right)\big|_{\varsigma \le \varsigma_0} \le 0 \text{ and } x_2\left(y-k_2,\ p_1,\ p_2,\ A,\ a_1,\ a_2\right)\big|_{\varsigma > \varsigma_0} > 0. \tag{3.3.7}$$ [186]

Plugging (3.3.5) and (3.3.6) into utility function (3.3.1) yields

$$u_{S_2} = \left(\frac{y-k_2}{p_1} - \frac{p_2}{p_1} \cdot \frac{A \cdot \frac{y-k_2}{p_1} - a_2 + A \cdot a_1}{1 + A \cdot \frac{p_2}{p_1}} + a_1\right)^d + \exp\left(m + \beta \cdot \varsigma\right) \cdot \left(\frac{A \cdot \frac{y-k_2}{p_1} - a_2 + A \cdot a_1}{1 + A \cdot \frac{p_2}{p_1}} + a_2\right)^d, \tag{3.3.8}$$

provided that $\varsigma \ge \varsigma_0$ so that $x_2\left(y-k_2,\ p_1,\ p_2,\ A,\ a_2\right) > 0$.

---

[185] This expression can be computed as follows:

$$\frac{d}{p_1} \cdot \frac{1}{\left(X_1 + a_1\right)^{1-d}} = \frac{d}{p_2} \cdot \exp\left(m + \beta \cdot \varsigma\right) \cdot \frac{1}{\left(X_2 + a_2\right)^{1-d}} \Leftrightarrow \left(\frac{X_2 + a_2}{X_1 + a_1}\right)^{1-d} = \frac{p_1}{p_2} \cdot \exp\left(m + \beta \cdot \varsigma\right) \Leftrightarrow \frac{X_2 + a_2}{X_1 + a_1} = \left(\frac{p_1}{p_2} \cdot \exp\left(m + \beta \cdot \varsigma\right)\right)^{\frac{1}{1-d}} \Leftrightarrow$$

$$\Leftrightarrow X_2 = \left(\frac{p_1}{p_2} \cdot \exp\left(m + \beta \cdot \varsigma\right)\right)^{\frac{1}{1-d}} \cdot X_1 - a_2 + \left(\frac{p_1}{p_2} \cdot \exp\left(m + \beta \cdot \varsigma\right)\right)^{\frac{1}{1-d}} \cdot a_1.$$ Plugging in $X_1 = \frac{y-k_2}{p_1} - \frac{p_2 \cdot X_2}{p_1}$ yields:

$$\Leftrightarrow X_2\left(1 + \left(\frac{p_1}{p_2} \cdot \exp\left(m + \beta \cdot \varsigma\right)\right)^{\frac{1}{1-d}} \frac{p_2}{p_1}\right) = \left(\frac{p_1}{p_2} \cdot \exp\left(m + \beta \cdot \varsigma\right)\right)^{\frac{1}{1-d}} \cdot \frac{y-k_2}{p_1} - a_2 + \left(\frac{p_1}{p_2} \cdot \exp\left(m + \beta \cdot \varsigma\right)\right)^{\frac{1}{1-d}} \cdot a_1 \Leftrightarrow$$

$$\Leftrightarrow X_2 = \left(\left(\frac{p_1}{p_2} \cdot \exp\left(m + \beta \cdot \varsigma\right)\right)^{\frac{1}{1-d}} \cdot \frac{y-k_2}{p_1} - a_2 + \left(\frac{p_1}{p_2} \cdot \exp\left(m + \beta \cdot \varsigma\right)\right)^{\frac{1}{1-d}} \cdot a_1\right) \cdot \left(1 + \left(\frac{p_1}{p_2} \cdot \exp\left(m + \beta \cdot \varsigma\right)\right)^{\frac{1}{1-d}} \frac{p_2}{p_1}\right)^{-1}.$$

[186] For proof, see Appendix A3.7.

Since composite good one is an essential good, again, I choose $a_1 = 0$, since this ensures that X_1 is greater than zero.[187] (3.3.6), (3.3.7) and (3.3.3) therefore simplify to:

$$X_2 = \max\left(x_2\left(y-k_2,\ p_1,\ p_2,\ A,\ a_2\right),0\right), \tag{3.3.9}$$

with $x_2\left(y-k_2,\ p_1,\ p_2,\ A,\ a_2\right) = \dfrac{A\cdot\dfrac{y-k_2}{p_1}-a_2}{1+A\cdot\dfrac{p_2}{p_1}}$ and $A = \left(\dfrac{p_1}{p_2}\cdot\exp\left(m+\beta\cdot\varsigma\right)\right)^{\frac{1}{1-d}}$.

$$u_{S_2} = \left(\frac{y-k_2}{p_1}-\frac{p_2}{p_1}\cdot\frac{A\cdot\dfrac{y-k_2}{p_1}-a_2}{1+A\cdot\dfrac{p_2}{p_1}}\right)^d + \exp\left(m+\beta\cdot\varsigma\right)\cdot\left(\frac{A\cdot\dfrac{y-k_2}{p_1}-a_2}{1+A\cdot\dfrac{p_2}{p_1}}+a_2\right)^d \tag{3.3.10}$$

and (3.3.3) to

$$u_{S_1} = \left(\frac{y}{p_1}\right)^d + \exp\left(m+\beta\cdot\varsigma\right)\cdot\left(a_2\right)^d. \tag{3.3.11}$$

Note that parameter a_2 is always greater than zero, i.e. $a_2 > 0$.

In the following, I discuss under which condition a household would choose to consume only composite good one and under which condition it would own and use a car. This decision is mapped by the model by comparing utilities u_{S_1} and u_{S_2}: if $u_{S_1} \geq u_{S_2}$, the household will decide not to own a car, and if $u_{S_1} < u_{S_2}$, the household will decide to keep a car. Assuming that all parameters are fixed, the decision depends on the value of the random term ς, since both u_{S_1} and u_{S_2} depend on that value. It can be shown that $\varsigma = \varsigma_c$ exists such that

$$u_{S_2} - u_{S_1}\,|_{\varsigma\geq\varsigma_c} \geq 0 \text{ and } u_{S_2} - u_{S_1}\,|_{\varsigma<\varsigma_c} < 0. \tag{3.3.12}$$ [188]

It can also be shown that

$$\varsigma_c > \varsigma_0. \tag{3.3.13}$$

---

[187] For proof, see Subchapter 3.1.

[188] For proof, see Appendix A3.8.

This implies that

$$x_2\left(y-k_2,\, p_1,\, p_2,\, A,\, a_1,\, a_2\right)\big|_{\varsigma=\varsigma_c} > 0 \qquad (3.3.14)$$

since $\partial x_2(\bullet)/\partial\varsigma > 0$ and $x_2\left(y-k_2,\, p_1,\, p_2,\, A,\, a_1,\, a_2\right)\big|_{\varsigma=\varsigma_0} = 0$.

This fact can also be illustrated as follows:

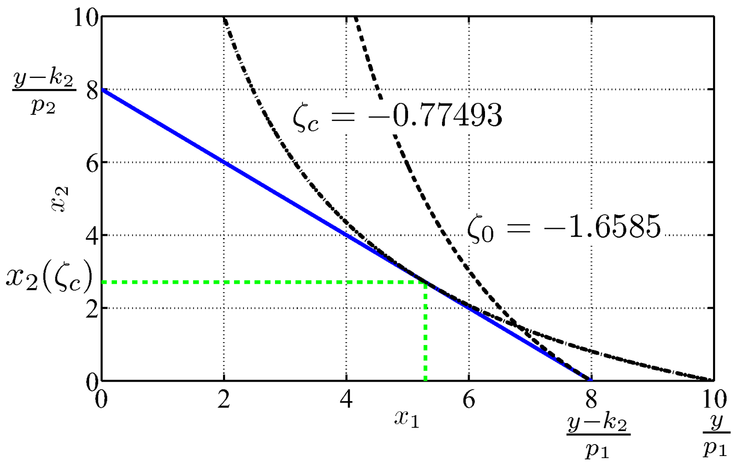


Figure 3.3.1: Indifference curves for two different utility functions.[189]

The indifference curves are illustrated in this diagram. The first curve corresponds to the critical relative preference $\varsigma = \varsigma_c$. If a household has this level of preference, it will be indifferent to the question of not owning a car and spending all its income on good one, which means consuming $\left(y/p_1, 0\right)$, or of consuming the optimal value of good one and driving a car $\left(\left(y-k_2-p_2\cdot x_2(\varsigma_c)\right)/p_1,\, x_2(\varsigma_c)\right)$. The illustration above also suggests why driving distances below $x_2(\varsigma_c)$ are not rational, even for very low relative preference parameters $m+\beta\cdot\varsigma$. This implies that the minimal driving distance $x_2(\varsigma_c) = x_2\left(y-k_2,\, p_1,\, p_2,\, A,\, a_1,\, a_2\right)\big|_{\varsigma=\varsigma_c}$ is always greater than zero, $x_2(\varsigma_c) > 0$.[190] Further, the diagram above shows the indifference curve corresponding to the case where the optimal level of consumption x_2 goes to zero for choice set S_2. It is also shown that $\varsigma_c > \varsigma_0$ for this specific case, namely $-0.775 > -1.659$.

---

[189] Note that this diagram is based on the following parameters: $a_1 = 0,\, a_2 = 4,\, m = 0,\, d = 0.05,\, k_2 = 0,\, y = 10,\, p_1 = 1,\, p_2 = 1$.

[190] There is one exception: if the utility function corresponds to the Leontief type, the critical preference $\varsigma_c = \varsigma_0$ is equal to $x_2 = 0$. Since in the Leontief case, the iso-utility function is vertical for any value $x_2 > 0$, the point $\left(\left(y-k_2\right)/p_1,\, x_2 = 0\right)$ is therefore located on the iso-utility function, and thus yields the same utility. This means that the household ignores fixed costs which is implausible. Thus, the Leontief case is infeasible when the model maps fixed costs. Note that my assumption on the utility function corresponds to Bhat(2005: 686) and does only allow for the Cobb-Douglas case in the extreme case $d \to 0$. Later, Bhat introduced a specification that allows for the Leontief case, Bhat(2008: 277). The empirical results will show that the restriction to the Cobb-Douglas $0 < d < 1$ is not binding and thus does not have any impact on the results.

Derivation of the Maximum Likelihood function

The key difference between this model that includes the fixed costs of car driving and the model that neglects these fixed costs, as described in Subchapter 3.2, is that probability $P\left(X_2=0\right)$ has to be computed differently. The demand for driving is always zero if $u_{S_2}<u_{S_1}$. As shown in (3.3.12), $u_{S_2}<u_{S_1}$ if and only if $\varsigma<\varsigma_c$. Therefore,

$$P\left(X_2=0\,|\,\theta,p_1,p_2,y,k_2,s\right)=P\left(\varsigma<\varsigma_c\,|\,\theta,p_1,p_2,y,k_2,s\right)=F_\varsigma\left(\varsigma_c\right), \qquad (3.3.15)$$

where $F_\varsigma\left(x\right)$ is the density function of the logistic distribution, i.e. $F_\varsigma\left(x\right)=\dfrac{1}{1+e^{-x}}$.

It is important to note that, in contrast to the case of the model without fixed costs, there is no explicit function by which the critical relative preference ς_c can be computed. In our case, this level has to be computed numerically. As I will prove in Appendix A3.8, there is a unique solution for ς_c. Note that ς_c depends not only on economic variables but also on parameters $\theta=\left\{d,a_2,\gamma,\beta\right\}$. This causes certain problems when estimating the parameters. I will address these problems and present solutions to them in the section entitled “Estimation routine”.

Since the density of X_2 can be computed based on the same first-order conditions – namely (3.2.11) and (3.2.12) – the same density function for any positive value of X_2 is yielded. The density of the driving demand of a single household can therefore be computed using (3.2.21) and (3.3.15):[191]

$$L\left(z=X_2\,|\,\theta,p_1,p_2,y,s\right)=f_{X_2\wedge(X_2>0)}\left(z\,|\,\theta,p_1,p_2,y-k_2,s\right)^{I(z>0)}\cdot P\left(X_2=0\,|\,\theta,p_1,p_2,y,k_2,s\right)^{I(z=0)}, \qquad (3.3.16)$$

where

$$f_{X_2\wedge(X_2>0)}\left(z\,|\,\theta,\ p_1,\ p_2,\ y-k_2,s\right)=\frac{1}{\beta}\cdot f_\varsigma\left(\frac{V_1-V_2}{\beta}\right)\cdot\left(\frac{1-d}{\dfrac{y-k_2-p_2z}{p_1}+a_1}\cdot\frac{p_2}{p_1}+\frac{1-d}{z+a_2}\right), \qquad (3.3.16a)[192]$$

$$V_1=\ln\left(d\right)-\ln\left(p_1\right)-\left(1-d\right)\cdot\ln\left(\frac{y-k_2-p_2z}{p_1}+a_1\right), \qquad (3.3.16b)$$

$$V_2=\ln\left(d\right)+\ln\left(p_2\right)+m-\left(1-d\right)\cdot\ln\left(z+a_2\right),\text{ with } m=\gamma\cdot s \qquad (3.3.16c)$$

---

[191] Note that (3.3.16a) corresponds to (3.2.16). The only difference is that available income is reduced from y to $y-k_2$. The same holds for (3.3.16b), which corresponds to (3.2.14a), (3.2.14b) and (3.2.15).

[192] Note that this function is identical to function (3.2.21), with the exception that y is replaced by $y-k_2$.

and $I(z>0)$ and $I(z=0)$ are indicator functions, being one when the argument is true and zero when it is false. The parameter-vector θ contains all parameters, $\theta=\{d,a_2,\gamma,\beta\}$. Probability $P(X_2=0\,|\,\theta,p_1,p_2,y,k_2,s)$ is defined in (3.3.15). Using probability (3.2.16), the probability of observing $(x_{2n},y_n,k_2,p_{2n},s_n)_{n=1,2,\ldots,N}$ can be computed by:

$$L_{MLE}\left((X_2=x_{2n})_{n=1,2,\ldots,N}\,|\,\theta,\,p_1,\,p_{2n=1,2,\ldots,N},y_{n=1,2,\ldots,N},k_2,s_{n=1,2,\ldots,N}\right)=$$

$$\prod_{n=1}^{N}P\left(x_{2n}=0\,|\,\theta,\,p_1,\,p_{2n},y_n,k_2,s_n\right)^{I(x_{2n}=0)}\cdot\prod_{i=n}^{N}f_{X_2\wedge(X_2>0)}\left(x_{2n}\,|\,\theta,\,p_1,\,p_{2n},y_n-k_2,s_n\right)^{I(x_{2n}>0)}. \quad (3.3.17)$$

Simulation

Again, as already mentioned in the section entitled “Simulation” in Subchapter 3.2, the principal aim of this study is to simulate the impact of a change in fuel price and household income on car ownership and travel behaviour. Again, the simulation is based on average changes in simulated probabilities of not owning a car and the expectation value of individual households' driving demand. The only differences to the model that does not include fixed costs are the functions that describe the probability of not owning a car and the expectation value of driving demand. These functions are now as follows:

$$P_{sim,n}\,|\,\tilde{\theta},p_1,\,p_{2n},y_n,s_n=P\left(x_{2n}=0\,|\,\tilde{\theta},p_1,\,p_{2n},k_2,y_n,s_n\right)=F_{\varsigma}\left(\varsigma_{cn}\right), \quad (3.3.18)$$

where $F_{\varsigma}(x)$ is the density function of the logistic distribution, $F_{\varsigma}(x)=\dfrac{1}{1+e^{-x}}$, and ς_{cn} is the value of the unobserved relative preference for car driving ς where household *n* would be indifferent to owning a car or not, and $\tilde{\theta}$ is a random vector with the distribution of the estimated parameter vector $\theta=\{d,a_2,\gamma,\beta\}$ with $m_n=\gamma\cdot s_n$. The distribution of $\tilde{\theta}$ is defined in the subsequent section “Estimation routine”.

The expectation value of a household’s driving distance is defined as:

$$E_{sim,n}\left(X_2\,|\,\tilde{\theta},p_1,\,p_{2n},y_n,k_2,s_n\right)=\int_{z=x_2(\varsigma_{cn})}^{z=y/p_2}z\cdot f_{X_2\wedge(X_2>0)}\left(z\,|\,\tilde{\theta},\,p_1,\,p_{2n},y_n-k_2,s_n\right)dz\,. \quad (3.3.19)$$

Since this integral is calculated numerically and the density function yields very small values for large values of *z*, this integral should also be transformed via integration by parts in the case of the model that includes fixed costs[193]:

---

[193] To understand why integration by parts is the better method of computing the integral than computing, see the footnote corresponding to (3.2.30).

$$E_{sim,n}\left(X_2 \mid \tilde{\theta}, p_1, p_{2n}, y_n, k_2, s_n\right) = \cdots$$

$$\cdots = \frac{y}{p_2} - x_2\left(\varsigma_2\right) \cdot P_{sim,n} \mid \tilde{\theta}, p_1,\ p_{2n}, y_n, k_2, s_n - \int_{z=x_2(\varsigma_{cn})}^{z=y/p_2} F_{X_2 \wedge (X_2>0)}\left(z \mid \tilde{\theta}, p_1, p_{2n}, y_n - k_2, s_n\right) dz, \quad (3.3.20)$$

where $F_{X_2 \wedge (X_2>0)}$ is the cumulated density of car driving, given that the household drives a positive amount. This function is defined as:

$$F_{X_2 \wedge (X_2>0)}\left(z \mid \tilde{\theta}, p_1, p_{2n}, y_n, s_n\right) = F_{\varsigma}\left(\frac{V_{1n} - V_{2n}}{\beta}\right), \quad (3.3.21)$$

where V_{1n} and V_{2n} are defined as

$$V_{1n} = \ln\left(\tilde{d}\right) - \ln\left(p_1\right) - \left(1 - \tilde{d}\right) \cdot \ln\left(\frac{y - k_2 - p_{2n} z}{p_1}\right), \quad (3.3.22a)$$

$$V_{2n} = \ln\left(\tilde{d}\right) - \ln\left(p_{2n}\right) - \tilde{m}_n - \left(1 - \tilde{d}\right) \cdot \ln\left(z + \tilde{a}_2\right), \text{ with } \tilde{m}_n = \tilde{\gamma} \cdot s_n . \quad (3.3.22b)$$

As in the case of the model without fixed costs, the population average marginal effects of changes in economic variables – here the case of the population average marginal effects of an increase in driving costs – are computed as follows:

$$\Delta P_{sim} \mid \tilde{\theta}, p_1,\ p_{2n}, \Delta p_{2n}, y_n, k_2, s_n = \frac{1}{N} \cdot \sum_{n=1}^{N} \Delta P_{sim,n} \mid \tilde{\theta}, p_1, p_{2n}, \Delta p_{2n}, y_n, k_2, s_n =$$
$$= \frac{1}{N} \cdot \sum_{n=1}^{N} P\left(x_{2n} = 0 \mid \tilde{\theta}, p_1,\ p_{2n} + \Delta p_{2n}, y_n, k_2, s_n\right) - \frac{1}{N} \cdot \sum_{n=1}^{N} P\left(x_{2n} = 0 \mid \tilde{\theta}, p_1,\ p_{2n}, y_n, k_2, s_n\right), \quad (3.3.23)$$

$$\Delta E_{sim}\left(X_2 \mid \tilde{\theta}, p_1, p_{2n}, \Delta p_{2n}, y_n, k_2, s_n\right) = \frac{1}{N} \cdot \sum_{n=1}^{N} \Delta E_{sim,n}\left(X_2 \mid \tilde{\theta}, p_1, p_{2n}, \Delta p_{2n}, y_n, k_2, s_n\right) =$$
$$= \frac{1}{N} \cdot \sum_{n=1}^{N} E_{sim,n}\left(X_2 \mid \tilde{\theta}, p_1, p_{2n} + \Delta p_{2n}, y_n, k_2, s_n\right) - \frac{1}{N} \cdot \sum_{n=1}^{N} E_{sim,n}\left(X_2 \mid \tilde{\theta}, p_1, p_{2n}, y_n, k_2, s_n\right). \quad (3.3.24)$$

Population average marginal effects (3.2.33) and (3.2.34) are conditional on the estimated parameter vector $\tilde{\theta}$, and are therefore random variables. The results will show that the variation of the simulated effects is rather small; see section entitled “Results”.

Estimation routine

As mentioned in the previous section, a problem arises when estimating parameters by MLE. The two diagrams below shed light on this problem. The diagrams show that parameters d and a_2 influence the minimal driving distance $x_2\left(\varsigma_c\right)$.

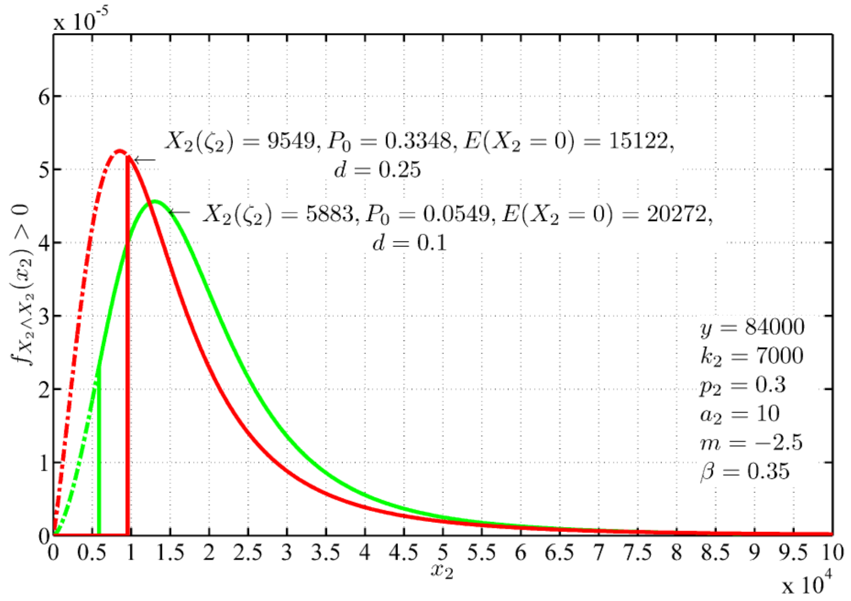


Figure 3.3.2: Density function of X_2 for different parameter values of *d*.

The diagram above shows that a decrease in parameter *d* decreases the minimum distance a household drives when deciding to own a car. This is due to the fact that *d* determines the decrease in marginal utility: the lower *d* is, the more rapidly the utility of car driving decreases. On the other hand, the first few kilometres yield a higher utility than in the case where *d* is high. For lower values of *d*, therefore, households attempt instead to spread consumption over both goods, even if their income is decreased by the fixed costs incurred when owning a car. In the extreme case, $d \to 0$, $x_2(\varsigma_c) \to 0$, since $\varsigma_c \to \varsigma_0$, as previously described in this section.

Another parameter that influences $x_2(\varsigma_c)$ is parameter a_2.

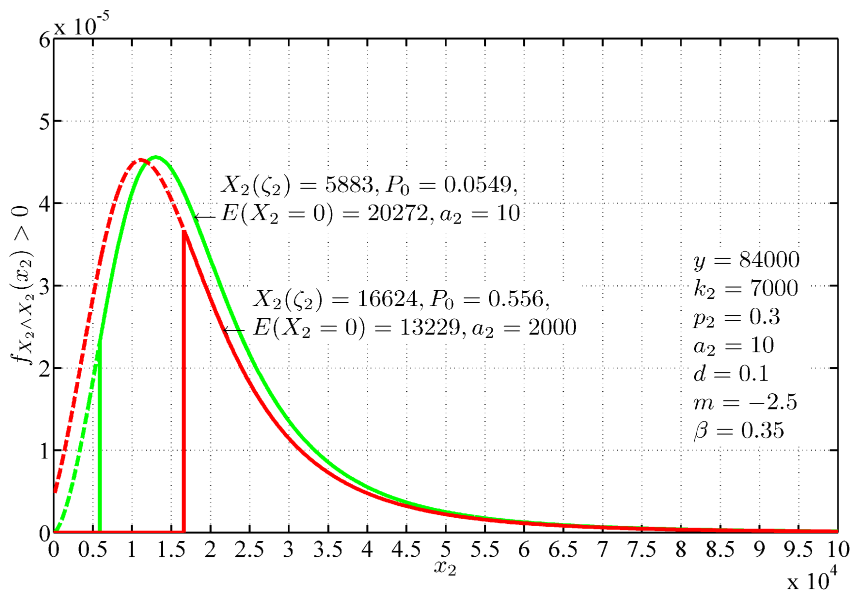


Figure 3.3.3: Density function of X_2 for different a_2.

The problem that can arise when estimating the parameters by MLE can be illustrated as follows: let us assume there is a household with an income of CHF 84,000 and an annual distance driven of 8,000 km. Let the parameters be the same as in the diagram above, and let a_2 be $a_2 = 10$. Let us now pretend that density function $f_{X_2 \wedge (X_2>0)}(\;)$, and therefore the ML function, increases when parameter a_2 increases. At some point $a_2 \approx 12$, the value of $f_{X_2 \wedge (X_2>0)}\left(8{,}000|\tilde{\theta}, p_1, p_{2n}, y_n, k_2, s_n\right)$ in this observation will be zero, since $x_2(\varsigma_c)$ is greater than 8,000 km, $x_2(\varsigma_c) | y, k_2, p_1, p_2, a_2, d > 8000$. This would mean that the MLE function is discontinuous at this point $a_2 \approx 12$ and that it yields zero for any value $a_2 > 12$, such that the optimization routine cannot find a maximum.

I circumvent this problem by the applying following estimation routine:

1. Choose values for d and a_2.
2. Compute $x_2(\varsigma_{cn}) | y_n, k_2, p_1, p_{2n}, a_2, d$ for each observation n.
3. Eliminate all observations where $0 < x_{2n} < x_2(\varsigma_{cn}) | y_n, k_2, p_1, p_{2n}, a_2, d$.
4. Estimate parameters γ and β by MLE conditional on d and a_2.
5. Compute a penalty function that depends a) positively on the proportion of eliminated datasets, b) positively on the relative error of the difference between the average simulated proportion of carless households, c) positively on the actual proportion of carless households and d) on the difference between the average simulated expectation value of driving demand and the actual average driving distance. Note that the actual proportion of carless households and the actual average driving distance refer to the measures based on the dataset after eliminating the observations according to Step 3.
6. Repeat Steps 1 - 5 for a number of different values for d and a_2 (grid search). Choose values d and a_2 so that the lowest value of the penalty function is yielded.

As the penalty function I chose

$$Q = \left(\frac{P_{sim} - P_{real}}{P_{real}}\right)^2 + c_1 \cdot \left(\frac{E_{sim}(X_2) - \text{mean}(x_{2\bullet})}{\text{mean}(x_{2\bullet})}\right)^2 + c_2 \cdot \left(\frac{\text{\# elim. observations}}{\text{size of initial datasets}}\right)^2, \qquad (3.3.25)$$

where P_{sim} is the average of the simulated probabilities, P_{real} is the actual proportion of carless households in dataset, $E_{sim}(X_2)$ is the average of the simulated expectation values of driving distance and $\text{mean}(x_{2\bullet})$ is the mean of the actual driving distance in the dataset. Expressions $\frac{P_{sim} - P_{real}}{P_{real}}$ and $\frac{E_{sim}(X_2) - \text{mean}(x_{2\bullet})}{\text{mean}(x_{2\bullet})}$ are the relative errors of the average of the simulated values, which could be called “replication errors”. Here “dataset” relates to the dataset after eliminating the dataset where

$0 < x_{2n} < x_2(\varsigma_{cn}) \mid y_n, k_2, p_1, p_{2n}, a_2, d$. Expression $\frac{\#\text{ elim. datasets}}{\text{size of initial datasets}}$ corresponds to the percentage of eliminated datasets with respect to the initial number of datasets. Parameters c_1 and c_2 are weighting parameters. I chose $c_1 = 1$, which means that both types of replication errors should be weighted about equally, and $c_2 = 0.5$. The latter choice yields a proportion of 8.8% of datasets that were eliminated; see the section on "Results". Despite the fact that some households may simply have stated a too low driving distance, a proportion of 8.8% of eliminated datasets seems to be quite high. The fact that a dataset is eliminated means that the corresponding households do not behave rationally according to that model. In this case, households drive less than the minimal driving distance that the model would predict. The reason for the high dropout rate could be that the model does not contain an option utility value of owning a car. An option utility value arises from the fact that owning a car simply gives the household an opportunity to drive anywhere at any time. The ignorance of this option value in the model could explain why there are so many households that own a car but drive shorter distances than would be rational according to the model. On the other hand, I am primarily interested in the changes in the total distance driven by households under different tax policies. Whether or not households that drive only very few kilometres are included in the dataset therefore does not influence the results much, since I do not expect these households to contribute significantly to the change in driving distance.[194] Note that since some datasets may be eliminated, depending on the choice of d and a_2, there will be discontinuities in the penalty function. For this reason, optimization routines for choosing optimal values d and a_2 that rely on smooth functions cannot be applied. I will therefore apply a grid search algorithm. Since it will transpire that the simulation results are not very sensitive to the choice of d and a_2, the grid does not need to contain very many points, once the grid range is located in the range where a maximum may be found. Despite the long computation time required to compute a single data point, results can therefore be computed in a reasonable time using the grid search algorithm. It is important to note that parameters m and β influence value ς_c, but not value $x_2(\varsigma_c)$. In other words, a change in parameters m and β does not change the number of datasets to be eliminated from the dataset.[195] This also enables parameters γ and β to be estimated by MLE.

[194] Since I exclude these datasets from the dataset when simulating changes, these households are treated as though they do not contribute anything to the change in driving distance. Since these households have a low mileage, I do not expect the simulated changes to be large in absolute terms. I therefore also do not expect a large error in the simulated average changes of driving demand caused by this exclusion.

[195] Recall that $m = \gamma \cdot s$.

Results

First, I shall present the three components that contribute to the penalty function as a function of the two parameters *d* and a_2, namely the share of datasets that have to be eliminated from the dataset, the relative deviation of the simulated proportion of carless households to the actual proportion and the relative deviation of the total simulated annual driving distance to the actual driving distance. Due to the high demand of computation time, I used a dataset that contains a random sample of 418 samples that amounts to 5% of the total dataset.

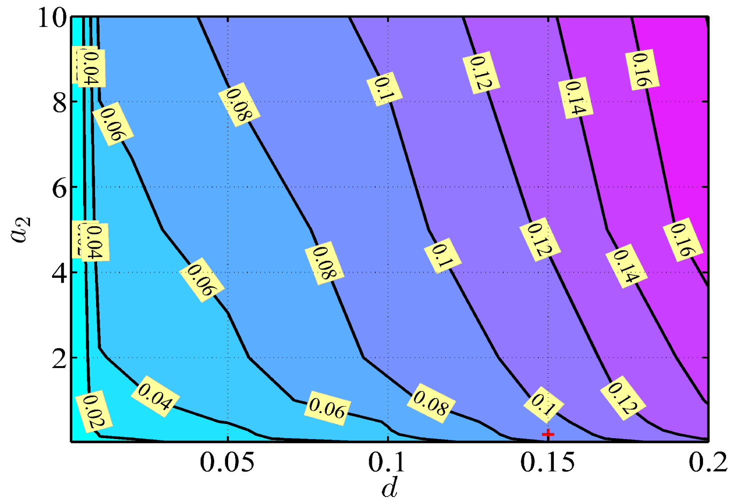


Figure 3.3.4: Proportion of datasets that to be eliminated from the dataset.[196]

This diagram shows that the greater parameter *d* and the greater parameter a_2, the more observations have to be eliminated due to $x_{2n} < x_2(\varsigma_c)|\cdots$, since value $x_2(\varsigma_c)$ increases in both *d* and a_2.[197] Since a proportion of eliminated datasets exceeding 10% is unrealistic, it can already be concluded from this diagram that the optimum solution has to be within the bottom left range of $d < 0.1$.

[196] This data is based on a random sample of the complete dataset. The sample size amounts to only 5% of the total sample for reasons of computation time.

[197] For a discussion on the impact of parameters a_2 and *d* on $x_2(\varsigma_c)$, see Mathematical Appendix MA2.

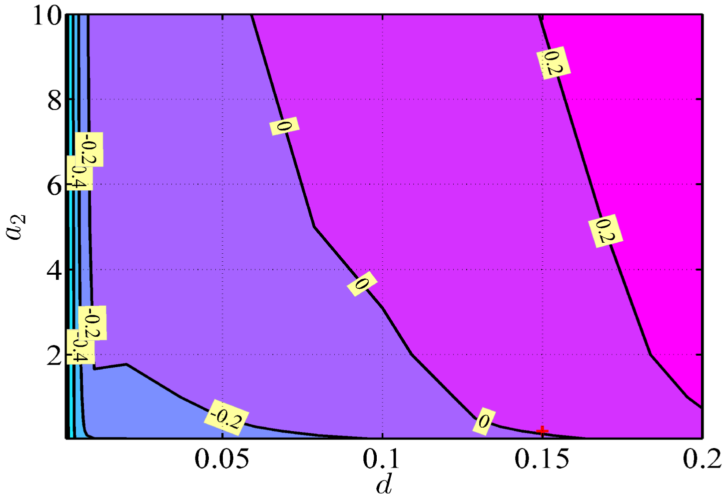


Figure 3.3.5: Relative deviation of the simulated proportion of carless households from the actual proportion.

This diagram shows that the smaller parameter *d*, the smaller also the difference between the actual proportion of carless households and the average forecast probability that households will not own a car. It is not possible to find an intuitive explanation for this observation, since both parameters *m* and β also influence the forecast probability that households will not own a car. The same holds for the following replication error of the dataset: the deviation of the average of the forecast expectation of driving distance from the average of observed driving distances in the dataset.

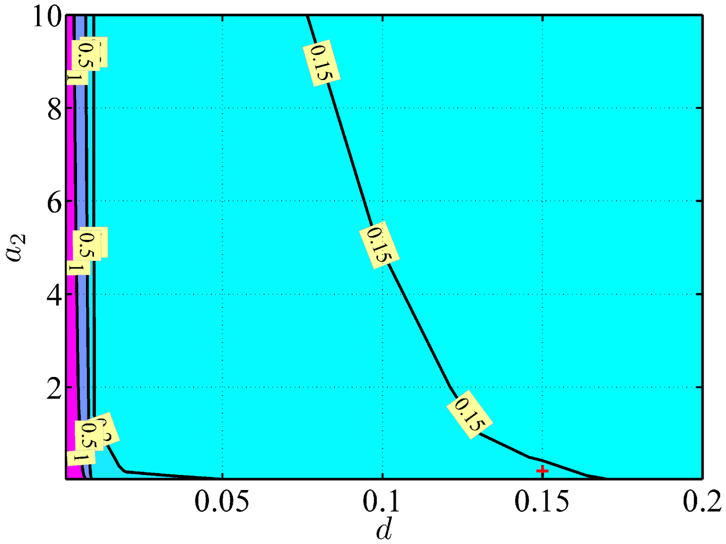


Figure 3.3.6: Relative deviation of the total simulated annual driving distance from the actual value.

This diagram shows that, irrespective of the choice of parameters d and a_2, the error of the forecast distance varies little. One assumption is that, again, as in the case of the model without fixed costs, the upper tail of the theoretical distribution is too heavy. I conclude from these three diagrams that the first two criteria, the proportion of eliminated datasets and the error of the forecast proportion of carless households, will determine the optimum parameter values d and a_2. This assumption is confirmed by the following diagram. The optimum result is located in the range where the proportion of eliminated datasets and the error of the forecast proportion of carless households varies considerably, whereas the error of forecast driving distance varies little and is quite high.

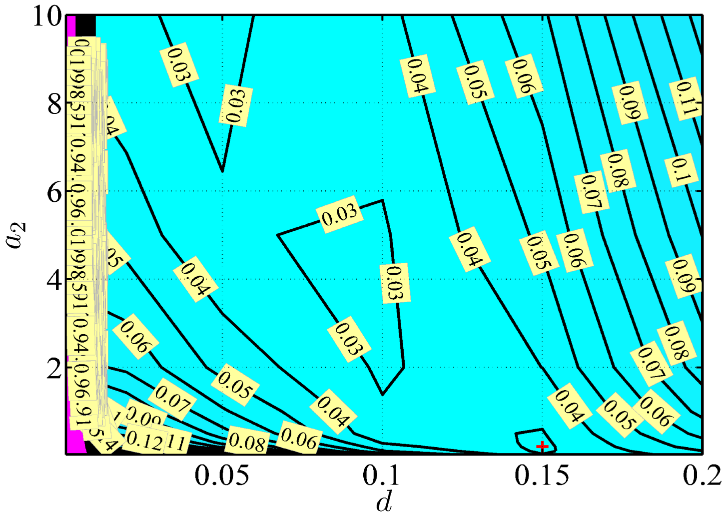


Figure 3.3.7: Value of the penalty function for different parameters d and a_2.

The diagram shows that penalty function Q is rather smooth, with the exception of very low values d and a_2. The minimum must be in the range around the line defined by points ($d = 0.15$, $a_2 = 0.02$) and ($d = 0.05$, $a_2 = 20$). Its exact location is at ($d = 0.15$, $a_2 = 0.2$), as indicated by the red cross "+".[198]

Note that the optimum solution corresponds to a drop-out rate of approximately 9%. This seems rather high, since it is implausible that this amount of people stated incorrect information or drove irrationally short distances each year. The reason for this rather high number of datasets could be that car ownership entails a fixed level of utility. This utility is provided by an option value of ownership. This value consists of the option of having a car available for a spontaneous trip or in the event of an emergency, e.g. when a family member has to be driven to hospital. I did not include this option value in the model for four reasons. First, the model would have become more complicated. Second, adding an additional parameter that captures this effect could have resulted in estimation problems, e.g. the existence of multiple local minima of the penalty function (3.3.25). Third, households that – according

[198] This outcome resulted from a grid defined by d = (0.02, 0.05, 0.1, 0.15, 0.2, 0.3) and a_2 = (0.3, 0.5, 1, 2, 5, 10, 20, 100).

to this model – drive irrationally short distances do not contribute a high proportion to the aggregate driving distance.[199] Fourth, I assume that these households do not change their driving demand differently in relative terms.[200] The effect of excluding the option value from the simulated effects on car ownership is unclear. The only aspect I found was that the choice of a combination of d and a_2 leading to fewer drop-outs reduces the simulated impact of p_2 on the probability that a households is carless, see Table A3.10.1. It is not possible, however, to determine whether this change in the simulated impact is due to the reduction of the number of eliminated datasets or due to parameters d and a_2, which have changed. However, the uncertainty surrounding the extent of this effect is not particularly relevant because I am more interested in the effect of policies on the change in aggregate driving distance.

I shall now present how optimal values d and a_2 are determined. Note that, conditional on these two values, parameters γ and β are estimated by MLE:

$$P_{MLE}\left(\left(X_2 = x_{2n}\right)_{n=1,2,..,N} \mid \theta_1, d, a_2, p_1, p_{2n=1,2,..,N}, y_{n=1,2,..,N}, k_2, s_{n=1,2,..,N}\right) = \tag{3.3.26}$$

$$= \prod^{N} P\left(x_{2n} = 0 \mid \theta_1, d, a_2, p_1, p_{2n}, y_n, k_2, s_n\right)^{I(x_{2n}=0)} \cdot \prod^{N} f_{X_2 \wedge (X_2>0)}\left(x_{2n} \mid \theta_1, d, a_2, p_1, p_{2n}, y_n - k_2, s_n\right)^{I(x_{2n}>0)},$$

with $\theta_1 = \{\gamma, \beta\}$.

---

[199] According to the model, the minimum driving distance that is rational amounts to 5,569 km for a household with an annual income of CHF 84,000. Assuming that households that drive less than this amount drive 4,000 km on average, this is much less than the average households that own a car drive, namely 17,127 km.

[200] The following example helps us understand arguments three and four. Assume that there are two individuals *A* and *B*. The aggregate relative change in aggregate driving demand is computed by:

$\frac{\Delta E(X_2)}{E(X_2)} = \frac{\Delta E(X_2)_A + \Delta E(X_2)_B}{E(X_2)_A + E(X_2)_B}$ (*). When assuming that $\Delta E(X_2)_B = k \cdot \left(\Delta E(X_2)_A + l\right)$ and $E(X_2)_B = k \cdot E(X_2)_A$, where k and l are constants, expression (*) can be written as $\frac{\Delta E(X_2)}{E(X_2)} = \frac{(1+k)\cdot \Delta E(X_2)_A + k \cdot l}{(1+k)\cdot E(X_2)_A}$ (*') and $\frac{\Delta E(X_2)_B}{E(X_2)_B}$ as $\frac{\Delta E(X_2)_B}{E(X_2)_B} = \frac{k \cdot \Delta E(X_2)_A + k \cdot l}{k \cdot E(X_2)_A} = \frac{\Delta E(X_2)_A + l}{E(X_2)_A}$ (**). The third argument refers to the case where constant k is very small. This means that both expectation value $E(X_2)_B$ and its change $\Delta E(X_2)_B = k \cdot \left(\Delta E(X_2)_A + l\right)$ are very small. Expression (*) is therefore approximately $\frac{\Delta E(X_2)}{E(X_2)} \approx \frac{\Delta E(X_2)_A}{E(X_2)_A}$ since $k\,l$ is very small compared to $\Delta E(X_2)_A$. This means that, in this case, neglecting the data of individual *B* implies almost no error. The fourth argument refers to the argument where l is very small. This implies $\frac{\Delta E(X_2)_B}{E(X_2)_B} \approx \frac{\Delta E(X_2)_A}{E(X_2)_A}$ and $\frac{\Delta E(X_2)}{E(X_2)} \approx \frac{(1+k)\cdot \Delta E(X_2)_A}{(1+k)\cdot E(X_2)_A} = \frac{\Delta E(X_2)_A}{E(X_2)_A}$. If at least one of arguments three and four holds, therefore, expression $\frac{\Delta E(X_2)_A}{E(X_2)_A}$ is a good approximation for $\frac{\Delta E(X_2)}{E(X_2)}$. Note that instead of individuals, $E(X_2)_A$ and $E(X_2)_B$ can be interpreted as expectation values of the sum of kilometres of a part of the population.

I have presented how parameters d, a_2, γ and β can be estimated.[201] Three aspects remain to be discussed: 1) the standard deviation of the parameters estimated and the simulated elasticities, 2) whether the choice of points of the grid and the width of the grid is not too large and 3) the error that may occur, since the result of the grid search algorithm used to determine optimum values d and a_2 is based on a sample containing only 20% of the observations of the original dataset.

I shall start by discussing the first aspect – the standard deviation of the parameters estimated and the standard deviation of the simulated elasticities. Since this estimation routine is not a standard routine like the MLE or the Method of Moments, there is also no direct numerical way to compute the variance-covariance matrix of the parameters estimated or the standard error of simulated elasticities, such as computing the Hessian matrix in the optimum of the penalty function Q. For this reason, I chose the bootstrapping method to compute these values. Since the computation time is very high using the complete dataset, I chose a dataset containing a random sample of the entire dataset, comprising 2% of the observations, namely 418. Then, 418 observations were always randomly drawn from this dataset. Each observation could be drawn more than once. This procedure was repeated 56 times based on a grid defined by $d = (0.002, 0.3, 0.002, 0.005, 0.01, 0.02, 0.05, 0.1, 0.15, 0.2, 0.3)$ and $a_2 = (0.02, 0.1, 0.2, 0.3, 0.5, 1, 2, 5, 10, 20, 100)$.[202] The following table shows for which values d and a_2 penalty function Q was minimal for each of these 56 dataset, given that the parameters for the penalty function were $c_1 = 1$ and $c_2 = 0.5$.

[201] Other estimation routines are, of course, also possible. For instance, a different penalty measure can be used, such as adding the final Maximum Likelihood value multiplied by a factor instead of expression (3.3.25). But not all of these potential estimation routines fit in a class of estimation procedures that have already been discussed in the literature. Second, the one I chose seems to be rather intuitive and to have sound properties, such as that the range of d and a_2 where penalty function (3.3.25) yields its minimal value is well known. Third, some observations may or may not be eliminated, depending on the choice of d and a_2. A measure based on a final Maximum Likelihood value would not be based on the same dataset; comparing such values is therefore problematic.

[202] Note that for the first 14 datasets, a grid defined by $d = (0.01, 0.02, 0.05, 0.1, 0.15, 0.2)$ and $a_2 = (0.02, 0.2, 0.5, 1, 2, 5, 10, 20, 100)$ was chosen. It transpired that some of the optimum solutions were close to the boundary of the grid, i.e. they were closed to $d = 0.01$. To ensure that all solutions were within the grid for the next draws and to test whether very small values could yield an optimal solution, as was the case in the model without fixed costs (see Subchapter 3.2, for iterations 15–20), the grid was expanded by adding lower values for d and a_2, namely $d = (0.002, 0.005, 0.01, 0.02, 0.05, 0.1, 0.15, 0.2, 0.3)$ and $a_2 = (0.02, 0.2, 0.5, 1, 2, 5, 10, 20, 100)$, and for iterations 21–56 $d = (0.001, 0.005, 0.01, 0.02, 0.05, 0.1, 0.15, 0.2, 0.3)$ and $a_2 = (0.02, 0.1, 0.2, 0.3, 0.5, 1, 2, 5, 10, 20, 100)$. The results showed that very small values of d were never optimal. In seven cases, however, a solution $a_2 = 0.02$ was optimal.

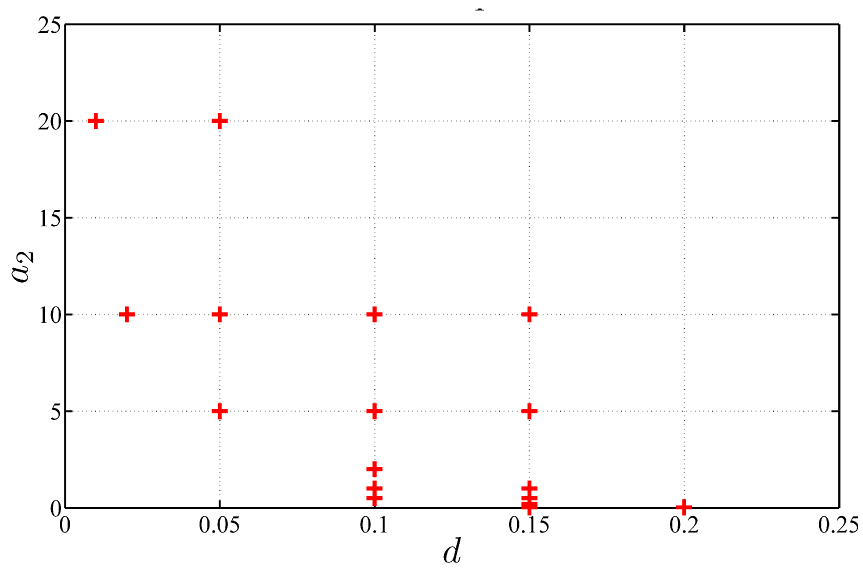


Figure 3.3.8: Optimal solutions for d and a_2 for different datasets generated by bootstrapping.

These results show that the optimum solutions vary quite considerably. Also, there are quite a number of solutions that yield a low value for parameter a_2. The standard deviation of the estimated parameters d and a_2 will therefore be rather large. Since these two parameters are not directly related to the effect of the explanatory variables, however, this fact is not very relevant. Further, it will be shown that the location of these optimal values d and a_2 does not have a significant effect on the simulated elasticities in which I am interested either, as shown in the following table and in Appendix A3.10. It is also shown in Appendix A3.10 that the effect of different choices of parameters c_1 and c_2 may change the optimum choice of values d and a_2. What is more, the choice of parameters c_1 and c_2 does not have a considerable effect on the value of the simulated elasticities.

The following table shows a summary of all estimated parameters and simulated values. Recall that these values were computed based on a dataset containing only 2% of the observations of the complete dataset.

	mean	median	min	max	stdev	$\frac{\text{stdev}}{\text{mean}}$	$\frac{\text{max}-\text{min}}{\text{mean}}$
d	0.089	0.100	0.010	0.200	0.057	0.639	2.141
a_2	8.329	10	0.020	20.000	7.077	0.850	2.399
γ_1	-2.818	-2.809	-3.019	-2.513	0.126	-0.045	-0.180
γ_2	0.364	0.363	0.209	0.538	0.085	0.234	0.904
β	0.408	0.405	0.327	0.480	0.043	0.106	0.376
Proportion of dropouts	0.105	0.108	0.060	0.153	0.024	0.230	0.889
Relative replication error of $P(X_2=0)$	-0.009	-0.007	-0.096	0.071	0.032	-3.389	-17.977
Relative replication error of $E(X_2)$	0.127	0.126	0.092	0.154	0.014	0.112	0.493
Q	0.023	0.023	0.016	0.032	0.004	0.161	0.720
$\varepsilon_{E(X_2),p_2}$	0.011	0.011	0.010	0.013	0.001	0.059	0.200
$\varepsilon_{P(X_2=0),p_2}$	-0.003	-0.003	-0.004	-0.002	0.000	-0.141	-0.614
$\varepsilon_{E(X_2),k_2}$	1.120	1.126	1.036	1.261	0.066	0.059	0.200
$\varepsilon_{P(X_2=0),k_2}$	-1.334	-1.327	-1.600	-1.188	0.093	-0.070	-0.309
$\varepsilon_{E(X_2),y}$	1.190	1.191	1.160	1.250	0.016	0.014	0.076
$\varepsilon_{P(X_2=0),y}$	-1.468	-1.468	-1.616	-1.323	0.073	-0.050	-0.200
$\frac{\Delta E(X_2)_{rural\to city}}{E(X_2)_{rural}}$ [%]	-31.07	-30.99	-44.83	-18.94	6.15	-0.20	-0.83
$\frac{\Delta P(X_2=0)_{rural\to city}}{P(X_2=0)_{rural}}$ [%]	73.86	69.63	37.94	150.62	25.08	0.34	1.53
$\frac{\Delta E(X_2)_{city\to rural}}{E(X_2)_{city}}$ [%]	46.25	45.27	23.52	80.60	12.99	0.28	1.23
$\frac{\Delta P(X_2=0)_{city\to rural}}{P(X_2=0)_{city}}$ [%]	-43.51	-42.19	-60.00	-27.67	8.20	-0.19	-0.74

Table 3.3.1: Estimated parameter values and simulated elasticities based on bootstrapped sample.[203]

This table shows that the values of interest – the simulated elasticities namely – vary little for the different datasets. With the exception of $\varepsilon_{P(X_2=0),p_2}$, the corresponding standard deviations are

[203] Note that these values are computed as defined by (3.3.23) and (3.3.24), where $\Delta p_{2n} = 0.01\cdot p_{2n}$ is chosen to compute $\varepsilon_{E(X_2),p_{2n}}$ and $\varepsilon_{P(X_2=0),p_{2n}}$. Values $\Delta y_n = 0.01\cdot y_n$ and $\Delta k_2 = 0.01\cdot k_2$ are used to compute the corresponding elasticities.

therefore rather small, particularly when comparing them with the absolute levels of the values of interest. Note that these values were computed based on a dataset containing only 2% of the observations of the complete dataset. For this reason, the standard deviations related to the complete dataset are smaller by a factor of $\sqrt{1/50}$:

	(mean)	(median)	stdev	$\frac{\text{stdev}}{\text{mean}}$
d	0.089	0.100	0.008	0.090
a_2	8.329	10.000	1.001	0.120
γ_1	-2.818	-2.809	0.018	-0.006
γ_2	0.364	0.363	0.012	0.033
β	0.408	0.405	0.006	0.015
Proportion of dropouts	0.105	0.108	0.003	0.033
Relative replication error of $P(X_2=0)$	-0.009	-0.007	0.004	-0.479
Relative replication error of $E(X_2)$	0.127	0.126	0.002	0.016
Q	0.023	0.023	0.001	0.023
$\varepsilon_{E(X_2),p_2}$	0.011	0.011	**0.0001**	**0.008**
$\varepsilon_{P(X_2=0),p_2}$	-0.003	-0.003	**0.0001**	**-0.020**
$\varepsilon_{E(X_2),k_2}$	1.120	1.126	**0.009**	**0.008**
$\varepsilon_{P(X_2=0),k_2}$	-1.334	-1.327	**0.013**	**-0.010**
$\varepsilon_{E(X_2),y}$	1.190	1.191	**0.002**	**0.002**
$\varepsilon_{P(X_2=0),y}$	-1.468	-1.468	**0.010**	**-0.007**
$\frac{\Delta E(X_2)_{rural\to city}}{E(X_2)_{rural}}$	-31.07%	-30.99%	**0.870**	**-0.028**
$\frac{\Delta P(X_2=0)_{rural\to city}}{P(X_2=0)_{rural}}$	73.86%	69.63%	**3.547**	**0.048**
$\frac{\Delta E(X_2)_{city\to rural}}{E(X_2)_{city}}$	46.25%	45.27%	**1.837**	**0.040**
$\frac{\Delta P(X_2=0)_{city\to rural}}{P(X_2=0)_{city}}$	-43.51%	-42.19%	**1.160**	**-0.027**

Table 3.3.2: Estimated parameter values and simulated elasticities corresponding to the complete dataset.

The results in this table show that all standard deviations of the simulated elasticities are rather small, namely less than 3.2% of their absolute level – with the exception of $\varepsilon_{P(X_2=0),p_2}$ – which is why the 95% confidence intervals are small, too. The actual model can therefore be considered to be rather accurate. Note that I place the mean and median values in parentheses, since they do not converge to the "true" value or to the value that is closer to the "true" value, namely the value we would arrive at when using the complete dataset listed in Table 3.3.3. Since the standard deviations of the simulated elasticities are very small, however, I assume that the values of the standard deviations "stdev" and the ratio stdev/mean would also be very small when using the complete dataset for computation purposes. I will therefore use these values for the standard deviations as the "best" estimate for the standard deviations in Table 3.3.3 below.

The second question I endeavour to answer is whether the grid width is too large. This question can be answered by an intuitive argument: let us consider the simulated elasticities of a set of grid points that yields less that 10% more value for penalty value Q than the grid point yielding the smallest value for Q. It transpires that all simulated elasticities corresponding to these grid points are very similar to that corresponding to the point yielding the smallest value for Q, see Figures A3.10.1 and A3.10.2. It can be concluded from this that there would only be a minor change in the results if the grid width was further reduced.

The third question, namely whether using a dataset containing only 20% of the total observations to determine optimum values d and a_2 could be answered using the same argument as in question two: although there may be a small deviation of d and a_2 from the "true" optimal values, this deviation would not affect the simulated elasticities much. For the same reason, I conclude that no major error is induced when using a dataset containing only 20% of the total observations.[204]

I shall now present the results based on an estimation using the complete dataset containing 19,038 observations. The results were computed conditional on values d = 0.15 and a_2 = 0.2. Recall that the values resulted from the grid search algorithm illustrated in Figure 3.3.8. The standard deviations of the point estimates are written in parentheses "()" below the values for the point estimates. Their values are based on the values from Table 3.3.1.

[204] The fact that the mean values of the elasticities computed using a dataset consisting of only 2% of all observations of the dataset are almost identical to the elasticities computed using the complete dataset.

$\varepsilon_{E(X_2),p_2} \mid (0.15,0.2) = -\underset{(0.0001)}{1.19}$	$\varepsilon_{P(X_2=0),p_2} \mid (0.15,0.2) = \underset{(0.0001)}{0.30}$
$\varepsilon_{E(X_2),p_2,tax\ neutral} \mid (0.15,0.2) = -\underset{(-)}{1.09}$	$\varepsilon_{P(X_2=0),p_2,tax\ neutral} \mid (0.15,0.2) = \underset{(-)}{0.21}$
$\varepsilon_{E(X_2),p_{fuel}} \mid (0.15,0.2) = -\underset{(0.00005)}{0.54}$	$\varepsilon_{P(X_2=0),p_{fuel}} \mid (0.15,0.2) = \underset{(0.00005)}{0.14}$
$\varepsilon_{E(X_2),p_{fuel},tax\ neutral} \mid (0.15,0.2) = -\underset{(-)}{0.10}$	$\varepsilon_{P(X_2=0),p_{fuel},tax\ neutral} \mid (0.15,0.2) = \underset{(-)}{0.10}$
$\varepsilon_{E(X_2),k_2} \mid (0.15,0.2) = -\underset{(0.009)}{0.18}$	$\varepsilon_{P(X_2=0),k_2} \mid (0.15,0.2) = \underset{(0.013)}{1.39}$
$\varepsilon_{E(X_2),k_2,tax\ neutral} \mid (0.15,0.2) = -\underset{(-)}{0.10}$	$\varepsilon_{P(X_2=0),k_2,tax\ neutral} \mid (0.15,0.2) = \underset{(-)}{1.29}$
$\varepsilon_{E(X_2),y} \mid (0.15,0.2) = \underset{(0.002)}{1.19}$	$\varepsilon_{P(X_2=0),y} \mid (0.15,0.2) = -\underset{(0.01)}{1.44}$

Table 3.3.3: Simulated elasticities on fixed parameters *d* and a_2[205]

The results show that driving costs p_2 have a high impact on driving demand, whereas their impact on car ownership, and consequently on the proportion of carless households, is rather small compared to a tax on car ownership, namely $\varepsilon_{P(X_2=0),p_2} = -0.30$ versus $\varepsilon_{P(X_2=0),k_2} = -1.39$. This is because for households with a low preference for car driving that drive only short distances – if they own a car at all – an increase in driving costs does not contribute much to the total costs of driving. These households do not therefore have a great economic incentive to sell the car and become carless. The difference of the effect of these taxes can be illustrated using the following example: let us consider a household with a rather low preference for car driving, which is sufficiently high, however, to induce it to own a car. Let us assume further that this household drives 6,000 km per year. Then, if a 1% tax on driving distance cost CHF 16 ≈ 6,000 · 0.27 · 0.01, a 1% tax on the fixed costs of the car would cost the household CHF 70 = 7,000 · 0.01, which is more than four times[206] the amount it had to pay for the extra taxes on fuel. For the households for which the decision whether or not to own a car is relevant

[205] The values in brackets "(...)" denote the standard deviations, see Table 3.3.1. The standard deviations were not computed for the values corresponding to government income-neutral tax policies. I assume that these values are almost identical to tax schemes that are not government income-neutral. Individual households' driving costs per kilometer were computed as 0.1601 + 0.077825 · (average fuel price during the period in which the household drove). The marginal effect of a tax on fuel was therefore computed using $p_{2n} = 0.1601 + 0.077825 \cdot p_{fuel,n} \cdot (1+0.01)$ to simulate driving distance and the probability of not owning a car. In order to save computation time, to compute the standard deviation of the simulated effects of a change in fuel price, I simply multiplied the standard deviations of the effects of a change in driving costs by 0.457. This factor results because the average fuel price in 2005 was about CHF 1.729, yielding fuel costs of 0.077825 l/km · CHF 1.729 /l = CHF 0.1346 /km. Thus, fuel costs as a percentage of the total driving costs are about CHF 0.1346 /km / (CHF 0.1346 /km + CHF 0.1601 /km) = 45.7%. A 1% increase in fuel price therefore increases the total costs of an additional kilometre by only 0.457. This implies that the fuel price elasticity is about 0.457 times the elasticity of the driving demand with respect to the driving costs, and so the standard deviations are also smaller by about that factor.

[206] Note that this is also about the ratio of $\varepsilon_{P(X_2=0),p_2}$ and $\varepsilon_{P(X_2=0),k_2}$.

due to their low preference for car driving, therefore, the effect of a tax on car ownership has a much higher effect due to higher costs. Note that for households with a high preference for driving, neither types of tax have an effect on car ownership, since these households will keep their car whatever. This argument leads to the explanation why the effect of a tax on driving costs is much more effective with respect to households' average driving distance than a tax on car ownership, namely $\varepsilon_{E(X_2),p_2} = -1.19$ versus $\varepsilon_{E(X_2),k_2} = -0.18$. This is because even households with a very high preference for car driving will reduce their driving distance significantly, whereas a tax on car ownership would have only a minor budget effect on such households. Since the behaviour of households with a high preference for driving is highly relevant to the impact of any policy on driving demand, the effect just described explains the main cause of the difference between the impact of the two policies. Note that both of these taxes yield around the same scale of revenue, namely an average of CHF 46.9 per capita for the 1% tax on driving and an average of CHF 54.7 per capita for the 1% tax on car ownership. Differences in terms of the effects of the two types of taxes can therefore not be explained by differences in tax revenue. This also implies that the effect of a driving distance tax on driving demand is much greater than the effect of a tax on car ownership per unit of tax revenue. Note that taxes on fuel have almost exactly the same effect per unit of tax revenue as taxes on driving distance, given that the fuel efficiency of cars remain unchanged.[207] For this standard car, approximately 45.7% of the driving costs are related to fuel costs. The fuel price elasticities could therefore be simply computed by multiplying the values of the elasticities of driving costs by 0.457.[208] The resulting fuel price elasticity is $\varepsilon_{E(X_2),p_{fuel}} = -0.54$. This value is smaller in magnitude than the value I computed for the model without fixed costs (-0.62). One explanation for this could be that, in this model here, the income effect of driving demand is no longer overestimated. For the same reason, the income elasticity is also smaller, namely $\varepsilon_{E(X_2),y} = 1.19$ versus 1.35. This value is also closer to those established by other studies; see Subchapter 1.3.

The term "tax neutral" stands for changes if price p_2 is increased solely by a tax proportional to p_{2n}, and the tax revenues are fully reimbursed to the household by an income subsidy proportional to their income. The results show that the magnitude of the elasticity of driving demand with respect to driving costs p_2 is reduced compared to the case where the tax revenue would not be reimbursed or the increase in marginal costs would simply be exogenous: the magnitude would drop from 1.19 to 1.09, a reduction of about 10%. This reduction results from the income effect. If income rises, driving demand

[207] Note that the ladder is assumed within this model framework.

[208] This is an approximation. A model that includes the adaptation of households towards more fuel efficient cars when fuel prices increase can be found in Appendix A3.14.

rises, too, and the probability of a household being carless decreases.[209] Due to this, the effect of a tax on driving distance is reduced. A similar effect occurs when examining the change in the probability that households do not hold a car. Due to the income effect, the increase in this proportion caused by the increase in driving costs is decreased, and the value of the elasticity for the tax income-neutral policy is lower. The impact of an increase in driving costs on the number of carless households would drop by almost one third, namely from 0.30 to 0.21. In relative terms, this reduction is higher than that of the elasticity of driving demand. This is because the income effect on car ownership is relatively higher than that of driving distance.

Similar results occur when we examine the tax income-neutral policy for a tax on car ownership. Since the income effect is equally strong on both the probability of not owning a car and on driving distance, the already weak effect of this tax on driving distance is almost halved, whereas the effect of this tax on the probability of being a carless household is reduced only by less than 10%. This is rather intuitive. As already mentioned above, households with a high preference for car driving that drive many kilometres contribute a large proportion of kilometres to the economy's total driving distance,[210] and therefore a tax on car ownership purely has an income effect for these households, since they are unwilling to sell their car whatever. Reimbursing the tax revenue therefore recompenses the income effect of these households to a high degree. The main effect of the tax income-neutral policy is that households with a low preference for driving will sell their car.

I have already argued in the previous paragraph that the reduction in the economy's driving demand is caused by either households selling their cars or by households that keep their cars but drive less because of higher costs. In the following, I wish to examine the extent to which these two forces drive the effect of the two different taxes. How these effects can be computed has already been described in Subchapter 3.2 in the section entitled "Results".[211]

---

[209] In fact, because of this effect, the tax revenue of the tax on driving distance will increase, meaning that the tax reimbursement also has to be increased. Again, the demand for driving will increase, and the probability of households being carless will decrease, etc. I therefore had to compute the equilibrium reimbursement quote δ_r for a given tax rate on driving distance t_{p_2}. I chose t_{p_2} to be $t_{p_2} = 0.01$. The reimbursement quote δ_r satisfies the following condition: $\delta_r \cdot \sum_{i=1}^{n} y_n = t_{p_2} \cdot \sum_{i=1}^{n} p_{2n} \cdot E_{sim,n}\left(X_2\right) | \tilde{\theta}, p_1, \left(1+t_{p_2}\right) p_{2n}, \delta_r \cdot y_n, k_2, s_n$, where the function $E_{sim,n}\left(X_2\right) | \tilde{\theta}, p_1, p_{2n}, y_n, k_2, s_n$ is as defined in (3.3.20).

[210] The descriptive analysis of the micro-census dataset from 2005 reveals that 50% of households account for about 85% of the economy's driving distance.

[211] Recall that these effects can be separated as follows: $\varepsilon_{E(X_2), p_{2n}, n} = \varepsilon_{E(X_2|X_2>0), p_2} - \dfrac{\varepsilon_{P(X_2=0), p_{2n}, n}}{P_n\left(X_2=0\right)^{-1}-1}$.

For the case of a tax on driving distance, the effect of households that keep their car but drive less amounts to

$$\varepsilon_{E(X_2|X_2>0),p_2} \;|\left(0.15,0.2\right)=-1.11\cdot \tag{3.3.27}$$

This means that the effect of people selling their car amounts to only 6.7%.[212] The main effect of a tax on driving demand is therefore that households that keep their car will drive less.

For the case of a tax on car ownership, the effect of households that keep their car but drive less amounts to

$$\varepsilon_{E(X_2|X_2>0),k_2} \;|\left(0.15,0.2\right)=-0.0972\,. \tag{3.3.28}$$

This means that the effect of people selling their car amounts only to 46%.[213] Half of the effect of a tax on car ownership therefore stems from households selling their car, and the other half from households that keep their car but drive less because of the budget effect of this tax. Note that this effect is quantified for the case in which the tax is not reimbursed. If it was reimbursed, however, the effect of households selling their car would be dominant.

A final interesting effect is how the driving demand changes if households move and change location. I distinguish between two types of locations: urban and rural areas. I consider all city areas, agglomerations and isolated cities to be urban areas.[214] In the following, the simulated effects when households move are as follows:

The first component refers to the effect of the households that keep their car but drive less and the second to the effect of households that sell their car.

[212] (1.19-1.11)/1.19 = 6.7%.

[213] (0.18-0.0972)/0.18 = 46%.

[214] This terminology is according to the definition of the Bundesamt für Statistik (2006a).

	Original urban area	Original rural area
Proportion of households	76.6%	23.4%
Simulated value $E(X_2)$ original area	15954	21569
Simulated value $E(X_2)$ after move	23395	14653
Change in absolute level $\Delta E(X_2)$	7441	-6916
Change in relative level $\frac{\Delta E(X_2)}{E(X_2)}$	$\frac{\Delta E(X_2)_{urban\to rural}}{E(X_2)_{urban}} = 46.6\%$	$\frac{\Delta E(X_2)_{rural\to urban}}{E(X_2)_{rural}} = -32.1\%$
Simulated value $P(X_2=0)$ original area	14.5%	23.8%
Simulated value $P(X_2=0)$ after move	25.3%	13.5%
Change in absolute level $\Delta P(X_2=0)$	10.8%	-10.3%
Change in relative level $\frac{\Delta P(X_2=0)}{P(X_2=0)}$	$\frac{\Delta P_{urban\to rural}}{P_{urban}} = -43.3\%$	$\frac{\Delta P_{rural\to urban}}{P_{rural}} = 74.3\%$

Table 3.3.4: Simulated changes when households change type of location.[215]

The effects of a change in household location is rather dramatic. Changing household location from an urban to a rural area, and vice versa, implies a change in driving distance of around 7,200 km and a change in the proportion of carless households of approximately 10.5% each.[216]

So far, I have presented results based on marginal changes of the explanatory variables. As a final result, I wish to present the simulated effects of non-marginal changes of the economic variables income y, the fixed costs of car ownership k_2, the marginal costs of driving p_2 and fuel prices p_{fuel}. It is important to compute the effect of significant changes in explanatory variables because, in the long run, economic variables may change quite dramatically. For instance, it would be interesting to know the effect of a high increase in fuel prices. The following diagram shows the relative changes in the

[215] Note that the term “original urban area” indicates that the corresponding households currently live in an urban area. The “proportion of households” of 76% is therefore the proportion of households that currently live in rural households. The “Change in absolute level” amounts to the average change in driving distance if households living in an urban area move to a rural area.

[216] The level of change is slightly higher when urban households move to rural areas than vice versa. An intuitive explanation for this is that the income of urban households is higher on average. Households with a higher income will increase their driving distance more strongly when moving from an urban to a rural area than households with a lower income, since their driving demand is less restricted by their budget.

proportion of carless households when economic variables y, k_2, p_2 and p_{fuel} change by a certain percentage.

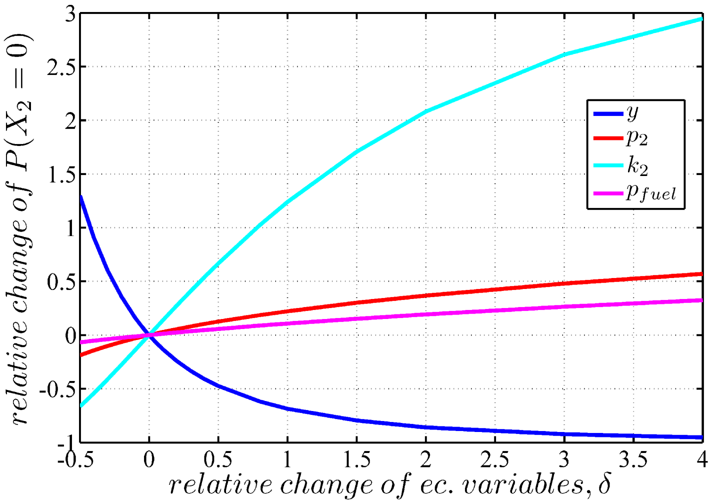


Figure 3.3.9: Relative changes in proportion of carless households when economic variables change.[217]

This diagram shows, for instance, that if the fixed costs of owning a car double, the proportion of carless households more than doubles. This would correspond to a rise from about 20% to more than 40%. Tripling the fixed cost would increase the proportion of carless households to about 56%. Doubling the marginal costs of driving would increase the proportion of carless households by around 30%, and even if the driving costs were four times as high as today, the proportion of carless households would only increase by 50%, which means that the level of carless households would increase only to a level of about 30%. Increasing the fuel price would have an even smaller effect: increasing fuel prices by the factor of four would yield a relative increase in the proportion of carless households of only about 13% to a level of about 23%. Interestingly, the effect of an increase in income is strongly decreasing. If income rises by 50%, the proportion of carless households almost halves. But even if income triples, still about 6% of households remain carless.

I also examined the effect of non-marginal changes on aggregate driving demand. The following diagram shows the relative changes in the aggregate driving demand when economic variables y, k_2, p_2 and p_{fuel} change by a certain percentage.

---

[217] Note that a relative change of $P(X_2 = 0)$, namely $\Delta P(X_2 = 0)/P(X_2 = 0)$, was observed on the y-axis, as well as a relative change δ of the economic variables with respect to the level of their value on the x-axis , e.g. $\Delta p_{2n} = \delta \cdot p_{2n}$. These values Δp_{2n} were plugged into Formula (3.3.23) to compute $\Delta P(X_2 = 0)$.

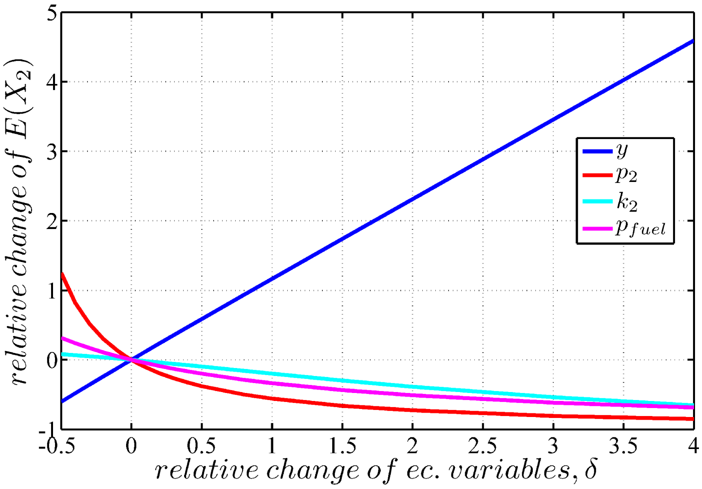


Figure 3.3.10: Relative changes in aggregate driving distance when economic variables change.[218]

Figures 3.3.9 and 3.3.10 show that the aggregate driving demand increases more or less proportionately to income. According to this model, there is therefore no satiation effect.[219] For increases in fixed costs by less than 100%, driving demand also decreases more or less linearly. This is the case because, in that range, the change in driving demand is mainly caused by the budget effect of households with a high preference for driving demand. If the fixed costs increase further, the effect of reducing driving demand by selling the car will become dominant. It seems that this effect is somehow less effective with respect to aggregate driving demand. The marginal change in driving demand with respect to fixed costs therefore decreases slightly when fixed costs increase. The change in driving demand with respect to driving costs p_2 also decreases. Whilst doubling driving costs results in about halving driving demand, an increase in driving costs by a factor of more than four is required to cut it to one quarter. Similarly, the change in driving demand with respect to the fuel price also decreases. Whilst doubling fuel prices results in a 30% decrease in driving demand, a tripling of fuel prices would be required to reduce driving demand by 50%. Note that the change with respect to fuel price decreases relatively less than the elasticity with respect to driving costs when the price increases by a certain factor, since the marginal driving costs contain a component that is independent of fuel price.[220]

---

[218] Note that a relative change of $E(X_2)$, namely $\Delta E(X_2)/E(X_2)$, was observed on the *y*-axis , as well as a relative change δ of the economic variables with respect to the level of their value, e.g. $\Delta p_{2n}=\delta \cdot p_{2n}$. These values Δp_{2n} were plugged into Formula (3.3.24) to compute $\Delta E(X_2)$.

[219] This is because this model does not capture the fact that at some point people no longer have the time to drive.

[220] Recall that the driving costs of an average car is given by p_2 = 0.1601 + 0.077825 p_{fuel}.

Comparison of results with the results of the model without fixed costs and other studies

Before comparing and discussing my results with those of the model without fixed costs, I shall first discuss the effect of a tax on car ownership, a tax on fuel and an increase in income on the share of households owning a car.

I shall start by discussing the effects of a tax on car ownership. The only study in which I could find a model where the effect of a tax on car ownership was examined was in Johansson and Shipper (1997). In their model, this tax was imposed by a tax on car purchase. Annualising one unit of this tax yields an increase in the fixed costs of car ownership of about 2% and will yield a decrease in car stock of 0.6%.[221] According to their model, a 1% increase in fixed costs would therefore reduce the vehicle stock by 0.3%. The number I obtain is

$$\varepsilon_{P(X_2>0),k_2} \,|(0.15,0.2) = \frac{P(X_2=0)}{1-P(X_2=0)} \cdot \varepsilon_{P(X_2=0),k_2} \,|(0.15,0.2) = -\frac{0.2162 \cdot 1.39}{1-0.2162} = -0.3835 \quad (3.3.29)^{222},$$

which is the relative change of the proportion of households with at least one car. The effect I obtain is larger. This could be the case because Swiss households have a higher propensity for substituting their car driving with public transport services than the average OECD countries on which the results of Johansson and Shipper are based. In contrast, the effects of income and fuel price on car ownership are much smaller than the results of Johansson and Shipper, namely

$$\varepsilon_{P(X_2=0),p_{fuel}} = \frac{P(X_2=0)}{1-P(X_2=0)} \cdot \varepsilon_{P(X_2=0),p_{fuel}} \,|(0.15,0.2) = -\frac{0.2162}{1-0.2162} \cdot 0.1235 = -0.0341 \qquad (3.3.30)$$

and

$$\varepsilon_{P(X_2>0),y} \,|(0.15,0.2) = \frac{P(X_2=0)}{1-P(X_2=0)} \cdot \varepsilon_{P(X_2=0),y} \,|(0.15,0.2) = \frac{0.2162}{1-0.2162} \cdot 1.437 = 0.3965 \qquad (3.3.31).$$

Johansson and Shipper (1997) find elasticities of the car stock of 1.26 with respect to income and of 0.12 with respect to fuel price. As already mentioned in the discussion of the results of the model without fixed costs, this difference can be explained by the fact that I do not explain the number of

---

[221] Johansson and Shipper (1997) simulate the effects of a tax on car purchase. Assuming that *the* average car costs about USD 25,000 and that the amortisation accounts for about half of the fixed costs, one unit of taxes of USD 1,000 constitutes an increase in fixed costs of about 1.1%. Note that this computation is based on Touring Club der Schweiz (2007). According to their cost structure, 31.2%/(31.2%+7.8%+13.2%+9.7%)=50.40% of the total fixed costs relate to amortisation. Assuming that the amortisation is proportional to the purchase price, increasing the purchase price by 1% increases the fixes costs by 0.504%. In other words, USD 1,000 correspond to an increase of about 1/25=4% in the purchase price, which implies an increase in the fixed costs of 2% = 4% · 0.504 .

[222] Note that $P(X_2=0)$ denotes the average simulated probability that a household is carless, which is $P(X_2=0)=0.2162$.

vehicles but only the proportion of households that own a car. If the average income increases, not only the share of households that own a car may increase but also the number of vehicles per households. The figure I obtain must therefore indicate the minimum of the effect. The fact that the latter effect is not captured by my model may explain the difference between the results. The same mechanism may be responsible for the differences in the effects of changes in fuel price on car ownership.

I now compare the results computed by the MDCEV model without fixed costs to the model that includes the fixed costs of car ownership. The following table lists all of the results computed by the two models.

Model that includes fixed costs	Model without fixed costs
$\varepsilon_{P(X_2=0),p_2} = \underset{(0.0001)}{0.30}$	$\varepsilon_{P(X_2=0),p_2} = \underset{(0.0088)}{1.63}$
$\varepsilon_{E(X_2),p_2} = -\underset{(0.0001)}{1.19}$	$\varepsilon_{E(X_2),p_2} = -\underset{(0.0043)}{1.36}$
$\varepsilon_{P(X_2=0),p_{fuel}} = \underset{(0.00005)}{0.14}$	$\varepsilon_{P(X_2=0),p_{fuel}} = \underset{(0.0040)}{0.68}$
$\varepsilon_{E(X_2),p_{fuel}} = -\underset{(0.00005)}{0.54}$	$\varepsilon_{E(X_2),p_{fuel}} = \underset{(0.0019)}{-0.56}$
$\varepsilon_{P(X_2=0),y} = -\underset{(0.01)}{1.44}$	$\varepsilon_{P(X_2=0),y} = \underset{(0.0087)}{1.62}$
$\varepsilon_{E(X_2),y} = \underset{(0.002)}{1.19}$	$\varepsilon_{E(X_2),y} = \underset{(0.0042)}{1.35}$
$\varepsilon_{P(X_2=0),k_2} = -\underset{(0.013)}{1.39}$	-
$\varepsilon_{E(X_2),k_2} = -\underset{(0.009)}{0.18}$	-
$\frac{\Delta P_{urban\to rural}}{P_{urban}} = \underset{(1.16/100)}{-43\%}$	$\frac{\Delta P_{urban\to rural}}{P_{urban}} = \underset{(1.00/100)}{-45\%}$
$\frac{\Delta E(X_2)_{urban\to rural}}{E(X_2)_{urban}} = \underset{(1.84/100)}{47\%}$	$\frac{\Delta E(X_2)_{urban\to rural}}{E(X_2)_{urban}} = \underset{(1.70/100)}{50\%}$
$\frac{\Delta P_{rural\to urban}}{P_{rural}} = \underset{(3.55/100)}{74\%}$	$\frac{\Delta P_{rural\to urban}}{P_{rural}} = \underset{(3.21/100)}{80\%}$
$\frac{\Delta E(X_2)_{rural\to urban}}{E(X_2)_{rural}} = \underset{(0.87/100)}{-32\%}$	$\frac{\Delta E(X_2)_{rural\to urban}}{E(X_2)_{rural}} = \underset{(0.76/100)}{-34\%}$

Table 3.3.5: Comparison of simulated elasticities of the two different models.

The most significant difference between these two results is that the effect of an increase in driving costs has about three times less an effect on the proportion of carless households than in the model without fixed costs. The same holds for the effect of an increase in fuel tax. The reason behind this becomes apparent when we look at the effect of an increase on the density function.

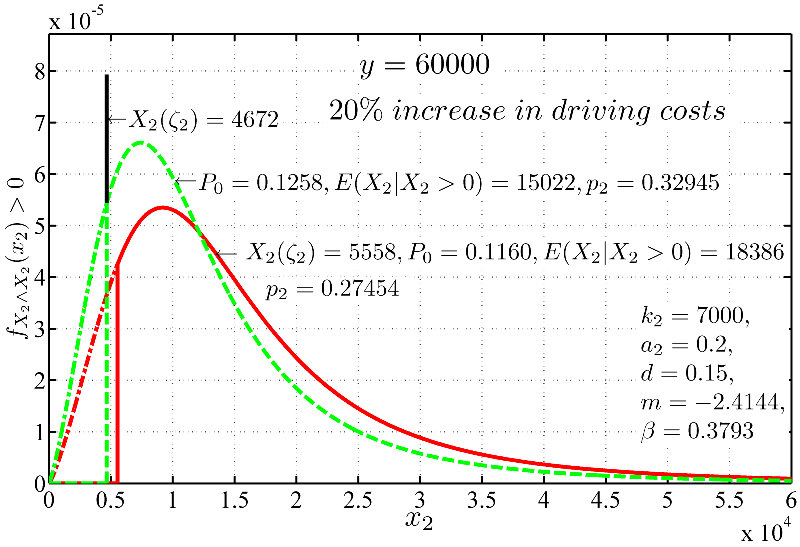


Figure 3.3.11: Illustration of the effect of an increase in driving costs.

The red solid curve shows the density function $f_{X_2 \wedge (X_2 > 0)}$ corresponding to the parameters estimated and an income level of CHF 60,000 for an urban household. Compared to the density function of the model without fixed costs (see Figure 3.2.10), the shape is very similar. Only the variance is slightly smaller, as we had assumed. The maximum is more to the right and, of course, the density function here approaches zero below a critical level $x_2(\varsigma_c)$. Interestingly, the critical level $x_2(\varsigma_c)$ decreases if fuel costs increase. But note that this critical level corresponds to a higher preference ς_c, namely $\varsigma_c = -0.7352$ versus $\varsigma_c = -0.7702$. The probability of not owning a car therefore also increases, namely from $P(X_2 = 0) = 0.1160$ to $P(X_2 = 0) = 0.1258$. Note that this decrease is relatively low compared to the decrease simulated by the model without fixed costs. There, an increase from $P(X_2 = 0) = 0.076$ to $P(X_2 = 0) = 0.116$ resulted. An intuitive explanation for this difference could be that, in this case, the critical travel distance decreases when fuel prices increase. The surface below the dashed and the dash-dotted curves that reflect $P(X_2 = 0)$ therefore do not change as much as if the limit remained unchanged as in the case of the model without fixed costs where this critical limit is always zero.

Another effect of this smaller change of the probability of not owning a car is twofold: first, the effect of selling the car has less impact on the change in aggregate kilometres. It contributes only about 6.7% to the total effect on driving demand. In the case of the model without fixed costs, however, this proportion constitutes about one third. Second, this could be the reason why the simulated effect of the fuel price on the aggregate driving demand is about 13% smaller than in the model without fixed costs.

The simulated level of the impact of an increase in income on the proportion of carless households is 11% lower; it is 13% lower on the aggregate driving distance than in the case of the model without fixed costs. Due to this, the results of the model including fixed costs of car ownership are closer to my results.[223]

Model quality

As in the case of the model without fixed costs, I wish to show that the model reflects the micro-census data of the Bundesamt für Statistik (2006a) quite well. It will transpire that this model here reflects the data better than the model without fixed costs, particularly in the case of high-income household segments.

However, I will first show that the aggregated data are also better fitted by this model here. Again, I computed the simulated data $P_{sim,avg}$ and $X_{2,sim,avg}$ as defined by (3.2.49) and (3.2.50), given estimates d, a_2, γ_1, γ_2 and β

Empirical values, original dataset	Empirical values	Simulated values
$\text{mean}\left(x_{2,n}=0\right)=18.90\%$	$\text{mean}\left(x_{2,n}=0\right)=20.72\%$	$P_{sim,avg}=21.62\%$
$\text{mean}\left(x_{2,n}\right)=13{,}890$	$\text{mean}\left(x_{2,n}\right)=14{,}868$	$X_{2,sim,avg}=17{,}281$

Table 3.3.6: Average simulated and average empirical driving distance and probability of not owing a car.

Note that the "Empirical values" refer to the dataset that from which observations with irrationally low values of driving distance had been removed. A total of 8.78% of the observations were eliminated, namely 1,832 observations from a total of 19,038. Since all of these observations included low driving distances, the average driving distance of the dataset used was higher. Also, since the eliminated observations contained data concerning households with a car, the proportion of carless households increased. The deviation of the simulated probability $P\left(X_2=0\right)$ from the actual proportion of carless households accounts for only 0.9%, a relative error of less than 5%. The deviation of the simulated average expectation value of car driving $E\left(X_2\right)$ from the actual average driving distance given by the data accounts to 2,413 km, a relative error of about 16%. Compared to the relative deviations corresponding to the model without fixed costs that account for 13% and 22%, this is a substantial improvement.

[223] For an overview of the results of international studies, see Subchapter 3.2, section "Results of other studies".

I also computed and illustrated the density functions for the different segments of households for this model, and compared them to the empirical distribution of the household segment.

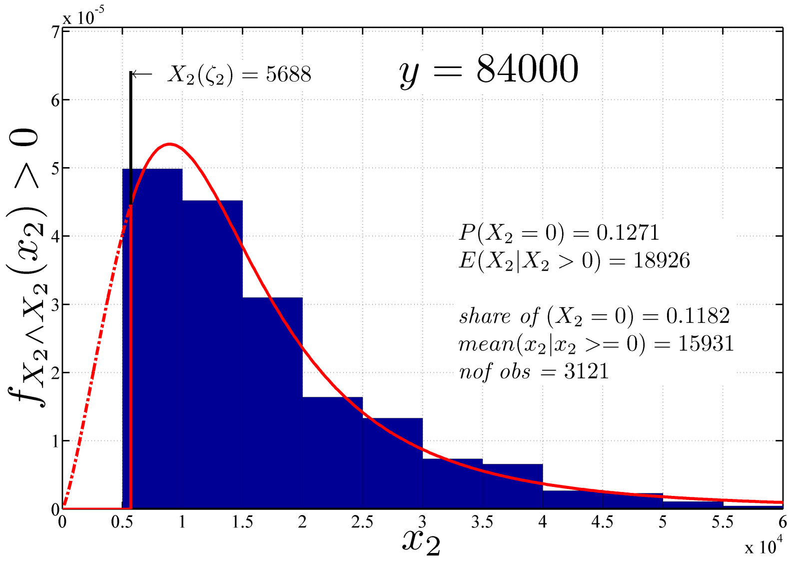


Figure 3.3.12: Histogram of households living in a urban area with an income of CHF 84,000.

This diagram shows that the density function $f_{X_2 \wedge (X_2 > 0)}$ determined by the model adapts the actual distribution of driving demand very well. The difference between the expectation value and the actual average of the observed driving distances occurs mainly because the density function has too much mass at the upper tail. On the other hand, probability $P(X_2 = 0)$ is almost exactly equal to the proportion of carless households in that segment.

One reason for the problem with the heavy tail of the density functions could be that the ML estimation routine – as a graphical interpretation – rather tries to adapt to the shape of the histograms than to the expectation values of individual household segments. In addition, the fact that the difference between the average expectation value and the mean of the observed data goes into the penalty function does not solve that problem, since the MLE routine is always the last step of the estimation routine. One cause of this problem could also be that the error term is standard Gumbel distributed, since the Gumbel distribution has a slightly heavier upper tail than a standard normal distribution. Note, however, that this assumption on the error term allows for a number of simplifications when deriving the ML function that can be computed very quickly. Another reason could be that the utility function does not account for the fact that households also have a time restriction, such that there is a satiation point above which the marginal utility of car driving would even be negative.

Comparison of the results of 2000 and 2005 datasets

I also estimated the model for the same kind of micro-census data collected in 2000, Bundesamt für Statistik (2006a). The resulting elasticities were almost identical for both datasets; see Appendix A3.11, Table A3.11.3. I also used the model to forecast the average driving demand and the proportion of carless households using the model estimated based on the dataset of 2000. To this end, I simulated the values of interest once using the 2000 data and then using the 2005 data. In both cases, I used parameter values corresponding to the 2000 dataset. This "forecast" yielded a change in the average driving distance by about 9.8%, whilst the actual increase was 5.8%; see Appendix A3.11, Table A3.11.3. I was able to show that this forecasting error is due to a reduction in household preferences – captured by the parameters of the model – for car driving. Taking this change of preferences into account, the forecast of the model is rather accurate. I also showed that such a reduction in preferences for car driving is rather plausible for mainly two reasons: first, in 2000-2005, the number of hours of congestion on Swiss motorways more than doubled, such that about 9.4% of all trips in 2005 were affected by traffic congestion. Second, in the same period the frequency of train services and the train speed increased significantly, particularly in regions where traffic jams prevail. For more details, see Appendix A3.11.

This finding is rather interesting and may also explain why the income and price elasticities in my model are almost twice those established by Baranzini et al. (2009): if fuel prices fall there will be more traffic and therefore road capacity will become short. This reduces the utility of car driving and, therefore, the rise in travelling demand is not as large as if there was no capacity restriction. The fuel price elasticity found by a model that ignores the effect of the road capacity on driving demand may therefore too low values of fuel price elasticities. The same mechanism holds for the income elasticity.

Since my model separates the effect of changes in economic variables and household preferences for car driving which are influenced by the offer of public transport and capacity limits of roads, the effect of policies can be forecast more precisely. For instance, it would be possible to determine the level of fuel price such that the level of traffic would remain constant or it would be possible to forecast the effect on traffic demand if road capacity were improved, such that the number of congestion hours would be halved.

A model based on a modified function for computing the expectation of driving demand

In the two models that include and exclude the fixed cost of car ownership, the distribution $f_{X_2 \wedge (X_2>0)}$ has too heavy tails. The forecast expectational values based on these density functions are therefore too high with high incomes, see Figures A3.11.4-19 and A3.11.21. This implies that the simulated income elasticity of driving demand is presumably too high, and there are suggestions that the same could be true for the fuel price elasticity of driving demand. A simple solution to avoid the problems arising from this heavy tail is to limit the upper boundary of the integral to 60,000 kilometres used when computing the expectational value $E(X_2)$, see (3.3.19). Using the same estimation routine as described above (3.3.25),[224] the following elasticities result:

Dataset	mz05	mz05
Limit of integrating, $E(X_2)$	∞	60,000km
$\varepsilon_{E(X_2),p_2}$	-1.19 (0.0001)	**-0.68**
$\varepsilon_{E(X_2),p_{fuel}}$	-0.54 (0.00005)	**0.28**
$\varepsilon_{E(X_2),y}$	1.19 (0.002)	**0.77**
$\varepsilon_{E(X_2),k_2}$	-0.18 (0.009)	**-0.17**
$\varepsilon_{P(X_2=0),p_2}$	0.30 (0.0001)	**0.26**
$\varepsilon_{P(X_2=0),p_{fuel}}$	0.14 (0.00005)	**0.11**
$\varepsilon_{P(X_2=0),y}$	-1.44 (0.01)	**-1.41**
$\varepsilon_{P(X_2=0),k_2}$	-1.39 (0.013)	**-1.31**
$\Delta E(X_2)_{rural\to city}/E(X_2)_{rural}$	-32% (0.87)	**-23%**
$\Delta P(X_2=0)_{rural\to city}/P(X_2=0)_{rural}$	74% (3.55)	**70%**
$\Delta E(X_2)_{city\to rural}/E(X_2)_{city}$	47% (1.84)	**27%**
$\Delta P(X_2=0)_{city\to rural}/P(X_2=0)_{city}$	-43% (1.16)	**-42%**

Table 3.3.7: Elasticities based on a model using a modified function for computing the expectation of driving demand.[225]

The resulting fuel price elasticity for driving demand is now smaller by a factor of 0.52 than in the case with the unmodified density function for computing the expectation value. The income elasticity of driving demand is only smaller by a factor of 0.65. Both these elasticities are now closer to the value reported by Baranzini et al. (2009). Since also, as illustrated in Figure A3.13.4, there is no strong

[224] Note that in this case the expectation value in step 5 is also based on this modified computation of the expectation value.

[225] Both these results are based on the complete dataset micro-census 2005, Bundesamt für Statistik (2006a).

trend in the deviation of the simulated expatiation value of driving demand from the empirical value when income increases, it can be followed that the income elasticity of driving demand is close to its true value.[226] This implies that the elasticities derived from this model with a modified function for computing the expectation of driving demand are closer to reality than those derived from the model with the unmodified demand function. The relative changes in driving demand when households move from rural to urban areas, and vice versa, are also smaller, namely by factors of 0.72 and 0.57. In this case, the values of the unmodified model are more realistic, since the relative changes computed by the empirical values, which denote a lower bound in magnitude,[227] are -0.29 and 0.40.[228] The elasticities with respect to the probability households being carless do not change much when using the modified function for computing the expectation of driving demand. All values decrease by only about 10%. A more extensive discussion on the differences between results due to this modification function for computing the expectation value of driving demand can be found in Appendix A3.13.

Comparison of results of the MDCEV and the Tobit model

So far, I have presented the results based on the MDCEV model in this chapter. An interesting question is whether the elasticities approximate those using the Tobit and the Probit model. The following table illustrates the results based on different models and datasets.

---

[226] Figure A3.13.4 shows that the difference between the simulated expectation value of driving demand and the empirical value slightly increases with income for urban households. For rural households, however, it strongly decreases. Since the proportion of rural households in the dataset is only about 23% (see Table A3.13.2) of the observations, there would be no trend in this difference with income at the level of the total population.

[227] The reason why they denote a lower bound is that the average income of households in urban areas is lower than that in rural areas.

[228] These two values are computed using the values in Table A3.13.2.

Type of model	MDCEV	MDCEV	MDCEV 60,000km	Tobit	MDCEV 60,000km	Tobit	Probit
Dataset	mz05	mz05	mz05	mz05	Erath	Erath	Erath
Fixed costs	no	yes	yes	--	yes	--	--
$\varepsilon_{E(X_2),p_2}$	-1.36	-1.19	-0.68	--	-0.69	-0.379 (0.02036)	--
$\varepsilon_{E(X_2),p_{fuel}}$	-0.564	-0.492	-0.28	--	-0.252 {-0.268}	-0.171 (0.01243)	--
$\varepsilon_{E(X_2),y}$	1.349	1.189	0.77	0.616 (0.0106)	0.822 {0.829}	0.786 (0.05809)	--
$\varepsilon_{E(X_2),k_2}$	--	-0.180	-0.17	--	-0.15	--	--
$\varepsilon_{P(X_2=0),p_2}$	1.63	0.297	0.26	--	0.21	0.472 (0.03908)	--
$\varepsilon_{P(X_2=0),p_{fuel}}$	0.682	0.124	0.11	--	0.0748 {0.093}	0.290 (0.0239)	0.2374 (0.072)
$\varepsilon_{P(X_2=0),y}$	-1.618	-1.437	-1.41	-0.850 (0.019)	-1.3	-0.950 (0.1109)	-1.248 (0.035)
$\varepsilon_{P(X_2=0),k_2}$	--	1.390	1.31	--	1.09	--	--
$\varepsilon_{x_{fuel},p_{fuel}}$	--	--	--	--	-0.31	--	--
$\varepsilon_{x_{fuel},y}$	--	--	--	--	0.978	--	--
Source:	Table A3.11.4	Table A3.11.4	Table A3.15.1	Table A3.15.1	Appendix A3.14	Appendix A3.14	Appendix A3.14

Table 3.3.8: Comparison of elasticities resulting from different models and different data.

The differences between the results generated by different models when the “Erath” datasets by Axhausen and Erath (2010) were used are of particular interest. This is because fuel prices in this stated preference dataset vary, meaning that the elasticities with respect to fuel price can also be computed. In the following I discuss the differences between the results yielded by the MDCEV and the Tobit model. The results based on the “Erath” dataset show that the fuel price elasticity of driving demand is about 32% lower in magnitude when the Tobit model is used. Both values are close to the value (-0.202) reported by Baranzin et al. (2009); the value of the MDCEV model is closer to the average found in international studies (-0.29), see Table 1.3.1. In contrast, the fuel price elasticities with respect to the probability of being carless is 3.9 time higher than when the Tobit model is used. Even though these elasticities cannot directly be compared to the values of the fuel price elasticity of the car stock found by international studies, the value found by the Tobit model is more realistic.[229] The fact that the value found by the Probit model is in a similar range to that established by the Tobit model also supports this view.

[229] See discussion below in the section on “Comparison with results of other studies” in Subchapter 3.2.

When considering the elasticities with respect to income, the MDCEV models yield higher values. In the case of the income elasticity with respect to driving distance, the values of the MDCEV model are closer to those reported by Baranzini et al. (2009) for the elasticity with respect to fuel demand, and can be therefore considered to be more realistic. In the case of the income elasticity with respect to the probability of a household being carless, the MDCEV model results are higher by a factor of 1.86 (mz05) and 1.37 (Erath). Again, even though these elasticities cannot be compared directly to the values of the income elasticity of the car stock found by international studies, the value reported by the MDCEV model is more realistic.[230]

Since the Erath dataset also contained information on the fuel efficiency of the car households would choose for different levels of fuel price, it was possible to compute the following three elasticities. First, the elasticities of fuel efficiency with respect to fuel price. Its value (-0.31) is smaller in magnitude than the average value found in international studies (-0.46), see Table 1.3.1. Second, the fuel price elasticity of driving demand including the effect of fuel prices on cars' fuel efficiency. This value becomes smaller (value in parentheses "{..}") in magnitude compared to the case where the effect of fuel prices on the cars' fuel efficiency is ignored (value in parentheses "{..}"). This difference is attributed to the so-called "rebound effect"[231]. Third, it is possible to compute the elasticities of fuel demand. The fuel price elasticity of fuel demand $\varepsilon_{x_{fuel}, p_{fuel}}$ is greater in magnitude than the fuel price elasticity of driving demand, due to the shift to more fuel-efficient cars when fuel prices increase. The "rebound effect" may not outperform the direct effect induced by the increase in the cars' fuel efficiency. The data also showed that the effect of an increase in income on the cars' fuel economy is negative. This is due to the fact that households with higher incomes tend to buy larger cars that are consequently less fuel-efficient. Including this effect leads to a very small decrease in the income elasticity of driving demand, namely 0.822 versus 0.829. The value of the income elasticities of fuel efficiency $\varepsilon_{x_{fuel}, y}$ is much higher than those found in international studies. One reason could be that Swiss households have a high preference for powerful cars and, therefore, the more they can afford them, the more likely they will buy such cars. A more extensive discussion on these differences between the results of these models and information on how these elasticities can be found in Appendices A3.14 and A3.15.

---

[230] The elasticity of the probability of owning a car would be $\varepsilon_{P(X_2=0), y} \approx -(-1.3) \cdot 0.25 = 0.32$. This value is much lower than the average value of the income elasticity of the car stock found by international studies (0.73), see Table 1.3.1.

[231] If fuel prices increase, households will buy cars that are more fuel-efficient. Therefore, the marginal costs – ceteris paribus – fall, compensating for part of the reduction in driving distance. This effect on driving distance induced by the use of more fuel-efficient cars is called the "rebound effect".

Final comments on the model and outlook for improvements

The results computed using the micro-census dataset showed that my model adapts very well to the data. A number of problems arose because the density functions corresponding to individual households had a heavy upper tail. This caused the problem that the expectation values corresponding to this density functions were higher than the average of the observed data. There was also doubt as to whether the forecast income and fuel price elasticities were rather too high due to this heavy tail problem. On the other hand, the elasticities I computed were in a range determined by international studies, but larger than those reported in the very recent studies of Axhausen and Erath (2010) and Baranzini et al. 2009 based on Swiss data. In these two studies, the elasticities with respect to fuel price are lower than those given in international studies. However, I have some doubt as to why the fuel price elasticity of Switzerland should be lower than in other countries, since due to the high standard public transportation system I assume that individuals may substitute private transportation for public transport more easily than individuals in other countries.

Another aspect is that a value for the fuel price elasticity could be processed despite the fact that the fuel price varied only very little. But, of course, if fuel prices variated more, the results would be more trustworthy. On the other hand, the elasticity of driving demand with respect to fuel prices was much smaller – namely 2.4 times smaller – than the elasticity with respect to income, which is in accordance with the results reported in international studies.

Another positive aspect of this model is that it is possible to simulate the effect on a tax on car ownership. I was previously unable to find any models that facilitated this. Further, it is possible to separate the effect of selling the car and driving less, given that the car is kept on the aggregated demand of different policies. It was also possible to compare the differences of the effects of two tax policies, namely a tax on fuel and a tax on car ownership, on car ownership and driving demand. Further, it was possible to examine the effect of the tax revenue being reimbursed to households.

So far, I have not applied the model to data of different countries or cross-sectional data from surveys prior to 2005. It would be interesting, for instance, to compare the differences of the relative preferences for car driving. If these relative preferences change over time, this could be a sign that the infrastructure of the public transportation sector has changed. The same cause could explain differences between countries.

So far, the model was restricted to the case where there is only one type of car available. An interesting extension would be to have the option of more than one car type. Each car type could be chosen or neglected, and ownership of each type would be connected with fixed costs. Preliminary research revealed that, for such a model, the density functions of driving demand, given a combination of ownership of car types, is still an explicit function of parameters and driving distances, but the probability that a household chooses such a combination can no longer be computed by computing at the root of a non-linear function. It would be necessary to compute these probabilities by simulation. For large datasets, computation time could therefore be prohibitively long. For small datasets up to 200-500 observations, however, it should be possible to compute results. Such a model would then also be interesting for other applications, such as the choice of a set of price plans of mobile phones that differ in fixed monthly rates and rates per minute of calling.

4. The willingness to pay for fuel efficiency

4.1 Introduction

In this chapter I examine how much households are willing to pay for more fuel-efficient cars. Firstly, knowledge of this willingness to pay is very important because switching to cars with lower fuel consumption will reduce carbon dioxide emissions per kilometre, as well as increase driving demand due to the lower marginal costs of driving. Secondly, it is important in order to estimate the demand for fuel-efficient cars, which is interesting to both policy makers and car manufacturers. Policy makers may be interested in knowing about the effectiveness of rebate systems[232] for fuel-efficient cars. Car manufacturers may wish to know whether consumers are willing to pay more for an improvement in fuel efficiency than it costs them to implement the technical solution required to realize it.

The dataset I will use is based on a stated preference survey.[233] Since each household had to answer a number of choice sets of three cars, the dataset has a panel structure.[234] Using a Multinomial Logit model (MNL), I will compute how car choice is affected by car and household attributes. In particular, I am interested in the impact of the cars' fuel economy and price on household choices. Using these results, I will be able to compute the willingness of different household segments to pay for fuel efficiency. Further, I intend to test whether households that drive long distances are willing to pay more for fuel efficiency because their fuel bills would be reduced more significantly than those of households that drive less. Since each household had to state several choices, I will use a MNL model that takes the panel structure of this data into account. There are basically two reasons why I evaluate this dataset again instead of simply commenting on the results of Wüstenhagen and Sammer (2007). First, Wüstenhagen and Sammer (2007) ignored the panel structure for their study. Second, I wish to examine whether or not capturing information on the energy label influences the model's results. In this context, the source of error is the poor design of the questionnaire in the sense that the information on the energy label presented to respondents did not correspond to the actual fuel consumption of the car. In reality, for a given car size the energy label (A, B, .. , G) relates directly to the fuel

[232] There are different types of rebate system. One type is to reduce the car ownership taxes of certain car types for a certain period after people have registered their new car. Another type of rebate is the payment of a bonus to an individual who buys a car of a certain type. Often these systems are coupled with additional taxes on certain car types, e.g. that have a low fuel efficiency – so-called gas guzzler taxes.

[233] This data was collected for a study that aimed to research consumer behaviour with respect to cars' fuel efficiency and fuel type (diesel versus petrol), see Wüstenhagen and Sammer (2007). I would like to thank Professor R. Wüstenhagen for providing this data for my research.

[234] A detailed description of the dataset can be found in Subchapter 1.4 and Appendix A1.1.

consumption of the car. Unfortunately, in this dataset the information on the energy label was simply randomly attributed to the cars. The information on the energy label therefore does not correspond to the information on the car's fuel consumption. For this reason, it is unclear whether respondents who care about a car's fuel efficiency took the information on actual fuel consumption or the information on the energy label into consideration. The parameters are therefore different to the case where households had been faced with realistic information, namely when the level of the energy label corresponds to the car's fuel consumption. Thus, in the following I do not only compute the results, but also endeavour to ascertain the extent to which this error could have biased them.

4.2 The model

The choice of car type from a given set corresponds to a discrete choice situation. A discrete-choice model is therefore used to evaluate this data. To shorten computation time, I use a Multinomial Logit model (MNL) rather than a Multinomial Probit model. The model is based on the following deterministic components of the utilities:

$$V_{ijn} = \alpha_i + \delta \cdot x_i + \gamma_i \cdot s_n . \qquad (4.2.1)$$

Index i stands for the car type and j indicates the choice set reported by household n. Parameter α_i is a constant specific to car type i, parameter vector δ captures weights relating to car attributes x_i, such as the car's price. For simplification, I determined that parameter vector δ does not vary between car types j, i.e. $\delta_i = \delta$. Parameter vector γ_i weights socio-demographic attributes of the households, including their income s_n. The error structure differs slightly from the standard MNL model due to the panel structure of the dataset.[235] The latent utilities U_{ijn} are defined as follows:

$$U_{ijn} = V_{ijn} + \eta_{in} + \zeta_{ijn} , \qquad (4.2.2)$$

where

$$\eta_{in} \sim iid\, N\left(0,\sigma_i\right) , \qquad (4.2.2a)$$

$$\zeta_{ijn} \sim iid\, F_\zeta\left(z\right) \text{ and} \qquad (4.2.2b)$$

$$i = 1,...,K , \;\; j = 1,...,J \text{ and } n = 1,...,N . \qquad (4.2.2c)$$

The cumulative density function $F_\zeta\left(z\right)$ corresponds to the standard extreme value distribution $F_\zeta\left(z\right) = \exp\left(-\exp\left(-z\right)\right)$. The random variable η_{in} is added due to the panel structure. The value of

[235] See also Appendix A2.1, section “The general structure of the OLS, the Probit and the Tobit model”.

η_{in} differs for each alternative and for each household, but remains the same for each choice set for which the household has to state its choice. It reflects the unobserved *a priori* preference of a household for a certain car type. I assume that variance σ_i can vary across alternatives *i*.[236] Note that considering values η_{in} as given, model (4.2.2) reflects a standard MNL model. Integer *K* reflects the total number of car types, *J* the total number of choice sets each household had to answer and *N* the total number of households. The set of cars in choice set number *j* for household *n* is denoted by S_{jn}.

Note that there is no explicit Maximum Likelihood function of model (4.2.2). To solve this problem, the probability for choosing car type *i* is computed conditional on random terms η_{in}. For this case, model (4.2.2) reflects a standard MNL model. The conditional probability $P_{Ijn} | \eta$ can then be computed as follows:

$$P_{I_{jn}jn} | \eta = \frac{e^{V_{I_{jn}jn} + \eta_{In}}}{\sum_{i \in S_{in}} e^{V_{ijn} + \eta_{in}}}, \tag{4.2.3}$$

with $I_{jn} = \arg\max_i U_{ijn}$,

where I_{jn} indicates the alternative chosen by household *n* in its choice set *j*, η denotes the set of all η_{in}, and $P_{I_{jn}jn}$ is the probability that household *n* chooses car type *i*. For computing the unconditional probability $P_{I_{jn}jn}$, terms η_{in} have to be integrated out:

$$P_{I_{jn}jn} = E_\eta \left(P_{I_{jn}jn} | \eta \right). \tag{4.2.4}$$

Since η_{in} are normally distributed and even more P_{Ijn} is a non-linear function of η_{in}, computing term $E_\eta(\bullet)$ by numerical integration is by far too demanding with respect to computation time. Probability $P_{I_{jn}jn}$ is therefore computed by simulation:

$$\hat{P}_{I_{jn}jn(\mathrm{SMLE})} = \frac{1}{S} \cdot \sum_{s=1}^{S} E_\eta \left(P_{I_{jn}jn} | \eta_{n(s)} \right), \tag{4.2.5}$$

where $\eta_{n(s)}$ is a vector containing one element $\eta_{in(s)}$ for each car type *i*. The $\eta_{n(s)}$ are independently drawn for each household *n* from the distribution $\eta_{in} \sim iid\, N(0, \sigma_i)$. The total number of draws is denoted by *S*. The simulated probability $\hat{P}_{I_{jn}jn}$ converges to the true probability $P_{I_{jn}n}$ when *S* approaches infinity. The parameters $\theta = \{ \alpha_{1..J}, \delta, \gamma, \sigma_{1..J} \}$ are computed using the Maximum

[236] It seems quite reasonable that preference for a Mini Cooper varies much more across households due to its fancy design: some may love it whilst others may strongly dislike it. In contrast, preference for a VW Golf may vary less, due to its rather modest design. The variance of unobserved preferences for car types is therefore expected to be lower for a VW Golf than for a Mini Cooper.

Simulated Likelihood (MSL) function.[237] The log-MSL function is based on the simulated probabilities $\hat{P}_{I_{jn}jn(\text{SMLE})}$. As shown by Walker (2002), all parameters $\theta = \{\alpha_{1..J}, \delta, \gamma, \sigma_{1..J}\}$ can be identified by solving the following maximisation problem:

$$\hat{\theta}_{MLE} = \arg\max_{\theta} \sum_{n=1}^{N} \ln\left(\hat{P}_{I_{jn}jn(\text{SMLE})} \mid \theta\right), \text{ where } \theta = \{\alpha_{1..J}, \delta, \gamma, \sigma_{1..J}\}. \quad (4.2.6)$$

In practice, however, convergence is usually very poor when numerically computing this problem. For this reason, it is recommended to set one parameter σ_i to a fixed value.[238]

4.3 Results

In this subchapter I will first state my hypothesis, namely that households are willing to formally pay for additional fuel efficiency exceeding the total amount of money saved on fuel expenditure. I will then proceed to present and discuss the results, followed by a proposal for further research with this dataset.

My hypothesis of economically rational behaviour

I start by presenting how my hypothesis that households are willing to pay for additional fuel efficiency exceeding the total amount of money saved on fuel expenditure can be tested using the model parameters. Let us assume that a household drives 10,000 kilometres per year. If this household's car fuel economy is increased by one litre per 100 kilometres, it would save 100 litres annually. Given a current fuel price of CHF 1.50 per litre and an expected car life time of nine years,

---

[237] For literature concerning the method of Simulated Maximum Likelihood estimation, see Train (2003: 237 ff.), Hajivassiliou and Ruud (1994: 2412 ff.), or Gourieroux and Monfort (1993: 18 ff.).

[238] As a standard routine, it is recommended to first run an optimisation where all parameters can be estimated. Since in most cases no stable optimum will be found, the computation process has to be broken up after a certain number of iterations. Next, parameter σ_i with the lowest estimated value is set to zero. This procedure was recommended by Professors Joan Walker and Michel Bierlaire. Unfortunately, I was unable to find any literature describing this procedure. Theory concerning the identifiability of parameters in choice models can be found in Walker et al. (2007) and Walker (2001).

and assuming that fuel prices increase annually by two percent less than the interest rate,[239] the net present value of an increase in fuel economy for each litre per 100 kilometres is CHF -1,292 for this household.[240] If this household behaves rationally, it would therefore pay at most CHF 1,292 more if a car consumes one litre per 100 kilometres less. The general formula that expresses the "willingness to pay" wtp_{rat}, given the household behaves rationally in the pure economic sense, is as follows:[241]

$$wtp_{rat} = \Delta e \cdot x_{km} \cdot p_{fuel} \cdot \frac{1}{100} \cdot \frac{1-q^T}{1-q}, \tag{4.3.1}$$

where $q = \dfrac{1}{1+r-\pi_{fuel}}$.

Variable Δe denotes the increase in fuel economy measured in litres per 100 kilometres of a certain car. x_{km} denotes the annual amount of kilometres driven by a household, and p_{fuel} is the fuel price at the time the car was purchased. The household expects the car to have a lifetime of T years. It also expects an interest rate of r and that fuel price will increase by an annual rate of π_{fuel}.

I now wish to show how the null-hypothesis, namely whether the households' willingness to pay corresponds to what is rational in a pure economic sense, can be tested. Assume that parameters δ_p and δ_e correspond to the cars' attributes "price" and "fuel economy". The willingness to pay for fuel economy can now be found by answering the following question: By how much can the price of a car increase if its fuel economy increases by one unit? The answer is simple: If the car's fuel economy increases by one unit, the utility V of that car increases by δ_e units. If the price of the car is increased by δ_e/δ_p units, the utility V of that car is again as at the beginning. The household is therefore indifferent to the initial car and the car that is more fuel efficient by one unit but that costs δ_e/δ_p units more. It follows that the willingness to pay for one unit of fuel efficiency is δ_e/δ_p. It can now be tested whether this is different to

$$\frac{\hat{\delta}_e}{\hat{\delta}_p} = wtp_{rat}, \tag{4.3.2}$$

[239] The average growth rate of the fuel price in the period 1975-2004 was 1.788%, see Bundesamt für Statistik, (2010b). The interest rate of a Swiss government bond with a duration of 30 years was 3.728% in 2004, see Schweizerische Nationalbank (2010). The difference was therefore approximately 2%. Note there is no consensus on the correct interest rate households impose when they discount their future expenditures. Values ranging from 3% to 6% are commonly used.

[240] This value can be computed using the following Formula (4.3.1).

[241] Note that the factor $1/100$ is needed, since the fuel economy is measured in litres per 100 kilometres and not per kilometre.

where $\hat{\delta}_p$ and $\hat{\delta}_e$ are the estimated parameters. Since the estimated parameters $\hat{\delta}_p$ and $\hat{\delta}_e$ are stochastic values, it has to be tested whether $\hat{\delta}_e/\hat{\delta}_p$ differs significantly from wtp_{rat}. I want to test whether coefficient $\left(\hat{\delta}_e/\hat{\delta}_p\right)/wtp_{rat}$ is one, which is equivalent to the following test:

$$\frac{\hat{\delta}_e}{\hat{\delta}_p}\cdot\frac{1}{wtp_{rat}}-1=0. \tag{4.3.3}$$

Since the parameters are estimated by a Maximum Likelihood estimation routine, their distribution is asymptotically normally distributed. The distribution of (4.3.3) can therefore be computed by using the covariances by which the joint normal distribution of the two parameters is defined. The simplest way to do this is to use the delta method, which yields the following:

$$\frac{\delta_e}{\delta_p}\cdot\frac{1}{wtp_{rat}}-1\sim\phi\left(\frac{\hat{\delta}_e}{\hat{\delta}_p}\cdot\frac{1}{wtp_{rat}},\ f_e^2\cdot\sigma_e^2+2\cdot f_e\cdot f_p\cdot\sigma_e^2+f_p^2\cdot\sigma_p^2\right), \tag{4.3.4}$$

where $f_e=\left[\partial\left(\frac{\delta_e}{\delta_p}\cdot\frac{1}{wtp_{rat}}\right)\Big/\partial\delta_e\right]_{\delta_e=\hat{\delta}_e,\,\delta_p=\hat{\delta}_p}=\frac{1}{\hat{\delta}_p}\cdot\frac{1}{wtp_{rat}}$ and $f_e=-\frac{\hat{\delta}_e}{\hat{\delta}_p^2}\cdot\frac{1}{wtp_{rat}}$.

Parameters σ_e^2, σ_p^2 and σ_{ep} denote the variances and the covariance of the covariance matrix that describes the joint normal distribution of parameters δ_p and δ_e.

I estimated the choice model using data concerning middle class cars.[242] Due to the limited number of 109 households – each responding to 21 choice sets – the number of household attributes that can be included is limited. I decided to capture each household's income, number of children, and the age and sex of the respondent.[243] As car attributes, I included the price of the car, fuel efficiency and the energy label. I omitted the attribute engine size, since the corresponding parameter was not significant.[244] Since it turned out that the impact of income and age on car choice is non-linear, I decided to use a splined function. I defined the spline points at a monthly income of CHF 7,000 and at an age of

[242] This dataset is presented in Subchapter 1.4.

[243] This choice resulted after comparing the estimation results of a number of different models. At the beginning, a vast number of explanatory variables were included in the model. I then eliminated the variables by stepwise extreme bound. To save computation time, this choice was based on a MNL model that ignored the panel structure of the data.

[244] Excluding this variable had virtually no impact on the estimated parameters δ_e and δ_p. Note that I propose why this parameter is insignificant later in this subchapter in the section entitled "Results".

45 years.[245] I computed parameter δ_e for three different household segments. The first segment "el" contains households which indicated that the information given on the energy label was very important to them when they recently decided to buy a car.[246] The second segment "work" consists of households for which the first priority for use of the car is to be able to commute to work. The third segment "others" consists of households that belong to neither group "el" or "work". The reason why I introduced this segmentation is that I expected the willingness to pay of households in these different segments to differ.[247] Note that some households belong to both segments "el" and "work". Table 4.3.1 shows some summary statistics of these segments.

	Share of observation	Mean driving distance
"others"	28.28%	16,167 km
"work"	36.36%	19,249 km
"el"	56.56%	21,583 km
Both "el" and "work"	21.21%	20,934 km

Table 4.3.1: Share of observations and mean driving distance of different household segments.

The values given in the table show that households in groups "el" and "work" clearly drive more kilometres on average than households in the group "others". This result is quite intuitive, since households that drive more kilometres have a greater incentive to care about their car's fuel consumption.

[245] The spline points define the points at which the slope of the corresponding variable can change. I chose the values 45 years and CHF 7,000 because I expected them to be in the range where preferences for certain car types, e.g. estate wagon versus hatch back, may change. In addition, these values do not differ greatly from the median of the corresponding variables, meaning that the problem of having insufficient observations in the lower or upper interval should not occur. An introduction to the theory of spline regression can be found in Greene (2003: 121 ff.).

[246] The household had to answer whether the information on the energy label played an important role when they recently purchased a car. They had to answer on a scale from 1 to 6, where "1" indicated "not important" and "6" "very important". I defined the group of households which indicated that the information on the energy label was important as households that ticked "5" or "6".

[247] I presume that households in segment "el" have a greater preference for fuel-efficient cars than households for which the information on the energy label was not important when they recently purchased a car. I therefore expect that households in segment "el" are more willing to pay for fuel efficiency. The same applies to households in the segment "work", which use their car to commute to work. I assume that these households are more aware of the car's cost since they may consider the car to be merely a means of transport. In contrast, households that do not use their car primarily for commuting to work may focus more on other car features, such as floorspace. I therefore expect that households in the segment "work" are more willing to pay for fuel efficiency than households from the other segments.

Results

Since I focus on the willingness to pay for car fuel efficiency, I will only present the results for parameters corresponding to fuel price δ_p and fuel efficiency δ_e, which are shown in Table 4.3.2.[248]

	$\hat{\delta}_i$	$\text{stdev}\left(\hat{\delta}_i\right)$	$\text{cov}\left(\hat{\delta}_{e,\cdot},\hat{\delta}_p\right)$	t-value	$P\left(\hat{\delta}_i\right) \neq 0$
$\delta_{e,others}$	-0.0327	0.0436	0.000122	-0.75	0.45
$\delta_{e,el}$	-0.1640	0.0601	0.000037	-2.72	0.01
$\delta_{e,work}$	-0.0740	0.0316	0.000104	-2.34	0.02
δ_p	-0.0357	0.0120	--	-2.97	0
Number of households: 109, number of observations: 1581					
Likelihood values (null, final) : (-1736.906, -1284.828)					
Likelihood ratio test: 904.157					
$\rho^2 = 0.260$, $\rho^2_{adj.} = 0.255$					
Number of random draws S per household: 800					

Note 1: All the standard deviations, covariances and t-values are "robust" estimators according to the software Biogeme.
Note 2: Parameter δ_p corresponds to the car price measured in CHF 1,000.

Table 4.3.2: Effects of fuel efficiency and price on car choice.

Note that the results in Table 4.3.2 show that all parameters, with the exception of $\delta_{e,others}$, are statistically significant. I now use these results to compute the willingness to pay for car fuel efficiency. All results are presented in Table 4.3.3. I also show the result of the willingness to pay wtp_{rat} that corresponds to a purely economically rational household. Finally, the table also contains the test statistics of the null-hypothesis, namely whether households in the different segments acts like a household that behaves purely economically rationally.

[248] I omit other interesting results, such as women's preferences for certain car models, those of high-income households or households with a large number of children.

				Test statistics		
Segment	wtp	wtp_{rat}	wtp/wtp_{rat}	sdev	t-value test	p-value
“others”	916	2,019	0.45368	0.58834	-0.92858	0.17655
“el”	4,594	2,695	1.70434	0.82546	0.85327	0.80325
“work”	2'073	2,404	0.86227	0.40131	-0.34320	0.36572

Note 1: The *wtp*, the wtp_{rat} and the test statistics are computed from (4.3.2), (4.3.3) and (4.3.4).
Note 2: The value wtp_{rat} is based on a fuel price of CHF 1.50 / litre, which was the price at the time of the survey, and the average driving distance of the corresponding household segment. I assumed that the household expected the annual rise of fuel prices to be 2% less than the interest rate. Further, I assumed that households do not plan to change their driving distance and that a car lifetime of nine years is expected.

Table 4.3.3: Effects of fuel efficiency and price on car choice.

The results show that the null-hypothesis cannot be rejected for all three segments on a 5% level. Nonetheless, I comment on the differences between wtp and wtp_{rat}. It is quite intuitive that households in segment “el” showed a greater willingness to pay for car fuel efficiency than those in the segment “other” that do not use the car to commute to work and that did not pay much attention to the information on the energy label. It seems that households in the group “other”, which primarily use their cars for other purposes than to commute to work, tend to value features that are relevant to practical utility, paying only little attention to car fuel consumption. In contrast, I do not find any obvious reason why the willingness to pay of households that use the car to commute to work as a first priority is lower than would be economically rational.

So far, I have shown that the willingness to pay for fuel economy differs between household segments “el”, “work” and “other” as defined above, but that the null-hypothesis – that households behave purely economically rationally – could not be rejected. Note that one problem of the previous hypothesis test could be that the null-hypothesis could not be rejected due to high standard errors of the estimated parameters.[249] Since this could have been caused by segmenting the households, I wish to test the hypothesis in a different way. This alternative test is based on the fact, according to (4.3.1), that the willingness to pay should be proportional to $e \cdot x_{km}$. In the next choice model, therefore, the term $e \cdot x_{km}$ will be used as an explanatory variable. The willingness to pay for fuel efficiency conditional on annual driving distance x_{km} can be computed using the corresponding parameter δ_{ex}. The hypothesis test is based on the following identity:

[249] There are basically two reasons why the standard deviation of the estimated parameters is high. First, the number of households in the dataset is rather low and the number of households in the various segments is even lower. To ensure the number of households in the segments not lower still, I did not estimate the coefficient for fuel consumption for the segments “el” and “work” separately. Second, there could be a high degree of heterogeneity in household behaviour within these segments.

$$\frac{\delta_p}{\delta_{ex}} = p_{fuel} \cdot \frac{1}{100} \cdot \frac{1-q^T}{1-q}, \text{ with } q = \frac{1}{1+r-\pi}. \tag{4.3.5}$$

To test whether this identity holds, I test whether the transformed identity

$$\frac{\delta_p}{\delta_{ex}} \cdot \frac{1}{wtr^*_{rat}} - 1 = 0, \text{ with } wtr^*_{rat} = p_{fuel} \cdot \frac{1}{100} \cdot \frac{1-q^T}{1-q}. \tag{4.3.6}$$

holds. Again, the test statistics is derived by the delta method, and is identical to (4.3.4).

The estimation results generated by this second model are presented in Table 4.3.4.

	$\hat{\delta}_i$	$\text{stdev}\left(\hat{\delta}_i\right)$	$\text{cov}\left(\hat{\delta}_{e,\cdot},\hat{\delta}_p\right)$	t-value	$P\left(\hat{\delta}_i\right) \neq 0$
δ_{ex}	-3.14E-06	7.64E-07	7.55E-10	-4.12	0
δ_p	-2.65E-02	8.70E-03		-3.05	0
Number of observations: 1544					
Likelihood values (null, final) : (-1696.257, -1419.024)					
Likelihood ratio test: 554.467					
$\rho^2 = 0.163$, $\rho^2_{adj.} = 0.132$					

Note 1: All the standard deviations, covariances and t-values are “robust” estimators according to the software Biogeme.

Note 2: Parameter δ_p corresponds to the car price measured in CHF 1,000. Since (4.3.6) relates to values in CHF, δ_p has to be divided by 1,000 when computing wtp^*.

Table 4.3.4: Effects of fuel efficiency multiplied by annual driving distance and price on car choice.

Note that the results in Table 4.3.4 show that both parameters are statistically significant. I now apply these results to compute the willingness to pay for car fuel efficiency per kilometre of driving wtp^*, which is presented in Table 4.3.5. I also show the result of the willingness to pay wtp^*_{rat} that corresponds to a purely economically rational household. Finally, the table also contains the test statistics of the null-hypothesis whether households in the different segments act like a household that behaves purely economically rationally.

				Test statistics		
	wtp^*	wtp^*_{rat}	wtp^* / wtp^*_{rat}	sdev	t-value test	p-value
	0.11360	0.12488	0.90982	0.4151	-0.2172	0.4140

Note 1: The wtp^* is equal to $wtp^* = \hat{\delta}_p / \hat{\delta}_{ex}$ and wtp^*_{rat} is equal to the right-hand side of (4.3.6).

Note 2: The value wtp_{rat} is based on a fuel price of CHF 1.50 / litre, which was the price at the time of the survey, and the average driving distance of the corresponding household segment. I assumed that households expected the annual rise of fuel prices to be 2% less than the interest rate. Further, I assumed that households do not plan to change their driving distance and that a car lifetime of nine years is expected.

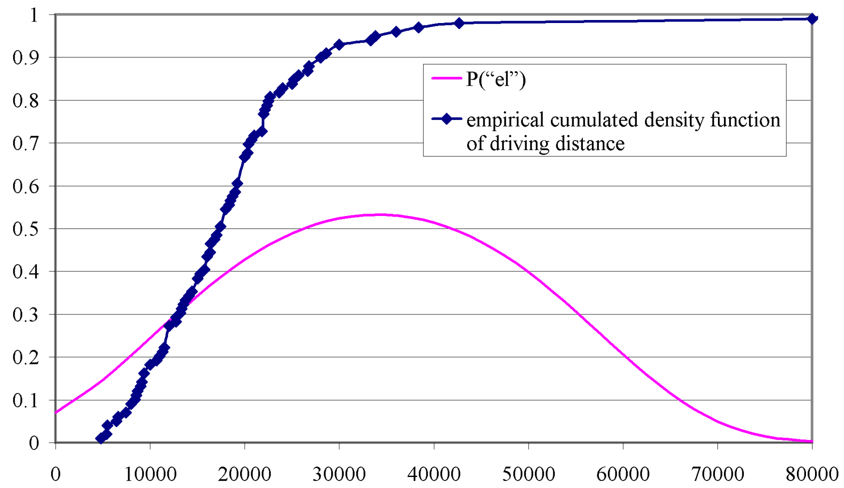
Note 3: The results are based on a pooled model. Based on the model that captures the panel structure, the software Biogeme did not find any values for the standard deviations of household-specific error terms.

Table 4.3.5: Effects of fuel efficiency and price on car choice.

The results show that the null-hypothesis cannot be rejected and that the willingness to pay is almost identical to the value of a household that behaves economically rationally. The hypothesis that households are willing to pay more for a more fuel-efficient car than they expect to save through reduced fuel expenditures throughout the lifetime of the car can therefore be rejected.

For completeness, I want to verify whether the specification of the alternative choice model is consistent. To this purpose, I wish to establish whether a household's willingness to pay is indeed proportional to the annual distance it drives. I therefore estimated a model that also contained the variable “fuel efficiency” *e,* measured in litres per 100 kilometres, in addition to the term $e \cdot x_{km}$. If the specification of the alternative choice model is correct, the parameter associated with fuel efficiency *e* should be zero. However, the result showed that this was not the case: this parameter was significantly negative and, moreover, the sign of the parameter associated with $e \cdot x_{km}$ even had the wrong sign, but was at least statistically insignificant. This is quite surprising, since, as expected, the share of household segment “el” increases with the households' driving distance. This relation is illustrated in Figure 4.3.1, which is based on the results of a Probit model that explains the probability of households belonging to segment “el”.[250]

[250] Only the driving distance and the square of the driving distance are used as explanatory variables. Note that when the household's income was also added to the model specification, the parameter values for the driving distance remained approximately the same, but became insignificant. The parameter accounting for the household's income showed a negative sign, but is also insignificant.



Note: The graph P("el") denotes the probability of a household that drives a certain annual distance belonging to segment "el". This value is forecast based on a Probit model.

Figure 4.3.1: The share of households that care about the energy label with respect to annual driving distance.

The result in this diagram shows that the share of households that care about the information contained in the energy label when they purchased their car increases with the households' driving distance. Nonetheless, some households with very low or very high driving distances behave very differently to what economical rationality would imply ("outliers"). To verify this claim, I ran a regression based on a dataset from which I eliminated households that usually drive less than 8,000 km or more than 60,000 km per year. The resulting parameter corresponding to term $e \cdot x_{km}$ based on this data showed the correct sign, even though it was not statistically significant.[251] The explanation for this difference is as follows: when the complete dataset was used, the model's parameters of interest are very sensitive with respect to the outcome of households driving very low or very high annual mileages. The result could therefore have been driven by the behaviour of a few households ("outliers"). The first group of households that do not behave economically rationally are the highly environmentally conscious households. Although such households drive very few kilometres, they tend to buy cars that consume low quantities of fuel. For this feature, they are willing to pay much more than the total fuel expenses they will save later. These households are the "outliers" with low kilometres. The second group I identify comprises two segments of households that drive many kilometres. Within this group, the first segment consists of households that drive many kilometres for professional reasons. Since such households are often reimbursed for their petrol, they do not care how much fuel the car consumes.[252]

[251] Note that the parameter referring to the variable energy efficiency e was still negative and significant.

[252] To verify this claim, I added the terms $e \cdot I_{at\_work}$ and $e \cdot x_{km} \cdot I_{at\_work}$ to the model. The term I_{at_work} equals one if the household uses the car for work. The results showed that the parameters corresponding to $e \cdot I_{at\_work}$ and $e \cdot x_{km} \cdot I_{at\_work}$ had a

The second segment comprises households that drive a lot simply because they like to drive. Such households tend to have little interest in environmental issues, and are therefore not willing to invest more in car fuel economy than is economically profitable. What is more, such households associate fuel-efficient cars with a lack of horsepower and comfort.[253] Their willingness to pay for fuel efficiency is therefore likely to be below what would be economically rational.[254] For these reasons, I conclude that some households which drive a lot tend to have a willingness to pay for fuel efficiency that is below the economic optimum. Further, I conclude that some households which drive only a few kilometres tend to have a willingness to pay for fuel efficiency that exceeds the economic optimum. I identify this as the reason why the parameters did not show the expected behaviour, and that not all households' behaviour corresponds to the economically rational behaviour as stated by identity (4.3.6).

Interestingly, however, in the aggregate the behaviour of household does not differ from what could be expected when assuming that all households behave economically rationally, which is a rather strong result of this survey.

---

positive sign. The sign of parameter δ_{ex} became negative. The sign of δ_{ex} also became negative when I omitted the observations of households that use their cars for work. Even though the estimated parameter δ_{ex} was statistically insignificant, I conclude that the claim is supported by these results. The claim that households that use their car for work drive much greater distances is also true. On average, such households drive 26,870 kilometres versus 17,377 kilometres driven by households that do not use their car for work. The share of households that use their car for work is 19%. If the results in Table 4.3.2 and 4.3.3 are re-estimated by a dataset from which households that use their car for work are eliminated the results change as follows: the willingness of households in the category “others” to pay for fuel efficiency corresponds exactly to the case that would result from economic rational behaviour. This is a sign that if households have to bear all the costs of driving, they would behave economically rationally. For more detailed results for this case, see Appendix A4.1.

[253] The assumption that many households care a lot about the power of their car is quite plausible: in this survey, 28% of respondents stated that the attribute “power” had top, second or third priority when they purchased their car. Note that I did not check this claim and the claim concerning environmentally friendly households due to a lack of variables that proxy “environmental friendy preferences” or “being a car lover with a high preference for car driving”. Note that an attempt to define such proxy variables to identify such groups can be made using so-called “Hybrid Choice Models”, see Börsch-Supan et al. (2002), Bolduc et al. (2005), Bolduc et al. (2008) and Bolduc et al. (2009).

[254] Note that another drawback of the questionnaire used by Wüstenhagen and Sammer (2007) is that they did not include any information concerning engine power. Instead, Wüstenhagen and Sammer (2007) provide information on the engine size. Although the engine size could be associated with engine power, it transpired that this variable had no statistically significant impact on car choice in the model. One reason for this could be that car taxes are based on engine size in many cantons. In the canton Aargau, for instance, car taxes increase by CHF 120 per additional litre of engine size. Assuming a car has a lifespan of nine year and that an interest rate of four percent is granted, the present net value of this tax is CHF 928. Hence the additional utility of the increase in engine power of larger engines is reduced by the additional taxes a consumer has to pay. Note that other cantons, e.g. the canton Solothurn, imposed similar tax amounts on engine size. For an overview on car taxes, see Schweizerischer Verband für elektrische und effiziente Strassenfahrzeuge e-mobile (2010).

Policy implications and the current trend of car demand in Switzerland

How much households are willing to pay for technologies that improve fuel efficiency can be concluded from the results. The key question is how much such technologies cost and by how much these technologies can improve car fuel efficiency. This question is rather difficult to answer, due to the many uncertainties with respect to the future costs of such technologies and future fuel prices. Different projections on the future market shares of technologies, such as hybrids or even pure electrically driven system, deviate strongly. The market share of cars driven by hybrid technology is currently very low, namely less than 2%.[255] It is quite understandable why this share is so low because, according to the calculations of an economically rational household, it would only pay off to buy such a car if the annual driving distance exceeds 44,000 kilometres.[256] It is also important to note that the technical possibilities to improve the fuel efficiency of conventional fuel combustion cars are very limited. There are two reasons for this: first, current driving technology is already close to the physically possible maximum. If the rather expensive hybrid technology is used, fuel efficiency can only be improved by an average of approximately 20% over conventional cars.[257] For motorway driving, the difference is even almost zero, since engines of conventional cars can run in a mode where efficiency is close to the theoretical maximal efficiency of that technology. Second, the demand for more aerodynamic car designs which are more narrow or lower to reduce the drag reference area is very limited due to customers' demand for spacious cars. I conclude from this that no improvements in fuel efficiency exceeding 20% for new car models can be expected in the near future, since hybrid technology seems to be the only one that may be cost effective to implement.[258]

---

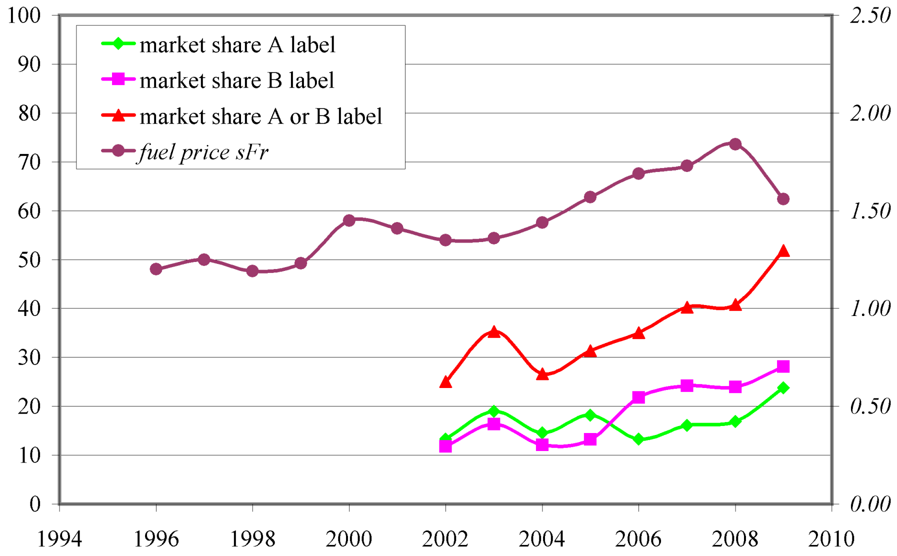
[255] The sales figures of the top sellers based on hybrid technology (2009) are Toyota Prius (1535), Honda Civic (245), Honda Insight (532) and Honda CR-Z (400). The number of these cars totals 2,712. The total number of cars imported to Switzerland is 294,239, see Vereinigung Schweizer Automobilimporteure (Auto-Schweiz) (2010). Note that only the total number (735) was listed for Honda Civic. I assumed that one third were Honda Civic hybrid models.

[256] I computed this figure by solving (4.3.1) for x_{km}. The data I used is based on the prices for a Honda Civic hybrid and a conventional Honda Civic. The current prices are CHF 35,700 versus a price of approximately CHF 29,700 for a model of the same car type with the same comfort features. The fuel economy of the two car types is 4.6 l / 100 km versus 5.7 l / 100 km, see www.honda.ch.

[257] For instance, the improvement in fuel economy of the Honda Civic hybrid compared to its conventional model is only 1.1 l / 100 kilometres, see previous footnote.

[258] The market share of purely electrically driven vehicles is almost zero, and the market potential in the near future is very low, since they cost much more than conventional vehicles. For instance, the purely electrically driven car Nissan Leaf will be sold for EUR 30,000 from autumn 2011, whereas the corresponding conventional model Nissan Tida costs EUR 16,000 and above. The difference is therefore at least EUR 10,000, so it will only pay off to buy this car in certain cases.

Finally, I present and discuss a number of recent trends visible on Swiss car markets. As can be seen in Figure 4.3.2, the share of imported cars with energy labels A and B, indicating high fuel efficiency, has increased in recent years.

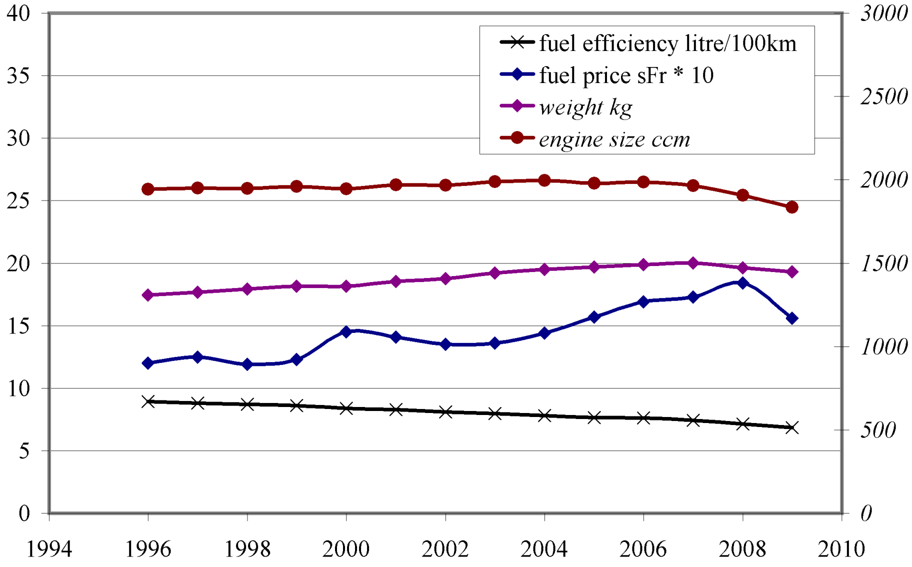


Data: Vereinigung Schweizer Automobilimporteure (Auto Schweiz) (2009).

Figure 4.3.2: The share of energy labels A and B of imported cars and fuel price.

Considering the trajectory of the fuel price, I conclude that households bought more cars with energy label A or B due to higher fuel prices and since households probably expect fuel prices to increase further. In addition, during the recession in late 2008 and 2009, households might have chosen to buy fuel-efficient cars. Further, during this period of recession, many car manufacturers expanded their offer of more fuel-efficient car models. Note that since the energy label level is corrected by the cars' weight such that large and heavy cars may also have an energy label A,[259] the diagram above does not necessarily imply that the average fuel consumption of cars decreased, since it could have been the case that households bought heavier cars.

[259] In the formula defining the energy label level, the car's weight stands in the denominator. For the exact definition of the energy label levels for cars, see Eidgenössisches Departement für Umwelt, Verkehr, Energie und Kommunikation (2010) and (2011).

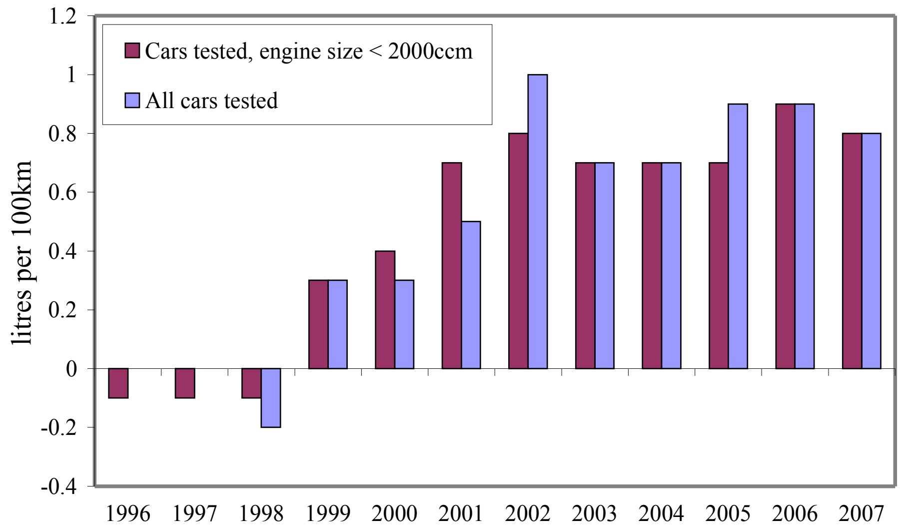


Data: Vereinigung Schweizer Automobilimporteure (Auto Schweiz) (2009).

Figure 4.3.3: The average car weight, engine size and fuel price.

Figure 4.3.3 shows that the average fuel consumption of the car fleet actually decreased, even during the period when the average car weight increased. Interestingly, the average car weight decreased only in the years of the recession, in late 2008 and 2009. The finding is similar for the average engine size. Note that the average fuel consumption in Figure 4.3.3 is based on a normalised driving cycle.[260] An interesting finding in this context is that the difference between the cars' actual fuel consumption and their consumption according to the normalised driving cycle has increased over time. Touring Club Schweiz (TCS) conducted a survey and summarised the actual measured fuel consumption under typical conditions for a vast number of vehicles over several years. They compared their results with the values corresponding to the normalised driving cycle, and computed the average difference for each year. Since cars that are tested by TCS form a roughly representative sample, I regard the difference they found to be realistic. The findings of TCS are shown in Figure 4.3.4.

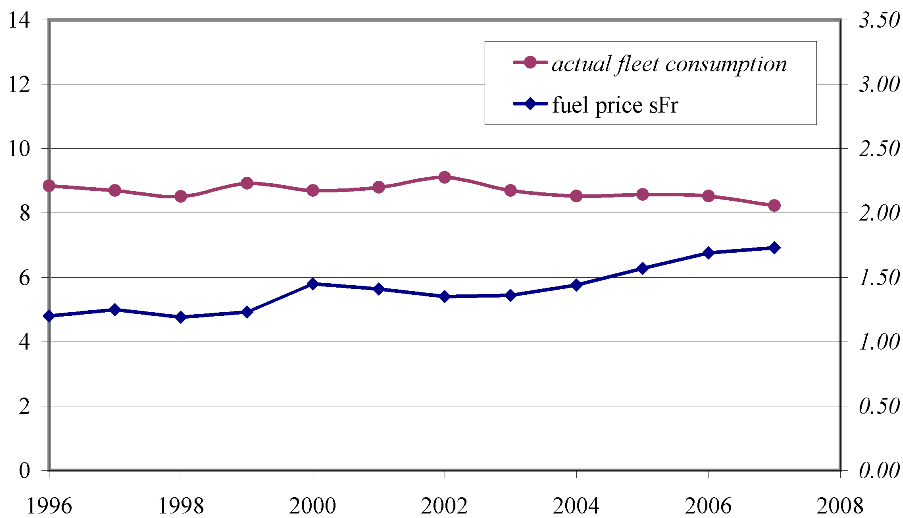
[260] It is the 1999/100/EC norm cycle defined by EU legislation.



Source: This diagram was taken from Touring Club der Schweiz (TCS 2008: 26).

Figure 4.3.4: Difference between norm and real consumption of cars.

These results imply that in the period from 1996 to 1999, the values corresponding to the normalized driving cycle met the values under real conditions. After 1999 the difference between the values under real and normalized conditions increased and remained at a level of about 0.8 litres / 100 kilometres. Using these differences, the actual consumption of the imported car fleet can be computed, see Figure 4.3.5.



Data: Vereinigung Schweizer Automobilimporteure (Auto Schweiz) (2009) and Touring Club der Schweiz (TCS 2008: 26).

Figure 4.3.5: The actual consumption of the imported car fleet.

It can be seen that the actual fuel consumption of imported cars remained almost unchanged in the period 1996 to 2007. Further, the diagram shows that when the fuel price increased, the average fuel consumption of the car fleet decreased.[261] This is again a strong argument for the hypothesis that households behave rationally on average. The question that remains is whether there has been technical progress in fuel economy. The answer is that small progress has been made which, however, is partly compensated by three trends. First, due to increasing incomes, households' need for additional functionalities that consume non-drive-related energy, such as air conditioning, has increased. Since at the same time these functionalities became cheaper, due to progress and economics of scale in manufacturing, they are now much more frequently built into cars. Second, households' demand for more spacious cars leads to greater weight – see Figure 4.3.3 – and thus to higher drag. Third, there is a trend towards four-wheel drive cars. The proportion of these cars has risen from 12% (1997) to 26% (2009), see Bundesamt für Statistik (2009b). Note that a four-wheel drive increases fuel consumption by 0.5 litre / 100 kilometres, see Touring Club der Schweiz (TCS) (2010).

I conclude from these trends shown in the diagrams and the results of the models presented in the this chapter that, given the current fuel prices and regulations and a continuing growth in income, there will be virtually no reduction in the average fleet consumption in future years.

[261] I also therefore question whether any reduction in fleet consumption can be attributed to the introduction of the energy label in 2002, after which car retailers have had to attach such a label to all cars in their show room. The slight drop in fuel consumption after 2003 can be attributed to the increase in fuel prices.

Literature

Allgemeiner Deutscher Automobil-Club (ADAC), 2010, "Durchschnittspreise Kraftstoffe seit 1950", Allgemeiner Deutscher Automobil-Club (ADAC), 2010. http://www1.adac.de/Auto_Motorrad/Tanken/zahlen_fakten/Die_Entwicklung_der_jaehrlichen_Durchschnittspreise_fuer_Kraftstoffe/default.aspComponentID=4252&SourcePageID=10100

American Petroleum Institute, 2010, "Gasoline Taxes", October 2010. http://www.api.org/statistics/fueltaxes/upload/October_2010_gasoline_and_diesel_summary_pages.pdf

Axhausen, Kay W. and L. Hurni (eds.), 2005, "Zeitkarten Schweiz 1950-2000", IVT and IKA, ETH Zürich, Zürich, 2005.

Axhausen, Kay W. and Alexander Erath, 2010, "Long term fuel price elasticity: Effects on mobility tool ownership and residential location choice", Bundesamt für Energie, 2010, Publikation 29015. http://www.bfe.admin.ch/php/modules/enet/streamfile.php?file=000000010334.pdf&name=000000290158

Baranzini, Andrea, David Neto and Sylvain Weber, 2009, "Élasticité-prix de la demande d'essence en Suisse", 2009, Bundesamt für Energie, 2009, Publikation 290036. http://www.bfe.admin.ch/forschungewg/02544/02807/02808/index.html?lang=de&dossier_id=03910

Bertrand, Marianne and Sendhil Mullainathan, 2001, "Do People Mean What They Say? Implications for Subjective Survey Data", The American Economic Review, Vol. 91, No. 2, Papers and Proceedings of the Hundred Thirteenth Annual Meeting of the American Economic Association (May, 2001), 67-72.

Bhat, Chandra R., 2005, "A multiple discrete-continuous extreme value model: formulation and application to discretionary time-use decisions", Transportation Research Part B 39 (2005), 679-707.

Bhat, Chandra R., 2006, "Household vehicle type holdings and usage: an application of the multiple discrete-continuous extreme value (MDCEV) model", Transportation Research Part B 40 (2006), 35-53.

Bhat, Chandra R., 2008. "The multiple discrete-continuous extreme value (MDCEV) model: role of utility function parameters, identification considerations, and model extensions", Transportation Research Part B 42 (3), 274-303.

Bichel, P.J. and David A. Freedman, 1981, “Some asymptotic theory for the bootstrap”, Annals of Statistics, 9, 1196-1217.

Bolduc Denis, Moshe Ben-Akiva, A. Michaud and J. Walker, 2005, “Hybrid Choice Models: Applicability to Large Scale Models”, in Integrated Land-Use and Transportation Models: Behavioral Foundations, M. E. H. Lee- Gosselin and Sean Doherty, Chapter 12, 275-302.

Bolduc Denis, Nathalie Boucher and Ricardo Alvarez-Daziano, 2008, “Hybrid Choice Modeling of New Technologies for Car Choice in Canada”, Transportation Research Record: Journal of the Transportation Research Board of the National Academies, Washington, D.C., Volume 2082 / 2008, 63-71.

Bolduc Denis and Ricardo Alvarez-Daziano, 2009, “On estimation of Hybrid Choice Models”, Working Paper, Département d’économique, Université Laval.

Board On Energy and Environmental Systems, 2002, “Effectiveness and Impact of Corporate Average Fuel Economy (CAFE) Standards (2002)”, The National Academies. http://books.nap.edu/openbook.php?record_id=10172&page=9

Börsch-Supan, Axel, Moshe Ben-Akiva, Kenneth Train, and Daniel L. McFadden, 2002, “Hybrid Choice Models: Progress and Challenges”, No. 02009, MEA discussion paper series, Mannheim Research Institute for the Economics of Aging, University of Mannheim. http://econpapers.repec.org/RePEc:xrs:meawpa:02009

Bundesamt für Statistik, 2001, “Mikrozensus 2000 zum Verkehrsverhalten”, Bundesamt für Statistik, Neuenburg (Switzerland) 2001.

Bundesamt für Statistik, 2006a, “Mikrozensus 2005 zum Verkehrsverhalten”, Bundesamt für Statistik, Neuenburg (Switzerland) 2006.

Bundesamt für Statistik, 2006b, “Mikrozensus 2005 zum Verkehrsverhalten, Kurzversion Fragebogen (Hauptbefragung)”, Bundesamt für Statistik, Neuenburg (Switzerland) 2006. http://www.portal-stat.admin.ch/mz05/files/de/00.xml

Bundesamt für Statistik, 2006c, “Schweizerische Strassenverkehrszählung 2005, Detailresultate - Kanton ZH, Zählstelle Brüttisellen N (AB), ZH 751N”. http://www.portal-stat.admin.ch/ssvz/docs/Zst0751N.pdf

Bundesamt für Statistik, 2007a, “Einkommens- und Verbrauchserhebung EVE 2000-2005”, Bundesamt für Statistik, Neuenburg (Switzerland) 2007.

Bundesamt für Statistik, 2007b, "Mobilität in der Schweiz, Ergebnisse des Mikrozensus 2005 zum Verkehrsverhalten". http://www.are.admin.ch/themen/verkehr/00256/00499/00502/index.html?lang=de

Bundesamt für Statistik, 2007c, "Mobilität in der Schweiz, Wichtigste Ergebnisse des Mikrozensus 2005 zum Verkehrsverhalten". http://www.are.admin.ch/themen/verkehr/00256/00499/00502/index.html?lang=de

Bundesamt für Statistik, 2008, "Leistungen des motorisierten privaten Personenverkehrs auf der Strasse, Aktualisierte Zeitreihe bis 2008 ", Bundesamt für Statistik, Neuenburg (Switzerland) 2008. http://www.bfs.admin.ch/bfs/portal/de/index/infothek/erhebungen__quellen/blank/blank/svps/02.Document.125582.pdf

Bundesamt für Statistik, 2009a, "Konsumentenpreisindex, LIK, Durchschnittspreise für Benzin und Diesel, Monatswerte", Bundesamt für Statistik, Neuenburg (Switzerland) 2009. http://www.bfs.admin.ch/bfs/portal/de/index/themen/05/02/blank/key/durchschnittspreise.Document.88015.xls

Bundesamt für Statistik, 2009b, Strassenfahrzeuge in der Schweiz. Inverkehrsetzung neuer Strassenfahrzeuge, 2009. http://www.bfs.admin.ch/bfs/portal/de/index/themen/11/03/blank/key/fahrzeuge_strasse/inverkehrsetzung.html#parsys_59381

Bundesamt für Statistik, 2010a, "Demografische Indikatoren in den EU- und EFTA-Staaten", Neuenburg (Switzerland) 2010.

Bundesamt für Statistik, 2010b, "LIK, Totalindex auf allen Basen seit Einführung", Bundesamt für Statistik, Neuenburg (Switzerland), 2010. http://www.bfs.admin.ch/bfs/portal/de/index/themen/05/02/blank/ml_data.html

Bundesamt für Statistik, 2010c, "Öffentlicher Verkehr, Eisenbahn, Zeitreihen, Eisenbahnen 1990-2007: Betriebsleistungen - Leistungen der Züge", Bundesamt für Statistik, Neuenburg (Switzerland) 2010. http://www.bfs.admin.ch/bfs/portal/de/index/themen/11/07/02/blank/03/02.html

Bundesamt für Statistik, 2010d, "Mobilität und Verkehr 2010", Bundesamt für Statistik, Neuenburg (Switzerland) 2010. http://www.bfs.admin.ch/bfs/portal/de/index/themen/11/07/02/blank/03/02.html

Bundesamt für Statistik, 2010e, "Statistische Grundlagen, die 7 Grossregionen der Schweiz", Bundesamt für Statistik, Neuenburg (Switzerland) 2010. http://www.bfs.admin.ch/bfs/portal/de/index/regionen/11/geo/analyse_regionen/02a.html

Bundesamt für Strassen ASTRA, 2007, "Verkehrsentwicklung und Verfügbarkeit der Nationalstrassen", Jahresbericht 2006, Bundesamt für Strassen ASTRA, 2007. www.news.admin.ch/NSBSubscriber/message/attachments/9208.pdf

Bundesamt für Strassen ASTRA, 2010, "Jahresberichte über den Verkehrsfluss und das Staugeschehen auf den Nationalstrassen 2004-2008 ", Jahresberichte 2004-2008, Bundesamt für Strassen ASTRA, 2010. http://www.astra.admin.ch/dokumentation/00119/00216/index.html?lang=de

Bundesamt für Verkehr, 2006, "Evaluation Bahn 2000, 1. Etappe Schlussbericht", Bern, 14 December 2006. http://www.infras.ch/downloadpdf.php?filename=B7081a-05a_Eval_B21_Schlussbericht_dt.pdf

Bundesministerium der Finanzen, 2005, "Kraftstoffpreise und Kraftstoffbesteuerung", September 2005. http://www.bundesfinanzministerium.de/nn_4134/DE/BMF__Startseite/Aktuelles/Monatsbericht__des__BMF/2005/09/050920agmb006,templateId=raw,property=publicationFile.pdf

Bundesministerium für Verkehr, Bau und Stadtentwicklung, 2006, "Verkehr in Zahlen", Bundesministerium für Verkehr, Bau und Stadtentwicklung, Berlin. http://www.gesis.org/fileadmin/upload/dienstleistung/daten/soz_indikatoren/Schluesselindikatoren/Verkehr.pdf

Bühler, Ralph and Uwe Kunert, 2008, "Trends und Determinanten des Verkehrsverhaltens in den USA und in Deutschland", Bundesministerium für Verkehr, Bau und Stadtentwicklung (BMVBS), Project number 70.0802/2006. http://www.bmv.de/Anlage/original_1060869/Trends-und-Determinanten-des-Verkehrsverhaltens-in-den-USA-und-in-Deutschland.pdf

Cameron, Colin A. and Pravin K. Trivedi, 2005, "Microeconometrics: Methods and Applications", New York: Cambridge University Press.

Dahl, C., 1995: "Demand for transportation fuels: a survey of demand elasticities and their components", The Journal of Energy Literature, 1, 3-27.

Dargay, Joyce M., 2001, "The effect of income on car ownership: evidence of asymmetry", Transportation Research Part A: Policy and Practice, Volume 35, Issue 9, November 2001, 807-821.

Dargay, Joyce and Dermot Gately, 1997, "Income's Effect On Car and Vehicle Ownership, Worldwide: 1960-2015", Working Papers 97-33, C.V. Starr Center for Applied Economics, New York University.

Dargay, Joyce M., Dermot Gately and Martin Sommer, 2007, "Vehicle Ownership and Income Growth, Worldwide: 1960-2030", The Energy Journal, International Association for Energy Economics, Vol. 28(4), 143-170.

De Jong, G. C., 1990, "An indirect utility model of car ownership and private car use", European Economic Review, Elsevier, Vol. 34(5), 971-985, July.

De Jong, G. C., 1997, "A micro-economic model of the joint decision on car ownership and car use", in: P. Stopher and M. Lee-Gosselin (eds.): Understanding travel behaviour in an era of change; Pergamon, Oxford.

Deutsches Institut für Wirtschaftsforschung (DIW), 2002, "Fahrleistungen und Kraftstoffverbrauch im Straßenverkehr", Wochenbericht, DIW, No. 51-52/2002.

Deutsches Institut für Wirtschaftsforschung (DIW), 2010, "Kraftfahrzeugverkehr - Fahrleistungen nach Kraftfahrzeugarten", DIW, 2010.

Dubin, Jeffrey A. and Daniel L. McFadden, 1984, "An Econometric Analysis of Residential Electric Appliance Holdings and Consumption", Econometrica, Vol. 52, No. 2 (Mar., 1984), 345-362.

Eidgenössisches Departement für Umwelt, Verkehr, Energie und Kommunikation, 2010, "Verordnung des UVEK über Angaben auf der Energieetikette von neuen Personenwagen of 8 September 2004" (as at 1 July 2010), 730.011.1. http://www.admin.ch/ch/d/sr/7/730.011.1.de.pdf

Eidgenössisches Departement für Umwelt, Verkehr, Energie und Kommunikation, 2011, "Energieverordnung (EnV)" of 7 December 1998 (as at 1 January 2011), 730.01. http://www.admin.ch/ch/d/sr/7/730.01.de.pdf

Eisenbahntechnische Rundsschau, 2007, "News: SBB: Täglich 180 Züge mit 200 km/h auf der NBS". http://www.eurailpress.de/search/article/view//sbb_taeglich_180_zuege_mit_200_kmh_auf_der_nbs.html

Energy Information Administration (EIA), 2010, "Gallon Real Gasoline Pump Price: Annual Average 1919-2011", Energy Information Administration 2010. www.eia.doe.gov/emeu/steo/pub/fsheets/real_prices.xls

Engle, R. F. and W. J. Granger, 1987, "Co-integration and Error Correction: Representation, Estimation, and Testing", Econometrica, 55(2): 251-276.

Fang, Hao Audrey, 2008, "A discrete-continuous model of households' vehicle choice and usage, with an application to the effects of residential density", Transportation Research Part B 42 (2008), 736-758.

Federal Highway Administration (FHA), 2010, "Historical Monthly VMT Report", United States Department of Transportation - Federal Highway Administration (FHA), 2010. http://www.fhwa.dot.gov/policyinformation/travel/tvt/history/

Franzen, M. and Thomas Sterner, 1995, "Long-run Demand Elasticities for Gasoline", in: Baker, T. et al. (eds.), Global Warming and Energy Demand", Routeledge, London and New York, 1995.

Goodwin, P and Dargay, J and Hanly, M., 2004, "Elasticities of road traffic and fuel consumption with respect to price and income: a review", Transport Reviews, 24 (3) , 275-292.

Gourieroux, C. and A. Monfort (1993), "Simulation-based inference: A survey with special reference to panel data models", Journal of Econometrics 59, 5-33.

Golob, Thomas F., David S. Bunch and David Brownstone, "A Vehicle Usage Forecasting Model Based on Revealed and Stated Vehicle Type Choice and Utilization Data", Journal of Transport Economics and Policy Vol. 31 (1997), 69-92.

Graham, Daniel and Stephen Glaister, 2002, "Review of income and price elasticities of demand for road traffic", Final Report. Centre for Transport Studies, Imperial College of Science, Technology and Medicine.

Greene, William H., 2003, "Econometric analysis", Prentice Hall, 5th edition, Upper Saddle River, NJ, 2003.

Grubb, M., 1995, "Asymetrical Price Elasticities of Energy Demand", in: Baker, T. et al. (eds.), Global Warming and Energy Demand, Routeledge, London and New York, 1995.

Hajivassiliou, V. and P. Ruud (1994), 'Classical estimation methods for LDV models using simulation', in R. Engle and D. McFadden, eds., Handbook of Econometrics, North-Holland, Amsterdam, 2383-2441.

Hanly, M., Dargay, J. and Goodwin, P., 2002, "Review of Income and Price Elasticities in the Demand for Road Traffic", Report 2002/13 (London: ESRC Transport Studies Unit, University College London). http://www.cts.ucl.ac.uk/tsu/elasfinweb.pdf

Hausman, Jerry A., 1981, "Exact Consumer's Surplus and Deadweight Loss", The American Economic Review, Vol. 71, No. 4. (Sep., 1981), 662-676.

Heston, Alan, Robert Summers and Bettina Aten, 2009, "Penn World Table Version 6.3, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania", August 2009. http://pwt.econ.upenn.edu/php_site/pwt63/pwt63_form.php

Institut für Energie und Umweltforschung, 2010, “EcoPassenger: Methodology and Data”. http://ecopassenger.hafas.de/hafas-res/download/Ecopassenger_Methodology_Report.pdf

Johansson, Olof and Lee Schipper, 1997, “Measuring the Long-Run Fuel Demand of Cars”, Journal of Transport Economics and Policy, September 1997.

Kremers, Hans, Peter Nijkamp and Piet Rietveld, 2002, “A meta-analysis of price elasticities of transport demand in a general equilibrium framework”, Economic Modelling 19 (2002), 463-485.

Litman, Todd, 2009, “Transportation Elasticities, How Prices and Other Factors Affect Travel Behavior”, Victoria Transport Policy Institute, British Columbia, Canada. http://www.vtpi.org/elasticities.pdf

Little, R.J.A. and D.B. Rubin, 1987, “Statistical Analysis with Missing Data”, New York: John Wiley & Sons.

Louviere, Jordan, Train, Kenneth, Moshe Ben-Akiva, Chandra Bhat, David Brownstone, Trudy Ann Cameron, Richard T. Carson, J. Deshazo, , Denzil G. Fiebig, William H Greene, David Hensher and Donald M. Waldman, (2005), “Recent Progress on Endogeneity in Choice Modeling”, Marketing Letters, 16, Issue 3, 255-265. http://econpapers.repec.org/RePEc:kap:mktlet:v:16:y:2005:i:3:p:255-265

Mas-Colell, Andreu, Michael D. Whinston and Jerry R. Green, 1995, “Microeconomic Theory”, Oxford Univ. Press, 1995.

Meyer, Grischa, 2009, “Analyse und technisch-ökonomische Bewertung von Gesetzesfolgen im Individualverkehr - dargestellt am Beispiel der Automobilindustrie Japans und Deutschlands”, Dissertation, Carl-Friedrich-Gauß-Fakultät an der Technischen Universität Carolo-Wilhelmina zu Braunschweig.

Mueller, Michel Ganesh, 2010, “Energy-efficient cars. Studies on consumer-technology interaction”, Dissertation, ETH (2010), ethz-a-006072780. http://e-collection.ethbib.ethz.ch/view/eth:1299

Mineralölverband e.V., 2010, “Entwicklung der Mineralölsteuersätze auf Kraftstoffe 1950–2007”. http://www.mwv.de/cms/front_content.php?idcat=14&idart=60

National Highway Traffic Safety Administration (NHTSA), 2010, “New passenger car fleet average characteristics”. http://www.nhtsa.gov/cars/rules/CAFE/NewPassengerCarFleet.htm

Oum, Tae Oum, W. G. Waters and Jong-Say Yong, 1992, "Concept of Price Elasticities of Transport Demand and Recent Empirical Estimates", Journal of Transport Economic and Policy", May 1992, 139-154.

Peter, Martin, N. Schmidt and M. Maibach, 2002, "Evaluation kurzfristiger Benzinpreiserhöhungen", Forschungsauftrag der Vereinigung Schweizerischer Verkehrsingenieure (SVI), SVI 2000/442.

Progtrans, 2004, "Analyse von Änderungen des Mobilitätsverhaltens - insbesondere der PKW Fahrleistung - als Reaktion auf geänderte Kraftstoffpreis", Research project No. 96.0756/2002/ e, 2004. http://www.ivtverkehrsforschung.de/pdf/Kraftstoffpreise_und_Mobilitaet.pdf

Puhani, Patrick A., 2000, "The Heckman Correction for Sample Selection and Its Critique", Journal of Economic Surveys, Blackwell Publishing, Vol. 14(1), 53-68.

Ramjerdi, F. and L. Rand, 1992, "The National Model System for Private Travel", Institute of Transport Economics. TØI rapport 150/1992. Oslo.

Research and Innovative Technology Administration, Bureau of Transportation Statistics, 2010, "Average Fuel Efficiency of U.S. Passenger Cars and Light Trucks". http://www.bts.gov/publications/national_transportation_statistics/html/table_04_23.html

Ruud, P., 2000, "An Introduction to Classical Econometric Theory", Oxford University Press, New York, NY: 2000.

Simma, Anja, 2003, "History of the Swiss travel surveys", conference paper, 3rd Swiss Transport Research Conference in Ascona, 2003. http://www.strc.ch/conferences/2003/simma.pdf

Schleiniger, Reto, 1995, "The Demand for Gasoline in Switzerland - in the Short and in the Long Run", working paper, Institute for Empirical Research in Economics 9503, 1995.

Schweizerische Bundesbahnen, 2004-2009, "annual reports", Schweizerische Bundesbahnen 2004-2009. http://mct.sbb.ch/mct/en/konzern_dienstleistungen/konzern_medien/konzern_publikationen/konzern_geschaeftsbericht.htm?=

Schweizerische Nationalbank (SNB), 2010, "Zinssätze und Devisenkurse Mai 2010", Schweizerische Nationalbank (SNB), 2010. http://www.snb.ch/de/iabout/stat/statpub/akziwe/stats/akziwe;jsessionid=3E5212F376153586F4210501C48B46BD.tomcat2-snb-prod

Schweizerischer Verband für elektrische und effiziente Strassenfahrzeuge e-mobile, 2010,"Kantonale Motorfahrzeuge". http://www.e-mobile.ch/index.php?pid=de,2,4

Shao, Jun, 2003, “Mathematical statistics”, Springer, New York.

Simma, Anja, 2004, “Geschichte des schweizerischen Mikrozensus zum Verkehrsverhalten”, Bundesamt für Raumentwicklung, 2004. http://www.are.admin.ch/themen/verkehr/00256/00499/00500/index.html?lang=de

Statistisches Bundesamt, 2010, “Verbraucherpreisindizes für Deutschland”, Statistisches Bundesamt, Wiesbaden 2010. http://www.destatis.de/jetspeed/portal/cms/Sites/destatis/Internet/DE/Navigation/Statistiken/Preise/Verbraucherpreise/Tabellen.psml

Sterner, T. and C. Dahl (1992): “Modelling transport fuel demand”, in T Sterner (ed.) International Energy Economics, Chapman and Hall, London, 65-79.

Sterner, T., C. Dahl and M. Franzen (1992): “Gasoline tax policy: carbon emissions and the global environment”, Journal of Transport Economics and Policy, 26, 109-19.

Train, K., (2003), ”Discrete Choice Methods with Simulation”, Cambridge University Press, 2003.

U.S. Department of Energy, 2010, “What determines the price of gasoline”. http://www.fueleconomy.gov/feg/gasprices/faq.shtml

Touring Club der Schweiz (TCS), 2001, “Kilometerkosten”, Touring Club der Schweiz.

Touring Club der Schweiz (TCS), 2007, “Kosten eines Musterautos”. http://www.tcs.ch/main/de/home/auto_moto/kosten/kilometer/musterauto.html

Touring Club der Schweiz (TCS), 2008, “Treibstoffverbrauch Werksangabe vs. Praxis”, 2008. http://www.tcs.ch/main/de/home/auto_moto/umwelt_energie/treibstoff.RightColumn.0002.CtxLinkDownloadFile4.pdf/KN63TreibstoffverbrauchPW_de.pdf

Touring Club der Schweiz (TCS), 2010, “Vergleich 4x4 vs. 4x2”, 2010. http://www.tcs.ch/etc/medialib/main/rubriken/auto_moto/pdf/andere.Par.0113.File.tmp/4065_infotechTCS-Vergleich_Antrieb_SUV_4x4_vs_4x2_de.pdf

Varian, Hal R., 1992, “Microeconomic Analysis”, 3rd edition, Norton, 1992.

Varian, Hal R., 1994, “Mikroökonomie“, 3rd edition, Oldenburg Verlag, 1994.

Verband öffentlicher Verkehr, 2007, “Die Geschichte des Generalabonnements”, Verband öffentlicher Verkehr, 2007. http://www.voev.ch/dcs/users/6/Geschichte_GA_d.pdf

Vereinigung Schweizer Automobilimporteure (Auto Schweiz), 2009, "14. Berichterstattung im Rahmen der Energieverordnung über die Absenkung des spezifischen Treibstoff-Normverbrauchs von Personenwagen", 2009. http://www.bfe.admin.ch/energieetikette/00886/index.html?dossier_id=00959&lang=de

Vereinigung Schweizer Automobilimporteure (Auto-Schweiz), 2010, "Modellstatistik Personenwagen 2010 ", 2010. http://www.auto-schweiz.ch/Personenwagen_nach_Modellen.html

Walker, Joan L., 2001, Extended Discrete Choice Models: Integrated Framework, Flexible Error Structures and Latent Variables, 2001, Dissertation, Massachusetts Institute of Technology, Transportation Systems. http://hdl.handle.net/1721.1/32704

Walker, Joan L., Ben-Akiva M., Bolduc D., 2007, "Identification of parameters in normal error component Logit-Mixture (NECLM) models", Journal of Applied Econometrics, Vol. 22, No. 6, 1095-1125.

Wüstenhagen, Rolf and Katharina Sammer, 2007, "Wirksamkeit umweltpolitischer Anreize zum Kauf energieeffizienter Fahrzeuge: Eine empirische Analyse Schweizer Automobilkunden", Zeitschrift für angewandte Umweltforschung (ZAU), Vol. 18 (2007), Issue 1, 61-78.

Appendix

A 1 Appendix to Chapter 1

A 1.1 The dataset of Wüstenhagen and Sammer (2007): A questionnaire sheet

The survey of Wüstenhagen and Sammer (2007) was based on the following questionnaire sheets:

Wenn Sie heute einen Neuwagen kaufen, für welches Modell würden Sie sich innerhalb dieser Auswahl entscheiden?		
Peugeot 206	Mercedes A-Klasse	Opel Zafira
Hubraum 1.2 l	Hubraum 2.0 l	Hubraum 1.6 l
Treibstoffverbrauch 5.6 (l/100km)	Treibstoffverbrauch 6.7 (l/100km)	Treibstoffverbrauch 7.8 (l/100km)
Treibstoffart Diesel	Treibstoffart Benzin	Treibstoffart Diesel
15'450 CHF	22'143 CHF	24'374 CHF
		
Welches dieser drei Modelle würden Sie kaufen? Bitte zutreffende Antwort ankreuzen!		
☐ 1	☐ 2	☐ 3
☐ 4 Keines dieser Modelle kommt für mich in Frage, weil ……………………………………………………………		

Table A1.1.1: A questionnaire sheet for car choice[1]

[1] See Table in Wüstenhagen and Sammer (2007: 78).

The choices may consist of the following car types and values of the attributes:

	Small cars	Middle class cars
Brand and type	VW Polo Opel Corsa Peugeot 206 Toyota Yaris Fiat Punto Renault Clio Mercedes A-Klasse Opel Zafira	VW Golf Peugeot 307 Audi A4 Toyota Corolla Ford Mondeo Skoda Octavia Renault Laguna Opel Zafira
Engine size in litre	1.2, 1.6, 2.0	1.8, 2.0, 2.2
Fuel type	Petrol / Diesel	Petrol / Diesel
Fuel consumption in litre/100km	4.5, 5.6, 6.7, 7.8	5.6, 7.4, 9.1, 10.8
Energy label	A, B, C, E, G	A, B, C, E, G
Price in CHF	15,450 17,681 19,912 22,143 24,374	26,540 30,373 34,206 38,039 41,872

Table A1.1.2: A questionnaire sheet for car choice, possible values of the attributes[2]

[2] See Table 1 in Wüstenhagen and Sammer (2007: 67).

A 2 Appendix of Chapter 2

A 2.1 The Maximum Likelihood function and the marginal effects of the OLS, Probit and Tobit models with panel structure

In the following, I first derive the Maximum Likelihood (ML) function of the OLS, the Probit[3] and the Tobit model used to estimate the coefficients. I then derive the ML function of the OLS and the Tobit model that can capture the panel structure of data. Finally, I show how the marginal effects of explanatory variables on the probability that the explained variable is zero and on the expected value of the explained variable can be computed. I also present the formulas used to compute the elasticities.

The general structure of the OLS, the Probit and the Tobit model

The deterministic component for the OLS, the Probit and the Tobit model is of the following form:

$$V_{jn} = \alpha + \beta \cdot x_{\bullet jn}, \tag{A2.1.1}$$

$$j = 1,...,J \text{ and } n = 1,...,N.$$

where index n represents the household, N is the total number of households, j indicates the fuel price level presented to the household and J the total fuel levels presented to them. Parameter α is constant; parameter vector β captures weights relating to x_i, which stands for variables such as fuel prices, the marginal costs of driving and household properties. With respect to the error term, I also assume the same structure in all three cases, namely:

$$y^*_{jn} = V_{jn} + \eta_n + \varepsilon_{jn}, \tag{A2.1.2}$$

where

$$\eta_n \sim iid\, N(0, \sigma_\eta), \tag{A2.1.2a}$$

$$\varepsilon_{jn} \sim iid\, N(0, \sigma_\varepsilon) \text{ and} \tag{A2.1.2b}$$

$$j = 1,...,J \text{ and } n = 1,...,N. \tag{A2.1.2c}$$

Note that in the case of the OLS and the Tobit model, variable $y_{jn} = V_{jn}$ denotes the observed driving distance where, in the case of the OLS model,

[3] Note that in contrast to the Tobit model, the Probit model examines whether or not a household owns a car, but not how far the household drives this car. The Probit model will only be used in Appendix A3.13 when I compare the results yielded by this model with those computed using the MDCEV model.

$$y_{jn} = y^*_{jn} \tag{A2.1.3}$$

and, in the case of the Tobit model,

$$y_{jn} = \left\{ \begin{array}{l} y^*_{jn} < 0 : y_{jn} = 0 \\ y^*_{jn} \geq 0 : y_{jn} = y^*_{jn} \end{array} \right\} . \tag{A2.1.4}$$

In the case of the Probit model, value $y_{jn} = 0$ denotes that the household is carless; the value $y_{jn} = 1$ denotes that the household owns at least one car, where

$$y_{jn} = \left\{ \begin{array}{l} y^*_{jn} \leq 0 : y_{jn} = 0 \\ y^*_{jn} > 0 : y_{jn} = 1 \end{array} \right\} . \tag{A2.1.5}$$

In the event that there is no panel structure or that the panel structure is neglected, all error terms η_n are considered to be zero. Such models are called "pooled" models.

The ML functions of the OLS, the Probit and the Tobit model

The Tobit models are solved by Maximum Likelihood Estimation (MLE). The ML function is as follows:

$$L_{MLE}\left(Y = y \mid \beta, x, \eta\right) = \prod_{n=1}^{N} \prod_{j=1}^{J} \Phi\left(y_{jn} - \beta \cdot x_{\bullet jn} - \eta_n\right)^{\left(y_{jn}=0\right)} \cdot \phi\left(y_{jn} - \beta \cdot x_{\bullet jn} - \eta_n\right)^{\left(y_{jn}>0\right)} . \tag{A2.1.6}$$

In the OLS model, the ML function reduces to

$$L_{MLE}\left(Y = y \mid \beta, x, \eta\right) = \prod_{n=1}^{N} \prod_{j=1}^{J} \phi\left(y_{jn} - \beta \cdot x_{\bullet jn} - \eta_n\right) , \tag{A2.1.7}$$

since the case $y_{jn} = 0$ does not have a discrete and therefore positive probability in the case of the OLS model.

Note that the ML function is conditional on the individual specific random terms η_n, $n = 1,...,N$. Variable y denotes the total number of annual kilometres households would drive given different fuel prices. The ML function is given by

$$L_{MLE}\left(Y = y \mid \beta, x\right) = E_\eta\left(L_{MLE}\left(Y = y \mid \beta, x, \eta\right)\right) . \tag{A2.1.8}$$

There is no function of a closed form such that the ML function can be computed. The random term η_n therefore has to be integrated out. This could be realised, for example, by the Simulated Maximum Likelihood Estimation method (SMLE):

$$L_{SMLE}\left(Y = y \mid \beta, x\right) \approx \frac{1}{S}\sum_{s=1}^{S}\prod_{n=1}^{N}\prod_{j=1}^{J}\Phi\left(y_{jn} - \beta \cdot x_{\bullet jn} - \eta_{n(s)}\right)^{\left(y_{jn}=0\right)} \cdot \phi\left(y_{in} - \beta \cdot x_{\bullet jn} - \eta_{n(s)}\right)^{\left(y_{jn}>0\right)}, \quad \text{(A2.1.9)}$$

where $\eta_{n(s)}$ are S samples that are independently drawn from distribution $\eta_n \sim iid\, N\left(0, \sigma_\eta\right)$.

Analogously, the results of the Probit model were computed:

$$L_{SMLE}\left(Y_{jn} = y_{jn} \mid \beta, x\right) \approx \frac{1}{S}\sum_{s=1}^{S}\prod_{n=1}^{N}\prod_{j=1}^{J}\Phi\left(y_{jn} - \beta \cdot x_{\bullet jn} - \eta_{n(s)}\right)^{\left(y_{jn}=0\right)} \cdot \phi\left(-y_{jn} + \beta \cdot x_{\bullet jn} + \eta_{n(s)}\right)^{\left(y_{jn}=1\right)},$$

(A2.1.10)

where $y_{jn} = 0$ denotes that household n chooses not to own a car given the fuel prices j and $y_{jn} = 0$ denotes that household n chooses to own at least one car. Note that the software used – STATA – is not based on the SMLE method. STATA uses the Gauss-Hermit method for numerically integrating out the random term.[4]

Formulas to compute the elasticity of driving demand

I compute the elasticities of driving demand in the OLS and the Tobit model, both of which capture the panel structure, as follows:

$$\sum_{s=1}^{S}\sum_{n=1}^{N}\sum_{j=1}^{J}\frac{\partial E\left(Y_{jn} \mid \beta, \sigma, \eta_{n(s)}, x_{\bullet jn}\right)}{\partial x_{jn}} \cdot x_{ijn} \cdot \left(\sum_{s=1}^{S}\sum_{n=1}^{N}\sum_{j=1}^{J} E\left(Y_{jn} \mid \beta, \eta_{n(s)}, x_{\bullet jn}\right)\right)^{-1}, \quad \text{(A2.1.11)}$$

where in the case of the Tobit model[5]

$$\frac{\partial E\left(Y_{jn} \mid \beta, \sigma, \eta_{n(s)}, x_{\bullet jn}\right)}{\partial x_{ijn}} = \beta_i \cdot \Phi\left(\frac{-x_{\bullet jn}\beta - \eta_{n(s)}}{\sigma}\right) \quad \text{(A2.1.11a)}$$

and in the case of the OLS model[6]

---

[4] I chose to use 196 integration points to integrate by quadrature. Note that the resulting parameter values no longer changed when the number of integration points exceeded 190.

[5] See Mathematical Appendix MA3, Formula (MA 3.2).

$$\frac{\partial E\left(Y_{jn} \mid \beta, \sigma, \eta_{n(s)}, x_{\bullet jn}\right)}{\partial x_{ijn}} = \beta_i . \tag{A2.1.11b}$$

$\eta_{n(s)}$ are iid draws from distribution $\eta_n \sim N\left(0, \sigma_\eta\right)$. I approximate the expression $\sum_{s=1}^{S} \sum_{n=1}^{N} \sum_{j=1}^{J} E\left(Y_{jn} \mid \beta, \eta_{n(s)}, x_{\bullet jn}\right)$ by the sum of the observed driving distances $S \cdot \sum_{n=1}^{N} \sum_{j=1}^{J} y_{jn}$.

It is important to note that I chose this type of definition for the elasticity because it denotes the relative effect on driving demand if the explanatory variable of each household increases by one percent. For instance, the elasticity with respect to household income denotes the relative change in driving demand if the income of each household increases by one percent.[7]

Note that the marginal effect of the OLS model that includes the panel structure can be computed in the same way as in the model that neglects the panel structure, see Formula (2.1.5).

If the Tobit and the Probit model both capture the panel structure, I compute the elasticities of the probability of a household being carless as follows:

$$\sum_{s=1}^{S} \sum_{n=1}^{N} \sum_{j=1}^{J} \frac{\partial P\left(Y_{jn} = 0 \mid \beta, \sigma, \eta_{n(s)}, x_{\bullet jn}\right)}{\partial x_{ijn}} \cdot x_{ijn} \cdot \left(\sum_{s=1}^{S} \sum_{n=1}^{N} \sum_{j=1}^{J} P\left(Y_{jn} = 0 \mid \beta, \eta_{n(s)}, x_{\bullet jn}\right) \right)^{-1}, \tag{A2.1.12}$$

where for both the Probit and the Tobit model[8]

$$\frac{\partial P\left(Y_{jn} = 0 \mid \beta, \sigma, \eta_{n(s)}, x_{\bullet jn}\right)}{\partial x_{ijn}} = \frac{\beta_i}{\sigma} \cdot \phi\left(\frac{-x_{\bullet jn}\beta - \eta_{n(s)}}{\sigma} \right). \tag{A2.1.12a}$$

In this case, I replace the sum of simulated probabilities $\sum_{s=1}^{S} \sum_{n=1}^{N} \sum_{j=1}^{J} P\left(Y_{jn} = 0 \mid \beta, \eta_{n(s)}, x_{\bullet jn}\right)$ that a household does not own a car by the actual proportion of carless households multiplied by the number of draws S: $S \cdot \sum_{n=1}^{N} \sum_{j=1}^{J} y_{jn}$.

Note that if all error terms η_n are set to zero and S is set to one, all formulas above correspond to the elasticities of the model that neglects the panel structure.

---

[6] Note that $E\left(Y_{jn} \mid \beta, \sigma, \eta_{n(s)}, x_{\bullet jn}\right) = \alpha + \beta \cdot x_{jn} + \eta_{n(s)}$.

[7] Note that the result would change if the same total amount of additional income were distributed differently amongst households.

[8] See Mathematical Appendix MA3, Formula (MA 3.4).

A 2.2 The mean marginal effect on driving distance versus the marginal effect at the mean

In this appendix, I discuss why the mean of the marginal effects of explanatory variables on the driving demand of individual households is unequal to the mean of the marginal effects.

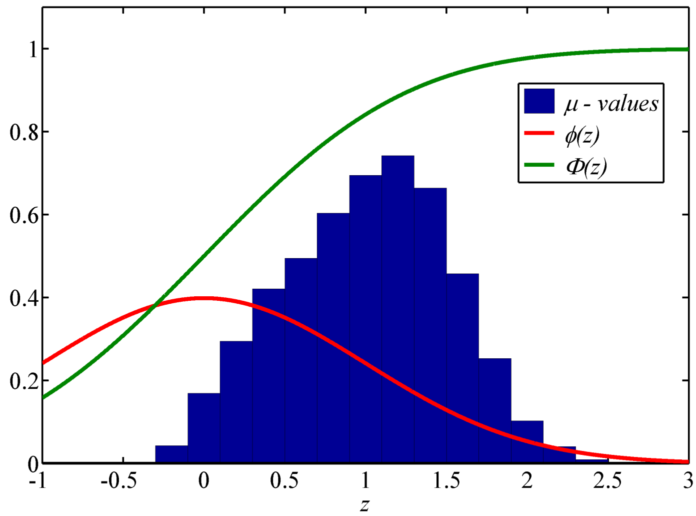


Figure A2.2.1: Distribution of μ values and probability functions

Figure A2.2.1 shows the histogram of household values $(\alpha + \beta \cdot x_n) \cdot \sigma^{-1}$. Note that these values go into non-linear functions to compute the marginal effects of explanatory variables on driving demand. In the case of the marginal effects of the expectation value on driving demand, it is the cumulated density function $\Phi(\bullet)$ of a standard normal distribution; in the case of the marginal effect on the probability of being carless, it is the probability density function $\phi(\bullet)$ of a standard normal distribution, see Formulas (A2.1.11a) and (A2.1.12a). I have also included the two functions $\phi(\bullet)$ and $\Phi(\bullet)$ in the diagram. It is now important to realise that if these functions were linear in the relevant range of $(\alpha + \beta \cdot x_n) \cdot \sigma^{-1}$, $\text{mean}\left(\phi\left((\alpha + \beta \cdot x_n) \cdot \sigma^{-1}\right)\right) = \phi\left(\text{mean}\left((\alpha + \beta \cdot x_n) \cdot \sigma^{-1}\right)\right)$ and therefore also elasticities $\varepsilon_{E(Y),income} | \bar{x}$ and $\text{mean}\left(\varepsilon_{E(Y),income} | x_n\right)$ would be equal:

$\varepsilon_{E(Y),income} | \beta, \sigma, \bar{x} = \text{mean}\left(\varepsilon_{E(Y),income} | \beta, \sigma, x_n\right)$.[9] The graphs in Figure A2.2.1 show that function $\Phi(\bullet)$ is more or less linear in the relevant range of $(\alpha + \beta \cdot x_n) \cdot \sigma^{-1}$ than function $\phi(\bullet)$. This is the reason why the difference between $\varepsilon_{P(Y=0),income} | \bar{x}$ and $\text{mean}\left(\varepsilon_{P(Y=0),income} | x_n\right)$ is smaller than that of $\varepsilon_{E(Y),income} | \bar{x}$ and $\text{mean}\left(\varepsilon_{E(Y),income} | x_n\right)$, see results in Table 2.2.1.

[9] Note that this follows from rule $E(g(X)) = g(E(X))$ if and only if $g(\bullet)$ is linear. Note that if x_n are random draws of the distribution $x_n \sim f_x(z)$, then $\lim_{n \to \infty} n^{-1} \cdot x_n = E(X)$, such that the following rule results $\lim_{n \to \infty} n^{-1} \cdot g(x_n) = E(g(X)) = g\left(\lim_{n \to \infty} n^{-1} \cdot x_n\right)$ if and only if $g(\bullet)$ is linear.

A 3 Appendix of Chapter 3

A 3.1 Effect of parameter changes on the conditional pdf

In this subsection, I illustrate the effect of changes of parameters d, a_2, m and β on the shape of the conditional probability density function (pdf) of driving demand $f_{X_2|(X_2>0)}$, the expected distance a household will travel, given it owns a car, $E(X_2 | X_2 > 0)$, and on the probability of a household being carless $P(X_2 = 0)$.

I shall start by illustrating the effect of a change in parameter *d*.

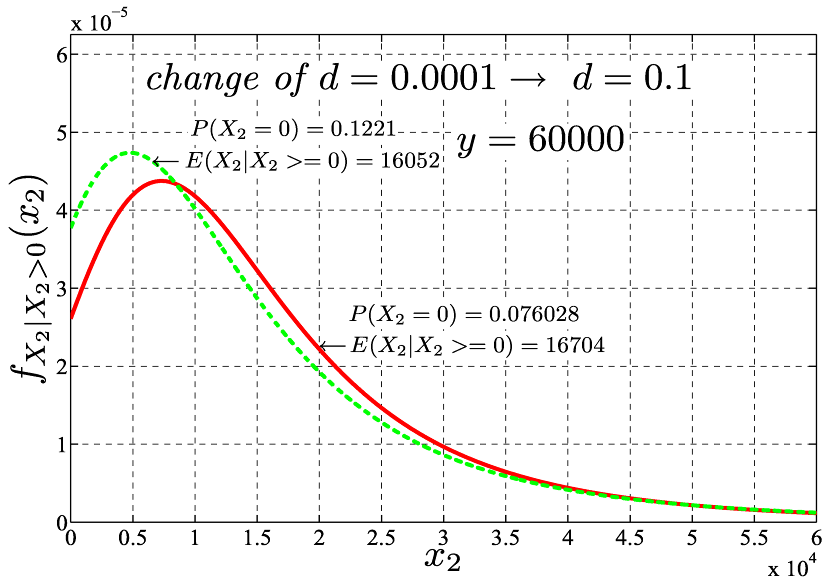


Figure A3.1.1: The effect of a change in parameter *d* on the conditional pdf of driving demand.

This diagram shows that an increase in parameter *d* leads to a shift of the conditional pdf towards the origin. At the same time, the tail at the right becomes heavier. This implies that both the probability of a household being carless $P(X_2 = 0)$ and the expectation value $E(X_2 | X_2 > 0)$ increase. It is important to note that decreasing parameter *d* to below the value $d = 0.0001$ will leave the conditional pdf almost unchanged.[8] This complies with the observation that further decreasing parameter *d* would not change any model property.

---

[8]If *d* is decreased from $d = 0.0001$ to $d = 0.00000001$, there will be no change in the "unconditional" pdf observable in the diagram, and the probability of the household being carless increases only from $P(X_2 = 0) = 0.076028$ to $P(X_2 = 0) = 0.075991$. The expected driving distance decreases only from $E(X_2) = 16705$ to $E(X_2) = 16704$.

I shall next illustrate the impact of a change in parameter a_2.

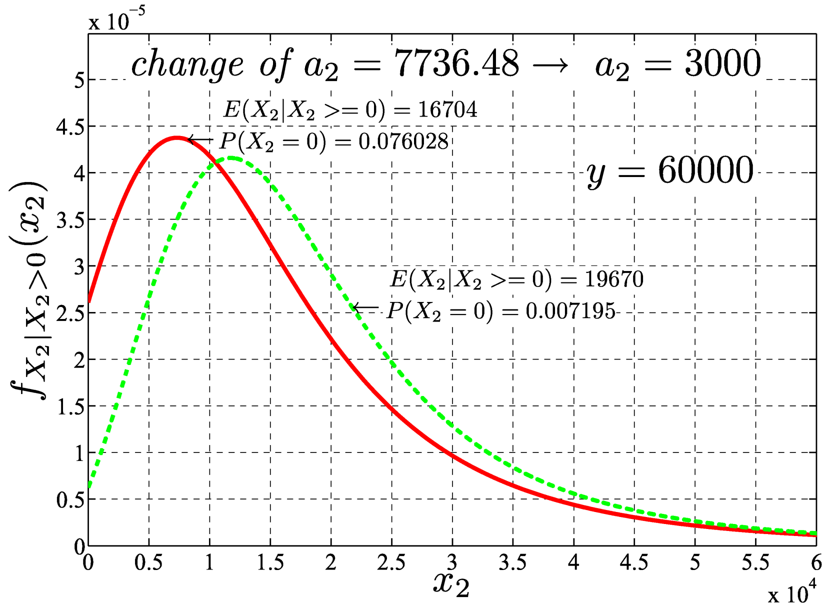


Figure A3.1.2: The effect of a change in parameter a_2 on the conditional pdf of driving demand.

This diagram shows that an increase in parameter a_2 leads to a shift of the conditional pdf towards the origin. This leads to a decrease in the expectation value of driving demand $E(X_2 | X_2 > 0)$ and to an increase in the probability of the household being carless $P(X_2 = 0)$.

The effect of a change in parameter *m* is illustrated in Figure A3.5.2. This diagram shows that when a household moves from a rural to an urban area – corresponding to a decrease in *m* – the conditional pdf shifts towards the origin. Expectation value $E(X_2 | X_2 > 0)$ therefore decreases and the expectation value of the probability $P(X_2 = 0)$ increases. Note that, compared to the effect of an increase in parameter a_2, the upper tail of the conditional pdf becomes less heavy. An increase in parameter a_2 and a decrease in parameter *m* therefore lead to differences in the changes of probability $P(X_2 = 0)$ and in expectation value $E(X_2 | X_2 > 0)$, even though both changes affect these two measures in the same direction.

Finally, I examine the effect of a change in increase in the scaling parameter of the random term β .

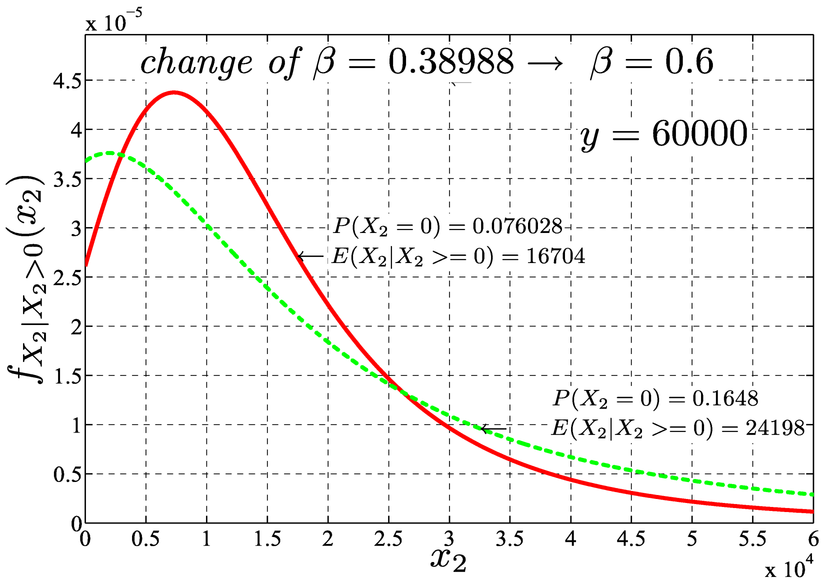


Figure A3.1.3: The effect of a change in parameter β on the conditional pdf of driving demand.

The diagram shows that the upper tail of the conditional pdf is heavier when the scaling parameter of the random term β increases. This is what I expected: if the variance of the relative preference for car driving increases, the variance of the actual driving distance must also increase. Since the upper tail is heavier, the expectation value $E(X_2 | X_2 > 0)$ increases.[11] In addition, probability $P(X_2 = 0)$ increases.[12] In contrast to changes in parameters a_2 and *m*, therefore, a change in parameter β does not affect values $E(X_2 | X_2 > 0)$ and $P(X_2 = 0)$ in the same direction.

---

[11] Note that, in this case, the unconditional expectation value $E(X_2) = (1 - P(X_2)) \cdot E(X_2 | X_2 > 0)$ would also increase: $(1 - 0.076028) \cdot 16704 = 14092$ is smaller than $(1 - 0.1648) \cdot 24198 = 20021$.

[12] To gain insight into this fact, see Figures A3.5.2 and A3.5.3.

A 3.2 Definition of formulas describing marginal effects

In this subsection, I will state the formulas not yet given in the "Results" section in Subchapter 3.2; see Formulas 3.2.43 and 3.2.44. I shall start with the relative changes with respect to changes in income:

$$\varepsilon_{P(X_2=0),y,sim} \,|\theta_{1,k}, d = \sum_{n=1}^{N} \frac{\partial P_{sim,n} \,|\{\theta_{1,k}, d, p_1, p_{2n}, y_n, s_n\}}{\partial y_n} \cdot \frac{y_n}{\sum_{n=1}^{N} P_{sim,n} \,|\{\theta_{1,k}, d, p_1, p_{2n}, y_n, s_n\}}, \qquad \text{(A3.2.1)}$$

$$\varepsilon_{E(X_2),y,sim} \,|\theta_{1,k}, d = \sum_{n=1}^{N} \frac{\partial E(X_2)_{sim,n} \,|\{\theta_{1,k}, d, p_1, p_{2n}, y_n, s_n\}}{\partial y_n} \cdot \frac{y_n}{\sum_{n=1}^{N} E(X_2)_{sim,n} \,|\{\theta_{1,k}, d, p_1, p_{2n}, y_n, s_n\}}. \qquad \text{(A3.2.2)}$$

The relative changes when moving from rural to urban areas are defined as

$$\varepsilon_{P(X_2=0),rural\rightarrow urban,sim} \,|\theta_{1,k}, d = 1 - \frac{\sum_{n=1}^{N} P_{sim,n,urban} \,|\{\theta_{1,k}, d, p_1, p_{2n}, y_n, s_n\}}{\sum_{n=1}^{N} P_{sim,n,rural} \,|\{\theta_{1,k}, d, p_1, p_{2n}, y_n, s_n\}}, \qquad \text{(A3.2.3)}$$

$$\varepsilon_{E(X_2),rural\rightarrow urban,sim} \,|\theta_{1,k}, d = 1 - \frac{\sum_{n=1}^{N} E(X_2)_{sim,urban,n} \,|\{\theta_{1,k}, d, p_1, p_{2n}, y_n, s_n\}}{\sum_{n=1}^{N} E(X_2)_{sim,rural,n} \,|\{\theta_{1,k}, d, p_1, p_{2n}, y_n, s_n\}}, \qquad \text{(A3.2.4)}$$

where for $P_{sim,n,urban} \,|\{\cdots\}$ and $E(X_2)_{sim,urban,n} \,|\{\cdots\}$ the location dummy of the socio-demographic vector s_{2n} is set to one: This indicates that the household is located in a rural area. For $P_{sim,n,rural} \,|\{\cdots\}$ and $E(X_2)_{sim,\,rural,n} \,|\{\cdots\}$, this dummy is set to zero. This means that in the case when the dummy s_{2n} is one, parameter $m_n = \gamma \cdot s_n$ is higher than if dummy s_{2n} is zero, since the corresponding parameter γ_2 is greater than zero; see (3.2.47a). This means that the relative preference of rural households for driving cars is higher than of households located in urban areas. This implies that the driving demand decreases, and the probability of households being carless increases if households move from urban to rural areas; see Table 3.2.3 .

The relative changes if households move from urban to rural areas would be defined analogously to (A3.2.3) and (A3.2.4).

A 3.3 Histograms of simulated relative changes

In the following, the histograms of income elasticities and the relative changes of the probability of choosing not to own a car with respect to a change in income are presented. The histograms result from computing the simulated elasticities and from the relative changes conditional on a random draw of the estimated parameters. The histogram of relative changes in driving demand and in the probability of choosing not to own a car when households change their location are then presented.

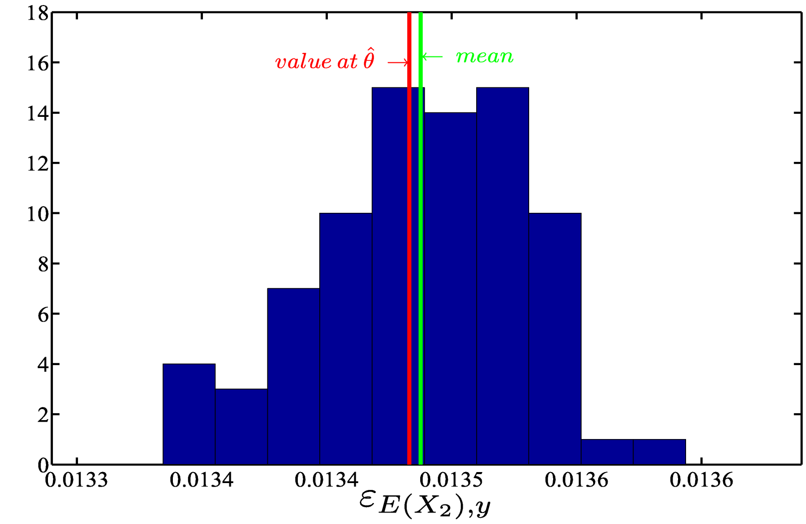


Figure A3.3.1: Distribution of income elasticity.

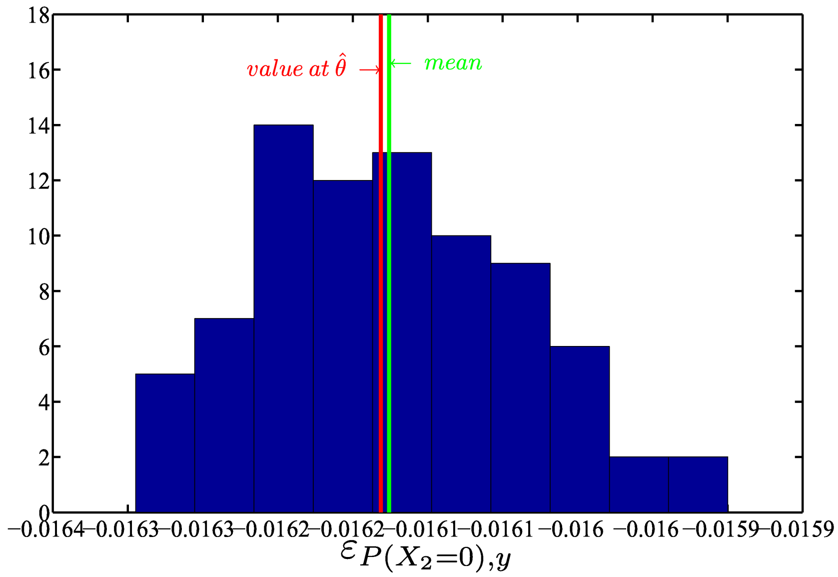


Figure A3.3.2: Distribution of the relative change in the probability of not owning a car with respect to income.

Again, as in the case of the influence of driving costs, the variance of the distributions is very small with respect to the level of effects. As a result, the means of the simulated values are also closed to the estimated value.

I now present the histogram of relative changes in driving demand and of the probability of choosing not to own a car when households change location.

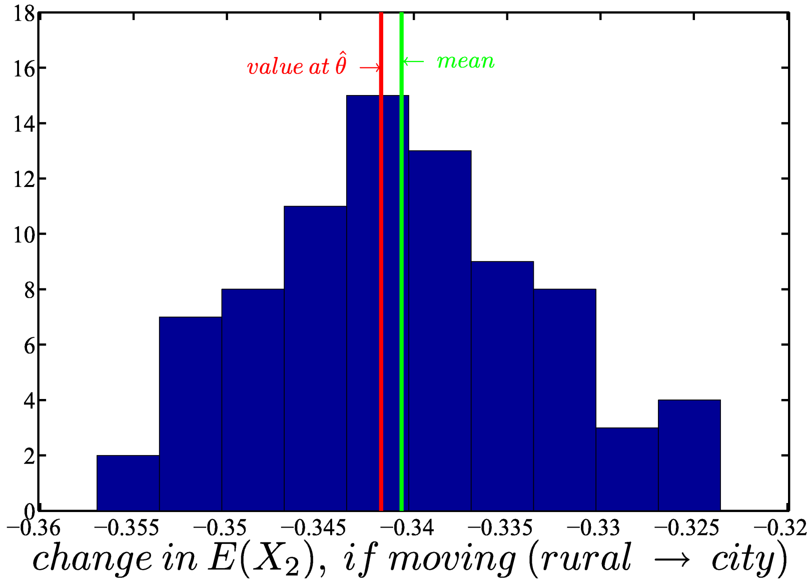


Figure A3.3.3: Distribution of the change in demand if moving from a rural to an urban area.

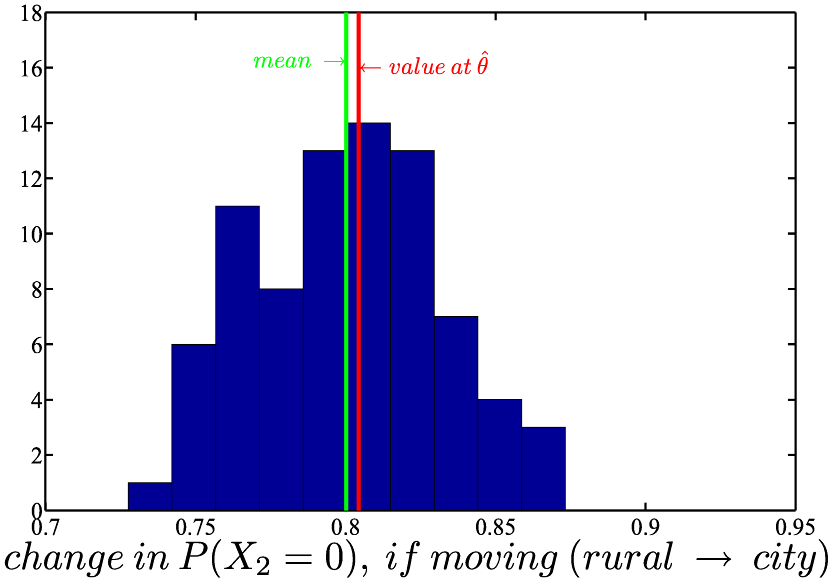


Figure A3.3.4: Distribution of the relative change of the probability of not owning a car if moving from an urban to a rural area.

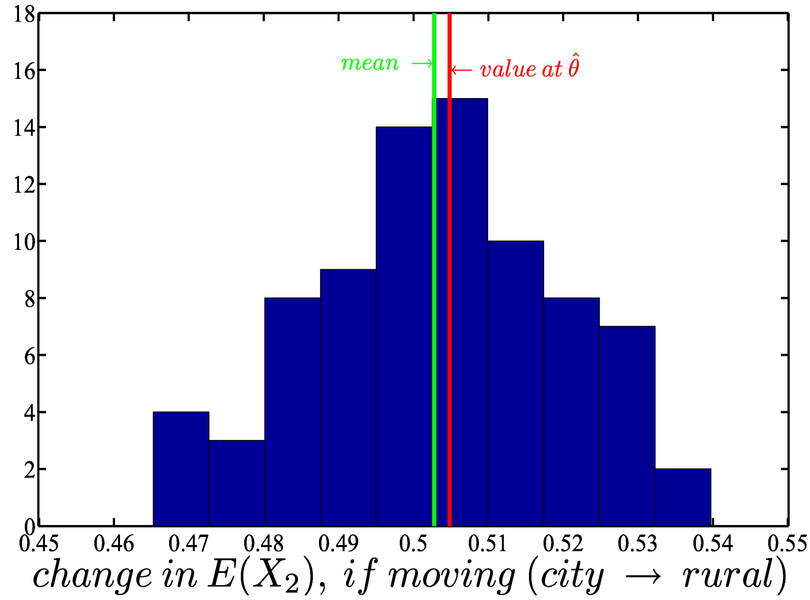


Figure A3.3.5: Distribution of the relative change in probability of car ownership if moving from an urban to a rural area.

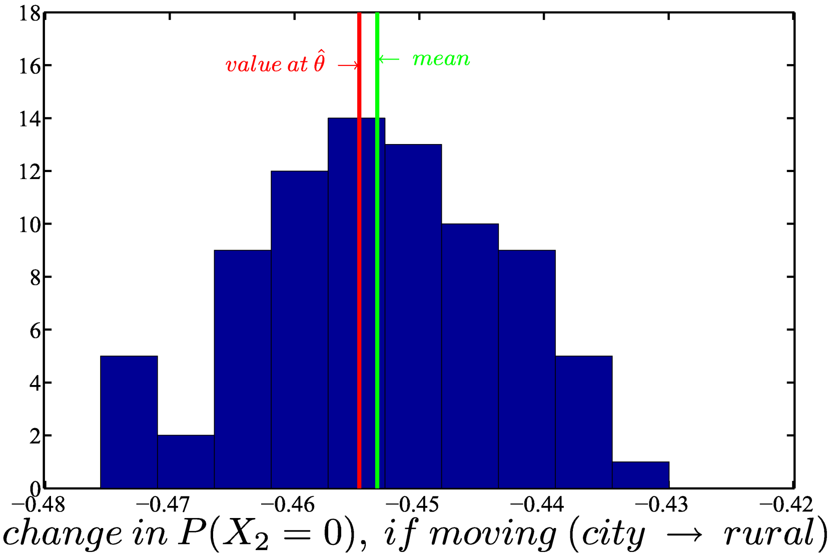


Figure A3.3.6: Distribution of the relative change in probability of car ownership if moving from a rural to an urban area.

Again, also in these cases, the distribution of the relative changes is relatively small with respect to their levels, and also the mean of the simulated changes is very much closed to the value at the estimated value $\hat{\theta}_1$. Note that only households located in urban areas were considered for Figures A3.3.3 and A3.3.4; correspondingly, only households located in rural areas were considered for Figures A3.3.5 and A3.3.6.

A 3.4 Distribution of driving demand of different household segments

In the following, I shall examine whether the model approximates the data well. To this end, I compare the empirical distribution of driving demand to the conditional pdf the model forecasts for different household segments. I shall do the same by comparing the forecast probability of households being carless to the actual proportion of households without a car. Households are segmented by the category income and by the type of area in which they reside, namely rural versus urban areas. I will compute these histograms and proportions for each combination of segmentation criteria.

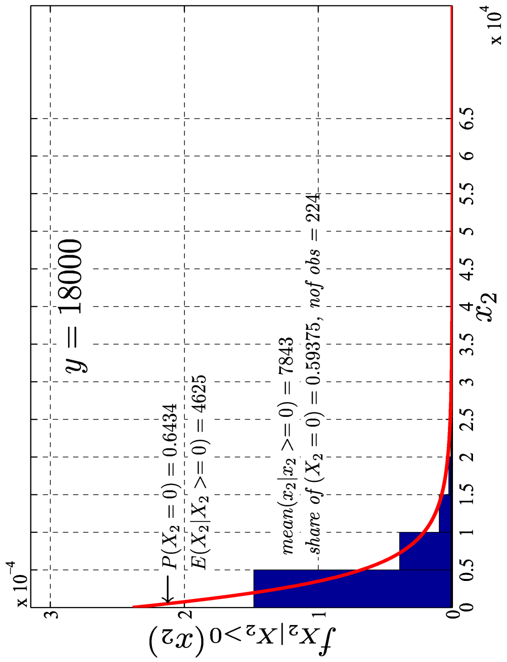


Figure A3.4.1: Histogram of households in rural areas with and income of CHF 18,000.

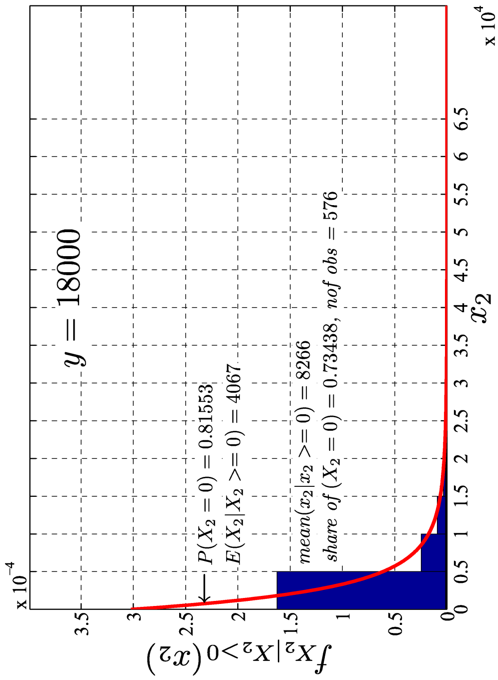


Figure A3.4.2: Histogram of households in urban areas with an income of CHF 18,000.

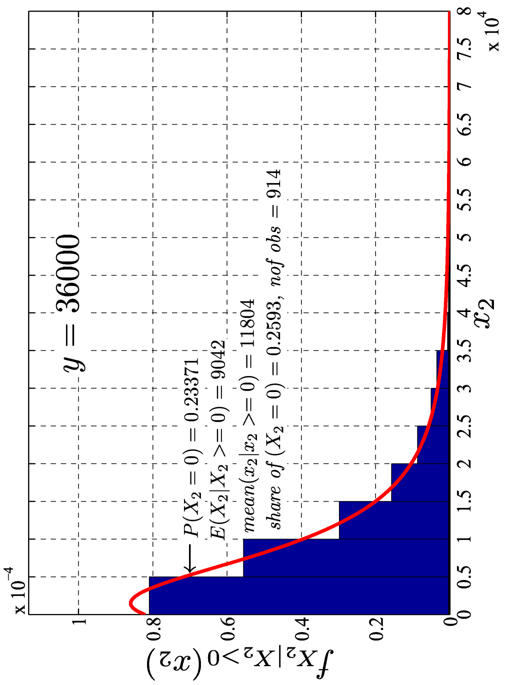


Figure A3.4.3: Histogram of households in rural areas with an income of CHF 36,000.

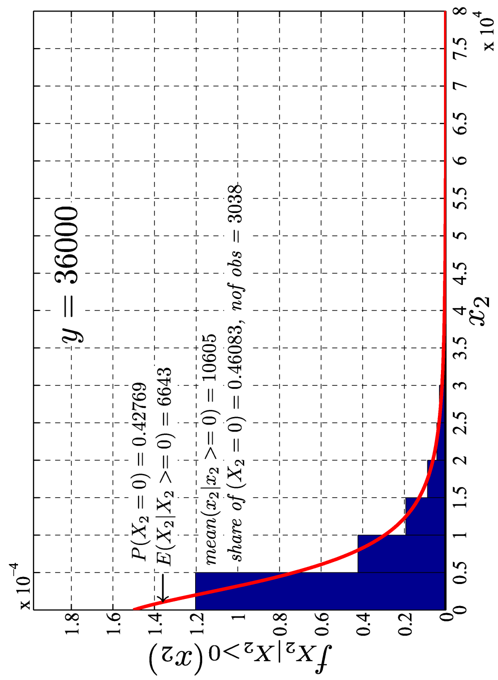


Figure A3.4.4: Histogram of households in urban areas with an income of CHF 36,000.

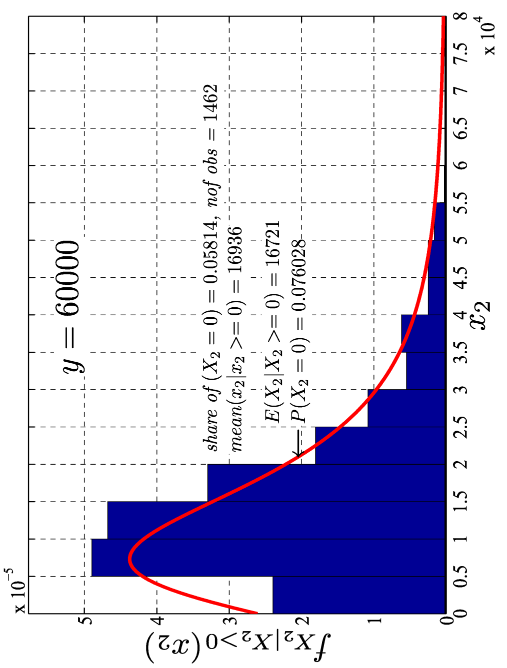


Figure A3.4.5: Histogram of households in rural areas with an income of CHF 60,000.

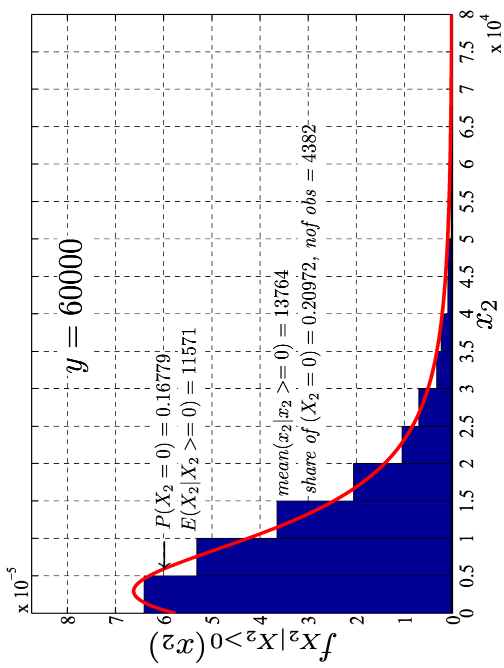


Figure A3.4.6: Histogram of households in urban areas with an income of CHF 60,000.

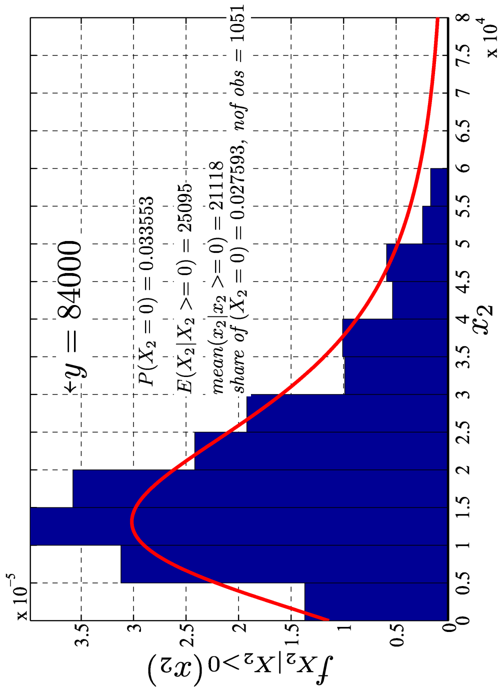


Figure A3.4.7: Histogram of households in rural areas with an income of CHF 84,000.

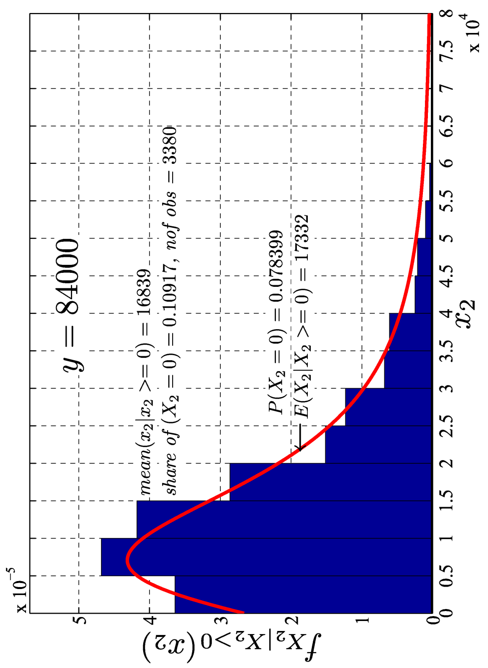


Figure A3.4.8: Histogram of households in urban areas with an income of CHF 84,000.

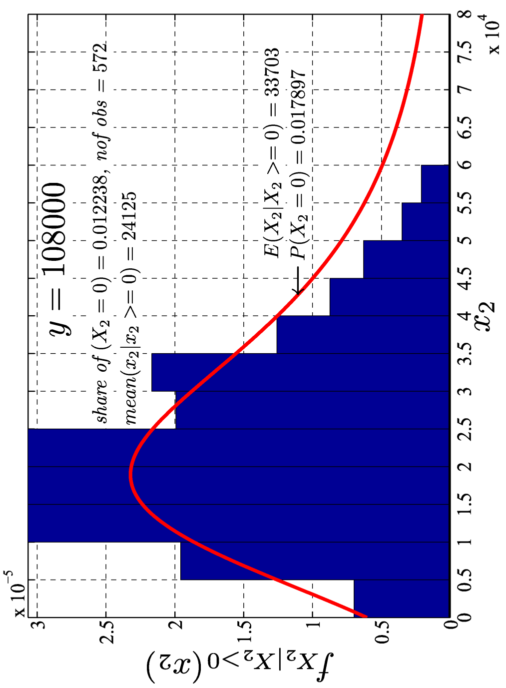


Figure A3.4.9: Histogram of households in rural areas with an income of CHF 108,000.

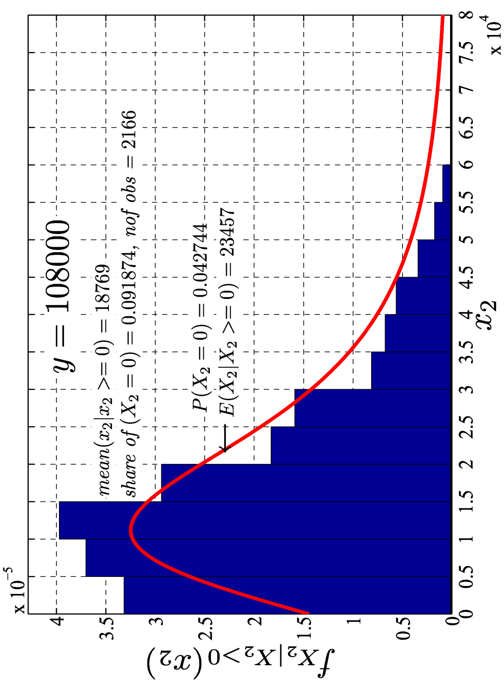


Figure A3.4.10: Histogram of households in urban areas with an income of CHF 108,000.

These diagrams show that the densities and probabilities predicted by the model fit the actual data of each household segment very well. The shape of the conditional pdfs are very similar to the shape of the histograms.[13] The only negative aspect that can be observed is that the conditional pdfs have too heavy tails, particularly for household income levels equal to and greater than CHF 84,000.[14] This leads to the fact that the forecast expected driving distances are greater than the actual driving distances. In the following table, the simulated expected values and simulated probabilities of households being carless are compared to the empirical values for each income and household location category.

y	location	Nobs.	share of $x_2 = 0$	$P(X_2 = 0)$	Δ	$mean(x_2)$	$E(X_2 = 0)$	Δ
18000	city	576	0.7344	0.8155	0.081	2083	750	-1333
18000	rural	224	0.5938	0.6434	0.050	3358	1649	-1708
36000	city	3038	0.4608	0.4277	-0.033	5718	3802	-1916
36000	rural	914	0.2593	0.2337	-0.026	8743	6929	-1814
60000	city	4382	0.2097	0.1678	-0.042	10878	9629	-1248
60000	rural	1462	0.0581	0.0760	0.018	15952	15540	-502
84000	city	3380	0.1092	0.0784	-0.031	15000	15973	973
84000	rural	1051	0.0276	0.0336	0.006	20535	24252	3717
108000	city	2166	0.0919	0.0427	-0.049	17044	22455	5411
108000	rural	572	0.0122	0.0179	0.006	23831	33100	9269
132000	city	1177	0.0671	0.0260	0.0411	19095	28992	9897
132000	rural	276	0.0145	0.0108	0.0037	26424	42052	15628
156000	city	617	0.0388	0.0171	0.0217	21674	35517	13843
156000	rural	117	0.0085	0.0070	0.0015	29953	50864	20911

Table A3.4.1: Empirical and simulated values of driving distance and proportions of carless households

Since it is hard to detect any systematic relationship between the differences in simulated and empirical values from this table, I shall plot these differences with respect to income. The results, illustrated in Figure A3.4.11, show that the difference between simulated probabilities $P(X_2 = 0)$ and actual proportions of carless households decreases slightly with income. I conclude from this that simulated changes of $P(X_2 = 0)$ with respect to income are slightly smaller in magnitude than is

[13] Note that the surface of the histogram is normalized to one.

[14] The problem of too heavy tails and too high values $E(X_2)$ can be observed to an ever stronger extent when plotting the same diagrams for income categories greater than CHF 10,8000.

actually the case. In contrast, the results illustrated in Figure A3.4.12 show that the difference between simulated expected driving demand $E(X_2)$ and actual values increases with income. I therefore assume that simulated income elasticities of driving demand are up to 0.58 too high.[15] Due to the budget effect of price changes, I also expect the simulated elasticities of driving demand to be rather too high in magnitude.[16]

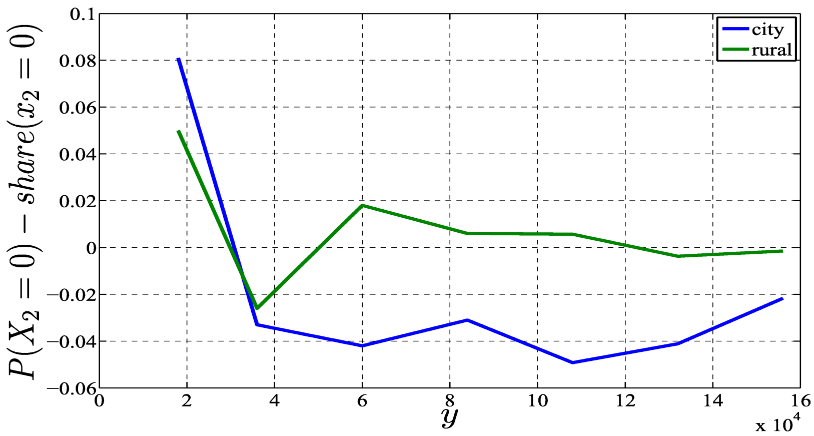


Figure A3.4.11: Deviation of simulated proportion of carless households from empirical value.

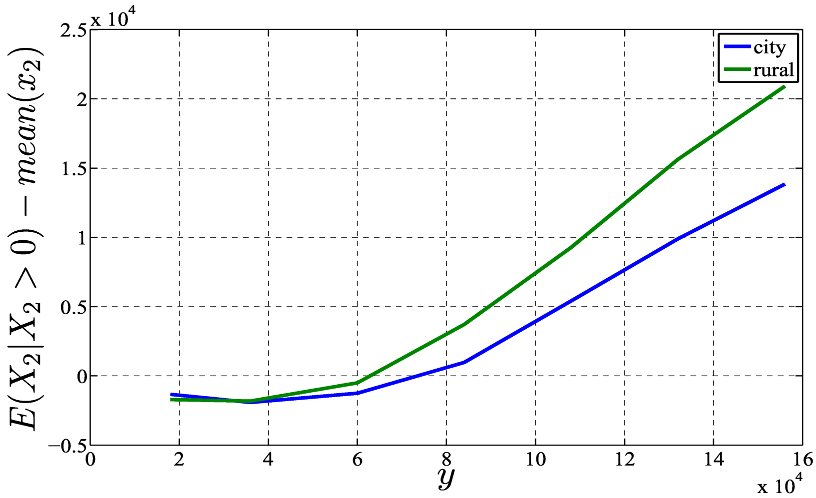


Figure A3.4.12: Deviation of simulated driving distance from empirical value.

---

[15] The diagram shows that for an income difference of CHF 90,000 in the interval of CHF 20,000 to CHF 110,000, the difference between forecast and actual values increases by about 9,000 km. This equals 0.1 km/CHF. This implies that the simulated income elasticity of driving demand might be 0.58 higher than it actually is: $0.1\text{km/CHF} \cdot 80187\text{sFr}/13890\text{km} = 0.577...$, where CHF 80,187 is the mean income and 13,890 km is the mean annual driving distance of households; see Table 3.2.1. Note that I chose the interval of CHF 20,000 to CHF 110,000 because more than 80% of households have an income in this range, as can be seen in Table 4.2.2.

[16] Note that since the driving demand is a normal good, the budget effect is positive. The effect shown in the case of income therefore leads in the same direction. This means that the simulated price elasticity of driving demand is also expected to be too high.

To conclude this subsection, I would like to present an overview of the frequencies of the household segments in the dataset presented above.

y	Nobs. urban	Nobs. rural	Share of urban	Nobs.	% of total
18000	576	224	72%	800	3.8%
36000	3038	914	77%	3952	18.9%
60000	4382	1462	75%	5844	28.0%
84000	3380	1051	76%	4431	21.2%
108000	2166	572	79%	2738	13.1%
132000	1177	276	81%	1453	7.0%
156000	617	117	84%	734	3.5%
180000	295	66	82%	361	1.7%
228000	475	82	85%	557	2.7%

Table A3.4.2: Empirical and simulated values of driving distance and proportions of carless households

This table shows that 81.2% of the households have an income between CHF 36,000 and CHF 108,000, and that the share of households in urban areas increases with income.

A 3.5 Illustration of simulated changes in driving demand

In the following, I shall illustrate the effect of an increase in income and a change in household properties on the conditional pdf of driving demand and on the probability of a household being carless. Note that the effect of an increase in the marginal costs of driving was already illustrated in Subchapter 3.2 in the section "Visualization of simulated changes"; see Figure 3.2.10.

I shall start by illustrating the effect of a change in income.

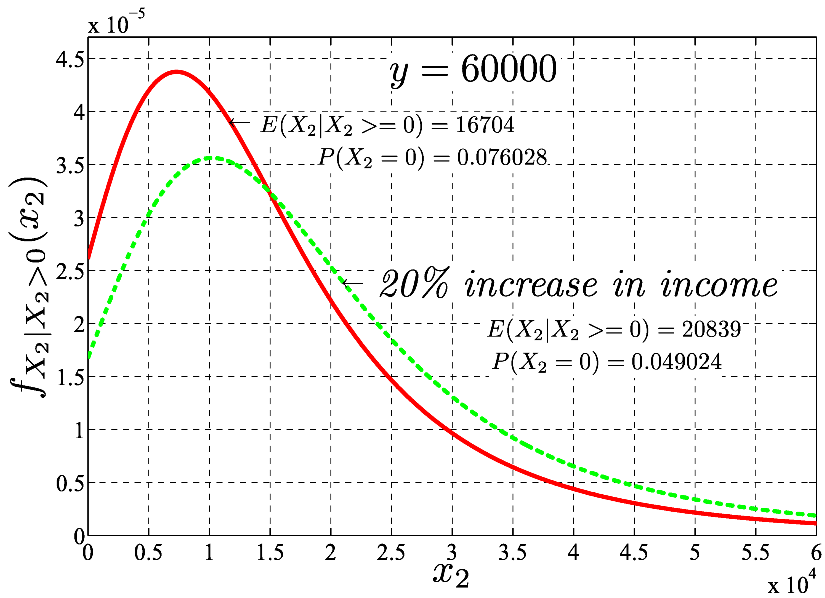


Figure A3.5.1: Change in distribution of driving demand if income increases.

This diagram shows that if income increases, the conditional pdf shifts to the right. Moreover, the upper tail becomes heavier. For this reason, expected driving demand increases. At the same time, probability $P(X_2 = 0)$ decreases. This result is very intuitive, since households with a higher income are more likely to be able to afford a car and to drive more.

I shall next examine the effect of a household moving from a rural to an urban area. Recall that when households moved from rural to urban areas, parameter $m_n = \gamma \cdot s_n$, which reflects the relative preference for driving, decreases. This leads to the effect shown by the following Figure A3.5.2: If a household moves from a rural to an urban area, the conditional pdf shifts towards the origin. Moreover, the upper tail becomes less heavy. The expectation value of driving demand $E(X_2 | X_2 > 0)$ therefore also increases. Also, probability $P(X_2 = 0)$ decreases. This result is intuitive, since households in rural areas have a greater need to own a car and require longer driving distances, since the facilities they want or have to visit are further away on average.

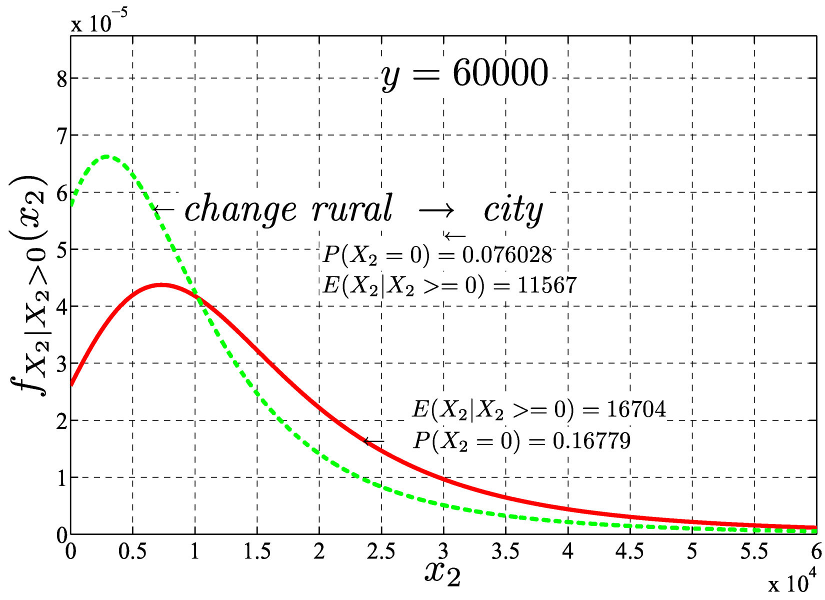


Figure A3.5.2: Change in distribution of driving demand if household moves from a rural to an urban area.

The following diagrams illustrate the same effect, namely the effect when a household moves from a rural to an urban area. This time, however, I computed the "unconditional" pdf. The "unconditional" pdf is not zero for any negative value x_2, which would economically be correct. This way, I can illustrate the effect of the change in household location on probability $P(X_2 = 0)$. It can be shown that probability $P(X_2 = 0)$ is equal to the area below the curve of the "unconditional" pdf in interval $-a_2 \leq x_2 < 0$.[17] The following diagrams show that this area is dramatically smaller when households are located in a rural area compared to the case where they are located in an urban area.

[17] Note that the value of a_2 was estimated to be $a_2 = 7736$.

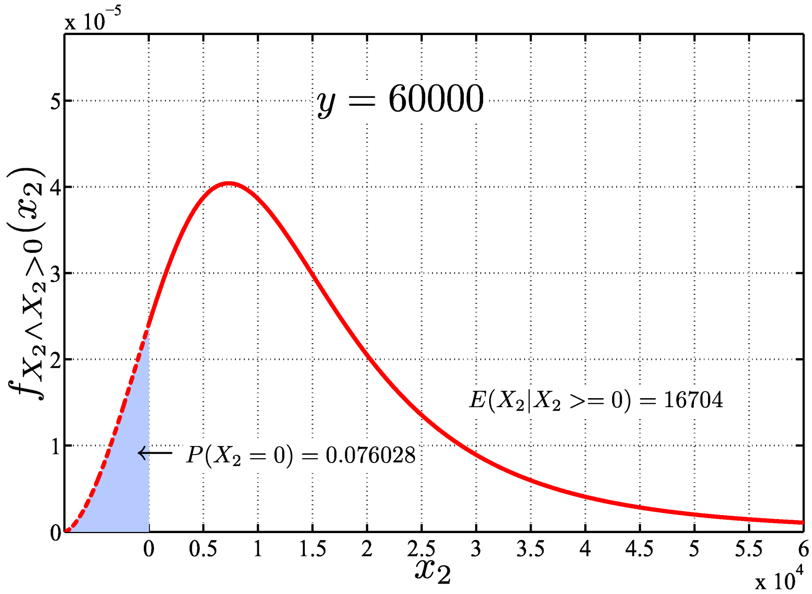


Figure A3.5.3: Illustration of the probability of observing a carless household.

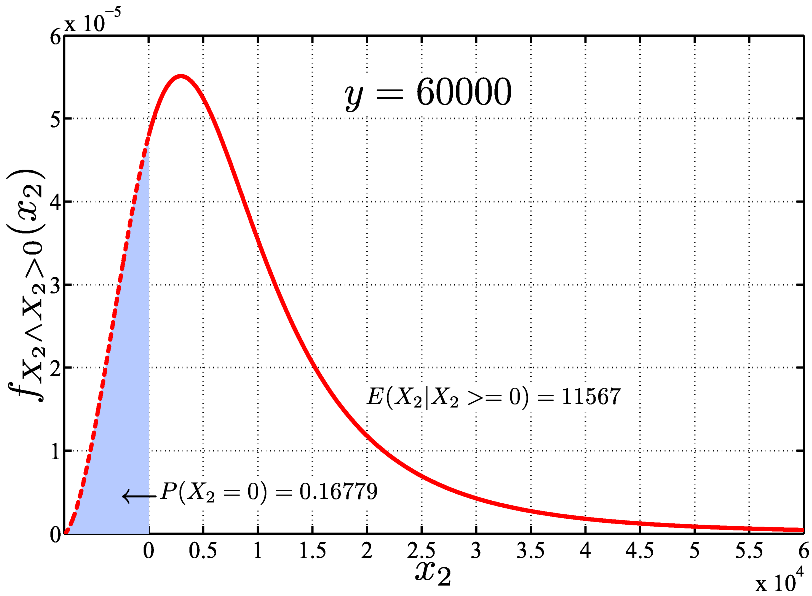


Figure A3.5.4: Illustration of the probability of observing a carless household in an urban area.

A 3.6 Estimation and simulation using simulated data

In the following, I wish to compute the t-values of parameters θ_1 by parametric bootstrapping.[18] I processed these t-values as follows:

1. I estimated parameters θ_1 by MLE using the data from the original dataset, $\hat{\theta}_1 \mid real\, data$.
2. With these parameters $\hat{\theta}_1 \mid real\, data$, I computed a dataset of the same size using simulated observations x_{2n}. This is performed by drawing a random variable ς_{2n} for each observation and then computing the Marshallian demand function (A 3.6.1).
3. Using these datasets containing simulated observations $x_{2n,sim}$, I computed parameters $\hat{\theta}_1 \mid sim.data$ by MLE.
4. Using the estimated parameters $\hat{\theta}_1 \mid sim.data$ and the simulated data $x_{2n,sim}$, I simulated the effects of changes in income, price and the change in household location.
5. I repeated steps 2, 3 and 4 *n*-times. I then computed the summary statistics of the estimated parameters $\hat{\theta}_1 \mid sim.data$ and the simulated effects, respectively; see Tables (A3.6.1) and (A3.6.2).

The Marshallian demand function used for simulating the data is defined as follows:[19]

$$X_2 = \frac{y - k_2 - p_1 \cdot (B \cdot a_2 - a_1)}{p_1 \cdot B + p_2}, \text{ with } B = \left(\frac{p_2}{p_1} \cdot \exp(-m - \varsigma)\right)^{\frac{1}{1-d}} = \left(\frac{p_2}{p_1}\right)^{\frac{1}{1-d}} \cdot \exp\left(-\frac{m + \varsigma}{1-d}\right). \quad \text{(A3.6.1)}$$

---

[18] The parametric bootstrapping method is described in Bichel and Freedman (1981).

[19] This Marshallian demand can be computed by setting Formula (3.2.11) equal to (3.2.12) and by solving for X_1. Plugging this result in the budget restriction (3.2.3) and solving for X_2 yields this Marshallian demand function.

Applying the procedure above, we arrive at the following results based on 150 simulated datasets:

	γ_1	γ_2	β	a_2
$mean(\hat{\theta}_1) \mid sim.data$	-2.717	0.350	0.390	7738.4
$stdev(\hat{\theta}_1) \mid sim.data$	0.01093	0.01141	0.00341	142.1
$t\text{-}value(\hat{\theta}_1) \mid sim.data$	-248.5	30.7	114.2	54.5
$\hat{\theta}_1 \mid real\,data$	-2.72	0.350	0.390	7736.5
$t\text{-}value(\hat{\theta}_1) \mid real\,data$	-294	31.5	126.3	65.25

Table A3.6.1: t-values of parameters using parametric bootstrapping

These results are very similar to the t-values computed based on the Hessian Matrix of the MLE function that was applied to the original dataset. In particular, the point estimates based on the original dataset are almost identical to the mean value of the estimates based on simulated data. I conclude from this that both the simulation and the ML estimation routines are correctly implemented.

$\varepsilon_{E(X_2),p_2} \mid d = 0.0001 = \underset{\substack{\langle -1.362 \rangle \\ (0.00755)}}{-1.3634}$	$\varepsilon_{P(X_2=0),p_2} \mid d = 0.0001 = \underset{\substack{\langle 1.6345 \rangle \\ (0.01146)}}{1.6395}$
$\varepsilon_{E(X_2),p_{fuel}} \mid d = 0.0001 = \underset{\substack{\langle -0.62243 \rangle \\ (0.00345)}}{-0.62307}$	$\varepsilon_{P(X_2=0),p_{fuel}} \mid d = 0.0001 = \underset{\substack{\langle 0.74697 \rangle \\ (0.005237)}}{0.7493}$
$\varepsilon_{E(X_2),y} \mid d = 0.0001 = \underset{\substack{\langle 1.3483 \rangle \\ (0.00748)}}{1.3497}$	$\varepsilon_{P(X_2=0),y} \mid d = 0.0001 = \underset{\substack{\langle -1.6181 \rangle \\ (0.01135)}}{-1.6231}$
$\Delta_{\ln(E(X_2)),city \rightarrow rural} \mid d = 0.0001 = \underset{\substack{\langle 0.50485 \rangle \\ (0.017862)}}{0.5006}$	$\Delta_{\ln(P(X_2=0)),city \rightarrow rural} \mid d = 0.0001 = \underset{\substack{\langle -0.45479 \rangle \\ (0.01055)}}{-0.4527}$
$\Delta_{\ln(E(X_2)),rural \rightarrow city} \mid d = 0.0001 = \underset{\substack{\langle -0.34175 \rangle \\ (0.0079257)}}{-0.33978}$	$\Delta_{\ln(P(X_2=0)),rural \rightarrow city} \mid d = 0.0001 = \underset{\substack{\langle 0.80436 \rangle \\ (0.003412)}}{0.7982}$

Table A3.6.2: Simulated parameters conditional on a fixed parameter *d* using simulated data[20]

These results show that the standard deviations are only slightly larger than those I computed using the delta method in Subchapter 3.2. This is yet another sign that it is correct to apply the delta method I used in Subchapter 3.2.

[20] These values were computed by generating 10 datasets and calculating the summary statistics of the resulting values of the elasticities. The values in the first line are the mean simulated values; the values in brackets “ $(\ldots)$ ” are the standard deviations based on the simulated datasets. The values in brackets “ $\langle\ldots\rangle$ ” are the simulated values at the point estimate $\hat{\theta}_1$ using the real data and the values.

A 3.7 Boundary solution in the case of the model with fixed costs

In this section, I shall show that there is a critical relative preference expressed by parameters $m+\varsigma$. If the relative preference is below this level, the demand for car driving will yield a boundary solution that means the driving distance is zero. If the relative preference is above this level, the driving demand will be positive. Since this is not evident when just looking at the utility function and the budget restriction, I prove (3.3.7), which maps the statement above in the following. To this end, I first restate (3.3.7):

A value $\varsigma = \varsigma_0$ exists such that

$$x_2(y-k_2, p_1, p_2, A, a_1, a_2)\,|_{\varsigma\le\varsigma_0} \le 0 \text{ and } x_2(y-k_2, p_1, p_2, A, a_1, a_2)\,|_{\varsigma>\varsigma_0} > 0\,, \qquad \text{(A3.7.1)}$$

where

$$x_2(y-k_2, p_1, p_2, A, a_1, a_2) = \frac{A\cdot\frac{y-k_2}{p_1} - a_2}{1 + A\cdot\frac{p_2}{p_1}} \text{ and } A = \left(\frac{p_1}{p_2}\cdot\exp(m+\beta\cdot\varsigma)\right)^{\frac{1}{1-d}}. \qquad \text{(A3.7.2)}$$

The proof that (A3.7.1) is correct follows from these conditions:

i. $\lim\limits_{\varsigma\to-\infty} x_2(y-k_2, p_1, p_2, A, a_1, a_2) \le 0\,,$

ii. $\lim\limits_{\varsigma\to\infty} x_2(y-k_2, p_1, p_2, A, a_1, a_2) > 0\,,$

iii. $\dfrac{\partial x_2(y-k_2, p_1, p_2, A, a_1, a_2)}{\partial A}\cdot\dfrac{\partial A}{\partial\varsigma} > 0\,.$

Only conditions i., ii. and iii. therefore need to be verified.
The proof of condition i. follows from

$$\lim_{\varsigma\to-\infty} A = \lim_{\varsigma\to-\infty}\left(\frac{p_1}{p_2}\cdot\exp(m+\beta\cdot\varsigma)\right)^{\frac{1}{1-d}} = 0\,.$$ This implies that $$\lim_{\varsigma\to-\infty} x_2(\bullet) = \lim_{\varsigma\to-\infty}\frac{A\cdot\frac{y-k_2}{p_1} - a_2}{1 + A\cdot\frac{p_2}{p_1}} = -a_2 < 0\,.$$

The proof of condition ii. follows from

$$\lim_{\varsigma\to\infty} A = \lim_{\varsigma\to\infty}\left(\frac{p_1}{p_2}\cdot\exp(m+\beta\cdot\varsigma)\right)^{\frac{1}{1-d}} = \infty\,.$$

This implies that $$\lim_{\varsigma\to-\infty} x_2(\bullet) = \lim_{\varsigma\to-\infty}\frac{A\cdot\frac{y-k_2}{p_1} - a_2}{1 + A\cdot\frac{p_2}{p_1}} = \frac{A\cdot\frac{y-k_2}{p_1}}{A\cdot\frac{p_2}{p_1}} = \frac{y-k_2}{p_2} < 0\,.$$

This result is rather intuitive. If $\varsigma \to \infty$, this means that the household has a very strong preference for car driving, and it is therefore plausible that it spends all income $y - k_2$ on car driving.

In order to prove condition iii., the derivative simply has to be computed. This yields:

$$\frac{\partial x_2(\bullet)}{\partial \varsigma} = \frac{\partial\left(\left(A \cdot \frac{y-k_2}{p_1} - a_2\right)\cdot\left(1 + A \cdot \frac{p_2}{p_1}\right)^{-1}\right)}{\partial \varsigma} = \frac{A' \cdot \frac{y-k_2}{p_1} \cdot \left(1 + A \cdot \frac{p_2}{p_1}\right) - \left(A \cdot \frac{y-k_2}{p_1} - a_2\right) \cdot A' \cdot \frac{p_2}{p_1}}{\left(1 + A \cdot \frac{p_2}{p_1}\right)^2} =$$

$$= A' \cdot \frac{\frac{y-k_2}{p_1} + A \cdot \frac{y-k_2}{p_1} \cdot \frac{p_2}{p_1} - A \cdot \frac{y-k_2}{p_1} \cdot \frac{p_2}{p_1} + a_2 \cdot \frac{p_2}{p_1}}{\left(1 + A \cdot \frac{p_2}{p_1}\right)^2} = A' \cdot \frac{\frac{y-k_2}{p_1} + a_2 \cdot \frac{p_2}{p_1}}{\left(1 + A \cdot \frac{p_2}{p_1}\right)^2} > 0 \text{, with}$$

$$A' = \frac{1}{1-d} \cdot \frac{p_1}{p_2} \cdot \left(\frac{p_1}{p_2} \cdot \exp(m + \beta \cdot \varsigma)\right)^{\frac{1}{1-d} - 1} = \frac{1}{1-d} \cdot \frac{p_1}{p_2} \cdot \left(\frac{p_1}{p_2} \cdot \exp(m + \beta \cdot \varsigma)\right)^{\frac{d}{1-d}} = \frac{\beta}{1-d} \cdot \frac{p_1}{p_2} \cdot A^d > 0 \cdot$$

A 3.8 Minimal driving distance

As shown in Subchapter 3.3, there is a minimal driving distance x_{2c} and a corresponding relative preference ς_c. I now prove that $\varsigma = \varsigma_c$ exists such that[19]

$$u_{S_2} - u_{S_1}\big|_{\varsigma \geq \varsigma_c} \geq 0 \text{ and } u_{S_2} - u_{S_1}\big|_{\varsigma < \varsigma_c} < 0. \qquad \text{(A3.8.1)}$$

The proof that (A3.8.1) is correct follows from these conditions:

i. There exists a $u_{S_2} - u_{S_1}\big|_{\varsigma = \varsigma_c} < 0$,

ii. $\lim_{\varsigma \to \infty} u_{S_2} - u_{S_1} > 0$,

iii. $\frac{\partial u_{S_2} - u_{S_1}}{\partial \varsigma} > 0,\ \varsigma \in (\varsigma_c, .., \infty]$.

Only conditions i., ii., and iii. therefore need to be verified.

I shall start by proving iii.

To compute $\frac{\partial u_{S_2} - u_{S_1}}{\partial \varsigma}$, I use a modified formula rather than (3.3.10). By plugging into (3.3.1), I get:

[19] This statement is equivalent to (3.3.12).

$u_{S_2} = \left(\frac{y - k_2}{p_1} - \frac{p_2}{p_1} \cdot X_2 \right)^a + \exp(m + \varsigma) \cdot (X_2 + a_2)^d$, with $X_2 = x_2(y - k_2, p_1, p_2, A, a_2)$, as defined in (3.3.6). Further, using (3.3.11), expression $u_{S_2} - u_{S_2}$ yields

$$u_{S_2} - u_{S_1} = \left(\frac{y - k_2}{p_1} - \frac{p_2}{p_1} \cdot X_2 \right)^d + \exp(m + \varsigma) \cdot (X_2 + a_2)^d - \left(\frac{y}{p_1} \right)^d - \exp(m + \varsigma) \cdot (a_2)^d . \qquad \text{(A3.8.2)}$$

The derivative $\frac{\partial u_{S_2} - u_{S_1}}{\partial \varsigma}$ can then be computed as follows:

$$\frac{\partial u_{S_2} - u_{S_1}}{\partial \varsigma} = \frac{\partial u_{S_2} - u_{S_1}}{\partial X_2} \cdot \frac{\partial X_2}{\partial \varsigma} + \frac{\partial u_{S_2} - u_{S_1}}{\partial \varsigma} = \ldots$$

$$\ldots = \left(d \cdot \left(\frac{y - k_2}{p_1} - \frac{p_2}{p_1} \cdot X_2 \right)^{d-1} \cdot \left(- \frac{p_2}{p_1} \right) + d \cdot \exp(m + \beta \cdot \varsigma) \cdot (X_2 + a_2)^{d-1} \right) \cdot \frac{\partial X_2}{\partial \varsigma} + \ldots$$

$$\ldots + \exp(m + \beta \cdot \varsigma) \cdot (X_2 + a_2)^d - \exp(m + \beta \cdot \varsigma) \cdot (a_2)^d . \qquad \text{(A3.8.3)}$$

I then choose $\varsigma = \varsigma_0$, which corresponds to $x_2(\bullet) = 0$. It follows from this that

$$\frac{\partial u_{S_2} - u_{S_1}}{\partial \varsigma}\Big|_{\varsigma = \varsigma_0} = \left(d \cdot \left(\frac{y - k_2}{p_1} \right)^{d-1} \cdot \left(- \frac{p_2}{p_1} \right) + d \cdot \exp(m + \beta \cdot \varsigma) \cdot (a_2)^{d-1} \right) \cdot \frac{\partial X_2}{\partial \varsigma} . \qquad \text{(A3.8.4)}$$

It follows from (3.3.4) that

$$\frac{X_2 + a_2}{X_1} = \left(\frac{p_1}{p_2} \cdot \exp(m + \beta \cdot \varsigma) \right)^{\frac{1}{1-d}} , \qquad \text{(A3.8.5)}$$

which I denote as *A*; see (3.3.6).

Since I chose $X_2 = 0$, it follows that

$$\frac{p_1 \cdot a_2}{y - k_2} = \left(\frac{p_1}{p_2} \cdot \exp(m + \beta \cdot \varsigma) \right)^{\frac{1}{1-d}} . \qquad \text{(A3.8.6)}$$

Plugging this into (A3.8.4) yields

$$\frac{\partial u_{S_2} - u_{S_1}}{\partial \varsigma}\Big|_{\varsigma = \varsigma_0} = \ldots$$

$$\ldots = \left(d \cdot \left(a_2 \cdot \left(\frac{p_1}{p_2} \cdot \exp(m + \beta \cdot \varsigma) \right)^{-\frac{1}{1-d}} \right)^{d-1} \cdot \left(- \frac{p_2}{p_1} \right) + d \cdot \exp(m + \beta \cdot \varsigma) \cdot (a_2)^{d-1} \right) \cdot \frac{\partial X_2}{\partial \varsigma} = 0. \qquad \text{(A3.8.7)}$$

If any value $\varsigma > \varsigma_0$ that corresponds to $x_2(\bullet) > 0$ is plugged into (A 3.8.2), derivative $\frac{\partial u_{S_2} - u_{S_1}}{\partial \varsigma}$ is greater than zero. The proof for this is as follows:

If X_2 increases, then also both expressions

$-\frac{p_2}{p_1} \cdot d \cdot \left(\frac{y-k_2}{p_1} - \frac{p_2}{p_1} \cdot X_2 \right)^{d-1} + d \cdot \exp(m + \beta \cdot \varsigma) \cdot (X_2 + a_2)^{d-1}$ and

$\exp(m + \beta \cdot \varsigma) \cdot (X_2 + a_2)^d - \exp(m + \beta \cdot \varsigma) \cdot (a_2)^d$ increase.

Since $\partial X_2 / \partial \varsigma > 0$ – as shown in Appendix A 3.7 – and from (A 3.8.7), it follows that $\frac{\partial u_{S_2} - u_{S_1}}{\partial \varsigma} > 0$ for all $\varsigma > \varsigma_0$.

I shall now prove i. The proof is straightforward: plugging $\varsigma = \varsigma_0$ that corresponds to $x_2(\bullet) = 0$ into (A 3.8.2) yields[20]

$$u_{S_2} - u_{S_1} = \left(\frac{y-k_2}{p_1} \right)^d - \left(\frac{y}{p_1} \right)^d < 0. \tag{A3.8.8}$$

The proof of ii. is also straightforward: Plugging $\varsigma = \infty$ that corresponds to $x_2(\bullet) = \frac{y-k_2}{p_2}$ – as shown previous in this Appendix A3.7 – in (A3.8.2) yields[21]

$$\lim_{\varsigma \to \infty} u_{S_2} - u_{S_1} = \lim_{\varsigma \to \infty} \exp(m + \varsigma) \cdot \left(\left(\frac{y-k_2}{p_2} + a_2 \right)^d - (a_2)^d \right) - \left(\frac{y}{p_1} \right)^d = \infty \ . \tag{A3.8.9}$$

Note that it also follows from i., ii., and iii. that

$$\varsigma_c > \varsigma_0 . \tag{A3.8.10}$$

The proof I have just presented can also be illustrated:

---

[20] Note that $u_{S_2} - u_{S_1} = \left(\frac{y-k_2}{p_1} \right)^d + \exp(m + \beta \cdot \varsigma) \cdot (a_2)^d - \left(\frac{y}{p_1} \right)^d - \exp(m + \beta \cdot \varsigma) \cdot (a_2)^d = \left(\frac{y-k_2}{p_1} \right)^d - \left(\frac{y}{p_1} \right)^d < 0$.

[21] Recall that parameter a_2 is always greater than zero.

$$\lim_{\varsigma \to \infty} u_{S_2} - u_{S_1} = \lim_{\varsigma \to \infty} \left(\frac{y-k_2}{p_1} - \frac{p_2}{p_1} \cdot \frac{y-k_2}{p_2} \right)^d + \exp(m + \beta \cdot \varsigma) \cdot \left(\frac{y-k_2}{p_2} + a_2 \right)^d - \left(\frac{y}{p_1} \right)^d - \exp(m + \beta \cdot \varsigma) \cdot (a_2)^d = \ldots$$

$$\ldots = \lim_{\varsigma \to \infty} \exp(m + \beta \cdot \varsigma) \cdot \left(\left(\frac{y-k_2}{p_2} + a_2 \right)^d - (a_2)^d \right) - \left(\frac{y}{p_1} \right)^d = \infty .$$

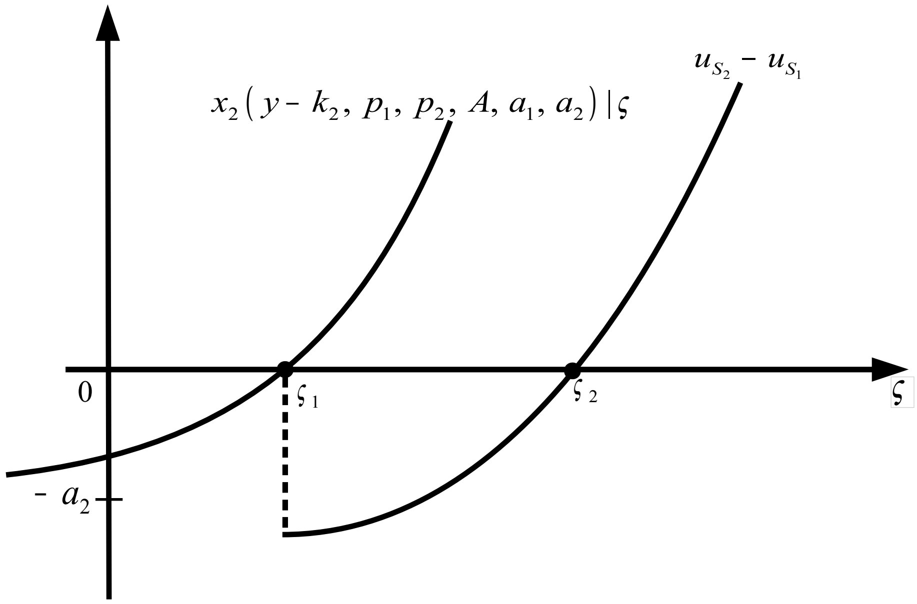


Figure A3.8.1: An illustration of the effect of the relative preference on choice.

A 3.9 The impact of model parameters on the minimal driving distance

In this subchapter, I shall illustrate the impact of changes in model parameters a_2 and d on the minimum driving distance $X_2(\varsigma_c)$, since this is one of the key points of this model. The impact of parameters d and a_2 of the pdf of driving distance X_2 has already been illustrated in Subchapter 3.3; see Figures 3.3.2 and 3.3.3. This impact can best be illustrated by an x_1/x_2-diagram.

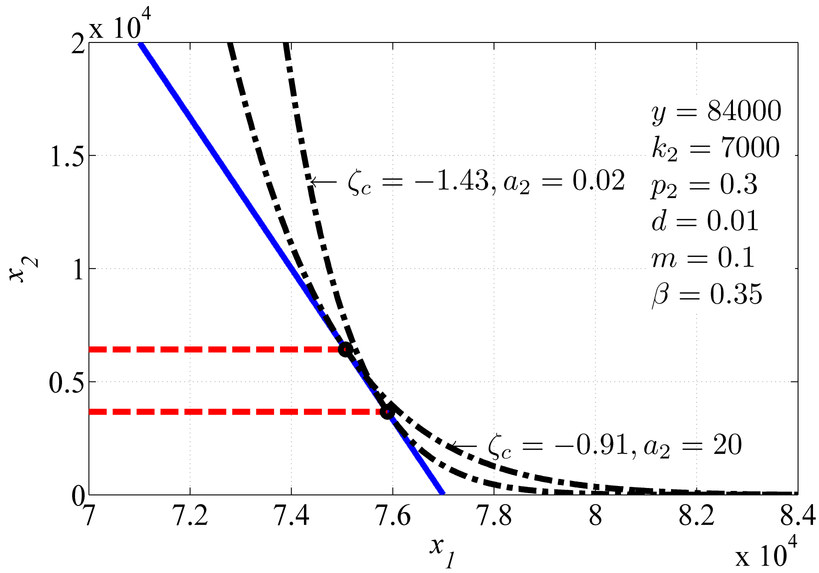


Figure A3.9.1: Indifference curves and minimum consumption for different parameters a_2.

This diagram shows that the minimum consumption levels indicated by the dashed lines increase if parameter a_2 increases. It also shows that for $a_2 = 0.02$ the preference where households would be indifferent between owing and not owning a car ζ_c is smaller than in the case $a_2 = 20$. This also explains why probability $P(X_2 = 0|\theta, p_1, p_2, y, k_2, s)$ becomes smaller if a_2 decreases. Note that if the indifference curves of the utility function were not restricted to $x_2 \geq 0$, they would approach the horizontal line at value -a_2.

The following diagram shows that the minimum consumption levels indicated by the dashed lines increase if parameter d increases. It also shows that for $d = 0.02$ the preference where households would be indifferent between owning and not owning a car ζ_c is smaller than in the case $d = 0.3$. This also explains why probability $P(X_2 = 0|\theta, p_1, p_2, y, k_2, s)$ becomes smaller if d decreases. Note also that the lower the value of d, the more cornered the indifference curve is. Note that if the indifference curves of the utility function were not restricted to $x_2 \geq 0$, they would approach the horizontal line at value -a_2.

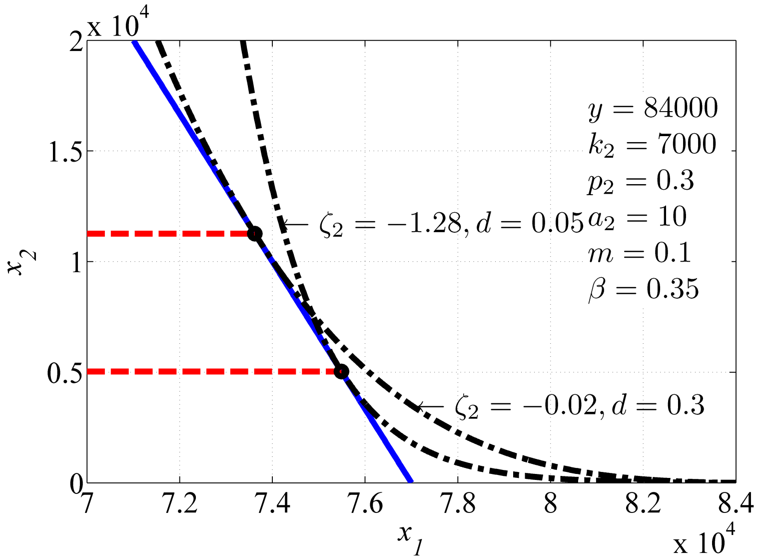


Figure A3.9.2: Indifference curves and minimum consumption for different parameters *d*.

A 3.10 The impact of the choice of penalty function on the results

In this appendix, I shall discuss the impact of the choice of parameters c_1 and c_2 of penalty function (3.3.25). Using the result of the grid defined by d = (0.001, 0.005, 0.01, 0.02, 0.05, 0.1, 0.15, 0.2, 0.3) and a_2 = (0.02,0.1, 0.2, 0.3, 0.5, 1, 2, 5, 10, 20, 100), all three error components as defined in penalty function (3.3.25), namely the proportion of datasets that have to be eliminated from the dataset, the relative deviation of the simulated proportion of carless households to the actual proportion, and the relative deviation of the total simulated annual driving distance to the actual driving distance, were computed. These results are illustrated in Figures 3.3.4, 3.3.5 and 3.3.6. From these results, I computed the penalty function for a set of parameter combinations. For combinations 2 to 4, one error component always had half the weight. For combinations 5 to 7, one error component always had double weight, while for combination one, all error components had the same weight.

	c_1	c_2	a_2	d	$\frac{drop.}{total}$	$\frac{\Delta P}{P}$	$\frac{\Delta E}{E}$	ε_{X_2,p_2}	ε_{P,p_2}	$\varepsilon_{X_2,y}$	$\varepsilon_{P,y}$
1	1	1	0.2	0.15	8.96	1.738	15.35	-1.189	0.302	1.188	-1.464
2	1	0.5	0.2	0.15	8.96	1.738	15.35	-1.189	0.302	1.188	-1.464
3	0.5	1	2	0.10	8.43	-2.649	15.56	-1.126	0.277	1.180	-1.482
4	2	2	5	0.05	6.68	-6.219	15.66	-1.071	0.247	1.171	-1.490
5	0.5	0.5	20	0.02	8.82	-0.964	15.56	-1.046	0.260	1.184	-1.474
6	2	1	0.3	0.15	9.06	3.531	15.17	-1.190	0.306	1.191	-1.459
7	1	2	5	0.05	6.68	-6.219	15.66	-1.071	0.247	1.171	-1.490
mean			4.671	0.096	8.23	-1.292	15.47	-1.126	0.277	1.182	-1.475
median			2	0.100	8.82	-0.964	15.56	-1.126	0.277	1.184	-1.474
min			0.2	0.020	6.68	-6.219	15.17	-1.190	0.247	1.171	-1.490
max			20	0.150	9.06	3.531	15.66	-1.046	0.306	1.191	-1.459
stddev			7.088	0.056	1.07	3.918	0.18	0.064	0.026	0.008	0.013
stddev/mean			1.517	0.584	0.13			-0.057	0.095	0.007	-0.009

Table A3.10.1: The effect of different penalty functions on simulation results.

The results show that although the weighting parameters were changed by a factor of two, the simulated elasticities remained within an interval with a length of less than 13% of its absolute magnitude.[21]

[21] This result stems from the following calculation: 0.129 = (1.190-1.046)/(1.190+1.046)·2.

Similar results are obtained if all simulated elasticities of all combinations of d and a_2 are computed, for which the value penalty function is no more than 10% higher than for the combination of d and a_2 that yields the lowest value of the penalty function.

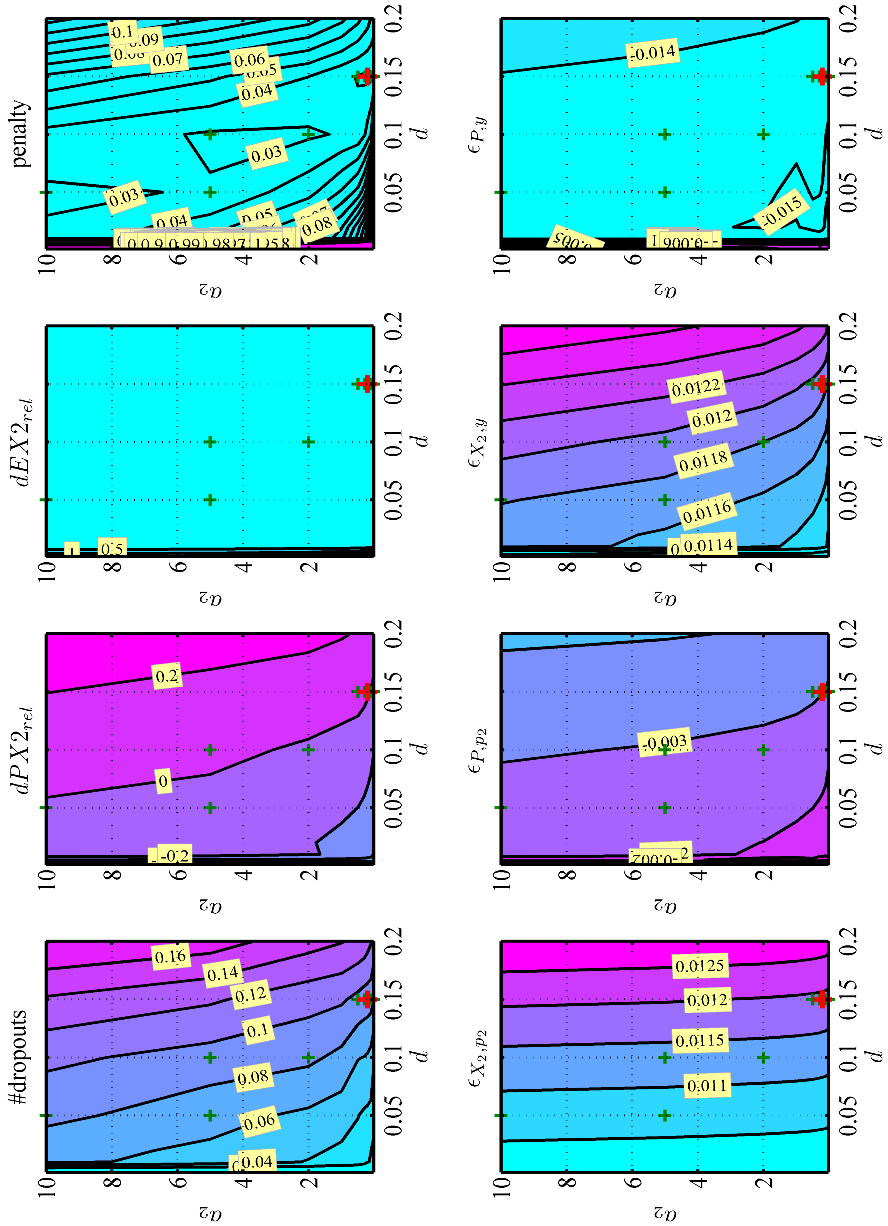


Figure A3.10.1: Simulation results for a set of points yielding low penalty values.

Note that the red cross indicates the optimum combination of d and a_2 that yields the minimal value of the penalty function, whereas the green crosses indicate the combinations of d and a_2 that yield higher values of the penalty function that are, however, less than 10% higher than the minimal value of the penalty function.

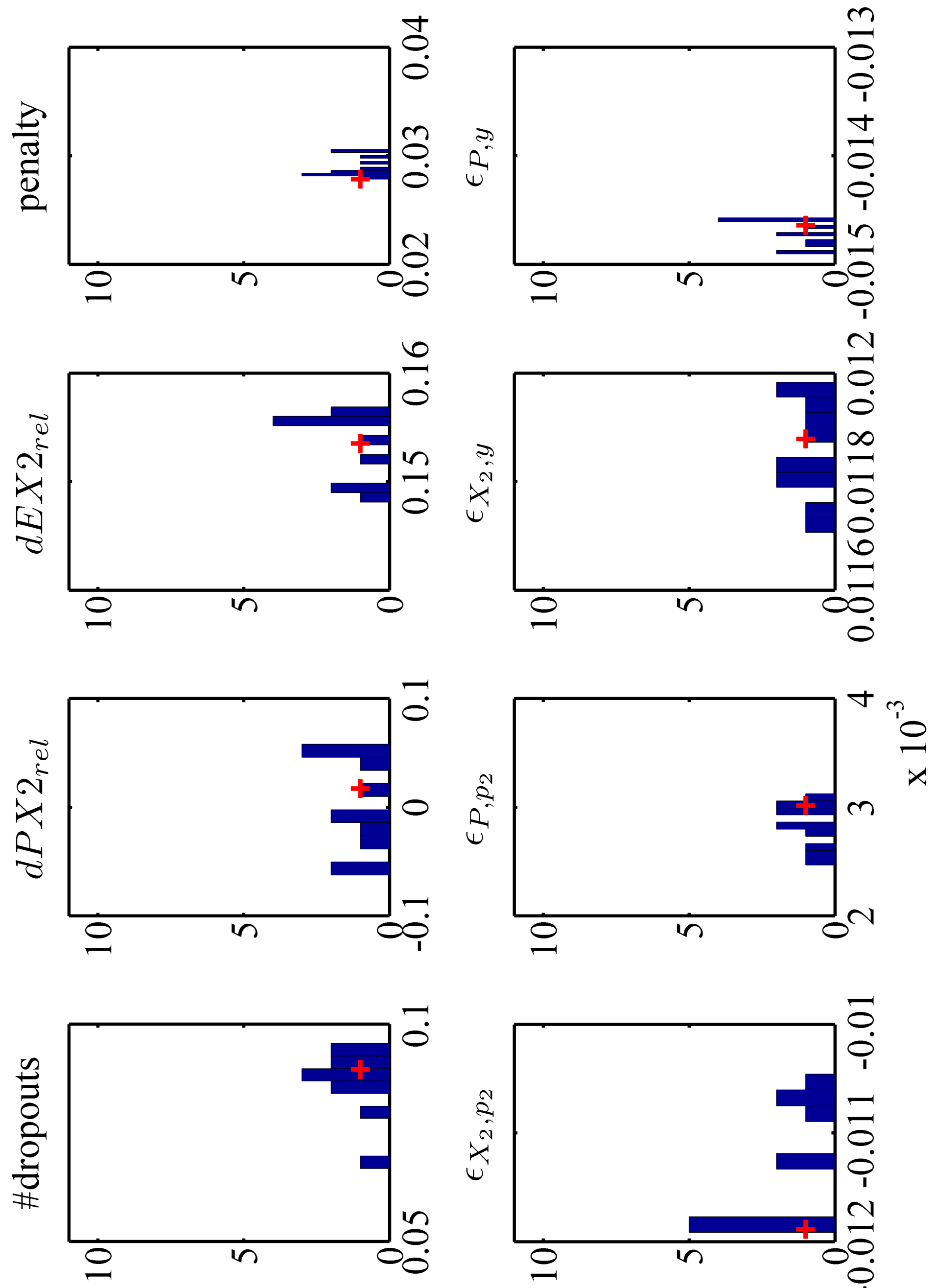


Figure A3.10.2: Simulation results for a set of points yielding low penalty values.

In this case, the resulting simulated values for the different combinations *d* and a_2 do not differ much either.

A 3.11 Comparison of results based on micro-censuses 2000 and 2005

I also applied the estimation and simulation to the 2000 micro-census data, Bundesamt für Statistik (2001). In the following, I will present the results and discuss the differences between the results based on the 2000 dataset and the results based on the 2005 dataset. Later I will examine how well the model fits the data in the same way as in Appendix A3.4 for the case where I estimated the model using data from 2005.

Comparison of the results

I start by comparing the point estimates of the estimated parameters:

Estimated parameters (Micro-census 2000)	Estimated parameters (Micro-census 2000)	Estimated parameters (Micro-census 2005)
$d = 0.1$	$d = 0.15$	$d = 0.15$
$a_2 = 1$	$a_2 = 0.2$	$a_2 = 0.2$
$\gamma_1 = -2.840$	$\gamma_1 = -2.726$	$\gamma_1 = -2.759$
$\gamma_2 = 0.3669$	$\gamma_2 = 0.3397$	$\gamma_2 = 0.3445$
$\beta = 0.4255$	$\beta = 0.3904$	$\beta = 0.3793$

Table A3.11.1: Comparison of estimated parameters.

The estimated parameters differ only slightly. Even if parameters a_2 and d are fixed at the same value, the difference between the two models with respect to the fuel price elasticity of the aggregate demand cannot be concluded from the differences in parameters. Nor can certain questions be answered, such as whether preference for driving demand increased between 2000 and 2005. This last question, however, can be answered by comparing the average simulated values for the share of carless households and the expectation values of driving demand presented in the following table:

	"Base" (simu-lated)	Effect of changes in economic variables and household location, re-lative changes	Effect of changes in preferences, relative changes	Sum of effects	Actual relative change (relative difference between simulated values)
Data used	mz00	mz05	mz00	–	mz05
Parameters used	θ_{00}	θ_{00}	θ_{05}	–	θ_{05}
$P(X_2 = 0)$	21.18% (21.20%) [19.81%]	– 9.80%	+ 13.08%	+ 2.28%	+ 2.19% (– 2.01%) [– 4.59%]
$E(X_2)$	14,056 (16,425) [13,359]	+ 9.77%	– 4.38%	+ 5.39%	+ 5.77% (+ 5.21%) [+ 3.97%]

Table A3.11.2: Comparison of simulated changes for 2005 with respect to 2000.[22]

To understand the figures in this table, a number of explanations are required. The second column contains the average driving distance $\text{mean}(x_2)$ and the proportion of carless households of the sample from which observations that, according to the model that includes the fixed costs, were "irrationally low" were eliminated. Parameters θ_{00} and θ_{05} are estimated based on the samples from which "irrationally low" observations have been eliminated. The values in parentheses "(...)" are the averages of the simulated values. The third column lists the changes in simulated values based on the micro-census data of 2005, "mz05", Bundesamt für Statistik (2006). These changes, driven by the changes in income, fuel prices and household location, can be considered to be forecasting values. Compared to the values that are actually observed, as listed in the last column, these forecast effects are too high. The questions now is why the model estimated by the micro-census data of 2000, "mz00", Bundesamt für Statistik (2001), results in inaccurate forecasts. One problem could be that the behaviour of households according to the model is defined by parameters $\theta_{00,}$ and is therefore assumed to be fixed. Since parameters θ_{05} reflect the behaviour of households in 2005, I wish to examine the effects of changes in behaviour on driving demand and on the proportion of carless households. To this end, I simulated the behaviour of households for parameters θ_{00} and parameters θ_{05}, respectively; based in both cases on the data "mz00". The results, presented in the fourth column, show that the preference for driving must have reduced between 2000 and 2005: the effect of this behaviour is that the number

[22] A similar picture emerges when we compare the simulated changes for 2000 with respect to 2005.
Note that the empirical values 21.18% and 14,056 km correspond to the dataset from which observations considered to be irrational according to that model were eliminated. The values in brackets "[...]" correspond to the values of the dataset from which only households that reported annual kilometres of more than 60,000 km or that would spend more than 30% to cover driving costs were eliminated.

of carless households increases by 13.09% and driving distance decreases by 4.38% due to the change in household behaviour. Adding these two effects, namely the effect of the changes in economic variables and changes in household properties and the effect of the change in households' preferences, should yield the "true" effects. The results (see the penultimate column) show that this is more or less the case: the forecast changes are rather closed to the changes actually observed. The interesting finding that household preferences for car driving has reduced over time, therefore seems to be quite trustworthy. The question is now whether such a change is plausible. We believe this is indeed the case. This change in preference could have occurred due to an improvement in public transport (see Figure 3.3.9) or due to less utility from car driving caused by denser traffic.

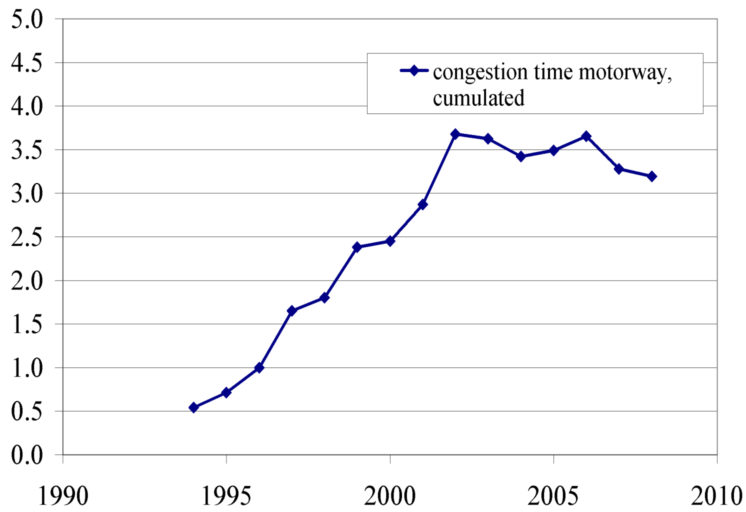


Figure A3.11.1: Congestion time on Swiss motorways.[23]

The above diagram shows that the cumulated congestion time on Swiss motorways increased by 42%, namely from 7,711 hours to 10,975 hours between 2000 and 2005.[24] In the same period, congestion caused by too much traffic more than doubled, namely from approximately 2,300 hours to about 4,800 hours.[25] These values indicate the total numbers of hours in which congestion could be observed. I estimate that about 9.4% of trips undertaken in 2005 were affected if the journeys are weighted by

[23] This data can be found in Bundesamt für Strassen ASTRA (2007) and Bundesamt für Strassen ASTRA (2010).

[24] This data can be found in Bundesamt für Strassen ASTRA (2010).

[25] See Figure 3 on page 7 in Bundesamt für Strassen ASTRA (2007).

length.[26] Since all these congestion points on motorways are situated parallel to train routes where trains run at least every half hour, as shown by the two following diagrams, I expect that the increase in such congestion will lead to a shift by some drivers towards using public transport.

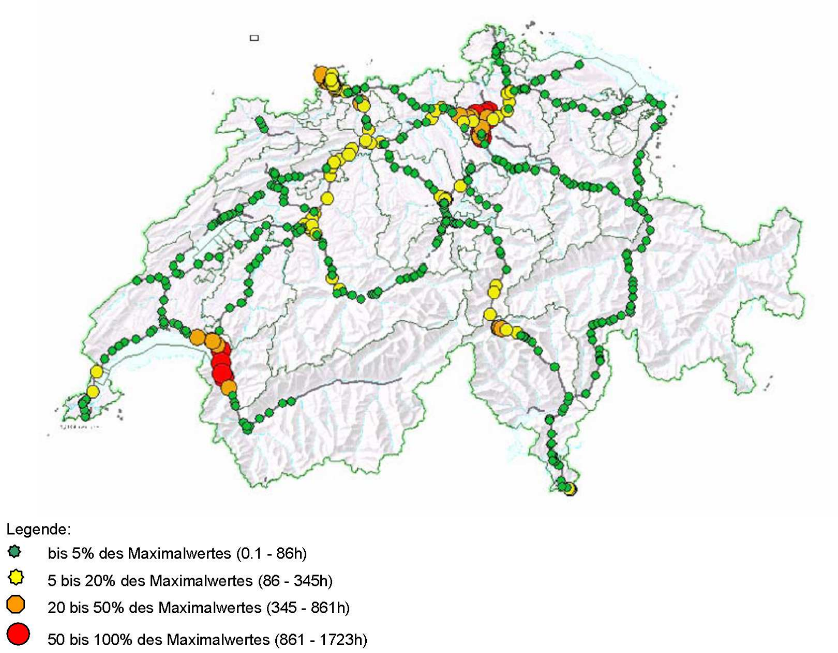


Figure A3.11.2: Congestion time on Swiss motorways.[27]

[26] Let us consider a spot of congestion on a motorway close to Zürich where about 50,000 cars pass in each direction every day. This congestion affects about 22% of cars travelling. I came to the value of 22% by examining the frequency of vehicles that use this motorway. I considered the three hours when car frequency is the highest to be the hours of congestion. 22% of the vehicles use this motorway at this particular time. The data is based on frequency data provided by Bundesamt für Statistik, 2006b, diagram "Tagesganglinien DTW".At least about $22{,}000 = 2 \cdot 0.22 \cdot 50{,}000$ vehicles are therefore involved in traffic congestion every day. Assuming that these cars drive about 60 km per direction on average, this corresponds to $22{,}000 \cdot 60 \cdot 360 = 0.475 \cdot 10^9$ annual trip kilometres that are affected by this congestion. Since there are about six other spots that contribute about half of this spot to "congestion affected trip kilometres", I estimate the total amount of "congestion affected trip kilometres" to be $4 \cdot 1.3 \cdot 10^9 = 5.2 \cdot 10^9$ kilometres. Since there are about four spots that contribute about the same amount and about six that generate about half of the amount of "congestion affected trip kilometres", the total of these "hot spots" might be about $3.8 \cdot 10^9 = 8 \cdot 0.475 \cdot 10^9$ "congestion affected trip kilometres". I assume the sum of "congestion affected trip kilometres" by congestion points with much fewer congestion hours to be about half that of these "hot spots". I therefore assume the total "congestion affected trip kilometres" to be about $5.7 \cdot 10^9$ kilometres. Since the total of car kilometres travelled by Swiss households is $60.464 \cdot 10^9$ kilometres, a total of about 9.4% of all trips – weighted by distance – were affected by traffic congestion.

[27] This diagram can be found in Bundesamt für Strassen ASTRA (2010: 10), in "Jahresbericht über den Verkehrsfluss und das Staugeschehen auf den Nationalstrassen 2005".

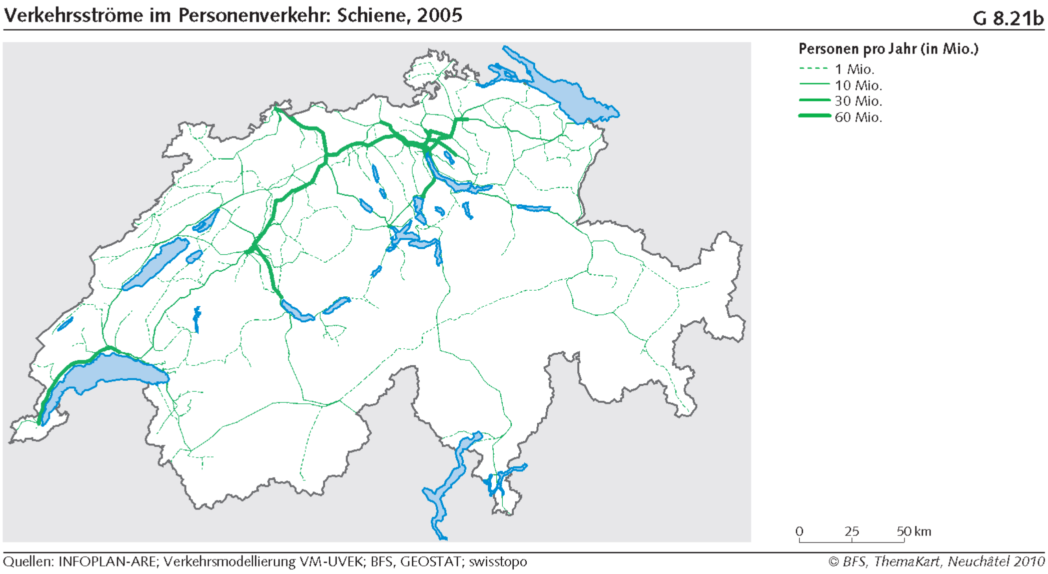


Figure A3.11.3: Number of passengers of different railway lines in Switzerland.[28]

To verify whether this preference effect really occurs, I also test whether the sum of the two effects also yields the actual observed changes in the case where the data from 2000 is simulated based on the model estimated using the 2000 data. Not surprisingly, the following change shows that the total change can be explained by the sum of the effect of the change in preferences and the change in preferences. In addition, the two tables also contain information on the empirical mean of distance driven per household and the proportion of carless households. Note also the effects that resulted from eliminating certain observations from the dataset, as visible in the table.

[28] This diagram can be found in Bundesamt für Statistik (2010d: 52).

	"Base" (simulated values)	Effect of changes in economic variables and household location, relative changes	Effect of changes in preferences, relative changes	Sum of effects	Actual relative change (relative difference between simulated values)
Data used	mz05	mz00	mz05	–	mz05
Parameters used	θ_{05}	θ_{00}	θ_{05}	–	θ_{05}
$P(X_2=0)$	20.72% (21.62%) [18.9%]	– 10.85%	+ 11.59%	+ 0.74%	+ 2.14% (– 2.01%) [+ 4.81%]
$E(X_2)$	14'868 km (17'281 km) [13'890 km]	+ 9.11%	– 4.34%	– 5.77%	– 5.46% (– 4.95%) [– 3.82%]

Table A3.11.3: Comparison of simulated changes for 2000 with respect to 2005.[29]

Note that all driving distances are measured per household. If we were interested in relativity changes per capita, a value of 1.41% would need to be subtracted[30] from the value corresponding to the growth between 2000 and 2005, as presented in Table A3.11.2. If we were interested in the relative changes in the driving demand of the whole population[31], a value of 1.65% would need to be added to the value corresponding to the growth between 2000 and 2005, as presented in Table A3.11.2.

I shall now present the elasticities that resulted based on the micro-census dataset of 2000 and will compare all values to the results based on the 2005 micro-census dataset. The results show that the

[29] Note that the empirical values 21.18% and 14,056 km correspond to the dataset from which observations that were considered to be irrational according to that model were eliminated. The values in brackets "[...]" correspond to the values of the dataset from which only households that reported annual kilometres of more than 60,000 km or that would spend more than 30% to cover the driving costs were eliminated. The values in brackets "(...)" correspond to the values of the dataset that no longer includes households that reported annual kilometres of more than 60,000 km or that would spend more than 30% to cover driving costs , or observations that contain driving distances that are considered as "irrationally" low according to the model including fixed costs of car ownership. Note that this elimination leads to an increase in the proportion of carless households and an increase in the average driving distance of the remaining data compared to the initial dataset.

[30] The reason for this is that the average number of people per household included in the dataset increased by 1.41%, see Bundesamt für Statistik (2001) and (2006a).

[31] This value is yielded because the Swiss population has grown by 3.06%, namely from 7,266,920 to 7,489,370; see Heston et al. (2010), and the average number of people per household included in the dataset increased by 1.41%; see Bundesamt für Statistik (2001) and (2006a). Each household in the 2005 dataset therefore represents 1.65% = (3.06 – 1.41)% more people than in the 2000 dataset.

elasticities changed little, despite the fact that preferences for car driving changed between these year as shown above.

Model that includes fixed costs (Micro-census 2000)	Model that includes fixed costs (Micro-census 2005)	Model without fixed costs (Micro-census 2005)
$\varepsilon_{P(X_2=0),p_2} = \underset{(-)}{0.249}$	$\varepsilon_{P(X_2=0),p_2} = \underset{(0.0001)}{0.297}$	$\varepsilon_{P(X_2=0),p_2} = \underset{(0.00884)}{1.6339}$
$\varepsilon_{E(X_2),p_2} = \underset{(-)}{-1.124}$	$\varepsilon_{E(X_2),p_2} = \underset{(0.0001)}{-1.189}$	$\varepsilon_{E(X_2),p_2} = -\underset{(0.0042703)}{1.3625}$
$\varepsilon_{P(X_2=0),p_{fuel}} = \underset{(-)}{0.1235}$	$\varepsilon_{P(X_2=0),p_{fuel}} = \underset{(0.00005)}{0.14}$	$\varepsilon_{P(X_2=0),p_{fuel}} = \underset{(0.00404)}{0.6819}$
$\varepsilon_{E(X_2),p_{fuel}} = \underset{(-)}{-0.492}$	$\varepsilon_{E(X_2),p_{fuel}} = -\underset{(0.00005)}{0.54}$	$\varepsilon_{E(X_2),p_{fuel}} = \underset{(0.00195)}{-0.5639}$
$\varepsilon_{P(X_2=0),y} = \underset{(-)}{-1.408}$	$\varepsilon_{P(X_2=0),y} = \underset{(0.01)}{-1.437}$	$\varepsilon_{P(X_2=0),y} = \underset{(0.0087525)}{1.6176}$
$\varepsilon_{E(X_2),y} = \underset{(-)}{1.186}$	$\varepsilon_{E(X_2),y} = \underset{(0.002)}{1.189}$	$\varepsilon_{E(X_2),y} = \underset{(0.004227)}{1.3488}$
$\varepsilon_{P(X_2=0),k_2} = \underset{(-)}{1.316}$	$\varepsilon_{P(X_2=0),k_2} = \underset{(0.013)}{1.39}$	-
$\varepsilon_{E(X_2),k_2} = \underset{(-)}{-0.179}$	$\varepsilon_{E(X_2),k_2} = -\underset{(0.009)}{0.18}$	-
$\frac{\Delta P_{urban \to rural}}{P_{urban}} = \underset{(-)}{-42.3\%}$	$\frac{\Delta P_{urban \to rural}}{P_{urban}} = -\underset{(1.84/100)}{43.3\%}$	$\frac{\Delta P_{urban \to rural}}{P_{urban}} = -\underset{(1.0019/100)}{45.48\%}$
$\frac{\Delta E(X_2)_{urban \to rural}}{E(X_2)_{urban}} = \underset{(-)}{45.9\%}$	$\frac{\Delta E(X_2)_{urban \to rural}}{E(X_2)_{urban}} = \underset{(1.84/100)}{46.6\%}$	$\frac{\Delta E(X_2)_{urban \to rural}}{E(X_2)_{urban}} = \underset{(1.7023/100)}{50.28\%}$
$\frac{\Delta P_{rural \to urban}}{P_{rural}} = \underset{(-)}{71.5\%}$	$\frac{\Delta P_{rural \to urban}}{P_{rural}} = \underset{(3.55/100)}{74.3\%}$	$\frac{\Delta P_{rural \to urban}}{P_{rural}} = \underset{(3.2088/100)}{80\%}$
$\frac{\Delta E(X_2)_{rural \to urban}}{E(X_2)_{rural}} = \underset{(-)}{-31.7\%}$	$\frac{\Delta E(X_2)_{rural \to urban}}{E(X_2)_{rural}} = -\underset{(0.87/100)}{32.1\%}$	$\frac{\Delta E(X_2)_{rural \to urban}}{E(X_2)_{rural}} = -\underset{(0.76374/100)}{34.06\%}$

Table A3.11.4: Comparison of simulated elasticities of the two different models.

Model quality

For the case where I estimated the model using the 2000 dataset,[32] I also computed the density functions $f_{X_2 \wedge (X_2 > 0)}$ for each combination of income category and type of place of residence to test if it adapts well to the empirical distribution. Again, as in the case when estimating the model using the 2005 dataset, the results show that the model approximates the data well, as the following diagrams show.

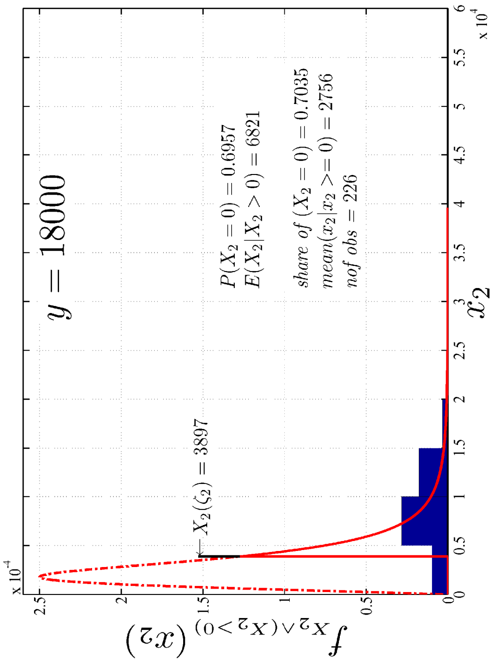


Figure A3.11.4: Histogram of households in rural areas with an income of CHF 18,000.

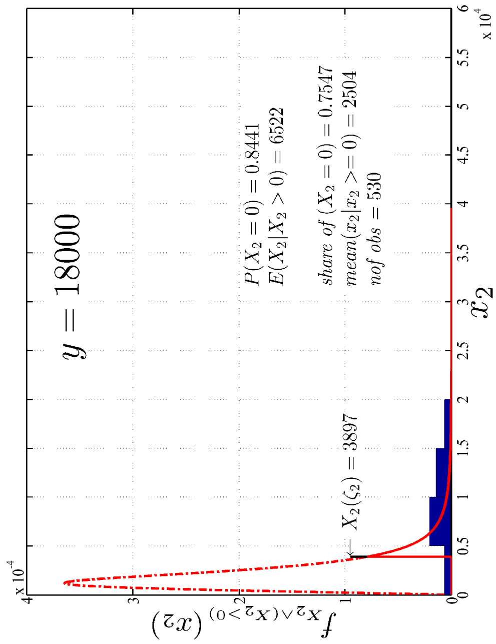


Figure A3.11.5: Histogram of households in urban areas with an income of CHF 18,000.

[32] See Bundesamt für Statistik (2001).

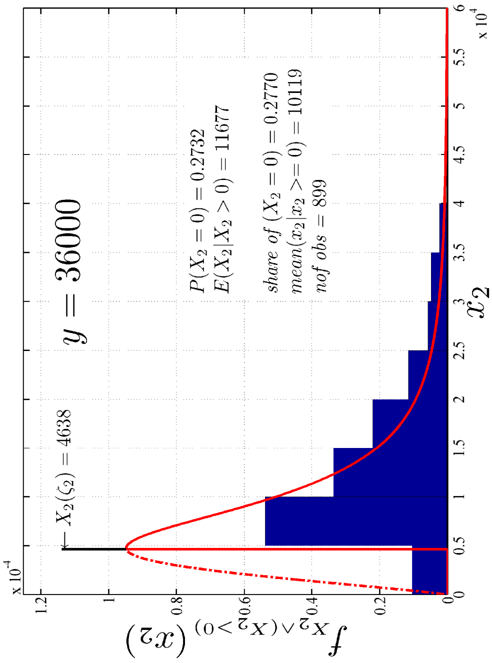


Figure A3.11.6: Histogram of households in rural areas with an income of CHF36,000.

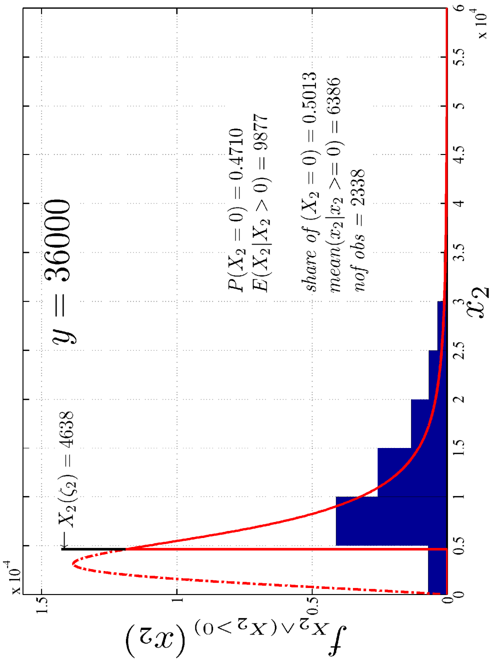


Figure A3.11.7: Histogram of households in urban areas with an income of CHF 36,000.

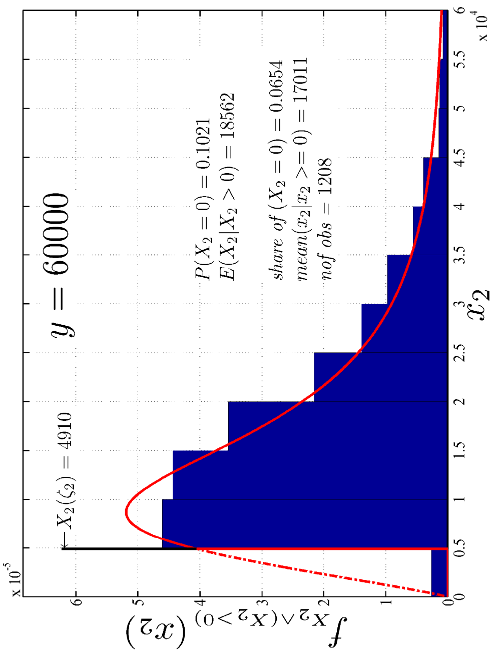


Figure A3.11.8: Histogram of households in rural areas with an income of CHF 60,000.

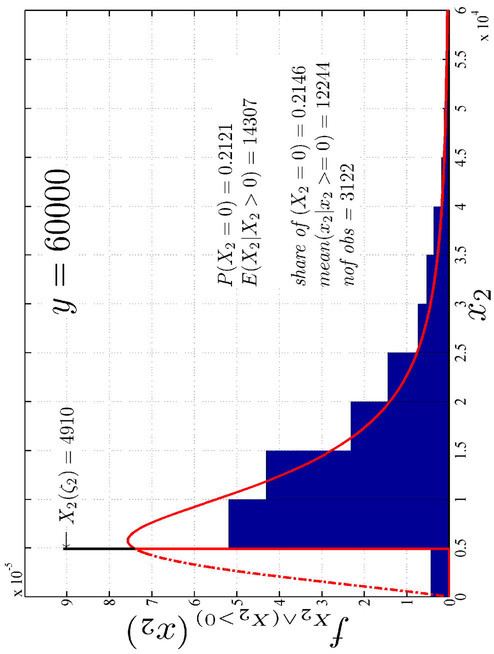


Figure A3.11.9: Histogram of households in urban areas with an income of CHF 60,000.

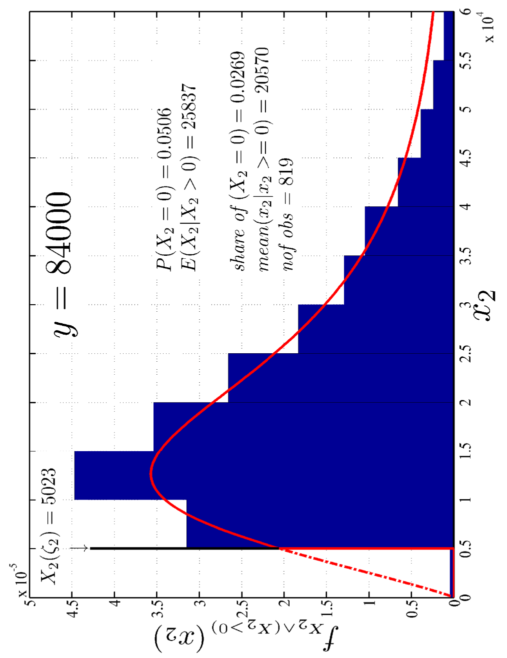


Figure A3.11.10: Histogram of households in rural areas with an income of CHF 84,000.

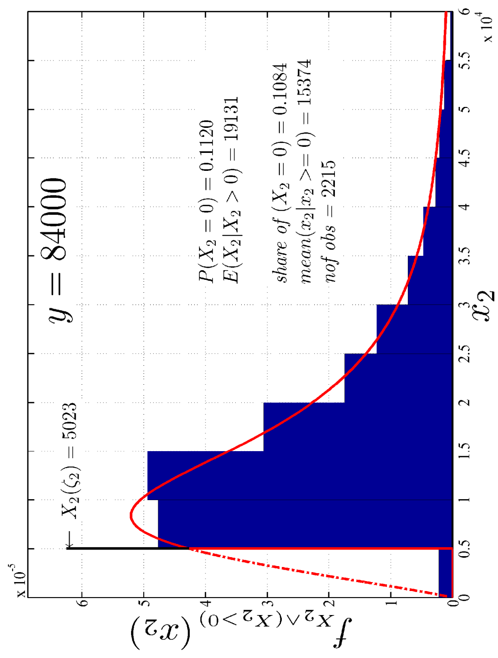


Figure A3.11.11: Histogram of households in urban areas with an income of CHF 84,000.

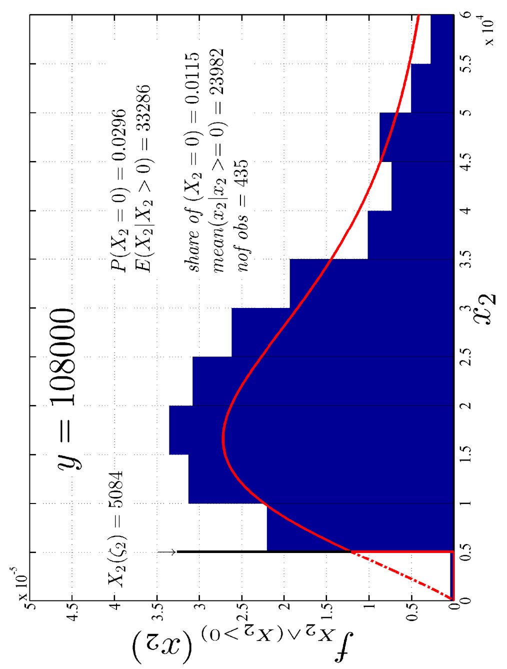


Figure A3.11.12: Histogram of households in rural areas with an income of CHF 108,000.

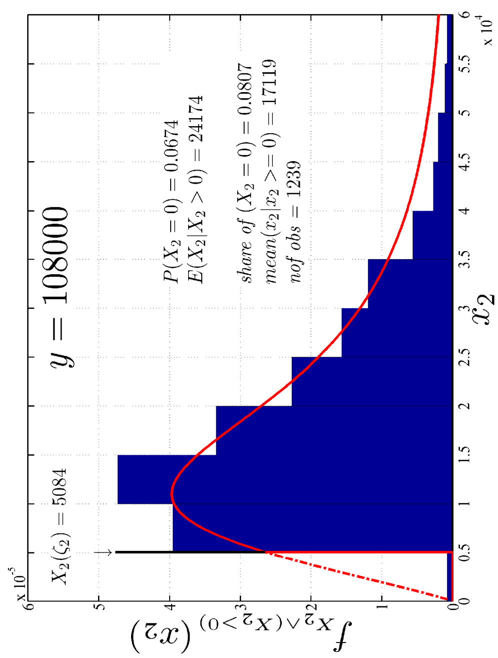


Figure A3.11.13: Histogram of households in urban areas with an income of CHF 108,000.

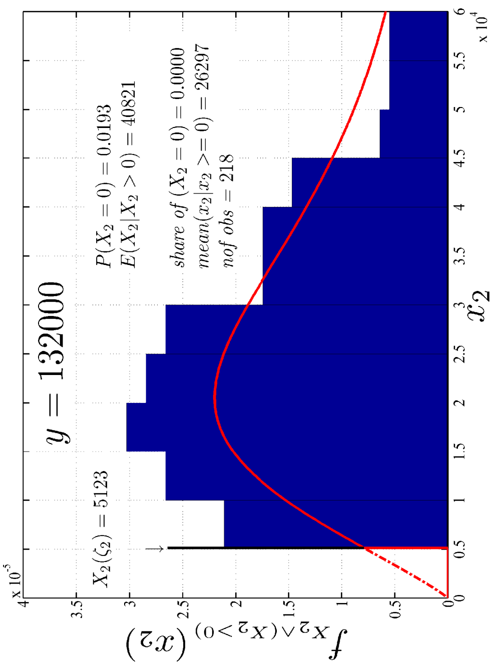


Figure A3.11.14: Histogram of households in rural areas with an income of CHF 132,000.

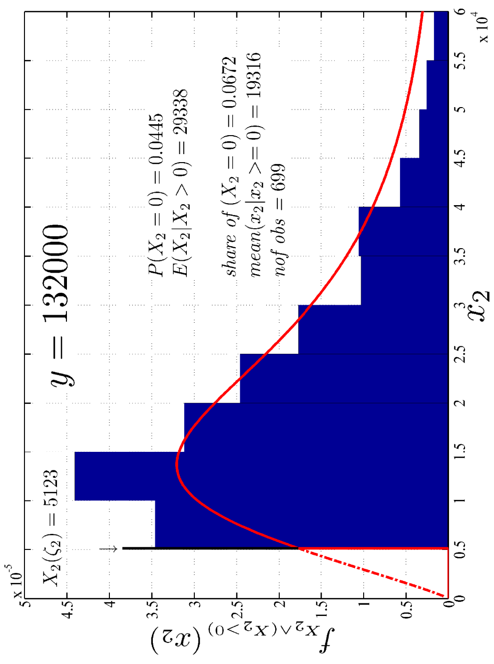


Figure A3.11.15: Histogram of households in urban areas with an income of CHF 132,000.

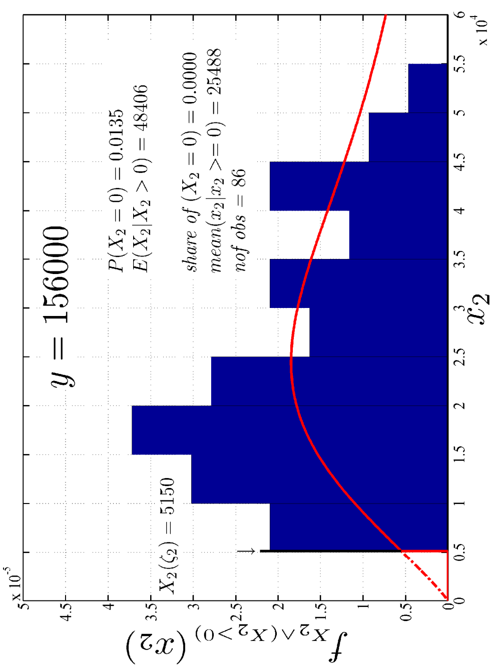


Figure A3.11.16: Histogram of households in rural areas with an income of CHF 156,000.

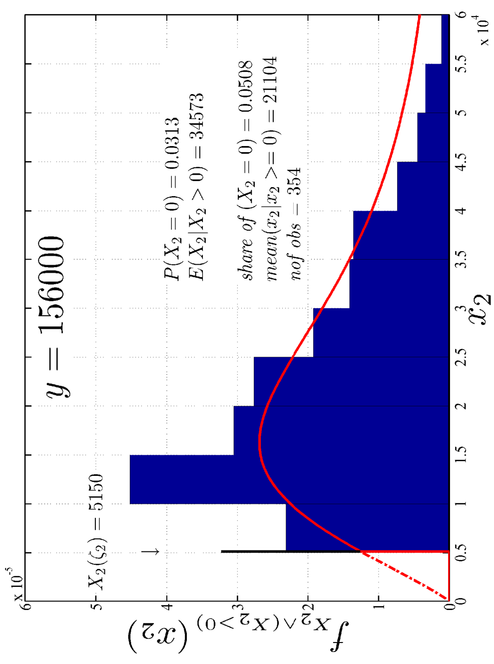


Figure A3.11.17: Histogram of households in urban areas with an income of CHF 156,000.

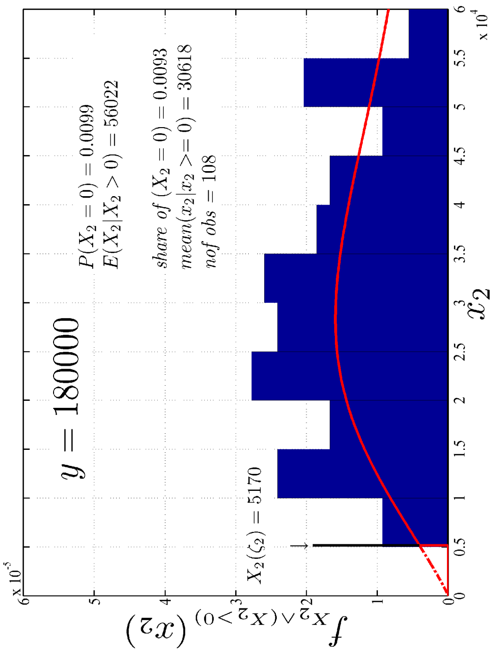


Figure A3.11.18: Histogram of households in rural areas with an income of CHF 180,000.

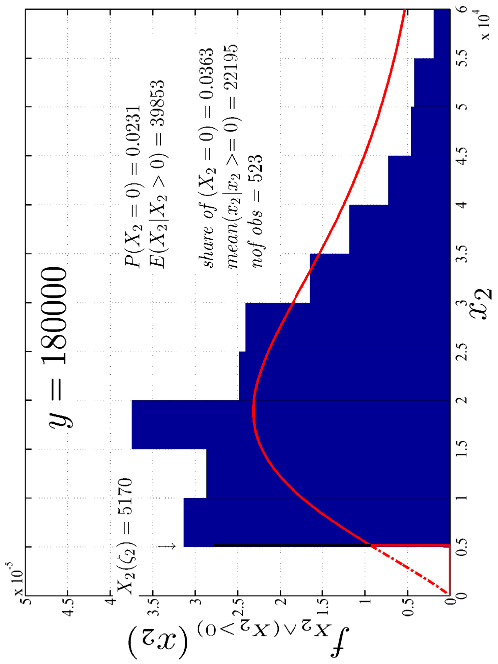


Figure A3.11.19: Histogram of households in urban areas with an income of CHF 180,000.

These diagrams show that the densities and probabilities predicted by the model correspond very well with the actual data of each household segment. The shapes of the pdfs are very similar to the shapes of the histograms.[33] The only negative aspect that can be observed is that the pdfs have too heavy tails, particularly for household income levels equal to and greater than CHF 84,000. However, these tails are less heavy than in the model without fixed costs. This also explains why the simulated values deviate less from the average empirical values of the driving distance. This can be seen when comparing the following Figures A3.11.12 and A3.11.13 with Figures A3.4.9 and A3.4.10.

[33] Note that the surface of these histograms is normalized to the share of households owning a car; the surface of the histograms in Appendix A3.4 is normalized to one.

y	Loca-tion	Nobs	Share of $x_2 = 0$	$P(X_2 = 0)$	Δ	$mean(x_2)$	$E(X_2 = 0)$	Δ	$X_2(\varsigma_c)$	ς_c
18000	urban	530	0.755	0.844	0.089	2504	1017	-1487	3897	
18000	rural	226	0.704	0.696	-0.008	2756	2075	-680	3897	0.352
36000	urban	2338	0.501	0.471	-0.030	6386	5225	-1161	4638	
36000	rural	899	0.277	0.273	-0.004	10119	8486	-1633	4638	-0.416
60000	urban	3122	0.215	0.212	-0.002	12244	11272	-972	4910	
60000	rural	1208	0.065	0.102	0.037	17011	16667	-343	4910	-0.925
84000	urban	2215	0.108	0.112	0.004	15374	16988	1614	5023	
84000	rural	819	0.027	0.051	0.024	20570	24531	3961	5023	-1.248
108000	urban	1239	0.081	0.067	-0.013	17119	22545	5426	5084	
108000	rural	435	0.011	0.030	0.018	23982	32300	8318	5084	-1.485
132000	urban	699	0.067	0.044	-0.023	19316	28034	8718	5123	
132000	rural	218	0.000	0.019	0.019	26297	40035	13738	5123	-1.672
156000	urban	354	0.051	0.031	-0.020	21104	33490	12386	5150	
156000	rural	86	0.000	0.013	0.013	25488	47755	22267	5150	-1.827
180000	urban	523	0.036	0.023	-0.013	22195	38931	16736	5170	
180000	rural	108	0.009	0.010	0.001	30618	55467	24849	5170	-1.959

Table A3.11.5: Empirical and simulated values of driving distance and proportions of carless households.

Since it is hard to detect any systematic relationship between the differences between the simulated and the empirical values from this table, I shall plot these differences with respect to income. The results, illustrated in Figure A3.11.20, show that the difference between the simulated probabilities $P(X_2 = 0)$ and the actual proportions of carless households decreases slightly with income. I conclude from this that simulated changes in $P(X_2 = 0)$ with respect to income are slightly smaller in magnitude than is actually the case. In contrast, the results illustrated in Figure A3.11.21 show that the difference between simulated expected driving demand $E(X_2)$ and actual values increases with income. I therefore expect simulated income elasticities of driving demand to be too high by up to 0.32.[34]

[34] The diagram shows that for an income difference of CHF 90,000 in the interval of CHF 20,000 to CHF 110,000, the difference between forecast and actual values increases by about 7,500 km. This equals 0.0833 km/CHF. This implies that the simulated income elasticity of driving demand may be 0.48 higher than it actually is: $0.0556\text{km/sFr} \cdot 75'598\text{sFr}/13'559\text{km} = 0.31...$, where $75'598\text{sFr}$ is the mean income and $13'559\text{km}$ is the mean annual driving distance of households. Note that I chose the interval of CHF 20,000 to CHF 110,000 because the income of more than 80% of households is within this range.

Due to the budget effect of price changes, I also expect the simulated elasticities of driving demand to be rather too high in magnitude.[35]

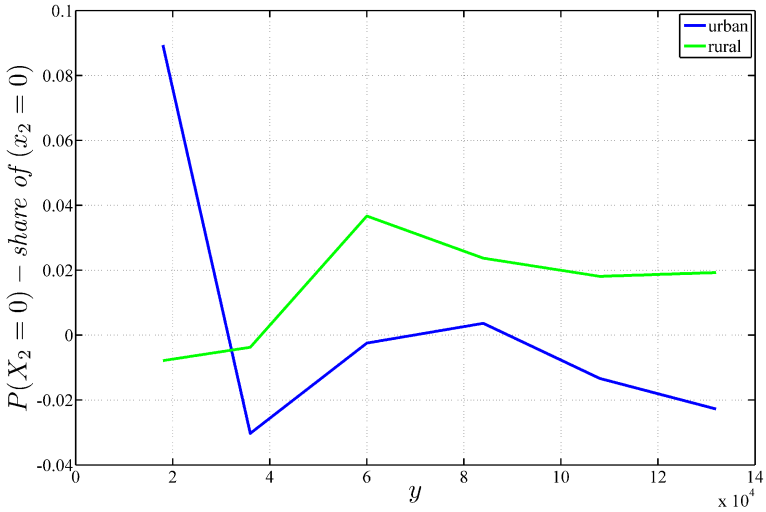


Figure A3.11.20: Deviation of the simulated proportion of carless households from the empirical value.

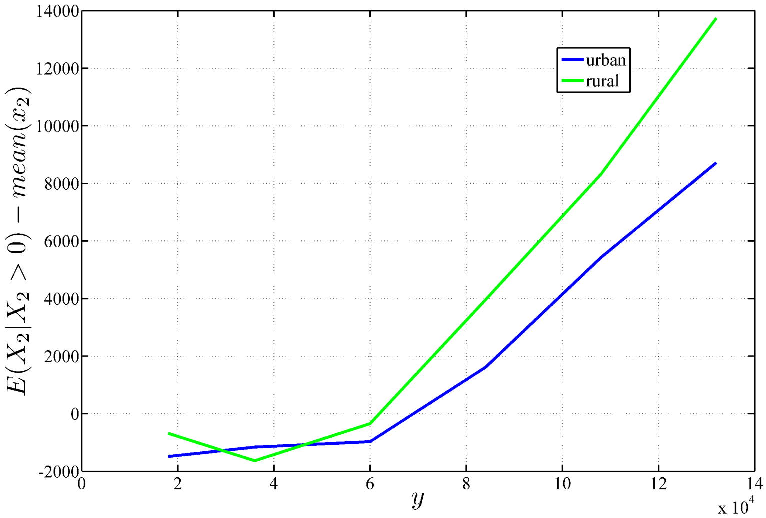


Figure A3.11.21: Deviation of the simulated driving distance from the empirical value.

---

[35] Note that since driving demand is a normal good, the budget effect is positive. The effect shown in the case of income therefore leads in the same direction. This means that the simulated price elasticity of driving demand is also expected to be too high.

A 3.12 The impact of including additional explanatory variables on the elasticities

Until now, I have not included any other variable than the type of place of residence to explain household preference for car driving. I did this to keep the model simple and to save computation time. I now wish to test whether the elasticities change if I include another variable in the model. I chose the variable "number of people", denoting the number of people belonging to a household. I chose this variable because the coefficient associated with it yielded significant positive parameter values in simple OLS and Tobit models that explained the driving demand, and because it variates quite a lot. This variable therefore explains quite a lot of the variation in driving demand. The following table shows the parameter values and the elasticities in one case for the model including the variable "number of people" and in the other case without including this variable, as was the case for the previously discussed model.

The results presented in the table show that there are only very small differences in the simulated elasticities. The estimated parameter values do not differ much either. I conclude from this that the income and type of place of residence are the dominant variables for explaining driving demand and car ownership, and that including even more socio-demographic variables of households would not considerably change the values of interest, namely the elasticities.

Model that includes the variable "number of people" (Micro-census 2000)	Model that does not include the variable "number of people" (Micro-census 2000)
$d = 0.1,\ a_2 = 1,$ $\gamma_1 = -2.982,\ \gamma_2 = 0.343,$ $\gamma_3 = 0.0616,\ \beta = 0.425$	$d = 0.1,\ a_2 = 1,$ $\gamma_1 = -2.840,\ \gamma_2 = 0.3669,$ $\beta = 0.4255$
$\varepsilon_{P(X_2=0),p_2} = \underset{(-)}{0.283}$	$\varepsilon_{P(X_2=0),p_2} = \underset{(-)}{0.249}$
$\varepsilon_{E(X_2),p_2} = \underset{(-)}{-1.186}$	$\varepsilon_{E(X_2),p_2} = \underset{(-)}{-1.124}$
	$\varepsilon_{P(X_2=0),p_{fuel}} = \underset{(-)}{0.1235}$
	$\varepsilon_{E(X_2),p_{fuel}} = \underset{(-)}{-0.492}$
$\varepsilon_{P(X_2=0),y} = \underset{(-)}{-1.406}$	$\varepsilon_{P(X_2=0),y} = \underset{(-)}{-1.408}$
$\varepsilon_{E(X_2),y} = \underset{(-)}{1.181}$	$\varepsilon_{E(X_2),y} = \underset{(-)}{1.186}$
$\varepsilon_{P(X_2=0),k_2} = \underset{(-)}{-1.365}$	$\varepsilon_{P(X_2=0),k_2} = \underset{(-)}{-1.316}$
$\varepsilon_{E(X_2),k_2} = \underset{(-)}{-0.178}$	$\varepsilon_{E(X_2),k_2} = \underset{(-)}{-0.179}$
$\frac{\Delta P_{urban\to rural}}{P_{urban}} = \underset{(-)}{-41.6\%}$	$\frac{\Delta P_{urban\to rural}}{P_{urban}} = \underset{(-)}{-42.3\%}$
$\frac{\Delta E(X_2)_{urban\to rural}}{E(X_2)_{urban}} = \underset{(-)}{44.8\%}$	$\frac{\Delta E(X_2)_{urban\to rural}}{E(X_2)_{urban}} = \underset{(-)}{45.9\%}$
$\frac{\Delta P_{rural\to urban}}{P_{rural}} = \underset{(-)}{69.5\%}$	$\frac{\Delta P_{rural\to urban}}{P_{rural}} = \underset{(-)}{71.5\%}$
$\frac{\Delta E(X_2)_{rural\to urb.}}{E(X_2)_{rural}} = \underset{(-)}{-31.1\%}$	$\frac{\Delta E(X_2)_{rural\to urb.}}{E(X_2)_{rural}} = \underset{(-)}{-31.7\%}$

Table A3.12.1: Simulated elasticities if the model includes the number of people per household.

A 3.13 Estimation results using a modified density function for computing the expectation value

As previously mentioned, computing its estimation value of the distribution of driving demand leads to the problem of a heavy tail, e.g. see Appendix A3.4, Figure 3.3.1ff. In particular, as shown in Table A3.4.1, the simulated expectation values are too high for almost all income groups and differences to the mean values of these income groups' observed driving distances increase with income. The latter effect leads to the presumption that the simulated elasticities with respect to income and driving costs might be too high. In the following, I will compute the integral based on the same pdf, but integrating only up to a distance of 60,000km.

$$E_{sim,n}\left(X_2\right) \mid \tilde{\theta}, p_1, p_{2n}, y_n, k_2, s_n = \int_{z=x_2(\varsigma_{cn})}^{z=60000} z \cdot f_{X_2 \wedge X_2 > 0}\left(z \mid \tilde{\theta}, p_1, p_{2n}, y_n - k_2, s_n\right) dz\,. \qquad \text{(A3.13.1)}$$

The estimation routine is the same as defined in the section entitled “Estimation routine” in Subchapter 3.3.[36]

Results

Compared to the results generated when computing the expectation value as defined in equation (3.3.19), the simulated expectation value is – as expected – no longer systematically too high with respect to the empirical values, see Figure A3.12.16 compared to Figure A3.13.4. Furthermore, the difference between the simulated expectation value and the mean value no longer increases in income. I conclude from this that the computed income elasticity using the modified density function of 0.78 is more realistic than that computed using the initial density function (1.19). Also, an elasticity of 0.78 is closer to the values determined in other studies. What is much more notable is the fact that the model with a modified density function yields an own price elasticity of driving demand that differs from the elasticity with respect to income. The value of -0.68 and the elasticity with respect to the fuel price of -0.28 is also much closer to the values found by other studies than the values established by the model based on the unmodified density function, which yielded values of -1.19 and -0.54, respectively. The reduction in elasticities when reimbursing tax revenues to households was about the same in both models.

---

[36] Note that although Function (3.3.16a), which is part of the ML function, is set to zero for any value above 60,000 km, the ML function is *not* multiplied by factor $P\left(X_2 \leq 60000\right)^{-1}$. Multiplying by factor $P\left(X_2 \leq 60000\right)^{-1}$ causes (3.3.16) to remain a density, which means that the sum $P\left(z = X_2 \mid \theta, p_1, p_2, y, s\right) + \int_{z=-\infty}^{z=60000} f_{X_2 \wedge X_2 > 0}\left(z \mid \theta, p_1, p_2, y - k_2, s\right) dz$ equals one. Since for most income levels the probability $P\left(X_2 > 60000\right)$ is very small, I do not expect this simplification to cause major errors in the results.

The differences between the forecast simulated probabilities of being carless and the actual proportion of carless households are about the same for both density functions, see A3.13.3 and A3.11.20.[37] But while this difference tends to decrease with income in the case of the model based on the initial density function, it increases in the case of the model based on the modified density function. In both cases, in the range of income that covers most households, the differences between the forecast simulated probabilities of being carless and the actual proportion of carless households do not change much. This also seems to be the reason why the elasticities with respect to being carless do not differ much between the two models.

When comparing the simulated change in driving demand when households move from rural to urban areas, and vice versa, the magnitude of the change is much lower for the model based on the modified density function. The reason for this seems to be that, in the case for rural households and high incomes, a lot of the distribution is cut off, as can particularly be seen by densities A3.13.15ff. In these cases, this effect even leads to the result that when income increases from CHF 156,000 to CHF 228,000, the expectation value of driving demand decreases, see also Figures A3.13.19 and A3.13.21. I conclude from this that this model is not suitable for predicting driving demand for households with high incomes. Further, since the simulated driving demand for households living in urban areas is rather too high, and too low for rural households, I conclude that the simulated effects of changes in households is rather too low in this model. On the other hand, this bias occurs only in high incomes. Since these high incomes amount for a rather small proportion of the total population, the bias at aggregate level might be rather small.

Finally, the model with the modified density function adapts better to the data. This is reflected in a lower penalty value *Q*. Adaptation to the shape of the histogram of the observed data is similarly accurate in both cases, see Figures A3.11.4 ff. versus A3.13.5 ff. Since the model based on the unmodified density function yields too high values for the forecast expectation value $E\left(X_2\right)$, due to the problem of the heavy tail of the distribution, a much higher penalty value is generated.

All of the tables and figures are presented in the following.

[37] Note that in the case where I examine the deviations for the model with the unmodified density function, I compare the simulated values with the empirical values based on the dataset, where the data containing “irrationally” low kilometre values were eliminated. In the case here, where I examine the deviations for the model with the modified density function, I compare the simulated values with the empirical values based on the original dataset. Note that since the empirical values of these two datasets do not differ much, this may not be a major cause for differences in deviations between the simulated and the empirical values of these two models.

Dataset	mz05	mz05
Limit of integrating, $E(X_2)$	∞	60,000km
d	0.15	0.1
a_2	0.2	2
Constant γ_1	-2.76	-2.955
Urban area γ_2	0.34	0.354
Number of people γ_3	-	0.044
β	0.38	0.406
Proportion of drop-outs	9.60%	9.00%
Relative replication error of $P(X_2=0)$	-0.04	0.03
Relative replication error of $E(X_2)$	0.1623	-0.09
Q	0.0327	0.009
$\varepsilon_{E(X_2),p_2}$	$\underset{(0.0001)}{-1.19}$	**-0.68**
$\varepsilon_{E(X_2),p_2,tax\ neutral}$	$\underset{(-)}{-1.09}$	**-0.64**
$\varepsilon_{E(X_2),p_{fuel}}$	$\underset{(0.00005)}{-0.54}$	**0.28**
$\varepsilon_{E(X_2),y}$	$\underset{(0.002)}{1.19}$	**0.78**
$\varepsilon_{E(X_2),k_2}$	$\underset{(0.009)}{-0.18}$	**-0.17**
$\varepsilon_{P(X_2=0),p_2}$	$\underset{(0.0001)}{0.30}$	**0.26**
$\varepsilon_{P(X_2=0),p_2,tax\ neutral}$	$\underset{(-)}{0.21}$	**0.2**
$\varepsilon_{P(X_2=0),p_{fuel}}$	$\underset{(0.00005)}{0.14}$	**0.11**
$\varepsilon_{P(X_2=0),y}$	$\underset{(0.01)}{-1.44}$	**-1.41**
$\varepsilon_{P(X_2=0),k_2}$	$\underset{(0.013)}{1.39}$	**1.31**
$\varepsilon_{P(X_2=0),k_2,tax\ neutral}$	$\underset{(-)}{1.29}$	**1.21**
$\Delta E(X_2)_{rural\to city} / E(X_2)_{rural}$	$\underset{(0.87)}{-32\%}$	**-23%**
$\Delta P(X_2=0)_{rural\to city} / P(X_2=0)_{rural}$	$\underset{(3.55)}{74\%}$	**70%**
$\Delta E(X_2)_{city\to rural} / E(X_2)_{city}$	$\underset{(1.84)}{47\%}$	**27%**
$\Delta P(X_2=0)_{city\to rural} / P(X_2=0)_{city}$	$\underset{(1.16)}{-43\%}$	**-42%**

Table A3.13.1: Changes in estimated parameter values and simulated elasticities when using a modified density function to compute the expectation value.[38]

[38] These results are based on the complete dataset.

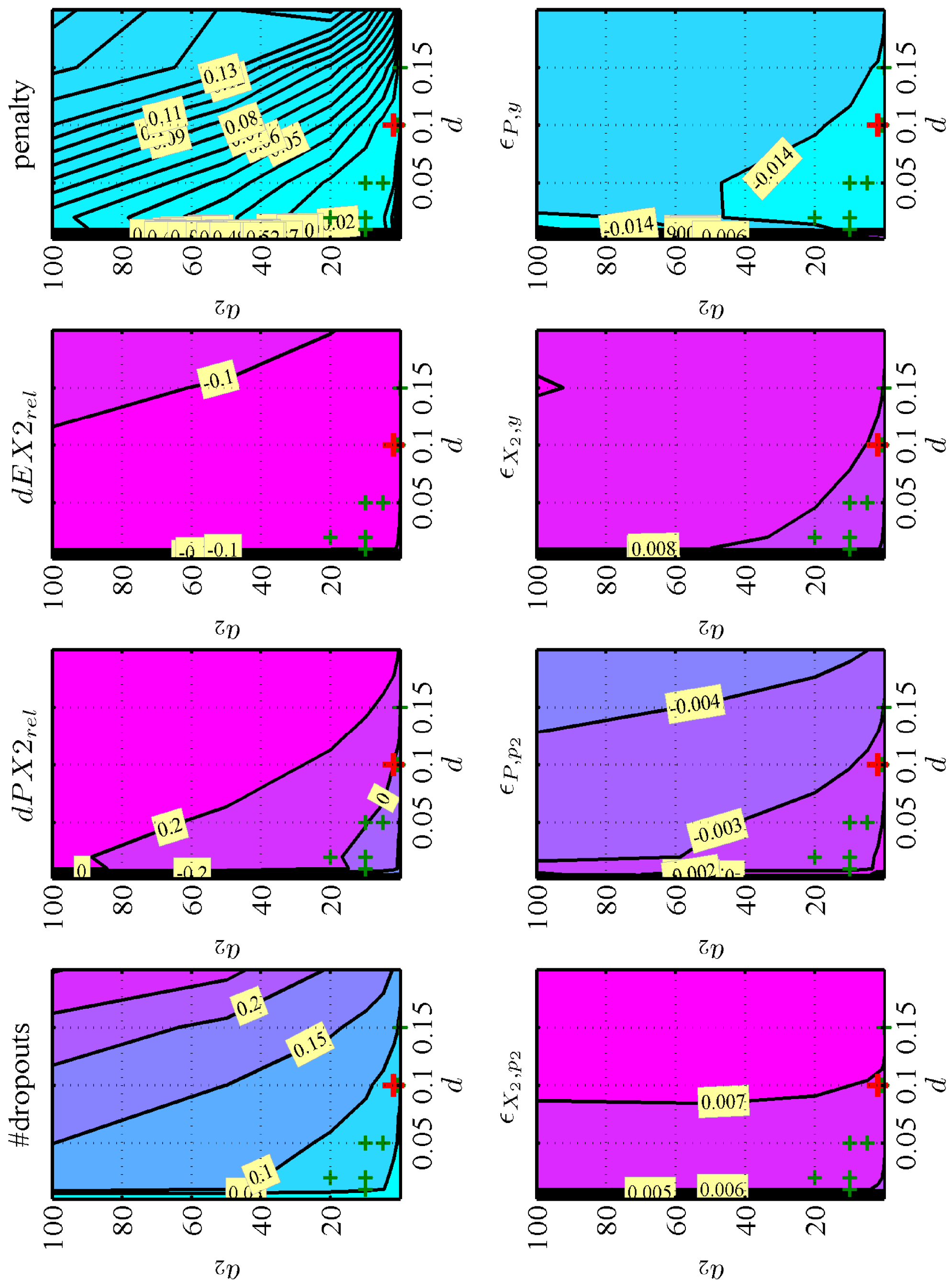


Figure A3.13.1: Elasticities of parameter combinations d and a_2 that yield low penalty values for the case that households adapt the fuel efficiency of the car they choose, mz05.

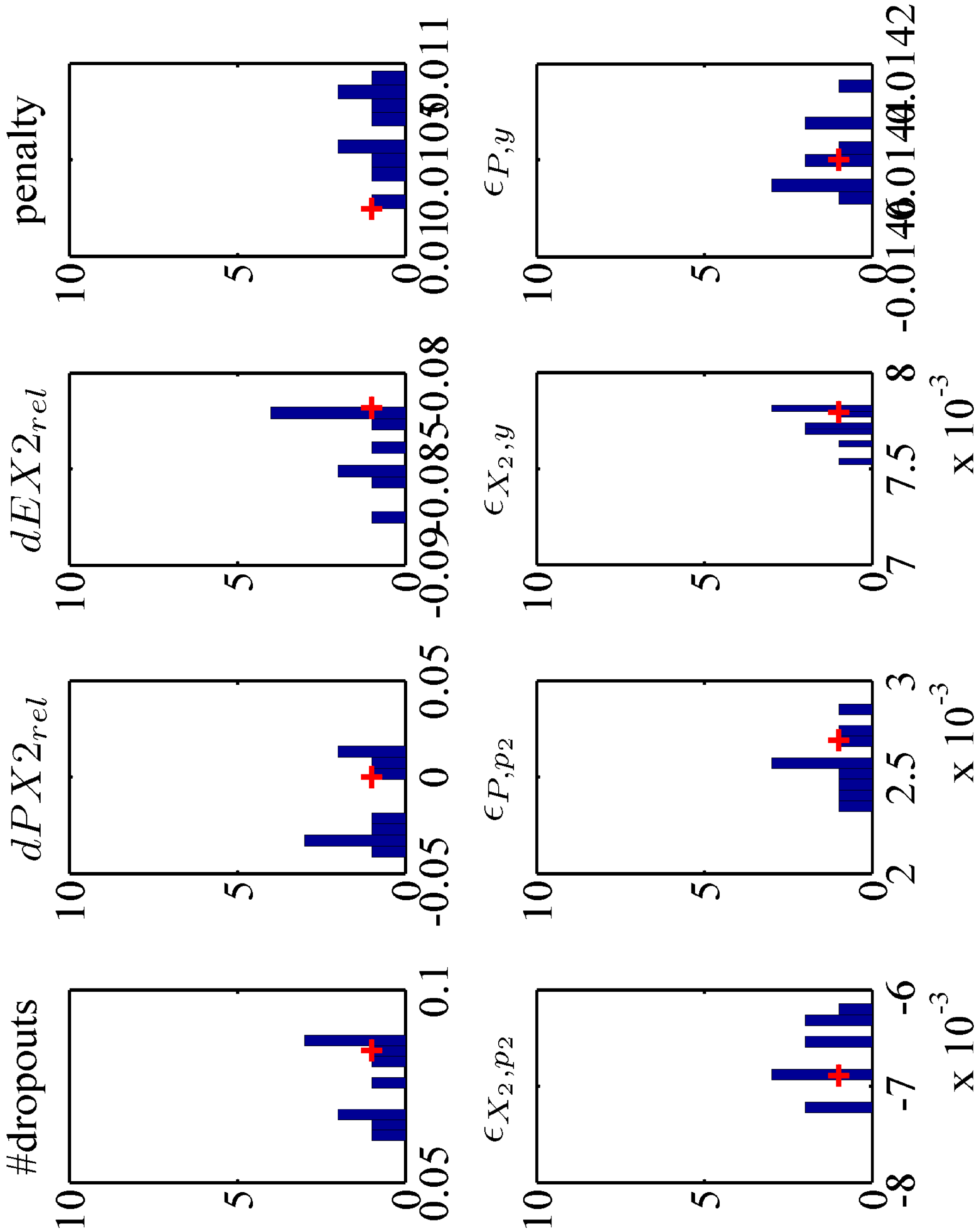


Figure A3.13.2: Elasticities of parameter combinations d and a_2 that yield low penalty values for the case that households adapt the fuel efficiency of the car they choose, mz05.

y	Lo-ca-tion	Nobs	Share of $x_2 = 0$	$P(X_2 = 0)$	Δ	$mean(x_2)$	$E(X_2 = 0)$	Δ	$X_2(\varsigma_c)$	ς_c
18,000	0	576	0.734	0.874	0.140	2,196	846	-1,350	4,169	0.788
18,000	1	224	0.594	0.739	0.145	3,186	1,823	-1,363	4,169	0.422
36,000	0	3,038	0.461	0.511	0.050	5,718	4,805	-913	4,970	0.018
36,000	1	914	0.259	0.298	0.039	8,743	8,025	-719	4,970	-0.348
60,000	0	4,382	0.210	0.230	0.020	10,877	10,299	-578	5,264	-0.492
60,000	1	1,462	0.058	0.108	0.050	15,952	14,867	-1,084	5,264	-0.857
84,000	0	3,380	0.109	0.119	0.010	15,001	14,681	-320	5,386	-0.815
84,000	1	1,051	0.028	0.052	0.024	20,536	19,584	-952	5,386	-1.180
108,000	0	2,166	0.092	0.070	-0.022	17,045	18,042	998	5,453	-1.052
108,000	1	572	0.012	0.030	0.017	23,830	22,412	-1,417	5,453	-1.417
132,000	0	1,177	0.067	0.045	-0.022	19,095	20,493	1,397	5,495	-1.239
132,000	1	276	0.014	0.019	0.004	26,425	23,647	-2,777	5,495	-1.605
156,000	0	617	0.039	0.031	-0.008	21,672	22,128	457	5,524	-1.394
156,000	1	117	0.009	0.013	0.004	29,953	23,658	-6,295	5,524	-1.760
180,000	0	295	0.047	0.023	-0.025	20,779	23,055	2,276	5,546	-1.526
180,000	1	66	0.000	0.009	0.009	32,125	22,822	-9,303	5,546	-1.892
228,000	0	475	0.040	0.014	-0.026	23,841	23,248	-593	5,575	-1.744
228,000	1	82	0.024	0.006	-0.019	30,635	19,850	-10,785	5,575	-2.109

Table A3.13.2: Empirical and simulated values of driving distance and proportions of carless households.

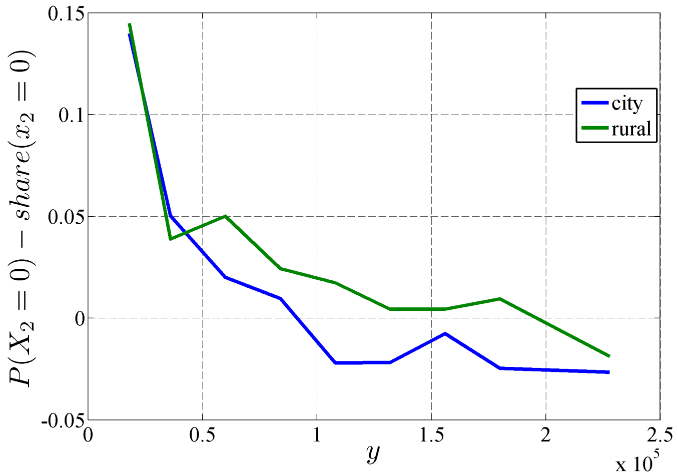


Figure A3.13.3: Deviation of simulated proportion of carless households from the empirical value.

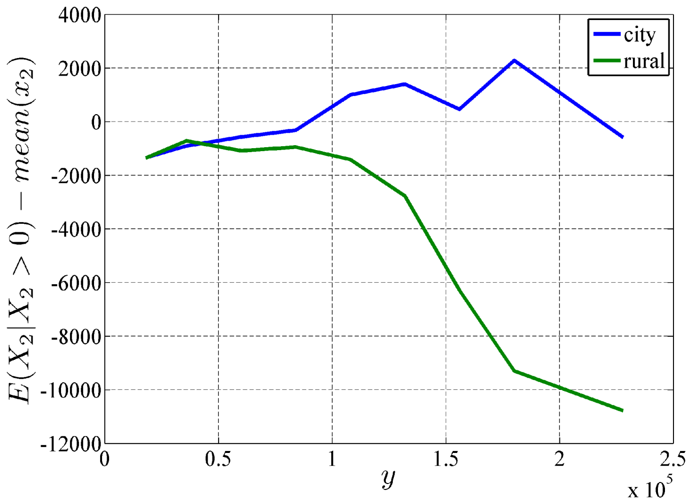


Figure A3.13.4: Deviation of simulated expatiation value of driving demand from the empirical value.

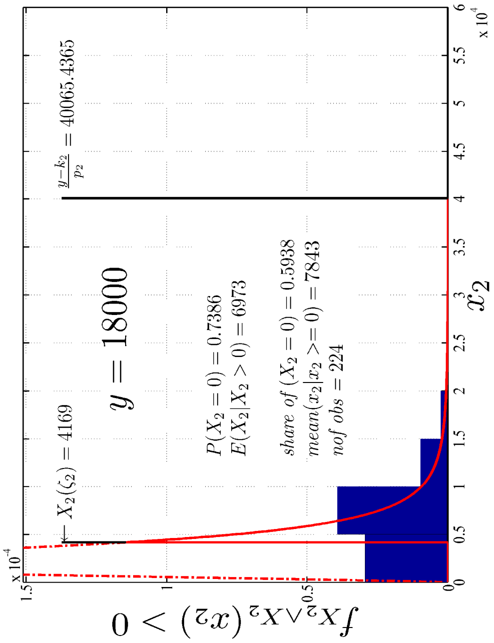


Figure A3.13.5: Histogram of households in rural areas with an income of CHF 18,000.

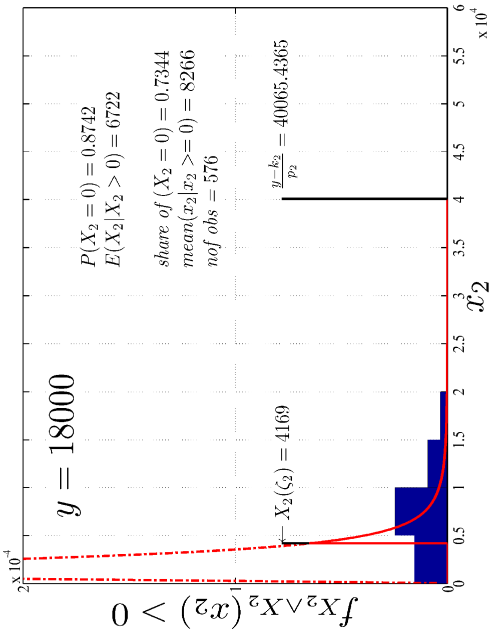


Figure A3.13.6: Histogram of households in urban areas with an income of CHF 18,000.

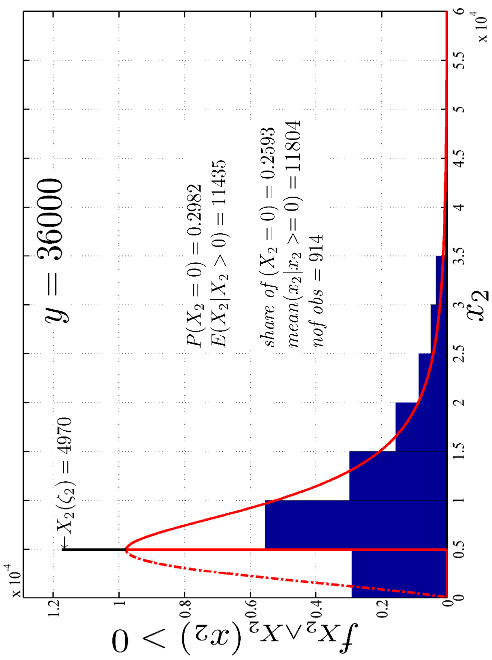


Figure A3.13.7: Histogram of households in rural areas with an income of CHF 36,000.

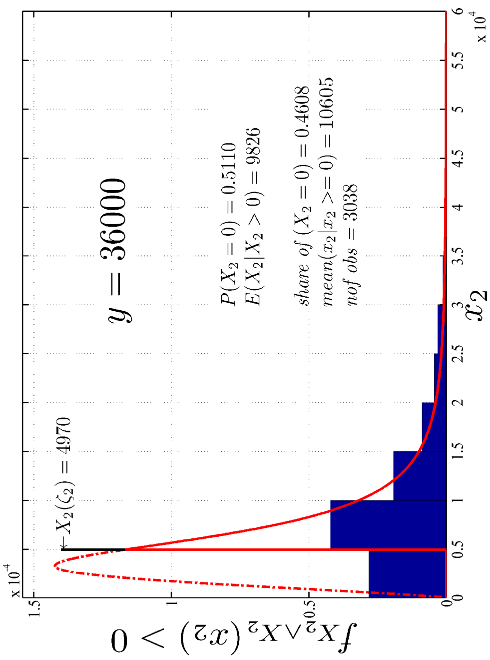


Figure A3.13.8: Histogram of households in urban areas with an income of CHF 36,000.

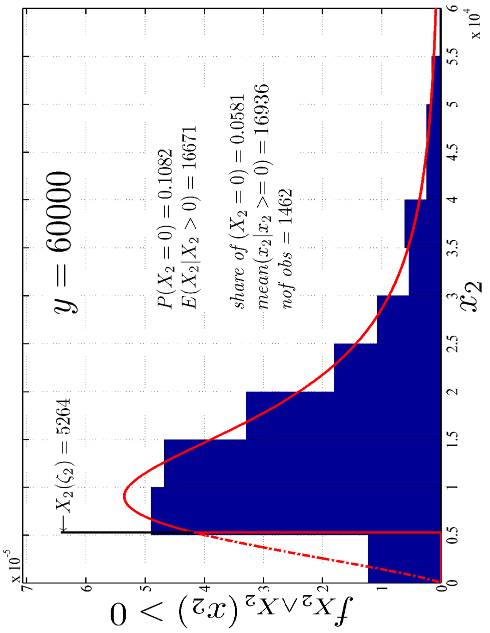


Figure A3.13.9: Histogram of households in rural areas with an income of CHF 60,000.

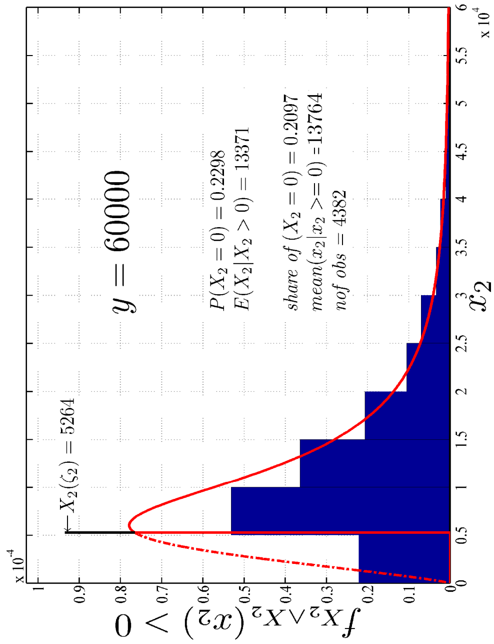


Figure A3.13.10: Histogram of households in urban areas with an income of CHF 60,000.

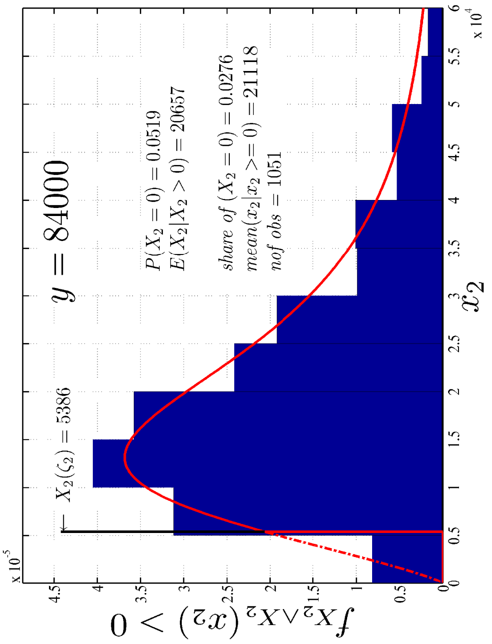


Figure A3.13.11: Histogram of households in rural areas with an income of CHF 84,000.

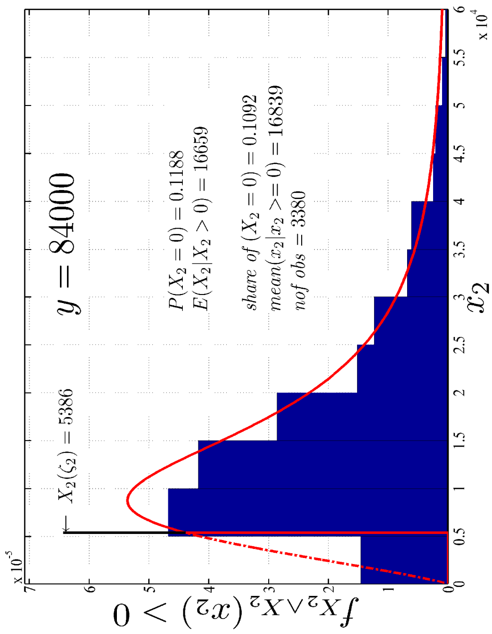


Figure A3.13.12: Histogram of households in urban areas with an income of CHF 84,000.

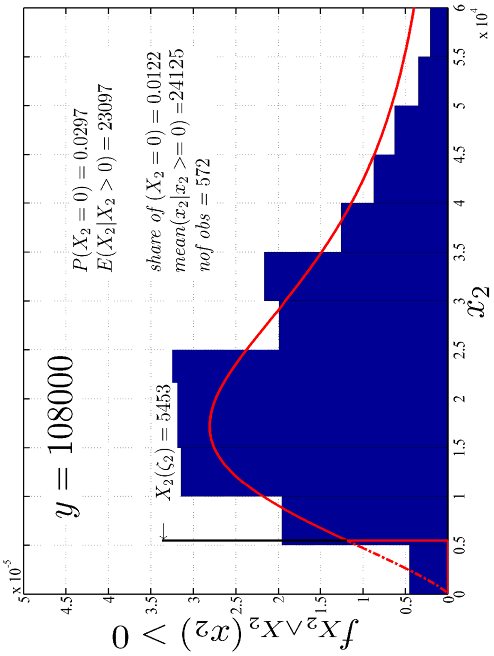


Figure A3.13.13: Histogram of households in rural areas with an income of CHF 108,000.

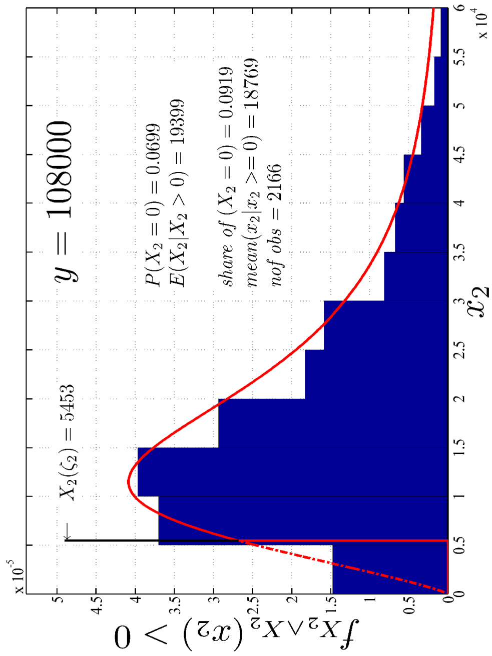


Figure A3.13.14: Histogram of households in urban areas with an income of CHF 108,000.

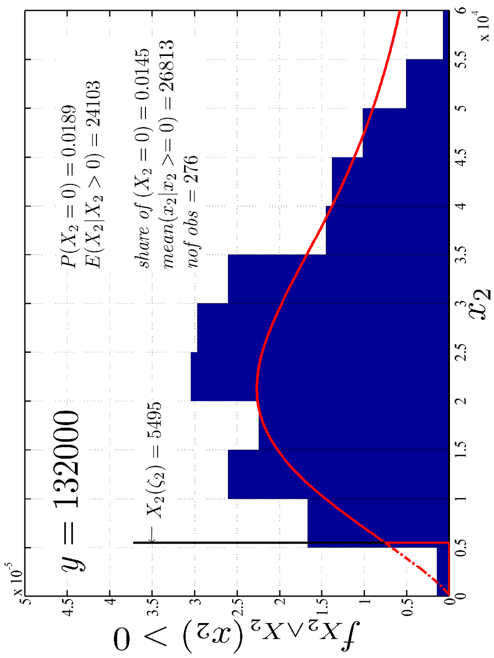


Figure A3.13.15: Histogram of households in rural areas with an income of CHF 132,000.

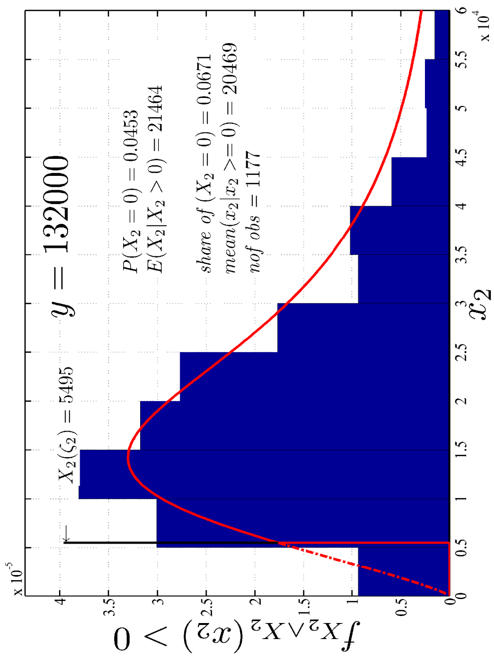


Figure A3.13.16: Histogram of households in urban areas with an income of CHF 132,000.

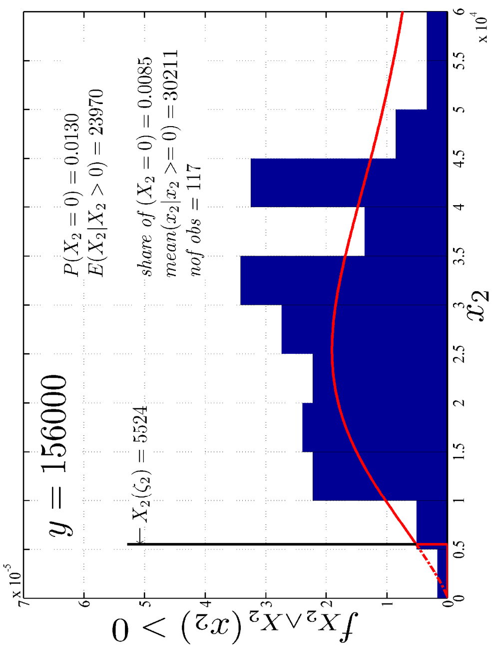


Figure A3.13.17: Histogram of households in rural areas with an income of CHF 156,000.

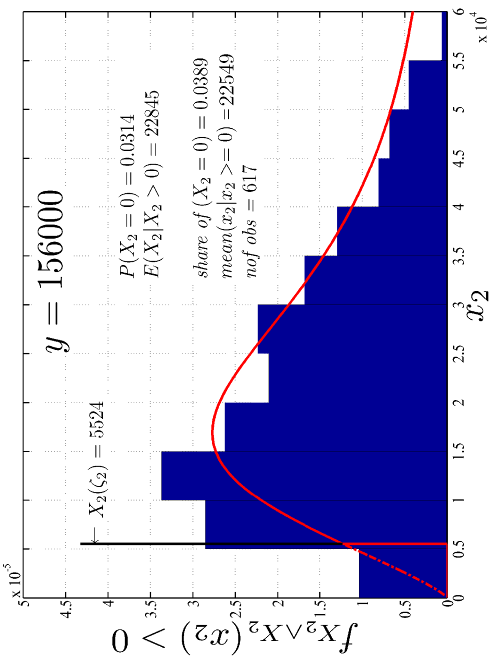


Figure A3.13.18: Histogram of households in urban areas with an income of CHF 156,000.

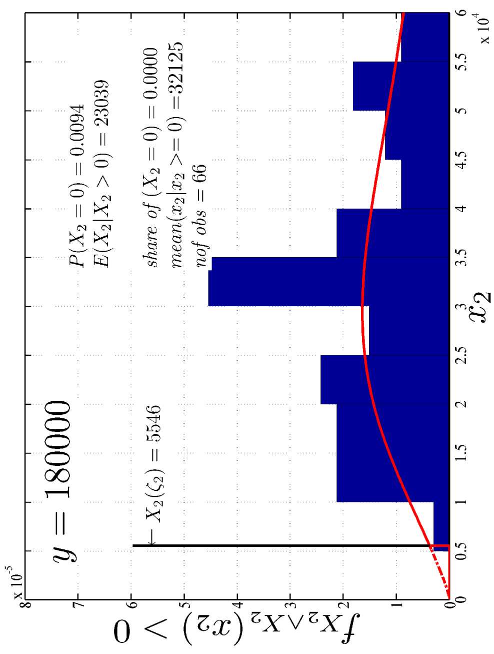


Figure A3.13.19: Histogram of households in rural areas with an income of CHF 180,000.

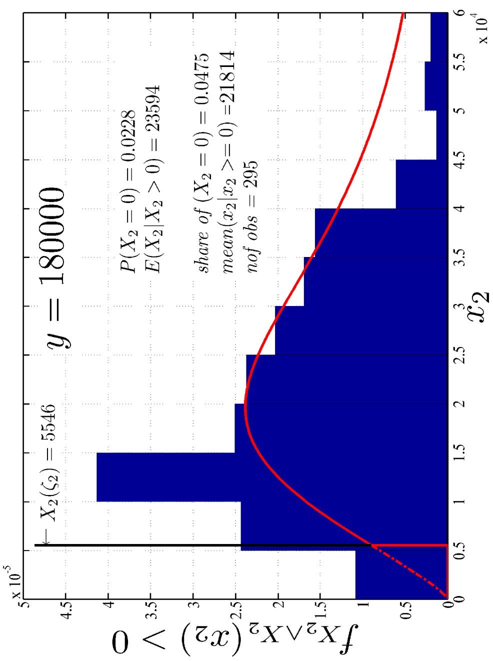


Figure A3.13.20: Histogram of households in urban areas with an income of CHF 180,000.

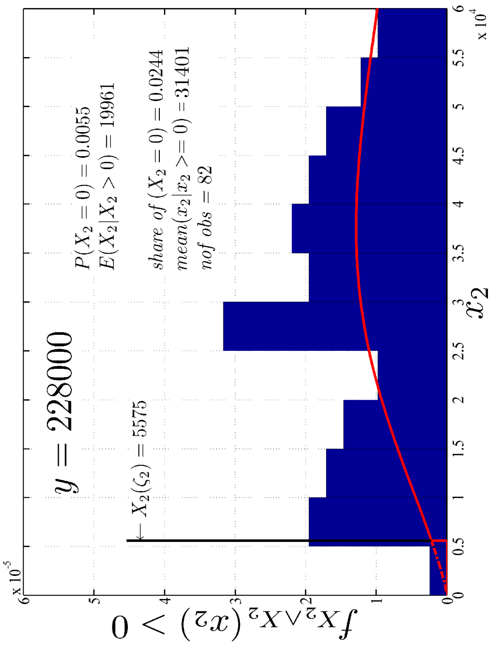


Figure A3.13.21: Histogram of households in rural areas with an income of CHF 228,000.

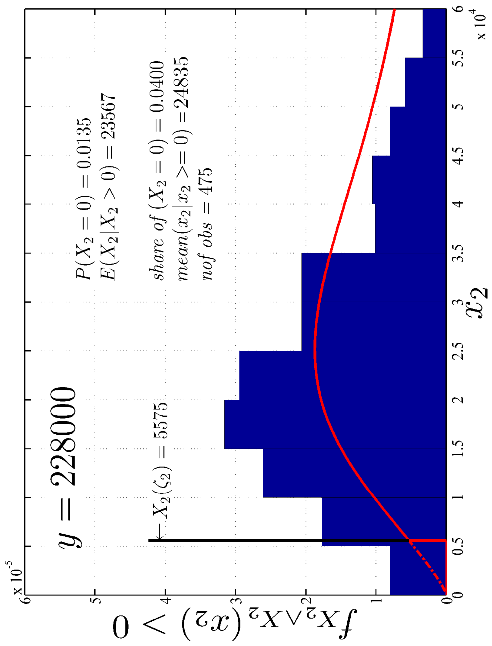


Figure A3.13.22: Histogram of households in urban areas with an income of CHF 228,000.

A 3.14 Estimation results based on a stated preference dataset: comparison of results based on different models

In this subchapter, I will compare the elasticities yielded by the OLS, the Tobit, the Probit and the MDCEV model. The purpose is to examine whether the results yielded by these models differ much from one another. I will always use the dataset from Axhausen and Erath (2010), which I call “Erath”. I chose this dataset because it is the only dataset where the variance of fuel price is as large as, in the case of the OLS, the Tobit and the Probit model, the estimated parameters relating to the fuel price yield significant values. Note that these three models cannot provide all of the elasticities that can be computed using the MDCEV model. The dataset has already been described in Subchapter 1.4; the OLS, the Tobit and the Probit model were presented in Subchapter 1.5 and Appendix A2.1. First, I shall present and discuss the results when applying the MDCEV model including fixed costs on the data. I will then compare the resulting elasticities with the elasticities yielded based on OLS, Tobit or Probit models to see whether the results yielded by the MDCEV model approximate those computed by “traditional” models. Further, the results of the MDCEV using the Erath dataset can be considered as more “trustworthy” than those yielded based on other datasets, since in the case of the Erath dataset, the variation of the ML function is no longer only caused by variables other than the driving costs.

Results of the MDCEV model

The results were computed using the same procedure described in the section entitled “Estimation routine” in Subchapter 3.2. First, I shall present the results based on data that ignore the fact that households own more fuel-efficient cars when fuel prices rise. For comparison, I then show the results based on data that capture the effect that households buy more fuel-efficient cars when income rises. I also compute the results applying the model that uses a modified function for computing the expectation value, as discussed in Appendix A3.13. Finally, I shall compare all of the results with those based on dataset Mz05.

The following table shows the values of a set of measures for different combinations of d and a_2. These results correspond to a model where the place of residence (urban versus rural area), the number of household members and the price of public transport explain the relative preference for car driving that is captured by parameter m, see Formula (3.2.4).

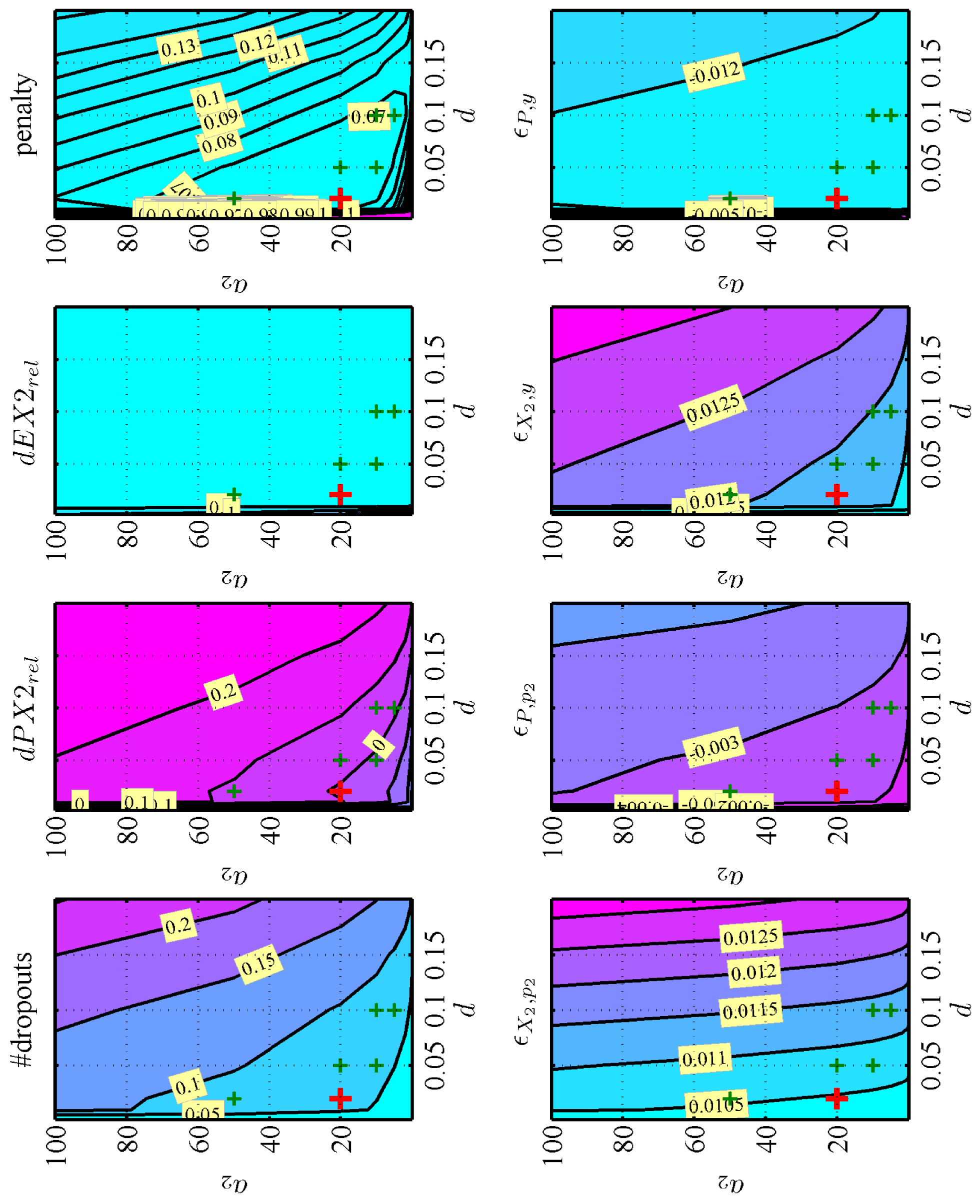


Figure A3.14.1: Simulated elasticities and penalty values for parameters a_2 and d.

The above results were computed based on the complete dataset. The diagrams above show that the penalty function is again convex in parameters a_2 and d. For this reason, the grid points indicated by small green crosses that correspond to penalty values Q less than 10% above the grid point corresponding to the minimal penalty value indicated by a large red cross, are also in the vicinity of that latter point. Also, in this case the simulated elasticities for these points do not differ much, as shown in Figure A3.14.2. Hence, also in this case the results based on the Erath dataset seem to be quite robust with respect to the choice of different penalty functions.

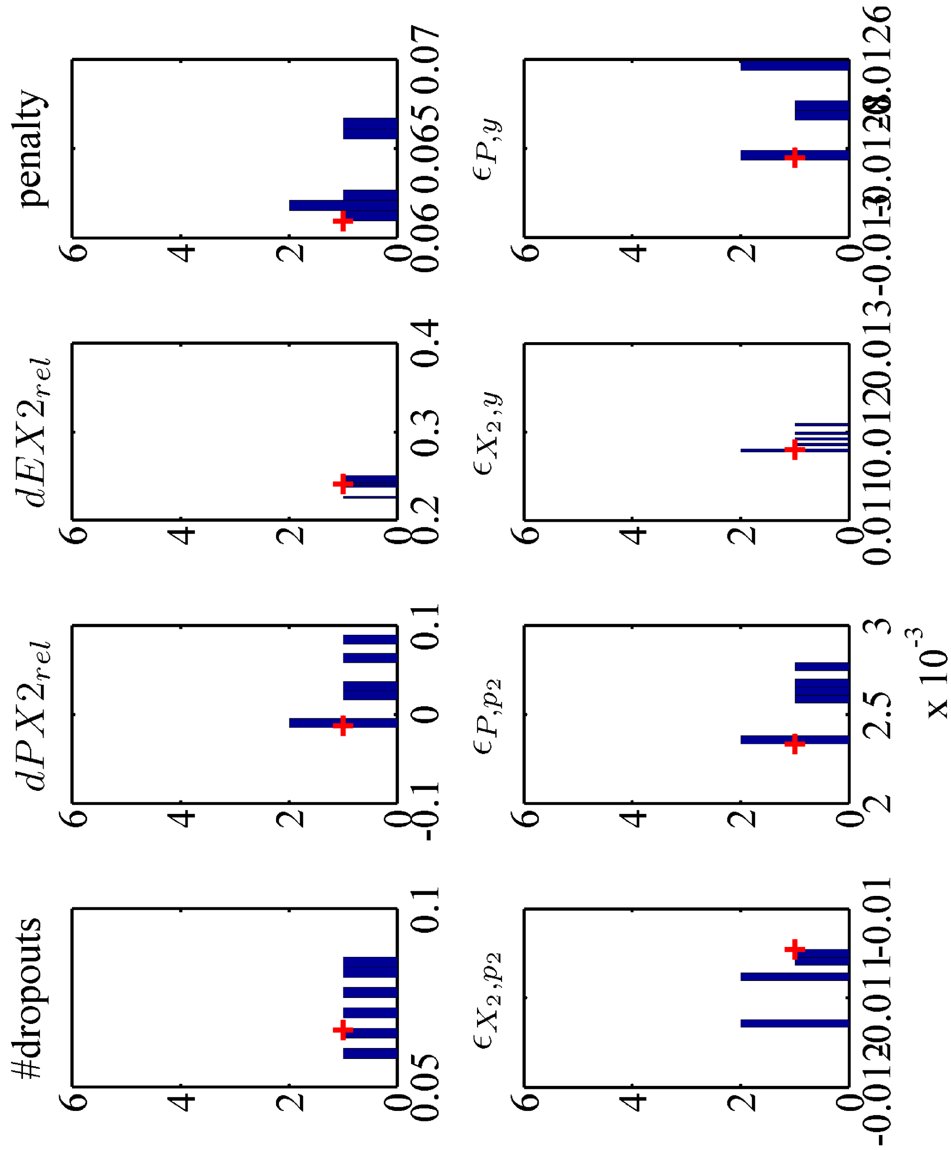


Figure A3.14.2: Elasticities of parameter combinations *d* and a_2 that yield low penalty values.

So far, I have neglected the fact that households can switch to more fuel-efficient cars when fuel prices rise. In the survey by Erath and Axhausen (2009), households could choose from different car models when asked how they would behave in the event of fuel prices changing. Since the dataset contains the information on the fuel efficiency of the car chosen by households, I was able to compute that households choose a car that consumes 0.026 litre / 100 km less when fuel prices increase by CHF 0.1 / litre using a regression model.

Dataset	**Erath**	**Erath**	**Erath**
	OLS, panel	OLS, pooled	OLS, pooled
(Constant) β_1	6.2980 (0.2508)	7.249 (0.1094)	5.393 (0.0907)
Income in CHF 1000 per month) β_2	0.1306 (0.0254)	--	0.1332 (0.0107)
(Fuel price) β_2	-0.2932 (0.0137)	--	--
(Rural area) β_3	-0.4550 (0.1896)	-0.2653 (0.0032)	--
(Annual km in 1000) β_4	-0.0136 (0.0059)	--	--
(Number of people in household) β_5	0.1831 (0.0843)	--	--
σ	1.75	2.00	1.97
σ_η	0.81	--	--
R^2 / pseudo R^2	0.1079	0.0290	0.0620
$\varepsilon_{fuel\ eff,\, p_{fuel}} \mid \bar{x}$	-0.1463 (0.0071)	-0.1324 (0.0159)	--
$\varepsilon_{fuel\ eff,\, p_{fuel}} \mid \bar{x},\ p_{fuel} = 1.50\,\mathrm{sFr}$	-0.0637 (0.0029)	-0.0581 (0.0065)	--
$\varepsilon_{fuel\ eff,\, p_{fuel}} \mid \bar{x},\ p_{fuel} = \overline{p_{fuel}} = 3.22\ \mathrm{sFr}$	--	-0.1324 (0.0159)	--
$\varepsilon_{fuel\ eff,\, income} \mid \bar{x}$	0.1545 (0.0301)	--	0.1576 (0.0127)

Note 1: The values in parentheses “(..)” denote standard deviations.

Table A3.14.1: Elasticities of the fuel efficiency with respect to income and fuel price.

With these results, I was able to compute the actual marginal costs of the car households would choose.[39] Using this modified marginal costs for the model, I expect that the model will yield higher elasticities with respect to marginal costs of driving. This is because, given the variation of the observed annual mileage, driving costs vary less because of households' tendency to switch to more fuel-efficient cars. Note that I ignore that cars with a higher fuel efficiency are either more expensive what would be reflected by higher fixed costs or they would provide less utility that could be captured by a change in parameter *m*. Again, I show the same diagrams as in the case where the shift towards more fuel-efficient cars is ignored in the data.

[39] Since adding additional control variables and capturing the panel structure of the data does not change the parameters and the elasticities of interest, I used the results of the simple univariate models: $fuel\ eff = 7.249 - 0.26531 \cdot p_{fuel}$ and $fuel\ eff = 5.393 - 0.1332 \cdot p_{fuel}$.

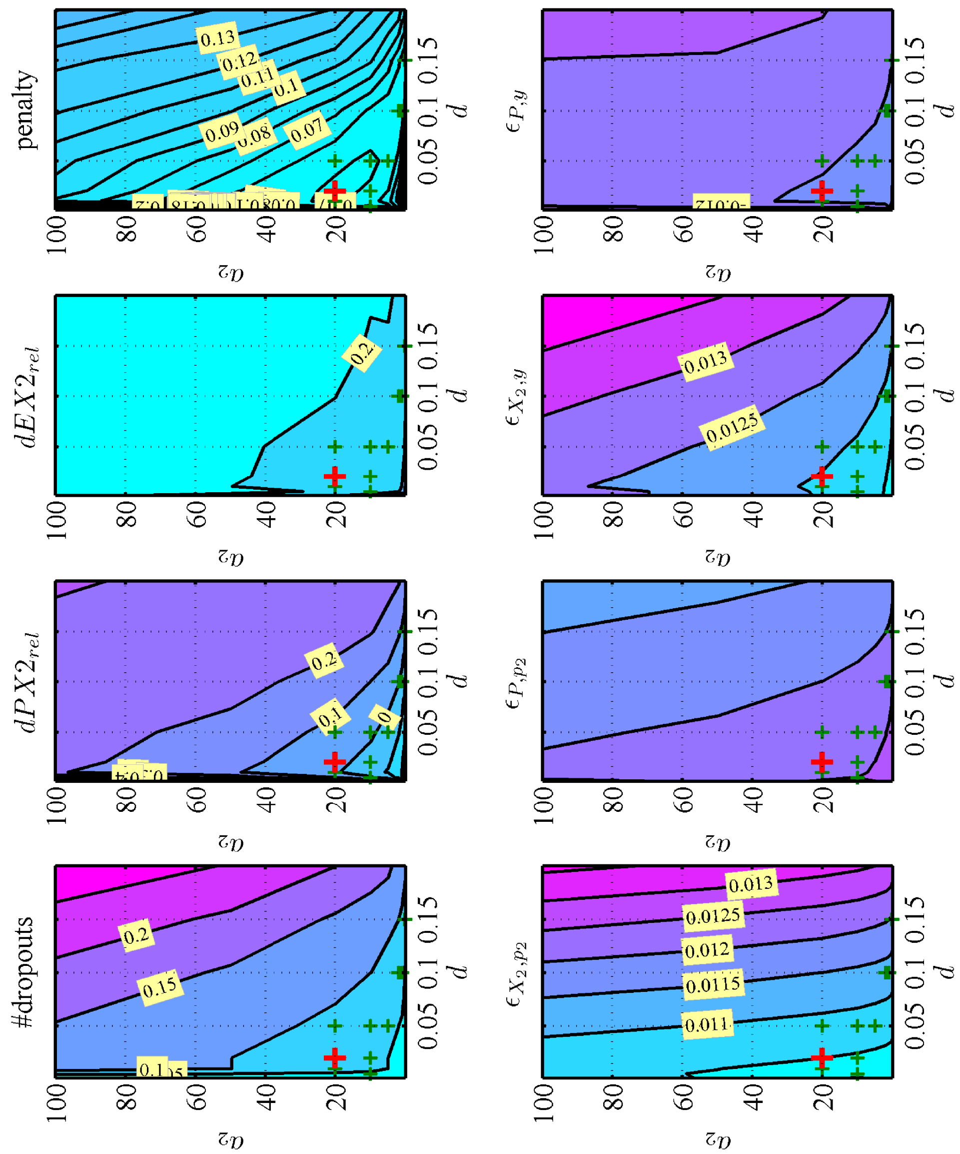


Figure A3.14.3: Simulated elasticities and penalty values for parameters a_2 and d for the case that households switch to more fuel-efficient cars.

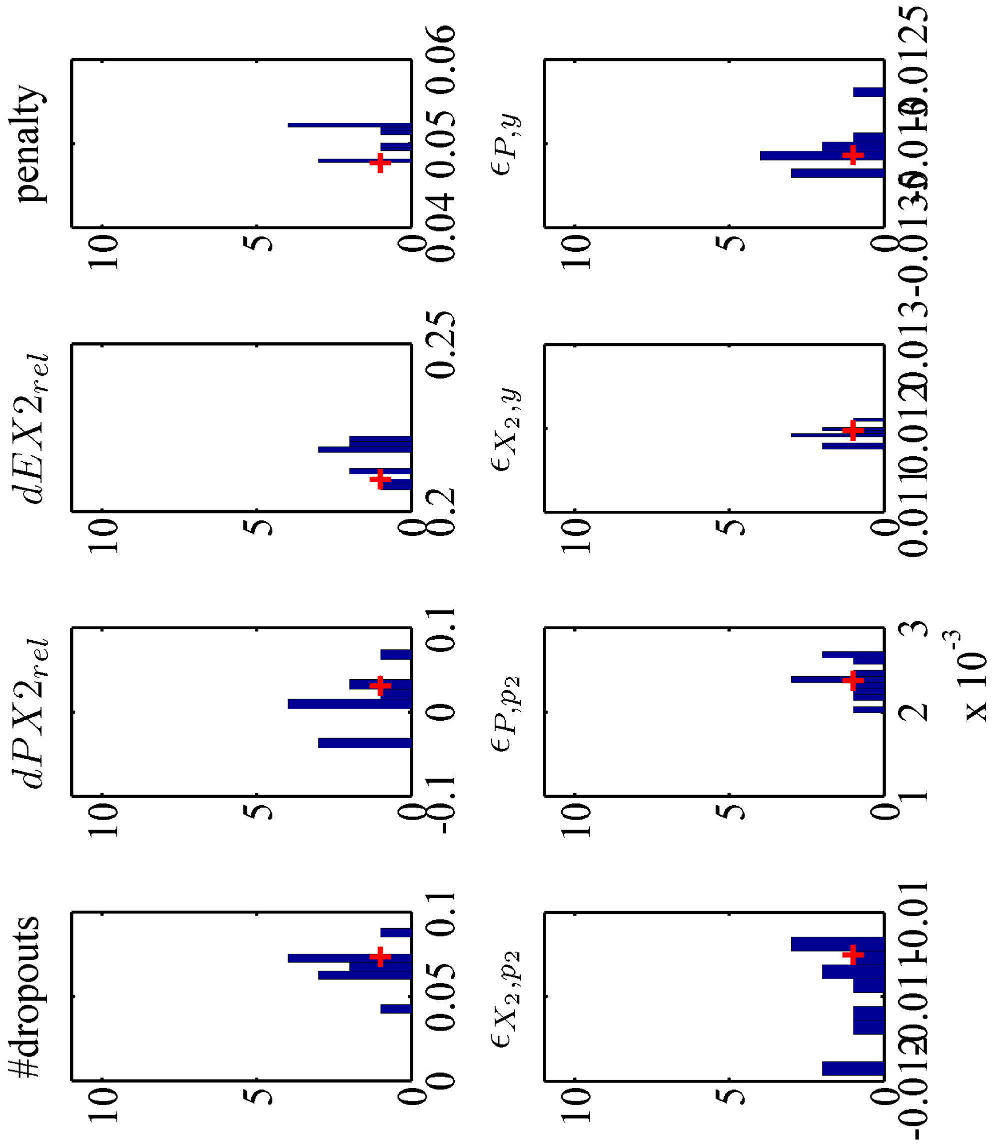


Figure A3.14.4: Elasticities of parameter combinations *d* and a_2 that yield low penalty values for the case that households switch to more fuel-efficient cars.

Also in this case, the results seem to be quite robust with respect to the choice of different penalty functions. The resulting elasticities differ only slightly from those yielded when using the data ignoring the shift towards more fuel-efficient cars when fuel prices rise.

For the model that uses a modified function for computing the expectation value as discussed in Appendix A3.13, I present the two diagrams differently to how the two last models were presented. I will list only the final results in the following table, which shows the results of all MDCEV models based on the Erath dataset plus the results based on the Mz05 dataset for comparison.

Dataset	**Erath**	**Erath**[(1)]	**Erath**[(1)]	**mz05**
Limit of integrating, $E(X_2)$	∞	∞	60000km	60000km
d, a_2	0.01, 20	0.02, 20	0.005, 10	0.1, 2
(constant) γ_1	-3.653	-3.865	-3.818	-2.955
(household location in rural area) γ_2	0.045	0.031	0.07	0.354
(number of people in household) γ_3	0.061	0.048	0.079	0.044
(price public transportation) γ_4	0.532	0.683	0.567	--
β	0.467	0.437	0.508	0.406
share of dropouts	0.065	0.081	0.029	0.09
relative replication error of $P(X_2=0)$	-0.034	0.033	-0.025	0.031
relative replication error of $E(X_2)$	0.19	0.169	-0.03	-0.09
Q	0.0371	0.0329	0.0011	0.0071
$\varepsilon_{E(X_2),p_2}$	-1.036	-1.05	-0.685	-0.679
$\varepsilon_{E(X_2),p_2,tax\ neutral}$	-0.944	-0.964	-0.637	-0.638
$\varepsilon_{E(X_2),p_{fuel}}$ [(2)]	-0.469	-0.387 {-0.411}	-0.252 {-0.268}	-0.282
$\varepsilon_{x_{fuel},p_{fuel}}$	--	-0.45	-0.310	--
$\varepsilon_{E(X_2),y}$	1.174	1.191	0.822 {0.829}	0.781
$\varepsilon_{x_{fuel},y}$	--	1.393	0.978	--
$\varepsilon_{fuel\,eff,y}$			-0.058	
$\varepsilon_{fuel\,eff,p_{fuel}}$			0.156	
$\varepsilon_{E(X_2),k_2}$	-0.157	0.173	-0.148	-0.175
$\varepsilon_{P(X_2=0),p_2}$	0.240	0.241	0.206	0.264
$\varepsilon_{P(X_2=0),p_{fuel}}$	0.109	0.0893 {0.109}	0.0748 {0.093}	0.108
$\varepsilon_{P(X_2=0),p_2,tax\ neutral}$	0.156	0.168	0.141	0.197
$\varepsilon_{P(X_2=0),p_{fuel}}$	0.11	0.112	0.095	0.11
$\varepsilon_{P(X_2=0),y}$	-1.35	-1.334	-1.3	-1.41
$\varepsilon_{P(X_2=0),k_2}$	1.112	1.107	1.09	1.307
$\varepsilon_{P(X_2=0),k_2,tax\ neutral}$	1.023	1.022	1.003	1.21

[(1)]: In these cases, I took into account that households switch to more fuel-efficient cars when fuel prices increase by 0.026 litre / 100 km when fuel prices increase by CHF 0.1 / litre.

[(2)]: The values in braces "{..}" denote the values of the elasticity of driving demand for the computation method that ignores the fact that households may buy cars that are more or less fuel-efficient if fuel prices or incomes change. The results therefore correspond to the values computed by the datasets that do not capture this effect, such as dataset Mz05. All values are based on a fuel price of CHF 1.50 / litre.

Table A3.14.2: Elasticities of parameter combinations *d* and a_2 that yield low penalty values for the case that households switch to more fuel-efficient cars.

Three main conclusions can be gained from the results listed in the table: first, the elasticities and parameters I obtain using dataset Mz05 are very similar to those using the dataset Erath. I conclude from this that these two datasets are quite comparable, despite the fact that the Erath dataset is a stated preference dataset, with the exception of one of the six observations per household that reflects the actual behaviour of the household. Therefore, I also Erath also expect that the results I will compute be the OLS, Tobit and Probit models would be very similar to those I would get using dataset Mz05. Second, the effect of capturing the fact that households switch to more fuel-efficient cars when fuel prices rise has virtually no effect on the estimated parameters and the elasticities, apart from the elasticities with respect to fuel price. However, this substitution effect strongly affects the elasticities with respect to fuel prices, since the marginal costs of driving increase less when fuel prices rise than when households have no option to switch to more fuel-efficient car types. This implies that the elasticities with respect to fuel demand are about 22% smaller than in the case when households may not switch car types.[40] As in the case when using dataset Mz05, all elasticities with respect to driving demand are smaller when the modified density function for computing the expected driving demand is used. Since I consider the model based on this modified demand function as the most realistic it can be followed that the model yields rather plausible values for the elasticities, since the value -0.252 of the elasticity of driving demand with respect to the fuel price demand is rather close to -0.202 that was found by Baranzini et al. (2009) for the own price elasticity of fuel demand. Note that if I consider also household behaviour with respect to car fuel economy to compute the own price elasticity of fuel demand, the value increases in absolute terms, namely -0.310.[41] Note that all elasticities referred to a fuel price of CHF 1.50 / litre, which was about the market price at the time of the survey. If I compute the fuel price elasticities of driving demand at the fuel price of CHF 3.22 / litre which is the sample

[40] In the event that households can switch to more fuel-efficient cars when fuel prices rise, the fuel price elasticities can be computed as follows: first, note that $\varepsilon_{E(X_2),p_2}$ is computed by $\varepsilon_{E(X_2),p_{fuel}} = \partial E(X_2)/\partial p_2 \cdot \partial p_2/\partial p_{fuel} \cdot p_{fuel}/p_2 \cdot p_2/E(X_2) = \partial p_2/\partial p_{fuel} \cdot p_{fuel}/p_2 \cdot \varepsilon_{E(X_2),p_2} = \varepsilon_{p_2,p_{fuel}} \cdot \varepsilon_{E(X_2),p_2}$. Since $p_2 = 0.1601 + (0.07249 - 0.00265 \cdot p_{fuel}) \cdot p_{fuel}$, where $fuel\,cons\,per\,km = 0.07249 - 0.00265 \cdot p_{fuel}$ and values 0.07249 and 0.00265 are yielded by regressing the fuel price on fuel efficiency, it follows that $\partial p_2/\partial p_{fuel} = 0.07249 - 2 \cdot 0.00265 \cdot p_{fuel}$. For the variables, I plug in the actual market price at the time of data collection, namely $p_{fuel} = \text{CHF } 1.50 \text{ / litre}$. This implies $\partial p_2/\partial p_{fuel} = 0.0645$, $p_{fuel}/p_2 = 1.50/0.361 = 5.706$ and therefore $\varepsilon_{p_2,p_{fuel}} = 5.706 \cdot 0.0645 = 0.3683$. Using formula $\varepsilon_{E(X_2),p_{fuel}} = \varepsilon_{p_2,p_{fuel}} \cdot \varepsilon_{E(X_2),p_2}$, the final results can be computed: $\varepsilon_{E(X_2),p_{fuel}} = 0.3683 \cdot (\text{-}0.685) = \text{-}0.3683$. Note that the same formula can be applied to compute the fuel price elasticities of the probability of being carless, $\varepsilon_{P(X_2=0),p_{fuel}} = \varepsilon_{p_{fuel},p_2} \cdot \varepsilon_{P(X_2=0),p_2} = 0.3683 \cdot 0.206 = 0.0759$. Note that $\varepsilon_{p_{fuel},p_2} = 0.4486$ in the event that households may not substitute their car. This value is 22% higher than when households may switch car type when fuel prices rise. If all values are computed at the mean of the fuel prices in the dataset, namely for $p_{fuel} = 3.22$, then the elasticities yield: $\varepsilon_{E(X_2),p_{fuel}} = \text{-}0.3341\{\text{-}0.3857\}$ and $\varepsilon_{x_{fuel},p_{fuel}} = \text{-}0.4679$.

[41] I computed this value as follows: assuming that all household change the car's fuel economy by the same percentage when fuel prices rises by one percent, the following relation does not only hold at the household but also at the aggregate level: $x_{fuel} = e \cdot E(X_2)$, where e is car fuel demand per kilometre. This implies . Evaluating $\varepsilon_{x_{fuel},p_{fuel}} = \varepsilon_{e,p_{fuel}} + \varepsilon_{E(X_2),p_{fuel}}$, where $\varepsilon_{e,p_{fuel}} = \partial e/\partial p_{fuel} \cdot p_{fuel}/e$ yields $\varepsilon_{e,p_{fuel}} = \text{-}0.00265 \cdot 1.5/0.068515 = \text{-}0.058$ and thus $\varepsilon_{x_{fuel},p_{fuel}} = \text{-}0.058 - 0.252 = \text{-}0.310$.

mean, the following values yield: $\varepsilon_{E(X_2),p_{fuel}} = \text{-}0.3341\{\text{-}0.3857\}$, $\varepsilon_{x_{fuel},p_{fuel}} = \text{-}0.4679$ and $\varepsilon_{fuel\ eff,p_{fuel}} = 0.1337$.[42] Values $\varepsilon_{E(X_2),p_{fuel}}$ and $\varepsilon_{x_{fuel},p_{fuel}}$ are now almost identical to the averages of international studies, see Table 1.3.2, and particularly the values generated by Brons et al. (2006), see Table 1.3.3. The value of $\varepsilon_{fuel\ eff,p_{fuel}}$ is still much smaller than those reported by international studies, see Tables 1.3.2, 1.3.3 and 1.3.4.

Note that changes in income also affect the marginal costs of driving, since households use less fuel-efficient cars when incomes rise. In other words, if income increases, the marginal costs of driving increases due to the decrease in car fuel efficiency. This leads to a slight reduction in the increase in driving demand namely from 0.829 to 0.822.[43] If the effect of income on car fuel efficiency is considered when computing the income elasticities of fuel demand, the elasticity has the value 1.056.[44] This value is almost identical to the average value found in international studies by Goodwin et al. (2004) and Graham and Glaister (2005), see Table 1.3.2.

The fuel price and income elasticities of car fuel efficiency are -0.058 ($p_{fuel} = 1.5$) and 0.1324 ($p_{fuel} = 3.22$), respectively.[45] The value I found for the elasticity with respect to income is about 40% lower than those given in international studies. In the case of the elasticity with respect to fuel price it is even much lower, since the average given in international studies range from -0.4 to -0.23, which is much higher. If I computed the fuel price elasticity at the average fuel price found in the dataset of CHF 3.22 / litre instead of CHF 1.50 / litre, it would be -0.132, which is still far below the value found in other countries. The reason for this could be similar to in the case of driving distance: in this survey

---

[42] Note that the value of $\varepsilon_{fuel\ eff,p_{fuel}} = 0.1337$ is identical to that which results when regressing $\log(p_{fuel})$ on $\log(fuel\ eff)$, which yields an elasticity of -0.127.

[43] The income elasticity of driving demand that includes the effect of changes in car fuel efficiency can be computed as follows: $\dfrac{\partial x_2(y,p_2(y))}{\partial y} = \dfrac{\partial x_2(y,p_2(y))}{\partial y} + \dfrac{\partial x_2(y,p_2(y))}{\partial p_2(y)} \cdot \dfrac{\partial p_2(y)}{\partial y}$.

It can be shown that $\varepsilon_{x_2(y,p_2(y)),y} = \dfrac{\partial x_2(y,p_2(y))}{\partial y} \cdot \dfrac{y}{x_2} = \varepsilon_{E(X_2),p_2} \cdot \varepsilon_{p_2,y} + \varepsilon_{E(X_2),y}$, where $\varepsilon_{p_2,y}$ denotes the elasticity of driving costs with respect to income, $\varepsilon_{p_2,y} = \beta_2 \cdot p_{fuel} \cdot y / p_2$, where $dp_2/dy = d\alpha_1 + (\beta_1 + \beta_2 \cdot y) \cdot p_{fuel}/dy = \beta_2 \cdot p_{fuel}$. Note that $\alpha_1 = 0.161$ and parameters β_1 and β_2 can be found in Table A3.14.1. Plugging in the values mean income $\bar{y} = 7{,}496$ and for the fuel price $p_{fuel} = 1.5$, $\varepsilon_{p_2,y}$ yields $\varepsilon_{p_2,y} = 0.001324 \cdot 1.5 \cdot 7{,}496/0.6392 = 0.00999$. This implies $\varepsilon_{x_2(y,p_2(y)),y} = \text{-}0.685 \cdot 0.00999 + 0.829 = 0.822$.

[44] Since $x_{fuel} = e \cdot x_2$, where e stands for the fuel efficiency measured in fuel consumption per kilometre, it follows that $\varepsilon_{x_{fuel},y} = \varepsilon_{e,y} + \varepsilon_{x_2(y,p_2(y)),y}$. The elasticity of the efficiency with respect to income $\varepsilon_{e,y}$ is $\varepsilon_{e,y} = \beta_2 \cdot y/e$, where $e = \beta_1 + \beta_2 \cdot y$. Plugging in the values mean income $\bar{y} = 7{,}496$ and for the fuel price $p_{fuel} = 1.5$, $\varepsilon_{e,y}$ yields $\varepsilon_{e,y} = 0.1563$. Therefore $\varepsilon_{x_{fuel},y} = 0.2344 + 0.822 = 0.9784$.

[45] To see how these values were computed, see footnotes above. The results are based on the simple OLS regression. The results of this regression are listed in Table 1.3.1.

based on stated preference information, households might not have been willing to make the effort to determine which car would be the best with respect to fuel economy when fuel prices rise, since they did not have to bear the real costs of choosing an inappropriate car, which would lead to high fuel expenditures. It seems that households tended to choose the car type they already knew, or a similar one.[46] If I computed the elasticity $\varepsilon_{x_{fuel}, p_{fuel}}$ based on the average fuel price of the dataset of CHF 3.22 / litre, a value of $\varepsilon_{x_{fuel}, p_{fuel}} = 0.4679$ would result. This value is far greater in absolute terms than the value -0.202 established by Baranzini et al. (2009), but it would be in the range reported in international studies, see Tables 1.3.2, 1.3.3 and 1.3.4.

It is important to note that, in the context of this model, households' choice of car fuel efficiency when income or fuel prices change is purely exogenous. It is assumed that there is no interaction between the decision related to driving distance and car fuel efficiency. Despite this simplification, I consider my method of computing the elasticities of the fuel demand to be a very good approximation.

Comparison of the results with results computed by other models

In the following, I also compute the elasticities of the fuel demand and the probability of a household being carless using the OLS, Tobit and Probit model. The aim is to examine whether the results yielded by these methods will be in a similar range to those I computed using the MDCEV model. Since I use the Erath dataset (2009), fuel prices and therefore the marginal costs of driving also variates sufficiently, such that I expect the value of the corresponding coefficients to be significant. Further, the methods OLS, Tobit and Probit may be able to capture the panel structure of the data. I will examine whether the results differ when I use a model adapted to the panel structure versus models that do not. If this difference in results is minor, I will conclude that the difference in results would also be minor in the case of the MDCEV model, where I ignore the panel structure. The results for the OLS and the Tobit model in the following table can also be found in Table 2.2.3, where the results are also discussed. Here, I will discuss only the differences between the results of the various models. I also added the results of a Probit model, which captures only car choice, in the table below. I added the results of the Probit model because I presume that the Probit yields the most realistic results with respect to car ownership. I only present the results yielded by the Probit model – the theory of the Probit model is presented in Appendix A2.1.

---

[46] Axhausen and Erath (2010) also come to this conclusion: “It can be argued, that the respondents, given the vast number of possible combinations of car types, engine types and sizes, might have had problems finding out which cars were eligible for the incentives,” Axhausen and Erath (2010: 70).

Type of model	**MDCEV**	**OLS**	**OLS**	**Tobit**	**Tobit**	**Probit**	**Probit**
Dataset	Erath	Erath	Erath	Erath	Erath	Erath	Erath
		pooled	panel (rand. eff.)	pooled	panel (rand. eff.)	pooled	panel (rand. eff.)
Number of model	(1)	(2)	(3)	(4)	(5)	(6)	(6)
Limit of integrating, $E(X_2)$	60000km	∞	∞	∞	∞	--	--
d, a_2	0.005, 10	--	--	--	--	--	--
(Constant) b_1	-3.818	-2225* (752)	-2893* (949)	-1693* (945)	-1856* (519)	0.73815 (0.1427)	0.73815 (0.1427)
(Income) b_2	--	0.1011 (0.00375)	0.1016* (0.00887)	0.1228* (0.00464)	0.1250* (0.00893)	-16.8e-6 (1.05e-6)	-119e-6 (13.6e-6)
(Driving costs) b_3	--	-6396* (1548)	-7768* (487.2)	-7835* (1931)	-9917* (604)	0.6786 (0.2859)	0.5742 (0.0848)
(Fuel price) b_3	--	-497.8* (120)	-605.5* (38.45)	-609.8* (150.4)	-771.8 (47.01)	0.0528 (0.0223)	0.0528 (0.0223)
(Place of living rural) b_4	0.070	529.2* (347)	347 (819)	1566* (430)	2112* (750)	-0.6583 (0.0711)	-4.7274 (1.216)
(Number of people in household) b_5	0.079	1354.5* (155)	1360.2* (366)	1796* (190)	2192* (345)	-26497 (0.0343)	-1.989 (0.5636)
β	0.508	--	--	--	--	--	--
Proportion of drop-outs	2.90%	0	0	0	0	0	0
σ	--	8577	2673	10347 (157.8)	2995 (48.18)	1	1
σ_η	--	--	8185	--	10876 (349.2)	--	7.597 (0.591)
Relative replication error, $P(X_2=0)$	-0.025	--	--	0.036	--	0.0014	--
Relative replication error, $E(X_2)$	-0.03	0	0	-0.005	--	--	--
Q	0.0011	--	--	0.0013	--	--	--
R^2 or *pseudo* R^2 or *rho*		0.28	-0.28	0.02	0.825 (0.0050)	0.22	0.983 (0.0026)
$\varepsilon_{E(X_2),p_2}$	-0.685	-0.245 (0.0593)	-0.298 (0.0187)	-0.231 (0.0765)	-0.279 (0.02036)	--	--
$\varepsilon_{E(X_2),p_{fuel}}$	-0.268	-0.150 (0.0362)	-0.182 (0.0114)	-0.140 (0.0467)	-0.171 (0.01243)	--	--
$\varepsilon_{E(X_2),y}$	0.829	0.774 (0.0287)	0.778 (0.0679)	0.790 (0.0374)	0.786 (0.05809)	--	--
$\varepsilon_{E(X_2),k_2}$	-0.148	--	--	--	--	--	--
$\varepsilon_{P(X_2=0),p_2}$	0.206	--	--	0.378 (0.8175)	0.472 (0.03908)		
$\varepsilon_{P(X_2=0),p_{fuel}}$	0.095	--	--	0.233 (0.0499)	0.290 (0.02387)	0.1787 (0.0721)	0.23741 (0.0721)
$\varepsilon_{P(X_2=0),y}$	-1.300	--	--	-0.913 (0.0435)	-0.950 (0.11089)	-1.4504 (0.0754)	-1.2480 (0.0351)
$\varepsilon_{P(X_2=0),k_2}$	1.090	--	--	--	--	--	--

Note 1: The marginal effects of the Probit and the OLS models are computed at the sample mean; in the case of the models that capture the panel structure of the data, it was additionally assumed that all η_n are zero.
Note 2: The standard errors of the simulated values were computed for the case where the elasticities were computed at the sample mean.
Note 3: The average of the simulated values was not computed for the case of models that capture the panel structure of the data.
Note 4: The values in parentheses "(..)" denote standard deviations.
Note 5: The levels of significance are denoted as: *: p<0.05, **: p<0.01 and ***: p<0.001.

Table A3.14.3: Elasticities based on different models.

The results listed in the table above show that the value of the MDCEV model of the elasticity of driving distance with respect to income $\varepsilon_{E(X_2),y}$ is only slightly greater that those yielded by the other models. In contrast, the value of the MDCEV of the elasticity of the probability of being carless car with respect to income $\varepsilon_{P(X_2=0),y}$ is about 30% greater than that yielded by the Tobit model, but virtually the same as the value yielded by the Probit model. Since the Probit model captures only the decision to be carless or own at least one car, I believe its results map this decision better than the Tobit model. One further reason why the results of Probit and the MDCEV model are more trustworthy than those yielded by the Tobit model is that the mechanism between the decision to own a car and driving demand is given by the model structure. In contrast, in the case of the MDCEV model, this link is not completely fixed because it depends on a number of parameter values. Regarding the elasticity of driving distance with respect to fuel price $\varepsilon_{E(X_2),p_2}$, the value of the MDCEV model is more that 50% greater than those yielded by the other model. It is therefore also greater than the value reported by Baranzini et al. (2009). On the other hand, it is closer to the range found for the average values by international studies, which is -0.30 with respect to driving distance and -0.64 and -0.43, respectively, with regard to fuel demand according to Brons et al. (2006) and Goodwin et al. (2004), see also Tables 1.3.2 and 1.3.3. Note that the elasticity of car fuel efficiency with respect to fuel prices I find based on the Erath dataset is much lower than that established by international studies. I explain this difference as follows: the time series data on average fuel economy reflect not only changes due to consumer preferences, as in the Erath dataset. It also contains the impact on a change in the car's supply, which was basically driven by the regulation. In the USA in particular, the so-called “Corporate Average Fuel Economy (CAFE)” is the standard imposed by regulations on the average fuel efficiency of a car manufacturer's car fleet.

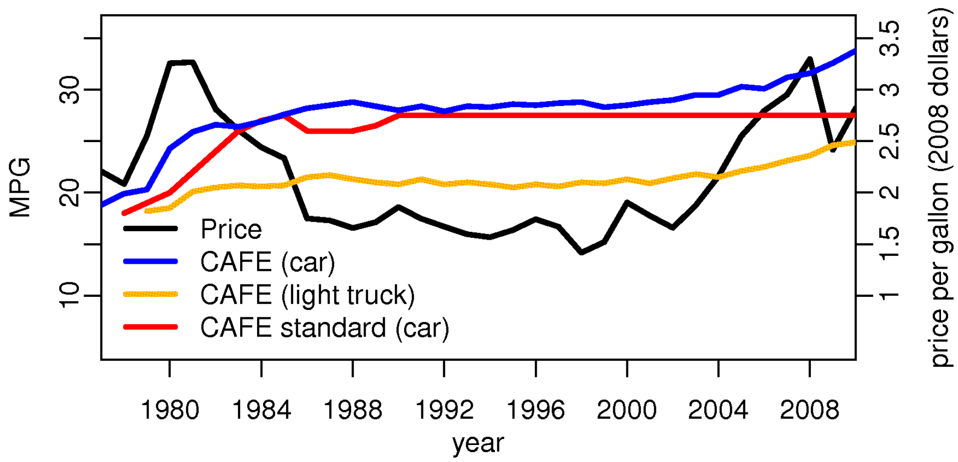


Figure A3.14.5: Fuel prices and the standard on average fleet consumption in the USA.[47]

[47] I found this diagram in http://en.wikipedia.org/wiki/Corporate_Average_Fuel_Economy. According to its author, this diagram is based on data supplied by the Energy Information Administration (EIA) (2010) and National Highway Traffic Safety Administration (NHTSA) (2010).

This diagram clearly shows that, during periods of high fuel prices, the standard on the average fleet consumption in the USA (CAFE standard) was increased. According to the Board On Energy and Environmental Systems (2002), the regulation by the CAFE standard led to an additional overall increase in 14%[48] from 1980 to 2002. Since during this period the overall improvement was about 33%[49], about 40% of this improvement was induced by the CAFE standard. Assuming that a share of 0.4 on the improvement in fuel economy is contributed by the supply side in all countries, the elasticity of fuel efficiency induced by consumer preferences amounts to about 0.14, Goodwin et al. (2004) and 0.18 Graham and Glaister (2005), see Table 1.3.2. These two values are now very close to that established using the Erath dataset (0.1324) when computing the elasticity at the sample mean fuel price.

I conclude from all this that the results yielded by the MDCEV model are very reasonable. Of course, not for all elasticities values that yield from other studies could be found to use as a benchmark. But it would be rather surprising if all results with regard to income and fuel prices yielded results in the range reported by other studies, implying that all results based on changes in fixed costs would be completely wrong.

[48] See Board On Energy and Environmental Systems (2002: 111).

[49] According to the figure, the average fuel economy rose by one third from 21 miles / gallon to 28 miles / gallon.

A 3.15 Comparing the Tobit model with the MDCEV model

In this subchapter, I shall examine the Tobit or the MDCEV model using the micro-census data of Swiss households from 2005, Bundesamt für Statistik (2006a). First, I shall examine whether these models yield different elasticities. Second, I shall explain why I excluded the fuel price as an explanatory variable from the Tobit model. Third, I will show that including additional household variables leads to virtually no change in the parameter related to income and the income elasticity of driving demand. Fourth, I shall examine how effectively these models adapt the data. I shall perform this quantitatively by checking ow accurately the model was able to forecast the average driving distance for each household segment and by comparing how well the density functions of the two models adapt the observed data.

The Tobit model I use is defined as follows:

$$y_n^* = \alpha + \beta \cdot x_n + \varepsilon_n \,. \qquad \text{(A3.15.1)}$$

where index n indicates the household and N the total number of households. Parameter α is constant; parameter vector β captures weights relating to x_i, which stands for variables such as fuel prices, the marginal costs of driving and household properties. Note that variable y_n^* is a latent variable and variable y_n:

$$y_n = \begin{Bmatrix} y_n^* < 0 : y_n = 0 \\ y_n^* \geq 0 : y_n = y_n^* \end{Bmatrix} \qquad \text{(A3.15.2)}$$

denotes the driving distance.

The following table shows that both the elasticity of driving and the relative change in the probability of being carless with respect to income of the Tobit model is approximately 23% smaller than the elasticity yielded by the MDCEV model based on a modified density function. Comparing the diagrams that illustrate the differences between the forecast and the actual average driving distances of each household segments, Figures A3.15.1 versus A3.13.3, I conclude that the value of the Tobit model is more likely to be correct. The reason why I eliminated the fuel price from the equation is that the corresponding parameter was determined by a difference in the average fuel price of carless households and households owning at least one car. I consider this difference[48] to be coincidental because I do not believe that the decision whether or not to buy a car depends on such small differences of the actual fuel price, which were rather small in the year during which the data was

[48] The average fuel price for carless households was CHF 1.4738 /litre; the price was CHF 1.4697 /litre for households with a car.

collected. A further justification for eliminating the fuel price from the mode is that, when considering only households with at least one car and explaining the driving demand using an OLS model, the parameter corresponding to fuel price is insignificant. It would therefore be strange if people reacted to the actual fuel price when deciding to buy a car but not when deciding how much to drive the car. I would actually expect the opposite. Considering all these arguments, I believe it is correct to exclude fuel price from the model, even though a significant coefficient resulted in the case of the Tobit model.

Type of model	**mz05**	**Tobit**	**Tobit**
Dataset	mz05	mz05	mz05
Number of model	(1)	(2)	(3)
Limit of integrating, $E(X_2)$	60000km	∞	∞
d	0.1	--	--
a_2	2	--	--
(constant) α	-2.955	-4012* (231)	264.1 (2809)
(income) β_1	--	0.127* (0.0018)	0.117* (0.0022)
(driving costs) β_2	--	--	--
(fuel price) β_3	--	--	--
(place of living rural) β_{3a}	0.354	5589* (212)	--
(number of people in household) β_4	0.044	2113* (74)	--
σ	--	12495* (70)	12180* (68)
share of dropouts	9.60%	--	--
relative replication error, $P(X_2=0)$	-0.03	0.04	0.06
relative replication error, $E(X_2)$	-0.09	0	0
R^2 or *pseudo* R^2	--	0.22	0.02
Penalty Q	0.009	0.0017	0.0036
$\varepsilon_{E(X_2),p_2}$	-0.68	--	--
$\varepsilon_{E(X_2),p_{fuel}}$	-0.28	--	--
$\varepsilon_{E(X_2),y}$	0.78	0.616 (0.0106)	0.574 (0.0107)
$\varepsilon_{E(X_2),k_2}$	-0.17	--	--
$\varepsilon_{P(X_2=0),p_2}$	0.26	--	--
$\varepsilon_{P(X_2=0),p_{fuel}}$	0.11	--	--
$\varepsilon_{P(X_2=0),y}$	-1.41	-0.8504 (0.019)	-0.759 (0.019)
$\varepsilon_{P(X_2=0),k_2}$	-1.31	--	--

Table A3.15.1: Estimated parameter values and simulated elasticities corresponding to the complete dataset.

In addition, I tested whether the inclusion of more household variables would change the magnitude of the income elasticity. I added two types of explanatory variables in model (3) in Table A3.15.1: one covers the place of living and the other information on the type of household. Including all these variables leads only to a small change in the income elasticity compared to the simple Tobit model (2). This finding is similar to that in Section A3.12 for the MDCEV model. But the inclusion of these variables reveals a number of interesting facts. First of all, the model shows that households living in agglomerations, towns or rural areas drive 4,614, 4,132 and 7,709 km, respectively, more than households living in city centres. The place of living can be further segmented in economic regions. It turns out that households living in the Middle lands, in the Lémanique and eastern parts drive about 1000 km more per year than households living in the economic region of Zürich or north-western Switzerland, which are the most urbanized regions of Switzerland. Households living in the mountainous region in the south of Switzerland, the Ticino region, even drive about 4,000 km more per year than people living in the economic region of Zürich or north-western Switzerland. Note for comparison that an additional income of CHF 40,500 also has an impact on driving demand of 4,000 km. I believe there are two reasons for this regional effect: first, facilities are more densely distributed in urban regions. This means that the average distances to get to work and leisure facilities are shorter. Second, the supply of public transportation is more frequent and stops are more dense in these regions. Interesting results were also found when I examined the impact of the family structure on driving demand. The results showed that families with only one parent do not drive more than single households while couples drive 4,693 km more than singles. I assume the reason for this is that these households have more to spend for housing than couples and the available budget is therefore smaller. The additional driving demand of traditional families with two parents and children is dramatically higher than that of single households, namely 7,180 km. A last interesting finding is that an additional child in a one-parent household causes 1,000 km additional driving demand while an additional child in a two-parent household does not cause any additional driving demand. I assume that in the latter case the effect of additional driving demand balances out with the effect of the reduction of available income due to higher housing costs. Furthermore, I assume that for most journeys all children are driven to the same place such that there is no additional driving if there is one child more. All these findings are summarized in Table A3.15.2.

	variable name	$\partial y/\partial x$	stddev	95% confidence interval		mean of variable
dep. var						
	driving distance y					13890
indep. var						
	income [CHF/year]	0.0979	(0.0019)	0.0943	0.1016	80169
type of location	(base = city)					
	agglo*	4,614	(175)	4,271	4,957	0.4352
	town*	4,137	(971)	2,234	6,041	0.0069
	rural*	7,709	(226)	7,267	8,152	0.2275
economic region	(base = Zürich)					
	Lémanique*	829	(244)	351	1,308	0.1782
	Middle lands*	949	(222)	513	1,384	0.2716
	north west*	14	(349)	-669	698	0.0579
	east*	1,165	(279)	618	1,712	0.1267
	Swiss-Central*	493	(271)	-37	1,023	0.1314
	Canton Ticino*	3,969	(449)	3,088	4,849	0.0365
family type	(base = single)					
	one parent with kids*	142	(1147)	-2,105	2,389	0.0537
	two parents with kids*	7,181	(672)	5,864	8,498	0.2988
	couple*	4,694	(205)	4,291	5,096	0.3132
	flat share*	4,091	(484)	3,141	5,041	0.0308
	subtenancy*	2,386	(5709)	-8,804	13,577	0.0002
effect of an addition kid						
	one parent with kids*	1,065	(418)	245	1,885	0.1391
	two parents with kids*	-59	(154)	-361	243	1.1743

* Note that these are dummy variables. The marginal effect of these variable corresponds to a discrete change from zero to one.

Table A3.15.2: Marginal effects on driving demand.

Note that all these effects were computed at the mean values of the independent variables.

I shall now discuss how well the Tobit model adapts to the data. To do so, I estimate a Tobit model capturing only the variables income and type of place of living (rural versus urban area). I then compare the values forecast by the model and the empirical values of the corresponding household segments. Note that the coefficient for income increases slightly to 0.14821 due to the omitted variable bias. The results are listed in the following Table A3.15.3 and in Figures A3.15.1 and A3.15.2. In both cases, the deviation of the forecast probability of being carless and the deviation from of the forecast driving distance from their empirical values do not increase or decrease in the income range CHF 60,000 to 128,000, which includes most households. I conclude from this that the value of the coefficient corresponding to income is more or less correct. Since the coefficient of the "large model" (3) in Table A3.15.1 amounts to only 0.117, which is 21% lower than 0.1482, I conclude that the income elasticities corresponding to model (3) are also rather a lower bound of the actual. Therefore, the true elasticity of driving demand could be approximately up to 0.535 · 1.2 = 0.642. This value is almost exactly the geometrical mean of the result of the MDCEV model (1) and the "large" Tobit model (3) in Table A3.15.1, and could be considered the best estimate for the income elasticity of driving demand. Analogously, the best estimate of the elasticity of the share of carless households could be approximately -0.953 · 1.2 = -1.144. These two values also seem realistic since, in the case of the MDCEV model with a modified density function (1), the differences of simulated and empirical values increase slightly in both cases, see Figures A3.13.3 and A3.13.4. Therefore, the elasticities supplied by the MDCEV model are rather upper bound values.

y	Location	Nobs	Share of $x_2 = 0$	$P(X_2 = 0)$	Δ	$mean(x_2)$	$E(X_2 = 0)$	Δ
18,000	0	576	0.7344	0.4462	-0.2881	2,196	5,979	3,783
18,000	1	224	0.5938	0.2623	-0.3315	3,186	10,113	6,926
36,000	0	3,038	0.4608	0.3650	-0.0958	5,718	7,565	1,847
36,000	1	914	0.2593	0.1987	-0.0606	8,743	12,168	3,425
60,000	0	4,382	0.2097	0.2661	0.0563	10,877	10,004	-873
60,000	1	1,462	0.0581	0.1301	0.0720	15,952	15,146	-806
84,000	0	3,380	0.1092	0.1829	0.0737	15,001	12,768	-2,233
84,000	1	1,051	0.0276	0.0799	0.0523	20,536	18,335	-2,201
108,000	0	2,166	0.0919	0.1182	0.0263	17,045	15,795	-1,250
108,000	1	572	0.0122	0.0460	0.0337	23,830	21,672	-2,158
132,000	0	1,177	0.0671	0.0716	0.0045	19,095	19,019	-76
132,000	1	276	0.0145	0.0247	0.0102	26,425	25,107	-1,318
156,000	0	617	0.0389	0.0406	0.0017	21,672	22,381	709
156,000	1	117	0.0085	0.0124	0.0038	29,953	28,600	-1,353
180,000	0	295	0.0475	0.0215	-0.0259	20,779	25,830	5,052
180,000	1	66	0.0000	0.0058	0.0058	32,125	32,126	1
228,000	0	475	0.0400	0.0049	-0.0351	23,841	32,863	9,022
228,000	1	82	0.0244	0.0010	-0.0234	30,635	39,220	8,584

Table A3.15.3: Empirical and simulated values of driving distance and shares of carless households

Figures A3.15.1 and A3.15.2 below illustrate the difference in the forecast and the empirical values.

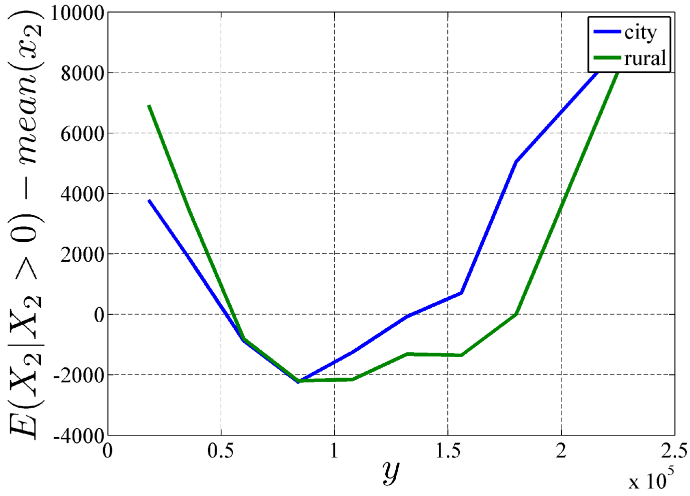


Figure A3.15.1: Deviation of simulated share of carless households from empirical value.

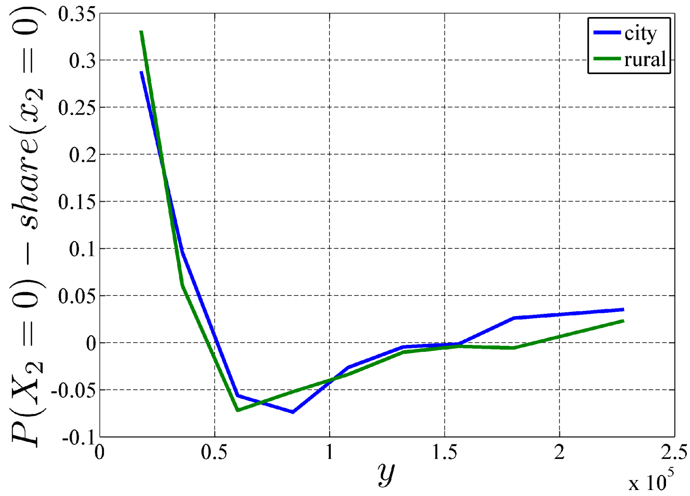


Figure A3.15.2: Deviation of simulated share of carless households from empirical value.

Finally, I examine how well the Tobit model adapts to the data. To do so, I compare the shape of the histograms of each household segment with the theoretical density functions defined by the model parameters.

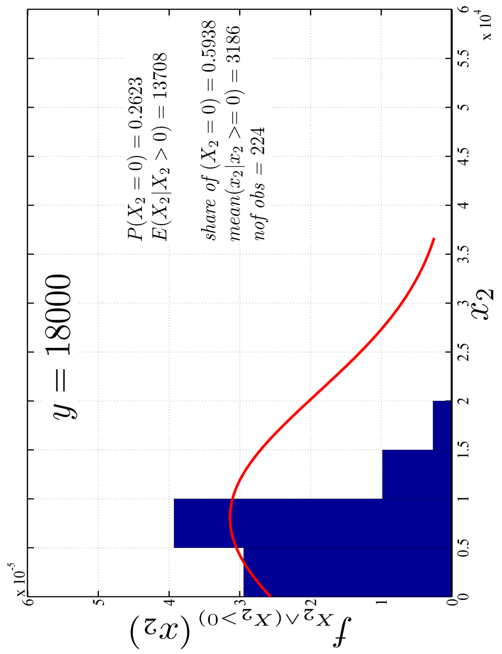


Figure A3.15.3: Histogram of households in rural areas with an income of CHF 18,000.

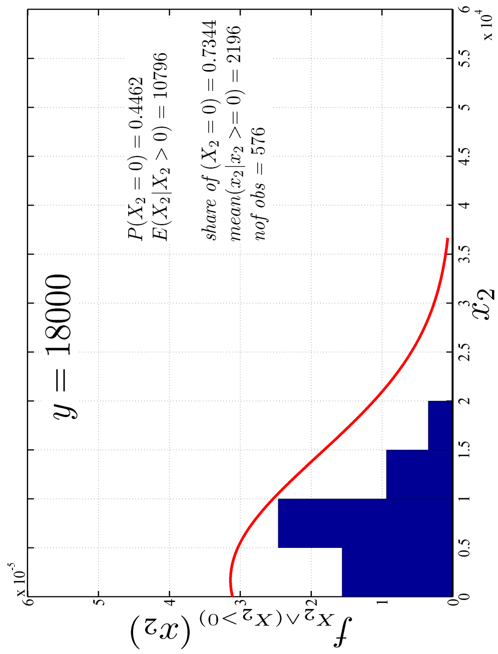


Figure A3.15.4: Histogram of households in urban areas with an income of CHF 18,000.

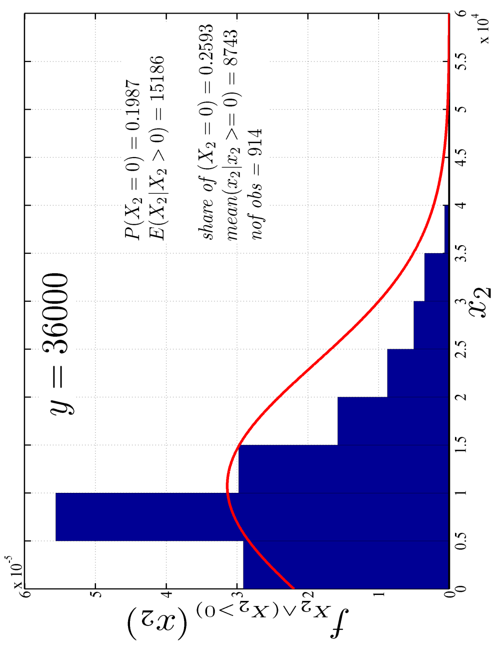


Figure A3.15.5: Histogram of households in rural areas with an income of CHF 36,000.

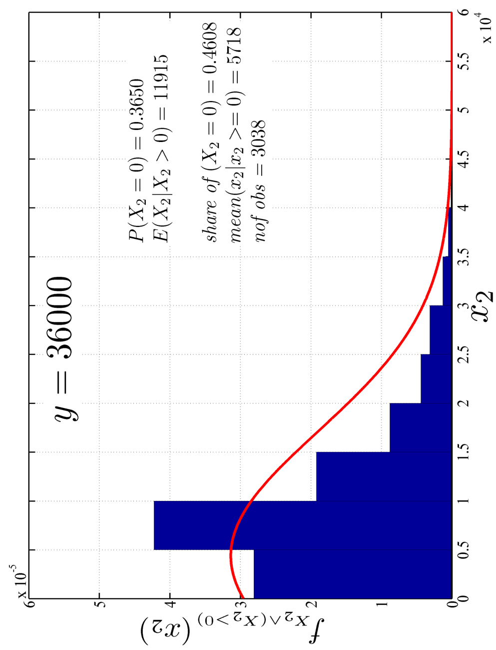


Figure A3.15.6: Histogram of households in urban areas with an income of CHF 36,000.

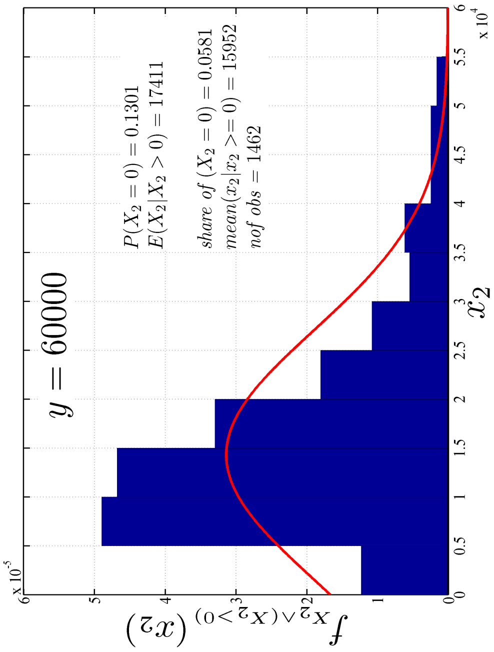


Figure A3.15.7: Histogram of households in rural areas with an income of CHF 60,000.

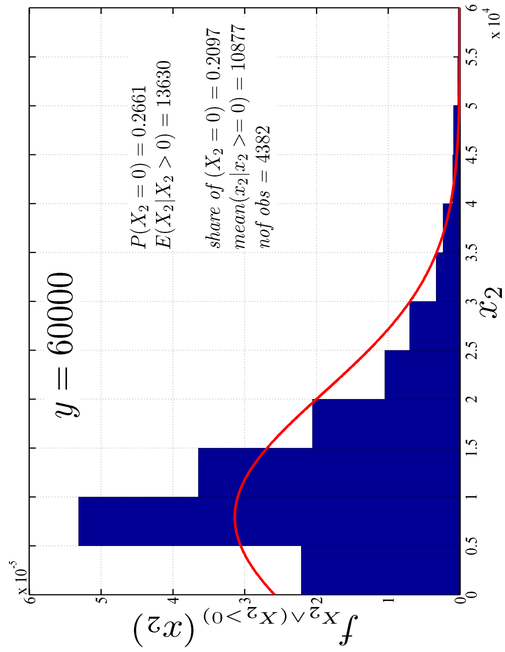


Figure A3.15.8: Histogram of households in urban areas with an income of CHF 60,000.

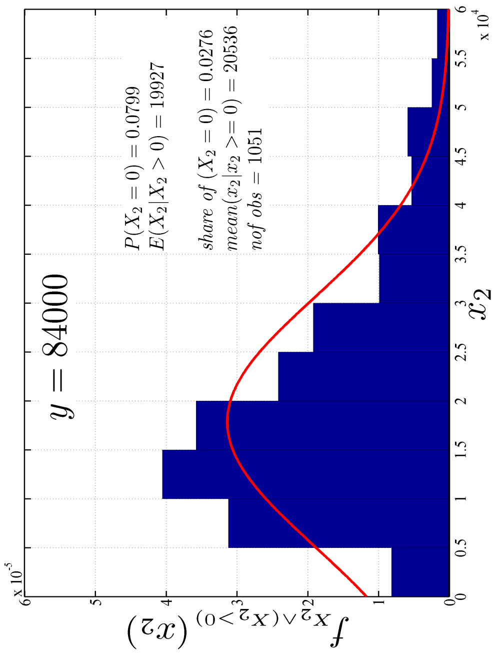


Figure A3.15.9: Histogram of households in rural areas with an income of CHF 84,000.

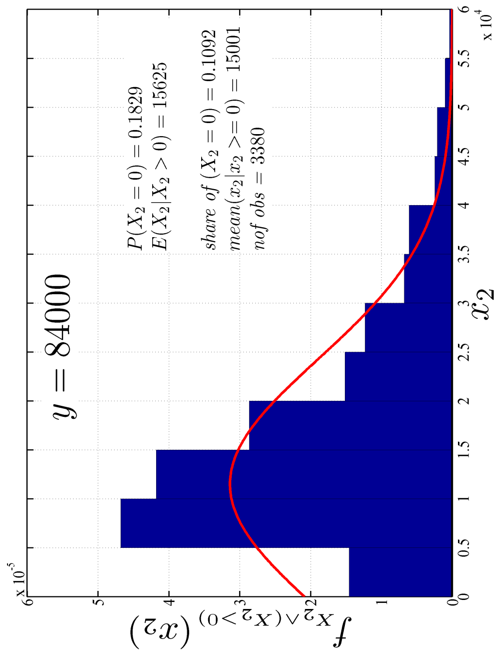
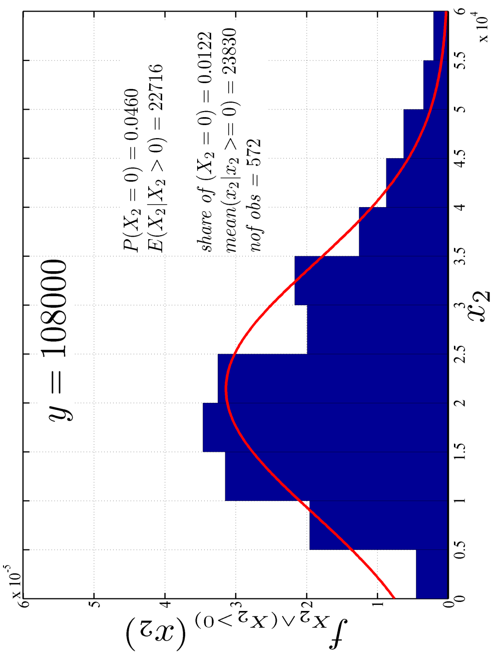


Figure A3.15.10: Histogram of households in urban areas with an income of CHF 84,000.



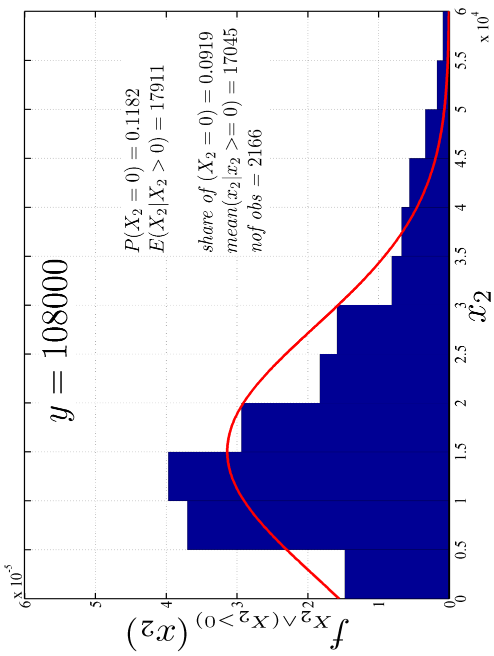
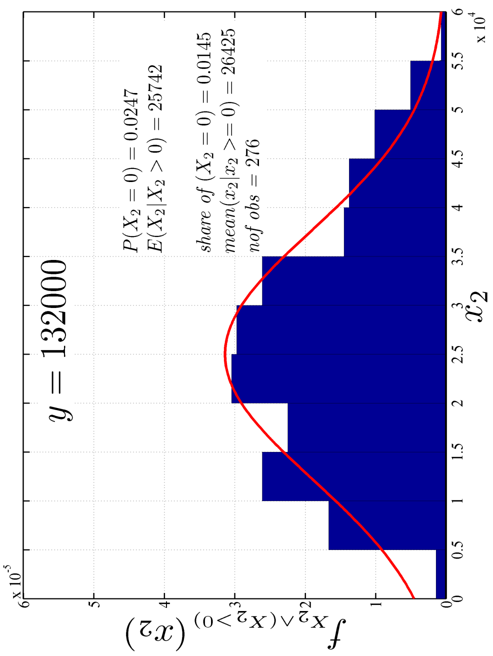


Figure A3.15.11: Histogram of households in rural areas with an income of CHF 108,000.

Figure A3.15.12: Histogram of households in urban areas with an income of CHF 108,000.



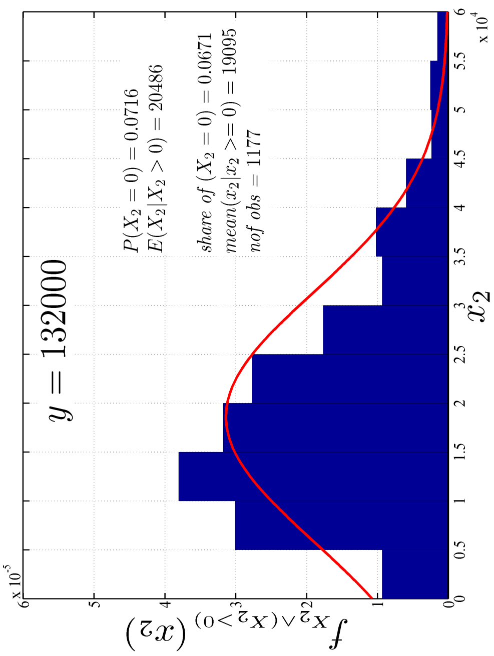


Figure A3.15.13: Histogram of households in rural areas with an income of CHF 132,000.

Figure A3.15.14: Histogram of households in urban areas with an income of CHF 132,000.

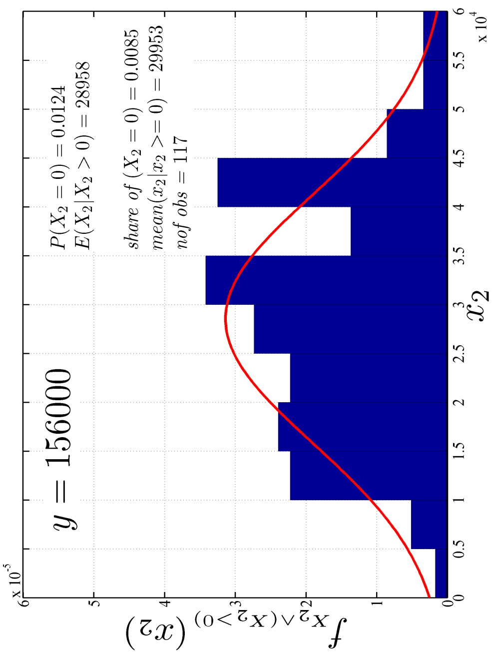


Figure A3.15.15: Histogram of households in rural areas with an income of CHF 156,000.

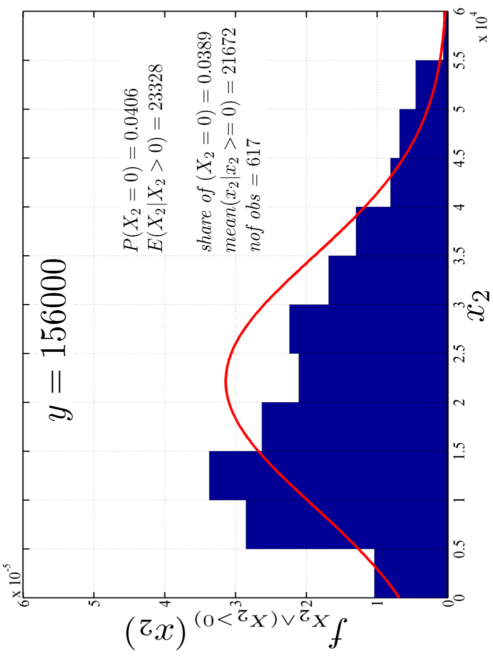


Figure A3.15.16: Histogram of households in urban areas with an income of CHF 156,000.

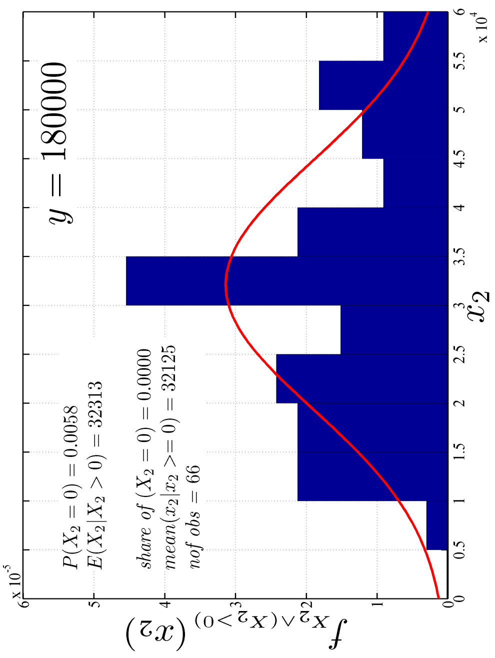


Figure A3.15.17: Histogram of households in rural areas with an income of CHF 180,000.

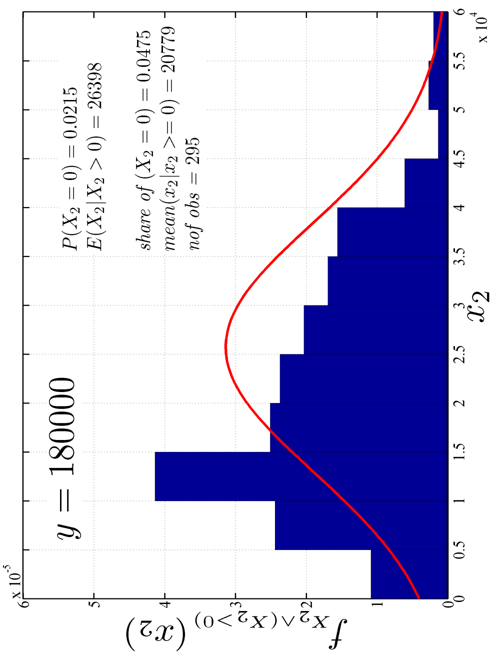


Figure A3.15.18: Histogram of households in urban areas with an income of CHF 180,000.

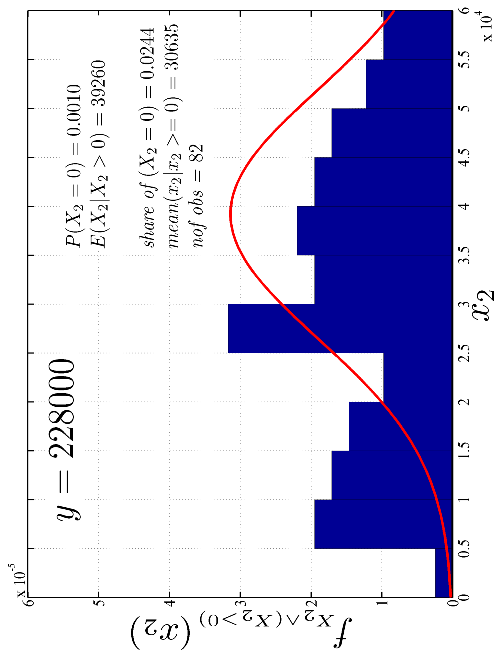


Figure A3.15.19: Histogram of households in rural areas with an income of CHF 228,000.

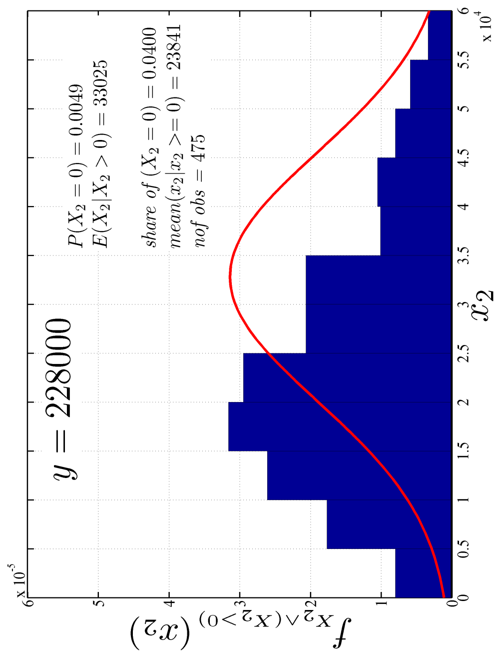


Figure A3.15.20: Histogram of households in urban areas with an income of CHF 228,000.

A 4 Appendix to Chapter 4

A 4.1 Additional results of Subchapter 4.3

In this appendix, I present the results generated when using the dataset from which households that use the car for work have been excluded. The model used is identical to the one for which the results are presented in Table 4.3.2.

	$\hat{\delta}_i$	$\text{stdev}\left(\hat{\delta}_i\right)$	$\text{cov}\left(\hat{\delta}_{e,\cdot},\hat{\delta}_p\right)$	t-value	$P\left(\hat{\delta}_i\right) \neq 0$
$\delta_{e,others}$	-0.0773	0.0444	0.00015	-1.74	0.08
$\delta_{e,el}$	-0.216	0.0783	1.33E-04	-2.76	0.01
$\delta_{e,work}$	-0.0588	0.0398	-1.78E-05	-1.48	0.14
δ_p	-0.0426	0.0146		-2.92	0
Number of households: 80, number of observations: 1249					
Likelihood values (null, final) : (-1372.167, -993.201)					
Likelihood ratio test: 757.931					
$\rho^2 = 0.276$, $\rho^2_{adj.} = 0.231$					
Number of random draws *S* per household: 800					

Note 1: All the standard deviations, covariances and t-values are "robust" estimators according to the software Biogeme.
Note 2: Parameter δ_p corresponds to the car price measured in CHF 1,000.

Table A4.1.1: Effects of fuel efficiency and price on car choice.

The corresponding willingness to pay *wtp* for fuel efficiency is shown in Table A4.1.2 for the different segments.

				Test statistics		
Segment	wtp	wtp_{rat}	wtp/wtp_{rat}	sdev	t-value test	p-value
"Others"	1,815	1,844	0.98	0.5875	-0.0269	0.4893
"El"	5,070	2,352	2.1559	1.0111	1.1432	0.8735
"Work"	1,380	2,412	0.5722	0.4394	-0.9735	0.1651

Note 1: The *wtp*, the wtp_{rat} and the test statistics are computed from Formulas (4.3.2) and (4.3.3) in Subchapter 4.3.
Note 2: The value wtp_{rat} is based on a fuel price of CHF 1.50 / litre, which was the price at the time of the survey and the average driving distance of the corresponding household segment. I assumed that the household expected the annual rise of fuel prices to be 2% less than the interest rate. Further, I assumed that households do not plan to change their driving distance and that a car lifetime of nine years is expected.

Table A4.1.2: Effects of fuel efficiency and price on car choice.

Again, as in the case where I used the complete dataset to estimate the model, the standard deviations are rather large. For this reason, the differences between these results and the results in Table 4.3.3 cannot be considered to be statistically significant. Nonetheless, it is interesting to note how the point estimates change. First, the willingness to pay *wtp* for fuel efficiency of households in the category "others" is higher and corresponds exactly to the case that would result from economic rational behaviour. This is a sign that if households have to bear all costs of driving, they behave economically rationally. In contrast, the *wtp* of the segment "work" is lower in this case, but not to such an extent that the estimated δ_{work} is insignificant. The *wtp* of the segment "el" is even higher than in Table 4.3.3. Altogether, the *wtp* have increased. This is because the *wtP* of the group that uses the car for work is lower than the one for the other households.

A 4.2 Diagrams of Swiss car stock

The following diagrams were copied from Vereinigung Schweizer Automobilimporteure (Auto Schweiz), (2009). Whilst the corresponding figures in Subchapter 4.3 cover only the years 1996-2009, the data illustrated in the following diagrams go back to 1989.

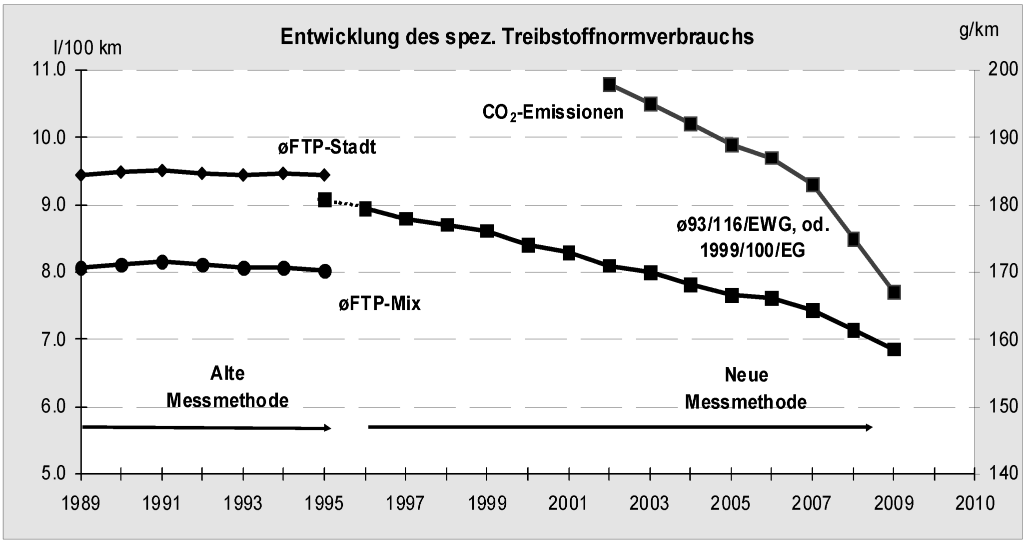


Figure A4.2.1: The average fleet consumption of cars imported to Switzerland.

Figure A4.2.1 shows that the average fleet consumption remained stable up to 1995, when a new method was introduced for measuring cars' fuel consumption. The aggregate fleet consumption of imported cars only started to decline noticeably after this point in time, a phenomenon which I am unable to explain.

Figure A4.2.2 shows that engine sizes and car weights increased gradually in the period 1989-2007. This trend only stopped after 2007; since then, both values have decreased. Comments on possible reasons for this decline can be found in Subchapter 4.3.

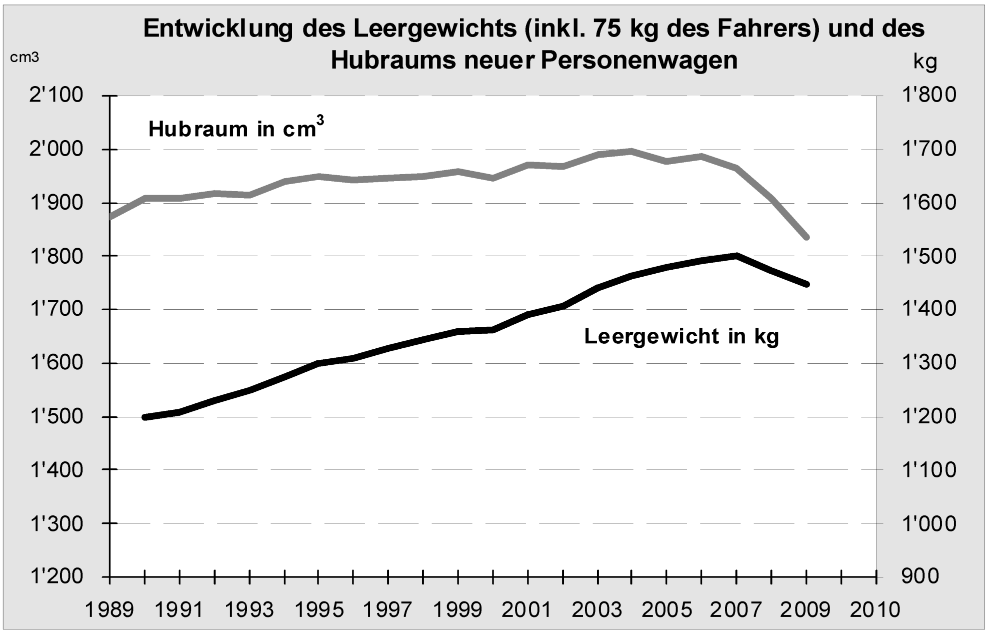


Figure A4.2.2: The average engine size and average weight of cars imported cars to Switzerland.

Finally, Figure A4.2.3 shows the target for the average fleet consumption defined by the Swiss government. The results show that the values of the actual fleet consumption of imported cars clearly fail to meet these target values, and that the gap is even increasing. Considering also the fact that the difference between car fuel consumption in practical use and the values given for normalized driving cycles has actually increased, this gap would be even larger, see Figures 4.3.3 and 4.3.4.

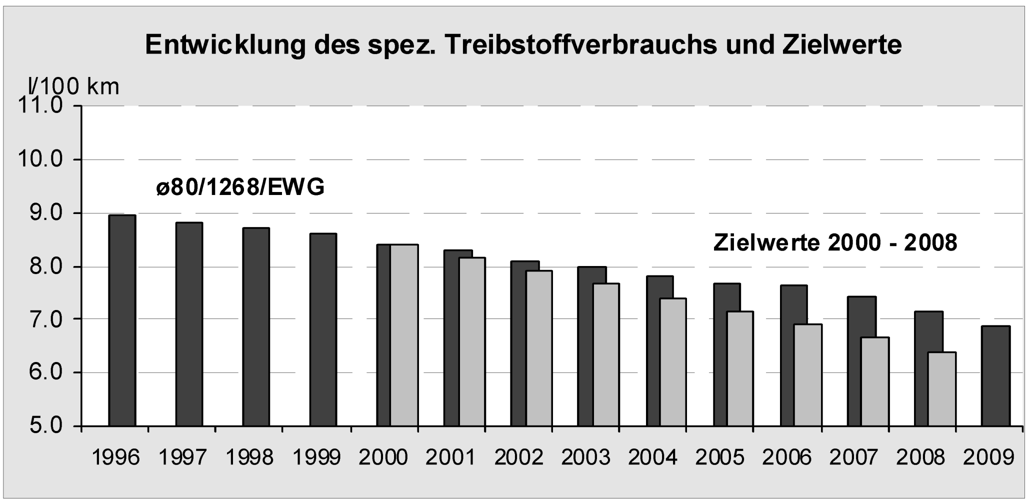


Figure A4.2.3: The average engine size and average weight of cars imported to Switzerland.

Mathematical Appendix

Mathematical Appendix

The Gumbel distribution is defined in Mathematical Appendix MA 1. In Chapters 3, 4 and 5, the models are based on random variables that are Gumbel distributed. I also noted a number of rules that will be useful when proving the error correction term used in the Discrete-Continuous choice model presented in Chapter 5.

In Mathematical Appendix MA 2, I determine the sign of the impact parameters and economic variables have on the minimum consumption level, which play an important role in the Multiple Discrete-Continuous Extreme Value Model (MDCEV) discussed in Subchapter 3.3. This Mathematical Appendix MA 2 will help readers understand the mechanism of the estimation routine proposed in Subchapter 3.3

MA 1 Gumbel distribution

The Gumbel distribution as used in the context of this paper is defined as follows:

$$X \sim gu(0,1)\ ,\ f_g(x) = e^{-x} \cdot \exp\left(-e^{-x}\right) \text{ and } F_g(x) = \exp\left(-e^{-x}\right)\ . \tag{MA1.1}$$

The shape of the probability density function is as follows:

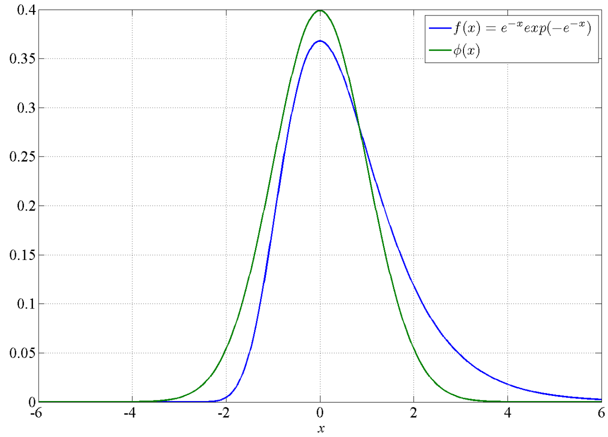


Figure MA1.1: Probability functions of the standard Gumbel and the standard normal distribution

The shape of the cumulated density function is as follows:

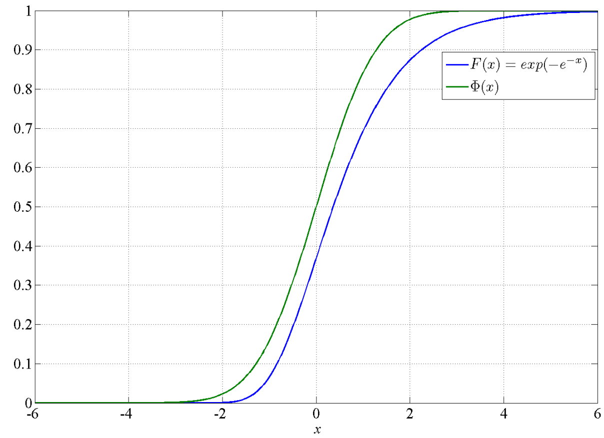


Figure MA1.2: Cumulated density functions of the standard Gumbel and the standard normal distribution

The diagram shows that the shapes of the density function of the standard Gumbel and the standard normal distribution are very similar. The standard Gumbel is non-symmetric and the mean is non-zero, namely equalling the Euler Mascherioni constant $E(X) = \lambda = 0.577..$. The variance is lower than that of the standard normal distribution, namely $\text{var}(X) = \pi^2/6 = 0.523..$.

In the following, the key properties of this distribution are listed. Further, some distributions of linear combinations and expectation values of transformed Gumbel distributed variables are listed. These will be useful when deriving some of the formulas in Chapters 3 and 4.

1. If X is standard Gumbel distributed, then $X \sim f_g(x) = e^{-z} \cdot e^{-e^{-z}}$
   a.) $E(X) = \lambda = 0.577..$,
   b.) $\text{var}(X) = \pi^2/6 = 0.523..$,
   c.) $\text{median}(X) = -\ln(\ln(2)) = 0.3665..$,[51]
   d.) $\text{mode}(X) = 0$.[52]

2. If Z is Gumbel distributed $Z \sim f(z) = \alpha \cdot e^{-\alpha z+\beta} \cdot e^{-e^{-\alpha z+\beta}}$, then
   a.) $E(Z) = \dfrac{\lambda + \beta}{\alpha}$,[53]
   b.) $\text{var}(Z) = \dfrac{\pi^2}{6 \cdot \alpha^2}$,
   c.) Z is distributed as $Z = \dfrac{X + \beta}{\alpha}$, where X is standard Gumbel distributed.

3. If X_1 and X_2 are iid standard Gumbel and are linear transformations with the same shifting parameter, namely $Y_1 = a_1 + bX_1$ and $Y_2 = a_2 + bX_2$, then $Z = Y_1 - Y_2 = a_1 - a_2 + b(X_1 - X_2)$ is distributed as:[54]

$$Z \sim F(z) = \frac{1}{1 + e^{-\frac{1}{b}(z - a_1 + a_2)}}.$$

[51] $F_Z(z) = \exp(-e^{-x}) = 0.5 \Leftrightarrow -x = \log(-\log(0.5)) \Leftrightarrow x = -\log(\log(2))$.

[52] $\dfrac{\partial f_Z(z)}{\partial z} = -e^{-x}\exp(-e^{-x}) + e^{-x}e^{-x}\exp(-e^{-x}) = 0 \Leftrightarrow e^{-x}\exp(-e^{-x}) = e^{-x}e^{-x}\exp(-e^{-x}) \Leftrightarrow e^{-x} = 1 \Leftrightarrow x = 0.$

[53] $\displaystyle\int_{z=-\infty}^{z=\infty} x \cdot \alpha \cdot e^{-\alpha x+\beta} \cdot e^{-e^{-\alpha x+\beta}} dz = \int_{z=-\infty}^{z=\infty} \frac{z+\beta}{\alpha} \cdot e^{-z} \cdot e^{-e^{-z}} dz = \frac{\beta}{\alpha} + \frac{1}{\alpha} \cdot \int_{z=-\infty}^{z=\infty} z \cdot e^{-z} \cdot e^{-e^{-z}} dz = \frac{\beta}{\alpha} + \frac{1}{\alpha} \cdot E(Z) = \frac{\lambda+\beta}{\alpha}$.

[54] Proof: First the cumulated density function of Z conditional on X_2 has to be calculated: $F_{Z|X_2}(z) = F_{Z|X_2}\left(\dfrac{z - a_1 + a_2}{b} + X_2\right)$. Since X_2 is standard Gumbel distributed and $F_{Z|X_2}(z)$ is the cdf of the standard Gumbel distribution, rule 6 of MA1 can be applied:
$E_{X_2}\left(F_{Z|X_2}(z)\right) = E_{X_2}\left(F_{Z|X_2}\left(\dfrac{z - a_1 + a_2}{b} + X_2\right)\right) = \dfrac{1}{1 + e^{-\frac{1}{b}(z - a_1 + a_2)}}$.

The corresponding density function is:

$$Z \sim f(z) = \frac{1}{\beta} \cdot \frac{e^{-\frac{1}{b}(z - a_1 + a_2)}}{\left(1 + e^{-\frac{1}{b}(z - a_1 + a_2)}\right)^2}.$$

4. If X_1 and X_2 are iid standard Gumbel and are linear transformations, namely $Y_1 = a_1 + bX_1$ and $Y_2 = a_2 + bX_2$, then $Z = \max(Y_1, Y_2)$ is distributed as:[55]

$$Z \sim F(z) = \exp\left(-e^{-\left(\left(\frac{z}{b}\right) - \ln\left(e^{\frac{a_1}{b}} + e^{\frac{a_2}{b}}\right)\right)}\right).$$

The corresponding density function is:

$$Z \sim f(z) = \frac{1}{\beta} \cdot e^{-\left(\left(\frac{z}{b}\right) - \ln\left(e^{\frac{a_1}{b}} + e^{\frac{a_2}{b}}\right)\right)} \cdot \exp\left(-e^{-\left(\left(\frac{z}{b}\right) - \ln\left(e^{\frac{a_1}{b}} + e^{\frac{a_2}{b}}\right)\right)}\right).$$

This means that *Z* is distributed as if

$Z = bX + b \cdot \ln\left(e^{\frac{a_1}{b}} + e^{\frac{a_2}{b}}\right)$, where *X* is standard Gumbel.[56]

[55] Proof:

$$F_Z(z) = P(X_1 \le z \wedge X_2 \le z) = \int_{x_1=-\infty}^{x_1=z} \int_{x_2=-\infty}^{x_2=z} f_{X_1,X_2}(x_1, x_2)\, dx_2 dx_1 = \int_{x_1=-\infty}^{x_1=z} \int_{x_2=-\infty}^{x_2=z} f_{X_1}(x_1) f_{X_2}(x_1)\, dx_2 dx_1 = \int_{x_1=-\infty}^{x_1=z} f_{X_1}(x_1)\, dx_1 \cdot \int_{x_2=-\infty}^{x_2=z} f_{X_2}(x_2) dx_2 =$$

$$= F_{X_1}(z) \cdot F_{X_2}(z) = \exp\left(-e^{-\left(\frac{z-a_1}{b}\right)} - e^{-\left(\frac{z-a_2}{b}\right)}\right) = \exp\left(-e^{-\left(\frac{z}{b}\right)}\left(e^{\frac{a_1}{b}} + e^{\frac{a_2}{b}}\right)\right) = \exp\left(-e^{-\left(\frac{z}{b}\right) + \ln\left(e^{\frac{a_1}{b}} + e^{\frac{a_2}{b}}\right)}\right) = \exp\left(-e^{-\left(\left(\frac{z}{b}\right) - \ln\left(e^{\frac{a_1}{b}} + e^{\frac{a_2}{b}}\right)\right)}\right).$$

[56] Note that if *X* is standard Gumbel distributed, then $X = \left(\frac{Z}{b}\right) - \ln\left(e^{\frac{a_1}{b}} + e^{\frac{a_2}{b}}\right) \Leftrightarrow Z = \beta X + \ln\left(e^{\frac{a_1}{b}} + e^{\frac{a_2}{b}}\right)$.

5. Applying property four to $Z = \max\left(Y_1, Y_2,..,Y_N\right)$, where $Y_i = a_i + bX_i$ and X_i is iid standard Gumbel yields:[57]

$$F_Z\left(z\right) = \exp\left(- e^{-\left(\left(\frac{z}{b}\right) - \ln\left(\sum_{i=1}^{N} e^{\frac{a_i}{b}}\right)\right)}\right).$$

This means that Z is distributed as $Z = \beta X + \ln\left(\sum_{i=1}^{N} e^{\frac{\alpha_i}{\beta}}\right)$, where X is standard Gumbel.[58]

6. $E_X\left(F_g\left(\frac{X+a}{\sigma}\right)\right) = \frac{1}{1+e^{-\frac{a}{\sigma}}} = \frac{e^{\frac{a}{\sigma}}}{1+e^{\frac{a}{\sigma}}}$, where $\frac{X}{\sigma}$ is standard Gumbel distributed and $F_g\left(\right)$ is the cdf of standard Gumbel.[59]

7. $E_X\left(F_g\left(\frac{X+a}{\sigma}\right)\cdot F_g\left(\frac{X+b}{\sigma}\right)\right) = \frac{1}{1+e^{-\frac{a}{\sigma}}+e^{-\frac{b}{\sigma}}}$, where $\frac{X}{\sigma}$ is standard Gumbel distributed and $F_g\left(\right)$ is the cdf of standard Gumbel.[60]

---

[57] The proof for this property is straight forward:

$$F_Z\left(z\right) = P\left(X_1 \le z \wedge X_2 \le z \wedge .. \wedge X_N \le z\right) = \int_{x_1=-\infty}^{x_1=z}\int_{x_2=-\infty}^{x_2=z}\cdots\int_{x_N=-\infty}^{x_N=z}\cdots f_{X_1,X_2,..,X_N}\left(x_1, x_2,..,x_N\right)dx_N\cdots dx_2 dx_1 =$$

$$= \int_{x_1=-\infty}^{x_1=z} f_{X_1}\left(x_1\right)dx_1\cdot\int_{x_2=-\infty}^{x_2=z} f_{X_2}\left(x_2\right)dx_2\cdots\int_{x_N=-\infty}^{x_N=z} f_{X_2}\left(x_N\right)dx_N = F_g\left(\frac{z-a_1}{b}\right)\cdot F_g\left(\frac{z-a_2}{b}\right)\cdots F_g\left(\frac{z-a_N}{b}\right).$$

Applying rule 12 of MA1 yields:

$$F_Z\left(z\right) = \exp\left(-\sum_{i=1}^{N} e^{-\left(\frac{z-a_i}{b}\right)}\right) = \exp\left(-e^{-\left(\frac{z}{b}\right)}\cdot\sum_{i=1}^{N} e^{\frac{a_i}{b}}\right) = \exp\left(-e^{-\left(\left(\frac{z}{b}\right)-\ln\left(\sum_{i=1}^{N} e^{\frac{a_i}{b}}\right)\right)}\right).$$

[58] Note that $X = \left(\frac{Z}{\beta}\right) - \ln\left(\sum_{i=1}^{N} e^{\frac{a_i}{b}}\right) \Leftrightarrow Z = bX + b\ln\left(\sum_{i=1}^{N} e^{\frac{a_i}{b}}\right)$.

[59] Proof: $E_X\left(F_g\left(\frac{X+a}{\sigma}\right)\right) = \int_{z=-\infty}^{z=\infty}\exp\left(e^{-\frac{z+a}{\sigma}}\right)\cdot\frac{1}{\sigma}\cdot e^{-\frac{z}{\sigma}}\cdot\frac{1}{\sigma}\exp\left(e^{-\frac{z}{\sigma}}\right)dz = \int_{z=-\infty}^{z=\infty}\frac{1}{\sigma}\cdot\exp\left(e^{-\frac{z+a}{\sigma}}+e^{-\frac{z}{\sigma}}\right)\cdot e^{-\frac{z}{\sigma}}dz =$

$$= \int_{z=-\infty}^{z=\infty}\frac{1}{\sigma}\cdot\exp\left(e^{-\frac{z}{\sigma}}\cdot\left(1+e^{-\frac{a}{\sigma}}\right)\right)\cdot e^{-\frac{z}{\sigma}}dz = \frac{1}{\sigma}\cdot\int_{z=-\infty}^{z=\infty}\exp\left(e^{-\frac{z}{\sigma}+\ln\left(1+e^{-\frac{a}{\sigma}}\right)}\right)\cdot e^{-\frac{z}{\sigma}}dz = \frac{1}{\sigma}\cdot e^{-\ln\left(1+e^{-\frac{a}{\sigma}}\right)}\cdot...$$

$$...\cdot\int_{z=-\infty}^{z=\infty}\exp\left(e^{-\frac{z}{\sigma}+\ln\left(1+e^{-\frac{a}{\sigma}}\right)}\right)\cdot e^{-\frac{z}{\sigma}+\ln\left(1+e^{-\frac{a}{\sigma}}\right)}dz = \frac{1}{1+e^{-\frac{a}{\sigma}}}\cdot\int_{z=-\infty}^{z=\infty}\frac{1}{\sigma}\cdot f_g\left(\frac{z}{\sigma}-\ln\left(1+e^{-\frac{a}{\sigma}}\right)\right)dz = \frac{e^{\frac{a}{\sigma}}}{1+e^{\frac{a}{\sigma}}}.$$

8. $E_X\left(f_g\left(X+a\right)\right)=\dfrac{e^a}{\left(1+e^a\right)^2}$,

   where X is standard Gumbel and $f_g\left(\ \right)$ is the pdf of a standard Gumbel distributed random variable.[61]

9. $E_X\left(\prod_{i=1}^{n} f_g\left(X+c_i\right)\right)=n!\cdot\dfrac{e^{-\sum_{i=1}^{n} c_i}}{\left(1+\sum_{i=1}^{n} e^{-c_i}\right)^{n+1}}$,

   where X is standard Gumbel and $f_g\left(\ \right)$ is the pdf of standard Gumbel.[62]

---

[60] Proof: $E_X\left(X\cdot F_\xi\left(X+c\right)\right)=\int_{x=-\infty}^{x=\infty} x\cdot f_\xi\left(x\right)\cdot F_\xi\left(x+c\right)dx$. Applying rule 12 of Ma1 yields:

$$E_X\left(X\cdot F_\xi\left(X+c\right)\right)=\int_{x=-\infty}^{x=\infty} x\cdot f_\xi\left(x\right)\cdot F_\xi\left(x+c\right)dx=\frac{1}{1+e^{-\alpha c}}\cdot\int_{x=-\infty}^{x=\infty} x\cdot f_\xi\left(x-\frac{1}{\alpha}\cdot\ln\left(1+e^{-\alpha c}\right)\right)dx.$$

Substituting $z=x-\frac{1}{\alpha}\cdot\ln\left(1+e^{-\alpha c}\right)$ yields:

$$E_X\left(X\cdot F_\xi\left(X+c\right)\right)=\frac{1}{1+e^{-\alpha c}}\cdot\int_{x=-\infty}^{x=\infty}\left(z+\frac{1}{\alpha}\cdot\ln\left(1+e^{-\alpha c}\right)\right)\cdot f_\xi\left(z\right)dz=\frac{\lambda-\frac{1}{\alpha}\cdot\ln\left(1+e^{-\alpha c}\right)}{1+e^{-\alpha c}}.$$

[61] For a proof, see the next footnote.

[62] The proof is as follows: I start by rewriting the integral as follows:

$$\int_{x=-\infty}^{x=\infty}\prod_{i=0}^{n} f_g\left(X+c_i\right)dx=\int_{x=-\infty}^{x=\infty}\prod_{i=0}^{n} e^{-\left(x+c_i\right)}\cdot\exp\left(-e^{-\left(x+c_i\right)}\right)dx=\int_{x=-\infty}^{x=\infty}\prod_{i=0}^{n} e^{-\left(X+c_i\right)}\cdot F_g\left(X+c_i\right)dx=\int_{x=-\infty}^{x=\infty}\prod_{i=0}^{n} e^{-\left(X+c_i\right)}\cdot\prod_{i=0}^{n} F_g\left(X+c_i\right)dx,$$

with $c_0=0$. Applying rule 11 of MA1 and reformulating $\prod_{i=1}^{n} e^{-\left(X+c_i\right)}$ yields:

$$\int_{x=-\infty}^{x=\infty}\prod_{i=0}^{n} f_g\left(x+c_i\right)dx=e^{\sum_{i=1}^{n}-c_i}\cdot\int_{x=-\infty}^{x=\infty}\left(e^{-x}\right)^{n+1}\cdot F_g\left(x-\ln\left(\sum_{i=0}^{n} e^{-c_i}\right)\right)dx=\left(-1\right)^{n+1}\cdot e^{\sum_{i=1}^{n}-c_i}\cdot\int_{x=-\infty}^{x=\infty}\left(e^{-x}\right)^{n+1}\cdot\exp\left(-\exp\left(-x\right)\cdot\left(\sum_{i=0}^{n} e^{-c_i}\right)\right)dx$$

. Substituting $z=-\exp\left(-x\right)\cdot\sum_{i=0}^{n} e^{-c_i}$, $dz=\exp\left(-x\right)\cdot\left(\sum_{i=0}^{n} e^{-c_i}\right)dx$ and plugging in limits $z=-\infty,...,0$ yields:

$$\int_{x=-\infty}^{x=\infty}\prod_{i=0}^{n} f_g\left(x+c_i\right)dx=e^{\sum_{i=1}^{n}-c_i}\cdot\left(\sum_{i=0}^{n} e^{-c_i}\right)^{-\left(n+1\right)}\cdot\int_{z=-\infty}^{z=0} z^n\cdot\exp\left(z\right)dx$$

. The integral can be computed by integrating by parts:

$$g_n=\int_{z=-\infty}^{z=0} z^n\cdot\exp\left(z\right)dx=\left(z^n\cdot\exp\left(z\right)\right)_{z=-\infty}^{z=0}-n\cdot\int_{z=-\infty}^{z=0} z^{n-1}\cdot\exp\left(z\right)dx=-n\cdot\int_{z=-\infty}^{z=0} z^{n-1}\cdot\exp\left(z\right)dx=-n\cdot g_{n-1}$$

. By use of iteration and $g_1=\int_{z=-\infty}^{z=0}\exp\left(z\right)dx=1$, the integral yields: $g_n=\int_{z=-\infty}^{z=0} z^n\cdot\exp\left(z\right)dx=\left(-1\right)^n\cdot n!$.

10. $E_X\left(X \cdot F_\xi\left(X+c\right)\right) = \frac{1}{1+e^{-\alpha c}} \cdot \frac{\gamma + \beta + \ln\left(1+e^{-\alpha c}\right)}{\alpha}$, where X is distributed as $X \sim f\left(x\right) = \alpha \cdot e^{-\alpha x+\beta} \cdot e^{-e^{-\alpha x+\beta}}$, γ is the Euler constant $\gamma = 0.577...$ and $F_\xi\left(x\right) = F_g\left(\alpha x - \beta\right) = \exp\left(e^{-\alpha x+\beta}\right)$.[63]

11. $\prod_{i=1}^{J} F_X\left(X+c_i\right) = F_X\left(X - \frac{1}{\alpha} \cdot \ln\left(\sum_{i=1}^{J} e^{-\alpha \cdot c_i}\right)\right)$, where X is distributed as $X \sim F_\xi\left(x\right) = \exp\left(e^{-\alpha x+\beta}\right)$.[64]

12. $f_X\left(X+c_1\right) \cdot F_X\left(X+c_2\right) = \frac{e^{-\alpha \cdot c_1}}{e^{-\alpha \cdot c_1} + e^{-\alpha \cdot c_2}} \cdot f_X\left(x - \frac{1}{\alpha} \cdot \ln\left(e^{-\alpha \cdot c_1} + e^{-\alpha \cdot c_2}\right)\right)$, where $F_\xi\left(x\right) = \exp\left(e^{-\alpha x+\beta}\right)$ and $f_\xi\left(x\right) = F'_\xi\left(x\right)$.[65]

---

[63] Proof: $E_X\left(X \cdot F_\xi\left(X+c\right)\right) = \int_{x=-\infty}^{x=\infty} x \cdot f_\xi\left(x\right) \cdot F_\xi\left(x+c\right) dx$. Applying rule 12 of MA1 yields:

$$E_X\left(X \cdot F_\xi\left(X+c\right)\right) = \int_{x=-\infty}^{x=\infty} x \cdot f_\xi\left(x\right) \cdot F_\xi\left(x+c\right) dx = \frac{1}{1+e^{-\alpha c}} \cdot \int_{x=-\infty}^{x=\infty} x \cdot f_\xi\left(x - \frac{1}{\alpha} \cdot \ln\left(1+e^{-\alpha c}\right)\right) dx.$$

Substituting $z = x - \frac{1}{\alpha} \cdot \ln\left(1+e^{-\alpha c}\right)$ yields:

$$E_X\left(X \cdot F_\xi\left(X+c\right)\right) = \frac{1}{1+e^{-\alpha c}} \cdot \int_{x=-\infty}^{x=\infty} \left(z + \frac{1}{\alpha} \cdot \ln\left(1+e^{-\alpha c}\right)\right) \cdot f_\xi\left(z\right) dz = \frac{\frac{\lambda+\beta}{\alpha} + \frac{1}{\alpha} \cdot \ln\left(1+e^{-\alpha c}\right)}{1+e^{-\alpha c}}.$$

[64] Proof: $\prod_{i=1}^{J} F_X\left(X+c_i\right) = \prod_{i=1}^{J} \exp\left(-e^{-\left(X+c_i\right)}\right) = \exp\left(-\sum_{i=1}^{J} e^{-\left(X+c_i\right)}\right) = \exp\left(-e^{-X} \cdot \sum_{i=1}^{J} e^{c_i}\right) = \exp\left(-e^{-X} \cdot \sum_{i=1}^{J} e^{c_i}\right) =$

$$= \exp\left(-e^{-\left(X - \ln\left(\sum_{i=1}^{J} e^{c_i}\right)\right)}\right) = F_X\left(X - \ln\left(\sum_{i=1}^{J} e^{c_i}\right)\right).$$

[65] Proof: $f_\xi\left(X+c_1\right) \cdot F_\xi\left(X+c_2\right) = e^{-\alpha \cdot \left(x+c_1\right)+\beta} \cdot \exp\left(-e^{-\alpha \cdot \left(x+c_1\right)+\beta}\right) \cdot F_\xi\left(X+c_2\right) = e^{-\alpha \cdot \left(x+c_1\right)+\beta} \cdot F_\xi\left(X+c_1\right) \cdot F_\xi\left(X+c_2\right)$.

Applying rule 11 of MA1 yields:

$$f_\xi\left(X+c_1\right) \cdot F_\xi\left(X+c_2\right) = e^{-\alpha \cdot \left(x+c_1\right)+\beta} \cdot F_\xi\left(X - \frac{1}{\alpha} \cdot \ln\left(e^{-c_1} + e^{-c_1}\right)\right).$$

Reformulating yields:

$$f_\xi\left(X+c_1\right) \cdot F_\xi\left(X+c_2\right) = e^{-\alpha \cdot \left(x+c_1\right)+\beta} \cdot F_\xi\left(X - \frac{1}{\alpha} \cdot \ln\left(e^{-\alpha \cdot c_1} + e^{-\alpha \cdot c_2}\right)\right) =$$

$$= e^{-\alpha \cdot \left(c_1 + \frac{1}{\alpha}\ln\left(e^{-c_1} + e^{-c_2}\right)\right)} \cdot e^{-\alpha \cdot \left(X - \frac{1}{\alpha}\ln\left(e^{-c_1} + e^{-c_2}\right)\right)+\beta} \cdot F_\xi\left(X - \frac{1}{\alpha} \cdot \ln\left(e^{-\alpha \cdot c_1} + e^{-\alpha \cdot c_2}\right)\right) = \frac{e^{-\alpha \cdot c_1}}{e^{-\alpha \cdot c_1} + e^{-\alpha \cdot c_2}} \cdot f_\xi\left(x - \frac{1}{\alpha} \cdot \ln\left(e^{-\alpha \cdot c_1} + e^{-\alpha \cdot c_2}\right)\right).$$

13. Theorem: (Densities of transformed random variables)[66]

If $X = (X_1, ..., X_k)$ is a random vector with density f_X, and if $Y_i = h_i(X_1, ..., X_k)$, for $i = 1,...,k$ such that

1. $h_1,...,h_k$ is continuous;
2. for every $x \in \mathbb{R}^k$, such that $y_i = h_i(x)$ for all $i = 1,...,k$. We write then $x_i = l_i(y)$, $i = 1,...,k$. $l = h^{-1}$ can also denoted "inverse function of *h*", where $l = (l_1, ..., l_k)$ and $h = (h_1, ..., h_k)$;
3. derivatives $\partial x_i / \partial y_i$ exist and are continuous;

then $Y_i = (Y_1, ..., Y_k)$ has density

$$f_Y(y) = f_X(l_1(y), ..., l_k(y)) \cdot |J(y)|,$$

where $|J(y)| = \det \begin{bmatrix} \frac{\partial x_1}{\partial y_1} & \cdots & \frac{\partial x_1}{\partial y_k} \\ \cdots & \cdots & \cdots \\ \frac{\partial x_k}{\partial y_1} & \cdots & \frac{\partial x_k}{\partial y_k} \end{bmatrix}$.

[66] See Shao (2003: 23).

MA 2 Minimum consumption threshold of driving distance

In this section I examine the sign of the influence of parameters d and a_2 on the minimum consumption threshold of good two $x_2(\varsigma_2)$. I will show that the impact on both d and a_2 on $x_2(\varsigma_2)$ is positive. It was not possible to prove the sign of these impacts formally. I therefore show graphically that these signs are as mentioned for a range d and a_2 conditional on economic variables y, k_2, p_1 and p_2.

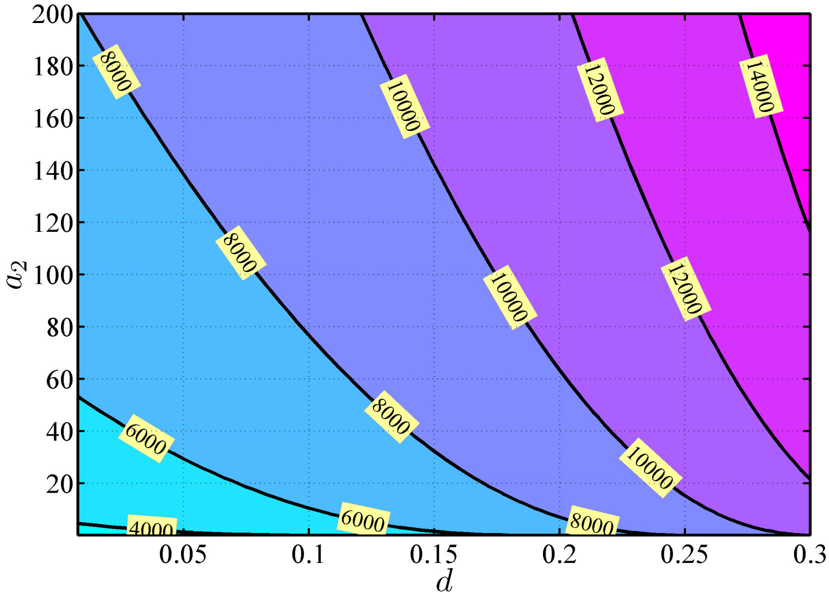


Figure MA2.1: Minimum consumption demand for different parameters d and a_2.

This diagram shows clearly that the minimum consumption threshold of good two $x_2(\varsigma_2)$ increases in both d and a_2. This diagram was plotted given values of the economic variables that are rather typical, namely given an income y = CHF 60,000, fixed costs of CHF 7,000 and marginal driving costs of p_2 = 0.3 CHF / km. I also tested this relation for a large range of economic variables and in all cases, $x_2(\varsigma_2)$ depended positively on both d and a_2.

I next show the values in the range of d and a_2 which comprises the optimum values.

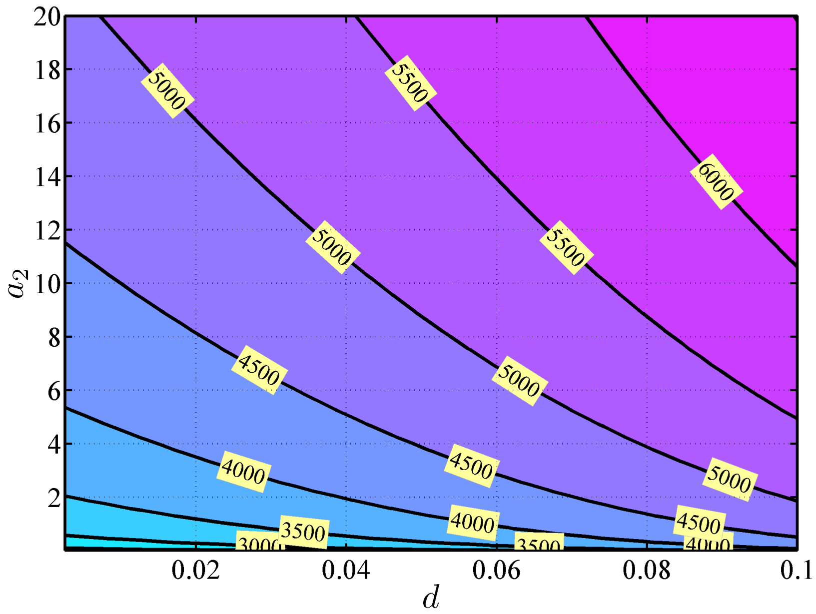


Figure MA2.2: Minimum consumption demand for different parameters d and a_2, range of optimal values for d and a_2.

This diagram shows that for all optimal values $(d, a_2) = \{(0.01, 20), (0.02, 20), (0.005, 10), (0.1, 2)\}$ – see Table A3.14.1 – the critical level $x_2(\varsigma_2)$ ranges from 4,200 to 5,000 km, given the values of the economic variables as mentioned. This range seems to be quite reasonable since the average kilometre cost given an annual driving distance of 5,000 km would be 1.70 CHF / km.[67] At this price, the cost of a taxi service is not much higher and the use of public transportation services is a much cheaper alternative.

In this section, I do not examine the impact of changes of economic variables on the critical level $x_2(\varsigma_2)$, since these effects are not relevant to the estimation routine.

[67] This value I computed as the total driving costs $k_2 + p_2 \cdot x_2(\zeta_2) = 7000 + 0.3 \cdot 5{,}000 = 8{,}500$ divided by the total annual kilometres driven $8{,}500/5{,}000 = 1.7$.

MA 3 Marginal effects of Probit and Tobit models

In the following I compute the marginal effects of variables on the probability of the outcome to be zero $P(X=0)$ and in the case of the Tobit model the expectation value $E(X=0)$. I require these results to compute the marginal effect for each household. I will then compute the average marginal effects of the households. I need to do this, since the software Stata only computes the marginal effects at the sample mean. Further, in the case of panel data with random effects, the software Stata computes the marginal effects at the sample mean, assuming all random effects to be zero. Since the variance of the random effects are more than double the variance of the error terms that are iid across the complete sample, a different procedure is needed to compute the average marginal effect.

I only compute the marginal effects for the case of the Tobit model. I will show that the marginal effect of the Probit model will be identical. The Tobit model is defined by the following density function:

$$f(z)=\left\{\begin{array}{l} z=0:\Phi\left(-\frac{x\beta}{\sigma}\right) \\ z\geq 0:\phi\left(\frac{z-x\beta}{\sigma}\right)\end{array}\right\}. \tag{MA 3.1}$$

The variable y is the dependent variable, parameters β and σ are to be estimated and variables x are explanatory variables. Note that x usually contains a column with ones.

I start by computing the marginal effects of the probability z being zero, $\Phi\left(-\frac{x\beta}{\sigma}\right)$:[68]

$$\frac{\partial\Phi\left(-\frac{x\beta}{\sigma}\right)}{\partial x_i}=\frac{\partial\Phi\left(-\frac{x\beta}{\sigma}\right)}{\partial\left(-\frac{x\beta}{\sigma}\right)}\cdot\frac{\partial\Phi\left(-\frac{x\beta}{\sigma}\right)}{\partial x_i}=-\phi\left(\frac{x\beta}{\sigma}\right)\cdot\frac{\beta_i}{\sigma}. \tag{MA 3.2}$$

Computing the marginal effects for the expectation value

$$E(z)=\int_{z=0}^{z=\infty} z\cdot\phi\left(\frac{z-x\beta}{\sigma}\right)dz \tag{MA 3.3}$$

[68] This formula was derived as follows: $\frac{\partial\Phi\left(-\frac{x\beta}{\sigma}\right)}{\partial x_i}=\frac{\partial\Phi\left(-\frac{x\beta}{\sigma}\right)}{\partial\left(-\frac{x\beta}{\sigma}\right)}\cdot\frac{\partial\Phi\left(-\frac{x\beta}{\sigma}\right)}{\partial x_i}=-\phi\left(-\frac{x\beta}{\sigma}\right)\cdot\frac{\beta_i}{\sigma}$. This formula can also be found in Greene (2003: 764). Note that $\varphi(-z)=\varphi(z)$.

by use of integration by parts yields the formula:[69]

$$\frac{\partial E(Z \mid x\beta, \sigma)}{\partial x_i} = \frac{\partial \int\limits_{z=0}^{z=\infty} z \cdot \phi\left(\frac{z - x\beta}{\sigma}\right) dz}{\partial x_i} = \beta_i \cdot \Phi\left(\frac{z - x\beta}{\sigma}\right). \tag{MA3.4}$$

Since the Probit model is defined as

$$f(z) = \begin{Bmatrix} z = 0: \ \Phi\left(-\frac{x\beta}{\sigma}\right) \\ z = 1: \ 1 - \Phi\left(-\frac{x\beta}{\sigma}\right) \end{Bmatrix}, \tag{MA 3.5}$$

it follows that the marginal effect of changes of explanatory variables x is the same as for the case of the Probit model, see (MA 3.2).

Further I will compute the expectation value of the dependent variable for a Tobit model. This formula is used for computing the average expectation values of the dependent variable of each household.[70]

$$E(z) = \int\limits_{z=0}^{z=\infty} z \phi\left(\frac{z - x\beta}{\sigma}\right) dz = x\beta \cdot \Phi\left(\frac{x\beta}{\sigma}\right) + \sigma \cdot \phi\left(\frac{x\beta}{\sigma}\right). \tag{MA 3.6}$$

---

[69] This formula was derived as follows:

$$\frac{\partial \int\limits_{z=0}^{z=\infty} z \cdot \phi\left(\frac{z - x\beta}{\sigma}\right) dz}{\partial x_i} = \frac{\partial \left[z \cdot \Phi\left(\frac{z - x\beta}{\sigma}\right)\right]_{z=0}^{z=\infty} - \int\limits_{z=0}^{z=\infty} \Phi\left(\frac{z - x\beta}{\sigma}\right) dz}{\partial x_i} = -\frac{\partial \int\limits_{z=0}^{z=\infty} \Phi\left(\frac{z - x\beta}{\sigma}\right) dz}{\partial x_i} = -\int\limits_{z=0}^{z=\infty} \frac{\partial \Phi\left(\frac{z - x\beta}{\sigma}\right)}{\partial x_i} dz = -\int\limits_{z=0}^{z=\infty} \phi\left(\frac{z - x\beta}{\sigma}\right) \cdot \frac{\beta_i}{\sigma} dz =$$

$$= -\frac{\beta_i}{\sigma} \cdot \int\limits_{q=\frac{x\beta}{\sigma}}^{q=\infty} \phi(q) dq = -\frac{\beta_i}{\sigma} \cdot \sigma \cdot \left(1 - \Phi\left(\frac{z - x\beta}{\sigma}\right)\right) = \beta_i \cdot \left(1 - \Phi\left(\frac{z - x\beta}{\sigma}\right)\right) = \beta_i \cdot \Phi\left(\frac{z - x\beta}{\sigma}\right)$$

. Note that this formula can also be found in Greene (2003: 763).

[70] This formula was derived as follows:

$$E(z) = \int\limits_{z=0}^{z=\infty} z \phi\left(\frac{z - x\beta}{\sigma}\right) dz = \int\limits_{q=-\frac{x\beta}{\sigma}}^{q=\infty} (\sigma q + x\beta) \cdot \phi(q) dq = x\beta \cdot \left(1 - \phi\left(-\frac{x\beta}{\sigma}\right)\right) + \frac{\sigma}{\sqrt{2\pi}} \cdot \int\limits_{q=\frac{x\beta}{\sigma}}^{q=\infty} q \cdot e^{-\frac{q^2}{2}} dq = x\beta \cdot \phi\left(\frac{x\beta}{\sigma}\right) - \ldots$$

$$\ldots - \sigma \cdot \left[\frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{q^2}{2}}\right]_{q=-\frac{x\beta}{\sigma}}^{q=\infty} = x\beta \cdot \Phi\left(\frac{x\beta}{\sigma}\right) + \sigma \cdot \left[\phi(q)\right]_{q=-\frac{x\beta}{\sigma}}^{q=\infty} = x\beta \cdot \Phi\left(\frac{x\beta}{\sigma}\right) + \sigma \cdot \phi\left(-\frac{x\beta}{\sigma}\right) = x\beta \cdot \Phi\left(\frac{x\beta}{\sigma}\right) + \sigma \cdot \phi\left(\frac{x\beta}{\sigma}\right).$$

Note that this formula can also be found in Greene (2003: 763).

MA 4 Proof of correction term of Discrete-Continuous choice model

MA 4.1 Introduction

In the paper by Dubin and McFadden "An econometric analysis of residential electric appliance holdings and consumption" (1984), the so-called "Discrete-Continuous choice model" was presented for the first time. This model captures a joint decision of deciding on one type of capital good and the intensity of using this capital good. Examples of such decisions are the choice of type of heating system and then the choice of room temperature that will define the energy costs as examined by Dubin and McFadden (1984) or the choice of car type and the annual mileage driven. In both cases, the choice of the type of capital good will influence the quality of service this good will provide and the marginal cost of its use. There is therefore an interaction between these two decisions. From a researcher's perspective, it is often interesting to explain the use decision, e.g. the total annual distance driven by households. The key issue of the model framework of Dubin and McFadden is how to get unbiased estimates of the parameters incorporated in the function that define the sign and magnitude of the impact of certain factors on the intensity of use of the capital good, e.g. the driving distance. The solution Dubin and McFadden found for solving this problem is to compute a correction term that can be added to the demand equation used to estimate the parameters. The genuine property of the model framework is that this correction term can be computed using the probabilities with which a household chooses the different types of capital goods. These probabilities can be computed by solving a simple multinomial choice model. So far I have not found any description in the literature why the functional form of this correction term is exactly how it is proposed by Dubin and McFadden. Further, I have not found any precise information on what kind of missing variables the error terms of their model actually capture. The key question is whether the error terms only capture missing socio-demographic variables of the households or only missing variables describing the properties of the capital goods, or both. It is important to answer this question since it will help to decide whether the model will yield unbiased results when using a certain dataset.

In the following, I first describe what Dubin and McFadden assume on household behaviour when they decide on the choice and use of a type of capital good. I then present the microeconomic demand system that maps this behaviour. Next, I show on what assumption Dubin and McFadden derived their assumptions on the error terms. I will then state the assumptions on the error terms (MA4.1.8). In Subchapter MA4.2 I will derive the correction term for the case of two goods. Further, I will prove that the assumptions on the error terms made by Dubin and McFadden can be derived from assumption (MA4.1.8). In Subchapter MA4.2 I will derive the correction term for the general case where a household can choose from several different types of the capital good.

Introduction of the Discrete-Continuous Choice Model of Dubin and MCFadden

In this paragraph, I first outline the household behaviour on which the model of Dubin and McFadden is based. I will then present their concrete microeconomic demand system, which should map this behaviour.

I start by describing how the interaction of the choice of a type of capital good and its use is captured by the model of Dubin and McFadden. This interaction is covered by assuming that a household computes the utility it could reach by holding and optimally using each type of capital good. It will then decide to choose the capital good for which this utility is highest. This decision corresponds to a microeconomic decision conditional on a type of capital good i:

$$u_i\left(x_1, x_{2,i}\right), \text{ s.t. } y - k_{2,i} = p_1 \cdot x_1 + p_{2,i} \cdot x_{2,i}. \qquad \text{(MA4.1.1)}$$

Good two is the capital good. The related fixed and marginal costs are denoted by $k_{2,i}$ and $p_{2,i}$, respectively, and $x_{2,i}$ denotes the intensity of use, e.g. the annual distance driven when researching car choice and use. Index *i* denotes the type of capital good, e.g. the car type when researching car choice and use. Good one is a composite good containing all goods but the capital good.

The household will then compute the optimal amounts consumed $\left(x_1^*, x_{2,i}^*\right)$. The optimal choice of $\left(x_1^*, x_{2,i}^*\right)$ maximizes utility u_i for the budget given $y - k_{2,i} = p_1 \cdot x_1 + p_{2,i} \cdot x_{2,i}$. Note that the available budget is given by income *y* minus fixed costs $k_{2,i}$. Since the fixed costs reduce the available budget, they also reduce the maximum utility. From this utility maximization conditional on capital good of type *i*, the corresponding utility levels yield u_i^*. These utility levels are equal to the value of the indirect utility function

$$u_i^* = v_i\left(p_1, p_{2,i}, y\right). \qquad \text{(MA4.1.2)}$$

The household will choose option *i* for which u_i^* reaches the highest value. Since we are not only interested in what choice *i* the household makes but also in the intensity of use $x_{2,i}^*$, I will now describe the interrelationship between the choice and the use of the capital good. This optimal value of use $x_{2,i}^*$ can be computed by the Marshallian demand function

$$x_{2,i}^* = x_{2,i}\left(p_1, p_{2,i}, y\right). \qquad \text{(MA4.1.3)}$$

Definition of the microeconomic demand system

One important feature of a microeconomic demand system is that the Marshallian demand function is linked to the indirect utility function by Roy's identity:

$$x_{2,i}\left(p_1, p_{2,i}, y\right) = -\frac{\partial v_i\left(p_1, p_{2,i}, y\right) / \partial p_{2,i}}{\partial v_i\left(p_1, p_{2,i}, y\right) / \partial y}. \tag{MA4.1.4}$$

This indicates that the Marshallian demand function cannot be chosen randomly since it directly follows from the indirect utility function according to this identity. Dubin and McFadden (1984) make the following assumptions on the demand system with a linear Marshallian demand function and its corresponding indirect utility function:[71]

$$v_i = e^{-\beta_i p_{2,i}} \cdot \left(\frac{\alpha_i}{\beta_i} + \beta_i \cdot \left(y - r_i\right) + \alpha p_{2,i} + \gamma_i s + \delta b_i\right) + \varepsilon_i \cdot e^{-\beta_i p_{2,i}} + \vartheta_i\,, \tag{MA4.1.5}$$

$$x_{2,i} = \alpha_i p_{2,i} + \beta_i\left(y - k_i\right) + \gamma_i s + \delta b_i + \varepsilon_i\,. \tag{MA4.1.6}$$

---

[71] This model is similar to that stated in Dubin and McFadden (1984) on page 349. Note that here good two is assumed to be the good related to the use of a capital good of type *i*, and good one is assumed to be the numeraire good with price one, whereas Dubin and McFadden assume good two to be related to the use of a capital good of type *i* and do not normalize to one the price of good one.

At this point I do not wish to present the complete proof of how the indirect utility function v_i can be derived from assuming a linear Marshallian demand function, but I would like to outline it.

In the first step, a function $y\left(p_2\right)$ shall be defined that relates any price p_2 to a budget *y*, such that the same utility level u_0 can be achieved when the utility maximization problem is solved. The starting point is to use a parametrized form $\left(y(t), p_2(t)\right)$: $u_0 = v\left(y, p_2\right) = v\left(y(t), p_2(t)\right)$.

The total integral yields $0 = \left[\partial v\left(y(t), p_2(t)\right) / \partial y(t)\right] \cdot \partial y(t)/\partial t + \left[\partial v\left(y(t), p_2(t)\right) / \partial\, p_2(t)\right] \cdot \partial p_2(t)/\partial t$. Reformulating and applying Roy's identity yields $-\frac{\partial v\left(y(t), p_2(t)\right) / \partial\, p_2(t)}{\partial v\left(y(t), p_2(t)\right) / \partial\, y(t)} = x_2\left(y, p_2\right) = \alpha p_2 + \beta y\left(p_2\right) + k = \frac{\partial y\left(p_2\right)}{\partial p_2}$. Solving this linear inhomogeneous differential equation yields $y\left(p_2\right) = c \cdot e^{\beta p_2} - \frac{1}{\beta}\left(\alpha \cdot p_2 + \frac{\alpha}{\beta} + k\right)$, where *c* is a constant that can be set to an arbitrary value. Setting $c = u_0 = v\left(y, p_2\right)$ and solving for $v\left(y, p_2\right)$ yields $v\left(y, p_2\right) = \left(y\left(p_2\right) + \frac{1}{\beta}\left(\alpha \cdot p_2 + \frac{\alpha}{\beta} + k\right)\right) \cdot e^{-\beta p_2}$.

Since the indirect utility function is defined up to any positive transformation, multiplying by β and adding ϑ yields $v\left(y, p_2\right) = \left(\frac{\alpha}{\beta} + \alpha \cdot p_2 + \beta \cdot y + k\right) \cdot e^{-\beta p_2} + \vartheta$.

Plugging in $k = \gamma_i s + \delta b_i + \varepsilon_i$ yields $v\left(y, p_2\right) = \left(\frac{\alpha}{\beta} + \alpha \cdot p_2 + \beta y + \gamma s + \delta b + \varepsilon_i\right) \cdot e^{-\beta p_2} + \vartheta$.

The outline of this proof can be found in Hausman (1981) on page 668. In Hausman (1981) other functional forms of the Marshallian demand function can also be found.

Here, s are socio-demographic variables of the household and b_i are observed car attributes. Note that the price of good one is set as numeraire. The random variable ε_i contains unobserved socio-demographic variables and unobserved car attributes. For the case of car type choice and use, unobserved car attributes could be car space, interior noise level, comfort features or car quality. Relevant unobserved households attributes could be the preference for car driving or a disability that prevents one member of the household from using public transportation. The error term ϑ_i accounts for the fact that the impact of an observed or unobserved variable relative to another variable can be different on car choice than on car use. For instance, consider the variable "car space". We could assume that car space has a positive impact on the probability to choose a certain car type. In other words, the larger the car space, the higher the deterministic part of the indirect utility function: the corresponding component of vector δ would be positive. By the functional form of the Marshallian demand function, the impact on driving demand will also be positive (MA4.1.6). In this case, this might be wrong: if the car space is small, for some type of shopping one would need to drive twice to the shopping mall to transport all of the goods home. For this case, therefore, even the sign of the parameter in the Marshallian demand function may be wrong. Therefore, this error has to be compensated by the extra error term added at the indirect utility function. A more general explanation is that the extra error term ϑ_i has to be added due to differences in the relative impact of observed or unobserved factors on car demand and car choice. Note that the same may hold for household-specific variables. For instance, the household location – e.g. city versus rural area – should strongly affect driving demand but it may not have a very strong influence on the choice of car type. Note that the error term ε is generally not independent from the error term ϑ_i.[72]

For simplicity, Dubin and McFadden replace expression $\varepsilon_i \cdot e^{-\beta p_{2,i}} + \vartheta_i$ by a random variable ξ_i.

$$v_i = e^{-\beta_i p_{2,i}} \cdot \left(\frac{\alpha_i}{\beta_i} + \beta_i \cdot (y - r_i) + \alpha_i p_{2,i} + \gamma_i s + \delta b_i \right) + \xi_i . \qquad \text{(MA4.1.7)}$$

This means that they assume that random variable ξ_i does not depend on the marginal cost $p_{2,i}$. This assumption is feasible when assuming that the contribution of $p_{2,i}$ to the variation of $\varepsilon_i \cdot e^{-\beta p_{2,i}} + \vartheta_i$ is negligible. At least for the case of car type choice and car use, this simplification seems reasonable, since the marginal costs of driving p_{2i} do not vary much between households. The reason for this is that the variations in marginal costs are caused only by changes in fuel price. Since in most datasets these prices do not vary much, neither between different regions the observed households live nor over time, the variation of $p_{2,i}$ can be regarded as very small and therefore this simplification seems to be feasible.

[72] One example can be derived from the example above: assume that car space is unobserved and is the only unobserved variable. Then its impact on driving demand would be negative and therefore so would be ε_i. Then the random variable ϑ_i needs to be positive, since car space would positively affect the probability of choosing this car type.

Assumptions on the error terms of the demand system

Dubin and McFadden make several assumptions on the relation between the error terms,[73] but give no information on what these assumptions are based and why some of them are necessary. In the following, I show that these assumptions are compatible with the assumption of a linear relationship between the error terms of the choice model ξ_i and the error terms of the Marshallian demand ε_i:

$$\varepsilon_i = \sum_{j=1}^{J} a_{ij} \cdot \xi_j + \kappa_i \,, \tag{MA4.1.8}$$

with $\xi_1, \xi_2, \xi_3, ..., \xi_J$ iid. with $F(x) = \exp\left(-e^{-h_1 \cdot x + h_2}\right)$, $h_1 = \frac{\pi}{\lambda\sqrt{3}}$, $h_2 = -\gamma$, where $\gamma = 0.577...$ is the Euler constant. From the assumptions on parameters h_1 and h_2 follows $E\left(\xi_j\right) = 0$ and $\text{var}\left(\xi_j\right) = \lambda^2/2$.

The random variable κ_i is independent of ξ_j for all j=1,...,J and a_j are constants. I assume that $\text{E}\left(\kappa_i\right) = 0$ and from this follows $\text{E}\left(\varepsilon_i\right) = 0$.

The key issue of this model is that observing the choice of a household implies that some information on the error terms $\xi = \left\{\xi_j\right\}_{j=1,...,J}$ is revealed. This can be seen when looking at the choice decision. As previously mentioned, the household will choose the capital good that yields the highest utility. The index of the good chosen is indicated by *j*.

$$i = \arg\max_j v_j \,. \tag{MA4.1.9}$$

For the chosen good *i*, the following condition holds:

$$v_i \geq v_j \quad \forall j = 1, ..., J \,. \tag{MA4.1.10}$$

This condition is equivalent to

[73] The conditions stated in Dubin and McFadden are the following:

a.) The stochastic terms ξ_i, $i = 1, ..., J$, are independent and identically Gumbel distributed:

$$F\left(\xi_i\right) = \exp\left(-\exp\left(-\frac{\xi_i \pi}{\lambda\sqrt{3}} - \gamma\right)\right).$$

b.) $E\left[\varepsilon_i \mid \xi_1, ..., \xi_J\right] = \frac{\sigma\sqrt{2}}{\lambda} \cdot \sum_{j=1}^{J} R_{ij} \xi_j$, $R_{ij} = \text{corr}\left(\varepsilon_i, \xi_j\right)$, $E\left[\varepsilon_i\right] = 0$ and $\sigma^2 = \text{var}\left[\varepsilon_i\right]$.

c.) The conditional variance of ε_i given $\xi_1, ..., \xi_J$ is $\text{var}\left[\varepsilon_i \mid \xi_1, ..., \xi_J\right] = \sigma^2 \left(1 - \sum_{j=1}^{J} R_{ij}^2\right)$.

d.) The correlation between ε_i and ξ_j, $R_{ij} = \text{corr}\left(\varepsilon_i, \xi_j\right)$, fulfils $\sum_{j=1}^{J} R_{ij}^2 < 1$ and $\sum_{j=1}^{J} R_{ij} = 0$.

$$\xi_i \geq V_j - V_i + \xi_j \ \ \forall j = 1,...,J \ , \tag{MA4.1.11}$$

with $V_j = e^{-\beta p_{2,j}} \cdot \left(\frac{\alpha_j}{\beta_j} + \beta_j \cdot (y - r_i) + \alpha_j p_{2,i} + \gamma_i s + \delta b_i \right)$.

Therefore, given choice *i*, the random terms ξ_j are no longer independent and, more importantly, their expected values are no longer zero:

$$E\left(\xi_j \mid I = i\right) \neq 0, \ \ \forall j = 1,...,J \ , \tag{MA4.1.12}$$

where *I* is an indicator $I = i$, if $\xi_i \geq V_j - V_i + \xi_j \ \ \forall j = 1,...,J$.

As will be shown later, it follows from this that the expected value of ε_i conditional on the choice *i*,

$$E\left(\varepsilon_i \mid I = i\right) = \sum_{j=1}^{J} a_{ij} \cdot E\left(\xi_j \mid I = i\right) \neq 0 \tag{MA4.1.13}$$

is also no longer zero. One intuitive explanation for this fact is that in the example of car choice and use, a person with a strong preference for car driving – an unobserved variable – tends to buy a large comfortable car instead of a small car. Therefore, the choice of a large comfortable car indicates that this person rather has also a strong preference for car driving, such that it can be expected that the expected value conditional on the choice of a big car is likely to be positive: $E\left(\varepsilon_i \mid I = i\right) > 0$.

Researchers are often interested in estimating the parameters of the Marshallian demand function (MA 3.1.6). The most simple way to do this is by OLS. Since OLS will only provide unbiased estimators for the parameters if the expected value of the error term is zero, we have to correct for $E\left(\varepsilon_i \mid I = i\right)$. Note that we do not have to know $E\left(\varepsilon_i \mid I = j\right), i \neq j$, since only the use of the type of capital good that was chosen is observed and the model excludes the possibility that the household can choose more than one capital good.

The key advantage of the model by Dubin and McFadden is that $E\left(\varepsilon_i \mid I = i\right)$ can be expressed in closed form as

$$E\left(\varepsilon_i \mid I = i\right) = \frac{\sigma\sqrt{6}}{\pi} \cdot \left(\left(\sum_{\substack{j=1,...,J \\ j \neq i}} R_{ij} \cdot \frac{P_j}{1 - P_j} \cdot \ln\left(P_j\right) \right) - R_{ii} \cdot \ln\left(P_i\right) \right), \tag{MA4.1.14}$$

where $R_{ij} = \operatorname{corr}\left(\varepsilon_i, \xi_j\right)$ and P_i is the probability that the household chooses the capital good of type *i*.

In the following I prove that the correction term (MA4.1.13) is correct. This proof cannot yet be found in the literature. Note that understanding this proof will also help when the model shall be extended or the functional form of the Marshallian demand function shall be modified.

Before I start with the general proof for many goods, I present the proof for the most basic case where households can only choose between two types of capital goods in order to present the key elements of the proof in a more comprehensive way.

MA 4.2 Proof of the case with two goods

The model

The model can be stated as follows:

$$\varepsilon_1 = a_{11} \cdot \xi_1 + a_{12} \cdot \xi_2 + \kappa_1,$$
$$\varepsilon_2 = a_{21} \cdot \xi_1 + a_{22} \cdot \xi_2 + \kappa_2,$$

with ξ_1, ξ_2 having common density function $f_\xi(x) = \alpha \cdot e^{-h_1 x + h_2} \cdot e^{-e^{-h_1 x + h_2}}$ and (MA4.2.1)

$$I = \begin{cases} 1: & \xi_1 + V_1 \geq \xi_2 + V_2 \\ 2: & \xi_1 + V_1 < \xi_2 + V_2 \end{cases}, \qquad \text{(MA4.2.2)}$$

where V_1, V_2 can be considered as constants.

Expected value of ε_1 conditional on the choice of capital good one

In order to correct for the selection bias, the expression $E\left(\varepsilon_1 \mid I = 1\right)$ has to be computed. Using (MA4.2.1), this expression can be rewritten as follows:[74]

$$E\left(\varepsilon_1 \mid I = 1\right) = a_{11} \cdot E\left(\xi_1 \mid I = 1\right) + a_{12} \cdot E\left(\xi_2 \mid I = 1\right). \qquad \text{(MA4.2.3)}$$

Therefore, the expressions $E\left(\xi_1 \mid I = 1\right)$ and $E\left(\xi_2 \mid I = 1\right)$ have to be determined.

[74] Note that $E\left(\kappa_1 \mid I = 1\right) = 0$, since indicator I depends only on random terms ξ and constants V_1 and V_2. Since random terms ξ are independent of κ_1 and $E\left(\kappa_1\right) = 0$, $E\left(\kappa_1 \mid I = 1\right) = 0$.

Expected value of ξ_1 conditional on the choice of capital good one

The value of $E(\varepsilon_1 \mid I=1)$ can be computed as follows:

$$E(\xi_1 \mid I=1) = \int_{x \in S_1} x_1 \cdot f_{\xi \mid I=1}(x_1, x_2)\, dx = \int_{x \in S_1} x_1 \cdot \frac{f_\xi(x_1, x_2)}{P_1} dx, \tag{MA4.2.4}$$

where S_1 is the set of $\xi = \{\xi_1, \xi_2\}$ that complies to $I=1$, namely $\xi_1 + V_1 \geq \xi_2 + V_2$. This set can be illustrated as follows:

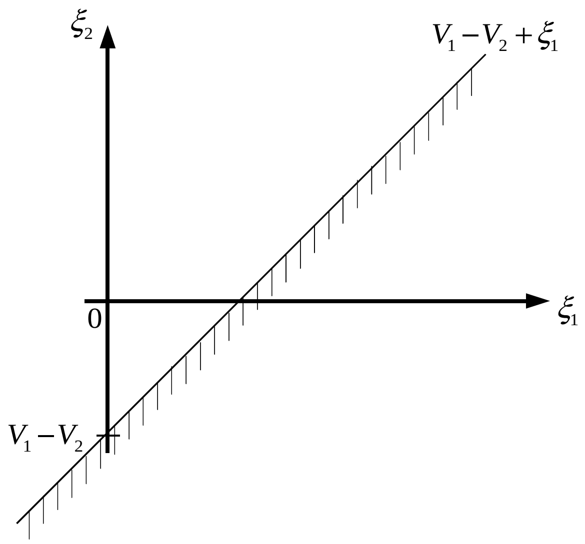


Figure MA4.2.1: Illustration of the space of random variables when option one is chosen.

The integral (MA4.2.4) can therefore be written as

$$E(\xi_1 \mid I=1) = \frac{1}{P_1} \cdot \int_{x_1=-\infty}^{x_1=\infty} \int_{x_2=-\infty}^{x_2=V_1-V_2+x_1} x_1 \cdot f_\xi(x_1, x_2)\, dx_2 dx_1 . \tag{MA4.2.5}$$

Since the two random variables ξ_1 and ξ_2 are independent, the common density function $f_\xi(x_1, x_2)$ can be written as $f_\xi(x_1, x_2) = f_\xi(x_1) \cdot f_\xi(x_2)$ and, therefore, integral (MA4.2.5) can be written as[75]

$$E(\xi_1 \mid I=1) = \frac{1}{P_1} \cdot \int_{x_1=-\infty}^{x_1=\infty} x_1 \cdot f_\xi(x_1) \cdot \int_{x_2=-\infty}^{x_2=V_1-V_2+x_1} f_\xi(x_2)\, dx_2 dx_1 . \tag{MA4.2.6}$$

75 $E(\xi_1 \mid I=1) = \frac{1}{P_1} \cdot \int_{x_1=-\infty}^{x_1=\infty} \int_{x_2=-\infty}^{x_2=V_1-V_2+x_1} x_1 \cdot f_\xi(x_1) \cdot f_\xi(x_2)\, dx_2 dx_1 = \frac{1}{P_1} \cdot \int_{x_1=-\infty}^{x_1=\infty} x_1 \cdot f_\xi(x_1) \cdot \int_{x_2=-\infty}^{x_2=V_1-V_2+x_1} f_\xi(x_2)\, dx_2 dx_1$.

Using $\int\limits_{x_2=-\infty}^{x_2=V_1-V_2+x_1} f_\xi(x_2)\,dx_2 = F_\xi(x_1+V_1-V_2)$ yields:

$$E(\xi_1 \mid I=1) = \frac{1}{P_1}\cdot \int\limits_{x_1=-\infty}^{x_1=\infty} x_1 \cdot f_\xi(x_1)\cdot F_\xi(x_1+V_1-V_2)\,dx_1 \,. \qquad \text{(MA4.2.7)}$$

This expression can be written as

$$E(\xi_1 \mid I=1) = \frac{1}{P_1}\cdot E_\xi\left(\xi\cdot F_\xi(\xi+V_1-V_2)\right). \qquad \text{(MA4.2.8)}$$

Using rule 10 of MA1, the expression above yields

$$E(\xi_1 \mid I=1) = \frac{1}{P_1}\cdot\frac{1}{1+e^{-h_1\cdot(V_1-V_2)}}\cdot\frac{\gamma+h_2+\ln\left(1+e^{-h_1\cdot(V_1-V_2)}\right)}{h_1}. \qquad \text{(MA4.2.9)}$$

Plugging in $h_1 = \frac{\pi}{\lambda\sqrt{3}}$ and $h_2 = -\gamma$ - see (MA4.1.8) - yields

$$E(\xi_1 \mid I=1) = \frac{1}{P_1}\cdot\frac{1}{1+e^{-h_1\cdot(V_1-V_2)}}\cdot\frac{\lambda\sqrt{3}}{\pi}\cdot\ln\left(1+e^{-h_1\cdot(V_1-V_2)}\right). \qquad \text{(MA4.2.10)}$$

Now I wish to compute the probability that a household chooses capital type one, P_1. Since the household decides on capital good one if the random variables $\{\xi_1, \xi_2\}$ are in set S_1, probability P_1 can be computed by integrating the joint density function of $\{\xi_1, \xi_2\}$ over set S_1. This yields[76]

---

[76] $P_1 = \int\limits_{x_2\in S_1} f_\xi(x_1,x_2)\,dx = \int\limits_{x_1=-\infty}^{x_1=\infty}\int\limits_{x_2=-\infty}^{x_2=V_1-V_2+x_1} f_\xi(x_1,x_2)\,dx_2dx_1 = \int\limits_{x_1=-\infty}^{x_1=\infty} f_\xi(x_1)\int\limits_{x_2=-\infty}^{x_2=V_1-V_2+x_1} f_\xi(x_2)\,dx_2dx_1 = \int\limits_{x_1=-\infty}^{x_1=\infty} f_\xi(x_1)\cdot F_\xi(x_1+V_1-V_2)\,dx_1.$

Substituting $z = h_1x_1+h_2 \Leftrightarrow x_1 = \frac{z-h_2}{h_1} \Rightarrow x_1 = \frac{1}{h_1}\cdot dz$ yields

$$P_1 = \int\limits_{z=-\infty}^{z=\infty}\frac{1}{h_1}\cdot f_\xi\left(\frac{z-h_2}{h_1}\right)\cdot F_\xi\left(\frac{z-h_2}{h_1}+V_1-V_2\right)dx_1 = \int\limits_{z=-\infty}^{z=\infty} f(z)\cdot F_\xi\left(\frac{z-h_2+h_1\cdot(V_1-V_2)}{h_1}\right)dz =$$
$$= \int\limits_{z=-\infty}^{z=\infty} f(z)\cdot\exp\left(-\exp\left(h_1\frac{z-h_2+h_1\cdot(V_1-V_2)}{h_1}+h_2\right)\right)dz = \int\limits_{z=-\infty}^{z=\infty} f(z)\cdot\exp\left(-\exp\left(z+h_1(V_1-V_2)\right)\right)dz = E_Z\left(F_g\left(Z+h_1(V_1-V_2)\right)\right),$$

where Z is a standard Gumbel distributed random variable.

Applying rule 6 of MA1, this expected value can be computed: $P_1 = \frac{1}{1+e^{-h_1\cdot(V_1-V_2)}}$.

$$P_1 = \int_{x_2 \in S_1} f_\xi(x_1, x_2)\,dx = \int_{x_1=-\infty}^{x_1=\infty} \int_{x_2=-\infty}^{x_2=V_1-V_2+x_1} f_\xi(x_1, x_2)\,dx_2 dx_1 = \frac{1}{1+e^{-h_1\cdot(V_1-V_2)}}\,. \qquad \text{(MA4.2.11)}$$

Plugging in this result in (MA4.2.10) yields:

$$E(\xi_1 \mid I=1) = -\frac{\lambda\sqrt{3}}{\pi}\cdot \ln(P_1)\,. \qquad \text{(MA4.2.12)}$$

Expected value of ξ_2 conditional on the choice of capital good one

Expression $E(\xi_2 \mid I=1)$ is only slightly different to $E(\xi_1 \mid I=1)$, namely

$$E(\xi_2 \mid I=1) = \int_{x \in S_1} x_2 \cdot f_{\xi|I=1}(x_1, x_2)\,dx = \frac{1}{P_1}\cdot \int_{x \in S_1} x_2 \cdot f_\xi(x_1, x_2)\,dx\,. \qquad \text{(MA4.2.13)}$$

Again, integrating over S_1 yields the integrals

$$E(\xi_2 \mid I=1) = \frac{1}{P_1}\cdot \int_{x_1=-\infty}^{x_1=\infty} \int_{x_2=-\infty}^{x_2=V_1-V_2+x_1} x_2 \cdot f_\xi(x_1, x_2)\,dx_2 dx_1\,. \qquad \text{(MA4.2.14)}$$

Stochastic independence of ξ_1 and ξ_2 yields:

$$E(\xi_2 \mid I=1) = \frac{1}{P_1}\cdot \int_{x_1=-\infty}^{x_1=\infty} \int_{x_2=-\infty}^{x_2=V_1-V_2+x_1} x_2 \cdot f_\xi(x_1)\cdot f_\xi(x_2)\,dx_2 dx_1\,. \qquad \text{(MA4.2.15)}$$

Since the inner integral $\int_{x_2=-\infty}^{x_2=V_1-V_2+x_1} x_2 \cdot f_\xi(x_2)\,dx_2$ cannot been solved explicitly, we write

$$E(\xi_2 \mid I=1) = \frac{1}{P_1}\cdot \int_{x_2=-\infty}^{x_2=\infty} \int_{x_1=-\infty}^{x_1=x_2+V_2-V_1} x_2 \cdot f_\xi(x_1)\cdot f_\xi(x_2)\,dx_1 dx_2 \qquad \text{(MA4.2.16)}$$

Regrouping the integrands yields

$$E(\xi_2 \mid I=1) = \frac{1}{P_1}\cdot \int_{x_2=-\infty}^{x_2=\infty} x_2 \cdot f_\xi(x_2)\cdot \int_{x_1=x_2+V_2-V_1}^{x_1=\infty} f_\xi(x_1)\,dx_1 dx_2 =$$

$$= \frac{1}{P_1}\cdot \int_{x_2=-\infty}^{x_2=\infty} x_2 \cdot f_\xi(x_2)\cdot\left(1-F_\xi(x_2+V_2-V_1)\right)dx_2 = \frac{1}{P_1}\cdot\left(E_\xi(\xi) - E_\xi\left(\xi\cdot F_\xi(\xi+V_2-V_1)\right)\right). \qquad \text{(MA4.2.17)}$$

Expression $E_\xi\left(\xi \cdot F_\xi\left(\xi + V_2 - V_1\right)\right)$ can be simplified by applying rule 10 from MA1 and the expected value $E_\xi\left(\xi\right)$ can be computed by applying rule 2 from MA1: $E\left(\xi\right) = \frac{\lambda + h_2}{h_1}$. Therefore,

$$E\left(\xi_2 \mid I = 1\right) = \frac{1}{P_1} \cdot \left(\frac{\lambda + h_2}{h_1} - \frac{1}{1 + e^{-\alpha(V_2 - V_1)}} \cdot \frac{\gamma + h_2 + \ln\left(1 + e^{-\alpha(V_2 - V_1)}\right)}{h_1} \right) =$$

$$= \frac{1}{P_1} \cdot \left(\frac{\lambda + h_2}{h_1} - \frac{1}{1 + e^{-h_1 \cdot (V_2 - V_1)}} \cdot \frac{\gamma + h_2 - \ln\left(\frac{1}{1 + e^{-h_1 \cdot (V_2 - V_1)}}\right)}{h_1} \right). \qquad \text{(MA4.2.18)}$$

For this case, I also want to rewrite the function such that it is a function of P_1. Thus, expression $\frac{1}{1 + e^{-h_1 \cdot (V_2 - V_1)}}$ has to be reformulated[77] and plugged into (MA4.2.18):

$$E\left(\xi_2 \mid I = 1\right) = \frac{1}{P_1} \cdot \left(\frac{\lambda + h_2}{h_1} - \left(1 - P_1\right) \cdot \frac{\gamma + h_2 - \ln\left(1 - P_1\right)}{h_1} \right). \qquad \text{(MA4.2.19)}$$

Plugging in $h_1 = \frac{\pi}{\lambda\sqrt{3}}$ and $h_2 = -\gamma$ yields

$$E\left(\xi_2 \mid I = 1\right) = \frac{\lambda\sqrt{3}}{\pi} \cdot \frac{1 - P_1}{P_1} \cdot \ln\left(1 - P_1\right). \qquad \text{(MA4.2.20)}$$

Linear relation between error terms and assumptions of Dubin and McFadden (1984)

In this paragraph it is shown that the assumption of a linear relationship between ξ and ε as defined by (MA4.1.8) is compatible with the assumptions of Dubin and McFadden (1984) as stated in Section 3.1. In the following, I test all these assumptions. I will do this even for the general case of many goods. I begin with assumption

$$E\left(\varepsilon_i \mid \xi\right) = \frac{\sigma\sqrt{2}}{\lambda} \cdot \sum_{j=1}^{J} R_{ij}\xi_j \,,\; R_{ij} = \operatorname{corr}\left(\varepsilon_i, \xi_j\right),\; E\left(\varepsilon_i\right) = 0 \text{ and } \sigma^2 = \operatorname{var}\left(\varepsilon_i\right).^{78} \qquad \text{(MA4.2.21)}$$

To prove this, I start with $E\left(\varepsilon_i \mid \xi\right)$:

[77] $\frac{1}{1 + e^{-h_1 \cdot (V_2 - V_1)}} = \frac{e^{-h_1 \cdot (V_1 - V_2)}}{1 + e^{-h_1 \cdot (V_1 - V_2)}} = 1 - \frac{1}{1 + e^{-h_1 \cdot (V_1 - V_2)}} = 1 - P_1$.

[78] See footnote 73 on page MA-21.

$$E(\varepsilon_i \mid \xi) = \sum_{j=1}^{J} a_{ij} \cdot \xi_j \ . \qquad \text{(MA4.2.22)}$$

From the independence of κ_i and ξ, and since $E(\kappa_i) = 0$, it follows that $E(\kappa_i \mid \xi) = 0$.

In the following, I need to prove that $a_{ij} = \frac{\sigma\sqrt{2}}{\lambda} \cdot R_{ij}$. To start, I need to compute $R_{ij} = \operatorname{corr}(\varepsilon_i, \xi_j)$:[79]

$$R_{ij} = \operatorname{corr}(\varepsilon_i, \xi_j) = \frac{1}{\sigma \cdot \sigma_\xi} E\left(\left(\sum_{k=1}^{J} a_{ik} \cdot \xi_k + \kappa_k \right) \cdot \xi_j \right) = \frac{\sigma_\xi}{\sigma} \cdot a_{ij} \text{, with } \sigma_\xi^2 = \operatorname{var}(\xi_j) = \lambda^2/2 \ . \qquad \text{(MA4.2.23)}$$

Solving for a_{ij} yields $a_{ij} = \sigma\sqrt{2}/\lambda \cdot R_{ij}$ and plugging in (MA4.2.22) yields

$$E(\varepsilon_i \mid \xi) = \frac{\sigma\sqrt{2}}{\lambda} \sum_{j=1}^{J} R_{ij} \cdot \xi_j \ , \qquad \text{(MA4.2.24)}$$

which corresponds to the assumption of Dubin and McFadden (1984).

Next, I want to check the compliance to the assumption of Dubin and McFadden (1984) on the conditional variance $\operatorname{var}(\varepsilon_i \mid \xi) = \sigma^2 \cdot \left(1 - \sum_{j=1}^{J} R_{ij}^2 \right)$:

$$\operatorname{var}(\varepsilon_{in} \mid \xi_{\cdot n}) = E\left(\left(\sum_{j=1}^{J} a_j \cdot \xi_j + \kappa_i - E(\varepsilon_i \mid \xi) \right)^2 \right) = E\left(\kappa_i^2\right) = \sigma_\kappa^2 \text{, with } \sigma_\kappa^2 = \operatorname{var}(\kappa_j) \ . \qquad \text{(MA4.2.25)}$$

Since it is not convenient to estimate σ_κ^2, Dubin and McFadden eliminated σ_κ^2. This can be done by solving $\sigma^2 = \operatorname{var}(\varepsilon_i)$:[80]

$$\operatorname{var}(\varepsilon_i) = \sigma^2 \cdot \sum_{j=1}^{J} a_{jj}^2 + \sigma_\kappa^2 = \sigma^2 \sum_{j=1}^{J} R_{ij} + \sigma_\kappa^2 \ . \qquad \text{(MA4.2.26)}$$

---

[79] $\operatorname{corr}(\varepsilon_i, \xi_j) = E\left(\left(\sum_{k=1}^{J} a_{ik} \cdot \xi_k + \kappa_k \right) \cdot \xi_j \right) = E\left(\left(\sum_{k=1}^{J} a_{ik} \cdot \xi_k \cdot \xi_j + \kappa_k \cdot \xi_j \right) \right) = a_{ij} \cdot E(\xi_j \cdot \xi_j) = a_{ij} \cdot \sigma_\xi$.

[80] $\operatorname{var}(\varepsilon_i) = E\left(\left(\sum_{j=1}^{J} a_{ij} \cdot \xi_j + \kappa_i \right)^2 \right) = E\left(\sum_{j=1}^{J} \sum_{k=1}^{J} a_j a_k \xi_k \xi_j + \sum_{j=1}^{J} a_{ij} \xi_j \kappa_i + \kappa_i^2 \right) = \sum_{j=1}^{J} a_{jj}^2 \cdot E\left(\xi_j^2\right) + E\left(\kappa_i^2\right) = \sigma_\xi^2 \cdot \sum_{j=1}^{J} a_{jj}^2 + \sigma_\kappa^2$. Plugging in $a_{ij} = \frac{\sigma\sqrt{2}}{\lambda} \cdot R_{ij}$ yields $\operatorname{var}(\varepsilon_i) = \frac{\lambda^2}{2} \cdot \sum_{j=1}^{J} \frac{2\sigma^2}{\lambda^2} \cdot R_{ij} + \sigma_\kappa^2 = \sigma^2 \sum_{j=1}^{J} R_{ij} + \sigma_\kappa^2 = \sigma^2$.

Solving for σ_κ^2 yields

$$\operatorname{var}\left(\varepsilon_{in} \mid \xi_{\bullet n}\right) = \sigma_\kappa^2 = \sigma^2 \left(1 - \sum_{j=1}^{J} R_{ij}^2\right). \tag{MA4.2.27}$$

Since variances are always positive, it follows that also $1 - \sum_{j=1}^{J} R_{ij}^2$ must be positive and therefore

$$\sum_{j=1}^{J} R_{ij}^2 < 1\,, \tag{MA4.2.28}$$

which corresponds to one of the restrictions imposed by Dubin and McFadden (1984).

Finally, I want to show why the last restriction $\sum_{j=1}^{J} R_{ij} = 0$ is reasonable. To justify this restriction, it is necessary to introduce a small model that explains the source of random terms ε_i and ξ_i. I assume that these random terms depend linearly on some unobserved socio-demographic variables $\tilde{s}$ and unobserved features of the capital good of type *i*, $\tilde{b}_i$:

$$\varepsilon_i = k_i \cdot \tilde{s} + l_i \cdot \tilde{b}_i\,, \tag{MA4.2.29}$$

$$\xi_j = \left(c_j - \overline{c}\right) \cdot \tilde{s} + d \cdot \tilde{b}_j + \varphi_j\,, \tag{MA4.2.30}$$

where φ_j accounts for the fact that some observed variables may not influence the indirect utility function as described by function (MA4.1.7) or that the impact of the random terms $\tilde{s}$ and $\tilde{b}_i$ on ξ_j may not be linear only. Vector $\overline{c}$ accounts for the fact that for the choice model, only the differences between the utility levels matter, but not the absolute levels, see (MA4.1.12). Therefore, independent of the values of vector $\overline{c}$, the choice of the capital good does not change. Assuming the model defined by (MA4.2.28) and (MA4.2.29), the correlation R_{ij} yields[81]

$$\begin{aligned} R_{ij} &= \frac{1}{\sigma_\varepsilon \cdot \sigma_\xi} \cdot \left(k_i \cdot \Sigma_{\tilde{s}} \cdot \left(c_j - \overline{c}\right)' + k_i \cdot \Sigma_{\tilde{s},\varphi_j} + l \cdot \Sigma_{\tilde{b}_i,\tilde{b}_j} \cdot d' + l \cdot \Sigma_{\tilde{b}_i,\varphi_j} \right) = \\ &= \frac{1}{\sigma_\varepsilon \cdot \sigma_\xi} \cdot \left(k_i \cdot \Sigma_{\tilde{s}} \cdot c_j' - k_i \cdot \Sigma_{\tilde{s}} \cdot \overline{c}' + k_i \cdot \Sigma_{\tilde{s},\varphi_j} + l \cdot \Sigma_{\tilde{b}_i,\tilde{b}_j} \cdot d' + l \cdot \Sigma_{\tilde{b}_i,\varphi_j} \right), \end{aligned} \tag{MA4.2.31}$$

---

[81] $R_{ij} = \operatorname{corr}\left(\varepsilon_i, \xi_j\right) = \frac{1}{\sigma_\varepsilon \cdot \sigma_\xi} \cdot E\left(\left(k_i \cdot \tilde{s} + l \cdot \tilde{b}_i\right) \cdot \left(\left(c_j - \overline{c}\right) \cdot \tilde{s} + d \cdot \tilde{b}_j + \kappa_j\right)\right) = \frac{1}{\sigma_\varepsilon \cdot \sigma_\xi} \cdot \ldots$

$\cdot \left(k_i \cdot E\left(\tilde{s} \cdot \tilde{s}'\right)\left(c_j - \overline{c}\right)' \cdot + k_i \cdot E\left(\tilde{s} \cdot \tilde{b}_j\right) \cdot d' + k_i \cdot E\left(\tilde{s} \cdot \kappa_j\right) + l \cdot E\left(\tilde{b}_i \cdot \tilde{s}\right) \cdot \left(c_j - \overline{c}\right)' + l \cdot E\left(\tilde{b}_i \cdot \tilde{b}_j\right) \cdot d' + l \cdot E\left(\tilde{b}_i \cdot \kappa_j\right) \right) =$

$= \frac{1}{\sigma_\varepsilon \cdot \sigma_\xi} \cdot \left(k_i \cdot \Sigma_{\tilde{s}} \cdot \left(c_j - \overline{c}\right)' \cdot + k_i \cdot \Sigma_{\tilde{s},\tilde{b}_j} \cdot d' + k_i \cdot \Sigma_{\tilde{s},\kappa_j} + l \cdot \Sigma'_{\tilde{s},\tilde{b}_i} \cdot \left(c_j - \overline{c}\right)' + l \cdot \Sigma_{\tilde{b}_i \cdot \tilde{b}_j} \cdot d' + l \cdot \Sigma_{\tilde{b}_i \cdot \kappa_j} \right).$

Since it is reasonable to assume that the correlation between unobserved socio-demographic variables and unobserved attributes of the capital goods are uncorrelated, $\Sigma_{\tilde{s},\tilde{b}_j}$ can be assumed to be zero and therefore the expression simplifies to expression (MA4.2.31).

where $\Sigma_{\tilde{s}}$ is the variance matrix of the unobserved socio-demographic variables $\tilde{s}$ and $\Sigma_{\cdot,\cdot}$ the covariance matrices of the variables indicated by the subscripts. Expression (MA4.2.31) shows that correlation R_{ij} is defined up to a constant term $k_i \cdot \Sigma_{\tilde{s}} \cdot \bar{c}'$. Dubin and McFadden assume that $\sum^{J} R_{ij} = 0$ for each *i*. This means that they impose *J* restrictions. It could be shown that, since vector $\bar{c}$ contains *J* elements, all these *J* restrictions can be met. Or to word it differently: the restrictions $\sum_{j=1}^{J} R_{ij} = 0$ are necessary, otherwise not all elements of R_{ij} could be identified when estimating the model.

The complete correction term for two goods

The complete correction term can be obtained by plugging (MA4.2.12), (MA4.2.20) and $a_{ij} = \frac{\sigma\sqrt{2}}{\lambda} \cdot R_{ij}$ (see (MA4.2.24)) in (MA4.1.13):

$$E(\varepsilon_1 \mid I = 1) = \frac{\sigma\sqrt{2}}{\lambda} \cdot \left(\frac{\lambda\sqrt{3}}{\pi} \cdot \frac{1-P_1}{P_1} \cdot \ln(1-P_1) - \frac{\lambda\sqrt{3}}{\pi} \cdot \ln(P_1) \right).$$

Cancelling out and reformulating probabilities by use of equality $P_1 + P_2 = 1$ yields

$$E(\varepsilon_1 \mid I = 1) = \frac{\sigma\sqrt{6}}{\pi} \cdot \left(\frac{P_2}{1-P_2} \cdot \ln(P_2) - \ln(P_1) \right). \tag{MA4.2.32}$$

Note that $E(\varepsilon_2 \mid I = 2)$ can be computed by simply replacing the indices of formula

$$E(\varepsilon_2 \mid I = 2) = \frac{\sigma\sqrt{6}}{\pi} \cdot \left(\frac{P_1}{1-P_1} \cdot \ln(P_1) - \ln(P_2) \right). \tag{MA4.2.33}$$

These results correspond exactly to the correction term as stated in Dubin and McFadden (1984: 355).

MA 4.3 Correction term for many goods

In this paragraph I will derive the correction term of Dubin and McFadden (1984) (MA4.1.14) for the case where a household has the choice between several types of capital goods. The elements I will need to derive the formula are similar to in the case where a household has only two goods from which to choose.

Again, from (MA4.1.13) it follows that I need to prove $E(\xi_j \mid I = i)$ for all *j*=1,...,*J*. It will turn out that it is sufficient to prove it for *i*=1 and for the two cases $j = 1$ and $j \neq 1$.

I start first by deriving the expression for P_1 and then expression $E(\xi_1 \mid I(\xi) = 1)$.

Probability of choosing capital good type one

The probability that the household chooses the capital good type one is the integral over set S_1 of ξ:

$$P_1 = \int_{x \in S_1} f_\xi(x)\,dx\,. \tag{MA4.3.1}$$

Set S_1 contains all combinations of ξ for which indicator *I* is equal to one:

$$I = 1, \text{ if } \xi_j < V_1 - V_j + \xi_1,\ \forall j \neq 1\,. \tag{MA4.3.2}$$

Therefore, integral (MA4.3.1) can be rewritten as:

$$P_1 = \int_{x \in S_1} x_1 \cdot f_{\xi|I=1}(x_1, x_2)\,dx = \int_{x \in S_1} x_1 \cdot \frac{f_\xi(x_1, x_2)}{P_1}\,dx\,. \tag{MA4.3.3}$$

$$P_1 = \int_{x_1=-\infty}^{x_1=\infty} \int_{x_2=-\infty}^{x_2=V_{1n}-V_{2n}+x_1} \cdots \int_{x_J=-\infty}^{x_J=V_{1n}-V_{Jn}+x_1} f_{\xi.}(x_1,\ldots,x_J)\,dx_J \ldots dx_2 dx_1\,. \tag{MA4.3.4}$$

Exploiting the fact that the random variables ξ_i are iid. yields

$$P_1 = \int_{x_1=-\infty}^{x_1=\infty} f_\xi(x_1) \prod_{i=2}^{J} F_\xi(V_1 - V_i + x_1)\,dx_1\,. \tag{MA4.3.5}$$

Applying rule 11 of MA1 and simplifying the expression yields[82]

---

[82] Note that (MA4.3.4) can rewritten as $E(\xi_1 \mid I = 1) = \frac{1}{P_1} \cdot E_\xi\left(X \cdot F_\xi\left(X - \frac{1}{h_1} \cdot \ln\left(\sum_{i=2}^{J} e^{-h_1 \cdot (V_1 - V_i)}\right)\right)\right)$ and on this formula, rule 10 of MA1 can be applied, yielding

$$E_X\left(X \cdot F_\xi\left(X - \frac{1}{h_1} \cdot \ln\left(\sum_{i=2}^{J} e^{-h_1 \cdot (V_1 - V_i)}\right)\right)\right) = \frac{1}{1 + e^{h_1 \cdot \frac{1}{h_1} \cdot \ln\left(\sum_{i=2}^{J} e^{-h_1 \cdot (V_1 - V_i)}\right)}} \cdot \frac{1}{h_1} \cdot \left(\gamma + h_2 + \ln\left(1 + e^{h_1 \cdot \frac{1}{h_1} \ln\left(\sum_{i=2}^{J} e^{-h_1 \cdot (V_1 - V_i)}\right)}\right)\right)$$

. Rewriting and simplifying yields

$$E_X\left(X \cdot F_\xi\left(X - \frac{1}{h_1} \cdot \ln\left(\sum_{i=2}^{J} e^{-h_1 \cdot (V_1 - V_i)}\right)\right)\right) = \frac{1}{1 + \sum_{i=2}^{J} e^{-h_1 \cdot (V_1 - V_i)}} \cdot \frac{1}{h_1} \cdot \left(\gamma + h_2 - \ln\left(\left(1 + \sum_{i=2}^{J} e^{-h_1 \cdot (V_1 - V_i)}\right)^{-1}\right)\right).$$

Using equality $\frac{1}{1 + e^{-h_1 \cdot V_1} \sum_{i=2}^{J} e^{h_1 \cdot V_i}} = \frac{e^{h_1 \cdot V_1}}{e^{h_1 \cdot V_1} + \sum_{i=2}^{J} e^{h_1 \cdot V_i}} = \frac{e^{h_1 \cdot V_1}}{\sum_{i=1}^{J} e^{h_1 \cdot V_i}} = P_1$ yields

$$E_X\left(X \cdot F_\xi\left(X - \frac{1}{h_1} \cdot \ln\left(\sum_{i=2}^{J} e^{-h_1 \cdot (V_1 - V_i)}\right)\right)\right) = P_1 \cdot \frac{\gamma + h_2 - \ln(P_1)}{h_1}.$$

$$P_1 = \int_{x_1=-\infty}^{x_1=\infty} f_\xi(x_1) \cdot F_\xi\left(x_1 - \frac{1}{h_1} \cdot \ln\left(\sum_{i=2}^{J} e^{-h_1 \cdot (V_1 - V_i)}\right)\right) dx_1 = E_\xi\left(F_\xi\left(X - \frac{1}{h_1} \cdot \ln\left(\sum_{i=2}^{J} e^{-h_1 \cdot (V_1 - V_i)}\right)\right)\right). \quad \text{(MA4.3.6)}$$

Applying rule 6 of MA1 yields[83]

$$P_1 = \frac{1}{1 + e^{-h_1 \cdot V_1} \sum_{i=2}^{J} e^{h_1 \cdot V_i}} = \frac{e^{h_1 \cdot V_1}}{\sum_{i=1}^{J} e^{h_1 \cdot V_i}} . \quad \text{(MA4.3.7)}$$

Expected value of ξ_1 conditional on the choice of capital good one

Again, as in the case with two types of capital goods, the expected value $E(\xi_1 \mid I = 1)$ can be computed as follows:

$$E(\xi_1 \mid I = 1) = \int_{x \in S_1} x_1 \cdot f_{\xi \mid I=1}(x_1, x_2)\, dx = \int_{x \in S_1} x_1 \cdot \frac{f_\xi(x_1, x_2)}{P_1} dx . \quad \text{(MA4.3.8)}$$

From the definition of the indicator function *I* (MA4.3.2) it follows that integral (MA4.3.8) can be written as

$$E(\xi_1 \mid I = 1) = \frac{1}{P_1} \cdot \int_{x_1=-\infty}^{x_1=\infty} \int_{x_2=-\infty}^{x_2=V_{1n}-V_{2n}+x_1} \cdots \int_{x_J=-\infty}^{x_J=V_{1n}-V_{Jn}+x_1} x_1 f_{\xi.}(x_1, ..., x_J)\, dx_J \cdots dx_2 dx_1 . \quad \text{(MA4.3.9)}$$

Exploiting the fact that the random variables ξ are iid. yields

$$E(\xi_1 \mid I = 1) = \frac{1}{P_1} \cdot \int_{x_1=-\infty}^{x_1=\infty} x_1 \cdot f_\xi(x_1) \cdot F_\xi(V_1 - V_2 + x_1) \cdot \cdots \cdot F_\xi(V_1 - V_J + x_1)\, dx_1 . \quad \text{(MA4.3.10)}$$

Applying rule 11 of MA1 yields

$$E(\xi_1 \mid I = 1) = \frac{1}{P_1} \cdot \int_{x_1=-\infty}^{x_1=\infty} x_1 \cdot f_\xi(x_1) \cdot F_\xi\left(x_1 - \frac{1}{h_1} \cdot \ln\left(\sum_{i=2}^{J} e^{-h_1 \cdot (V_1 - V_i)}\right)\right) dx_1 . \quad \text{(MA4.3.11)}$$

[83] $P_1 = \frac{1}{1 + e^{h_1 \cdot \frac{1}{h_1} \cdot \ln\left(\sum_{i=2}^{J} e^{-h_1 \cdot (V_1 - V_i)}\right)}} = \frac{1}{1 + \sum_{i=2}^{J} e^{-h_1 \cdot (V_1 - V_i)}} = \frac{1}{1 + e^{-h_1 \cdot V_1} \sum_{i=2}^{J} e^{h_1 \cdot V_i}} = \frac{e^{h_1 \cdot V_1}}{e^{h_1 \cdot V_1} + \sum_{i=2}^{J} e^{h_1 \cdot V_i}} = \frac{e^{h_1 \cdot V_1}}{\sum_{i=1}^{J} e^{h_1 \cdot V_i}}$.

Applying rule 10 of MA1 and simplifying the expression yields[84]

$$E(\xi_1 \mid I=1) = \frac{1}{h_1} \cdot \left(\gamma + h_2 - \ln(P_1)\right). \tag{MA4.3.12}$$

Plugging in the values for h_1 and h_2, see (MA4.1.8), yields

$$E(\xi_1 \mid I=1) = -\frac{\lambda\sqrt{3}}{\pi} \cdot \ln(P_1), \tag{MA4.3.13}$$

and by symmetry,

$$E(\xi_j \mid I=j) = -\frac{\lambda\sqrt{3}}{\pi} \cdot \ln(P_j). \tag{MA4.3.14}$$

Expected value ξ_2 conditional on the choice of capital good one

Again, as in the case with two types of capital goods, the expected value $E(\xi_1 \mid I=1)$ can be computed as follows:

$$E(\xi_2 \mid I=1) = \int_{x \in S_j} x_2 \cdot f_{\xi \mid I=1}(x_1,...,x_J)\, dx = \int_{x \in S_1} x_2 \cdot \frac{x_2 \cdot f_{\xi.}(x_1,...,x_J)}{P_1} dx. \tag{MA4.3.15}$$

From the definition of the indicator function I (MA4.3.2) it follows that integral (MA4.3.8) can be written as

$$E(\xi_2 \mid I=1) = \frac{1}{P_1} \cdot \int_{x_1=-\infty}^{x_1=\infty} \int_{x_2=-\infty}^{x_2=V_1-V_2+x_1} \int_{x_3=-\infty}^{x_3=V_1-V_3+x_1} \cdots \int_{x_J=-\infty}^{x_J=V_1-V_J+x_1} x_2 \cdot f_{\xi.}(x_1,...,x_J)\, dx_J \cdots dx_3 dx_1. \tag{MA4.3.16}$$

---

[84] Note that (MA4.3.4) can rewritten as $E(\xi_1 \mid I=1) = \frac{1}{P_1} \cdot E_\xi\left(X \cdot F_\xi\left(X - \frac{1}{h_1} \ln\left(\sum_{i=2}^{J} e^{-h_1 \cdot (V_1 - V_i)} \right)\right)\right)$ and rule 10 of MA1 can be applied to this formula, yielding

$$E_X\left(X \cdot F_\xi\left(X - \frac{1}{h_1} \cdot \ln\left(\sum_{i=2}^{J} e^{-h_1 \cdot (V_1 - V_i)} \right)\right)\right) = \frac{1}{1 + e^{h_1 \cdot \frac{1}{h_1} \cdot \ln\left(\sum_{i=2}^{J} e^{-h_1 \cdot (V_1 - V_i)} \right)}} \cdot \frac{1}{h_1} \cdot \left(\gamma + \beta + \ln\left(1 + e^{h_1 \cdot \frac{1}{h_1} \cdot \ln\left(\sum_{i=2}^{J} e^{-h_1 \cdot (V_1 - V_i)} \right)} \right)\right).$$

Rewriting and simplifying yields

$$E_X\left(X \cdot F_\xi\left(X - \frac{1}{h_1} \cdot \ln\left(\sum_{i=2}^{J} e^{-h_1 \cdot (V_1 - V_i)} \right)\right)\right) = \frac{1}{1 + \sum_{i=2}^{J} e^{-h_1 \cdot (V_1 - V_i)}} \cdot \frac{1}{h_1} \cdot \left(\gamma + h_2 - \ln\left(\left(1 + \sum_{i=2}^{J} e^{-h_1 \cdot (V_1 - V_i)} \right)^{-1} \right)\right).$$

Using the following equality $\frac{1}{1 + e^{-h_1 \cdot V_1} \sum_{i=2}^{J} e^{\alpha \cdot V_i}} = \frac{e^{h_1 \cdot V_1}}{e^{h_1 \cdot V_1} + \sum_{i=2}^{J} e^{h_1 \cdot V_i}} = \frac{e^{h_1 \cdot V_1}}{\sum_{i=1}^{J} e^{h_1 \cdot V_i}} = P_1$ yields

$$E_X\left(X \cdot F_\xi\left(X - \frac{1}{h_1} \cdot \ln\left(\sum_{i=2}^{J} e^{-h_1 \cdot (V_1 - V_i)} \right)\right)\right) = P_1 \cdot \frac{\gamma + h_2 - \ln(P_1)}{h_1}.$$

Exploiting the fact that random variables ξ are iid. yields

$$E\left(\xi_2 \mid I=1\right)=\frac{1}{P_1}\cdot\int_{x_1=-\infty}^{x_1=\infty} f_\xi\left(x_1\right)\cdot\int_{x_2=-\infty}^{x_2=V_2-V_1+x_1} x_2\cdot f_\xi\left(x_2\right)\cdot\prod_{j=3,..,J} F_\xi\left(V_1-V_j+x_1\right)dx_2dx_1\ . \tag{MA4.3.17}$$

It is necessary to rewrite the integral such that the limits of x_2 are $-\infty$ and ∞. This will allow to compute the expected value of a linear transformation of ξ_2.

$$E\left(\xi_2 \mid I=1\right)=\frac{1}{P_1}\cdot\int_{x_2=-\infty}^{x_2=\infty} x_2\cdot f_\xi\left(x_2\right)\cdot\int_{x_1=V_2-V_1+x_2}^{x_1=\infty} f_\xi\left(x_1\right)\cdot\prod_{j=3,..,J} F_\xi\left(V_1-V_j+x_1\right)dx_1dx_2\ . \tag{MA4.3.18}$$

Applying rule 11 of MA1 yields

$$E\left(\xi_2 \mid I=1\right)=\frac{1}{P_1}\cdot\int_{x_2=-\infty}^{x_2=\infty} x_2\cdot f_\xi\left(x_2\right)\cdot\int_{x_1=V_2-V_1+x_2}^{x_1=\infty} f_\xi\left(x_1\right)\cdot F_\xi\left(x_1-\frac{1}{h_1}\cdot\ln\left(\sum_{j=3,...,J}^{J} e^{-h_1\cdot\left(V_1-V_j\right)}\right)\right)dx_1dx_2\ . \tag{MA4.3.19}$$

Applying rule 12 of MA1 yields

$$E\left(\xi_2 \mid I=1\right)=\frac{1}{P_1}\cdot\frac{1}{1+\sum\limits_{j=3,...,J}^{J} e^{-h_1\cdot\left(V_1-V_j\right)}}\cdot\cdots$$

$$\cdots\int_{x_2=-\infty}^{x_2=\infty} x_2\cdot f_\xi\left(x_2\right)\cdot\int_{x_1=V_2-V_1+x_2}^{x_1=\infty} f_X\left(x_1-\frac{1}{h_1}\cdot\ln\left(1+\sum_{j=3,...,J}^{J} e^{-h_1\cdot\left(V_1-V_j\right)}\right)\right)dx_1dx_2\ . \tag{MA4.3.20}$$

Substituting $z=x_1-\frac{1}{h_1}\cdot\ln\left(1+\sum\limits_{j=3,...,J}^{J} e^{-h_1\cdot\left(V_1-V_j\right)}\right)$ and solving the above integral $\int\limits_{x_1=V_2-V_1+x_2}^{x_1=\infty}\cdots dx_1$ yields

$$E\left(\xi_2 \mid I=1\right)=\frac{1}{P_1}\cdot\frac{1}{1+\sum\limits_{j=3,..,J}^{J} e^{-h_1\cdot\left(V_1-V_j\right)}}\cdot\cdots$$

$$\cdots\int_{x_2=-\infty}^{x_2=\infty} x_2\cdot\left(1-F_\xi\left(x_2+V_2-V_1-\frac{1}{h_1}\cdot\ln\left(1+\sum_{j=3,...,J}^{J} e^{-h_1\cdot\left(V_1-V_j\right)}\right)\right)\right)\cdot f_\xi\left(x_2\right)dx_2\ . \tag{MA4.3.21}$$

This expression can be written using expectation operators:

$$E\left(\xi_2 \mid I=1\right)=\frac{1}{P_1}\cdot\frac{1}{1+\sum\limits_{j=3,...,J}^{J} e^{-h_1\cdot\left(V_1-V_j\right)}}\cdot\cdots$$

$$\cdots\left(E_\xi\left(\xi\right)-E_\xi\left(F_\xi\left(x_2+V_2-V_1-\frac{1}{h_1}\cdot\ln\left(1+\sum_{j=3,...,J}^{J} e^{-h_1\cdot\left(V_1-V_j\right)}\right)\right)\right)\right). \tag{MA4.3.22}$$

This expression can be simplified by rewriting[85] $\frac{1}{1+\sum_{j=3,\ldots,J}^{J} e^{-h_1\cdot(V_1-V_j)}}=\frac{P_1}{1-P_2}$, plugging in $E_\xi(\xi)=0$, applying rule 10 of MA1[86] $E_\xi\left(\xi\cdot F_\xi\left(\xi+V_2-V_1-\frac{1}{h_1}\cdot\ln\left(1+\sum_{j=3,\ldots,J}^{J} e^{-h_1\cdot(V_1-V_j)}\right)\right)\right)=P_2\cdot\frac{\gamma+h_2-\ln(P_2)}{h_1}$

and plugging in the values for h_1 and h_2, see (MA4.1.8):

$$E(\xi_2 \mid I=1)=-\frac{1}{P_1}\cdot\frac{P_1}{1-P_2}\cdot\frac{\lambda\sqrt{3}}{\pi}\cdot P_2\cdot\ln(P_2)=-\frac{\lambda\sqrt{3}}{\pi}\cdot\frac{P_2}{1-P_2}\cdot\ln(P_2), \tag{MA4.3.23}$$

and by symmetry for $i \neq j$,

$$E(\xi_j \mid I=i)=-\frac{\lambda\sqrt{3}}{\pi}\cdot\frac{P_j}{1-P_j}\cdot\ln(P_j). \tag{MA4.3.24}$$

The complete correction term

The complete correction term is obtained by plugging (MA4.3.14), (MA4.3.24) and $a_{ij}=\frac{\sigma\sqrt{2}}{\lambda}\cdot R_{ij}$ (see (MA4.2.23)) in (MA4.1.13):

$$E(\varepsilon_i \mid I=i)=\frac{\sigma\sqrt{2}}{\lambda}\cdot\sum_{j=1}^{J} R_{ij}\cdot\frac{\lambda\sqrt{3}}{\pi}\cdot\frac{P_j-\delta_{ij}}{1-P_j}\cdot\ln(P_j).$$

Cancelling out yields

$$E(\varepsilon_i \mid I=i)=\frac{\sigma\sqrt{6}}{\pi}\cdot\sum_{j=1}^{J} R_{ij}\cdot\frac{P_j-\delta_{ij}}{1-P_j}\cdot\ln(P_j), \tag{MA4.3.25}$$

---

[85] $$\frac{1}{1+\sum_{j=3,..,J}^{J} e^{-h_1\cdot(V_1-V_j)}}=\frac{1}{\sum_{j=1,..,J}^{J} e^{-h_1\cdot(V_1-V_j)}-e^{-h_1\cdot(V_1-V_2)}}=\frac{e^{h_1\cdot V_1}}{\sum_{j=1,..,J}^{J} e^{h_1\cdot V_j}-e^{h_1\cdot V_2}}=\frac{e^{h_1\cdot V_1}\cdot\left(\sum_{j=1,..,J}^{J} e^{h_1\cdot V_j}\right)^{-1}}{1-e^{h_1\cdot V_2}\cdot\left(\sum_{j=1,..,J}^{J} e^{h_1\cdot V_j}\right)^{-1}}=\frac{P_1}{1-P_2}.$$

[86]

$$E_\xi\left(\xi\cdot F_\xi\left(\xi+V_2-V_1-\frac{1}{h_1}\ln\left(1+\sum_{j=3,..,J}^{J} e^{-h_1\cdot(V_1-V_j)}\right)\right)\right)=\frac{1}{1+e^{-h_1\cdot\left(V_2-V_1-\frac{1}{h_1}\ln\left(1+\sum_{j=3,..,J}^{J} e^{-h_1(V_1-V_j)}\right)\right)}}\cdot\frac{\gamma+h_2+\ln\left(1+e^{-h_1\cdot\left(V_2-V_1-\frac{1}{\alpha}\cdot\ln\left(1+\sum_{j=3,..,J}^{J} e^{-h_1(V_1-V_j)}\right)\right)}\right)}{h_1}.$$

Note that

$$\frac{1}{1+e^{-h_1\cdot\left(V_2-V_1-\frac{1}{h_1}\cdot\ln\left(1+\sum_{j=3,..,J}^{J} e^{-\alpha\cdot(V_1-V_j)}\right)\right)}}=\frac{1}{1+e^{-h_1\cdot(V_2-V_1)}\cdot\left(1+\sum_{j=3,..,J}^{J} e^{-h_1\cdot(V_1-V_j)}\right)}=\frac{e^{h_1\cdot(V_2-V_1)}}{1+e^{h_1\cdot(V_2-V_1)}+\sum_{j=3,..,J}^{J} e^{-h_1\cdot(V_1-V_j)}}=\frac{e^{h_1\cdot(V_2-V_1)}}{\sum_{j=1,..,J}^{J} e^{h_1\cdot(V_j-V_1)}}=\frac{e^{h_1\cdot V_2}}{\sum_{j=1,..,J}^{J} e^{h_1\cdot V_j}}=P_2.$$

where δ_{ij} is the indicator $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ if $i \neq j$.

This corresponds exactly to the correction term as stated in Dubin and McFadden (1984: 355).

Selbständigkeitserklärung

Ich erkläre hiermit, dass ich diese Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen benutzt habe. Alle Koautorenschaften sowie alle Stellen, die wörtlich oder sinngemäß aus Quellen entnommen wurden, habe ich als solche gekennzeichnet. Mir ist bekannt, dass andernfalls der Senat gemäss Artikel 36 Absatz 1 Buchstabe o des Gesetzes vom 5. September 1996 über die Universität zum Entzug des aufgrund dieser Arbeit verliehenen Titels berechtigt ist.

Bern, 16. September 2011

Reto Tanner