The Determinants of Energy Demand of the Swiss Private Transportation Sector

Reto Tanner, Universität Bern

Conference paper STRC 2009

Reto Tanner
Universität Bern
Schanzeneckstrasse 1
CH-3001 Bern
Phone: +41 31 631 47 75
Fax: +41 31 631 39 92
email: retanner@gmx.ch

Last modification: 28th August 2009
R:\bhat\bhat_doku_rt\Bhat_STRC_2009_ver2.odt, Sat 29 Aug 09, 00:10:10

Abstract

This paper examines whether a household buys a car, and if, how much it drives that car. In this paper an approach based on the Multiple Discrete-Continuous Extreme Value Model (MDCEV). The MDCEV has been developed by Chandra R. Bhat (2005). In Bhat (2006) he adapted the Model to car choice and use. In this paper households could choose to own several cars and how much to use them. His approach has two drawbacks: First the total annual number of kilometres a household drives is considered to be fixed and second the fact that holding cars causes fixed costs is neglected.

I now adapted the model, such that households decision is based on an economic rational decision. This decision incorporates that owing cars causes fixed costs and that households decide on the number of kilometres they want to drive per year. So far, the model has been developed to the case where households may choose between none or one car and on the annual distance they want to drive.

Model parameters were estimated by use of Swiss data on car use on household level. Policy simulations yield similar fuel price elasticities as found in international studies. The model shows further, that reduction of fuel demand by higher fuel prices is mainly caused by households owning cars but using them less. The contribution of households switching from owning a car to not owning a car to the reduction in fuel demand is very low. The first reason for this is, that not many households will switch to not owning a car due to higher fuel prices. The second reason is, that these people did not drive many kilometres before when they owned a car. Further results show that household location - urban versus rural area - plays an important role both on demand for driving and on the decision whether to own a car or not. With respect to the choice of policies for reducing fuel demand, results show that not only the type and height of taxes on fuel and cars may play an important role, but also spatial planning.

Keywords

Fuel demand – private transportation sector – car demand – travel demand - Multiple Discrete-Continuous Extreme Value Model (MDCEV)

The Determinants of Energy Demand of the Swiss Private Transportation Sector

1.1 Introduction of the Model

In contrast to the discrete-continuous choice approach of Dubin and McFadden (1984), which can only capture exclusive choices between car types, or if several car types are involved, bundles of cars, this model can capture non-exclusive choices of a set of cars including the choice of not owning a car. This extension is very relevant for Swiss data, since in Switzerland 19% of households do not own a car and 30% of households own two or more cars.

The model based on the Multiple Discrete-Continuous Extreme Value Model (MDCEV) introduced by Bhat (2004) is applied to car choice and use in Bhat (2006). Bhat assumes that total driving distance is given for each household and is equal to the sum of kilometers driven by the vehicles that households declared in a survey. Further Bhat's approach contains the assumption that households are not restricted by either the households' budget or the fixed costs when owning one or several cars. Therefore, that model only captures households' preference for car types but does not capture the households' economic behavior. For instance, Bhat's model does not capture that it is economically not rational for households to own a vast number of cars. The purpose of the extension of Bhat's model is therefore to transform the model so that it represents the economic behavior of the household.

In the following, the foundations of the models are presented in section 1.1. In section 1.2 the most simple model where households can choose between owning a car and choose the driving distance, or not owning a car is derived for the case where fixed costs of car ownership are neglected. In section 1.3 the model is extended to the case where owning a car implies a fixed cost.

This introduction is structured as follows: First the basic principle of the model framework is presented. Second the microeconomic optimization problem is stated in a general form where households may choose between several cars. Third some illustration for the two good case is presented. Fourth the problem of unobservability of household preferences is by the researcher is stated and how this problem is captured by the model is described. Fifth, the utility function used in this framework is presented and it is discussed, why this distinct function was chosen.

The basic principle of the model framework

The model presented in the following describes the microeconomic decision of a household with respect to car ownership and use. In its general form it is assumed that households can choose from a set of cars one or several cars. The choice is only restricted in the way, that households may not choose two cars from of the same type. It is assumed, that the decision of deciding for one or several car types and the choice of distance driven with the corresponding cars is simultaneous. The household can also decide not to own a car. The household is assumed to maximize its utility for a given budget. The utility function values the utility the household yields by driving cars of different types and from a consumption good. The consumption good includes all goods apart from driving cars: Housing, health, food, insurances etc. The household can choose zero, one or several cars out of a set of different car types. Each car type can only be chosen one time. 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. The budget constraint also contains - apart from the simplified cases, where fixed costs are neglected - the fixed costs of owning one or several cars. The remaining expenditure is spent on the consumption good. The inclusion of fixed cost of car holding allows for a realistic description of households' behavior with respect to the decision whether to own or not to own a car. This decision is especially relevant for household with low income. It is assumed that the consumer has perfect information. This means hath the consumer exactly knows its preferences and is informed about the features of all car types it could choose. In contrast, from the researchers' perspective the utility function of the households are not known exactly and are therefore 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 Maximum Likelihood estimation. In order to get a simple formula for the Maximum Likelihood 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 last part of this introduction.

The microeconomic optimization problem

Now, the microeconomic problem solved by a household is described. The household is considered to behave as if it maximizes a utility function

$$\max_{x} u(x), \tag{1.1.1}$$

subject to:

$$y \geq p_1 x_1 + \sum_{i=2}^{J} p_i x_i + \sum_{i=2}^{J} I_{x_i>0}(x_i) k_i \text{ and } x_i \geq 0 \;\; \forall i = 1..J . \tag{1.2.1}$$

The amount of consumption of good one is denoted by x_1. Index i=2..J is an index for car types. The annual distance in kilometers driven by car type i is denoted by x_i. Vector x contains all x_i. Variable

p_i denotes the cost per kilometer for car driving by car *i*. The costs per kilometer consist of fuel costs and depreciation caused by driving the car, eg. wear of the mechanical components and tires. Variable k_i refers to the annual fixed of holding a car. These fixed costs consist of parking cost, insurance costs, taxes and depreciation. In this context, depreciation only captures the loss of value caused by factors unrelated to the use of cars, like rusting and loss in value due to technical obsolescence. It is assumed that if a household owns a car household will also drive by this car and therefore annual distance x_i is assumed to be strictly positive in this case. When a household decides not to own a certain car type *i*, then the corresponding distance x_i is zero. Therefore ownership of car type *i* is equivalent to a positive value x_i. Since fixed costs only arise if a car is owned, an indicator is needed when summing up the fixed cost of the different car types. Indicator $I_{x_i>0}(x_i)$ is one if x_i is greater than zero and zero otherwise.

The microeconomic 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 restriction is now non-linear in $x_i \; \forall i = 2..N$ due to the indicator function $I_{x_i>0}(x_i)$. Therefore the optimization problem cannot be solved by standard Lagrangian approach. Instead, the maximization problem has to be solved by the Kuhn Tucker approach. To show that, it is necessary to restate the constraints (1.1.2):

$$g_{J+1}(x) = \sum_{i=1}^{J} q_i(x_i) - y \leq 0, \tag{1.1.3a}$$

where $q_i(x_i) = I_{x_i>0}(x_i) \cdot k_i + p_i \cdot x_i$, with $k_1 = 0$ and

$$g_j(x) = -x_j \leq 0 \;, j = 1,2,3,..,J. \tag{1.1.3b}$$

This problem can be stated as Lagrangian:[1]

In the following it is shown, that the Kuhn-Tucker approach can be applied to the maximization problem stated by (1.1.1) and (1.1.2). This will be done by the Kuhn-Tucker Theorem. This theorem relates to the Lagrangian representation of (1.1.3):

$$L = u(x) - \sum_{j=1}^{J+1} \lambda_j g_j(x). \tag{1.14}$$

The Kuhn-Tucker Theorem states, that if x^* solves (1.1.1) and the constraints (1.1.3) hold at x^*, then there exists a set of Kuhn-Tucker multipliers $\lambda_j \geq 0$, for *i* = 1, 2,.., $J+1$ such that

[1]Varian (1992), page 505.

$$\frac{\partial u(x^*)}{\partial x_i} = \sum_{j=1}^{J+1} \lambda_j \cdot \frac{\partial g_j(x^*)}{\partial x_i},\ i = 1, 2,.., J\ . \quad (1.1.5)$$

Furthermore, there are the so called complementary slackness conditions for:

$$\lambda_j > 0 \text{, if } g_j(x^*) = 0, \quad (1.1.6a)$$

$$\lambda_j = 0 \text{, if } g_j(x^*) < 0, \quad (1.1.6b)$$

where j=1,2,..J+1.

The Kuhn-Tucker sufficiency[2] theorem states, that if x^* complies (1.1.5) and (1.1.6) if solves the maximization problem stated by (1.1.1) and (1.1.2) if $u(x)$ is a quasi-concave function[3] and $g_i(x^*), i = 1, 2,..J+1$ are convex functions.

Note that any utility function must be quasi-convex to satisfy the fundamental axioms on preference relations[4]. Since $g_1(x)$ is not a convex function, the Kuhn-Tucker sufficiency[5] theorem does not apply to the problem stated by (1.1.4) and (1.1.5). Therefore, the problem has to be restated as follows: Assume that the household first chooses zero, one or several cars out of the set S_k of cars plus always the consumer good at choice set S_c. Each car of the set S_c may only be chosen once. The household then maximizes the utility conditional on the choice S_k, where $k = 1,2,..,K$ is an index for all possible choices S_k. That means that its budget will be reduced by the fixed costs of the cars in the set S_k causes. Given this reduced budget, household will then solve the following maximization problem:

$$\max_{x_i,\, i\in S_k} u(x) \text{, with } x_j = 0,\ j \in S_c / S_k\ , \quad (1.1.7)$$

subject to:

$$y \geq p_1 x_1 + \sum_{i=1}^{J} p_i x_i + \sum_{i\in S_k} k_i \text{ and } x_i \geq 0\ \forall i \in S_k\ . \quad (1.1.8)$$

---

[2]See Varian (1992), page 503.

[3] See Mas-Colell, page 49.

[4]The property of strictly-convex preferences implies that a utility function is strict quasi-concave, see Mas-Colell et al (1995), page 49. Strictly-convex preferences are defined as strictly convex if for every *x*, we have that $x \succsim y, x \succsim x$, and $x \neq y$ implies $\alpha y + (1-\alpha) z \succ x$ for all $\alpha \in (0,1)$, see Mas- Colell et al (1995), page 44. Strict quasi-concavity is defined as if $u(\alpha x + (1-\alpha) y) > \min(u(x), u(y))$ for any *x*, *y* and $\alpha \in (0,1)$, then $u(\bullet)$ is a ", Mas-Colell, page 44. The assumption of a strict-convex utility function is an additional restriction, but most commonly used utility functions are strictly-convex. The utility function that will be used for the model presented here is also strictly-convex . Note that every function that is strictly-convex is also quasi-convex .

[5]See Varian (1992), page 503.

Now, for each S_k the Lagrangian has to be set up:

$$L = u(x) - \sum_{i \in S_k} \lambda_i g_i(x), \text{ with } x_j = 0,\ j \in S_c / S_k . \qquad (1.1.9)$$

Restrictions $g_j(x) = -x_j \leq 0, i \in S_k$ are defined as

$$g_j(x) = -x_j \leq 0,\ j \in S_c , \qquad (1.1.10a)$$

$$g_j(x) = 0,\ j \in S_k / S_c . \qquad (1.1.10b)$$

Restriction $g_{J+1}(x)$ is now defined as

$$g_{J+1}(x) = \sum_{i \in S_k} p_i x_i + \sum_{i \in S_k} k_i - y \leq 0 . \qquad (1.1.10c)$$

Note that $g_{J+1}(x)$ does now depend linearly on $x_i, i \in S_k$ and in this case it is a convex function. The same holds for $g_j(x), i \in S_k$. Conditions $g_j(x) = 0,\ j \in S_k / S_c$ (1.1.10b) are not relevant, since they are always fulfilled.

Since all restrictions $g_i(x)$ are now convex, the Kuhn-Tucker sufficiency Theorem holds. Therefore,

if x^* is feasible and solves (1.1.11) and fulfills (1.1.12), then x^* solves the maximization problems stated in (1.1.7) and (1.1.8).

$$\frac{\partial u(x^*)}{\partial x_i} = \sum_{j=1}^{J+1} \lambda_j \cdot \frac{\partial g_j(x^*)}{\partial x_i},\ i \in S_k , \qquad (1.1.11)$$

with the corresponding complementary slackness conditions:

$$\lambda_j > 0 \text{, if } g_j(x^*) = 0 \text{, if} \qquad (1.1.12a)$$

$$\lambda_j = 0 \text{, if } g_j(x^*) < 0 , \qquad (1.1.12b)$$

where $i \in S_k$.

Solving this maximization problem the household will yield optimal consumption level and driving distance for the set S_k. The utility household gets form this consumption shall be denoted by $u_k = u(x^*)$. The household will now compute u_k for any possible choice S_k. The household will then choose the choice set S_k that yields the highest utility.

Illustration of the maximization principle in the case of two goods

In the following, there is some intuitive presentation of the model where households can choose between two goods with one of them causing fixed costs for any positive consumption level. It will be shown, how changes in prices, income and preferences on optimal solution is. This is done for the case where only one car or no car can be chosen.

I want to start by illustrating the maximization routine described in paragraph above:

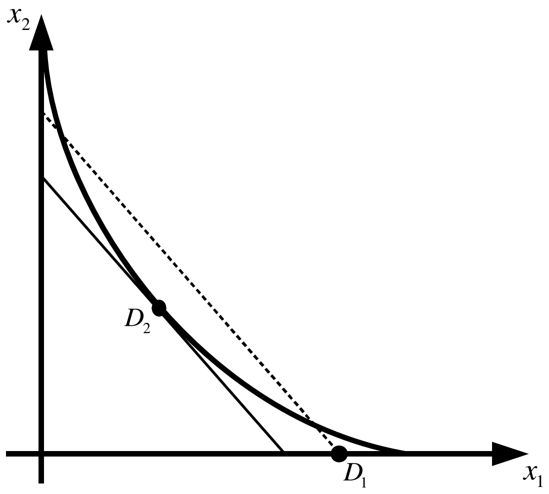


Diagram 1.1.1: Optimal consumption for given fixed cost for car driving

The dashed line represents the budget constraint when the household decides not to own a car and therefore does not have to bear fixed cost. In this situation all income is spent on the consumption good. The optimal consumption bundle when deciding not to own a car is represented by D_1. The household will compare the utility when spending all income only for the consumption good to the situation where it decides also to drive by car. In this situation the budget line is on a lower level because a part of the income is spent for the fixed cost of owning a car. The household now decides the optimal amount of driving. The slope of the budget line is equal to the price of the consuming good divided by the marginal cost of driving. Since price of the consumption good is normalized to one, the slope of the budget line is equal to one divided by price of good two. Utility maximization calculus for the case where the household is supposed to own and use a car yields optimal consumption bundle D_2. The solid line represents the iso-utility level of the maximal utility that can be reached with the income net the fixed costs. The household will now compare the utility level yielded given the case when spending the whole budget on the consumption good to the case when owing and using a car. In the case as illustrated in the diagram above, the household would yield a higher utility when owning a car, since D_1 is below the utility level yielded by D_2. Note that for the same utility function, this decision can change, when fixed costs are increasing: When fixed costs are increasing, solid budget line will shift towards to the origin of the diagram and therefore also D_2 does. From this it follows that the crossing of the solid lined iso-utility function will shift towards the left. At some point, D_2 will be above the iso-utility function and for this case it is optimal not to own a car. Diagram below illustrates such a situation.

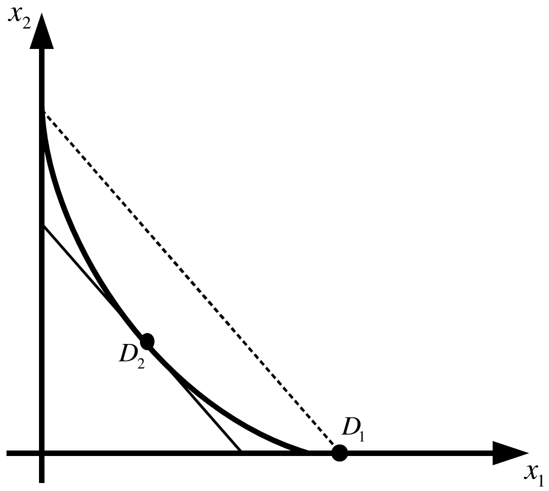


Diagram 1.1.2: Optimal consumption for high fixed cost for good two

Compared to the situation illustrated in diagram (1.1.1) decision of owning a car can change when income is lower, since fixed costs will decrease available income by a larger share.

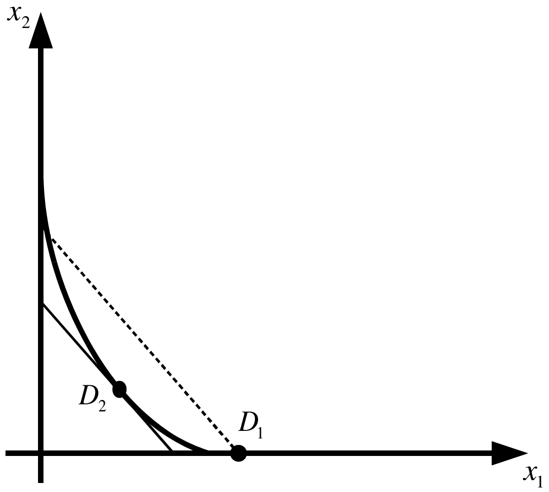


Diagram 1.1.3: Optimal consumption for given fixed cost for good two when income is low

When income is decreasing, optimum consumption bundle D_2 shifts towards the origin and therefore also iso-utility function does. Since the distance between the two budget lines remains the same, at some point, D_1 will be above the iso-utility function and therefore it will be optimal not to own a car. This is quite intuitive, since for lower income the fixed costs are getting relatively higher. In the extreme case, fixed cost would lower available income to zero and household could not consume any goods.

In the following case the effect of an increase of the price of good two is examined.

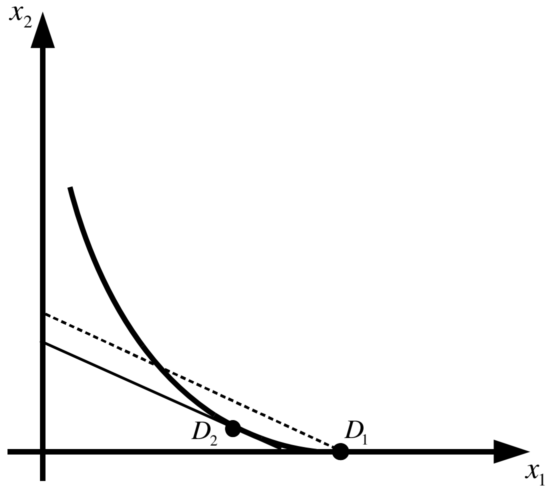


Diagram 1.1.4: Optimal consumption for given fixed cost for good two when price of good two is high

The above diagram shows that if price of good two increases, optimal consumption of good two given good two is decreasing. From this, the iso-utility line is crossing the x_1 axis closer to the budget line. Since the distance between the crossing of the budget lines on the x_1 axis is still the same, D_1 is now in the better set of iso-utility line corresponding to the case where good two is consumed. Therefore in the case illustrated above, household will choose not to buy a car.

The model does now not only capture the effect of changes in economic variables income *y*, fixed cost of owning a car k_2, and cost of driving a kilometer with that car p_2 as presented above but also individual preferences.

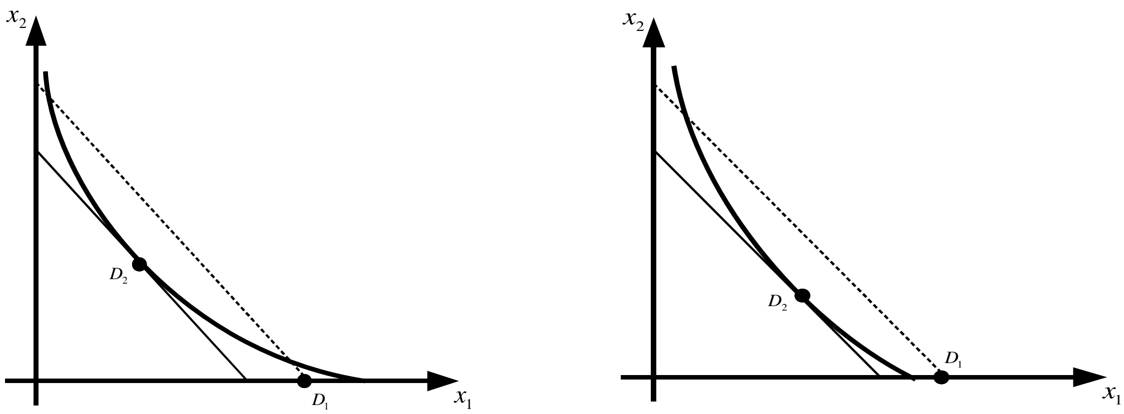


Diagram 1.1.5: Optimal consumption with strong and weak preference for car driving

On the left, optimal solution of a household with strong preference for car driving is illustrated. This household will choose to own a car and consume bundle D_2. On the right, optimal solution of a household with weak preference for car driving is illustrated. This household will choose not to own a car and consume bundle D_1.

The key issue of the model is now to explain these relative preferences of households by data. These preferences can depend on household characteristics, like the number of household members or on the type of household location, like city versus rural areas. Given these preferences, change in behavior of households behavior when economic variable like income and prices changes can be simulated. The following diagram shows the predicted outcome when a households with a given preference but varying income could decide for no car, a small car, a big car or both a small and a big car. In contrast to the preceding examples households can now decide between no car, a small car with low fixed and variable costs and a big car with high fixed an variable cost. It is assumed, that the big car yields higher utility than the small car, since the big car is more comfortable an provides more space for transporting goods.

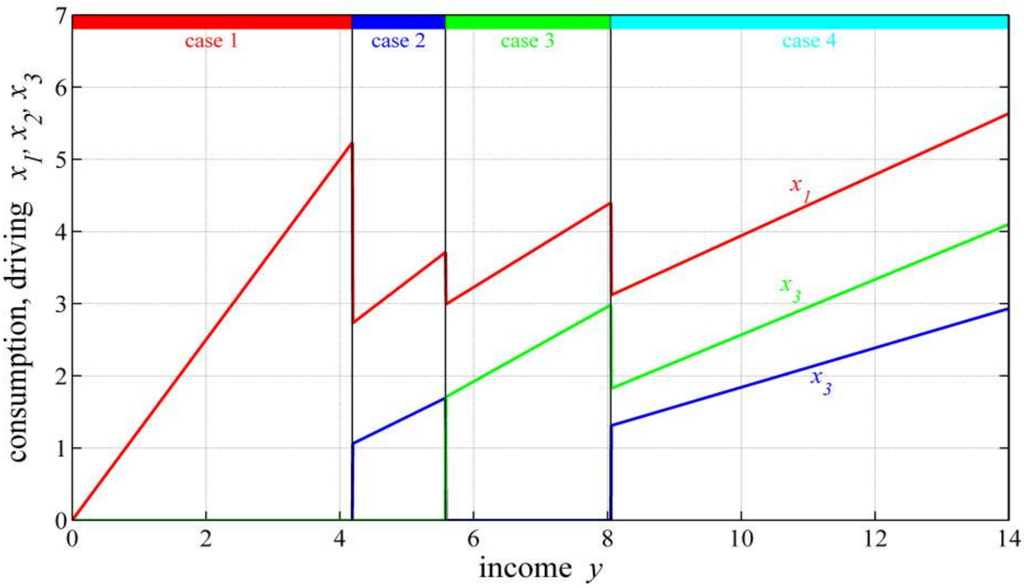


Diagram 1.1.4: Engel-curves for consumption and driving for given preferences[6]

[6]Diagram was computed based on the utility function $U = w_1 \cdot (x_1 + a_1)^d + w_2 \cdot (x_2 + a_2)^d + w_3 \cdot (x_3 + a_3)^d$ subject to fixed and marginal cost. Parameters were $a_1 = 0, a_2 = 0.7, a_3 = 1, w_1 = 1, w_2 = 0.8, w_3 = 1.2, d = 0.1$ and economic variables were $p_1 = 0.95, p_2 = 1.05, p_3 = 0.8, k_1 = 0, k_2 = 1, k_3 = 1.4$. Note that $w_3 > w_2$ denotes relative preference of driving a big car (good three) being greater than relative preference of driving a small car (good two). Utility was maximized given different income levels *y*. Price $p_3 > p_2$ denotes, that marginal costs of driving a big car are higher that driving a small car. The same holds for fixed costs: $k_3 > k_2$. Relation $a_3 > a_2$ implies that marginal utility of driving a small car decreases faster that driving a small car. This is quite reasonable, since for instance driving long distances by a small car gets faster tiring than driving the same distance by a big car.
For a detailed description of the maximization routine see in the following sections.

In this case when household income increases, first no car would be chosen and the whole income would be spent on consumption good x_1. This choice is denoted as "case 1" above. When income increases first the small car will be chosen to drive with - x_2 - due to its lower fixed cost, "case 2". When income increases further, the household will choose to drive the big car - x_3 - due to its higher preference for it, despite to its higher fix costs, "case 3". When the income is high enough, the household will choose to drive both cars, "case 4". In this case the household will drive more by the more expensive car due to its higher preference for this car type, despite of higher marginal costs. Note the steps in the Engel curves are due to the fact that fixed costs arise when owning a car. These fixed costs will reduce the disposable income. An interesting detail is the jump in distance driven from zero to a positive value when the car is bought due to a higher income. This seems reasonable since only owning but not driving a car yields utility and no one would bare the fixed costs when the only change would be, that the consumption level of the consumer good has to be reduced due to the decrease of available income.

Until now, preferences were assumed to be known. In reality however, not all factors that determine individual preferences are available and second it is not exactly known how the factors influence the individual preferences. From this follows that individual preferences cannot be exactly described from researchers perspective. Hence, whether a household chooses to own a car or not - given observed household characteristics and economic variables - can be predicted by the researcher only with a certain probability. Also for the case a household chooses to own a car, the amount consumed of good one and the distance driven cannot exactly be determined. It is only possible to compute probabilities that these amounts are within certain intervals.

The chapter is structured as follows: First, in section 1.2, the model is introduced for the simple case when there are only two goods and no fixed costs. Then in section 1.3. the two good case is extended to the case with fixed cost. After that, the model is expanded to the case of three goods, first in section 1.4 without fixed cost and later including fixed costs in section 1.5.

1.2 Model with two goods and no fixed cost

I first start with this very simple model because the basic implications can be studied in a comprehensive way. Further the basic problems when deriving the ML function can be examined and presented in the most simple way. The ML function describes the probability observing the household given model parameters. It will be used to determine the model parameters by use of the MLE procedure. In this section the model with two goods and *no* fixed cost is described. First the basic assumptions are presented and some implications of the model are illustrated by diagrams. Finally the ML function necessary for estimating the parameters when observing data is derived.

The model is based on a additive utility function that includes a random term. This random term stands for the fact that researchers may not be able to determine consumers true utility function. There are two goods. Good two X_2 denotes the kilometres driven by the car. In this model owning a car does not implicate fixed costs. Good one X_1 denotes a bundle of goods containing any goods apart form driving car. The utility function is defined as follows:

$$U = u_1(X_1) + u_2(X_2) = \exp(m_1 + \xi_1)\cdot(X_1 + a_1)^{d_1} + \exp(m_2 + \xi_2)\cdot(X_2 + a_2)^{d_2}, \qquad (1.2.1)$$

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

Since the marginal utility is decreasing in X_1 and X_2 both d_1 and d_2 are bound to the interval between zero and one:[7] $0 < d_j < 1,\ j = 1,2$. The lower d_j the more rapid the marginal utility of good *j* decreases when X_2 increases. Parameters a_1 and a_2 are shifting parameters. Since marginal utility of X_1 and X_2 are infinite at $X_1 = -a_1$ and $X_2 = -a_2$ respectively, $-a_1$ and $-a_2$ define lower limits for X_1 and X_2 for optimal solutions for X_1 and X_2 when the ranges of the solutions are not bounded. Since good one X_1 is essential, a_1 must be non negative in order to insure that the solution for X_1 is always positive. Since good two is not essential a_2 must be positive. This allows for negative solutions for X_2 that will be bounded to zero. Expression $\exp(m_j + \xi_j)$ is weighting $(X_j + a_j)^{d_j}$. The higher $\exp(m_j + \xi_j)$, the stronger the preference for good *j*. This weight is determined by sociodemographic variables s and characteristics 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 with many members usually have a higher preference for driving a car. Therefore it can be expected that the number of people of a household would increase m_2 and therefore utility of good X_2 would be weighted higher compared to good X_1 for such households. The random terms ξ_j represent sociodemographic variables $\tilde{s}$ and vehicle characteristics $\tilde{b}$ that cannot be observed by the researcher. These random terms are assumed to be iid Gumbel distributed[8]:

---

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

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

, 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 additive separable type.

[8]The Gumbel distribution is a non symmetric distribution but is shaped similarly to the normal distribution, see figure in the appendix A1. The Gumbel distribution also has has some useful properties necessary for getting a ML function that is an

$$\xi_j \sim iid\, gu(0,1),\ f_\xi(x) = e^{-x} \cdot \exp\left(-e^{-x}\right). \tag{1.2.2}$$

Both the choice of this special form of the utility function and the assumption on the error terms are done because it allows for some formal simplifications when derivating the ML function. These simplifications will yield a ML function in closed form. Further, the cumulated density function that will appear in the ML function is of a simple form and therefore permits a short computation time. It is assumed that the household maximizes its utility by choosing optimal values for X_1 and X_2. For this it has to take into account its budget constraint:

$$y = p_1 \cdot X_1 + p_2 \cdot X_2. \tag{1.2.3}$$

Note that for this case there are no fixed costs assumed for good two. The maximization of the household can be represented by solving the following Lagrangian:

$$L = \exp(m_1 + \xi_1) \cdot (X_1 + a_1)^{d_1} + \exp(m_2 + \xi_2) \cdot (X_2 + a_2)^{d_2} + \lambda \cdot (y - p_1 \cdot X_1 - p_2 \cdot X_2), \tag{1.2.4}$$

$$X_1 > 0,\ X_2 \geq 0.$$

Note that in the following ξ_1 and ξ_2 that represent unobserved characteristics are treated as given. This is for formal reasons but is also reasonable, since households known their characteristics. Therefore random terms ξ_1 and ξ_2 are considered as known by the households.

The corresponding first order conditions are:

$$\partial U / \partial X_1 - \lambda \cdot p_1 = 0, \tag{1.2.5}$$

$$\partial U / \partial X_2 - \lambda \cdot p_2 \leq 0. \tag{1.2.6}$$

The third first order condition is budget constraint (1.2.3). By plugging in the expressions for the marginal utility functions and reformulating, (1.2.5) and (1.2.6) can be written as follows:

$$d_1 \cdot \exp(m_1 + \xi_1) \cdot \frac{1}{(x_1 + a_1)^{1-d_1}} = \lambda \cdot p_1, \tag{1.2.7}$$

$$d_2 \cdot \exp(m_2 + \xi_2) \cdot \frac{1}{(x_2 + a_2)^{1-d_2}} \leq \lambda \cdot p_2. \tag{1.2.8}$$

explicit function of the parameters, see in the appendix. The Gumbel distribution is also often called extreme value distribution of type I.

The reason why (1.2.6) is an inequality but (1.2.5) is an equality is that $a_1 = 0$ that implies optimal $X_1 > 0$ solutions while as $a_2 > 0$ that allows for negative solutions X_2, that have to be bound too zero. Because of a_2 being greater than zero, solutions for X_2 that are bound to zero are possible. For this case the marginal utility of X_2 is strictly smaller than the Lagrangian multiplier λ times p_2. The reason why only X_2 can yield bounded solution will become clear when this case will be illustrated below. Intuitively it can be imagined that households start solving the system of equations by choosing a very high positive value for the Lagrangian multiplier λ. Given this value λ, the household solves (1.2.7) for $X_1 = x_1(\lambda)$ and (1.2.7) for $X_2 = \max(0, x_2(\lambda))$. Household will then check if the budget constraint is violated: $y \geq p_1 \cdot x_1(\lambda) + p_2 \cdot x_2(\lambda)$. Since for very high values λ the values $x_1(\lambda)$ and $x_2(\lambda)$ are small, the budget constraint will not be violated. Since both $x_1(\lambda)$ and $x_2(\lambda)$ depend negatively on λ,[9] the household will lower λ. It will do this until the whole budget is used for consumption due to increasing values for both $x_1(\lambda)$ and $x_2(\lambda)$. This optimization process can be illustrated by the following diagrams:

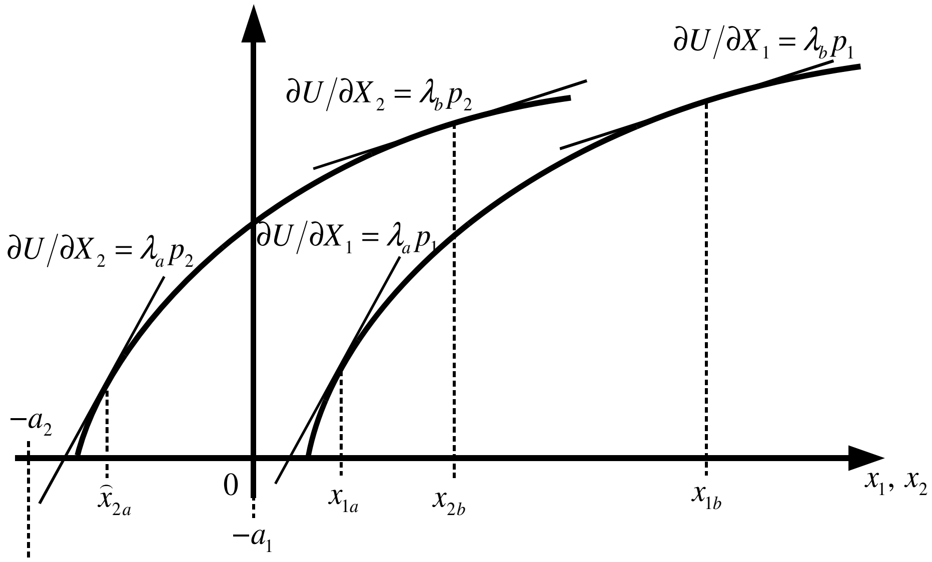


Diagram 1.2.1: The maximization calculus[10]

---

[9] $$\frac{dx_1}{d\lambda} = -\frac{\partial d_1 \cdot \exp(m_1+\xi_1)\cdot \frac{1}{(x_1+a_1)^{1-d_1}} - \lambda \cdot p_1 \Big/ \partial\lambda}{\partial d_1 \cdot \exp(m_1+\xi_1)\cdot \frac{1}{(x_1+a_1)^{1-d_1}} - \lambda \cdot p_1 \Big/ \partial x_1} = -\frac{-p_1}{-d_1(1-d_1)\exp(m_1+\xi_1)\cdot \frac{1}{(x_1+a_1)^{2-d_1}}} < 0,$$

$$\frac{dx_2}{d\lambda} = -\frac{\partial d_2 \cdot \exp(m_2+\xi_2)\cdot \frac{1}{(x_2+a_2)^{1-d_2}} - \lambda \cdot p_2 \Big/ \partial\lambda}{\partial d_2 \cdot \exp(m_2+\xi_2)\cdot \frac{1}{(x_2+a_2)^{1-d_2}} - \lambda \cdot p_2 \Big/ \partial x_2} = \left\{ \begin{array}{l} x_2 > 0 \;:\; -\dfrac{p_2}{d_2(1-d_2)\exp(m_2+\xi_2)\cdot \frac{1}{(x_2+a_2)^{2-d_2}}} \\ x_2 = 0 \;:\; 0 \end{array} \right\}.$$

Here the results for two values λ are illustrated, $\lambda_a > \lambda_b$. For λ_a the optimal value of X_2 has to be bound to zero, since for $\partial U/\partial x_2 = \lambda_a p_2$ the value for X_2 would be smaller than zero. Because X_2 is bound to zero and therefore increased compared to the non bounded solution, marginal utility of X_2 is smaller than $\lambda_a p_2$, $\partial U/\partial x_2 (X_2 = 0) < \lambda_a p_2$. Further it is assumed that for λ_a budget is not completely used up, $y > p_1 \cdot x_1(\lambda_a) + p_2 \cdot \max(0, x_2(\lambda_a))$. Therefore the household can still increase its utility by decreasing λ. This will increase in each case X_1 and in the case as illustrated above, at some point X_2 will change from being zero to a positive value. When λ has been decreased to λ_b, the whole budget is used up and household has maximized its utility. Note that diagram 1.2.1 does also illustrate the role of the shifting parameters a_1 and a_2: $-a_1$ and $-a_2$ define the minimum value of the consumption of good one and good two when maximizing without restricting values to be positive. From diagram 1.2.1 and formula (1.2.7) it can be seen that marginal utility of X_1 goes to infinity when X_1 goes to zero. Therefore optimal solution of X_1 is strictly positive for any finite λ. Contrary to a_1, a_2 is greater than zero for allowing negative non restricted solutions for $X_2 > -a_2$ that have to be bound to zero. Since in the context of the applications, X_1 always has to be positive, a_1 is chosen to be zero for insuring non-bound solutions, $X_1 > 0$. In contrast a_2 is chosen to be greater than zero what allows for solutions of X_2 that are bound to zero, $X_2 = 0$ as described above. Here the shape of the utility function, the prices and the income are such that for this case both X_1 and X_2 are greater than zero when choice is optimal. When income would be such that it would be used up at level λ_a then the optimal solution for X_2 would be equal to zero. When the corresponding income levels for these two cases are denoted y_a and y_b, the two optimal solutions for these cases can be illustrated as follows:

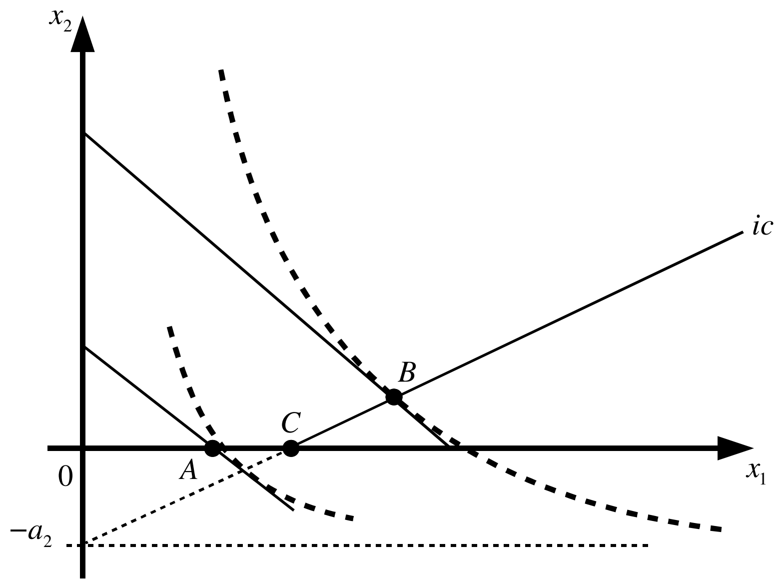


Diagram 1.2.2: Optimal consumption for different levels of income

[10]An alternative illustration of the maximization calculus by choosing a λ such that the expenditures is equal to income, is presented in the appendix A2.

Solution denoted by *A* indicates the optimal solution bundle $(x_{1a}, x_{2a})^*$ for income y_a, in that case $(x_{1a}, x_{2a})^* = (y_a / p_1, 0)$. The solutions denoted by *B* indicates the optimum solution levels $(x_{1b}, x_{2b})^*$ that are both positive. Both solutions correspond to scenario described above and illustrated in diagram 1.2.1 in a different way. Again a_1 and a_2 play an important role. As illustrated by diagram 1.2.2, the slope of the isoquant goes to infinity when X_1 goes to $-a_1 = 0$ and goes to zero when X_2 goes to $-a_2$. The reason for this is $\lim_{x_1 \to 0} \partial U / \partial X_1 = \infty$ and $\lim_{x_2 \to -a_2} \partial U / \partial x_2 = \infty$. Due to this, X_1 can never be zero for any income and price, while X_2 can be zero for some prices and incomes.
Curve ic is the income consumption path. Since households preferences are homothetic[11], the income consumption part is linear increasing above a certain level of income, denoted by y_c. The corresponding budget line would cross point *C*. Below income level y_c consumption path is horizontal, since X_2 is always bound to zero.

Up until now all solutions were presented for given values ξ_1 and ξ_2. When a couple of households would be observed, then even if all parameters a_1, a_2, m_1, m_2, d_1, d_2, prices p_1, p_2 and income y would be given, optimal values (X_1, X_2) will vary between the households. This is because the utility functions of the households depends on ξ_1 and ξ_2 and therefore also the optimal values (X_1, X_2) will be random variables too. The following diagram illustrates the solution for twenty households with equal parameters but different realizations of random terms ξ_1 and ξ_2. In top left diagram the optimal consumption bundle for a given preference is illustrated. Diagrams in the second columns are histograms of the amounts consumed by the different households. Top right diagram is just for projecting the consumption values of realized values x_2 from the vertical axis to the horizontal axis.

[11] Note, that the utility function is homothetic with respect to $(x_i + a_i)$ and therefore the income consumption curve for $(x_i + a_i)$ is a straight line increasing line in the case where x_i is not bounded. The kink at point C is because of x_2 is bounded to zero for income levels lower than y_c.

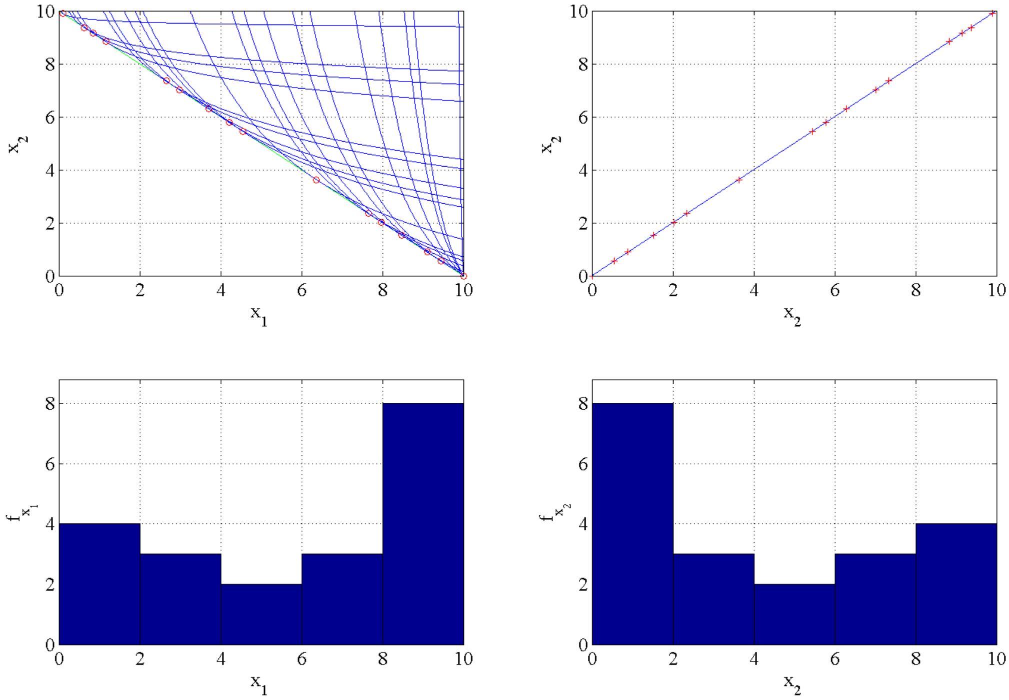


Diagram 1.2.3: Distribution of optimal consumption for twenty households with identical parameters[12]

Diagram above shows optimal consumption values for households that all have the same income *y*=10, are faced with the same prices, $p_1 = p_2 = 1$, have the same deterministic part of the preference, $m_1 = m_2 = 1$ and have the same shape parameters a_1, a_2, d_1 and d_2. Difference in optimal consumption only arises due to different realizations of the random terms ξ_1 and ξ_2. The solution shows a wide range of optimal values. Note that there are a couple of bounded solutions, $x_2 = 0$. Bounded solutions result if relative preference for driving a car - $\exp(m_2 + \xi_2)/\exp(m_1 + \xi_1)$ [13] - is so small that households will choose not to own a car.

The aim of this model is now to estimate the values of the parameters a_1, a_2, m_1, m_2, d_1, d_2, when the prices p_1, p_2 and the incomes y_n and the values $(x_1, x_2)_n$ of households $n = 1,2,..,N$ are observed. This will be done by Maximum Likelihood estimation: Parameters have to be changed such that the probability of observing $(x_1, x_2)_{n=1,2,..,N}$ is maximized,

---

[12]The parameters were chosen as follows: $a_1 = 0, a_2 = 2, d_1 = d_2 = 0.001, m_1 = m_2 = 1$ and the prices $p_1 = p_2 = 1$.

[13]Note, that utility function is ordinal and therefore equivalent to any positive transformation. Applying transformation $f(u) = (\exp(m_1 + \xi_1))^{-1} \cdot u$ yields $u(x_1, x_2) = (x_1 + a_1)^{d_1} + \exp(m_2 + \xi_2)/\exp(m_1 + \xi_1) \cdot (x_2 + a_2)^{d_2}$. Therefore $\exp(m_2 + \xi_2)/\exp(m_1 + \xi_1)$ can be considered as weight of $(x_2 + a_2)^{d_2}$ relative to $(x_1 + a_1)^{d_1}$.

$P\left(\left(X_1 = x_{1n},\, X_2 = x_{2n}\right)_{n=1,2,..,N} \mid \theta,\, p_1,\, p_2,\, y_{n=1,2,..,N}\right)$, where $\theta = \left(a_1, a_2, d_1, d_2, a_1, a_2, m_1, m_2\right)$. (1.2.9) [14]

Since it is assumed that the observations of the households are independent, probability (1.2.9) can be rewritten as:

$$P\left(\left(X_1 = x_{1n},\, X_2 = x_{2n}\right)_{n=1,2,..,N} \mid \theta,\, p_1,\, p_2,\, y_{n=1,2,..,N}\right) = \prod_{n=1}^{N} P\left(\left(X_1 = x_{1n},\, X_2 = x_{2n}\right) \mid \theta,\, p_1,\, p_2,\, y_n\right). \quad (1.2.10)$$

This means, that probability $P\left(\left(X_1 = x_{1n},\, X_2 = x_{2n}\right) \mid \theta,\, p_1,\, p_2,\, y_n\right)$ has to be calculated. To do this, two cases have to be distinguished: Case one that gives the probability of observing a bounded solution for X_2 and case two where both optimal values X_1 and X_2 are positive. For case one, a probability and for case two a density function has to be calculated. Before calculating the probability functions for these two cases the first order conditions (1.2.7) and (1.2.8) have to be reformulated.

$$V_1 + \xi_1 = \ln(\lambda) \quad (1.2.11)$$

$$V_2 + \xi_2 \le \ln(\lambda) \quad (1.2.12)$$

with:

$$V_1 = \ln(d_1) - \ln(p_1) + m_1 - (1 - d_1) \cdot \ln(X_1 + a_1), \quad (1.2.13)$$

$$V_2 = \ln(d_2) - \ln(p_2) + m_2 - (1 - d_2) \cdot \ln(X_2 + a_2). \quad (1.2.14)$$

Case 1: Only good one is consumed

This means, that the realisation of good two is bounded, $X_2 = 0$. Therefore (1.2.12) is a strict inequality. By plugging (1.2.11) in (1.2.12), the following inequality has to be fulfilled:

$$\xi_2 < V_1 - V_2 + \xi_1, \quad (1.2.15)$$

Since from $X_2 = 0$ follows that the whole budget is spent for good one yielding $X_1 = y/p_1$. Therefore values V_1 and V_2 are known and fixed. Derivating probability of observing $X_2 = 0$ will be done in two steps. First step is to compute probability of (1.2.13) conditional on ξ_1:

$$P\left(\xi_2 < V_1 - V_2 + \xi_1 \mid \xi_1\right) = F_{\xi}\left(V_1 - V_2 + \xi_1\right) \quad (1.2.16)$$

---

[14] Note that in fact both both X_1 and X_2 are continuous random variables, but with some discrete probability at $X_2 = 0$, $X_1 = y/p_1$ respectively. For a simpler formulation, notation $P\left(\left(X_1 = x_{1n},\, X_2 = x_{2n}\right)_{n=1,2,..,N} \mid \theta,\, p_1,\, p_2,\, y_{n=1,2,..,N}\right)$ is used.

By integrating out ξ_1, the unconditional probability $P(\xi_2 < V_1 - V_2 + \xi_1)$ can be calculated:[15]

$$P(\xi_2 < V_1 - V_2 + \xi_1) = \int_{z=-\infty}^{\infty} F_\xi\left(V_1(x_1) - V_2(x_1) + z\right) \cdot f_\xi(z)\,dz = \frac{e^{V_1}}{e^{V_1} + e^{V_2}}. \quad (1.2.17)\ ^{16}$$

Therefore probability of observing $X_1 = y/p_1$ and $X_2 = 0$ is

$$P\left(X_1 = \frac{y}{p_1} \wedge X_2 = 0\right) = \frac{e^{V_1}}{e^{V_1} + e^{V_2}}. \quad (1.2.18)\ ^{17}$$

Note that V_1 and V_2 are equal to (1.2.13) and (1.2.14) evaluated at $X_1 = y/p_1$ and $X_2 = 0$.

Case 2: Both goods are consumed

In this case, both first order conditions (1.2.14) and (1.2.15) are equalities since optimal solution is non-bounded. From this follows

$$\xi_2 = V_1 - V_2 + \xi_1. \quad (1.2.19)$$

In a first step X_1 has to be expressed as a function of X_2 by use of budget restriction (1.2.3). Therefore V_1 becomes a function X_2:

$$\breve{V}_1(X_2) = \ln(d_1) - \ln(p_1) + m_1 - (1 - d_1) \cdot \ln\left(\frac{y - p_2 \cdot X_2}{p_1} + a_1\right), \quad (1.2.20)$$

In a second step probability of observing $X_2 = x_2$ has to be determined. Since $\breve{V}_1(x_2) - V_2(x_2)$ is strictly increasing in x_2,[18] (1.1.10), can be solved as follows:[19]

---

[15]This follows by applying the law of total probability: $P(A) = \sum_{i=1}^{n} P(A \mid B_i) \cdot P(B_i)$, if $B_i, i \in \{1,2,..,n\}$ is a finite or countably infinite partition of a probability space and each set B_n is measurable.

[16] Apply rule 6 in appendix A1.

[17]This result can also been directly yielded by applying rule 3 of appendix A1 to $\xi_2 < V_1 - V_2 + \xi_1 \Leftrightarrow \xi_2 - \xi_1 < V_1 - V_2$. Since by this rule the cumulated density function of $Z = \xi_2 - \xi_1$ is $Z \sim F_{\xi_2 - \xi_1}(z) = 1/(1+e^{z})$ it follows, that the probability of case 1 is: $P(\xi_2 - \xi_1 < V_1 - V_2) = F_{Z_1}(V_1 - V_2) = \frac{1}{1+e^{-(V_1 - V_2)}} = \frac{e^{V_1}}{e^{V_1} + e^{V_2}}$.

[18] $\frac{\partial(\breve{V}_1(z) - V_2(z))}{\partial z} = \frac{\partial \breve{V}_1(z)}{\partial z} - \frac{\partial V_2(z)}{\partial z} = \frac{\partial \breve{V}_1(X_1)}{\partial X_1} \cdot \frac{\partial X_1}{\partial z} - \frac{\partial V_2(z)}{\partial z} = \frac{1-d_1}{x_1(z) + a_1} \cdot \frac{-p_2}{p_1} - \frac{1-d_2}{z + a_2} < 0.$

[19]Applying theorem "Densities of transformed random variables" yields (1.2.21), with
$h_2(X_2) = \breve{V}_1(X_2) - V_2(X_2) + \xi_1$ and

$$f_{X_2\wedge(X_2>0)|\xi_1}(z)=\left(\frac{1-d_1}{\frac{y-p_2z}{p_1}+a_1}\cdot\frac{p_2}{p_1}+\frac{1-d_2}{z+a_2}\right)\cdot f_\xi\left(\breve{V}_1(z)-V_2(z)+\xi_1\right). \qquad (1.2.21)$$

Probability density function $f_{X_2\wedge(X_2>0)}(z)$ can be obtained by integrating out ξ_1 from (1.2.21):[20]

$$f_{X_2\wedge(X_2>0)}(z)=\left(\frac{1-d_1}{\frac{y-p_2z}{p_1}+a_1}\cdot\frac{p_2}{p_1}+\frac{1-d_2}{z+a_2}\right)\cdot\frac{e^{V_2(z)-\breve{V}_1(z)}}{\left(1+e^{V_2(z)-\breve{V}_1(z)}\right)^2}. \qquad (1.2.22)$$

Note that $f_{X_2}(z)$ refers to the case where X_2 is not bounded, which means that $\xi_2=V_1(X_2)-\breve{V}_2(X_2)+\xi_1$ is always true. Note that $f_{X_2}(y/p_2)$ goes to infinity, but that does not imply, that the probability for $P(X_2=y/p_2)=P(X_1=0)$ does have a finite value, what would be in contrast to that parameter a_1 forces the possible range for optimal solutions of X_1 to values greater than zero, $-a_1=0$.[21] Note also, that integrating (1.2..22) yields probability $X_2>0$, $P(X_2>0)=1-P(X_2=0)$.[22]

$$|J(x_2)|=\left|\frac{\partial\xi_2}{\partial x_2}\right|=\left|\frac{\partial\left(\breve{V}_1(x_2)-V_2(x_2)\right)}{\partial x_2}\right|=\left|\breve{V}_1'(x_2)-V_2'(x_2)\right|=\frac{1-d_1}{\frac{y-p_2z}{p_1}+a_1}\cdot\frac{p_2}{p_1}+\frac{1-d_2}{z+a_2}$$ yields (1.2.21).

[20] Applying rule 8 in A.1 on $f_\xi\left(\breve{V}_1(z)-V_2(z)+\xi_1\right)$ yields this results.

[21] For a proof see appendix A3.

[22]Proof: Since (1.2.19), $\xi_2=\breve{V}_1(X_2)-V_2(X_2)+\xi_1$. Since $\breve{V}_1(X_2)-V_2(X_2)$ is strictly increasing in X_2, $\breve{V}_1(X_2)-V_2(X_2)$ is greater than $\breve{V}_1(0)-V_2(0)$ for any positive value X_2. Therefore ξ_2 that always has to match $\xi_2=\breve{V}_1(X_2)-V_2(X_2)+\xi_1$, will take any value greater than $\breve{V}_1(0)-V_2(0)$, when X_2 is increasing from zero to its maximal value $X_2=y/p_2$. Therefore integrating (1.2.22) over X_2 is equivalent to probability $\xi_2>\breve{V}_1(0)-V_2(0)+\xi_1$. This probability is counter probability of (1.2.17) and therefore $P(X_2>0)=1-P(X_2=0)$.

The graph of function $f_{X_2|(X_2>0)}(z) = f_{X_2\wedge(X_2>0)}(z) \cdot P(X_2 > 0)^{-1}$ is as follows:

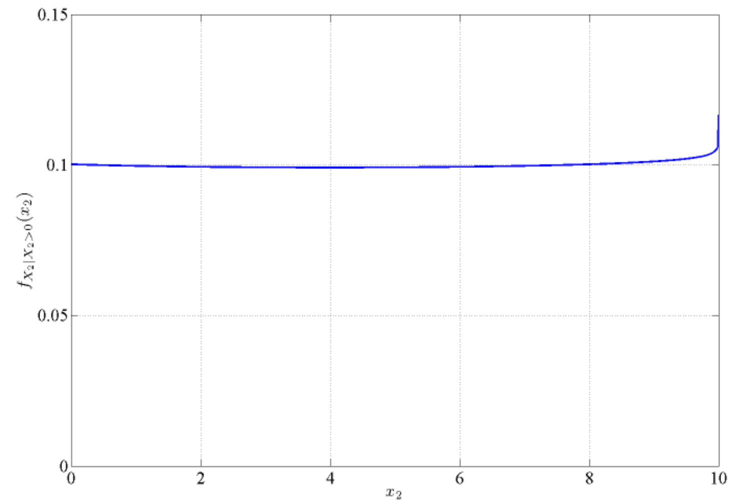


Diagram 1.2.4: Probability density function for good two

The density is rather like a uniform distribution with a sharp increase at the boundary of the feasible range of X_2. In the next section it will be shown how this shape changes when the variance of the error term is changed. Before the effect of such a change is discussed, first this case will be brought to an end. From previous results, $P((X_1 = x_{1n}, X_2 = x_{2n}) | \theta, p_1, p_2, y_n)$ can now be calculated.[23] Before this is done, this probability can be restated, since X_1 is a function of X_2 and y. Therefore the probability measure (1.2.9) for the ML estimation is:

$$P(z = X_2 | \theta, p_1, p_2, y) = f_{X_2\wedge X_2>0}(z) \cdot I(z>0) + P(X_2 = 0) \cdot I(z=0) \qquad (1.2.23),^{24}$$

where $I(z>0)$ and $I(z=0)$ are indicator functions being one, when the argument is true and zero otherwise. The probability $P(X_2 = 0)$ is defined in (1.2.18) and the density $f_{X_2|X_2>0}(z)$ is defined in (1.2.22). By plugging (1.2.23) in the ML function (1.2.10) all parameters can be estimated for given data $(x_{1n}, x_{2n}, y_n)_{n=1,2,..,N}$.

[23] Note that in the following the observed variables are noted without index *n*.

[24] Note that also the probabilities $P(X_2 > 0)$, $P(X_2 = 0)$ and the probability density function $f_{X_2|X_2>0}(z)$ are also conditional on the parameter values but in this notation it is not explicitly noted.

Change of the variance of the error term

As diagram 1.2.4 shows the solutions of was shown in X_2 conditional on that they are positive is almost uniformly distributed over the interval of feasible solutions. This means that unobserved preferences have a large impact on the outcomes even when income and prices are the same for all households. It seems reasonable that the spread of solutions could be smaller in reality, i.e. there is more concentration of solutions around a value within the interval of feasible solutions of X_2. This outcome should now also replicated by the model. In the following it is shown that a concentration of realisations X_2 around a certain value - means that the density function will reach a maximum at this value - can be realised by decreasing the variance of the error terms ξ_1, ξ_2. The following diagram shows the outcome, when the variance of both ξ_1, ξ_2 is reduced by factor 0.36 all parameters, prices and income being the same as in diagram 1.2.3.

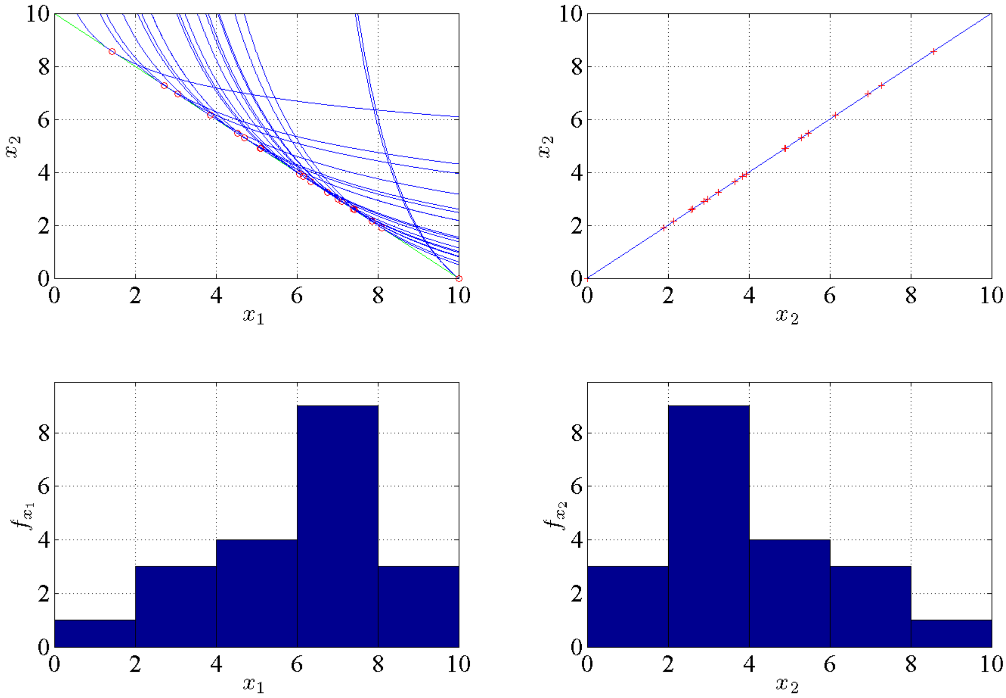


Diagram 1.2.5: Distribution of optimal consumption when variance in error terms is reduced[25]

This diagram shows that the realizations of X_2 are now more concentrated around the value $X_2 = 3$. Further the probability of X_2 being zero is strongly reduced. Again the density function of X_2 and the

[25]The parameters were chosen as follows: $a_1 = 0, a_2 = 2, d_1 = d_2 = 0.001, m_1 = m_2 = 1$ and the prices $p_1 = p_2 = 1$. The variance of variance of both ξ_1, ξ_2 - that are iid standard Gumbel distributed - is reduced by factor 0.36.

probability X_2 being zero is calculated for use for the ML function. In the following the error term will be $\beta\xi_1, \beta\xi_2$ instead of ξ_1, ξ_2.

Case 1: Only good one is consumed, changed variance of error terms

This means, that the realisation of good two is bounded, $X_2 = 0$. Again the situation is the same as in the case where the error terms are ξ_1, ξ_2, but now the error terms are $\beta\xi_1, \beta\xi_2$.

$$\beta\xi_2 < V_1 - V_2 + \beta\xi_1, \tag{1.2.24}$$

with V_1 and V_2 being the same functions as in the case before, namely:

$$V_1 = \ln(d_1) - \ln(p_1) + m_1 - (1-d_1)\cdot\ln(X_1 + a_1), \tag{1.2.25}$$

$$V_2 = \ln(d_2) - \ln(p_2) + m_2 - (1-d_2)\cdot\ln(X_2 + a_2). \tag{1.2.26}$$

By use of rule 3 of Gumbel distributed random variables as stated in the attachment A 1, the probability $P(\beta\xi_1 - \beta\xi_2 < V_1 - V_2)$ can be calculated straight forward:

$$P(\beta\xi_1 - \beta\xi_2 < V_1 - V_2) = F_{\beta\xi_1 - \beta\xi_2}(V_1 - V_2) = \frac{1}{1 + e^{-\frac{1}{\beta}(V_1 - V_2)}} \tag{1.2.27}$$

Case 2: Both goods are consumed

Also in this case the situation is the same as in the case where the error terms are ξ_1, ξ_2, but now with error terms are $\beta\xi_1, \beta\xi_2$.

$$\beta\xi_2 - \beta\xi_1 = V_1 - V_2 \tag{1.2.28}$$

In a first step X_1 has to be expressed as a function of X_2 by use of the budget restriction. Therefore V_1 and V_2 are again like in the previous case:

$$V_1(X_2) = \ln(d_1) - \ln(p_1) + m_1 - (1-d_1)\cdot\ln\left(\frac{y - p_2\cdot X_2}{p_1} + a_1\right), \tag{1.2.29}$$

$$V(X_2) = \ln(d_2) - \ln(p_2) + m_2 - (1-d_2)\cdot\ln(X_2 + a_2). \tag{1.2.30}$$

Now the density $f_{X_2 \wedge (X_2>0)}$ can be calculated by applying property 3 of attachment A 1.[26]

$$f_{X_2 \wedge (X_2>0)}(z) = f_{\beta\xi_1-\beta\xi_2}\left(V_1(z)-V_2(z)\right)\cdot\frac{d\left(V_1(z)-V_2(z)\right)}{dz} = \tag{1.2.31}$$

$$=\frac{1}{\beta}\cdot\frac{e^{-\frac{1}{\beta}\left(V_1(z)-V_2(z)\right)}}{\left(1+e^{-\frac{1}{\beta}\left(V_1(z)-V_2(z)\right)}\right)^2}\cdot\left(\frac{1-d_1}{\left(\frac{y-p_2\cdot X_2}{p_1}+a_1\right)^{d_1}}\cdot\frac{p_2}{p_1}+\frac{1-d_2}{z+a_2}\cdot\left(\frac{y-p_2\cdot X_2}{p_1}+a_1\right)^{1-d_1}\right).$$

The shape of the density of X_2 changes now from u-shape for high values β to hump shape for low values β. For $\beta=2$, $\beta=1$ and $\beta=0.6$ the probability density functions are as follows:

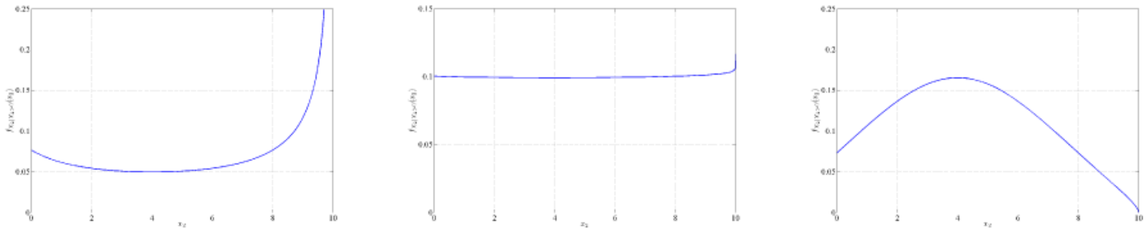


Diagram 1.2.6: Probability density function for good two for different variances of the error term

This diagram shows that the smaller the variance of the error term the more concentrated the realisations are around the level $X_2=4$. When the variance of the error term goes to zero, the realisations will converge to level $X_2=4$ - the level when there are no stochastic components ξ_1 and ξ_2 - with probability one.

[26] $f_{X_2 \wedge (X_1>0)}(z)=\lim_{dz\to 0}\left(\frac{P(z<X_2<z+dz)}{dz}\right)=\lim_{dz\to 0}\left(\frac{P\left(V_1(z+dz)-V_2(z+dz)<V_1(X_2)-V_2(X_2)<V_1(z)-V_2(z)\right)}{dz}\right)=$

$=\lim_{dz\to 0}\left(\frac{P\left(V_1(z+dz)-V_2(z+dz)<\beta\xi_1-\beta\xi_2<V_1(z)-V_2(z)\right)}{dz}\right)=\lim_{dz\to 0}\left(\frac{F_{\beta\xi_1-\beta\xi_2}\left(V_1(z)-V_2(z)\right)-F_{\beta\xi_1-\beta\xi_2}\left(V_1(z+dz)-V_2(z+dz)\right)}{dz}\right)=$

$=-\frac{dF_{\beta\xi_1-\beta\xi_2}\left(V_1(z)-V_2(z)\right)}{dz}=-f_{\beta\xi_1-\beta\xi_2}\left(V_1(z)-V_2(z)\right)\cdot\frac{d\left(V_1(z)-V_2(z)\right)}{dz}.$

troduction of the Model, (This file: R:\bhat\bhat_doku_rt\Bhat_STRC_2009_II_ver2.odt, Sat 29 Aug 09, 00:10:33)
1.2 Model with two goods and no fixed cost

1.3 Model with two goods and fixed cost for good two

1.3.1 Model setup and calculation of the ML function

In this case car driving, denoted by good two, is connected with fixed cost when consuming any positive amount. The household has to decide if it wants to bear the fixed costs or if it wants to spend all income on the consumption good one only. In latter case it has the disadvantage of decreasing marginal utility of consumption good and in former case it can spread the income on two goods and yield higher marginal utility but has to bear disutility by disposable income lowered by the fixed cost caused by car ownership. The impact of changes in income, fixed cost and prices on household's decision on consumption has already discussed in the introductory section. In the following, the focus shall now be on the impact of changes of household's preference for car driving. To understand this impact and its implications is crucial when deriving the observation probabilities of household's consumption choices. In the following optimal consumption decision for households with equal income at given prices but with different preferences for good two car driving, is illustrated. Note, that variable x_2 denotes the annual kilometers driven by the household.

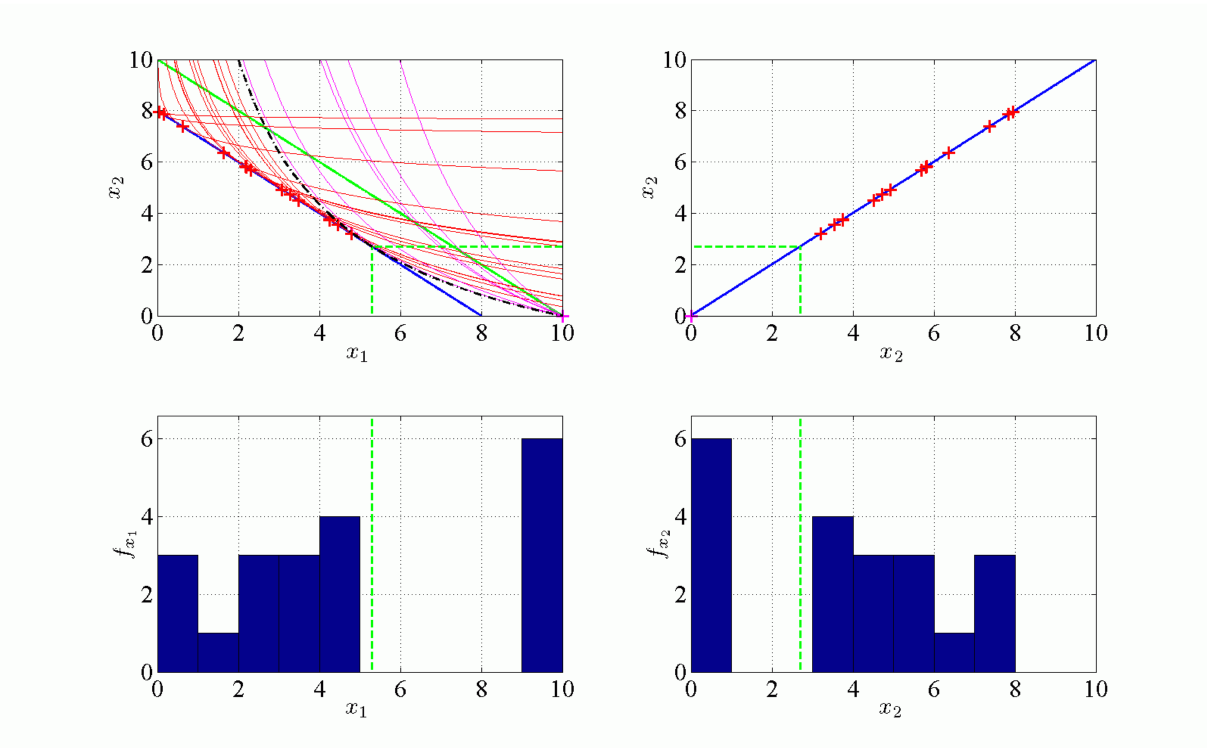


Diagram 1.3.1: Optimal consumption for households with different preferences

Top left diagram shows budget line and iso-utility functions at maximal utility of households with different preference for car driving. The higher this preference the higher are optimal values x_1 and therefore the lower are optimal values x_2 and the more to the bottom-right the iso-utility functions shift. When preference for car driving gets below a certain level, households will decide not to own a car any more and to spend all income on consumption good one. The point, where households switch between these two types of consumption schemes is indicated by the green dashed line and the dash-dotted iso-utility function. This case shows also, that for given income and fixed cost the realizations of good two will either be zero or above a certain limit that is indicated by this boundary. This is plausible: For instance no household would bear the fixed costs of holding a car and then not driving it. Iso-utility functions corresponding to utility functions with preferences lower than this threshold level are illustrated by the magenta coloured curves crossing at consumption bundle $x_1 = 10, x_2 = 0$.

Before deriving probabilities for observing household's consumption decisions, household's maximization problem shall be formulated for this case. Utility function and distribution are still the same as in the case without fixed cost described in previous section, see (1.2.1) and (1.2.2). For this model, budget restriction has changed:

$$y = p_1 \cdot x_1 + k_2 \cdot F_\kappa(x_2) + p_2 \cdot x_2 \tag{1.3.3}$$

Note that for this case there are now fixed costs assumed for good two. Function $F_\kappa(\bullet)$ is an indicator function, being one if the argument is positive and being zero otherwise. Maximization of the household can be represented by solving the following Lagrangian:

$$\begin{aligned} L = \exp(m_1 + \xi_1) \cdot (X_1 + a_1)^{d_1} + \exp(m_2 + \xi_2) \cdot (X_2 + a_2)^{d_2} \\ + \lambda \cdot (y - p_1 \cdot X_1 - k_2 \cdot F_\kappa(X_2) - p_2 \cdot X_2), \end{aligned} \tag{1.3.4}$$

$X_1 > 0, X_2 \geq 0$.

As in the previous section, ξ_1 and ξ_2 represent unobserved characteristics and are treated as given since it is assumed that these values are known by the households.

The corresponding first order conditions can be written as:

$$d_1 \cdot \exp(m_1 + \xi_1) \cdot \frac{1}{(X_1 + a_1)^{1-d_1}} = \lambda \cdot p_1 \tag{1.3.5}$$

$$d_2 \cdot \exp(m_2 + \xi_2) \cdot \frac{1}{(X_2 + a_2)^{1-d_2}} \leq \lambda \cdot (k_2 \cdot f_\kappa(X_2) + p_2) \tag{1.3.6}$$

Function $f_\kappa(X_2)$ is the derivative of function $F_\kappa(X_2)$ and goes to infinity at $X_2 = 0$, $\lim_{X_2 \to \infty} f_\kappa(X_2) = \infty$. Therefore, there are always two solutions, $X_2 = 0$ and $X_2 \neq 0$. These two types of solutions are denoted - like in previous section - as “case 1” and “case 2”. Note that in case two, solutions $X_2 < 0$ will be bound to zero. Indicator function $F_\kappa(X_2)$ provides nothing else than that for solutions $X_2 > 0$ household's available income is reduced by fixed costs k_2.

In the following, the decision between consuming both goods or only good one is discussed in more detail. The two diagrams below show the two cases for which households decide for case one, where all budget is used to consume good one, and case two, where households decide to buy a car and to bear its fixed cost. Below diagram is already presented in top-left part of diagram (1.3.1) in a similar way. Again in both diagrams below, income, fixed cost and price are identical for both households. On the left side, household has a higher preference for driving, while as on the left side household has lower preference for car driving.

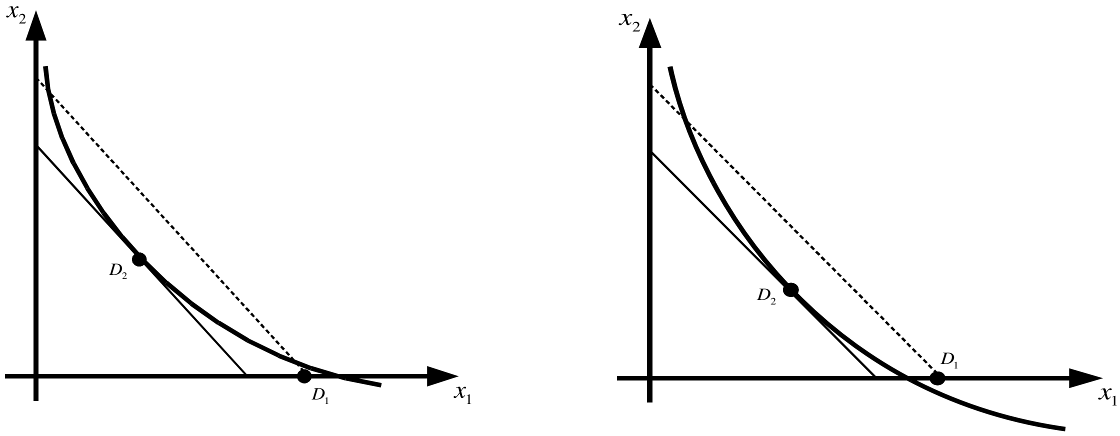


Diagram 1.3.2: Optimal consumption for given fixed cost for good two x_2

The two straight lines depict the budget lines. The solid budget line corresponds to the income net the fixed cost for good two for the case of positive consumption of good two and the dashed line corresponds to the income that is relevant when x_2 is chosen to be zero and therefore there are also no fixed cost for that good. In the left diagram a household with strong preference for car driving chooses to buy a car (case 2). This is due to decreasing marginal utility of good one. Because of this, spending the income on both goods but accepting the loss of available income due to fixed costs yields a higher utility. In the situation illustrated in diagram on the right household has low preference for car driving. The loss of available income when owning a car has a stronger effect than the gain by spreading consumption on both goods. Therefore spending the whole income on good one yields a higher utility (case 1).

The following diagram offers an alternative illustration of that decision:

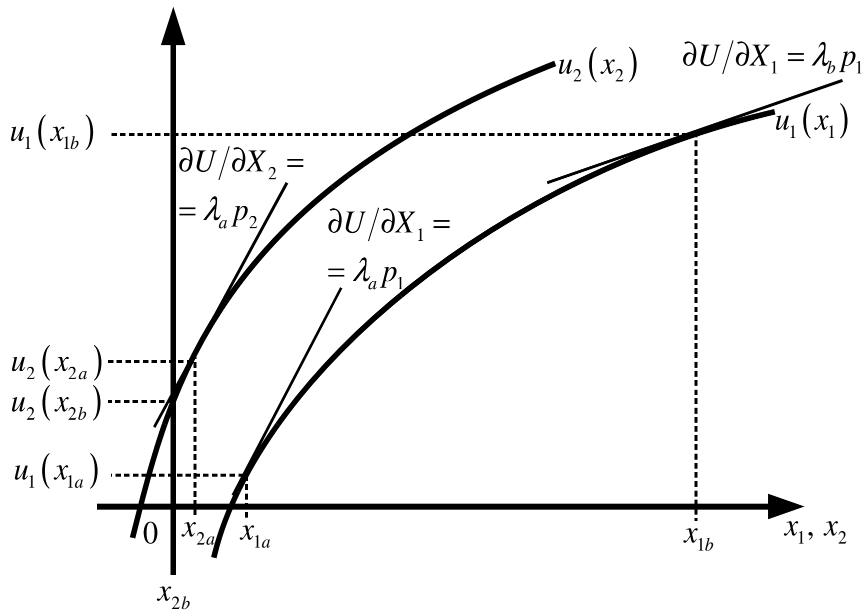


Diagram 1.3.3: The maximization calculus[1]

Here the results for two values λ are illustrated, $\lambda_a > \lambda_b$. For λ_a the optimal value of x_2 is positive since for $\partial U/\partial X_1 = \lambda_a p_1$ the value for x_2 is positive. Variable λ_a is chosen so that the whole disposable income $y - k_2$ is used up for consumption. Variable λ_b is the value corresponding the solution when the whole income is spent for good one and there are no fixed cost, $X_1 = y/p_1$, $X_2 = 0$. In the example illustrated the case when all income is spent on good one yields higher utility than case when the income is spent for both goods and fixed cost have to be beared: Diagram above shows, that sum of partial utilities of amounts x_{1a} and x_{2a} corresponding to λ_a (case 2) yields lower utility than sum of partial utilities of amounts x_{1b} and x_{2b} corresponding to “case 1”:

$$u_1(x_{1b}) + u_2(x_{2b}) > u_1(x_{1a}) + u_2(x_{2a}) \Leftrightarrow u(X_1 = x_{1a} = y/p_1, X_2 = 0) > u(X_1 = x_{1a}, X_2 = x_{2a}).$$

Calculation of the Maximum Likelihood function

Again probability for observing $X_2 = 0$ and the density function for observing $X_2 > 0$ has to be calculated for ML function (1.2.10). Also when fixed costs are included in the model, probability function $P(X_2 = 0)$ (case 1) and density function $f_{X_2 \wedge (X_2 > 0)}(z)$ (case 2) have to be computed. The difficulty in this extended model is that X_2 being zero cannot only result when the realization of the random parameters ξ_1 and ξ_2 are such that for given disposable income yields a boundary solution for X_2. For this model, X_2 can also be zero when optimal solution of X_2 at disposable income $y - k_2$ is an interior solution (case 2) but saving the fixed cost for good one and spending the whole income on

[1]Note that $u_i(x_i) = \exp(m_i + \xi_i) \cdot (x_i + a_i)^{d_i}$ is partial utility of good *i*.

good one (case 1) would yield a higher income. For this latter comparison utility levels have to be computed for both cases. In the following first optimal consumption values and utilities are computed. After that, condition when utility for case 1 is greater than utility for case 2 is stated and probabilities are computed. By use of this condition it will be possible to compute the probability that case 1 is observed.

Case 1: Only good one is consumed

This means, that the realisation of good two is bounded, $X_2 = 0$. For this case, utility is

$$U\left(y/p_1, 0\right) \mid \theta, \xi_1, \xi_2 = \exp\left(m_1 + \xi_1\right) \cdot \left(y/p_1 + a_1\right)^{d_1}. \tag{1.3.7}$$

Case 2: Both goods are consumed

In this case, like in the case without fixed cost, condition (1.2.19) holds but now available income is reduced to income minus fixed cost, $y - k_2$. Therefore function (1.2.20) changes to

$$\breve{\breve{V}}_1\left(X_2\right) = \ln\left(d_1\right) - \ln\left(p_1\right) + m_1 - \left(1 - d_1\right) \cdot \ln\left(\frac{y - k_2 - p_2 \cdot X_2}{p_1} + a_1\right), \tag{1.3.8}$$

and density (1.2.22) changes to

$$f_{X_2 \wedge (X_2 > 0)}\left(z\right) = \left(\frac{1 - d_1}{\dfrac{y - k_2 - p_2 z}{p_1} + a_1} \cdot \frac{p_2}{p_1} + \frac{1 - d_2}{z + a_2}\right) \cdot \frac{e^{V_2(z) - \breve{\breve{V}}_1(z)}}{\left(1 + e^{V_2(z) - \breve{\breve{V}}_1(z)}\right)^2}. \tag{1.3.9}$$

Now, probability $P\left(X_2 = 0\right)$ has to be determined. Note, that X_2 can be zero for two reasons. The first reason is, that interior solution in case 2 can yield negative values for X_2. In the following, this first reason is denoted as condition one and corresponds to[2]

$$\xi_2 < \breve{\breve{V}}_1 - V_2 + \xi_1. \tag{1.3.10}$$

The second reason is that consumption scheme in case 2 can yield lower utility than spending all income for the consumption good number one as in case 1. In the following, this second reason is denoted as condition two. Condition two is equivalent to comparing the sum of partial utilities $u_1\left(x_{1b}\right) + u_2\left(x_{2b}\right)$ to $u_1\left(x_{1a}\right) + u_2\left(x_{2a}\right)$ in diagram 1.3.3. The function that would yield the utility of

[2]See also "case 1" in section 1.2.

case 2 for any income, price and preference parameters is the indirect utility function. Using this function, condition two would be then:

$$U\left(y/p_1,0\right)|\,\theta,\xi_1,\xi_2 > v\left(y-k_2,p_1,p_2\right)|\,\theta,\xi_1,\xi_2\,, \qquad (1.3.11)$$

where $\theta=\left(a_1,\,a_2,\,d_1,\,d_2,m_1,\,m_2\right)$.

$U(\bullet)$ denotes the direct utility function and $v(\bullet)$ the indirect utility function provided that optimal consumption levels of the two goods are not bound to zero.

It is important to note, that condition two is not equivalent to condition one. The following diagram where the decision of two households is presented illustrates this fact. Both households face the same income and prices and differ only in preference for car driving. First household on the left has preference for car driving such that it is indifferent between owning a car an bearing the fixed cost and spending all income for consumption good one. Second household's preference for car driving is so that it would not drive car when being forced to hold a car. This household's preference corresponds to the case, where case two would yield a boundary solution $X_2=0$.[3]

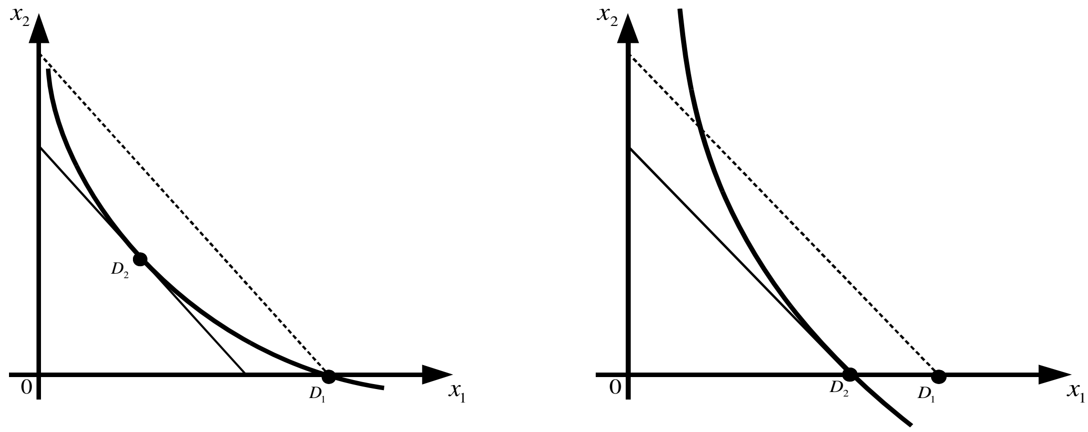


Diagram 1.3.4: Optimal consumption at critical preference levels for car driving

Diagrams above show that preference for car driving where households would switch from bounded to interior solutions given they own a car is much lower than preference where household would switch from not owning to owning a car, given they take fixed costs of car ownership into account.

Note that indirect utility function is based on non-bound solutions for optimal consumption of the two goods. Therefore, negative solutions for car driving are not excluded. This leads to the question, if for

[3]Note, that the lower preference for car driving is, the more optimal consumption D_2 shifts to the bottom along the solid budget line, since optimal amount of car driving is decreasing in this case.

very low preference for car driving the value of utility function in (1.3.11) could become again larger than the value of the direct utility function at some point. The following diagram illustrates such a situation.

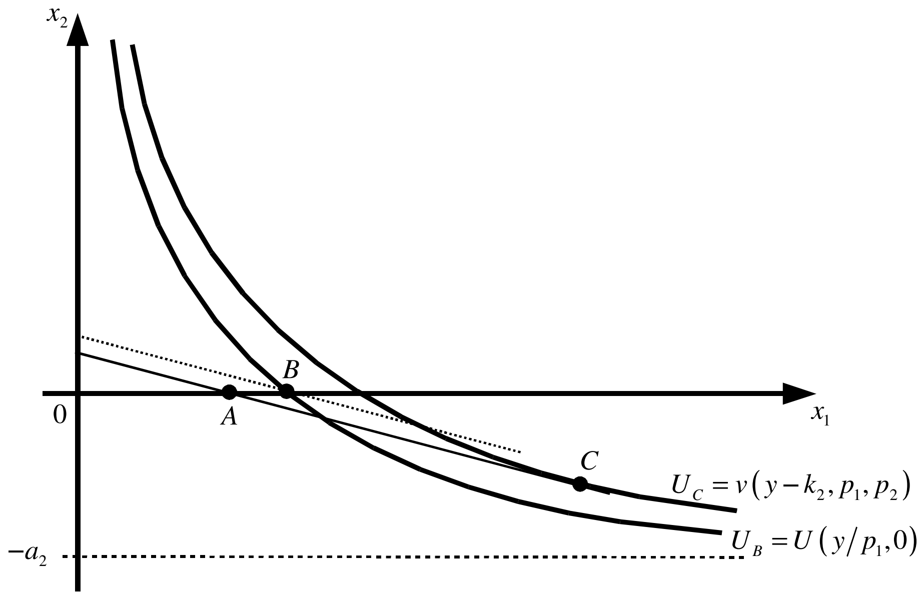


Diagram 1.3.4: Optimal consumption at critical preference levels for car driving

This diagram illustrates that for some preference level for car driving it was optimal to choose solution *A* that was inferior to solution *B*, that illustrates case 1, where all income is spent on consumption good one. Despite of that, for very small preferences for car driving, relation (1.3.11) can change again.

The question is now, if there are any parameter values of utility function (1.2.1) [4] for which a solution as illustrated in diagram 1.3.4 is possible. In the following it will be shown that for $d = d_1 = d_2$ the answer is no. For proving this, first the indirect utility function $v(y-k_2, p_1, p_2) | \theta, \xi_1, \xi_2$ has to be determined. This will be done by deriving the non-bounded Marshallian demand functions and then plugging in these in the direct utility function.

Marshallian demand function can be computed by solving Lagrangian (1.3.4). Note that this time parameters are choosen to be $d = d_1 = d_2$. This is necessary since only for $d = d_1 = d_2$ the indirect utility will be of explicit functional form. Later it will become obvious why this is important.

Solving Lagrangian (1.3.4) yields the following first order conditions:

[4] Note that (1.2.1) is $U = u_1(X_1) + u_2(X_2) = \exp(m_1 + \xi_1) \cdot (X_1 + a_1)^{d_1} + \exp(m_2 + \xi_2) \cdot (X_2 + a_2)^{d_2}$ and that this utility function is always used for MDCEV models.

$$d \cdot \exp(m_1 + \xi_1) \cdot (X_1 + a_1)^{d-1} = \lambda p_1 , \tag{1.3.12}$$

$$d \cdot \exp(m_2 + \xi_2) \cdot (X_2 + a_2)^{d-1} = \lambda p_2 , \tag{1.3.13}$$

Form these first order conditions follows[5]

$$X_1 = B \cdot (X_2 + a_2) - a_1 \text{, with } B = \left(\frac{p_2}{p_1} \cdot \exp(-m - \varsigma) \right)^{\frac{1}{1-d}} = \left(\frac{p_2}{p_1} \right)^{\frac{1}{1-d}} \cdot \exp\left(-\frac{m+\varsigma}{1-d} \right), \tag{1.3.14}$$

where $\varsigma = \xi_2 - \xi_1$ and $m = m_2 - m_1$.

Note that expression $\exp(m+\varsigma)$ denotes preference for car driving relative to preference for consumption good.[6]

Plugging (1.3.14) in the budget restriction and solving for X_2 yields:[7]

$$X_2 = \frac{y - k_2 - p_1 \cdot (B \cdot a_2 - a_1)}{p_1 \cdot B + p_2} . \tag{1.3.15}$$ [8]

Solution for X_1 can be computed by plugging (1.3.15) in (1.3.14):

$$X_1 = \frac{y - k_2 - p_1 \cdot (B \cdot a_2 - a_1)}{p_1 + p_2 \cdot B^{-1}} + B \cdot a_2 - a_1 , \tag{1.3.16}$$

with *B* as defined in (1.3.14) for both (1.3.15) and (1.3.16).

---

[5] $\frac{d \cdot \exp(m_2 + \xi_2) \cdot (X_2 + a_2)^{d-1}}{d \cdot \exp(m_1 + \xi_1) \cdot (X_1 + a_1)^{d-1}} = \frac{\lambda p_2}{\lambda p_1} \Leftrightarrow \left(\frac{X_2 + a_2}{X_1 + a_1} \right)^{1-d} = \frac{p_1}{p_2} \cdot \exp(m_2 - m_1 + \xi_2 - \xi_1) \Leftrightarrow$

$\Leftrightarrow \frac{X_2 + a_2}{X_1 + a_1} = \left(\frac{p_1}{p_2} \cdot \exp(m_2 - m_1 + \xi_2 - \xi_1) \right)^{\frac{1}{1-d}} \Leftrightarrow \frac{X_2 + a_2}{X_1 + a_1} = \left(\frac{p_1}{p_2} \right)^{\frac{1}{1-d}} \cdot \exp\left(\frac{m_2 - m_1 + \xi_2 - \xi_1}{1-d} \right).$

[6] Note that any positive transformation of utility function (1.2.1) yields the same Marschallian demand functions and leave relation same relation (1.3.17) unchanged. Dividing (1.2.1) by $\exp(m_1 + \xi_1)$ is a linear positive transformation that yields: $U(x_1, x_2) = (x_1 + a_1)^d + \exp(m_1 - m_2 + \xi_1 - \xi_2) \cdot (x_2 + a_2)^d$, if it is assumed that $d = d_1 = d_2$.

[7] $X_1 = B \cdot (X_2 + a_2) - a_1$,

$y - k_2 = p_1 \cdot (B \cdot (X_2 + a_2) - a_1) + p_2 \cdot X_2 \Leftrightarrow y - k_2 - p_1 \cdot (B \cdot a_2 - a_1) = p_1 \cdot B \cdot X_2 + p_2 \cdot X_2 \Leftrightarrow X_2 = \frac{y - k_2 - p_1 \cdot (B \cdot a_2 - a_1)}{p_1 \cdot B + p_2}$

[8] It is important to note, that relation between X_1 and X_2 would be non-linear for $d_1 \neq d_2$. Consequently expression that would yield when plugging (1.3.14) in

Plugging in these two Marshallian demand functions in the utility function yields the indirect utility function:

$$v\left(y-k_2,p_1,p_2\right)\mid\theta,\varsigma=\left(\frac{y-k_2+p_1\cdot a_1+p_2\cdot a_2}{p_1+p_2\cdot B^{-1}}\right)^d+\exp\left(m+\varsigma\right)\cdot\left(\frac{y-k_2+p_1\cdot a_1+p_2\cdot a_2}{p_1\cdot B+p_2}\right)^d. \quad (1.3.17)$$

Therefore condition two $v\left(y-k_2,p_1,p_2\right)\mid\theta,\varsigma>U\left(y/p_1,0\right)\mid\theta,\varsigma$ (1.3.11) can be rewritten as:

$$\begin{aligned}&\left(\frac{y-k_2+p_1\cdot a_1+p_2\cdot a_2}{p_1+p_2\cdot B^{-1}}\right)^d+\exp\left(m+\varsigma\right)\cdot\left(\frac{y-k_2+p_1\cdot a_1+p_2\cdot a_2}{p_1\cdot B+p_2}\right)^d>\\&>\left(\frac{y}{p_1}+a_1\right)^d+\exp\left(m+\varsigma\right)\cdot a_2{}^d.\end{aligned} \quad (1.3.18)$$

Now, the probability that these conditions are fulfilled simultaneously has to be calculated. One important feature is that both conditions depend only on one random variable, namely ς. Before the probability that both conditions are fulfilled is calculated, conditions one and two have to be transformed:

$$g_1\left(\varsigma\right)=\frac{y-k_2+p_1\cdot a_1+p_2\cdot a_2}{p_1\cdot B+p_2}-a_2, \quad (1.3.23)$$

$$\begin{aligned}g_2\left(\varsigma\right)=&\left(\frac{y-k_2-p_1\cdot\left(B\cdot a_2-a_1\right)}{p_1+p_2\cdot B^{-1}}+B\cdot a_2\right)^d+\\&+\exp\left(m+\varsigma\right)\cdot\left(\frac{y-k_2-p_1\cdot\left(B\cdot a_2-a_1\right)}{p_1\cdot B+p_2}\right)^d-\left(\frac{y}{p_1}+a_1\right)^d-\exp\left(m+\varsigma\right)\cdot a_2{}^d.\end{aligned} \quad (1.3.24)$$

Condition one and two are fulfilled, when $g_1\left(\varsigma\right)>0$, $g_2\left(\varsigma\right)>0$ respectively. It can be shown, that $g_1\left(\varsigma\right)$ and $g_2\left(\varsigma\right)$ have the following shape:[9]

[9]For a proof, see A5.

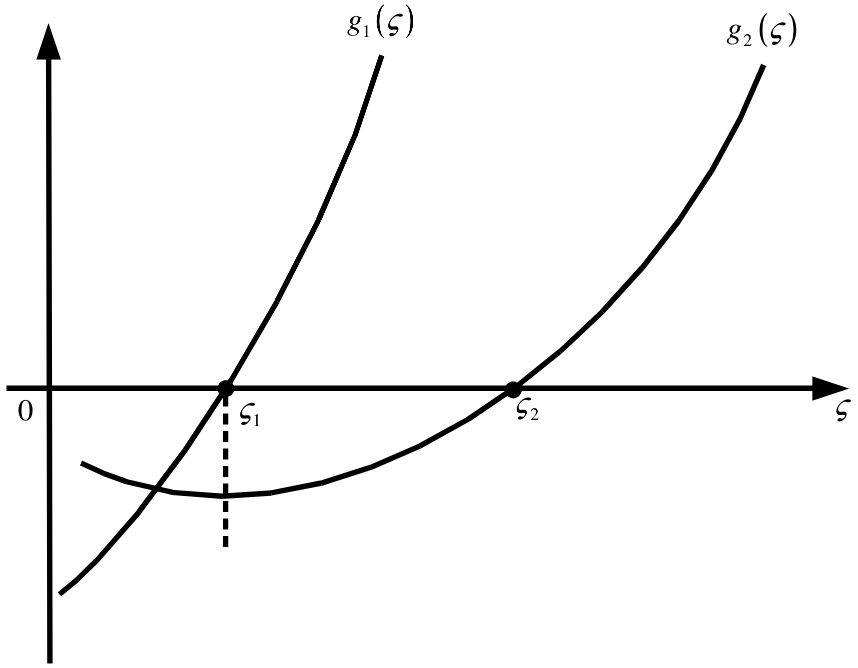


Diagram 1.3.7: The principle of calculating the probability of good two being zero

Since both function are increasing in ς in the relevant range $\varsigma > \varsigma_1$, both condition one and two are fulfilled in the range ς being above $\max(\varsigma_1, \varsigma_2)$. It can be proven,[10] that $g_2(\varsigma_2)$ is always smaller than zero and both $g_1(\varsigma)$ and $g_2(\varsigma)$ are increasing for any $\varsigma > \varsigma_1$. From this it follows that $\varsigma_2 > \varsigma_1$ and therefore $\max(\varsigma_1, \varsigma_2) = \varsigma_2$.[11] This result is rather intuitive: Condition one stands for the case when households choose whether X_2 shall be consumed regardless of fixed costs, while as in condition two, they also take into account that consuming good two implies fixed costs. It quite natural, that in latter case households will switch to consume good two only at a higher level of relative preference ς, namely at preference level $\varsigma_2 > \varsigma_1$. Given preference level ς_2, probability that households choose $X_2 = 0$ is equal to the probability that ς is smaller than this critical level:

$$P(X_2 = 0) \mid (y_D = y - k_2, p_1, p_2 \mid \theta) = \int_{\varsigma=-\infty}^{\varsigma=\varsigma_2} f_\varsigma(\varsigma) d\varsigma = F_\varsigma(\varsigma_2), \quad (1.3.25)$$

[10]See appendix A5.

[11]Note, that it is also necessary to proof that $g_2(\varsigma) < 0$ for any $\varsigma < \varsigma_1$. If this was not the case, then a case where owning a car but not driving it would be a rational choice. Since it can be proven that $g_2(\varsigma) < 0$ for any $\varsigma < \varsigma_1$ this counterintuitive case is excluded in this framework.

with $F_{\varsigma}(z)=\frac{1}{1+e^{-z}}$ being the cumulative density function of ς.

Since ς_2 is the root of $g_2(\varsigma)$, which can not be solved for ς_2 as an explicit function, ς_2 has to be computed numerically.

What remains to calculate is the density function of the X_2 being a positive solution. Again this density is the same like in the case with no fixed cost but in this case with parameter values $d_1 = d_2 = d$:

$$f_{X_2 \wedge (X_2>0)}(z) \mid (y_D = y-k_2, p_1, p_2 \mid \theta) = \left(\frac{1-d}{\frac{y-k_2-p_2 z}{p_1}+a_1} \cdot \frac{p_2}{p_1} + \frac{1-d}{z+a_2} \right) \cdot \frac{e^{V_2(z)-V_1(z)}}{\left(1+e^{V_2(z)-V_1(z)}\right)^2}, \quad (1.3.26)$$

with $V_1(z) = -\ln(p_1) - (1-d) \cdot \ln\left(\frac{y-k_2-p_2 z}{p_1}+a_1\right)$ and

$V_2(z) = \ln(d) - \ln(p_2) + m - (1-d) \cdot \ln(z+a_2)$.

Therefore the Likelihood function is as follows:

$$\ell(\theta, x_{2i}) \mid (\{y_i\}_{i=1..N}, k_2, p_1, p_2) = P(X_{2i}=0)^{(1-I_{X_{2i}})} \cdot \left(f_{X_2 \wedge (X_2>0)}(x_{2i})\right)^{I_{X_{2i}}}, \quad (1.3.27)$$

where $I_{x_2}(X_2) = \left\{ \begin{matrix} X_2 > 0 : 1 \\ X_2 \le 0 : 0 \end{matrix} \right\}$.

Note that for notational simplicity $P(X_{2i}=0)$ was not explicitly written to be conditional on parameters θ as done correctly in (1.3.25) and (1.3.26).

The shape of density (1.3.26) depends on all parameters a_1, a_2, d and economic variables.

Change of the variance of the error term

Again, also in the case where fixed costs are considered, changes in variance of error terms change the shape of probability density function. A smaller variance than in the standard Gumbel case a computed above would lead to a more realistic distribution of values in diagram 1.3.1:

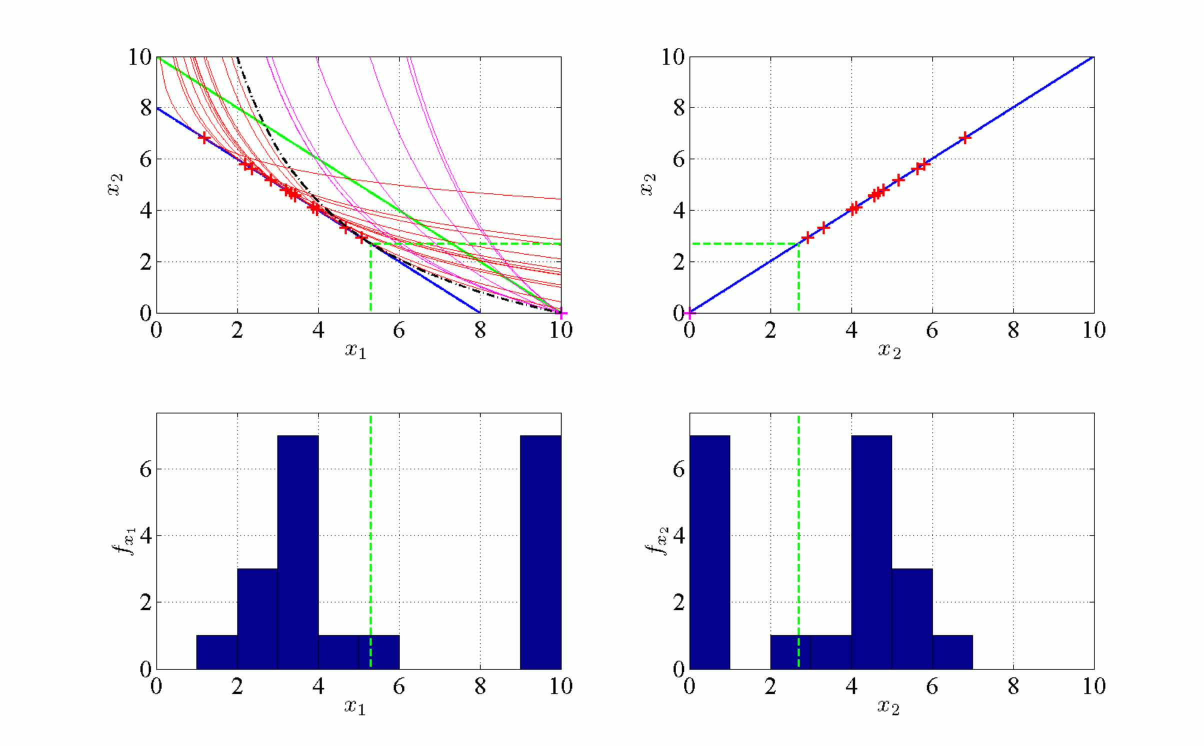


Diagram 1.3.8: Optimal consumption for households with small differences in preferences

Probability $P(X_2=0)$ and density function $f_{X_2 \wedge (X_2>0)}(z)$ can be computed as illustrated in chapter 1.2. The functions yield then:

$$P(X_2=0)\,|\,(y_D=y-k_2,p_1,p_2\,|\,\theta)=\int_{\varsigma=-\infty}^{\varsigma=\varsigma_2}\frac{1}{\beta}f_{\varsigma}\left(\frac{\varsigma}{\beta}\right)d\varsigma=F_{\varsigma}\left(\frac{\varsigma_2}{\beta}\right), \qquad (1.3.28)$$

with $F_{\varsigma}\left(\frac{z}{\beta}\right)=\frac{1}{1+e^{-\frac{z}{\beta}}}$ being the cumulative density function of ς and

$$f_{X_2 \wedge (X_2>0)}(z)\,|\,(y_D=y-k_2,p_1,p_2\,|\,\theta)= \qquad (1.3.28)$$

$$=\frac{1}{\beta}\cdot\left(\frac{1-d}{\frac{y-k_2-p_2z}{p_1}+a_1}\cdot\frac{p_2}{p_1}+\frac{1-d}{z+a_2}\right)\cdot\frac{e^{\frac{1}{\beta}(V_2(z)-V_1(z))}}{\left(1+e^{\frac{1}{\beta}(V_2(z)-V_1(z))}\right)^2}.$$

1.3.2 Parameter estimation and empirical results

In this section first the basic principle and some specific problems in ML estimating the parameters of this model are discussed. Second the estimation procedure is described. Third some information on data and their use for the model are given and estimation results (of different model specifications: not done yet) are presented.

The basic principle of ML estimation

ML estimation yields to maximize probability of observing data of a dataset. This is done by changing the parameters such that the probability function of the model fits the observed data the "best".

Following histogram shows distance driven from Swiss households in an income category about 84'000 swiss francs per year. Households live in a urban area.

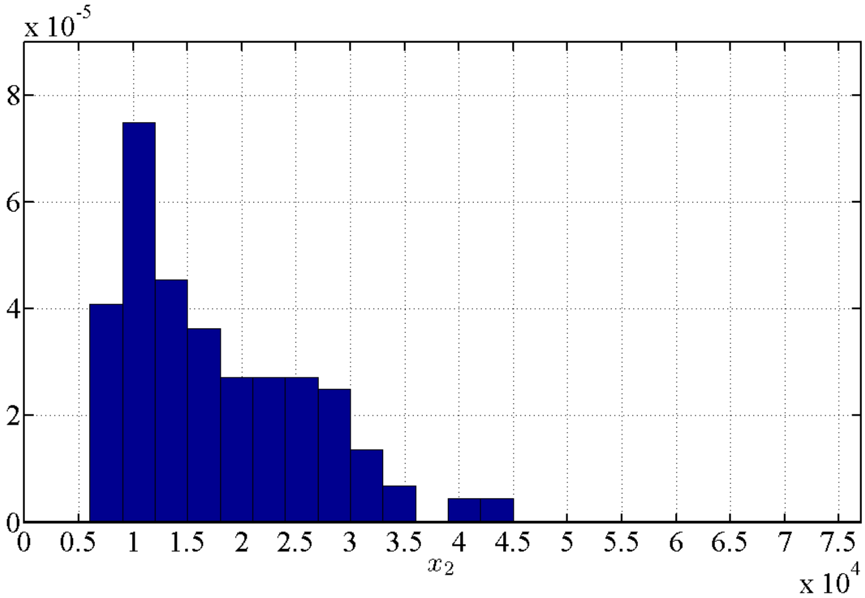


Diagram 1.3.9: Observed data of X_2 of households with income 84000Fr living in urban area

Note that in diagram above observations where households do not drive a car are filtered out. The aim is now in principle to choose parameters so that the shape of density function (1.3.28) fits best this histogram and at the same time that probability (1.3.28) is as close as possible the share of households not owning a car. It is important to note that ML estimation tries - in procíple - to fit all histograms for all income categories and all places of living. The advantage of this model compared to more simple structures like the Tobit model is that the different model parameters affect the shape of the densities and the value of the probability for observing households without cars in quite different ways. This property allows for a good fit of the model with the data despite of the small number of parameters.

In the following the impact on the shape of density and the value $P(X_2 = 0) | (y_D = y - k_2, p_1, p_2 | \theta)$ is illustrated for all parameters of the set $\theta = (a_2, d, m, \beta)$ [12]. Note that the driving force of the densities below is random variable ς.

First the effect of a change of *m* shall be examined:

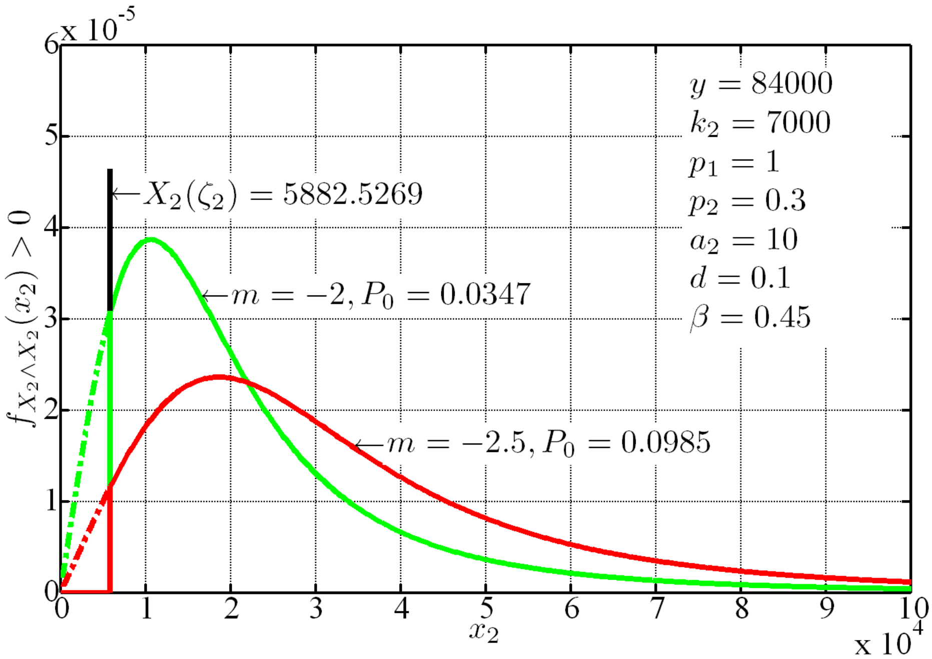


Diagram 1.3.10: pdf of X_2 or different relative (deterministic) preference *m* for car driving

This diagram shows that an increase in relative preference for car driving shifts the density function to towards the right. Higher annual kilometres driven become more likely. At the same time, probability that households do not own a car is reduced. Note that probabilities that households do not on own a car is represented by the surface below the dash-dotted lines. One important feature of a change in *m* is that it does not affect the minimum distance a household does drive when deciding to own a car.

---

[12] Note that $\theta = (a_2, d, m)$ is assumed to be zero and transformation of utility function - as presented in footnote corresponding to (1.3.14) - showed, that only difference $m = m_2 - m_1$ matters. Therefore parameter set θ has been reduced by some parameters.

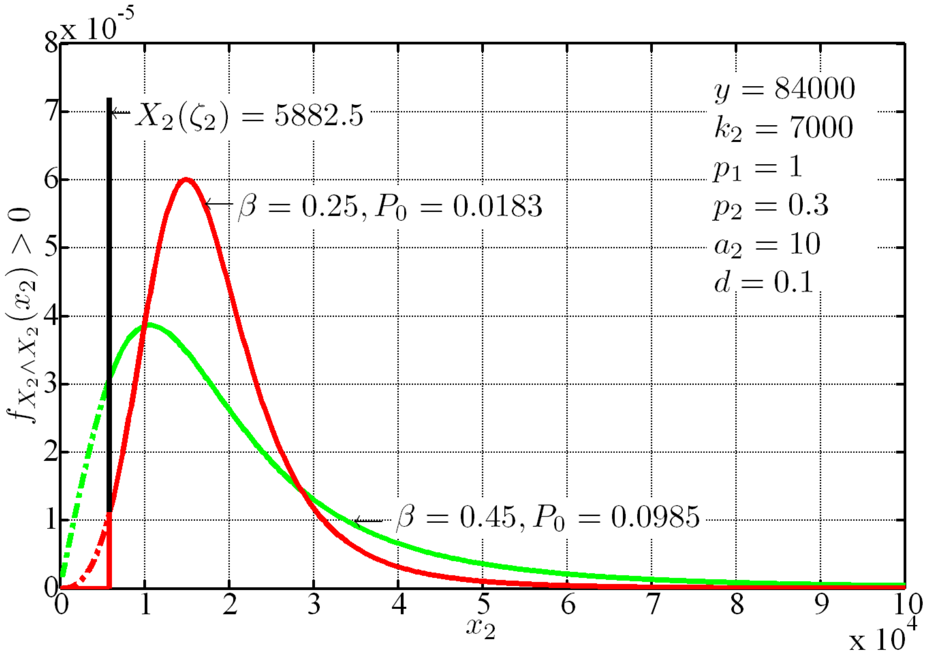


Diagram 1.3.10: pdf of X_2 for different „variance parameter" β

This diagram shows that a decrease in „variance parameters" β concentrates density function to a certain value and reduces density for high annual kilometres dramatically. At the same time, probability that households do not own a car is reduced. Again, one important feature of a change in β is that it does not affect the minimum distance a household does drive when deciding to own a car.

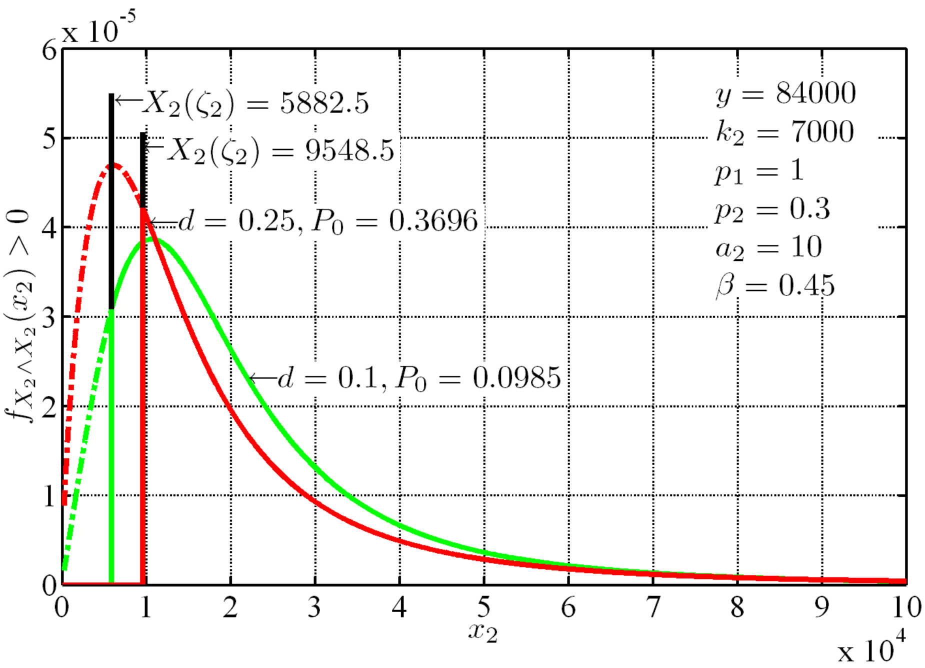


Diagram 1.3.11: pdf of X_2 for different „shape parameter" d

Diagram above shows that an decrease in „shape parameter" d decreases the minimum distance a household does drive when deciding to own a car. This is due to the fact that d determines the decrease of marginal utility: The lower d, the faster the utility of car driving decreases. On the other hand the first some kilometres yield higher utility than in the case where d is high. Therefore for lower values of d households rather try to spread consumption on both goods even if their income is decreased by the fixed costs arising when owning a car. That is what explains the decrease in the minimum distance a household does drive when deciding to own a car. Further, decreasing d strongly decreases the probability that households do not own a car.

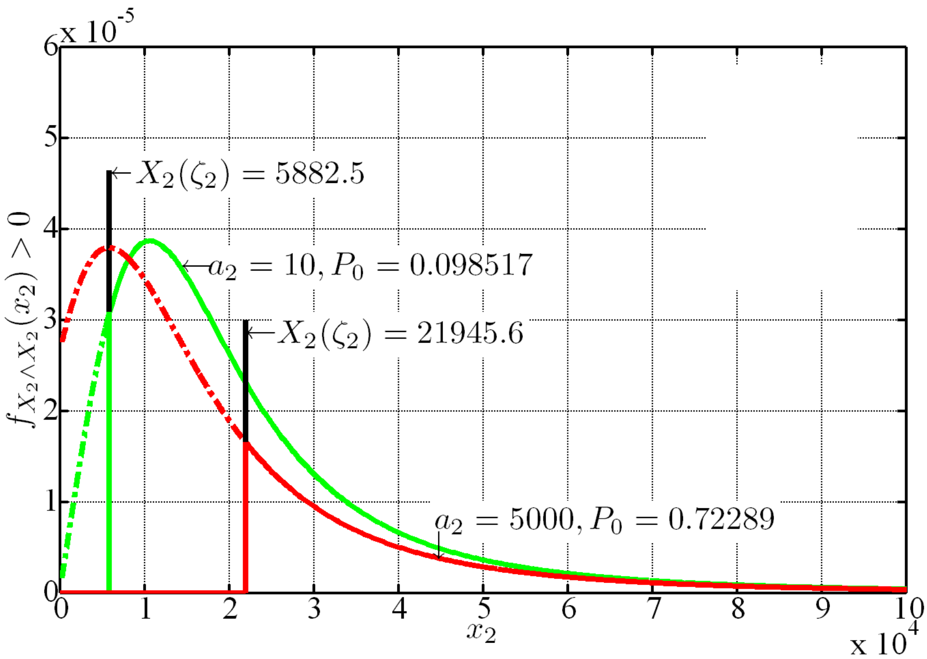


Diagram 1.3.12: pdf of X_2 for different „shift parameters" a_2

Diagram above shows that an increase in „shift parameter" a_2 only increases the minimum distance of a household does drive when deciding to own a car. This is due to the fact that a_2 pushes expression $(x_2 + a_2)^d$ to a range, where the first kilometre driven yields a lower marginal utility. Also here, shifting this minimum distance to a higher value increases the probability of not owning a car.

The problem of discontinuity

Diagram above shows that if $x_2 = 10000$ is observed, then if a_2 was $a_2 = 10$ this observation has a positive probability. When increasing a_2, at some point probability of observing $x_2 = 10000$ switches to zero. The problem is, that such a discontinuity is not feasible when computing MLE: As soon as one observations probability of one observation yields zero the probability for observing all data get zero and therefore any change of parameters do not change the probability of observing the dataset. The problem that one observation probability can get zero can arise, when varying *d*. Therefore MLE procedure must be modified.

Estimation procedure

As mentioned above, standard MLE procedure may not be applied due to the discontinuity problem. For solving this problem, the following estimation routine is applied:

1. Choose some value for a_2 and *d.*

2. Eliminate all positive observations that are in the interval $\left(0,\ X_2(\varsigma_2)\right)$.

3. Estimate β and γ, where $m = \gamma \cdot s$ with *s* being sociodemographic variables

4. Check if simulated $\hat{P}\left(X_2 = 0 \mid \hat{\theta}, \text{data}\right)$ and $\hat{E}\left(X_2 \mid \hat{\theta}, \text{data}\right)$ are the same like in data. $\hat{\theta}$ is the set of parameters set and estimated respectively as described above.

5. Evaluate the result by a valuation function *M*. This function is increasing in difference of simulated and empirical probability of X_2 being $X_2 = 0$ and simulated expectation value and empirical mean of distance driven and increasing in the number of eliminated observation in step 2.

6. Change a_2 and *d* until *M* has reached its minimum value.

Empirical results

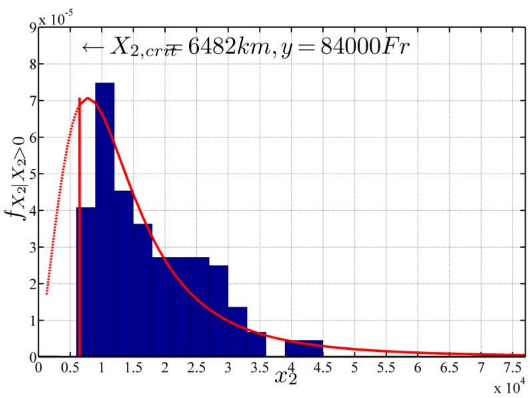
In the following survey data from Mikrozensus zum Verkehrsverhalten 2005, BfS (2005), was used. These data are cross section data of more than 30'000 Swiss households. Among other information, these households reported the total amount of kilometres they drove by their car or their cars, in case they had more than one. Theses distances were summed up for each household and is considered as driving distance x_2 in the model. According to assumptions of the model it was assumed that households have no cost, when switching from no car to one car an vice versa. This means that the loss of value of the car when selling it due to information symmetry between seller and buyer and transactions costs were neglected. In principle the economic environment of the households is considered as if households would rent their cars. Since there is only one car type captured in the model it is assumed that households can choose between no car and a standard car. This standard car has a fixed cost of 7000sFr per year. Variable costs are assumed to be 0.2sFr plus the fuel costs based on a fuel consumption of 8 l/100km. Both fixed cost and variable costs are based on calculations of TCS (2009) for a typical middle class car. Fuel price is based on the average fuel price of the last 12 month before the household was interviewed. Households owning more than one car where considered as if they had only one car. The dataset was cleaned from entries where households would spend more than one third of their income just for driving, drove more than 65'000 km a year and of households that stated they were driving less than 1000 km's per year. In all those cases it was assumed that households were giving wrong information on their driving distance or that they use the car professionally. After removing these dataset, some random sample entries of households owning no car were removed in order to keep the share of households owning no car remains the same. From this sample, again a random subsample was taken in order to keep computation time low. As household specific variables, only a dummy for household living in rural area was used. This variable turned out to be most influential in other models.

Estimation routine was performed as described in paragraph above. Resulting parameters where:

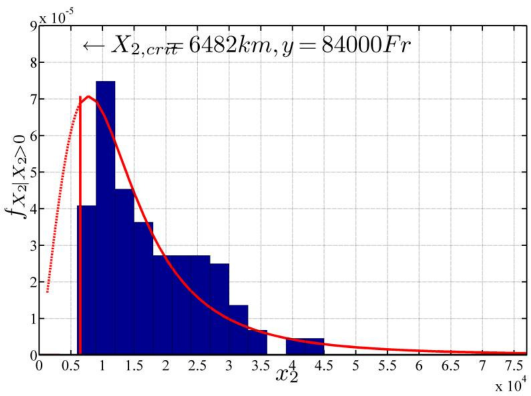
$\hat{\hat{d}} = 0.18951,\ \hat{\hat{a}}_1 = 0,\ \hat{\hat{a}}_2 = 0.81824$,

and the parameters determining the relative preference for driving cars *m* are

$\hat{\gamma}_1 = \underset{(0.0663)}{-2.2696},\ \hat{\gamma}_2 = \underset{(0.1389)}{0.34882},\ \hat{\beta} = \underset{(0.03)}{0.30662}$, with $m = \gamma_1 + \gamma_2 \cdot rural$ and $\hat{\beta}$ being a parameter determining the variance of the unobserved preference of the households as defined in (1.3.28). Since parameters γ_1, γ_2 and β are estimated by Maximum Likelihood estimation, standard errors could be computed. Standard errors of parameters estimated are in brackets. All values are highly significant. Parameter $\hat{\gamma}_2$ is greater than zero what implies that households living in rural areas have stronger preference for driving. This is a rather intuitive finding since distance to facilities is larger in average and offer of public transportation is less in such areas. Estimated parameters yield density functions for demand for driving X_2 that fits data quite well, as diagrams below show.



rural



urban

Diagram 11: Observed data of X_2 and pdf for income 84000Fr: Rural and urban areas

Simulation results

Given these estimated parameters it is now interesting to compute changes of household behaviour if variables like prices and fixed costs change. In the context of determining fuel demand and car holding, it is interesting to compute changes in expected demand of driving and probability of deciding not to own a car:

$$X_{2,sim} = E_{\varsigma}\left(X_2\left(p_1, p_2, y, k_2\right) \mid \hat{a}_1, \hat{a}_2, \hat{\hat{d}}, \hat{m}, \varsigma \right), \tag{1.3.29}$$

$$P\left(X_2 = 0\right)_{sim} = P_{X_2}\left(X_2\left(p_1, p_2, y, k_2\right) \mid \hat{a}_1, \hat{a}_2, \hat{\hat{d}}, \hat{m}, \varsigma = 0 \right). \tag{1.3.30}$$

Summing up the change in unconditional expected value of demand in car driving for each household will yield the expected total change of driving of the whole population.

Policy effects

In this paragraph the results of some policy changes are presented. The first policy of interest is an increase in costs per kilometres, p_2:

A change in costs per kilometres, p_2 by one percent yields a change in total kilometres driven of

1.32%. Since fuel price contribute only about one third of the variable costs of the car[13], and increase of fuel prices would decrease total kilometers driven only by 0.44%. Therefore fuel price elasticity is about 0.44. This value is in the range as found in other international studies.[14] An interesting result is also that in this case the share of households that do not own a car only decreases by 0.119% from 21.295% to 21.176%. This is a very small share and implies that the reduction in aggregate driving distance and therefore in aggregate fuel demand mainly is contributed by households still using a car but using that car less. Moreover this effect is very small since model implies that only households that already drove low annual distances would sell their car.

Another interesting policy is to increase taxes on car ownership. A tax that would increase fixed cost of cars by one percent, would presumably decrease the share of households owning a car and reduce their budget available. A one percent increase in fixed cost would increase the share of households not owning a car from 21.295% to 21.65% and total distance driven would decrease by fuel consumption would decrease by 0.266%. Later figure is hard to interpret, since no elasticity can derived from this figure. One possibility to compare the effect of this tax to an increase in fuel taxes is to calculate the total tax revenue. The effect on fuel demand can then be related to the tax revenue. The tax having more effect per tax revenue, may then be the "better" tax.

[13]The average fuel price was 1.46, share of fuel cost was only about one third: $1.26 \cdot 0.08 / \left(1.26 \cdot 0.08 + 0.2\right) = 0.335$.

[14]See Dahl et al. (1991), Graham et al. (2002) or Schleiniger (1996).

Comparing these taxes yields:

	1% tax on fuel	1% tax on fix cost
abs. effect one average driving distance	**67.67km**	**39.03km**
abs. effect one car ownership driving distance	**3.99%**	**35.30%**
rel. effect one average driving distance	**0.46%**	**0.27%**
Average tax revenue per household	**18.3**	**55.1**
abs. effect one average driving distance per 1sFr tax revenue	**3.69km /sFr**	**0.71km /sFr**
rel. effect one average driving distance per 1sFr tax revenue	**0.025% /sFr**	**0.005% /sFr**
abs. effect on share of car ownership per 1sFr tax revenue	**0.22% /sFr**	**0.64% /sFr**

Table 1.3.1: Effects of a tax on fuel to a tax on car ownership and driving distance

This table shows that per tax revenue a tax on fuel is about five times more effective than a tax on car ownership with respect to a reduction in annual kilometres driven. On the other hand, a tax on car ownership is three times more effective when the aim is to reduce the share of household owning a car.

Another interesting information is how much people would drive more, if they moved from an urban area to a rural area and vice versa. The model predicts the following changes:

present location	urban	rural
share of household on total population	79%	21%
average annual km's before move	13420	19262
average annual km's after move	22181	11784
absolute change in average annual km's	8761	-7478
share of no-car households before move	24%	12%
share of no-car households after move	6%	26%
absolute change in share of no-car households	18%	-14%

Table 1.3.2: Effects of household location on car ownership and driving distance

The model predicts a huge change in car ownership and car use when household change from urban to rural areas and vice versa as is shown in table above. An interesting detail is, that urban households would drive more additional kilometers if they move to a rural than rural households would drive less.

The reason for this might be that average income of urban households is higher than that of rural households, namely 81'759sFr compared to 75'900sFr.

With respect to policies for reducing fuel demand, results above show that not only the type and height of taxes on fuel and cars may play an important role, but also spatial planning.

Literature

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), pp. 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), pp. 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), pp. 274–303.

Dahl, Carol und Thomas Sterner, 1991, „Analyzing gasoline demand elasticities: a survey”, Energy Economics, July 1991, pp. 203-210.

Dubin, Jeffrey A. und Daniel L. McFadden, 1984, „An Econometric Analysis of Residential Electric Appliance Holdings and Consumption“, Econometrica, Vol. 52, No. 2 (Mar., 1984), pp. 345-362.

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.

Graham, Daniel und 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.

Graham, Daniel und Stephen Glaister, 2005, „Decomposing the determinants of road traffic demand“, Applied Economics 37(1): 19-28.

Mas-Colell Andreu, Michael D. Whinston and Jerry R. Green, 1995, ”Microeconomic Theory”, Oxford Univ. Press, 1995.

Shao, Jun, 2003, „Mathematical statistics”, Springer, New York.

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.

Varian, Hal R., 1992, „Microeconomic Analysis“, 3rd edition, Norton, 1992.

Appendix

A 1 Gumbel distribution

The Gumbel distribution as used in the context of this paper is defined as follows:

$$X \sim gu(0,1),\ f_{\xi}(x) = e^{-x} \cdot \exp\left(-e^{-x}\right). \qquad \text{(A1.1)}$$

The shape of the probability density function is as follows:

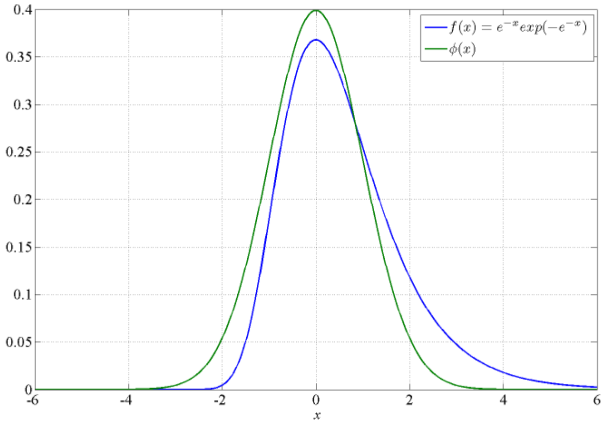


Diagram A.1.1: Probability functions of the standard Gumbel and the standard normal distribution

The shape of the cumulated density function is as follows:

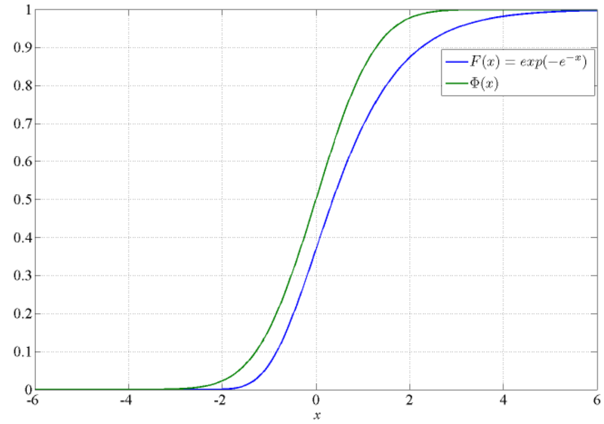


Diagram A.1.2: Cumulated density functions of the standard Gumbel and the standard normal distribution

The diagram show, 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 an the mean is non zero, namely equal the Euler Mascherioni constant $E[X] = \lambda = 0.577..$. The variance is lower, namely $\text{var}[X] = \pi^2/6 = 0.523..$. In contrast to the standard normal distribution has some very properties. Here the most important of them are listed:

1. $median[X] = -\ln(\ln(2)) = 0.3665..$.[1]

2. $mode[X] = 0$.[2]

3. If X_1 and X_2 are iid standard Gumbel and are linear transformations with the same shifting parameter, namely $Y_1 = \alpha_1 + \beta X_1$ and $Y_2 = \alpha_2 + \beta X_2$, then $Z = Y_1 - Y_2 = \alpha_1 - \alpha_2 + \beta(X_1 - X_2)$ is distributed as:[3]

$$Z \sim F(z) = \frac{1}{1 + e^{-\frac{1}{\beta}(z-\alpha_1+\alpha_2)}}.$$

---

[1] $F_Z(z) = \exp(-e^{-x}) = 0.5 \Leftrightarrow -x = \log(-\log(0.5)) \Leftrightarrow x = -\log(\log(2))$.

[2] $\frac{\partial f_Z(z)}{\partial z} = -e^{-x}\exp(-e^{-x}) + e^{-x}e^{-x}\exp(-e^{-x}) = 0 \Leftrightarrow e^{-x}\exp(-e^{-x}) = e^{-x}e^{-x}\exp(-e^{-x}) \Leftrightarrow e^{-x} = 1 \Leftrightarrow x = 0.$

[3] Proof: First the cumulated density function of Z conditional on x2 has to be calculated:

$$F_{Z|X_2}(z) = F_{Z|X_2}\left(\frac{z-\alpha_1+\alpha_2}{\beta} + X_2\right).$$

Note, that from $Z = Y_1 - Y_2 = \alpha_1 + \beta X_1 - \alpha_2 - \beta X_2$ it follows $X_1 = \frac{Z-\alpha_1+\alpha_2}{\beta} + X_2$.

Using $F_Z(z) = E_{X_2}\left[F_{Z|X_2}(z)\right]$ it follows that:

$$F_Z(z) = \int_{x_2=-\infty}^{x_2=\infty} e^{-x_2}\cdot\exp(-e^{-x_2})\cdot\exp\left(-e^{-\left(\frac{z-\alpha_1+\alpha_2}{\beta}+x_2\right)}\right)dz = \int_{x_2=-\infty}^{x_2=\infty} e^{-x_2}\cdot\exp\left(-e^{-x_2} - e^{-\left(\frac{z-\alpha_1+\alpha_2}{\beta}+x_2\right)}\right)dx_2 =$$

$$= \int_{x_2=-\infty}^{x_2=\infty} e^{-x_2}\cdot\exp\left(-e^{-x_2}\left(1+e^{-\left(\frac{z-\alpha_1+\alpha_2}{\beta}+x_2\right)}\right)\right)dx_2 = \int_{x_2=-\infty}^{x_2=\infty} e^{-x_2}\cdot\exp\left(-e^{-x_2+\ln\left(1+e^{-\left(\frac{z-\alpha_1+\alpha_2}{\beta}+x_2\right)}\right)}\right)dx_2 =$$

$$= e^{-\ln\left(1+e^{-\left(\frac{z-\alpha_1+\alpha_2}{\beta}\right)}\right)}\cdot\int_{x_2=-\infty}^{x_2=\infty} e^{-x_2+\ln\left(1+e^{-\left(\frac{z-\alpha_1+\alpha_2}{\beta}+x_2\right)}\right)}\cdot\exp\left(-e^{-x_2+\ln\left(1+e^{-\left(\frac{z-\alpha_1+\alpha_2}{\beta}+x_2\right)}\right)}\right)dx_2 = e^{-\ln\left(1+e^{-\left(\frac{z-\alpha_1+\alpha_2}{\beta}+x_2\right)}\right)} =$$

$$= \frac{1}{1+e^{-\left(\frac{z-\alpha_1+\alpha_2}{\beta}\right)}}.$$

The corresponding density function is:

$$Z \sim f(z) = \frac{1}{\beta} \cdot \frac{e^{-\frac{1}{\beta}(z-\alpha_1+\alpha_2)}}{\left(1+e^{-\frac{1}{\beta}(z-\alpha_1+\alpha_2)}\right)^2}.$$

4. If X_1 and X_2 are iid standard Gumbel and are linear transformations, namely $Y_1 = \alpha_1 + \beta X_1$ and $Y_2 = \alpha_2 + \beta X_2$, then $Z = \max(Y_1, Y_2)$ is distributed as:[4]

$$Z \sim F(z) = \exp\left(-e^{-\left(\left(\frac{z}{\beta}\right)-\ln\left(e^{\frac{\alpha_1}{\beta}}+e^{\frac{\alpha_2}{\beta}}\right)\right)}\right).$$

The corresponding density function is:

$$Z \sim f(z) = \frac{1}{\beta} \cdot e^{-\left(\left(\frac{z}{\beta}\right)-\ln\left(e^{\frac{\alpha_1}{\beta}}+e^{\frac{\alpha_2}{\beta}}\right)\right)} \cdot \exp\left(-e^{-\left(\left(\frac{z}{\beta}\right)-\ln\left(e^{\frac{\alpha_1}{\beta}}+e^{\frac{\alpha_2}{\beta}}\right)\right)}\right).$$

This means that Z is distributed as if

$Z = \beta X + \ln\left(e^{\frac{\alpha_1}{\beta}} + e^{\frac{\alpha_2}{\beta}}\right)$, where X is standard Gumbel.[5]

5. Applying property four to $Z = \max(Y_1, Y_2, .., Y_N)$, where $Y_i = \alpha_i + \beta X_i$ and X_i is iid standard Gumbel yields:[6]

---

[4]Proof:

$$F_Z(z) = P(X_1 \le z \wedge X_2 \le z) = \int_{x_1=-\infty}^{x_1=z} \int_{x_2=-\infty}^{x_2=z} f_{x_1,x_2}(x_1, x_2)\,dx_2 dx_1 = \int_{x_1=-\infty}^{x_1=z} \int_{x_2=-\infty}^{x_2=z} f_{X_1}(x_1) f_{X_2}(x_1)\,dx_2 dx_1 = \int_{x_1=-\infty}^{x_1=z} f_{X_1}(x_1)\,dx_1 \cdot \int_{x_2=-\infty}^{x_2=z} f_{X_2}(x_2)\,dx_2 =$$

$$= F_{X_1}(z) \cdot F_{X_2}(z) = \exp\left(-e^{-\left(\frac{z-\alpha_1}{\beta}\right)} - e^{-\left(\frac{z-\alpha_2}{\beta}\right)}\right) = \exp\left(-e^{-\left(\frac{z}{\beta}\right)}\left(e^{\frac{\alpha_1}{\beta}}+e^{\frac{\alpha_2}{\beta}}\right)\right) = \exp\left(-e^{-\left(\frac{z}{\beta}\right)+\ln\left(e^{\frac{\alpha_1}{\beta}}+e^{\frac{\alpha_2}{\beta}}\right)}\right) = \exp\left(-e^{-\left(\left(\frac{z}{\beta}\right)-\ln\left(e^{\frac{\alpha_1}{\beta}}+e^{\frac{\alpha_2}{\beta}}\right)\right)}\right).$$

[5]Note that $X = \left(\frac{Z}{\beta}\right) - \ln\left(e^{\frac{\alpha_1}{\beta}} + e^{\frac{\alpha_2}{\beta}}\right) \Leftrightarrow X + \ln\left(e^{\frac{\alpha_1}{\beta}} + e^{\frac{\alpha_2}{\beta}}\right) = \left(\frac{Z}{\beta}\right) \Leftrightarrow Z = \beta X + \ln\left(e^{\frac{\alpha_1}{\beta}} + e^{\frac{\alpha_2}{\beta}}\right)$.

[6]The proof for this property is straight forward:

$Z \sim F(z) = \dfrac{1}{1+e^{-\frac{1}{\beta}(z-\alpha_1+\alpha_2)}}$. This means that Z is distributed as $Z = \beta X + \ln\left(e^{\frac{\alpha_1}{\beta}} + e^{\frac{\alpha_2}{\beta}}\right)$, where X is standard Gumbel.[7]

6. $E_X\left(F_\xi\left(\dfrac{X+a}{\sigma}\right)\right) = \dfrac{1}{1+e^{-\frac{a}{\sigma}}} = \dfrac{e^{\frac{a}{\sigma}}}{1+e^{\frac{a}{\sigma}}}$, where X is standard Gumbel $F_\xi(\ )$ is the CDF of standard Gumbel.[8]

7. $E_X\left(F_\xi\left(\dfrac{X+a}{\sigma}\right)\cdot F_\xi\left(\dfrac{X+b}{\sigma}\right)\right) = \dfrac{1}{1+e^{-\frac{a}{\sigma}}+e^{-\frac{b}{\sigma}}}$, where X is standard Gumbel and $F_\xi(\ )$ is the CDF of standard Gumbel.[9]

8. $E_X\left(f_{a|X}(X,a)\right) = \dfrac{\partial E_X\left(F_{a|X}(X,a)\right)}{\partial a} = f_a(a)$.

   Applying this rule to the case where $E_X\left(f_{a|X}(X,a)\right) = E_X\left(\dfrac{1}{\sigma}\cdot f_\xi\left(\dfrac{X+a}{\sigma}\right)\right)$, $f_\xi(\bullet)$ being the density of a standard Gumbel distributed random variable, yields:

$$E_X\left(\frac{1}{\sigma}\cdot f_\xi\left(\frac{X+a}{\sigma}\right)\right) = \frac{\partial E_X\left(F_\xi\left(\frac{X+a}{\sigma}\right)\right)}{\partial a} = \frac{\partial \frac{1}{1+e^{-\frac{a}{\sigma}}}}{\partial a} = \frac{1}{\sigma}\cdot\left(\frac{1}{1+e^{-\frac{a}{\sigma}}}\right)^2\cdot e^{-\frac{a}{\sigma}} = \frac{1}{\sigma}\cdot\frac{e^{\frac{a}{\sigma}}}{\left(1+e^{\frac{a}{\sigma}}\right)^2}.$$

$$F_Z(z) = P\left(X_1 \le z \wedge X_2 \le z \wedge .. \wedge X_N \le z\right) = \int_{x_1=-\infty}^{x_1=z}\int_{x_2=-\infty}^{x_2=z}\cdots\int_{x_N=-\infty}^{x_N=z}\cdots f_{X_1,X_2,..,X_N}(x_1, x_2,..,x_N)\,dx_N\cdots dx_2 dx_1 =$$

$$= \int_{x_1=-\infty}^{x_1=z} f_{X_1}(x_1)\,dx_1\cdot\int_{x_2=-\infty}^{x_2=z} f_{X_2}(x_2)dx_2\cdots\int_{x_N=-\infty}^{x_N=z} f_{X_2}(x_N)dx_N = F_{X_1}(z)\cdot F_{X_2}(z)\cdots F_{X_N}(z) = \exp\left(-\sum_{i=1}^{N} e^{-\left(\frac{z-\alpha_i}{\beta}\right)}\right) = \exp\left(-e^{-\left(\left(\frac{z}{\beta}\right)-\ln\left(\sum_{i=1}^{N} e^{\frac{\alpha_i}{\beta}}\right)\right)}\right).$$

[7]Note that $X = \left(\dfrac{Z}{\beta}\right) - \ln\left(\sum_{i=1}^{N} e^{\frac{\alpha_i}{\beta}}\right) \Leftrightarrow X + \ln\left(\sum_{i=1}^{N} e^{\frac{\alpha_i}{\beta}}\right) = \left(\dfrac{Z}{\beta}\right) \Leftrightarrow Z = \beta X + \ln\left(\sum_{i=1}^{N} e^{\frac{\alpha_i}{\beta}}\right)$.

[8]Proof: $E_X\left(F_\xi\left(\dfrac{X+a}{\sigma}\right)\right) = \int_{z=-\infty}^{z=\infty} \exp\left(e^{-\frac{z+a}{\sigma}}\right)\cdot\dfrac{1}{\sigma}\cdot e^{-\frac{z}{\sigma}}\cdot\dfrac{1}{\sigma}\exp\left(e^{-\frac{z}{\sigma}}\right)dz = \int_{z=-\infty}^{z=\infty}\dfrac{1}{\sigma}\cdot\exp\left(e^{-\frac{z+a}{\sigma}} + e^{-\frac{z}{\sigma}}\right)\cdot e^{-\frac{z}{\sigma}}dz =$

$$= \int_{z=-\infty}^{z=\infty}\frac{1}{\sigma}\cdot\exp\left(e^{-\frac{z}{\sigma}}\cdot\left(1+e^{-\frac{a}{\sigma}}\right)\right)\cdot e^{-\frac{z}{\sigma}}dz = \frac{1}{\sigma}\cdot\int_{z=-\infty}^{z=\infty}\exp\left(e^{-\frac{z}{\sigma}+\ln\left(1+e^{-\frac{a}{\sigma}}\right)}\right)\cdot e^{-\frac{z}{\sigma}}dz =$$

$$= \frac{1}{\sigma}\cdot e^{-\ln\left(1+e^{-\frac{a}{\sigma}}\right)}\int_{z=-\infty}^{z=\infty}\exp\left(e^{-\frac{z}{\sigma}+\ln\left(1+e^{-\frac{a}{\sigma}}\right)}\right)\cdot e^{-\frac{z}{\sigma}+\ln\left(1+e^{-\frac{a}{\sigma}}\right)}dz = e^{-\ln\left(1+e^{-\frac{a}{\sigma}}\right)}\cdot\int_{z=-\infty}^{z=\infty}\frac{1}{\sigma}\cdot f_\xi\left(\frac{z}{\sigma}-\ln\left(1+e^{-\frac{a}{\sigma}}\right)\right)dz =$$

$$= e^{-\ln\left(1+e^{-\frac{a}{\sigma}}\right)} = \frac{1}{1+e^{-\frac{a}{\sigma}}} = \frac{e^{\frac{a}{\sigma}}}{1+e^{\frac{a}{\sigma}}}.$$

where X is standard Gumbel, $f_\xi(\ )$ is the PDF and $F_\xi(\ )$ is the CDF of a standard Gumbel distributed random variable. Note that here: $f_{a|X}(X,a) = f_\xi(X+a)$.

9. $$E_X\left(\prod_{i=1}^{n} f_\xi(X+a_i)\right) = n! \cdot \frac{e^{-\sum_{i=1}^{n} a_i}}{\left(1+\sum_{i=1}^{n} e^{-a_i}\right)^{n+1}},$$

where X is standard Gumbel and $F_\xi(\ )$ is the CDF of standard Gumbel.

10. Theorem: (Densities of transformed random variables)[10]

Shall $X = (X_1, ..., X_k)$ a random vector with density f_X, and shall $Y_i = h_i(X_1, ..., X_k)$, 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

$$f_Y(y) = f_X(l_1(y), ..., l_k(y)) \cdot |J(y)|,$$

---

[9]Proof: $$E_X\left(F_\xi\left(\frac{X+a}{\sigma}\right)\cdot F_\xi\left(\frac{X+b}{\sigma}\right)\right) = E_X\left(F_\xi\left(\frac{X+a}{\sigma}\right)\cdot F_\xi\left(\frac{X+b}{\sigma}\right)\right) = E_X\left(\exp\left(e^{-\left(\frac{X+a}{\sigma}\right)}\right)\cdot\exp\left(e^{-\left(\frac{X+b}{\sigma}\right)}\right)\right) =$$

$$= E_X\left(F_\xi\left(\frac{X+a}{\sigma}\right)\cdot F_\xi\left(\frac{X+b}{\sigma}\right)\right) = E_X\left(F_\xi\left(\frac{X+a}{\sigma}\right)\cdot F_\xi\left(\frac{X+b}{\sigma}\right)\right) = E_X\left(\exp\left(e^{-\left(\frac{X+a}{\sigma}\right)}\right)\cdot\exp\left(e^{-\left(\frac{X+b}{\sigma}\right)}\right)\right) =$$

$$= E_X\left(\exp\left(e^{-\left(\frac{X+a}{\sigma}\right)}\right)\cdot\exp\left(e^{-\left(\frac{X+b}{\sigma}\right)}\right)\right) = E_X\left(\exp\left(e^{-\left(\frac{X+a}{\sigma}\right)} + e^{-\left(\frac{X+b}{\sigma}\right)}\right)\right) = E_X\left(\exp\left(e^{-\frac{X}{\sigma}}\cdot\left(e^{-\frac{a}{\sigma}} + e^{-\frac{b}{\sigma}}\right)\right)\right) =$$

$$= E_X\left(\exp\left(e^{-\left(\frac{X}{\sigma}-\ln\left(e^{-\frac{a}{\sigma}}+e^{-\frac{b}{\sigma}}\right)\right)}\right)\right) = E_X\left(F_\xi\left(\frac{X-\ln\left(e^{-\frac{a}{\sigma}}+e^{-\frac{b}{\sigma}}\right)}{\sigma}\right)\right) = \frac{1}{1+e^{\ln\left(e^{-\frac{a}{\sigma}}+e^{-\frac{b}{\sigma}}\right)}} = \frac{1}{1+e^{-\frac{a}{\sigma}}+e^{-\frac{b}{\sigma}}}.$$

[10]See Shao (2003), page 23.

where $\left|J\left(y\right)\right| = \det\begin{bmatrix} \frac{\partial x_1}{\partial y_1} & \cdots & \frac{\partial x_1}{\partial y_k} \\ \cdots & \cdots & \cdots \\ \frac{\partial x_k}{\partial y_1} & \cdots & \frac{\partial x_k}{\partial y_k} \end{bmatrix}$.

A 2 The maximum utility calculus: An alternative illustration

In the following the maximum utility calculus is again illustrated. In contrast to illustration in section 1.2 the illustration is based on the first order conditions (1.2.7) and (1.2.8). For the illustration the first order conditions are reformulated first:

$$\frac{d_1}{p_1} \cdot \exp(m_1 + \xi_1) \cdot \frac{1}{(x_1 + a_1)^{1-d_1}} = \lambda, \qquad \text{(A.2.1)}$$

$$\frac{d_2}{p_2} \cdot \exp(m_2 + \xi_2) \cdot \frac{1}{(x_2 + a_2)^{1-d_2}} \leq \lambda. \qquad \text{(A.2.2)}$$

The reformulated first order condition (A.1.1) can be interpreted as "the marginal utility of spending an additional unit of the income for good one has to be equal to λ for an optimal choice of x_1 ". The same holds for x_2 when x_2 is positive. When x_2 is bound to zero, the marginal utility of spending an additional unit of the income for good two will be smaller than λ. Using (A 1.1) and (A 1.2) the choice can be illustrated as follows:

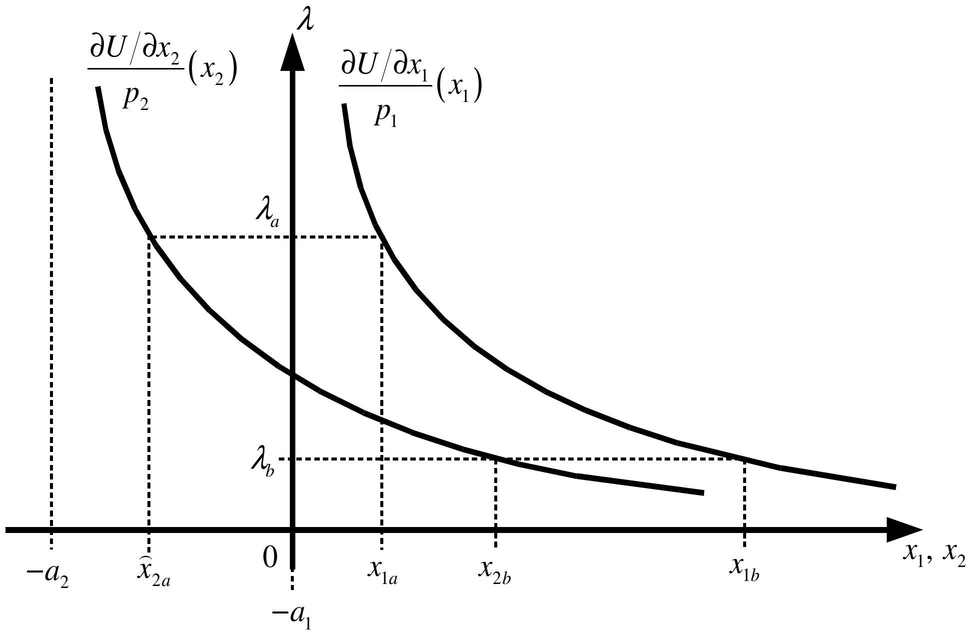


Diagram A 2.1: The maximization calculus

Again this illustration shows, that the parameters a_1 and a_2 define the minimum consumption level of x_1 and x_2, since for $x_1 = -a_1$ and $x_2 = -a_2$ marginal utilities of spending one additional unit of income goes to infinity. This illustration also shows that expenditures increase when λ decreases.

A 3 Probability of observing good one to be zero

This section relates to the two good case without fixed cost as describer in section 1.2. In the following it is proven, why probability $P(X_1 = 0) = P\left(X_2 = \frac{y}{p_2}\right)$ is zero. It is important that despite $f_{X_2|X_2>0}(z)$ goes to infinity for $z = y/p_2$ the discrete probability for $X_2 = y/p_2$, $X_1 = 0$ is zero. This would not be compatible with the fact that $x_1 = -a_1 = 0$ is a lower limit of the possible range of realisations of X_1. For prove $X_2 = y/p_2$, $X_1 = 0$ being zero, first the functional form of $f_{X_2|X_2>0}(z)$ in the limit $z = y/p_2$ has to be calculated:

$$e^{-V_1(X_2)} = e^{-\left(\ln(d_1)-\ln(p_1)+m_1-(1-d_1)\cdot\ln\left(\frac{y-p_2\cdot X_2}{p_1}+a_1\right)\right)} = e^{-\ln(d_1)+\ln(p_1)-m_1} \cdot e^{\left(\ln\left(\frac{y-p_2\cdot X_2}{p_1}+a_1\right)^{1-d_1}\right)} =$$

$$= e^{-\ln(d_1)+\ln(p_1)-m_1} \cdot \left(\frac{y-p_2\cdot X_2}{p_1}+a_1\right)^{1-d_1} \Rightarrow$$

$$\Rightarrow \lim_{z\to y/p_2} e^{-V_1(X_2)} = 0 \Rightarrow \lim_{z\to y/p_2} \left(1+e^{V_2(z)-V_1(z)}\right)^2 = 0$$

$$\Rightarrow \lim_{z\to y/p_2} f_{X_2|X_2>0}(z) = \lim_{z\to y/p_2} e^{V_2(z)} \left(\frac{1-d_1}{\frac{y-p_2 z}{p_1}+a_1}\cdot\frac{p_2}{p_1}+\frac{1-d_2}{z+a_2}\right)\cdot e^{-V_1(z)}\cdot\left(\frac{e^{V_2(0)}}{e^{V_1(0)}+e^{V_2(0)}}\right)^{-1} =$$

$$= \lim_{z\to y/p_2} e^{V_2(z)} \left(\frac{1-d_1}{\frac{y-p_2 z}{p_1}+a_1}\cdot\frac{p_2}{p_1}+\frac{1-d_2}{z+a_2}\right)\cdot e^{-\ln(d_1)+\ln(p_1)-m_1}\cdot\left(\frac{y-p_2\cdot X_2}{p_1}+a_1\right)^{1-d_1}\cdot\left(\frac{e^{V_2(0)}}{e^{V_1(0)}+e^{V_2(0)}}\right)^{-1} =$$

$$= \lim_{z\to y/p_2} e^{V_2(z)}\cdot e^{-\ln(d_1)+\ln(p_1)-m_1}\cdot\left(\frac{e^{V_2(0)}}{e^{V_1(0)}+e^{V_2(0)}}\right)^{-1}\cdot\left(\frac{1-d_1}{\left(\frac{y-p_2\cdot X_2}{p_1}+a_1\right)^{d_1}}\cdot\frac{p_2}{p_1}+\frac{1-d_2}{z+a_2}\cdot\left(\frac{y-p_2\cdot X_2}{p_1}+a_1\right)^{1-d_1}\right) =$$

$$= \lim_{z\to y/p_2} e^{V_2(z)}\cdot e^{-\ln(d_1)+\ln(p_1)-m_1}\cdot\left(\frac{e^{V_2(0)}}{e^{V_1(0)}+e^{V_2(0)}}\right)^{-1}\cdot\left(\frac{1-d_1}{\left(\frac{y-p_2\cdot X_2}{p_1}+a_1\right)^{d_1}}\cdot\frac{p_2}{p_1}\right) = \infty$$

Since $\lim_{z\to y/p_2} f_{X_2|X_2>0}(z) = \infty$ there is some doubt that $X_2 = y/p_2$, $X_1 = 0$ could be greater than zero. But this has not necessarily have to be: The "area" below the $f_{X_2|X_2>0}(z)$ can still be zero, when the length of the interval considered around $z = y/p_2$ goes to zero. This means that the following integral has to be calculated:

$\lim_{z\to\varepsilon}\int_{z=y/p_2-\varepsilon}^{z=y/p_2} f_{X_2|X_2>0}(z)dz$. For checking if this integral yields a finite value it makes the following transformation is feasible:

$\lim_{z\to\varepsilon}\int_{z=y/p_2-\varepsilon}^{z=y/p_2} f_{X_2|X_2>0}(z)dz = \lim_{z\to\varepsilon}\varepsilon\cdot f_{X_2|X_2>0}(z=y/p_2-\varepsilon)$. (Check that !!!) Plugging in the functional form of $f_{X_2|X_2>0}(z)$ in the limit $z=y/p_2$, this formula yields:

$$\lim_{z\to\varepsilon}\int_{z=y/p_2-\varepsilon}^{z=y/p_2} f_{X_2|X_2>0}(z)dz = \lim_{z\to\varepsilon}\varepsilon\cdot f_{X_2|X_2>0}(z=y/p_2-\varepsilon)=$$

$$=\varepsilon\cdot e^{V_2(y/p_2-\varepsilon)}\cdot e^{-\ln(d_1)+\ln(p_1)-m_1}\cdot\left(\frac{e^{V_2(0)}}{e^{V_1(0)}+e^{V_2(0)}}\right)^{-1}\cdot\left(\frac{1-d_1}{\left(\frac{y-p_2\cdot(y/p_2-\varepsilon)}{p_1}+a_1\right)^{d_1}}\cdot\frac{p_2}{p_1}\right)\overset{a_1=0}{=}$$

$$=\varepsilon\cdot e^{V_2(y/p_2-\varepsilon)}\cdot e^{-\ln(d_1)+\ln(p_1)-m_1}\cdot\left(\frac{e^{V_2(0)}}{e^{V_1(0)}+e^{V_2(0)}}\right)^{-1}\cdot\left(\frac{1-d_1}{\left(\frac{\varepsilon}{p_1}\right)^{d_1}}\cdot\frac{p_2}{p_1}\right)=$$

$$=\varepsilon^{1-d_1}\cdot e^{V_2(y/p_2-\varepsilon)}\cdot e^{-\ln(d_1)+\ln(p_1)-m_1}\cdot\left(\frac{e^{V_2(0)}}{e^{V_1(0)}+e^{V_2(0)}}\right)^{-1}\cdot\left((1-d_1)\cdot\frac{p_2}{p_1^{1-d_1}}\right)=0.$$

Therefore $P(X_2=y/p_2, X_1=0)=0$ is proven.

A 4 Minimum consumption threshold of good two

In this section the sign of the influence of y, k_2, p_1, p_2 and parameters d, a_1, a_2 on the minimum consumption threshold of good two $X_2(\varsigma_2)$ shall be examined. All proves are based on applying the implicit function theorem on $g_2(y,k_2,p_1,p_2,d,a_1,a_2)=0$.

The first prove that the influence of k_2 on $X_2(\varsigma_2)$ is positive. This is done by

$$\frac{\partial X_2(\varsigma_2)}{\partial k_2} = -\frac{\partial g_2(y,k_2,p_1,p_2,d,a_1,a_2,\varsigma)/\partial k_2}{\partial g_2(y,k_2,p_1,p_2,d,a_1,a_2,\varsigma)/\partial \varsigma} \cdot \frac{\partial X_2}{\partial \varsigma} - \frac{\partial X_2}{\partial k_2}. \tag{A 4.1}$$

First, $\partial g_2(y,k_2,p_1,p_2,d,a_1,a_2,\varsigma)/\partial \varsigma$ is calculated:

$$\frac{\partial g_2(y,k_2,p_1,p_2,d,a_1,a_2,\varsigma)}{\partial \varsigma} = \frac{d\breve{g}_2(B)}{dB} \cdot \frac{\partial B}{\partial \varsigma} > 0,$$

$$\frac{d\breve{g}_2(B)}{dB} = (1-d) \cdot \frac{p_2}{p_1} \cdot B^{-d} \cdot \left(Q_2(B)^d - a_2{}^d\right) > 0 \text{ for } Q_2(B(\varsigma_2)), \text{ since } Q_2(B(\varsigma_2)) > a_2.$$

$$\frac{\partial B(\varsigma)}{\partial \varsigma} = \left(\frac{p_1}{p_2}\right)^{\frac{1}{1-d}} \cdot \exp\left(\frac{m+\varsigma}{1-d}\right) \cdot \frac{1}{1-d} = \frac{1}{1-d} \cdot \frac{p_1}{p_2} \cdot B^d > 0.$$

Therefore, in the relevant range, namely for $X_2 > X_2(\varsigma_1)$ the sign of $\partial g_2(y,k_2,p_1,p_2,d,a_1,a_2,\varsigma)/\partial \varsigma$ is positive. Now, the sign of $\partial g_2(y,k_2,p_1,p_2,d,a_1,a_2,\varsigma)/\partial k_2$ has to be determined. Again derivation by applying chain rule is used:

$$\frac{\partial g_2}{\partial k_2} = \frac{\partial Q_1{}^d + \frac{p_2}{p_1} \cdot B^{1-d} \cdot Q_2{}^d}{\partial k_2} = \frac{\partial Q_1{}^d + \frac{p_2}{p_1} \cdot B \cdot Q_1{}^d}{\partial k_2} = \left(1 + \frac{p_2}{p_1} \cdot B\right) \cdot d \cdot Q_1{}^{d-1} \cdot \frac{\partial Q_1}{\partial k_2} < 0, \tag{A 4.2}$$

since $\dfrac{\partial Q_1}{\partial k_2} = -\dfrac{k_2}{p_1 + p_2 \cdot B} < 0.$

Further[11]

$$\frac{\partial X_2}{\partial \varsigma} = Q_2(B) \cdot \frac{1}{p_1 B^{-1} + p_2} \cdot \frac{p_1}{B^2} \cdot \frac{1}{1-d} \cdot \frac{p_1}{p_2} \cdot B^d > 0 \tag{A 4.3}$$

[11] $\dfrac{\partial X_2}{\partial \varsigma} = \dfrac{\partial Q_2(B)}{\partial B} \cdot \dfrac{\partial B}{\partial \varsigma} = Q_2(B) \cdot \dfrac{1}{p_1 B^{-1} + p_2} \cdot \dfrac{p_1}{B^2} \cdot \dfrac{1}{1-d} \cdot \dfrac{p_1}{p_2} \cdot B^d > 0$, since $\dfrac{\partial B}{\partial \varsigma} = \dfrac{1}{1-d} \cdot \dfrac{p_1}{p_2} \cdot B^d > 0$.

and

$$\frac{\partial X_2}{\partial k_2} = \frac{-1}{p_1 + p_2 \cdot B} < 0 \qquad \text{(A 4.4)}$$

Therefore $\partial X_2(\varsigma_2)/\partial k_2 > 0$.

Next the negative influence of a_2 on $X_2(\varsigma_2)$ is proven. This is done by

$$\frac{\partial X_2(\varsigma_2)}{\partial a_2} = -\frac{\partial g_2(y,k_2,p_1,p_2,d,a_1,a_2,\varsigma)/\partial a_2}{\partial g_2(y,k_2,p_1,p_2,d,a_1,a_2,\varsigma)/\partial \varsigma} \cdot \frac{\partial X_2}{\partial \varsigma} + \frac{\partial X_2}{\partial a_2}. \qquad \text{(A 4.5)}$$

The only component that has not been calculated is $\partial g_2(y,k_2,p_1,p_2,d,a_1,a_2,\varsigma)/\partial a_2$:

$$\frac{\partial g_2(y,k_2,p_1,p_2,d,a_1,a_2,\varsigma)}{\partial a_2} = d \cdot \frac{p_2}{p_1} \cdot B^{1-d}\left(Q_2^{d-1} - a_2^{d-1}\right) > 0 \qquad \text{(A 4.6)[12]}$$

since $Q_2(\varsigma) > a_2 \ \forall \varsigma > \varsigma_2$.

Further

$$\frac{\partial X_2}{\partial a_2} = \frac{\partial Q_2}{\partial a_2} - 1 = \frac{p_2}{p_1 \cdot B^{-1} + p_2} - 1 = -\frac{p_1 \cdot B^{-1}}{p_1 \cdot B^{-1} + p_2} < 0 \qquad \text{(A 4.7)}$$

Plugging these results in (A 4.5) yields:

12

$$\frac{\partial g_2(y,k_2,p_1,p_2,d,a_1,a_2,\varsigma)}{\partial a_2} = d \cdot Q_1^{d-1} \cdot \frac{dQ_1}{da_2} + \frac{p_2}{p_1} \cdot B^{1-d} \cdot d \cdot Q_2^{d-1} \cdot \frac{dQ_2}{da_2} - d \cdot \frac{p_2}{p_1} \cdot B^{1-d} \cdot a_2^{d-1} =$$

$$= d \cdot Q_1^{d-1} \cdot \frac{p_2}{p_1 + p_2 \cdot B} + d \cdot \frac{p_2}{p_1} \cdot B^{1-d} \cdot Q_2^{d-1} \cdot \frac{p_2}{p_1 \cdot B^{-1} + p_2} - d \cdot \frac{p_2}{p_1} \cdot B^{1-d} \cdot a_2^{d-1} =$$

$$= d \cdot B^{1-d} \cdot Q_2^{d-1} \cdot \frac{p_2}{p_1 + p_2 \cdot B} + d \cdot \frac{p_2}{p_1} \cdot B^{1-d} \cdot Q_2^{d-1} \cdot \frac{p_2}{p_1 \cdot B^{-1} + p_2} - d \cdot \frac{p_2}{p_1} \cdot B^{1-d} \cdot a_2^{d-1} =$$

$$= d \cdot B^{1-d} \cdot Q_2^{d-1} \cdot \frac{p_2}{p_1} \cdot \frac{p_1}{p_1 + p_2 \cdot B} + d \cdot \frac{p_2}{p_1} \cdot B^{1-d} \cdot Q_2^{d-1} \cdot \frac{p_2 \cdot B}{p_1 + p_2 \cdot B} - d \cdot \frac{p_2}{p_1} \cdot B^{1-d} \cdot a_2^{d-1} =$$

$$= d \cdot B^{1-d} \cdot Q_2^{d-1} \cdot \frac{p_2}{p_1} \cdot \left(\frac{p_1}{p_1 + p_2 \cdot B} + \frac{p_2 \cdot B}{p_1 + p_2 \cdot B}\right) - d \cdot \frac{p_2}{p_1} \cdot B^{1-d} \cdot a_2^{d-1} = d \cdot \frac{p_2}{p_1} \cdot B^{1-d} \cdot Q_2^{d-1} - d \cdot \frac{p_2}{p_1} \cdot B^{1-d} \cdot a_2^{d-1} =$$

$$= d \cdot \frac{p_2}{p_1} \cdot B^{1-d}\left(Q_2^{d-1} - a_2^{d-1}\right) > 0 \text{ for } Q_2 > a_2, X_2 > 0 \text{ respectively.}$$

$$\frac{\partial X_2(\varsigma_2)}{\partial a_2} = -\frac{d \cdot \frac{p_2}{p_1} \cdot B^{1-d}\left(Q_2^{\ d-1} - a_2^{\ d-1}\right)}{(1-d) \cdot \frac{p_2}{p_1} \cdot B^{-d} \cdot \left(Q_2(B)^d - a_2^{\ d}\right)} \cdot Q_2(B) \cdot \frac{1}{p_1 B^{-1} + p_2} \cdot \frac{p_1}{B^2} \cdot \frac{1}{1-d} \cdot \frac{p_1}{p_2} \cdot B^d - \frac{p_1 \cdot B^{-1}}{p_1 \cdot B^{-1} + p_2} =$$

$$= -\frac{d}{1-d} \cdot B \cdot Q_2(B) \cdot \frac{1}{p_1 B^{-1} + p_2} \cdot \frac{p_1}{B^2} \cdot \frac{1}{1-d} \cdot \frac{p_1}{p_2} \cdot B^d - \frac{p_1 \cdot B^{-1}}{p_1 \cdot B^{-1} + p_2} =$$

$$= -\frac{d}{1-d} \cdot Q_2(B) \cdot \frac{1}{p_1 B^{-1} + p_2} \cdot \frac{p_1}{B} \cdot \frac{1}{1-d} \cdot \frac{p_1}{p_2} \cdot B^d - \frac{p_1 \cdot B^{-1}}{p_1 \cdot B^{-1} + p_2} =$$

$$= -\frac{d}{1-d} \cdot Q_2(B) \cdot \frac{p_1 \cdot B^{-1}}{p_1 B^{-1} + p_2} \cdot \frac{1}{1-d} \cdot \frac{p_1}{p_2} \cdot B^d - \frac{p_1 \cdot B^{-1}}{p_1 \cdot B^{-1} + p_2} =$$

$$= -\frac{p_1 \cdot B^{-1}}{p_1 \cdot B^{-1} + p_2} \cdot \left(\frac{d}{(1-d)^2} \cdot Q_2(B) \cdot \frac{p_1}{p_2} \cdot B^d + 1 \right) < 0$$

since $\partial g_2 / \partial \varsigma > 0$, $\partial g_2 / \partial a_2 > 0$ and

Basically the influence of the parameters are interesting, since it has to examined which parameters influences $X_2(\varsigma_2)$. This is important to know, since a shift of $X_2(\varsigma_2)$ leads to the problem, that the Maximum Likelihood function will not be differentiable, it there is any observation within the interval $(0..X_2(\varsigma_2))$.

Next the influence of m on $X_2(\varsigma_2)$ is proven. This is done by

$$\frac{\partial X_2(\varsigma_2)}{\partial m} = -\frac{\partial g_2(y, k_2, p_1, p_2, d, a_1, a_2, \varsigma) / \partial B \cdot \partial B / \partial m}{\partial g_2(y, k_2, p_1, p_2, d, a_1, a_2, \varsigma) / \partial B \cdot \partial B / \partial \varsigma} \cdot \frac{\partial X_2}{\partial \varsigma} + \frac{\partial X_2}{\partial m} = -\frac{\partial X_2}{\partial \varsigma} + \frac{\partial X_2}{\partial m} = 0,$$

since

etc...

A 5 Calculating the observation probability for three goods without fixed costs

In the following it is shown, how random variable of formula (1.4.17) is integrated out.

First, for a simpler notation, (1.4.17) has to be rewritten as

$$E_{\xi_1}\left(f_{X_2 \wedge case2}(z) \mid \xi_1\right) = l_z(z) \cdot \int_{x=-\infty}^{x=\infty} f_\xi\left(V_1 - V_2 + x\right) \cdot F_\xi\left(V_1 - V_3 + x\right) \cdot f_\xi(x)\, dx, \qquad \text{(A 5.1)}$$

where $V_1\left(\frac{y - p_2 z}{p_1}\right) - V_2(z) = l\left(\frac{y - p_2 z}{p_1}, z\right)$, $V_1 = V_1(0)$ and $l_z(z) = \frac{\partial l\left(\frac{y - p_2 z}{p_1}, z\right)}{\partial z}$.

Now, the explicit expressions of $f_\xi\left(V_1 - V_2 + x\right)$ and $F_\xi\left(V_1 - V_3 + x\right)$ has to be plugged in and then integral (A 5.1) has to be solved:

$$\begin{aligned}
&E_{\xi_1}\left(f_{X_2 \wedge case2}(z) \mid \xi_1\right) = \\
&= l_z(z) \cdot \int_{x=-\infty}^{x=\infty} \exp\left(-e^{-(V_1 - V_2 + x)}\right) \cdot e^{-(V_1 - V_2 + x)} \cdot \exp\left(-e^{-(V_1 - V_3 + x)}\right) \cdot \exp\left(-e^{-x}\right) \cdot e^{-x}\, dx = \\
&= l_z(z) \cdot \int_{x=-\infty}^{x=\infty} \exp\left(-\left(e^{-(V_1 - V_2)} + e^{-(V_1 - V_3)} + 1\right) \cdot e^{-x}\right) \cdot e^{-(V_1 - V_2)} \cdot \left(e^{-x}\right)^2 dx = \\
&= l_z(z) \cdot e^{-(V_1 - V_2)} \cdot \int_{x=-\infty}^{x=\infty} \exp\left(-\left(e^{-(V_1 - V_2)} + e^{-(V_1 - V_3)} + 1\right) \cdot e^{-x}\right) \cdot \left(e^{-x}\right)^2 dx.
\end{aligned}$$

Now substitution $y = -e^{-x}$, $dy = e^{-x}dx$, $y(-\infty) = -\infty$, $y(\infty) = 0$ allows for simplifications. The new integral limits are $y(-\infty) = -\infty$, $y(\infty) = 0$:

$$\begin{aligned}
&E_{\xi_1}\left(f_{X_2 \wedge case2}(z) \mid \xi_1\right) = \\
&= l_z(z) \cdot e^{-(V_1 - V_2)} \cdot \int_{y=-\infty}^{y=0} \exp\left(\left(e^{-(V_1 - V_2)} + e^{-(V_1 - V_3)} + 1\right) \cdot y\right) \cdot (-y)^2 \cdot (-y)^{-1}\, dy = \\
&= -l_z(z) \cdot e^{-(V_1 - V_2)} \cdot \int_{y=-\infty}^{y=0} \exp\left(\left(e^{-(V_1 - V_2)} + e^{-(V_1 - V_3)} + 1\right) \cdot y\right) \cdot y^2 dy = \\
&= -l_z(z) \cdot e^{-(V_1 - V_2)} \cdot \int_{y=-\infty}^{y=0} \exp\left(\left(e^{-(V_1 - V_2)} + e^{-(V_1 - V_3)} + 1\right) \cdot y\right) \cdot y^2 dy.
\end{aligned}$$

Substitution

$$s=\left(e^{-(V_1-V_2)}+e^{-(V_1-V_3)}+1\right)\cdot y,\ ds=\left(e^{-(V_1-V_2)}+e^{-(V_1-V_3)}+1\right)\cdot dy,$$
$$y=\left(e^{-(V_1-V_2)}+e^{-(V_1-V_3)}+1\right)^{-1}\cdot s,\ dy=\left(e^{-(V_1-V_2)}+e^{-(V_1-V_3)}+1\right)^{-1}\cdot ds$$

with the corresponding integration limits $s(-\infty)=-\infty,\ s(0)=0$ yields further simplifications:

$$E_{\xi_1}\left(f_{X_2\wedge case2}(z)\mid\xi_1\right)=$$
$$=-l_z(z)\cdot e^{-(V_1-V_2)}\cdot\int_{s=-\infty}^{s=0}\exp(s)\cdot\left(e^{-(V_1-V_2)}+e^{-(V_1-V_3)}+1\right)^{-2}\cdot s^2\cdot\left(e^{-(V_1-V_2)}+e^{-(V_1-V_3)}+1\right)^{-1}ds=$$
$$=-l_z(z)\cdot\frac{e^{-(V_1-V_2)}}{\left(e^{-(V_1-V_2)}+e^{-(V_1-V_3)}+1\right)^3}\cdot\int_{s=-\infty}^{s=0}\exp(s)\cdot sds.$$

Integral $\int_{s=-\infty}^{s=0}\exp(s)\cdot sds$ is now be solved by integration by parts, $\int_{x=a}^{x=b}uv'dx=(uv)_{x=a}^{x=b}-\int_{x=a}^{x=b}u'vdx$:

$$E_{\xi_1}\left(f_{X_2\wedge case2}(z)\mid\xi_1\right)=$$
$$=-l_z(z)\cdot\frac{e^{-(V_1-V_2)}}{\left(e^{-(V_1-V_2)}+e^{-(V_1-V_3)}+1\right)^3}\cdot\left(\left(\exp(s)\cdot s\right)_{s=-\infty}^{s=0}-\int_{s=-\infty}^{s=0}\exp(s)ds\right)=$$
$$=-l_z(z)\cdot\frac{e^{-(V_1-V_2)}}{\left(e^{-(V_1-V_2)}+e^{-(V_1-V_3)}+1\right)^3}\cdot\left(0-\left(\exp(s)\right)_{s=-\infty}^{s=0}\right)=$$
$$=-l_z(z)\cdot\frac{e^{-(V_1-V_2)}}{\left(e^{-(V_1-V_2)}+e^{-(V_1-V_3)}+1\right)^3}\cdot(0-1)=$$
$$=l_z(z)\cdot\frac{e^{-(V_1-V_2)}}{\left(e^{-(V_1-V_2)}+e^{-(V_1-V_3)}+1\right)^3}.$$

The final result is therefore:

$$f_{X_2\wedge case2}(z)=l_z(z)\cdot\frac{e^{-(V_1-V_2)}}{\left(e^{-(V_1-V_2)}+e^{-(V_1-V_3)}+1\right)^3},\qquad\text{(A 5.1)}$$

where $V_1=V_1\left(\frac{y-p_2z}{p_1}\right),\ V_2=V_2(z)$ and $V_3=V_3(0)$.

A 6 Calculating the observation probability for three goods when there are fixed costs

In contrast to the case where there are no fixed cost the probabilities for the different cases one to four are not functions of closed form, but have to be calculated by simulation routine. Therefore, densities have now to be determined conditional on a certain case:

$$f_{X_2 \wedge case_i}(z) = f_{X_2|case2}(z) \cdot P(case_i) \tag{A 6.1}$$

Now for case 2, conditional probability density for X_2 (1.5.12) has to be proven. First, formulate the density for *l* instead for *z*.

From $l = l\left(\frac{y - p_2 z}{p_1}, z\right)$ and (1.5.12) follows

$$f_L(l) \mid case2, \xi_1 = f_\xi(l + \xi_1). \tag{A 6.2}$$

Applying rule 8 of appendix A 1 yields

$$f_L(l) \mid case2 = E_{\xi_1}\left(f_L(l) \mid case2, \xi_1\right) = f_\xi(l + \xi_1) = \frac{e^{-l}}{\left(1 + e^{-l}\right)^2}. \tag{A 6.3}$$

From this, density of X_2 can be calculated:

$$f_{X_2}(l) \mid case2 = \frac{e^{-l}}{\left(1 + e^{-l}\right)^2} \cdot l_z(z) \text{, where } l_z(z) = \frac{dl(z)}{dz}. \tag{A 6.4}$$

$$E_X\left(\prod_{i=1}^{n} f_\xi(X + a_i)\right) = E_X\left(\prod_{i=1}^{n} \exp\left(-e^{-(X+a_i)}\right) \cdot e^{-(X+a_i)}\right) = E_X\left(\prod_{i=1}^{n} \exp\left(-e^{-X} \cdot e^{-a_i}\right) \cdot e^{-X} \cdot e^{-a_i}\right) =$$

$$= E_X\left(\left(e^{-X}\right)^n \cdot \exp\left(-e^{-X} \cdot \sum_{i=1}^{n} e^{-a_i}\right) \cdot e^{-\sum_{i=1}^{n} a_i}\right) = e^{-\sum_{i=1}^{n} a_i} \cdot E_X\left(\left(e^{-X}\right)^n \cdot \exp\left(-e^{-X} \cdot \sum_{i=1}^{n} e^{-a_i}\right)\right) =$$

$$= e^{-\sum_{i=1}^{n} a_i} \cdot \int_{x=-\infty}^{x=\infty} \left(e^{-x}\right)^n \cdot \exp\left(-e^{-x} \cdot \sum_{i=1}^{n} e^{-a_i}\right) \cdot \exp\left(-e^{-x}\right) \cdot e^{-x} dx =$$

$$= e^{-\sum_{i=1}^{n} a_i} \cdot \int_{x=-\infty}^{x=\infty} \left(e^{-x}\right)^{n+1} \cdot \exp\left(-e^{-x} \cdot \left(1 + \sum_{i=1}^{n} e^{-a_i}\right)\right) dx =$$

$$= e^{-\sum_{i=1}^{n} a_i} \cdot \left(1 + \sum_{i=1}^{n} e^{-a_i}\right)^{-(n+1)} \cdot \int_{x=-\infty}^{x=\infty} \left(e^{-x} \cdot \left(1 + \sum_{i=1}^{n} e^{-a_i}\right)\right)^{n+1} \cdot \exp\left(-e^{-x} \cdot \left(1 + \sum_{i=1}^{n} e^{-a_i}\right)\right) dx.$$

Substituting $t = e^{-x} \cdot \left(1 + \sum_{i=1}^{n} e^{-a_i}\right),\ dt = -e^{-x} \cdot \sum_{i=1}^{n} e^{-a_i} dx,\ t = \infty..0$ yields

$$E_X\left(\prod_{i=1}^{n} f_\xi\left(X + a_i\right)\right) = -e^{-\sum_{i=1}^{n} a_i} \cdot \left(1 + \sum_{i=1}^{n} e^{-a_i}\right)^{-(n+1)} \cdot \int_{x=\infty}^{x=0} t^n \cdot e^{-t} dt =$$

$$= e^{-\sum_{i=1}^{n} a_i} \cdot \left(1 + \sum_{i=1}^{n} e^{-a_i}\right)^{-(n+1)} \cdot \int_{x=0}^{x=\infty} t^n \cdot e^{-t} dt. \qquad \text{(A 6.5)}$$

Integral $\int_{x=0}^{x=\infty} t^n \cdot e^{-t} dt$ has to be solved by integration by parts in a recursive way:

$S_n = \int_{x=0}^{x=\infty} t^n \cdot e^{-t} dt = \left[-t^n e^{-t}\right]_{t=-\infty}^{t=0} - n \cdot \int_{x=0}^{x=\infty} t^{n-1} \cdot (-1) \cdot e^{-t} dt = n \cdot \int_{x=0}^{x=\infty} t^{n-1} \cdot e^{-t} dt = n \cdot S_{n-1}$, where

$S_0 = \int_{x=0}^{x=\infty} e^{-t} dt = \left[-e^{-t}\right]_{t=-\infty}^{t=0} = 1$. From this follows

$$S_n = \int_{x=0}^{x=\infty} t^n \cdot e^{-t} dt = n!. \qquad \text{(A 6.6)}$$

By use of this result, equation (A 6.5) can be solved:

$$E_X\left(\prod_{i=1}^{n} f_\xi\left(X + a_i\right)\right) = n! \cdot \frac{e^{-\sum_{i=1}^{n} a_i}}{\left(1 + \sum_{i=1}^{n} e^{-a_i}\right)^{n+1}}. \qquad \text{(A 6.7)}$$

A 5 Two goods with fixed costs: Critical relative preference

In the following the critical relative preference where households switch from spending all budget for good one to spending budget for both goods one and two, where good two is the good with fixed cost, is computed. It will be proven, that critical value in the case where there are no fixed cost is lower, than in the case where good has positive fixed cost. As presented in section 1.3, roots of functions g1 and g2 are relevant, namely

Since the density of ς is known as will be shown later, the values ς_1 and ς_2 have now to be calculated. Further it has to be proven that $g_1(\varsigma)$ and $g_2(\varsigma)$ have a unique solutions, means that their first derivative is always negative.

First, condition one is considered. The solution for ς_1 is:[1]

$$\varsigma_1 = (1-d)\cdot\ln(B_1) - \ln\left(\frac{p_1}{p_2}\right) - m \text{, where } B_1 = \frac{p_1 \cdot a_2}{y - k_2 + p_1 \cdot a_1}. \qquad \text{(A3.1)}$$

Next it has to be proven, if this solution is unique. This is the case, when the first derivative of $g_1(\varsigma)$ does not change sign.[2]

$$\frac{\partial g_1(\varsigma)}{\partial \varsigma} = \frac{\partial \breve{g}_1(B(\varsigma))}{\partial B(\varsigma)} \cdot \frac{\partial B(\varsigma)}{\partial \varsigma} > 0 \text{, with } g_1(\varsigma) = \breve{g}_1(B(\varsigma)), \qquad \text{(A3.2)}$$

since $\frac{\partial g_1(B(\varsigma))}{\partial B(\varsigma)} = -\frac{y - k_2 + p_1 \cdot a_1 + p_2 \cdot a_2}{\left(p_1 \cdot B^{-1} + p_2\right)^2} \cdot p_1 \cdot (-1) \cdot B^{-2} > 0$ and

[1]
$$0 = \frac{y - k_2 + p_1 \cdot a_1 + p_2 \cdot a_2}{p_1 \cdot B^{-1} + p_2} - a_2 \Leftrightarrow \frac{y - k_2 + p_1 \cdot a_1 + p_2 \cdot a_2}{p_1 \cdot B^{-1} + p_2} = a_2 \Leftrightarrow y - k_2 + p_1 \cdot a_1 + p_2 \cdot a_2 = p_1 \cdot a_2 \cdot B^{-1} + p_2 \cdot a_2 \Leftrightarrow$$
$$\Leftrightarrow y - k_2 + p_1 \cdot a_1 = p_1 \cdot a_2 \cdot B^{-1} \Leftrightarrow B = \frac{p_1 \cdot a_2}{y - k_2 + p_1 \cdot a_1}.$$
$$B = \left(\frac{p_1}{p_2} \cdot \exp(m + \varsigma)\right)^{\frac{1}{1-d}} \Leftrightarrow B^{1-d} = \frac{p_1}{p_2} \cdot \exp(m+\varsigma) \Leftrightarrow \frac{p_2}{p_1} \cdot B^{1-d} = \exp(m+\varsigma) \Leftrightarrow \ln\left(\frac{p_2}{p_1} \cdot B^{1-d}\right) = m + \varsigma \Leftrightarrow$$
$$\Leftrightarrow \varsigma = \ln\left(\frac{p_2}{p_1} \cdot B^{1-d}\right) - m.$$

[2]
$$\frac{\partial g_1(\varsigma)}{\partial \varsigma} = \frac{\partial X_1}{\partial \varsigma} = \frac{\partial Q_1}{\partial \varsigma} = -\frac{y - k_2 + p_1 \cdot a_1 + p_2 \cdot a_2}{\left(p_1 \cdot B^{-1} + p_2\right)^2} \cdot p_1 \cdot (-1) \cdot B^{-2} \cdot \frac{\partial B}{\partial \varsigma} =$$
$$= -\frac{y - k_2 + p_1 \cdot a_1 + p_2 \cdot a_2}{\left(p_1 \cdot B^{-1} + p_2\right)^2} \cdot p_1 \cdot (-1) \cdot B^{-2} \cdot \left(\frac{p_1}{p_2}\right)^{\frac{1}{1-d}} \cdot \exp\left(\frac{m+\varsigma}{1-d}\right) \cdot \frac{1}{1-d} > 0$$

$$\frac{\partial B(\varsigma)}{\partial \varsigma} = \left(\frac{p_1}{p_2}\right)^{\frac{1}{1-d}} \cdot \exp\left(\frac{m+\varsigma}{1-d}\right) \cdot \frac{1}{1-d} = \frac{1}{1-d} \cdot B > 0 \cdot$$

Therefore, solution ς_1 of condition one is unique.

Since for condition two an explicit solution of ς_2 does not exist and therefore needs to be calculated numerically, it has to be proven that $g_2'(\varsigma)$ is positive in the relevant range to assure for a unique solution ς_2. Again the prove is done by derivation first with respect to *B* and then *B* with respect to ς:

$$\frac{\partial g_2(\varsigma)}{\partial \varsigma} = \frac{\partial \breve{g}_2(B(\varsigma))}{\partial B(\varsigma)} \cdot \frac{\partial B(\varsigma)}{\partial \varsigma}. \tag{A3.3}$$

Since $\partial B(\varsigma)/\partial \varsigma$ is greater than zero as shown above, it is sufficient to prove $\partial \breve{g}_2(B(\varsigma))/\partial B(\varsigma) > 0$. Therefore, expression $g_2(\bullet)$ has first to be expressed as a function of *B* by plugging in $\exp(m+\varsigma) = \frac{p_2}{p_1} \cdot B^{1-d}$:[3]

$$\breve{g}_2(B) = Q_1(B)^d + \frac{p_2}{p_1} \cdot B^{1-d} \cdot Q_2(B)^d - \left(\frac{y}{p_1}\right)^d - \frac{p_2}{p_1} \cdot B^{1-d} \cdot a_2{}^d, \tag{A3.4}$$

with $Q_1(B) = \frac{y - k_2 + p_1 \cdot a_1 + p_2 \cdot a_2}{p_1 \cdot B^{-1} + p_2}$ and $Q_2(B) = \frac{y - k_2 + p_1 \cdot a_1 + p_2 \cdot a_2}{p_1 + p_2 \cdot B}$.

Derivation $\partial \breve{g}_2(B)/\partial B$ yields now:[4]

---

[3] $B = \left(\frac{p_1}{p_2} \cdot \exp(m+\varsigma)\right)^{\frac{1}{1-d}} \Leftrightarrow \frac{p_1}{p_2} \cdot \exp(m+\varsigma) = B^{1-d} \Leftrightarrow \exp(m+\varsigma) = \frac{p_2}{p_1} \cdot B^{1-d}$.

[4] $$\frac{d\breve{g}_2(B)}{dB} = d \cdot Q_1(B)^{d-1} \cdot \frac{dQ_1(B)}{dB} + (1-d) \cdot \frac{p_2}{p_1} \cdot B^{-d} \cdot Q_2(B)^d + \frac{p_2}{p_1} \cdot B^{1-d} \cdot d \cdot Q_2(B)^{d-1} \cdot \frac{dQ_2(B)}{dB} - (1-d) \cdot \frac{p_2}{p_1} \cdot B^{-d} \cdot a_2{}^d,$$

$Q_1(B) = B^{-1} \cdot Q_2(B)$,

$$\frac{\partial Q_1(B)}{\partial B} = -\frac{y - k_2 + p_1 \cdot a_1 + p_2 \cdot a_2}{(p_1 + p_2 \cdot B)^2} \cdot p_2 = -Q_1(B) \cdot \frac{p_2}{p_1 + p_2 \cdot B},$$

$$\frac{\partial Q_2(B)}{\partial B} = -\frac{y - k_2 + p_1 \cdot a_1 + p_2 \cdot a_2}{(p_1 \cdot B^{-1} + p_2)^2} \cdot \frac{-p_1}{B^2} = Q_2(B) \cdot \frac{1}{p_1 \cdot B^{-1} + p_2} \cdot \frac{p_1}{B^2} = Q_2(B) \cdot \frac{1}{p_1 + p_2 \cdot B} \cdot \frac{p_1}{B},$$

$$\frac{d\breve{g}_2(B)}{dB} = (1-d)\cdot\frac{p_2}{p_1}\cdot B^{-d}\cdot\left(Q_2(B)^d - a_2{}^d\right). \qquad (1.3.30)$$

This means that $d\breve{g}_2(B)/dB > 0$ for $B < B_1$ and $d\breve{g}_2(B)/dB \le 0$ for $B \ge B_1$, respectively $dg_2(\varsigma)/d\varsigma > 0$ for $\varsigma < \varsigma_1$ and $dg_2(\varsigma)/d\varsigma \le 0$ for $\varsigma \ge \varsigma_1$ and that $g_2(\varsigma_2)$ is the minimum[5] of function $g_2(\varsigma)$. Therefore $g_1(\varsigma)$ and $g_2(\varsigma)$ diagram look about like:

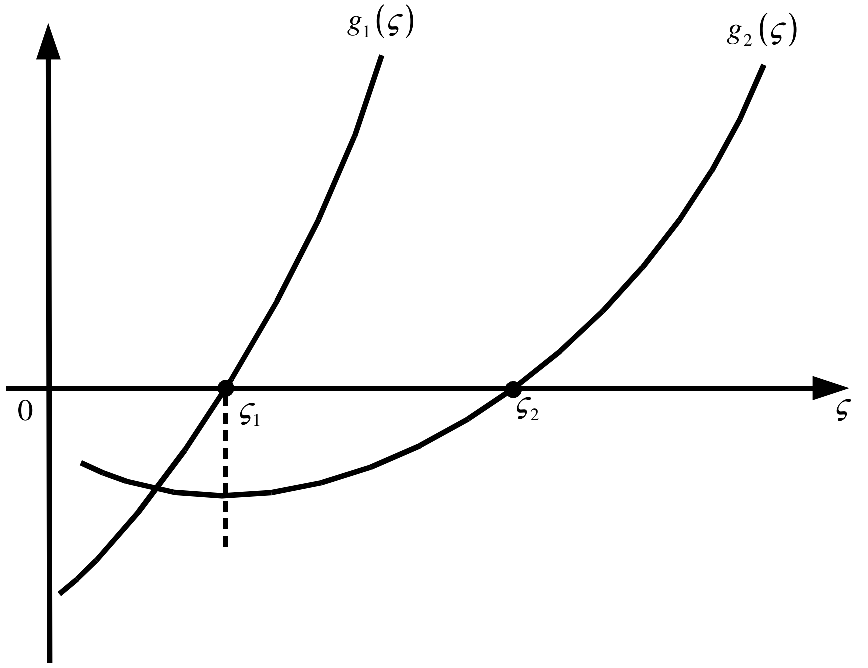


Diagram A.3.1: The principle of calculating the probability of good two being zero

The prove $\varsigma_2 > \varsigma_1$ is now as follows:

Since $g_2(\varsigma_1) < 0$, $dg_2(\varsigma)/d\varsigma > 0$ and finite for any $\varsigma > \varsigma_1$ and $\lim_{\varsigma\to\infty} g_2(\varsigma) = \infty$[6] it follows, that there exists a unique ς_2, such that $g_2(\varsigma_2) = 0$ if and only if the income is greater than the fixed cost k_2, $y > k_2 - p_1\cdot a_1$ where parameter a_1 is assumed to be zero. Therefore in this case there exists a unique solution for $g_2(\varsigma)$ being zero, namely $g_2(\varsigma_2) = 0$.

---

[5] $g_2(\varsigma_1) = \breve{g}_2(B_1) = Q_1(B_1)^d + \frac{p_2}{p_1}\cdot B_1^{1-d}\cdot Q_2(B_1)^d - \left(\frac{y}{p_1}\right)^d - \frac{p_2}{p_1}\cdot B_1^{1-d}\cdot a_2{}^d =$
$= \left(\frac{y-k_2}{p_1}\right)^d + \frac{p_2}{p_1}\cdot B_1^{1-d}\cdot a_2{}^d - \left(\frac{y}{p_1}\right)^d - \frac{p_2}{p_1}\cdot B_1^{1-d}\cdot a_2{}^d = \left(\frac{y-k_2}{p_1}\right)^d - \left(\frac{y}{p_1}\right)^d < 0.$

If the income is smaller than the fixed cost k_2 then household cannot consume good two independent of its preference and therefore in this case probability of choosing a positive amount of good two is zero.

Therefore the probability of X_2 being zero is

$$P(X_2 = 0) \mid (y_D = y - k_2, p_1, p_2 \mid \theta) = F_\varsigma(\varsigma_2) \cdot I_Y + (1 - I_Y), \qquad (1.3.31)$$

where $\varsigma_2 = \ln\left(\frac{p_2}{p_1} \cdot B_2^{1-d}\right) - m$, where $\breve{g}_2(B_2) = 0$, and $I_Y = \begin{cases} y > k_2 - p_1 \cdot a_1 : 1 \\ y \le k_2 - p_1 \cdot a_1 : 0 \end{cases}$.

A 6 Three goods with fixed costs: Probabilities for the four cases

In the following it is shown, how the probabilities for the case one to four depend on parameters and economic variables. To remind, the cases are defined as follows:

Case 1:	X1>0	X2=0	X3=0
Case 2:	X1>0	X2>0	X3=0
Case 3:	X1>0	X2=0	X3>0
Case 4:	X1>0	X2>0	X3>0

Table A6.1: List of cases for the three good case

$$\lim_{\varsigma\to\infty} g_2(\varsigma) = \lim_{\varsigma\to\infty}\left(\frac{y - k_2 + p_1 \cdot a_1 + p_2 \cdot a_2}{p_1 + p_2 \cdot B}\right)^d + \exp(m+\varsigma)\cdot\left(\frac{y - k_2 + p_1 \cdot a_1 + p_2 \cdot a_2}{p_1 \cdot B^{-1} + p_2}\right)^d - \left(\frac{y}{p_1}\right)^d - \exp(m+\varsigma)\cdot a_2^{\,d} =$$

$$= \lim_{\varsigma\to\infty} 0^d + \exp(m+\varsigma)\cdot\left(\frac{(y - k_2 + p_1 \cdot a_1 + p_2 \cdot a_2)\cdot B}{p_1 + p_2 \cdot B}\right)^d - \left(\frac{y}{p_1}\right)^d - \exp(m+\varsigma)\cdot a_2^{\,d} =$$

$$= \lim_{\varsigma\to\infty} \exp(m+\varsigma)\cdot\left(\frac{y - k_2 + p_1 \cdot a_1 + p_2 \cdot a_2}{p_2}\right)^d - \left(\frac{y}{p_1}\right)^d - \exp(m+\varsigma)\cdot a_2^{\,d} =$$

[6] $$= \lim_{\varsigma\to\infty} \exp(m+\varsigma)\cdot\left(\left(\frac{y - k_2 + p_1 \cdot a_1 + p_2 \cdot a_2}{p_2}\right)^d - a_2^{\,d}\right) - \left(\frac{y}{p_1}\right)^d = \infty$$

since

$$\lim_{\varsigma\to\infty}\left(\frac{p_1}{p_2}\right)^{\frac{1}{1-d}} \cdot \exp\left(\frac{m+\varsigma}{1-d}\right) = \infty,$$

$\frac{y - k_2 + p_1 \cdot a_1 + p_2 \cdot a_2}{p_2} > a_2$ if and only if $y > k_2 - p_1 \cdot a_1$.

Proof: $\frac{y - k_2 + p_1 \cdot a_1 + p_2 \cdot a_2}{p_2} > a_2 \Leftrightarrow y - k_2 + p_1 \cdot a_1 + p_2 \cdot a_2 > p_2 \cdot a_2 \Leftrightarrow y > k_2 - p_1 \cdot a_1$.

In the following, first effects of changes of the economic variables k_2, p_2 and y on probabilities of the different cases are examined. This is done by illustrating how the sets of case one to four in (ξ_2, ξ_3) space are changing compared to the baseline case with economic variables $(k_1, k_2, k_3) = (0, 0.8, 1.2)$, $y = 10$, $(p_1, p_2, p_3) = (1,1,1)$ and parameters $(m_1, m_2, m_3) = (0,0,0)$, $(a_1, a_2, a_3) = (0,1,1)$, $(d_1, d_2, d_3) = (0.5, 0.5, 0.5)$. Random variable ξ_1 is set $\xi_1 = 0$.

Note that the brighter boundaries always indicate the baseline case, while as the black boundaries indicate the boundaries for he modified parameters and economic variables.

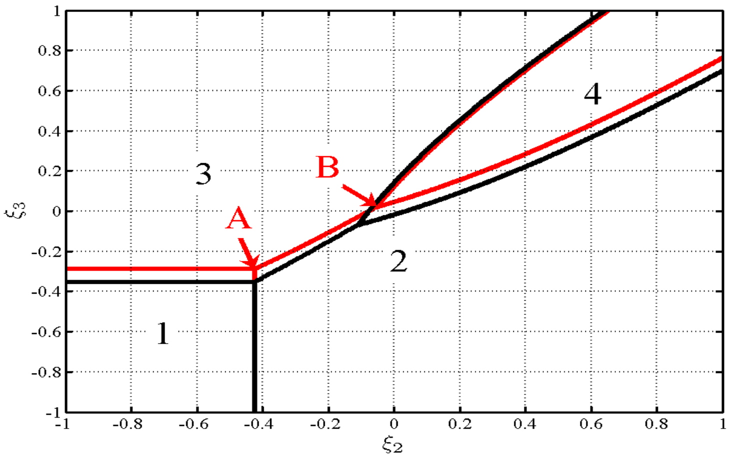


Figure A6.1: Effect of a decrease of fixed cost of good three[7]

A decrease of fixed cost of good three shift point A in direction of the x- axis. This means that the level of relative preference for good three where household switch from case one to case three is now lower. Also the line separating case three and case two is now also shifting towards the x-axis. Over all, the area for case 4 also becomes larger. In total, the following on probabilities can be summarized:

$P(case1)\downarrow$	$P(case2)\downarrow$
$P(case3)\uparrow$	$P(case4)\uparrow$

Table A 5.2 : Table of changes in probabilities for a decrease of fixed cost of good three

Next the case where price of good three decreases is examined.

[7]Fixed cost of good three, k_3, has decreased from 1.2 to 1.0.

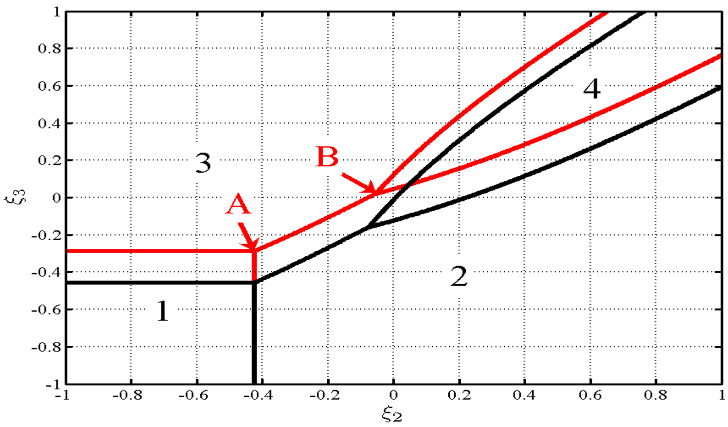


Figure A6.2: Effect of a decrease of price of good three[8]

A decrease in price of good three leads to a shift of all boundary levels of preference towards the x-axis. This implies, that probabilities both for case one and case two are decreasing. Probability of case four is also increasing, since the area of case four is shifted towards an area, where density of probability is higher.[9] Probability of case three is increasing: If mileage cost for the big car is decreasing, consumer tends to buy this type of car.

$P(case1)\downarrow$	$P(case2)\downarrow$
$P(case3)\uparrow$	$P(case4)\uparrow\downarrow$

Table A 5.3 : Table of changes in probabilities for a decrease of marginal costs of good three

When income increases it can be expected that fixed cost matter less and therefore that probability of case one decreases while as all other probabilities are increasing.

[8]Fixed costs of good three, k_3, have decreased from 1.0 to 0.8.

[9]Assuming that ξ_2 and ξ_3 are standard Gumbel distributed and independent. This finding also holds for other distributions with density centred around zero, like for instant Normal distribution with mean zero. For other distributions, this finding can change.

Lesefehler

Figure A6.3: Effect of an increase in income[10]

The effects are like presumed. An interesting detail is, that the line A-B has shifted towards the x-axis. The reason for this is, that for higher income it need critical preference for the good three relative to preference for good two is less. Applied to cars it means, that due to higher income, households switch rather from small to big cars. Or with other words: Given certain preferences, a household switches from owning a small car to a big car. Whether probabilities for case two and case three are increasing depends on several factors. Both case two and case th ntervailing effects: First the switch to consume one of the two goods is at t on the other hand the switch to consume both goods is also on lower f the thwo effects prevails depends on parameters and the assumption on t f (ξ_2, ξ_3).

$P(case1)\downarrow$	$P(case2)\uparrow\downarrow$
$P(case3)\uparrow\downarrow$	$P(case4)\uparrow$

ood three

examined. First, a

( ) ( ) ( ) can be seen by the

( ) in m will just shift

( ) boundaries will be

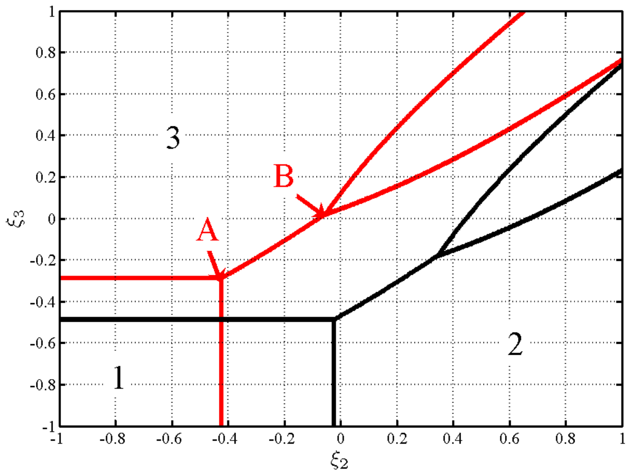


Figure A6.4: Effect of a change in parameters *m*

For this change, an increase in m_2 and a decrease in m_3, the boundaries of the cases are shifted towards the top left. This means that the probability for case three has decreased and that probability for case two has increased. For the change in probabilities of case one and case four there the sign of change will depend on the assumptions on the distribution of (ξ_2, ξ_3).

$P(case1)\uparrow\downarrow$	$P(case2)\uparrow$
$P(case3)\uparrow$	$P(case4)\uparrow\downarrow$

Table A6.5 : Table of changes in probabilities for a change in parameters *m*

Next, all parameters *d* are decreased. That implies, that the decrease of marginal utilities of all goods are getting smaller. Therefore spreading consumption over more goods tends to yield higher utility. This change is equivalent to a shift of the boundaries of cases one to four towards smaller ξ_2 and ξ_3.

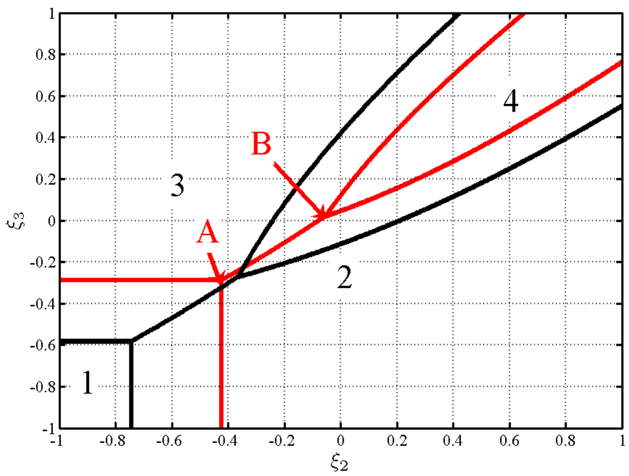


Figure A6.5: Effect of a decrease of all parameters *d*[11]

Therefore probability of case one gets smaller and probability of case four gets higher. For case two and three the sign of change will depend on the assumptions on the distribution of (ξ_2, ξ_3).

$P(case1)\downarrow$	$P(case2)(\uparrow\downarrow)$
$P(case3)(\uparrow\downarrow)$	$P(case4)\uparrow$

Table A6.6: Table of changes in probabilities for a change in parameters *m*

Last an increase of parameter a3 is examined. An increase of parameter a3 is increasing (x3+a3) for given x3. Therefore it will shift partial utility of good three towards a range, where marginal utility is decreasing. This effect will lead to a situation where probability of choosing case three is decreasing.

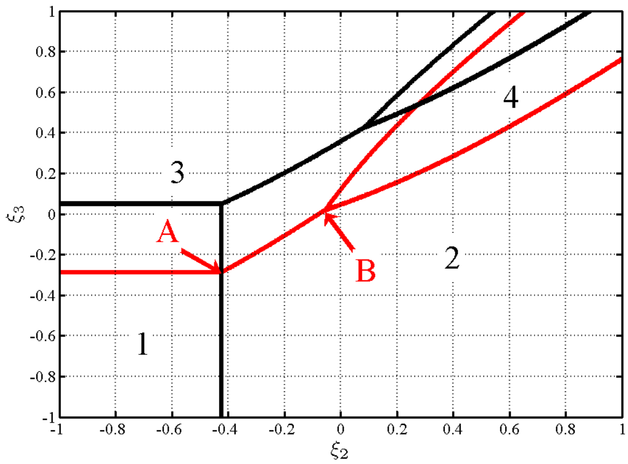


Figure A6.6: Effect of an increase of parameter *a* of good three[12]

Diagram above show, that the both boundaries fore cases two and three are sifted away from the x-axis. Therefore probability for case two is increasing while as probability for case three is decreasing. For distributions of (ξ_2, ξ_3) that have most mass around zero, probability for case four is decreasing, but for other distributions also more a increase could be possible.

$P(case1)\uparrow$	$P(case2)\uparrow$
$P(case3)\downarrow$	$P(case4)(\downarrow)$

Table A6.7: Table of changes in probabilities for a change in parameters *m*

[11]Parameter *d* changed from $d=(d_1,d_2,d_3)=(0.5,0.5,0.5)$ to $d=(0.3,0.3,0.3)$.

[12]Parameter *a* changed from $a=(a_1,a_2,a_3)=(0,1,1)$ to $a=(0,1,3)$.

A 7 Three goods with fixed costs: Optimal consumption in case 4

In the following, optimal consumption and utility for case 4 are computed. It is assumed, that solutions will be interior.
First order conditions of Lagrangian

$$L = \exp(m_1+\xi_1)\cdot(X_1+a_1)^d + \exp(m_2+\xi_2)\cdot(X_2+a_2)^d + \exp(m_3+\xi_3)\cdot(X_3+a_3)^d + \\ +\lambda\cdot\left(y-p_1\cdot X_1-p_2\cdot X_2-p_3\cdot X_3-I(X_2)\cdot k_2-I(X_3)\cdot k\right) \quad \text{(A7.1)}$$

are as follows:

$$\frac{d}{p_i}\cdot\exp(m_i+\xi_i)\cdot(X_i+a_i)^{d-1}=\lambda,\ i=1,2,3. \quad \text{(A7.2)}$$

Setting these first order condition equal, yields:[13]

$$X_i = B_{ij}\cdot\left(X_j+a_j\right)-a_i,\ \text{with } B_{ij}=\left(\frac{p_j}{p_i}\right)^{\frac{1}{1-d}}\cdot\exp\left(\frac{m_i-m_j+\xi_i-\xi_j}{1-d}\right). \quad \text{(A7.3)}$$

Using this result, budget restriction can be written as follows:

$$y-k_2-k_3=p_1\cdot\left(B_{1j}\cdot(X_3+a_3)-a_1\right)-p_2\cdot\left(B_{2j}\cdot(X_3+a_3)-a_2\right)-p_3\cdot X_3\,. \quad \text{(A7.4)}$$

Solving for X_3 yields:[14]

$$X_3=\frac{y-k_2-k_3-p_1\cdot(B_{13}\cdot a_3-a_1)-p_2\cdot(B_{23}\cdot a_3-a_2)}{p_1\cdot B_{13}+p_2\cdot B_{23}+p_3}\,. \quad \text{(A7.5)}$$

13
$$\left(\frac{d}{p_i}\right)^{\frac{1}{d-1}}\cdot\exp(m_i+\xi_i)^{\frac{1}{d-1}}\cdot(X_i+a_i)=\left(\frac{d}{p_j}\right)^{\frac{1}{d-1}}\cdot\exp\left(m_j+\xi_j\right)^{\frac{1}{d-1}}\cdot\left(X_j+a_j\right)\Leftrightarrow$$
$$\Leftrightarrow(X_i+a_i)=\left(\frac{p_i}{p_j}\right)^{\frac{1}{d-1}}\cdot\exp\left(m_j-m_i+\xi_j-\xi_i\right)^{\frac{1}{d-1}}\cdot\left(X_j+a_j\right)\Leftrightarrow$$
$$\Leftrightarrow(X_i+a_i)=\left(\frac{p_j}{p_i}\right)^{\frac{1}{1-d}}\cdot\exp\left(\frac{m_i-m_j+\xi_i-\xi_j}{1-d}\right)\cdot\left(X_j+a_j\right)\Leftrightarrow$$
$$\Leftrightarrow(X_i+a_i)=B_{ij}\cdot\left(X_j+a_j\right),\ \text{with } B_{ij}=\left(\frac{p_j}{p_i}\right)^{\frac{1}{1-d}}\cdot\exp\left(\frac{m_i-m_j+\xi_i-\xi_j}{1-d}\right).$$

14
$$y-k_2-k_3-p_1\cdot(B_{13}\cdot a_3-a_1)-p_2\cdot(B_{23}\cdot a_3-a_2)=p_1\cdot B_{13}\cdot X_3+p_2\cdot B_{23}\cdot X_3+p_3\cdot X_3\Leftrightarrow$$
$$\Leftrightarrow X_3=\frac{y-k_2-k_3-p_1\cdot(B_{13}\cdot a_3-a_1)-p_2\cdot(B_{23}\cdot a_3-a_2)}{p_1\cdot B_{13}+p_2\cdot B_{23}+p_3}$$