ISBN: 978-82-575-1453-2
ISSN: 1894-6402

Norwegian University
of Life Sciences

Postboks 5003
NO-1432 Ås, Norway
+47 67 23 00 00
www.nmbu.no

95807 / ANDVORDGRAFISK.NO

Philosophiae Doctor (PhD), Thesis 2017:54

Arnar Mar Buason

Norwegian University of Life Sciences
Faculty: School of Economics and Business

Philosophiae Doctor (PhD)
Thesis 2017:54

Time and Money: A study of Purchasing Decisions

Tid og Penger: En Studie av Innkjøpsbeslutninger

Arnar Mar Buason

Time and Money: A Study of Purchasing Decisions

Tid og Penger: En Studie av Innkjøpsbeslutninger

Philosophiae Doctor (PhD) Thesis

Arnar Mar Buason

Norwegian University of Life Sciences
School of Economics and Business

Ås (2017)

Thesis number 2017:54
ISSN 1894-6402
ISBN 978-82-575-1453-2

Acknowledgements

I am very grateful for the unwavering support of my supervisors Kyrre Rickertsen and Dadi Kristofersson who provided excellent supervision and constructive feedback throughout the whole PhD period. They dedicated countless hours to this task where they contributed significantly to the thesis and my progress of becoming an academic. They are also co-authors on two of the five papers included in the thesis.

I would like to thank Sveinn Agnarsson who is a co-author on two of the five papers and provided significant contributions to both papers. He furthermore provided me with additional funding to conclude my thesis and gave me the chance to travel all over Europe for conferences and meetings where I had the chance to meet other researchers and receive constructive feedback on my work from my pears around the world. I would also like to thank my other two co-authors Kristin Eiriksdottir and Audur Hermannsdottir who are co-authors on one paper each and provided significant contributions to the papers we worked together on.

Finally, I would like to thank my wife, Adalheidur Osk Gudlaugsdottir, who has supported me through all the good times and all the tough times during the whole PhD process. This would not have been possible without her.

Content

List of Papers

The thesis contains the following papers:

Paper 1: **Demand Systems and Frequency of Purchase Models**
Co-authors: Dadi Kristofersson and Kyrre Rickertsen

Paper 2: **Frequency and Time in Recreational Demand**
Co-authors: Kristin Eiriksdottir, Dadi Kristofersson, and Kyrre Rickertsen

Paper 3: **Habits in Frequency of Purchase Models**

Paper 4: **Fond of Fish? A Count Data Analysis of How Frequently French Consumers Purchase Seafood**
Co-author: Sveinn Agnarsson

Paper 5: **How often, how much? Analysis of Consumption of Label Rouge Salmon in France**
Co-authors: Audur Hermannsdottir and Sveinn Agnarsson

Abstract

A consumer does not just choose how much to purchase of each good on a shopping trip, but also how frequently he shops during a certain period of time. The first paper introduces a microeconomic model, which accounts for purchase frequencies. This is done by incorporating how often a consumer goes shopping into a utility maximization model. The total quantities purchased are given as the product of the purchase frequency and the average quantity purchased on each shopping trip. The model thus results in three demand systems; a system for frequencies of purchase, a system for average purchased quantities, and a system for total purchases. The paper also presents and proofs that the fundamental properties of demand systems hold for the derived systems, i.e., homogeneity, symmetry, and negativity.

In the second paper, the microeconomic model of purchase frequencies of the first paper is extended to a travel cost model. In this model, the consumer chooses how often to visit a recreational site and how much time to spend on site on an average visit. The corresponding willingness-to-pay (WTP) estimates are derived.

The third paper introduces habits into a theoretical model that resembles the model in the first paper. Habits are separated into habits related to shopping frequencies and habits related to average purchased quantities. The suggested model can be used to show to what extent habits in purchased quantities are mainly driven by purchasing frequencies or average purchases on each trip to the store.

The fourth paper focuses on consumer heterogeneity and analyzes whether distinct groups of consumers purchase different types of fish. This is done by estimating a system of purchase frequencies for different types of fish. The frequencies are derived from a model, which is similar to the models that are presented in papers 1-3.

In the fifth paper, we introduce a theory-consistent estimation method of a flexible infrequency of purchase model. In this model, the actual purchase frequencies are modelled instead of only considering whether a consumer buys a product or not. The model is used to compare consumer perception and loyalty to fresh salmon bearing the Label Rouge label and non-labelled fresh salmon.

An econometric model for estimating the demand systems derived in papers 1, 2, and 3 is also developed. This model is a hurdle type model. In the first step, consumers decide how often to go shopping and in the second step they decide how much to purchase on an average shopping trip. This model extends conventional hurdle models where consumers choose only whether to buy a good or not in the first step. The functional form of the estimated demand systems is restricted such that the product of the two steps results in a demand system for total purchased quantities. In the system for total quantities, the properties of homogeneity, symmetry, and negativity are maintained. It is assumed that the data generating process (DGP) for frequencies is given by a count data distribution and the DGP for average quantities is given by a continuous distribution, which is only defined over positive values. To allow for a non-restricted covariance structure between demand equations within each demand system, multivariate random effects are introduced. This results in likelihood functions, which contain integrals that cannot be solved analytically. To solve these integrals both Gaussian-Quadrature and simulation based methods are used.

The empirical applications are based on three data sets: a scanner data set for fresh fish consumption by French households during the years 2005-2008, a scanner data set for fresh fish consumption by French households during the years 2010-2013, and a stated preference data set from on-site sampling in an urban park in Iceland, Heiðmörk. Two different scanner data sets are used because only the first data set was available for the work with papers 1 and 2. The newer data set was used in papers 4 and 5.

There are five key empirical results of the thesis. First, it is shown that the demand for frequencies of purchase and average purchases can be used to formulate profitable loss-leader marketing strategies using fresh white fish as a loss-leader product group. Second, it is shown that WTP in conventional travel cost models produce significantly lower estimates as compared to the estimates provided by the model introduced in this thesis. Third, it is shown that habits of fresh fish purchases are mainly driven by habits in shopping frequencies, while price effects on total purchased quantities mainly originate from effects on average quantities purchased. Fourth, consumers have quite heterogeneous preferences for purchasing different types of fish. As an example, the typical consumer of fresh salmon is a healthy upper-class individual with university education, who comes from a small household in Paris or the north of France. On the other hand, the typical consumer of frozen white fish is an older, lower, middle to lower class individual, who comes from a large household in the south France. Fifth, consumers' perceptions and loyalty differ substantially between fresh salmon bearing the Label Rouge label and non-labelled salmon, which demonstrates that the Label Rouge is effective for product differentiation.

Sammendrag

En forbruker velger ikke bare hvor mye han skal kjøpe av en vare på en tur til butikken, men også hvor ofte han skal handle varen i løpet av en viss tidsperiode. Den første artikkelen introduserer en mikroøkonomisk modell, som inkluderer kjøpsfrekvenser. Dette gjøres ved å inkorporere hvor ofte en forbruker handler inn i en modell for nyttemaksimering. Den totale mengden som blir kjøpt er gitt som produktet av kjøpsfrekvensen og den gjennomsnittlige mengden som blir kjøpt på hver shoppingtur. Modellen gir derfor tre etterspørselssystemer; et system for innkjøpsfrekvenser, et system for gjennomsnittlig kjøpte mengder og et system for totale innkjøp. Det blir også bevist at de grunnleggende egenskapene til etterspørselssystemer holder i de avledede systemene, dvs. homogenitet, symmetri og negativitet.

I den andre artikkelen blir den mikroøkonomiske modellen for innkjøpsfrekvenser, som ble utledet i den første artikkelen, brukt i en reisekostnadsmodell hvor forbrukeren velger hvor ofte han skal besøke et rekreasjonssted og hvor mye tid han i gjennomsnitt skal bruke på stedet. Betalingsvilligheten for et besøk blir videre estimert.

Den tredje artikkelen introduserer vaner i en teoretisk modell som ligner modellen i den første artikkelen. Vaner er delt i vaner som er relatert til innkjøpsfrekvens og vaner som er relatert til gjennomsnittlige innkjøpte mengder. Den foreslåtte modellen brukes til å vise i hvor stor grad vaner i totale innkjøp kan tilskrives henholdsvis vaner i innkjøpsfrekvenser og vaner i gjennomsnittlige innkjøpte mengder på hver tur til butikken.

I den fjerde artikkelen analyserer vi hvilke forbrukergrupper som kjøper ulike typer av fisk. Dette gjøres ved å estimere innkjøpsfrekvensene i et system av forskjellige typer av fisk. Frekvensene er utledet i en modell som ligner på modellene i de tre første artiklene.

I den femte artikkelen introduserer vi en teori-konsistent estimeringsmetode for en fleksibel frekvensmodell. Her modelleres de faktiske innkjøpsfrekvensene, og ikke bare

hvorvidt en forbruker kjøper en vare eller ikke. Modellen brukes til å sammenligne forbrukernes oppfatning og lojalitet for fersk laks merket som "Label Rouge" og umerket fersk laks.

En økonometrisk modell for estimering av etterspørselssystemene utledet i de tre første artiklene blir også utviklet. Denne modellen er en såkalt «hurdle» modell. I det første trinnet bestemmer forbrukerne hvor ofte man skal handle. I det andre trinnet bestemmer forbrukerne hvor mye man skal kjøpes i gjennomsnitt på hver handletur. Denne modellen utvider konvensjonelle «hurdle» modeller hvor forbrukerne bare velger hvorvidt de skal kjøpe en vare eller ikke i det første trinnet. Den funksjonelle formen for de estimerte etterspørselssystemene er slik at produktet av de to trinnene resulterer i et etterspørselssystem for de totalt innkjøpte mengdene. I systemet for totale mengder holder egenskapene homogenitet, symmetri og negativitet. Det antas at den data genererende prosessen for frekvensene er gitt ved en fordeling for telledata og at prosessen for gjennomsnittlige mengder er gitt ved en kontinuerlig fordeling som bare er definert for positive innkjøp. For å tillate en ubegrenset kovariansstruktur mellom etterspørselslikningene i hvert etterspørselssystem innfører vi multivariate tilfeldige effekter. Resultatet er sannsynlighetsfunksjoner som inneholder integraler og ikke kan løses analytisk. For å løse disse integralene benyttes både Gaussisk-kvadratur og simuleringsbaserte metoder.

De empiriske eksemplene er basert på tre datasett; et skannerdatasett for fiskeforbruk i franske husholdninger for årene 2005-2008, et skannerdatasett for fiskeforbruk i franske husholdninger for årene 2010-2013 og et datasett basert på survey data blant brukere av friluftsområdet Heiðmörk i Reykjavik. To forskjellige skannerdatasett ble brukt fordi bare det eldste datasettet var tilgjengelig under arbeidet med de to første artiklene. Det nyere datasettet ble brukt i de to siste artiklene.

Avhandlingen inneholder fem viktige empiriske resultater. For det første blir det vist at de utviklete metodene kan brukes til å formulere lønnsomme markedsstrategier basert på «loss-leader» modellen. Her brukes hvitfisk som en «loss-leader» produktgruppe. For det andre blir det vist at estimater for betalingsvillighet i den vanlig brukte reisekostnadsmodellen gir betydelig lavere verdier enn dem vi finner ved å bruke den nye modellen. For det tredje blir det vist at vaner i forhold til kjøp av fersk fisk hovedsakelig er drevet av vaner i innkjøpsfrekvenser, mens priseffekter hovedsakelig kommer fra effekter på de gjennomsnittlige innkjøp på hver tur til butikken. For det fjerde blir det vist at forbrukerne har heterogene preferanser når det gjelder kjøp av fisk. Den typiske forbrukeren av fersk laks er et sunt individ med overklassebakgrunn og universitetsutdanning, som kommer fra en liten husholdning i Paris eller Nord-Frankrike. På den annen side er den typiske forbrukeren av frossen hvit fisk en eldre person som tilhører den lavere del av middelklassen og kommer fra en stor husholdning i Sør-Frankrike. For det femte så er forbrukernes oppfatninger og lojalitet vesentlig forskjellig for fersk laks merket med «Label Rouge» etiketten og umerket laks, noe som viser at «Label Rouge» er et effektivt virkemiddel for produktdifferensiering.

Introduction and Summary

The thesis consists of five papers and an introduction. Four papers are co-authored and one paper is single authored. The thesis focuses on demonstrating the importance of purchase- and travel frequency in demand analysis and developing its theoretical foundations in microeconomic and econometric theory. The empirical analysis is carried out using three separate data sets; a scanner data set for fresh fish consumption by French households during the years 2005-2008, a scanner data set for fresh fish consumption by French households during the years 2010-2013[1][2], and a stated preference data set from on-site sampling in an urban park in Iceland, Heiðmörk.

The first paper develops a theoretical model where a consumer does not only decide how much to purchase of a good when he goes to the store, but also how often to go shopping. The model is designed for everyday and storable consumer goods. Furthermore, a theory-consistent econometric method is presented to estimate the resulting demand functions. Finally, an empirical illustration is provided by using the French scanner data for purchases of fish for the years 2005-2008, and it is explained how the estimation results can be used to form sophisticated loss-leader pricing strategies.

The second paper introduces a travel cost model where time spent on site is endogenous. The corresponding welfare estimates are also derived. Furthermore, a theory-consistent method is developed for the estimation of the demand for frequency of travel and time spent on site. Finally, the model is estimated and the results are compared with the standard travel cost model, using the stated preference data. The results show that the

[1] The data type has received the name scanner data since it is collected by a hand-held scanner. Households included in the data collection process receive a hand-held scanner along with other relevant equipment to record purchases from home.

[2] Two scanner data sets are used since the writing process started out with the older data set and the second data set was received later. This is not believed to be a significant problem since the articles who use the older data set only use empirical examples as demonstrations.

standard model produces a significant downward bias of the welfare estimates compared with the model that explisitly accounts for time spent on site.

The third paper presents a theoretical model, valid for every day consumer goods and storable goods, where the consumer chooses how much service stock to hold of a good in each time period, where underlying these decisions are the choices of how much to purchase of goods and the frequency of shopping. These underlying decisions allow for a separation of the effects of habits, into habits related to the frequency of shopping and habits related to average quantities purchased. These two habit effects can then be combined to find effect of habits on total quantities purchased. An empirical demonstration is included, using the same French scanner data as in the first paper. The results show that for fresh fish in France the effects of habits on purchased quantities almost solely originate from the habits of shopping frequencies and not from habits related to average purchased quantities. The price effects on total purchased quantities mainly originate from price effects on average quantities purchased and not on the frequencies of purchase.

The fourth paper combines conventional demand analysis and the marketing approach of analyzing purchase frequencies in order to determine the characteristics of consumers buying different types of fish. The estimation uses French scanner data from 2010-2013, and the results show that consumers purchasing different types of fish are vastly heterogeneous. Furthermore, we present a possible way of using the results to increase store traffic and profits.

The fifth paper explores whether the well-known food quality label “the Label Rouge” has been able to differentiate between high-quality Label Rouge and other salmon. The results are consistent with previous results and show that the consumers of Label Rouge salmon are significantly more loyal than the average consumer of standard fresh salmon, and the Label Rouge salmon is less sensitive to price changes. Therefore, the label has produced the desired effect.

Motivation and Background

Demand analysis typically aims at determining how consumers change their consumption in response to changes in relative prices, income, sociodemographics, and other variables. The literature has a long history, starting in the 1950s with the work of J.R.N. Stone, see for example Stone (1954) where a linear expenditure system is applied to the pattern of British demand. Following Stone multiple articles were written on the subject, such as Barten (1969), Jorgenson and Lau (1975), and Lau (1976). Then after the publication of the seminal paper of Deaton and Muellbauer (1980) where they derived the almost ideal demand system (AIDS) countless articles have estimated this system in order to do predictions, estimate welfare effects, and doing other analyses, see for example; Xie and Myrland, 2011, Allais et al. (2010), Bertail, and Caillavet (2008), West and Williams (2004), Gustavsen and Rickertsen (2003), Larivière et al. (2000), Rickertsen (1998), Rickertsen (1996), and Rickertsen and Chalfant (1995). The decision being analyzed in these models, depending on how it is formulated, are essentially how much to spend on each good or how much to buy, in terms of total expenditure shares. However, these decisions account for only a part of a consumers' decision making process when it comes to purchasing disaggregate goods and therefore only provide some of the information desired. For disaggregate goods, the consumer also has to decide whether to buy a good or not, and given a positive purchase[3] - how often to go shopping.

[3] See for example, Tobin (1958), Cragg (1971), Heckman (1974), Lee and Pitt (1987), Wales and Woodland (1983), and Phaneuf et al. (2000)

In demand analysis, most data is aggregated over time. This might leave the researcher with monthly data of households which, for example, purchase an average of one kg of fresh fish each month. In that situation, it would significantly increase the information available, for example, for retailers if they also knew that the average household in the data purchases fresh fish four times each month, especially when the retailer needs to know the size of inventory to hold at each point in time. In such situations, the decision makers are not only concerned with how purchased quantities change in response to price or income changes but also how the shopping frequencies of different goods change in response to same variables.

Data comes in different forms and does not necessarily include all the information the researcher desires. Household expenditure data can be aggregated over months and could not include any information at all regarding how many shopping trips were made during the whole period, relevant household characteristics might be missing, products may be aggregated, or the data is usually not panel data. However, good detailed data such as household scanner data has become increasingly more available to researchers doing micro-econometric work in recent years, due to companies such as Kantar Worldpanel, previously TNS Worldpanel[4], who collect high quality micro panel data. The data is collected through the use of bar code scanners, which Kantar provides to all participating households to record all purchases of certain group of goods for up to four years. Each panel includes weekly purchase information for up to 20,000 households and their sociodemographic characteristics. This type of data opens up for many possibilities, including the analysis of how often households purchase their goods. Such analysis can improve both the accuracy of conventional demand analysis, which will be discussed later, and provide essential information for marketing decisions.

[4]For an example of an article using data from this firm see Allais et al. (2010).

Retailers are not only interested in how purchases change on average in response to prices or income changes, but also many other factors, which have been thoroughly investigated in the marketing literature. Retailers are, for example, often concerned with measuring store choice, store loyalty, store traffic, product choice, etc., to make informed marketing decisions, see for example Keng and Ehrenberg (1984), Uncles et al. (1995), Bhattacharya (1997) and Uncles and Lee (2006).[5] Keng and Ehrenber (1984) analyzed store loyalty in the U.K., for a few products, using a negative binomial model for purchase frequency and a multivariate beta for brand choice. The results show that store loyalty in the U.K. is very low. Uncles and Lee (2006) used stated preference data to estimate purchase frequencies for a few different products to measure the importance of older consumers. These examples further demonstrate the importance of analyzing more than just purchase volume or expenditure shares. Incorporating purchase frequency into the framework of demand systems is an important step towards deeper understanding of consumer behavior. When consumers are loyal to a brand or a particular good it could be reflected by low price elasticity with respect to purchase frequency but high price elasticity with respect to average quantities purchased, and not the other way around. Thus, it is of considerable theoretical and empirical interest for the development of sophisticated pricing strategies to extract as much of the valuable information in the data as possible.[6]

Demand analysis in the economics literature does not focus on understanding consumer purchase frequency. However, a set of models have been developed by Deaton and Irish (1984), Kay et al. (1984), Pudney (1985, 1986), Blundell and Meghir (1987), Meghir and Robin (1992) and Robin (1993), which are referred to as infrequency of purchase models

[5]Multiple of these issues have also been thoroughly studied in the economics literature. For a discussion of labelling and habits of product choices see for example Teisl et al. (2002) and Dhar and Foltz (2005), Adamowicz and Swait (2012).

[6] Here we define sophisticated pricing strategies as multiple price increases and / or decreases with the intention of increasing profits.

(IPM). The literature developed from the idea of recognizing that zero purchases observed in the data set are not only generated through corner solutions from utility maximizing behavior, as was unrealistically assumed in previous articles, for example, Tobin (1958), Cragg (1971), Heckman (1974), Lee and Pitt (1987), Wales and Woodland (1983), and Phaneuf et al. (2000), but also from purchase fluctuations over time and non-preference for a good. In Meghir and Robin (1992), the standard two-part model was extended by including the actual purchase frequency and not just whether a consumer buys a good or not. They use predicted purchase frequencies, from a Poisson model, to adjust an AIDS model with the probability of purchases, instead of the results from a logit or probit model. Robin (1993) then extended the use of this model to predict the share of consumer who will not purchase the good under any circumstances. Furthermore, his article shows that adjusting the model with the actual purchase frequency significantly improves the precision of the estimates. However, none of these articles really explored the microeconomic foundations of purchase frequency. Meghir and Robin (1992) presented a micro-economic model of a forward-looking consumer who chooses both how much to buy in terms of quantities and how often to purchase these goods. However, the problem is only presented and not derived in any detail. Without any formal model, it is unclear what theoretical restrictions should be implemented and how the econometric model should be formulated. Thus, devloping the idea of Meghir and Robin (1992) of separating the choice of quantity demanded into a formal model with the two decisions of how much to purchase and how often to go shopping. This formulation then provides a first step towards a satisfactory treatment of consumer purchasing behavior.

The analysis of the frequency of use is also important in recreational demand and health care. Examples from the recreational demand literature are Creel and Loomis (1990), Egan and Herriges (2006), and Hynes and Greene (2013). However, in these applications average time spent on site is assumed to be constant. Although this approach may suffice for some application it is unlikely to hold when time spent on site varies significantly. This is usually the case of urban parks, which often provide a wide range of activities including playgrounds for children; picnic areas; trails for hiking, running and walking; restrooms and sports facilities (McCormack et al. 2010, Konijnendijk et al. 2013). A few articles have allowed for endogenous time spent on site, for example, McConnell (1992), Larson (1993), and Hellström (2006). However, the models introduced in these articles do not allow for a satisfactory formulation of welfare estimates in the case of endogenous on site time. Using the above suggested extension to the IPMs, where quantity demanded is separeteds into average quantity demanded and frequency demanded, can also be applied to recreational demand. In this case, the total time spent on site is separated into the choice of frequency of trips and average time spent on site. From this formulation, it is then possible to derive welfare estimates which account for both time and travel frequency.

Count data models have also been used in health economics, for example, Deb and Trivedi (2002) who estimated the number of doctor visits, and Munkin and Trivedi (1999) and Wang (2003) who estimated the demand for health care. However, these models do not account for the expenditures corresponding to each visit. Applying the same framework, as suggest in the IPMs and travel cost models, would provide a more complete representation of the consumer's choices regarding health care. Then, total expenditures on healthcare would be separated into the two decisions: (i) how often to go to the doctor and (ii) how much to spend on average in each trip, that is in cases where the individual has some choice of treatment and tests. The second decision is an actual choice in countries such as the U.S. where healthcare is not fully provided by the government. However, these issues in health economics will not be pursued in this thesis.

The Thesis

The main objective of this thesis is to introduce a more theoretically appropriate treatment of how the choice of frequency of purchase can be incorporated into a microeconomic model of utility maximizing behavior, as well as into the corresponding econometric models. It does so by formulating and deriving microeconomic models of utility maximizing behavior which do not only include consumption as a choice variable, but also divide the consumption decision into (i) how often to purchase the goods and (ii) how much to purchase on average in each shopping trip. The model proposes that consumption is given by an identity, which is the product of purchase frequency and average quantity. The model includes a non-linear budget / time constraint where the resulting demand systems still maintain the microeconomic restrictions of homogeneity of degree zero, symmetry, and negativity.

A corresponding econometric framework is developed which allows for the joint estimation of these demand systems while incorporating the theoretical restrictions from microeconomic theory. The econometric model is a hurdle model which replaces the conventional first stage of binary outcomes with a count data model. The probability of passing this hurdle is thus given by predictions from the count data model. The second quantity stage is only observed when the frequency is at least equal to one. The quantities are therefore only defined over positive outcomes.

Previous research on purchase frequencies in microeconomic and econometric models is limited. The key references are Meghir and Robin (1992) and Robin (1993), who introduced a microeconomic model including the frequency of purchases. However, their model was not formally derived.

The main objectives of each of the five papers are:

- Paper 1: To specify a utility maximization problem, which incorporates how often consumers purchase goods and derive the associated demand systems. An econometric model, which jointly estimates the frequency of shopping trips and average quantity demanded on each trip is presented and estimated using Bayesian methods. The estimates are used to show how the developed framework can be used for the formulation of profitable loss-leader marketing strategies.
- Paper 2: To develop a travel cost model where the consumer jointly chooses the number of visits to a recreational site and how much time to spend at the site. The model is essentially the same as in Paper 1 except that it is fit to the framework of travel cost, and therefore the consumer chooses how often to visit a recreational site and how much time to spend their, instead of how often a consumer goes shopping and how much he purchases of each good in each trip. There are also minor differences in the specification of the time endowment. An econometric model, which jointly estimates the two decisions is presented and estimated. This model is essentially the same econometric model as in Paper 1 except that it is not estimated byusing Bayesian methods. Finally, the welfare estimates of the presented model are compared with the estimates of a standard travel cost model.
- Paper 3: To develop a microeconomic model, which incorporates habits to explain the dynamics of consumption of everyday consumer goods and storable goods. The underlying choice of service stock consists of two choices: (i) how often to purchase a good and (ii) how much to purchase in each shopping trip. An econometric model where both these decisions can be estimated is introduced. As in Paper 1, Bayesian estimation is used to be able to allow for non-restricted covariance matrix between demand equations. The results are used to discuss how habits in the demand for fresh fish in total purchased quantities can be separated into habits that originate in

frequencies of purchase and habits in average quantities purchased. Then the information is used to show how to formulate profitable pricing strategies.

- Paper 4: To explore which French consumers, who purchase different categories of fish in order to locate which consumer groups to target with marketing strategies. The microeconomic model used in Paper 4 is a simplified version of the micro-model in Paper 1, and it is not assumed that total quantity purchased can be separated into the two decisions of how often to purchase and how much to purchase on average. However, the consumer still faces the decision of how often to purchase the good. This simplification is useful for this paper, which focuses on the demand for shopping trips. The econometric model is different from the models in the other papers, and we use a negative binomial model to represent the data generating process of shopping trips.
- Paper 5. To compare demand for fresh Label Rouge salmon and other fresh salmon in France to investigate if this label can produce the desired product differentiation and increased loyalty. The microeconomic model used in this article is the same as in Paper 4. The econometric model is a Infrequency of Purchase model (IPM), which has two stages. First, a count data model is estimated and from this model the probabilities are calculated, which are used to adjust a LAIDS model where the error term is assumed to be normally distributed.

Paper 1: Demand Systems and Frequency of Purchase Models *(Co-authored with Kyrre Rickertsen and Dadi Kristofersson)*

A well-known marketing strategy is the loss-leader strategy, which is based on the consumers limited information regarding market prices. One or several products are priced and advertised to lure consumers to go shopping in a store, but then a range of other products in the store are priced marginally above the market price observed in other stores and this leads to an increased profit for the store. This strategy might work given that the travel cost and the opportunity cost of time is sufficiently high and the consumer possesses limited information regarding the market prices of different goods in different stores.

To locate good loss-leader products, knowledge of consumers' demand for shopping trips as well as the average quantities purchased on each shopping trip are important. For example, a product with a relatively high own-price purchase-frequency elasticity may bring many consumers to the store. The product will perform even better as a loss leader if it has a relatively low own-price elasticity in terms of average quantity. Good candidates for price increases have a high and positive cross-price elasticity with the other products who have their price lowered.

The paper presents a microeconomic model where the choice of quantity demanded is separated into how often to purchase each good and how much to purchase of each good on an average shopping trip. The paper derives a demand system where homogeneity of degree zero holds and the matrix of compensated substitution effects is symmetric and negative semi-definite for the total quantity demanded. Furthermore, the demand for shopping trips and the demand for average quantities purchased are also homogeneous of degree zero in prices and income, but not necessarily symmetric as long as the total effects are symmetric.

An econometric model is introduced to estimate the frequency of purchase and average quantity purchased in such a way that the results of these two demand systems can be multiplied to derive a demand system for total quantities purchased of each good. Homogeneity, symmetry and negativity also hold for the estimated system. The method is based on a two-step procedure where the first step determines the frequencies of purchase as generated from a count data model. The second step models the average purchased quantities given that a purchase takes place. The model accounts for zero inflation[7] generally observed in microdata by imposing truncation on the frequencies. The data generating process for the frequency and average quantity parts of the model are a multivariate Poisson log-normal distribution and a multivariate gamma log-normal distribution, respectively.

The method is applied to French scanner data of fresh fish purchases from 2005-2008. The data was collected by TNS Worldpanel. The data contains weekly information regarding purchases of fresh fish. The number of participants recording fresh fish purchases were 3291 in 2005, 3234 in 2006, 3165 in 2007, and 4479 in 2008. To reduce the number of zero observation the weekly purchases were aggregated to yearly purchases. Fresh salmon is the most popular fish type in France and was included in the demand system. The other included products were fresh white fish, which represent the most popular wild fish such as cod, saithe, whiting etc., and a group of other fish types.

A numerical example is included, which shows that fresh white fish is a good product group for a loss-leader strategy. The example shows that large retailers can marginally lower the price of fresh white fish and simultaneously increase the price fresh salmon and make a profit by increasing the purchase frequency of other goods. This strategy works best on infrequent consumers who have little information regarding market prices in different stores. The the strategy may also be used for driving smaller fish sellers out of the market by pricing a large fraction of their product below marginal cost.

Paper 2: Frequency and Time in Recreational Demand *(Co-authored with Kristin Eiriksdottir, Kyrre Rickertsen, and Dadi Kristofersson)*

[7] When the share of zeros in a data set significantly exeeds the share predicted by conventional distributions, such as the Poisson or negative binomial, it is called zero inflation.

Frequently it is assumed that time spent on site is constant while travel cost is allowed to vary between sites and individuals in the recreational demand literature. This is a perfectly acceptable assumption in cases where travel cost varies substantially, but time spent on site has little variation. This might be the case for national parks where people travel from different locations and therefore have significantly different travel costs, but might on average spend a similar amount of time at the location once they are there, for example, one whole day. However, in the case of open access urban parks, such as Central Park in New York and Boston Common in Boston, the travel cost is close to zero relative to the cost of time spent on site, and the variation in time spent on site might be substantial. One individual could be there for jogging for a half an hour and another for a picnic for three hours. In these situations, the assumption of a constant time spent on site may be rather farfetched.

This article introduces a microeconomic model where an individual chooses how often to travel to a recreational site as well as how much time to spend there on average. From this specification willingness to pay estimates, that account for endogenous on-site time, are derived. These estimates are consistent with microeconomic theory, through the homogeneity restrictions imposed. Symmetry conditions are not needed since the estimated model is not a demand system.

Furthermore, an econometric model, which jointly estimates the demand for recreational trips and their respective on site time is presented. In this hurdle model, time spent on site is only observed when a trip takes place. The data generating process for recreational trips is assumed to follow a Poisson distribution, and a gamma distribution is assumed for time spent on site.[8] The two stages of the model are likely to be stochastically correlated, and the model accounts for this correlation by having random effects in both

[8] Since the frequency part of recreational demand takes the form of non-negative discrete integers, it is modelled by a Poisson distribution. Time spent on-site can only take positive non-zero values and although it is observed discretely, it is generated continuously and is therefore modelled with a gamma distribution.

stages of the model, which allows for non-zero covariance between the stages. The likelihood function does not have a closed form solution and therefore a Gaussian quadrature is used. The model is estimated by a Dual Quasi-Newton method.

The model is estimated using data gathered on-site at an urban park in Iceland, Heiðmörk. Willingness to pay estimates are calculated and compared with the estimates from a single site travel cost model. The results show that the conventional model significantly underestimates consumers' willingness to pay, and thus demonstrating the importance of allowing for endogenous on site time.

Paper 3: Habits in Frequency of Purchase Models

In the demand literature, models allowing for habits have been widely applied. However, how habits influence consumers' shopping frequencies for different goods have not been analyzed in the demand system framework. Habits are negatively related to utility since consumers do not respond optimally to changes in relative prices, compared with the standard utility maximization problem without habits. Habits can be exploited by sellers when forming their marketing strategies. Furthermore, another key factor influencing repeated purchases is duration. Duration consists of the biological duration of the product purchased and consumer's saturation, that is how long it takes until the consumer wishes to purchase the product again, i.e., a duration of psychological nature.

The article introduces a microeconomic model where consumers derive utility from the flow of services provided by the consumption of the stock of goods. The reason for this is to introduce dynamics into a static model. Furthermore, the quantity demand underlying the service stock is separated into how often to purchase the good and how much to purchase each time. This separation is done in such a way that the effects of duration and habits can be separately identified, which allows for the analysis of the origin of these effects, that is to say whether the effect comes mainly from the habits of shopping frequency or purchased quantities.

An econometric model is proposed where the demand for shopping trips and average quantity demanded is estimated jointly, using Bayesian methods. The data generating process for the frequencies of shopping is assumed to follow a multivariate Poisson log-normal distribution and the average frequencies are assumed to follow a multivariate gamma log-normal distribution, as in Paper 1. The mixture distributions are introduced to take account of the panel structure of our data set and to allow for a non-restricted covariance matrix between demand equations. The loglikelihood function for these complicated distributions does not have a closed form solution and for a system larger than two equations a simple Gaussian quadrature is not desirable due to the course of dimensionality and therefore the simulation methods provided by the Bayesian framework are used for our estimation.

An empirical example is provided using French scanner data of fresh fish purchases. The demand system includes three types of fish: wild, farmed, and fish produced with unknown technology. The reason for this commodity specification is that we use monthly data for the analysis and categories such as fresh salmon or fresh cod simply provide too many zeros for the econometric model in this paper. Moreover, the comparison between farmed and wild fish is important for many consumers and has been widely analyzed, see for example Hermann et al. (1993), Asche et al. (2005), and Asche and Guttormsen (2014).

Our results show that net habits of total purchased quantities almost solely originate from habits in shopping frequencies, and the price effects on total purchases mainly originate from the effects on average quantities purchased. Finally, we show that if the average per kilo price of wild and farmed fish is decreased in a particular way it is possible to make a profit from sales of other products, due to increased purchase frequency, which leads to a net increase in revenue.

Paper 4: Fond of Fish? A Count Data Analysis of How Frequently French Consumers Purchase Seafood *(Co-authored with Sveinn Agnarsson)*

The EU is one of the largest producers of seafood in the world and the largest fish trader in terms of value. The popularity of salmon, in terms of quantity sold, in the EU has been increasing over time and the imports of salmon have continued to increase in response, where the largest exporter of seafood to the EU is Norway. Other important exporters of seafood to the EU are China, Ecuador, Morocco, United States, Vietnam, and Iceland.

Among the EU countries, France is one of the largest consumer markets for seafood and the largest for salmon. France is now a net importer of fish after the fisheries have slowly been declining. This decline has been offset by imports but the market has also been moving from white fish to salmon. A great range of seafood products are available in the French market and salmon and cod are the most popular. Furthermore, the market constantly moves from frozen, salted, or dried fish towards more valuable fresh seafood.

This article analyzes the French fish market and specifically try to understand which consumers who purchase which products, and how often and how much they purchase. This objective is achieved by combining the economic demand and marketing literature. The demand literature focuses on prediction and the changes in demand as a result of relative price changes for different products and changes in income. The dependent variable is either quantity, expenditure, or budget share. The marketing literature focuses on whether consumers purchase or not and how often they go shopping. Such predictions are then used to shed light on product loyalty, store traffic, etc.

The econometric model used in this paper is a negative binomial model, which is assumed to be the data generating process for the frequency of shopping trips. In this paper, we only estimate the demand for shopping trips and not the demand for purchased quantities, since it is not necessary to analyse which groups of consumer purchase which fish types. The econometric model is thus only related to the econometric models used in papers 1-3 by being a model of frequencies.

An empirical example is included where the model is estimated using French scanner data of fish purchases from 2010-2013. More recent data is used in this paper than the previous papers because it was written after these papers and new data were available. The data contains a large variety of socioeconomic variables, everything from income to the number of cats in the household. The demand system specified consists of five categories; fresh salmon, frozen Salmonidae, fresh cod, frozen cod, and all other fish. Where the category Salmonide includes both trout and salmon. In this paper, a larger system of demand equations is used to better demonstrate which consumers purchase which types of fish. An extension to this paper could be to include even more categories of fish. The results show that consumers purchasing different types of fish are vastly heterogeneous.

Paper 5: How Often, How Much? Analysis of Consumption of Label Rouge Salmon in France *(Co-authored with Audur Hermannsdottir and Sveinn Agnarsson)*

France is among the most important seafood markets in Europe and has been progressing towards higher quality fresh fish, where the most important product is fresh salmon. One of the best-known quality label for food in France is the Label Rouge. The label ensures that the product and its production process satisfies the strictest criteria of quality. Labeling can therefore inform consumer, who wish to buy good products, of product quality. Among the aims of the label is to ensure certain quality attributes and therefore differentiate this high quality products from other types of fresh salmon. Previous research has shown that those consumers who purchase Label Rouge are willing to pay a higher price and are more loyal than others, see for example Monfort (2006). Thus, demand for Label Rouge products is rather stable over time compared to the demand other unlabeled products due to the consistent quality of the products.

The article introduces a theory-consistent way of estimating a flexible infrequency of purchase model. The model has been estimated but not with the restrictions imposed by economic theory. This model is derived from a utility maximization which does not separate total purchased quantity into the product of purchase frequencies and average quantities as in papers 1 and 3, but assumes that frequencies are chosen separately from total purchased quantities as in paper 4. The econometric model is a two-stage model where the first stage is a truncated count data model of purchase frequencies and the second stage is an LAIDS model where the error terms are assumed to be normally distributed. The second stage is adjusted by the probabilities of purchases estimated in the first stage. This model is thus more closely related to a standard hurdle model than the econometric models used in papers 1-3. The data used for the estimation is a French scanner data set of 20,000 households from 2011 to 2013. A more recent data set is used in this paper for the same reason as in Paper 4. The system consists of fresh Rouge Labeled salmon, other fresh salmon, and all other fish. We use a different commodity specification than Paper 4 since here the aim is to compare Rouge Labeled salmon and other salmon. The specification generates a large share of zero observations, as is common in micro data, but the infrequency of purchase model accounts for whether a purchase takes place and also how often, thus utilizing the information in the data as far as possible.

The results show that consumers´ perception and loyalty differs substantially between fresh salmon bearing the Label Rouge label and non-labelled salmon, demonstrating that the Label Rouge is able to produce the desired effect of product differentiation. The loyalty towards the Label Rouge salmon introduces possibilities for sellers to marginally increase prices for a higher profit. However, it is very costly for producers to receive the label and an interesting question for future research is to estimate how profitable it is to acquire the Label Rouge for salmon in the French fish market.

Contributions, Implications and Limitations of the Thesis

The five papers contribute to the demand literature, specifically to the frequency of purchase and travel cost literature. The results and implications of the thesis must be viewed in light of its limitations. The first four papers combine shopping frequencies or the frequency of recreational activities with the average quantity purchased or average time spent at a recreational site into a utility maximizing framework. This formulation aims at providing a more realistic approach to consumer choice. By adding the choice of purchase frequency and the separation of total quantities purchased into purchase frequenceies and average quantities not only introduces a more realistic framework of purchaseing decisions, but also results in two additional demand systems which can be used to further understand the purchaseing behavior of consumers. As has been discussed above, the additional information enables the formulation of more sophisticated pricing strategies which not only aim at changing price of one or more good to increase quantites purchased but also to utilize differences in elasticities in purchase frequency and average quantity in order to increase quantities purchased of specific goods and shopping frequencies and thus increase quantites purchased of an even wider range of products. In the case of recreational demand, allowing for endogenous time spent on site not only introduces a more realistic framework of consumer decisions, but also leads to improved estimates of willingness to pay for recreational activities. However, one potential weakneses of the model is that it does not look at how much time is spent on site on each trip separately or how much is purchased on each shopping occatition separately but only on averages across trips. In many cases this is a simplification of reality. Paper 4 and 5 have a slightly different focus and assume that the quantity decision and the frequency decision are not directly related through the identity described above. In this case, the consumer is assumed to maximize utility with respect to frequency, total quantity and leisure, instead of frequencies, average quantity and leisure. This approach builds on other assumptions concerning consumer behavior. The reason for using this approach in Paper 5 instead of the one used in Papers 1-3 is to allow for a more flexible demand specification, i.e. LAIDS, and since modeling the frequencies of purchase are not the main focus of this paper we deem this model as more suitable to achieve our goal. One limitation is that we use a LAIDS spefcification instead of the non-linear AIDS model. The reason is that the model is already highly non-linear and by introducing a new source of non-linearity would significantly incease computational difficulties. In Paper 4 the aim was to analyse the differences in groups of different fish types by looking at their shopping behavior and therefore the approach in Papers 1-3 was not necessary. However, both methods provide utility maximization which accounts for frequencies and provide an improvement over the conventional hurdle models.

The first three papers develop variations of econometric models to estimate the demand systems derived from the aforementioned microeconomic models for different situations, such as purchase frequency models with and without habit formation, and travel cost models. The econometric model in all three papers is a hurdle model with two stages, where the first stage determines the frequency of purchase and the second stage determines how much the consumer purchases given that at least one purchase takes place. An optimal model would allow for correlation between stages as well as within demand systems. However, this is computationally troublesome and simulation based methods are needed for this to be feasible. Paper 1 and three allows for stochastic correlation within each system but not between stages. These articles use Bayesian methods to be able to allow for these complicated likelihood functions which do not poses an analytical solution. Article two allows for correlation between stages, but is only estimated for one choice so there is no need for correlation between demand functions within each stage. This article uses Gaussian quadrature for the numerical approximation of the integral in the log-likelihood function, which is usually not feasible for higher dimensionality than two. With constantly increasing computational power these types of models should not pose any significant difficulty in estimation.

As previously discussed, four out of five papers use French scanner data for their applications, but for different years, and paper two uses data gathered on-site at an urban park in Iceland, Heiðmörk. A range of different goals where set for these applications even though the data for four of them was similar. Paper 1 focuses on how the estimated demand for shopping trips and average quantity demanded could be used to form loss-leader pricing strategies, where white fish is found to be a good loss-leader product, that is the loss product for the strategy, but the data set does not include good products to use as leaders, such products should optimally have low own-price elasticity. It would thus be interesting to apply the model to a different dataset. Article three uses the same data set as article one and is an application of habit formation in the case of fish demand in France where the consumption habits are divided between frequencies and quantities purchased and the results show that the most of the habits formation originates from frequencies, but the difference in habit formation between fish categories is close to none. This is a limitation due to the data set which would be interesting to explore with a different data set, for example of meat and fish. Article two uses data gathered on-site at an urban park in Iceland, Heiðmörk, and shows that travel cost models which do not allow for endogenous time spent on site will significantly underestimate willingness to pay in the case of urban parks. This is due to the significant variation in recreational activities which urban parks provide. The main limitation of the analysis is the lack of alternative recreational sites. In the case of the Reykjavik area there is no comparable alternative to Heiðmörk, but it would be interesting to apply the method to different date where a demand system could be estimated. Paper 4 and 5 use French scanner data for years 2010-2013 instead of 2005-2008 as article one and three. Paper 4 estimates a system of demand equations, consistent with economic theory, in terms of purchase frequencies, in the attempt to understand the different groups of consumers who desire different categories of fish. The results show that the consumers that demand different types of fish are vastly different, which results in important information for marketers. The model is however estimated with standard maximum likelihood methods, where it is not feasible to estimate such a system with stochastic correlation. An extension would be to estimate the system with simulation based methods. The fifth and final paper makes a comparison of the demand for fresh salmon in France and Rouge Labeled salmon. The results were consistent with previous results and showed that the consumers of Rouge Labeled salmon are significantly more loyal than the average consumer of standard fresh salmon. The results therefore show that sellers can make some marginal price increase to Rouge labeled salmon and increase profit. There are however some limitations to the estimation. The model is estimated as a twostep procedure which results in a some inefficiency and due to the high nonlinearity of the statistical model the demand system is specified as a linear AIDS model.

In conclusion, even though the data is not always perfect to answered the questions we pose they are more than efficient to show the importance and implications of what this thesis is trying to achieve. The microeconomic and econometric models introduced all have their limitations and could be improved but still provide a significant step towards a more satisfactory treatment of frequencies in microeconomic and econometric demand analysis. The methods introduced are not only for theoretical curiosity but provide important empirical implications as is the case for example of welfare estimates for urban parks and the estimation of loyalty to Label Rouge products. There are however many unanswered questions and important improvements still to be made.

To summarize, the main contributions of the thesis are:

1) A microeconomic model is developed where total purchases of goods are determined by two decisions: (i) the frequencies of purchases, and (ii) the average quantities purchased conditional on positive purchase frequencies.

2) An econometric model is developed to estimate the purchase frequency and average quantity demanded as demand systems, and it is estimated by Bayesian methods to allow for unconstrained covariance within each demand system. The data generating processes of the frequency and average quantity purchased of each good is, respectively, assumed to follow a multivariate Poisson log-normal and a multivariate gamma log-normal distribution.

3) It is shown using the models described under 1) and 2) that it is possible to create profitable loss-leader pricing strategies from results generated by the models 1 and 3.

4) A travel cost model is developed where the consumer jointly chooses the number of visits to a recreational site and how much time to spend there. Corresponding willingness to pay estimates are also derived.

5) The mean and variance of the marginal distribution of the multivariate gamma log-normal distribution is derived.

6) An econometric model is developed which capable of estimating the demand for duration as a two-part model that allows for correlation between the two underlying parts; the decision of how many trips to take and the decision of how much time to spend on-site on each trip. The frequency part is modeled with a Poisson log-normal count model and the length of stay part is modeled with a gamma log-normal model that only allows non-negative values. The likelihood function of this model does not

have a closed form solution and is therefore approximated using a Gauss-Hermite integration, and it is optimized with the numerical DQN method.

7) It is shown that the conventional single site travel cost model produces a significant downward bias of welfare estimates in the case of urban parks, relative to results generated from the models described under 3) and 4).

8) A demand system, which incorporates habits to explain the dynamics of consumption of everyday consumer goods and storable goods, is introduced. In this model, the quantities purchased are modelled as a result of two decisions; how often to purchase each good and how much to purchase in each shopping occasion.

9) We find that habits in total fish purchases in France almost solely originate from habits in purchase frequencies, while habits in average purchased quantities are of minor importance. Contrary to the effects of habits, changes in total purchases in response to price changes is mainly determined by the changes in average quantities purchased, and not frequencies of purchase.

10) It is shown that the average consumer of fresh salmon differs substantially from the average consumer of frozen white fish. The typical consumer of fresh salmon is a healthy upper-class individual with university education who comes from a small household in Paris or the north of France. The average frozen white fish consumer is an older, lower, middle to lower class individual who comes from a large household in the south France.

11) It is shown that average consumers of fresh salmon and fresh salmon bearing the Label Rouge are significantly different. Furthermore, consumers who purchase Label Rouge salmon are more loyal towards the Label Rouge salmon, in terms of sensitivity to price changes, than those who buy non-label fresh salmon are towards other

salmon. The label is thus able to reach its goal of differentiating between fresh salmon and fresh Label Rouge salmon, both in terms of consumer perception and loyalty.

References

Adamowicz, W.L., and Swait, J.D. (2012). Are Food Choices Really Habitual? Integrating Habits, Variety-Seeking, and Compensatory Choice in a Utility-Maximizing Framework. *American Journal of Agricultural Economics* 95, 17-41.

Allais, O., Bertail, P., and Nichele, V. (2010). The Effects of a Fat Tax on French Households' Purchases: A Nutritional Approach. *American Journal of Agricultural Economics* 92, 228-245.

Asche, F., Guttormsen, A.G., Sebulonsen, T., and Sissener, E. H. (2005). Competition Between Farmed and Wild Salmon: The Japanese Salmon Market. *Agricultural Economics* 33, 333-340.

Asche, F., and Guttormsen, A.G. (2014). Seafood Markets and Aquaculture Production: Special Issue Introduction. *Marine Resource Economics* 29, 301-304.

Barten, A.P. (1996). Maximum Likelihood Estimation of a Complete System of Demand Equations. *European Economic Review* 1, 7-73.

Bertail, P., and, Caillavet, F. (2008). Fruit and Vegetable Consumption Patterns: A Segmentation Approach. *American Journal of Agricultural Economics* 90, 827-842.

Bhattacharya, C.B. (1997). Is Your Brand's Loyalty too Much, too Little, or Just Right? Explaining Deviations in Loyalty from the Dirichlet Norm. *International Journal of Research in Marketing* 14, 421-435.

Blundell, R., and Meghir, C. (1987). Bivariate Alternatives to The Tobit Model. *Journal of Econometrics* 34, 179-200.

Cragg, (1971). Some Statistical Models for Limited Dependent Variables with Application to Durable Goods. *Econometrica*, 39, 829-844.

Creel, M.D., and Loomis, J.B. (1990). Theoretical and Empirical Advantages of Truncated Count Data Estimators for Analysis of Deer Hunting in California. *American Journal of Agricultural Economics* 72, 434–441.

Deaton, A., and Irish, M. (1984). A Statistical Model for Zero Expenditures in Household Budgets. *Journal of Public Economics* 23 (1): 59-80.

Deaton, A., and Muellbauer, J. (1980). An Almost Ideal Demand System. *The American Economic Review* 70, 312-326.

Deb, P., and Trivedi, P.K. (2002). The Structure of Demand for Health Care: Latent Class versus Two-Part Models. *Journal of Health Economics* 21, 601-625.

Dhar, T. and, Foltz, J.D. (2005). Milk by Any Other Name… Consumer Benefits from Labeled Milk. *American Journal of Agricultural Economics* 87, 214-228.

Egan, K., and Herriges, J. (2006). Multivariate Count data Regression Models with Individual Panel Data from an On-Site Sample. *Journal of Environmental Economics and Management* 52, 567-581.

Gustavsen, W. G., and Rickertsen, K. (2003). Forcasting Ability of Theory-Constrained Two-Stage Demand Systems. *European Review of Agricultural Economics* 30, 539-558.

Heckman, J. (1974). Shadow Prices, Market Wages, and Labor Supply. *Econometrica* 42, 679-694.

Hellerstein, D.M. (1991). Using Count Data Models in Travel Cost Analysis with Aggregate Data. *American Journal of Agricultural Economics* 73, 860–866.

Hellström, Jörgen. (2006). A Bivariate Count Data Model for Household Tourism Demand. *Journal of Applied Econometrics* 21, 213–226.

Herrmann, M.L. Mittelhammer, R.C., and Lin, B. (1993). Import Demands for Norwegian Farmed Atlantic Salmon and Wild Pacific Salmon in North America, Japan and the EC. *Canadian Journal of Agricultural Economics* 41, 111-125.

Hynes, Stephen, and William Greene. 2013. A panel travel cost model accounting for endogenous stratification and truncation: A latent class approach. *Land Economics* 89 (1): 177-192.

Jorgensen, D.W. and Lau, L.J. (1975). The Structur of Consumer Preferences. *Annals of Economic and Social Measurement* 4, 49-101.

Kay, J. A., Keen, M. J., and Morris, C. N. (1984). Estimating Consumption from Expenditure Data. *Journal of Public Economics* 23 (1), 169-181.

Keng, K.A., and Ehrenberg, A.S.C. (1984). Patterns of Store Choice. *Journal of Marketing Research* 21, 399-409.

Konijnendijk, Cecil C., Annerstedt, Matilda, Nielsen, Anders B., and Sreetheran Maruthaveeran. (2013). Benefits of urban parks: A systematic review. A report for IFPRA. http://ign.ku.dk/english/employees/landscape-architecture-planning/?pure=files%2F44944034%2FIfpra_park_benefits_review_final_version.pdf Accessed December 1, 2015.

Lau, L. (1976). Complete Systems of Consumer Demand Functions Through Duality in M. Intriligator. *Frontiers in Quantitiative Economics* 71, 262-274.

Larivière, É., Larue, B., and Chalfant, J. (2000). Modeling the Demand for Alcoholic Beverages and Advertising Specifications. *Agricultural Economics* 22, 147-162.

Larson, Douglas M. (1993). Joint Recreation Choices and Implied Values of Time. *Land Economics* 69, 270-286.

Lee, L. F. and Pitt, M. M. (1986). Microeconomic Demand Systems with Binding Nonnegativity Constraints: The Dual Approach. *Econometrica* 54, 1237-1242.

McConnell, Kenneth. E. (1992). On-Site Time in the Demand for Recreation. *American Journal of Agricultural Economics* 74, 918-925.

McCormack, Gavin R., Rock, Melanie, Toohey, Ann M., and Hignell, Danica. (2010). Characteristics of Urban Parks Associated with Park use and Physical Activity. *A review of qualitative research. Health & Place* 16, 712-726.

Meghir, C., and Robin, J.M. (1992). Frequency of Purchase and the Estimation of Demand Systems. *Journal of Econometrics* 53, 53–85.

Monfort, C.M. (2006). Adding value to salmon helps capturing market shares. *IIFET 2006 Portsmouth Proceedings*. Corvallis, Oregon: International Institute of Fisheries Economics & Trade.

Munkin, M.K., and Trivedi, P.K. (1999). Simulated Maximum Likelihood Estimation of Multivariate Mixed-Poisson Regression Models, with Application. *Econometrics Journal* 2, 29-48.

Phaneuf, D.J. Kling, C.L, and, Herriges, J.A, (2000). Estimation and Welfare Calculations in a Generalized Corner Solution Model with an Application to Recreation Demand. *The Review of Economics and Statistics* 82, 83-92.

Pudney, S. (1985). Frequency of Purchase and Engel Curve Estimation. (Discussion Paper A56, LSE Econometrics Program).

Pudney, S. (1986). Modelling Individual Choice: The Econometrics of Corners, Kinks and Holes (Oxford: Basil Blackwell).

Rickertsen K., J.A. Chalfant, and M. Steen (1995). The Effects of Advertising on the Demand for Vegetables. *European Review of Agricultural Economics* 22, 481-494.

Rickertsen, K. (1996). Structural Change and the Demand for Meat and Fish in Norway. *European Review of Agricultural Economics* 23, 316-330.

Rickertsen, K. (1998). The Demand for Food and Beverages in Norway. *Agricultural Economics* 18, 89-100.

Robin, J.M. (1993). Econometric Analysis of the Short-Run Fluctuations of Households' Purchases. *The Review of Economic Studies* 60 (205): 923-934.

Smith, V.K. (1988). Selection and Recreation Demand. *American Journal of Agricultural Economics* 70 (1): 29-36.

Stone, J.R.N. (1954). Expenditure Systems and Demand Analysis: An Application to the Pattern of British Demand. *The Economic Journal* 64: 511-527.

Teisl, M.F., Roe, B., and Hicks, R. L. (2002). Can Eco-Labels Tune a Market? Evidence from Dolphin-Safe Labeling. *Journal of Environmental Economics and Management* 43, 339-359.

Tobin, J. (1958). Estimation of Relationships for Limited Dependent Variables. *Econometrica* 26, 24-36.

Uncles, M., Ehrenberg, A.S.C., and Hammond, K. (1995). Patterns of Buyer Behavior: Regularities, Models, and Extensions. *Marketing Science* 14, 71-78.

Uncles, M., and Lee, D. (2006). Brand Purchasing by Older Consumers: An Investigation Using the Juster Scale and the Dirichlet Model. *Marketing Letters* 17, 17-29.

Wales, T. J. and Woodland, A. D. (1983). Estimation of Consumer Demand Systems with Binding Nonnegativity Constraints. *Journal of Econometrics* 21, 437-468.

Wang, P. (2003). A Bivariate Zero-Inflated Negative Binomial Regression Model for Count Data With Excess Zeros. *Economics Letters* 78, 373-378.

West, S. E., and Williams III, R. C. (2004). Estimates from a Consumer Demand System: Implications for the Incidence of Environmental Taxes. *Environmental Economics and Management* 47, 535-558.

Xie, J., and Myrland, Ø. (2011). Consistent Aggregation in Fish Demand: A Study of French Salmon Demand. *Marine Resource Economics* 26, 267-280.

Paper 1

Demand Systems and Frequency of Purchase Models

Arnar Buason,[1] Dadi Kristofersson[2] and Kyrre Rickertsen[1]

[1] School of Economics and Business, Norwegian University of Life Sciences, P.O. Box 5003, 1432 Ås, Norway. Corresponding author is Arnar Buason: arnar.buason@nmbu.no
[2] Department of Economics, University of Iceland, Reykjavik, Iceland.

Abstract

Understanding both the frequencies of purchase and the average purchased quantities is very important for formulating a loss-leader pricing strategy. We develop a microeconomic model where total purchases of goods are determined by two decisions: (i) the frequencies of purchases, and (ii) the average quantities purchased conditional on positive purchase frequencies. An econometric model is developed to estimate the two demand systems. A Bayesian estimation method is used to allow for an unrestricted covariance structure within the demand systems. The potential usefulness of the model is illustrated by using French scanner data for purchases of fresh fish. The results demonstrate that fresh white fish is a suitable loss-leader product group.

Key words: Demand system, fish, frequency of purchase, multivariate gamma log-normal distribution, multivariate truncated Poisson log-normal distribution.

JEL: C11, C24, D11, D12.

1. Introduction

A widely used marketing strategy is the loss-leader pricing strategy (e.g., Kemp, 1955; Hess and Gerstner, 1987; Lal and Matutes, 1994; Ellison, 2005; Chen and Rey, 2012; In and Wright, 2014). This strategy implies that a retailer reduces the price of a product below its marginal cost to attract customers to the store. These customers will also purchase other and more profitable products and the result may be increased profits. According to Chen and Rey (2012), the loss-leader strategy is mainly used by large retailers who are competing with smaller retailers with a limited variety of products. In this situation, large retailers may find it profitable to sell some products for prices below the marginal costs while other products are sold for prices above the marginal costs.[9]

Knowledge about purchase frequencies and purchased quantities can be used to formulate profitable loss-leader pricing strategies. For example, by reducing the price of a frequently purchased product that is price elastic in the frequency of purchase, a retailer can attract more customers to the store. By simultaneously increasing the price of another product with a high and positive cross-price elasticity in the frequency of purchase with respect to the first product, the retailer can increase the profits. The strategy is most profitable if both products are price inelastic in the quantity of purchase.[10]

Meghir and Robin (1992) presented a theoretical model for a consumer who chooses the purchasing frequency as well as the purchased quantities on each purchasing occasion. However, they never derived their model from a constrained optimization problem. The model was used to specify a demand system. They assumed that any zero purchases are generated from infrequency of purchases, and would have been recorded as positive purchases given a longer observation period. Consequently, observed purchases differ from desired purchases, and they developed a two-step estimator to estimate desired purchases. In their first step, frequencies of purchase were estimated by a generalized linear model to obtain weight parameters[11], which were used to calculate the desired purchases. In their second step, the desired purchases were used to estimate the purchased quantities within a demand system.

---

[9] Consumer purchase frequencies have frequently been modeled by count data models like the Poison and negative binomial models to estimate brand success, brand loyalty, and store choice (e.g., Kau and Ehrenberg, 1984; Uncles *et al*., 1995; Bhattacharya, 1997; Uncles and Lee, 2006). Count data models, such as the Poisson and the negative binomial, have also been used to study demand issues in environmental economics, health economics and finance. Examples include Smith (1988), Creel and Loomis (1990) and Hellerstein (1991) who all used count data models in the estimation of the demand for recreation; Deb and Trivedi (2002) who estimated the number of doctor visits; Munkin and Trivedi (1999) and Wang (2003) who estimated the demand for health care; and Davutyan (1989) who estimated the elasticities of important factors of bank failures. However, a demand system was not used in these applications.

[10] This statement is demonstrated in our empirical example below.

[11] Generally, a Poisson or negative binomial model would be preferred for such an estimation.

We extend the work of Meghir and Robin (1992) in two ways. First, we develop a theoretical model based on a constrained utility maximization problem. Our model results in a frequency of purchase model of the form that was used by Meghir and Robin (1992). We follow Meghir and Robin (1992) and divide the consumer's purchase decision into a decision of frequency of purchase and a decision of quantity to purchase conditional on the purchase frequency. Total purchased quantity of a good is then given as the product of the frequency of purchase and the average quantity purchased. The consumer's choice variables are thus how often to buy different products and how much to buy on average of the products on each shopping trip. In our model, homogeneity of degree zero holds for purchase frequencies, average quantities and total quantities. Furthermore. the matrix of compensated substitution effects is symmetric and negative semi-definite.

Second, Meghir and Robin (1992) focused on how to adjust for the frequency of purchase to obtain consistent parameter estimates in a demand system under an infrequency of purchase assumption.[12] However, their frequency of purchase model was a basic Poisson system. In this system, homogeneity and symmetry were not imposed, the problem of zero inflation was not addressed,[13] and the covariance structure was assumed to be zero. We extend their count data estimation framework by: (i) accounting for homogeneity and symmetry in the frequency of purchase demand system, (ii) accounting for zero inflation by assuming a truncated data generating process for the counts, and (iii) allowing for an unrestricted covariance structure within the two demand systems. We assume a truncated multivariate Poisson log-normal (TMPLN) distribution for the counts and a multivariate

[12] In applied demand analysis for disaggregate products, zero purchases represent a major problem. Usually it is assumed that zero purchases either represent traditional corner solutions, infrequency of purchase or non-preference for a product (e.g., Wales and Woodland, 1983; Tiffin and Arnoult, 2010). To account for zero purchases in demand system estimation, a two-step model is usually estimated. In the first step, it is estimated whether a product is purchased or not. The results of this step are used to correct the estimated demand system in the second step to obtain consistent parameter estimates. However, the actual frequencies of purchase are not estimated.

[13] When the share of zeros in a data set significantly exeeds the amount predicted by conventional distributions, such as the Poisson, it is called zero inflation.

gamma log-normal (MGLN) distribution for the average quantities. To estimate these distributions, we use Bayesian estimation methods, specifically a random walk Metropolitan simulation algorithm.

We provide an empirical illustration on the potential usefulness of this type of models. Our example uses French scanner data for purchases of fresh fish, and the demand system includes fresh salmon, fresh white fish, and other types of fresh fish. Our results indicate that fresh white fish is a good loss-leader candidate. Large retailers who sell a wide variety of products, including a large selection of fresh white fish, could price fresh white fish below marginal cost and price other products such as coffee at a monopoly price. A smaller retailer, who only sell fresh salmon and fresh white fish, may be unable to compete with these prices and he would be driven out of the market.

The paper is organized as follows. In section two, our theoretical model is specified. In section three, our statistical model is developed. In section four, our data set and empirical specification are described. In section five, our empirical results are presented before we conclude.

2. Theoretical Model

We follow the specification and notation in Meghir and Robin (1992) as far as possible. However, we divide the decision of the quantity purchased of each good $x = (x_1, x_2, \dots, x_M)$ into two parts: the frequencies of purchase $n = (n_1, n_2, \dots, n_M)$ and the average quantities purchased on each occasion, $q = (q_1, q_2, \dots, q_M)$. By definition, the identity $x_i(n_i q_i) \equiv n_i q_i$ holds for $i = 1,2, \dots, M$. For convenience, it is assumed that *n* is continuous.[14] The consumer is assumed to obtain utility from purchased goods, leisure *l*, and frequency of purchase, and

[14] We assume that *n* is a latent variable of actual observed purchase frequencies, which can take any positive value and, consequently, we can take the derivatives with respect to *n*.

the utility function is specified as $v(x, n, l) = v(nq, n, l)$ where $nq = (n_1q_1, n_2q_2, \ldots, n_Mq_M)$. The utility function is assumed to be strictly quasiconcave in *n*, *q*, *l*.[15]

The consumer has wage income $y = wk$ where *w* is the hourly wage rate and k is the number of hours spent at work. The consumer may also have other types of income *R*, which we assume is exogenously given. The consumers budget constraint is $y + R = p'nq$ where $p = (p_1, p_2, \ldots, p_M)$ is the price vector.

The consumer has a constant time endowment *T*, which can be allocated between leisure, work and purchasing goods. The time spent on purchasing goods is given by the function $g(n)$, which is assumed to be increasing in *n*.[16] The consumer's time constraint is $T = l + k + g(n)$.

The consumer's utility maximization problem is specified as:[17]

$$\max_{n,q,l}\{v(x(nq), n, l): wT + R = p'x(nq) + wl + wg(n), n > 0, q > 0, l > 0\}. \quad (1)$$

As shown in Appendix 1, it follows from the first-order conditions to this problem that the solution satisfies:

$$\frac{\partial v/\partial x_i}{\partial v/\partial x_j} = \frac{p_i}{p_j} \text{ for } \forall i, j \quad (2)$$

$$\frac{\partial v/\partial x_i}{\partial v/\partial l} = \frac{p_i}{w} \text{ for } \forall i, j \quad (3)$$

$$\frac{\partial v/\partial n_i}{\partial v/\partial l} = \frac{\partial g}{\partial n_i} \text{ for } \forall i \quad (4)$$

---

[15] Meghir and Robin (1992) provide two arguments for positive marginal utilities in frequencies of purchase. First, there is a benefit in the form of saved space by having to hold smaller stocks of various goods. Second, freshness is important for many types of food including fish. The costs of frequent purchases will be reflected as lost leisure time.

[16] The consumer can purchase many goods at each occasion and the function $g(n)$ is increasing at a decreasing rate with the number of products purchased.

[17] For simplicity, corner solutions are not considered in the theoretical model. The zeros in the data are assumed to be generated by infrequency of purchase and non-preference. The first-order conditions to problem (1) is given in Appendix A1. Furthermore, from the budget constraint it follows that the wage rate is also the opportunity cost of time, which either is used for leisure or shopping.

$$\frac{\partial v/\partial n_i}{\partial v/\partial n_j} = \frac{\partial g/\partial n_i}{\partial g/\partial n_j} \text{for } \forall i, j. \tag{5}$$

Equations (2) and (3) are the standard first-order conditions of the consumer utility maximization problem with leisure. Equation (4) implies that the ratio between the marginal utility of shopping frequency for good *i* and the marginal utility of leisure equals the marginal cost of shopping frequency, in terms of time, for good *i*.[18] Equation (5) implies that the ratio of marginal utility of shopping frequencies for goods *i* and *j* are equal to the ratio of marginal costs of shopping goods *i* and *j*, in terms of time.

The solution of the first-order conditions of Equation (1) results in three sets of uncompensated demand functions $n(p, w, R)$, $q(p, w, R)$ and $l(p, w, R)$. The total purchased quantities, $x(p, w, R)$ are found by substituting $n(p, w, R)$ and $q(p, w, R)$ into the identity $x_i(n_i, q_i) \equiv n_i q_i$.

The dual problem to the consumer's utility maximization problem (1) is: $\min_{n,q,l}\{R = p'nq - w(T - l - g(n)) \colon v(x(nq), n, l) = v^*, n > 0, q > 0, l > 0\}$. The solution to the first-order conditions of this problem results in three sets of compensated demand functions $n^c(p, w, v^*)$, $q^c(p, w, v^*)$ and $l^c(p, w, v^*)$.[19] The compensated demand functions for total purchased quantities, $x^c(p, w, v^*)$ are found by substituting $n^c(p, w, v^*)$ and $q^c(p, w, v^*)$ into the identity $x_i^c \equiv n_i^c q_i^c$.

The conditions for homogeneity of degree zero in prices and income, symmetry, and negativity for the compensated and uncompensated demand functions are summarized in:

Proposition 1

(1.1) The demand equations $n(p, w, R)$, $q(p, w, R)$, $l(p, w, R)$, and $x(p, w, R)$ are homogeneous of degree zero in (p, w, R).

---

[18] The marginal cost of increased shopping frequency is given by the forgone wage, $w(\partial g/\partial n_i)$, i.e., from spending additional time on shopping.

[19] The first-order conditions to the minimization problem is given in Appendix A2.

(1.2) The matrix of compensated substitution effects for $x^c(p, w, v^*)$ is symmetric and negative semidefinite.

(1.3) The matrix of compensated substitution effect for the product $n^c(p, w, v^*)q^c(p, w, v^*)$ is symmetric and negative semidefinite.

These propositions are proved in Appendix A3.

3.1 Statistical Model

The frequency of shopping $n_i = (n_{i11}, n_{i12}, \dots, n_{iKT})$ is assumed to follow a discrete distribution $f_{Ni}(n_i|\beta_i, C)$, for $n = 0,1,2, \dots$ where β_i is a vector of parameters and C is a matrix of explanatory variables. *N* represents a random variable and n_{ikt} is an observed value of *N*, where the subscript *i* denotes the product, $k = 1,2, \dots, K$, and $t = 1,2, \dots, T$ denote household and time period, respectively. The average purchases $q_i = (q_{i11}, q_{i12}, \dots, q_{iKT})$ are only observed after a trip to the shop. Thus, the variable $q_i|n_i > 0$ is assumed to follow a continuous distribution $f_{Qi|n_i>0}(q_i|\alpha_i, C, n_i > 0)$, defined only over positive values, where α_i is a vector of parameters. The interpretation of Q and q_{ikt} are the same as for *N* and n_{ikt} above. The data generating process (DGP) for average quantity purchased is therefore represented by the following two-part model:

$$f_Q(q_i|\alpha_i, C) = \begin{pmatrix} \Pr(N = 0|\beta_i, C) & \text{if } q_i = 0 \\ \Pr(N > 0|\beta_i, C) f_{Qi|n>0}(q_i|\alpha_i, C, n_i > 0) & \text{if } q_i > 0 \end{pmatrix}. \quad (6)$$

The decision to purchase a good and how much to purchase in each trip are likely to be related choices, and it is desirable to model them as stochastically correlated. Furthermore, the demand for one good is usually related to the demand for other goods and it is important to allow for correlation between equations within each of the two systems. To allow for these correlations, random effects are introduced to both densities $f_{Ni}(n_i|\beta_i, C, b_{Ni})$ and

$f_{Qi|n>0}\left(q_i \middle| \alpha_i, C, n > 0, b_{Qi}\right)$, where b_{Ni} and b_{Qi} are random effects which are assumed to follow a multivariate normal distribution:

$$\begin{bmatrix} b_N \\ b_Q \end{bmatrix} \Bigg| D \sim MVN\left(\begin{bmatrix} 0^m \\ 0^m \end{bmatrix}, \begin{bmatrix} D_N & D_{NQ} \\ D_{NQ} & D_Q \end{bmatrix}\right), \tag{7}$$

where $b_N = (b_{N1}, \dots, b_{Nm})$, $b_Q = \left(b_{Q1}, \dots, b_{Qm}\right)$, and D is an unrestricted block covariance matrix. The conditional means of n_i and q_i as well as their corresponding marginal effects are given in appendix A4. The joint probability density function for n_i and q_i is given as:

$$p(n_i, q_i | \beta, \alpha, D, C) =$$

$$\int \prod_{t=1}^{T} f_{Ni}(n_{ikt} | \beta_i, C, b_{Nikt})\, f_{Qi|n>0}(q_i | \alpha_i, C, n_i > 0) \phi(b_{ik} | 0, D) db_{ik}. \tag{8}$$

The product operator is inside the integral since b_N and b_Q each have one draw for the $2T$ random variables $n_{ik1}, n_{ik2}, \dots, n_{ikT}$ and $q_{ik1}, q_{ik2}, \dots, q_{ikT}$, respectively. Thus, there is a new draw for each cluster, but not for each time period within a cluster. The likelihood is then given by:

$$L = \prod_{k=1}^{K} \prod_{i=1}^{M} p(n_i, q_i | \beta, \alpha, D, C). \tag{9}$$

Since the joint density $p(n_i, q_i | \beta_i, \alpha_i, D, C)$ does not have a closed form solution, the likelihood L can not be optimized with conventional Newton methods, and we use simulation methods.

3.2 Distribution Assumptions

To account for the large share of zeros in the data, we assume that the frequency of purchase, n_i, is generated from $n_i | n_i > 0, \beta_i, C, \sim$ truncated $\text{Poisson}(\mu_i)$, where $\mu_i = \exp(C\beta_i + b_{Ni})$. This results in a multivariate truncated Poisson log-normal distribution:

$$f_{Ni}(n_i|n_i > 0, \beta, D_N, C) = \int \frac{\exp\left(-\mu_i(b_{Ni})\right)\left(\mu_i(b_{Ni})\right)^{n_i}}{\left[1-\exp\left(-\mu_i(b_{Ni})\right)\right]n_i!}\phi(b_{Ni}|0, D_N)db_{Ni}. \tag{10}$$

where $\phi(b_{Ni}|0, D_N)$ is the multivariate normal distribution for b_{Ni}, with a covariance matrix D_N. We assume that the average quantity of purchases, q_i, is generated from $q_i|\alpha_i, C, n_i > 0 \sim \text{Gamma}(\kappa_i, \eta_i)$, where the mean of the gamma distribution is specified as $\kappa_i\eta_i = \exp\left(C_i\alpha_i + b_{Qi}\right)$. This specification results in a multivariate gamma log-normal distribution:

$$f_{Qi}(q_i|\alpha_i, C, n_i > 0) = \int \frac{q_i^{\kappa_i - 1}\exp(-q_i/\eta_i)}{\eta_i^{\kappa}\Gamma(\kappa_i)}\phi\left(b_{Qi}\middle|0, D_Q\right)db_{Qi} \tag{11}$$

where $\phi\left(b_{Qi}\middle|0, D_Q\right)$ is the multivariate normal distribution for b_{Qi}, with a covariance matrix D_Q.

3.3 Priors and MCMC Sampling

We assume uninformative priors, which is a common practice, see for example Chib and Winkelmann (2001). Let $\beta \sim \text{N}(\beta_0, B_0^{-1})$, $\alpha \sim \text{N}(\alpha_0, A_0^{-1})$, $\kappa \sim \text{Gamma}(k_0, s_0)$ $D_N^{-1} \sim \text{Wishart}(v_{N0}, R_{N0})$, $D_Q^{-1} \sim \text{Wishart}\left(v_{Q0}, R_{Q0}\right)$, where $\beta_0, B_0, \alpha_0, A_0, k_0, s_0, v_{N0}, R_{N0}, v_{Q0}, R_{Q0}$ are known hyperparameters and $\text{Wishart}(v_{o0}, R_{o0})$ is the Wishart distribution with v_{o0} degrees of freedom and a scale matrix R_{o0}, where $o = N, Q$.[20] By the Bayes theorem the posterior density of the two parts of the model are proportional to the following expressions:

$$\phi(\beta|\beta_0, B_0^{-1})f_W(D_N^{-1}|v_{N0}, R_{N0})\prod_{k=1}^{K}\prod_{i=1}^{M} f_N(n_{ik}|\beta, b_{Nik})\,\phi(b_{Nik}|0, D_N), \tag{12}$$

$$\phi(\alpha|\alpha_0, A_0^{-1})f_W\left(D_Q^{-1}\middle|v_{Q0}, R_{Q0}\right)\prod_{k=1}^{K}\prod_{i=1}^{M} f_Q\left(q_{ik}\middle|\alpha, b_{Qik}\right)\phi\left(b_{Qik}\middle|0, D_Q\right), \tag{13}$$

where f_W is the Wishart density. We then construct Markov chains using the blocks of parameters b_N, b_Q, β, α, D_N and D_Q and the full conditional distributions:

[20] For an application of the Wishart distribution see for example example Chib and Winkelmann (2001).

$$[b_N|n,\beta,D]; \quad [\beta|n,b_N]; \quad [D|b_N], \tag{14}$$

$$\left[b_Q|q,\alpha,D\right]; \quad \left[\alpha|q,b_Q\right]; \quad \left[D|b_Q\right]. \tag{15}$$

The blocks of parameters $[\beta|n,b_N]$ and $\left[\alpha|q,b_Q\right]$ are then separated into smaller blocks to help with convergence. The simulation output is generated by recursively simulating these distributions, using the most recent values of the conditioning variables at each step.

The sampling of b_N and b_Q starts with specifying the target densities:

$$\pi(b_N|n,\beta,D) = \prod_{k=1}^{K}\prod_{i=1}^{M}\pi(b_{Nik}|n_{ik},\beta,D), \tag{16}$$

$$\pi\left(b_Q|q,\alpha,D\right) = \prod_{k=1}^{K}\prod_{i=1}^{M}\pi\left(b_{Qik}|q_{ik},\alpha,D\right). \tag{17}$$

To sample the density of the k^{th} household of the i^{th} cluster we have:

$$\pi(b_{Nik}|n_{ik},\beta,D) = c_{ik}\phi(b_{Nik}|0,D_N)\prod_{t=1}^{T}\frac{\exp\left(-\exp(C_{ikt}\beta_t + b_{Nikt})\right)}{[1-\exp(C_{ikt}\beta_t + b_{Nikt})]} \tag{18}$$

$$\times\,[-\exp(C_{ikt}\beta_t + b_{Nikt})]^{n_{ikt}}$$

$$\equiv c_{ik}\pi^{+}(b_{Nik}|n_{ik},\beta,D_N),$$

$$\pi\left(b_{Qik}|q_{ik},\alpha,\kappa,D\right) = i_{ik}\phi\left(b_{Qik}|0,D_Q\right)\prod_{t=1}^{T}\frac{q_{ikt}^{C_{ikt}\alpha_t+b_{Nikt}}}{\Gamma\left(C_{ikt}\alpha_t + b_{Qikt}\right)\kappa^{C_{ikt}\alpha_t+b_{Qikt}}} \tag{19}$$

$$\times\;\exp[-q_{ikt}/\kappa]$$

$$\equiv i\pi^{+}\left(b_{Qik}|q_{ik},\alpha,\kappa,D_Q\right).$$

The likelihood of the target distributions are a truncated Poisson log-normal mixture and a gamma log-lognormal mixture, respectively, as specified in Section 3.2. The simulation algorithm is a random-walk Metropolis algorithm. This algorithm requires little computation power in each step but suffers from high autocorrelation and more Markov Chain Monte Carlo (MCMC) steps are required. The proposal density is found by approximating the target density around the modal value by a multivariate *t*-distribution. Let $\hat{b}_{Nik} =$ $\operatorname{argmax} \ln \pi^+(b_{Nik}|n_{ik}, \beta, D_N)$ and $V_{b_{Nik}} = \left(H_{b_{Nik}}\right)^{-1}$ be the inverse of the Hessian of $\ln \pi^+(b_{Nik}|n_{ik}, \beta, D_N)$ at the mode $\hat{b}_{Nik}$. To find these quantities, we use the Newton-Raphson algorithm. Then, our proposal density is $b_{Nik}^{(s)} \left| b_{Nik}^{(s-1)} \sim t\left(b_{Nik}^{(s-1)}, V_{b_{Nik}}, v\right)\right.$, where *v* is the degrees of freedom and *s* indicates the number of the draw. We make a draw e_{ik} from $t\left(0, V_{b_{Nik}}, v\right)$, where $b_{Nik}^{(s)} = b_{Nik}^{(s-1)} + e_{ik}$, and we move from $b_{Nik}^{(s-1)}$ to $b_{Nik}^{(s)}$ with probability:

$$r = \min\left\{\frac{\pi^+\left(b_{Nik}^{(s)}\middle|n_{ik}, \beta, D_N\right)}{\pi^+\left(b_{Nik}^{(s-1)}\middle|n_{ik}, \beta, D_N\right)}, 1\right\}. \tag{20}$$

Next, we sample *u* from a uniform distribution $\mathrm{U}(0,1)$ and if $u < r$ then $b_{Nik}^{(s)} = b_{Nik}^{*}$ otherwise $b_{Nik}^{(s-1)} = b_{Nik}^{*}$. We use the same steps for the sampling of b_Q. The sampling of β and α follow the approach above. The respective target distributions are given as follows:

$$\pi(\beta|n, b_N, D_N) = \phi(\beta|\beta_0, B_0^{-1}) \prod_{k=1}^{K}\prod_{i=1}^{M}\prod_{t=1}^{T} \frac{\exp(-\exp(C_{ikt}\beta_t + b_{Nikt}))}{[1 - \exp(C_{ikt}\beta_t + b_{Nikt})]} \times [-\exp(C_{ikt}\beta_t + b_{Nikt})]^{n_{ikt}} \tag{21}$$

$$\pi\left(\alpha\middle|q, b_Q, \kappa, D_Q\right) = \phi(\alpha|\alpha_0, A_0^{-1}) \prod_{k=1}^{K}\prod_{i=1}^{M}\prod_{t=1}^{T} \frac{q_{ikt}^{C_{ikt}\alpha_t + b_{Nikt}}}{\Gamma(C_{ikt}\alpha_t + b_{Nikt})\kappa^{C_{ikt}\alpha_t + b_{Qikt}}} \times \exp[-q_{ikt}/\kappa]. \tag{22}$$

Sampling D_N^{-1} and D_Q^{-1} is a simpler process than the other two blocks of parameters, since we specified a hyperprior, which resulted in a Wishart distribution. We sample D_o^{-1}, $o = N, Q$, from a distribution proportional to:

$$f_W(\,D_o^{-1}|v_{o0}, R_{o0}^{-1}) \prod_{k=1}^{K}\prod_{i=1}^{M} \phi(b_{oik}|0, D_o). \tag{23}$$

Combining terms the expression results in the Wishart distribution:

$$D_o^{-1}|b_0 \sim \text{Wishart}\left(M + v_{o0}, \left[R_{o0}^{-1} + \sum_{k=1}^{K}\sum_{i=1}^{M}(b'_{oik}b_{oik})\right]^{-1}\right), \tag{24}$$

with degrees of freedom $M + v_{o0}$ and scale matrix $[R_{o0}^{-1} + \sum_{k=1}^{K}\sum_{i=1}^{M}(b'_{oik}b_{oik})]^{-1}$.

4. Data and Empirical Example

French scanner data for the purchases of fresh fish as recorded by Kantar Worldpanel for the period 2005 – 2008 are used. The data is a rotating panel, i.e., when households drop out new households take their place. The number of participants recording fresh fish purchases were 3,291 in 2005, 3,234 in 2006, 3,165 in 2007, and 4,479 in 2008, which show a significant rotation of households in 2008. Purchased quantities and total expenditures are recorded, and the associated unit prices are constructed by dividing total expenditure by quantities. This approach is used in other demand studies such as Allais et al. 2010 and Bertail and Caillavet (2008). The data are weekly and includes many zero purchases. To reduce the number of zero purchases, weekly purchases were aggregated to yearly purchases, and the panel structure is accounted for by using random effects as discussed in the previous section. Demand for salmon has been increasing during the past years, and fresh fish is the most common product form in France (Xie and Myrland, 2011). Therefore, we include fresh salmon, fresh white fish, and other fresh fish in our demand system.

In Table 1, the yearly purchase frequencies in the sample for fresh salmon, fresh white fish, and other fresh fish over the four year survey period are presented. Between 14% and 45% did not purchase each of the three types of fish. Many households with a positive purchase of one fish type only purchased this type once or twice over the four years. However, more than 10% of the households purchased white fish more than ten times and almost 24% purchased other fish more than ten times.

(Table 1 about here)

Table 2 shows descriptive statistics of the variables used in our empirical model. The average frequency of purchases and average quantities for all three fish types are quite low due to the large share of zeros in the data set. However, the maximum frequency is 172 times for the group other fish. The average purchase of each type of fish is around 500 grams and the maximum average purchase of each type of fish is around 9 kilograms. The unit prices and total expenditures on fresh fish are divided by the average French CPI to impose homogeneity of degree zero in frequencies of purchases, average quantities, and total quantities.

(Table 2 about here)

Total quantities purchased are specified as:

$$x_i = \exp\big(b(p, w, R, \theta_i)\big), \tag{25}$$

where the function $b(\cdot)$ is linear in its parameters $\theta_i = (\theta_{1i}, \theta_{2i}, \dots, \theta_{ni})$. We decompose $b(\cdot)$ into a frequency part $f(p, w, R, \beta_i)$ and an average quantity part $z(p, w, R, \alpha_i)$ with the associated functions and , where $\beta_i = (\beta_{1i}, \beta_{2i}, \dots, \beta_{ni})$ and $\alpha_i = (\alpha_{1i}, \alpha_{2i}, \dots, \alpha_{ni})$ are parameter vectors. Using this decomposition, Equation (25) becomes:

$$x_i = n_i q_i = \exp\big(f(p, w, R, \beta_i)\big)\exp\big(z(p, w, R, \alpha_i)\big). \tag{26}$$

As is conventional when estimating count data models the conditional expectation is defined as a semi-logarithmic function, see for example Shonkwiler and Englin (2005). The purchase frequency demand function for each fish type is thus given by the following expression:

$$\mathrm{E}\big(n_{ij} | Z_i\big) = \exp\left(\beta_{ij} + \sum_{s=1}^{M} \beta_{is}\big(p_{sj}/CPI\big) + \theta_i\left(\big(R_j + y_j\big)/CPI\right) + b_{Nij}\right) \tag{27}$$

for $s = 1,2, \dots, M$ fish categories and $j = 1,2, \dots, J$ households. The price of fish category i for household j is denoted by p_{ij}. The total fish expenditure of household j is given by

$R_j + y_j$, *CPI* is the French consumer price index, and b_{Nij} is the random effects described in Section 3. For the demand system to be consistent with economic theory, homogeneity and symmetry are imposed on the parameters. LaFrance and Hanemann (1989) and LaFrance (1990) derived these restrictions for the semi logarthmnic functional form. In our case, we must have $\beta_{is} = 0 \; \forall \; i \neq s$ and $\theta_i = \theta \; \forall \; i$. To allow for additional explanatory variables we define β_{ij} as follows:

$$\beta_{ij} = \eta_{0i} + \sum_{k=1}^{K} z_{kj}\eta_{ki}. \tag{28}$$

The average quantity demand function for each good is given by the following expression:

$$\mathrm{E}\left(q_{ij}|Z_i\right) = \exp\left(\alpha_{ij} + \sum_{s=1}^{M} \alpha_{is}\left(p_{sj}/CPI\right) + \theta_i\left(\left(R_j + y_j\right)/CPI\right) + b_{Qij}\right) \tag{29}$$

The same parameter restrictions are imposed on this system as on the system for frequencies of purchase, and α_{ij} is defined as β_{ij} in Equation (28).

$$e_{isj} = n_{ij}\frac{\partial n_{sj}}{\partial y_j} = \theta n_{ij} n_{sj} \tag{30}$$

where e_{isj} is the compensated substitution effect between products *i* and *s* for household *j*, and the *n*'s are purchase frequencies of different product groups by household *j*. The same relationship holds for the compensated substitution effect of average purchase frequencies.

5. Empirical Results

Table 3 shows the posterior summary for the fresh salmon equations based on the TMPLN and MGLN models. The purchase frequencies are estimated by the TMPLN model, and the average purchases are estimated by the MGLN model. All parameters in the frequency of purchase model are significant except for the time dummies of 2007 and 2008, only the signs can be interpreted directly in the posterior summaries, which are all as expected. The Geweke

Z-test of Markov chain stationarity is not rejected at the 5% level except for the time dummy for 2006.[21] All the parameters in the equation for average purchases are significant and generated from a stationary Markov process. The time dummies show that average quantities have increased for all years relative to 2005. Kappa is the shape parameter of the gamma distribution and is highly significant.[22]

(Table 3 about here)

Table 4 shows the posterior summary for the fresh white fish equations. All parameters in the frequency of purchase model are significant except the 2008 dummy. Furthermore, it is not rejected that they are generated from a stationary Markov process. The time dummies show that frequency of purchase has decreased relative to 2005. All parameters in the equation for average purchases are significant and their Markov chains are stationary except for the constant. As for salmon, the time dummies show an increase relative to 2005.

(Table 4 about here)

Table 5 shows the posterior summary for the other fresh fish equations. All the parameters of both models are significant and generated from a stationary Markov process. The frequency of purchase has been decreasing over time whereas the average quantity demanded has been increasing relative to 2005.

(Table 5 about here)

Table 6 shows the cross-equations covariance matrix for the models, that is the covariance matrix of the random effects parameters. Sigma11 is thus the variance of the random effects in equation one, Sigma22 is the variance of the random effects in equation two and Sigma33 is the variance of the random effects in equation three. Sigma12 and Sigma21 show the

[21] Geweke Z values are normally distributed. For an application of the Geweke convergence test see for example Nylander et al. (2008).

[22] Many probability distributions are not a single distribution, but are in fact a family of distributions since they have a shape parameter, or shape parameters. Such parameters allow distribution to take on different shapes, depending on the value of the shape parameter. These distributions are very useful in modeling since they are flexible and can fit to many different datasets.

covariance between random effects of equation one and two, and Sigma13 and Sigma31 show the covariance between equation one and three. Finally, Sigma23 and Sigma32 show the covariance between equation two and three. All parameters in both systems are significant and generated from a stationary Markov process. The significant covariances demonstrate the importance to allow for an unrestricted covariance matrix.

(Table 6 about here)

Table 7 shows the own-price and total expenditure elasticities of both models and the also the elasticities regarding total quantities. The own-price elasticities for total quantities purchased are very similar for the three types of fish whereas the elasticities for purchase frequencies and average quantities are very different for the three types of fish. These differences make it possible to find a good loss-leader candidate.

The results in Table 7 and Table 8 indicate that fresh white fish is the best loss-leader candidate since it has almost the same own-price elasticity with respect to frequency as other fresh fish but has the lowest own-price elasticity with respect to average quantity. Now let us assume that a retailer, who does not only sell fresh fish, wants to create a profitable pricing strategy that also hurts fish retailers who compete with him in the market for fresh fish. If the retailer reduces the price of fresh white fish by 0.5%, and simultaneously increases the price of fresh salmon by 5% the net revenue gain will be around 3.57 Euros per customer on average over the four-years period. In addition, the average shopping frequency per individual increases by 0.17%. Most consumers will not only purchase fresh fish at this larger retailer, and let us assume that on average customers buy other goods for 100 Euros in each trip then the retailer would gain an extra 0.17 Euros, which exceeds the cost of 0.16 Euros from lowering the average price of fresh white fish. In total, the retailer would gain a 3.74 Euros revenue increase per customer over the four-years period by applying this loss-leader strategy. For a large retailer with 10,000 customers, this price strategy would increase revenues by 37,400 Euros over the period, from the marginal changes of these two prices.

(Table 7 about here)

(Table 8 about here)

6. Conclusions

For marketing strategies such as the loss-leader pricing strategy, the frequency of purchase and the average quantities purchased are of interest. In this paper, the frequencies of purchase, average purchases conditional on purchasing frequencies, and total purchased quantities are estimated. The paper contains two contributions. First, we extend the model in Meghir and Robin (1992) by assuming that total purchased quantity can be expressed as the product of average purchased quantity and the frequency of purchase. Furthermore, we introduce more elaborate restrictions and derive the model as well as providing proofs of homogeneity, symmetry, and negativity.

Second, we introduce a new estimation method, which is consistent with the microeconomic model. The econometric model by Meghir and Robin (1992) did not focus on the requirements imposed by theory on the selected functional form. Furthermore, the problem of zero frequencies was not addressed in their article, and they estimated a Poisson model. The covariance structure in their Poisson system is assumed to be zero. We extend their count data estimation framework in three ways. (i) We account for homogeneity and symmetry in the count data demand system. (ii) The problem of zero purchases is accounted for by assuming a truncated data generating process for the counts. (iii) The covariance structure is unrestricted within the two demand systems by assuming a multivariate Poisson log-normal distribution for the counts, and a multivariate gamma log-normal distribution for the average quantities. To be able to estimate these complicated distributions, we use Bayesian estimation methods, specifically a random walk Metropolitan simulation algorithm.

The proposed method is applied to French fish purchasing data to estimate the demand for fresh salmon, fresh white fish, and other fresh fish. The results show that fresh white fish could be a good loss-leader category for large retailers to increase store traffic and sales of inelastic products sold with a markup above the market price.

Future research could be focused on providing a more elaborate empirical example with a different product group and a larger range of products. Such a study could be a great help for large retailers to from their pricing strategies.

References

Allais, O., Bertail, P., and Nichele V. (2010). The Effect of a Fat Tax on French Households' Purchases: A Nutrition Approach. *American Journal of Agricultural Economics* 92, 228-245.

Bertail, P., and Caillavet, F. (2008). Fruit and Vegetable Consumption Patterns: A Segmentation Approach. *American Journal of Agricultural Economics* 90, 827-842.

Bhattacharya, C.B. (1997). Is Your Brand' s Loyalty too Much, too Little, or Just Right? Explaining Deviations in Loyalty from the Dirichlet Norm. *International Journal of Research in Marketing* 14, 421-435.

Chen, Z., and Rey, P. (2012). Loss Leading as an Exploitative Practice. *The American Economic Review* 102, 3462-3482.

Chib, S. and Winkelmann, R. (2001). Markov Chain Monte Carlo Analysis of Correlated Count Data. *Journal of Business and Economic Statistics* 19, 428-435.

Creel, M.D., and Loomis, J.B. (1990). Theoretical and Empirical Advantages of Truncated Count Data Estimators for Analysis of Deer Hunting in California. *American Journal of Agricultural Economics* 72, 434–441.

Davutyan, N. (1989). Bank Failures as Poisson Variates. *Economics Letters* 29, 333-338.

Deb, P., and Trivedi, P.K. (2002). The Structure of Demand for Health Care: Latent Class versus Two-Part Models. *Journal of Health Economics* 21, 601-625.

Ellison, G. (2005). A Model of Add-On Pricing. *The Quarterly Journal of Economics* 120, 585-637.

Hellerstein, D.M. (1991). Using Count Data Models in Travel Cost Analysis with Aggregate Data. *American Journal of Agricultural Economics* 73, 860–866.

Hess, J.D., and Gerstner, E. (1987). Loss Leader Pricing and Rain Check Policy. *Marketing Science* 6, 358-374.

In, Y., and Wright, J. (2014). Loss-Leader Pricing and Upgrades. *Economics Letters* 122, 19-22.

Kemp, M.C. (1955). An Appraisal of Loss Leader Selling. *The Canadian Journal of Economics and Political Science* 21, 245-250.

LaFrance, J.T., and Hanemann, W.M. (1989).The Dual Structure of Incomplete Demand Systems. *American Journal of Agricultural Economics* 71, 262-274.

LaFrance, J.T. (1990). Incomplete Demand Systems and Semilogarithmic Demand Models. *Australian Journal of Agricultural Economics* 34, 118-131.

Lal, R., and Matutes, C. (1994). Retail Pricing and Advertising Strategies. *Journal of Business* 67, 345-370.

Meghir, C., and Robin, J.M. (1992). Frequency of Purchase and the Estimation of Demand Systems. *Journal of Econometrics* 53, 53–85.

Munkin, M.K., and Trivedi, P.K. (1999). Simulated Maximum Likelihood Estimation of Multivariate Mixed-Poisson Regression Models, with Application. *Econometrics Journal* 2, 29-48.

Nylander, J.A.A., Wilgenbusch, J.C., Warren, D.L., and Swofford, D.L. (2008). AWTY (Are We There Yet?): A System for Graphical Exploration of MCMC Convergence in Bayesian Phylogenetics. *Bioinformatics* 24, 581-583.

Shonkwiler, J.S., and Englin, J. (2005). Welfare Losses Due to Livestock Grazing on Public Lands: A Count Data Systemwide Treatment. *American Journal of Agricultural Economics* 87, 302-313.

Smith, V.K. (1988). Selection and Recreation Demand. *American Journal of Agricultural Economics* 70, 29-36.

Tiffin, R., and Arnoult, M. (2010). The Demand for a Healthy Diet: Estimating the Almost Ideal Demand System with Infrequency of Purchase. *European Review of Agricultural Economics* 37, 501-521.

Uncles, M., Ehrenberg, A.S.C., and Hammond, K. (1995). Patterns of Buyer Behavior: Regularities, Models, and Extensions. *Marketing Science* 14, 71-78.

Uncles, M., and Lee, D. (2006). Brand Purchasing by Older Consumers: An Investigation Using the Juster Scale and the Dirichlet Model. *Marketing Letters* 17, 17-29.

Wales, T.J., and Woodland, A.D. (1983). Estimation of Consumer Demand Systems with Binding Non-Negativity Constraints. *Journal of Econometrics* 21, 263-285.

Wang, P. (2003). A Bivariate Zero-Inflated Negative Binomial Regression Model for Count Data With Excess Zeros. *Economics Letters* 78, 373-378.

Xie, J., and Myrland, Ø. (2011). Consistent Aggregation in Fish Demand: A Study of French Salmon Demand. *Marine Resource Economics* 26, 267-280.

Table 1. Purchase Frequencies

No of Purchases	Salmon		White Fish		Other Fish	
	Frequency	%	Frequency	%	Frequency	%
0	6326	44.64	4126	29.11	1982	13.99
1	2784	19.64	2738	19.32	2357	16.63
2	1514	10.68	1654	11.67	1603	11.31
3	1001	7.06	1135	8.01	1223	8.63
4	685	4.83	871	6.15	953	6.72
5	473	3.34	635	4.48	741	5.23
6	337	2.38	489	3.45	580	4.09
7	254	1.79	402	2.84	549	3.87
8	151	1.07	342	2.41	446	3.15
9	131	0.92	248	1.75	345	2.43
10≥	516	3.65	1532	10.84	3393	23.96

Table 2. Descriptive Statistics

Variable	Mean	Std. Dev.	Min	Max
Salmon (frequency)	1.96	3.39	0.00	52.00
White fish (frequency)	3.77	6.03	0.00	87.00
Other fish (frequency)	7.43	11.25	0.00	172.00
Salmon (average quantity)	421.78	608.24	0.00	9600.00
White fish (average quantity)	435.03	480.61	0.00	8537.70
Other fresh fish (average quantity)	552.08	467.96	0.00	8100.00
Real price of salmon	0.13	0.04	0.01	0.75
Real price of fresh white fish	0.14	0.05	0.02	0.71
Real price of other fish	0.12	0.05	0.01	0.68
Real total expenditures on fresh fish	0.98	1.33	0.00	15.57

Table 3. Posterior Summary for Salmon Based on the TMPLN and MGLN

	Frequency			Average Quantity		
Variable	Mean	t-value	Geweke Z	Mean	t-value	Geweke Z
Constant	-0.53	-16.00	1.07	6.66	369.87	-0.71
Price	-0.37	-1.97	-1.17	-3.02	-28.73	0.61
Expenditure	0.09	62.41	1.75	0.07	19.59	-0.25
Time06	-0.05	-3.80	2.10	0.11	9.30	-0.50
Time07·10	-0.07	-0.62	0.91	1.26	10.99	-0.03
Time08·100	-0.06	-0.05	0.15	14.27	12.63	-0.32
Kappa	-	-	-	2.36	136.60	-0.84

Notes: Time06, Time07 and Time08 are annual dummy variables, which takes the value of 1 in the indicated years. For the ease of reading, some of these dummy variables are scaled. Geweke *Z* provides the *Z*-value for a test of stationarity of the Markov chains. Kappa is the shape parameter of the gamma distribution.

Table 4. Posterior Summary for White Fish Based on the TMPLN and MGLN

	Frequency			Average Quantity		
Variable	Mean	t-value	Geweke Z	Mean	t-value	Geweke Z
Constant	-0.12	-5.94	0.35	6.35	402.04	-2.57
Price	-0.69	-6.84	-0.60	-1.38	-14.40	1.70
Expenditure	0.09	62.41	1.75	0.07	19.59	-0.25
Time06	-0.04	-5.66	-0.38	0.06	5.15	0.35
Time07	-0.04	-6.18	0.57	0.05	4.56	0.75
Time08	-0.01	-1.80	-0.16	0.06	5.55	1.35
Kappa	-	-	-	2.06	105.80	1.14

Notes: Time06, Time07 and Time08 are annual dummy variables, which takes the value of 1 in the indicated years. Geweke *Z* provides the *Z*-value for a test of stationarity of the Markov chains.

Table 5. Posterior Summary for Other Fish Based on the TMPLN and MGLN

	Frequency			Average Quantity		
Variable	Mean	t-value	Geweke Z	Mean	t-value	Geweke Z
Constant	0.32	24.00	-0.04	6.43	351.16	1.08
Price	-0.81	-10.55	0.24	-1.22	-10.50	-1.32
Expenditure	0.09	62.41	1.75	0.07	19.59	-0.25
Time06	-0.05	-15.47	-0.41	0.06	5.06	-1.72
Time07	-0.05	-15.42	-0.80	0.11	8.99	-1.80
Time08	-0.07	-18.67	1.18	0.13	10.48	-1.69
Kappa	-	-	-	2.22	123.39	-0.15

Notes: Time06, Time07 and Time08 are annual dummy variables, which takes the value of 1 in the indicated years. Geweke *Z* provides the *Z*-value for a test of stationarity of the Markov chains.

Table 6. Posterior Summary Cross-Equations Covariance Matrix Based on TMPLN and MGLN

	Frequency			Average Quantity		
Variable	Mean	t-value	Geweke Z	Mean	t-value	Geweke Z
Sigma11	0.49	18.65	-0.55	0.29	37.95	0.21
Sigma12	0.17	11.53	-0.05	0.13	24.04	0.90
Sigma13	0.17	14.53	-0.92	0.31	40.83	-0.33
Sigma21	0.17	11.53	-0.05	0.13	24.04	0.90
Sigma22	0.40	23.74	-0.23	0.21	36.01	2.37
Sigma23	0.21	21.79	-0.64	0.14	23.59	0.07
Sigma31	0.17	14.53	-0.92	0.31	40.83	-0.33
Sigma32	0.21	21.79	-0.64	0.14	23.59	0.07
Sigma33	0.26	29.50	-0.90	0.34	38.48	-0.56

Note: Sigma11 Sigma22, and Sigma33 represent the variance of the random effects of demand equations 1-3, respectively. The other Sigma estimates are covariance parameters of the random effects between equations. Geweke *Z* provides the *Z*-value for a test of stationarity of the Markov chains.

Table 7. Own-Price and Total Expenditure Elasticities

	Frequency					
	Salmon		White Fish		Other Fish	
	Estimate	t-value	Estimate	t-value	Estimate	t-value
Own-price	-0.05	-1.97	-0.10	-6.84	-0.10	-10.55
Total expenditure	0.09	62.41	0.09	62.41	0.09	62.41
	Average Quantities					
	Salmon		White Fish		Other Fish	
	Estimate	t-value	Estimate	t-value	Estimate	t-value
Own-price	-0.19	-28.73	-0.11	-14.40	-0.30	-10.50
Total expenditure	0.03	19.59	0.04	19.59	0.07	19.59
	Total Quantities					
	Salmon		White Fish		Other Fish	
	Estimate	t-value	Estimate	t-value	Estimate	t-value
Own-price	-0.24	-8.77	-0.21	-7.50	-0.21	-20.51
Total expenditure	0.12	42.65	0.13	53.85	0.14	81.57

Table 8. Compensated Substitution Effects

Frequency	Estimate
Salmon versus white fish	0.001941
Salmon versus other fish	0.004828
White fish versus other fish	0.011824
Average Quantity	
Salmon versus white fish	38.37530
Salmon versus other fish	48.70118
White fish versus other fish	80.78116

Note: The *t*-value for all estimates are 62.41 due to Equation (30).

Appendix A1. The First-Order Conditions of the Utility Maximization Problem

The consumer maximization problem is given by:

$$\max_{n,q,l}\{v(x(nq),n,l): wT + R = p'x(nq) + wl + wg(n), n > 0, q > 0, l > 0\}. \quad (1)$$

$$\mathcal{L} = v(x(nq),n,l) + \lambda\big[R - p'x(nq) + w\big(T - l - g(n)\big)\big]$$

$$\frac{\partial \mathcal{L}}{\partial q_i} = \frac{\partial v}{\partial x_i} n_i - \lambda p_i n_i = 0 \quad (2)$$

$$\frac{\partial \mathcal{L}}{\partial n_i} = \frac{\partial v}{\partial x_i} q_i + \frac{\partial v}{\partial n_i} - \lambda p_i q_i - \lambda w \frac{\partial g}{\partial n_i} = 0 \quad (3)$$

$$\frac{\partial \mathcal{L}}{\partial l} = \frac{\partial v}{\partial l} - \lambda w = 0 \quad (4)$$

From equations (2) and (4) we have:

$$\frac{\partial v}{\partial x_i}\frac{1}{p_i} = \lambda \quad (5)$$

$$\frac{\partial v}{\partial l}\frac{1}{w} = \lambda \quad (6)$$

Then by inserting Equation (5) into (3) we get:

$$\frac{\partial v}{\partial x_i} q_i + \frac{\partial v}{\partial n_i} - \frac{\partial v}{\partial x_i}\frac{p_i q_i}{p_i} - \lambda w \frac{\partial g}{\partial n_i} = 0$$

$$\frac{\partial v}{\partial x_i} q_i + \frac{\partial v}{\partial n_i} - \frac{\partial v}{\partial x_i} q_i - \lambda w \frac{\partial g}{\partial n_i} = 0$$

$$\frac{\partial v}{\partial n_i} - \lambda w \frac{\partial g}{\partial n_i} = 0$$

$$\frac{\partial v/\partial n_i}{\partial g/\partial n_i} = \lambda w \quad (7)$$

Then it follows from equations (5), (6), and (7) that the necessary conditions for

$$\frac{\partial v/\partial x_i}{\partial v/\partial x_j} = \frac{p_i}{p_j} \tag{8}$$

$$\frac{\partial v/\partial x_i}{\partial v/\partial l} = \frac{p_i}{w} \tag{9}$$

$$\frac{\partial v/\partial n_i}{\partial v/\partial l} = \frac{\partial g}{\partial n_i} \tag{10}$$

$$\frac{\partial v/\partial n_i}{\partial v/\partial n_j} = \frac{\partial g/\partial n_i}{\partial g/\partial n_j} \tag{11}$$

The solution to (1) is then given by three sets of demand equations $n(p, w, R)$, $q(p, w, R)$, and $l(p, w, R)$. Then the total demand for goods is given by $n(p, w, R)\ q(p, w, R) = x(p, w, R)$.

Appendix A2. The Dual Problem

The minimization problem is given by:

$$\min_{n,q,l}\left\{R = p'nq - w\big(T - l - g(n)\big): v(nq, n, l) = v^*, n > 0, q > 0, l > 0\right\} \tag{12}$$

$$\mathcal{L} = p'nq - w\big(T - l - g(n)\big) + \lambda[v^* - v(nq, n, l)]$$

$$\frac{\partial \mathcal{L}}{\partial q_i} = p_i n_i - \lambda \frac{\partial v}{\partial x_i} n_i = 0 \tag{13}$$

$$\frac{\partial \mathcal{L}}{\partial n_i} = p_i q_i + w \frac{\partial g}{\partial n_i} - \lambda \frac{\partial v}{\partial x_i} q_i - \lambda \frac{\partial v}{\partial n_i} = 0 \tag{14}$$

$$\frac{\partial \mathcal{L}}{\partial l} = w - \lambda \frac{\partial v}{\partial l} = 0 \tag{15}$$

From equations (13) and (15) we have:

$$\frac{p_i}{\partial v/\partial x_i} = \lambda \tag{16}$$

$$\frac{w}{\partial v/\partial l} = \lambda \tag{17}$$

Then by inserting Equation (16) into (14) we get:

$$p_i q_i + w\frac{\partial g}{\partial n_i} - \frac{\frac{\partial v}{\partial x_i}}{\frac{\partial v}{\partial x_i}} p_i q_i - \lambda\frac{\partial v}{\partial n_i} = 0$$

$$p_i q_i + w\frac{\partial g}{\partial n_i} - p_i q_i - \lambda\frac{\partial v}{\partial n_i} = 0$$

$$w\frac{\partial g}{\partial n_i} - \lambda\frac{\partial v}{\partial n_i} = 0$$

$$\frac{\partial v/\partial n_i}{\partial g/\partial n_i} = \frac{w}{\lambda} \tag{18}$$

Then it follows from equations (16), (17), and (18) that the necessary conditions for minimizing (12) are given by:

$$\frac{\partial v/\partial x_i}{\partial v/\partial x_j} = \frac{p_i}{p_j} \tag{19}$$

$$\frac{\partial v/\partial x_i}{\partial v/\partial l} = \frac{p_i}{w} \tag{20}$$

$$\frac{\partial v/\partial n_i}{\partial v/\partial l} = \frac{\partial g}{\partial n_i} \tag{21}$$

$$\frac{\partial v/\partial n_i}{\partial v/\partial n_j} = \frac{\partial g/\partial n_i}{\partial g/\partial n_j} \tag{22}$$

The solution to Equation (12) is given by three sets of demand equations $n^c(p, w, v^*)$, $q^c(p, w, v^*)$, and $l^c(p, w, v^*)$. Then total quantity demanded is given by $n^c(p, w, v^*)q^c(p, w, v^*) = x^c(p, w, v^*)$.

Appendix A3. Proofs of Homogeneity, Symmetry, and Negativity

Proof of Proposition (1.1)

i) Reorganize the constraint in Equation (1) and for any scalar $\rho > 0$, we have:

$$\{(n, q, l)' \in \mathbb{R}_+^{2M+1}: \rho R \geq \rho p'x - \rho w\big(T - l - g(n)\big), x \equiv nq\}$$

$$= \{(n, q, l)' \in \mathbb{R}_+^{2M+1}: R \geq p'x - w\big(T - l - g(n)\big), x \equiv nq\}$$

Thus, when p, w, and R increase by the same percentage ρ, the choice set and the objective function are unchanged and, therefore, the optimal choices of n, q, and l are unchanged.Q.E.D.

ii) For any scalar $\rho > 0$,

$$\{(n, x, l)' \in \mathbb{R}_+^{2M+1}: \rho R \geq \rho p'x - \rho w\big(T - l - g(n)\big), x \equiv nq\}$$

$$= \{(n, x, l)' \in \mathbb{R}_+^{2M+1}: R \geq p'x - w\big(T - l - g(n)\big), x \equiv nq\}$$

Q.E.D.

Proof of Proposition (1.2)

First, we need to show that the expenditure function $c(p, w, v^*)$ associated with Equation (1) is concave in p and w. For concavity, fix the utility level at $\bar{v}$, and let $p'' = \rho p + (1-\rho)p'$ and $w'' = \rho w + (1-\rho)w'$ for $\rho \in [0,1]$. Suppose that x'', n'' and l'' are optimal solutions to the expenditure minimization problem when prices are p'' and wages are w''. If so,

$$c(p'', w'', \bar{v}) = p''x'' - w''\big(T - l'' - g(n'')\big)$$

$$= \rho p x'' + (1-\rho)p'x'' - \rho w\big(T - l'' - g(n'')\big) - (1-\rho)w'\big(T - l'' - g(n'')\big)$$

$$= \rho\big[px'' - w\big(T - l'' - g(n'')\big)\big] + (1-\rho)\big[p'x'' - w'\big(T - l'' - g(n'')\big)\big]$$

$$\geq \rho c(p, w, \bar{v}) + (1-\rho)c(p', w', \bar{v})$$

where the inequality follows from $v(x'', n'', l'') \geq \bar{v}$. The definition of the expenditure function imply that $px'' - w(T - l'' - g(n'')) \geq c(p, w, \bar{v})$ and $p'x'' - w'(T - l'' - g(n'')) \geq c(p', w', \bar{v})$.

Next, we need to show that for all p and v^*, the compensated demand $x^c(p, w, v^*)$ is the derivative vector of the expenditure function with respect to p. That is $x^c(p, w, v^*) = \partial c(p, w, v^*)/\partial p_i$ for all $i = 1, \dots, M$. This follows directly from the Envelope Theorem. Thus,

$$\frac{\partial c(p, w, v^*)}{\partial p_i} = \frac{\partial \mathcal{L}(x^c, n^c, l^c)}{\partial p_i} = x^c(p, w, v^*).$$

The second derivative of the expenditure function gives

$$\frac{\partial^2 c(p, w, v^*)}{\partial p_i^2} = \frac{\partial x^c(p, w, v^*)}{\partial p_i}.$$

Since the expenditure function is concave, the matrix of compensated substitution effects for $x^c(p, w, v^*)$, is symmetric and negative semidefinite. Q.E.D.

Proof of Proposition (1.3)

This is a corollary of the previous result follows from the identity $n^c q^c \equiv x^c$. From the Envelope Theorem we have $\partial c(p, w, v^*)/\partial p_i = \partial \mathcal{L}(x^c, n^c, l^c)/\partial p_i = n^c q^c$. Then taking the second derivative we get $\partial^2 c(p, w, v^*)/\partial p_i^2 = (\partial n^c/\partial p_i) q_i + (\partial q^c/\partial p_i) n_i$. Thus, following from the concavity of the expenditure function, the matrix of compensated substitution effect for the product $n_i^c(p, w, v^*) q_i^c(p, w, v^*) \forall i = 1,2, \dots, M$ or $NQ_D^c = (\partial n^c/\partial p') q_i + (\partial q^c/\partial p') n_i$ is symmetric and negative semidefinite. Q.E.D.

Appendix A4. Conditional Means and Marginal Effects

The conditional mean of n_i and q_i are given as follows:

$$\mathrm{E}(n_i|\beta_i, C, b_{Ni}) \tag{23}$$

and

$$\mathrm{E}\big(q_i\big|\alpha_i, C, b_{Qi}\big) = \Pr(N > 0|\beta_i, C)\mathrm{E}\big(q_i\big|\alpha_i, C, n_i > 0, b_{Qi}\big). \tag{24}$$

The marginal effects of $\mathrm{E}\big(q_i\big|\alpha_i, C, b_{Qi}\big)$ are given by:

$$\frac{\partial \mathrm{E}\big(q_i\big|\alpha_i, C, b_{Qi}\big)}{\partial C_i} = \frac{\partial \mathrm{E}\big(q_i\big|\alpha_i, C, n_i > 0, b_{Qi}\big)}{\partial C_i}\Pr(N > 0|\beta_i, C). \tag{25}$$

Paper 2

Frequency and Time in Recreational Demand

Arnar Buason[1], Kristin Eiriksdottir[2], Dadi Kristofersson[2], and Kyrre Rickertsen[1]

[1] School of Economics and Business, Norwegian University of Life Sciences, P.O. Box 5003, 1432 Ås, Norway. Corresponding author is Arnar Buason: arnar.buason@nmbu.no
[2] Department of Economics, University of Iceland, Reykjavik, Iceland.

Abstract

We develop a travel cost model where the consumer jointly chooses the number of visits to a recreational site and how much time to spend on-site. Based on the model we derive the corresponding willingness to pay (WTP) for access estimates that accounts for not only the changes in the price of trips but also for the price of time spent on-site and the substitution and income effects. We show that under certain conditions the WTP of standard single site model is a special case of our measure and provide a context as to why the standard single site model can in some cases, especially when it comes to urban parks, provide biased welfare estimates. Furthermore, we introduce a two-part hurdle model with non-zero correlation of trips and time spent on-site to estimate the joint recreational decision. We assume that the number of trips are generated from a Poisson log-normal distribution and the data generating process for time spent on site is a gamma log-normal distribution. To demonstrate the model's usefulness, we apply it to data gathered on-site in Heiðmörk, an urban park in Iceland. Compared to the welfare estimates of the standard single site travel cost model the results derived from our model indicate a significant downward bias of welfare estimate using the standard approach.

JEL codes: Q51, Q26, C51

Key words: Travel cost model, endogenous on-site time, willingness to pay, urban parks, Poisson, gamma, hurdle model, mixture distribution

1. Introduction

In the recreational demand literature, time spent on-site is usually treated as an exogenous variable that is constant across individuals and sites (e.g., Creel and Loomis 1990, Egan and Herriges 2006, Hynes and Greene 2013). If time spent on-site does not vary greatly for the same individual between recreational sites or the variation in time spent on-site is small compared with the variation in travel costs, this endogeneity may be a minor problem.

However, endogeneity of time becomes increasingly more important when the cost of a trip is negligible compared with the opportunity cost of time spent on-site, which is often the case for open access urban parks. In such cases, it is a strong assumption that an individual derives the same benefit from a trip that lasts fifteen minutes as from a trip that lasts five hours. Consequently, welfare estimates that do not account for the endogeneity of the opportunity cost of the time spent on-site are likely to be biased downwards compared with the true welfare from visiting the recreational area.

Urban parks are generally large vegetated open areas that have been reserved for public use in local zoning plans (Konijnendijk et al. 2013). Historically, some urban parks were used as grazing land for livestock, e.g. Central Park in New York and Boston Common in Boston. Today urban parks have generally been designed by landscape architects and are maintained by local authorities to provide diverse recreational opportunities and to promote health and social well-being in urbanized areas. Urban parks will often include playgrounds for children; picnic areas; trails for hiking, running and walking; restrooms and sports facilities (McCormack et al. 2010, Konijnendijk et al. 2013). It is vital that economic valuation methods accurately measure all the benefits, provided by urban parks. Otherwise development pressure may lead to socially sub-optimal decisions about their long run conservation (More, Stevens and Allen 1988).[23]

Few empirical applications of the single site travel cost model to urban parks exist. Exceptions include Lockwood and Tracy's (1995) application of a zonal travel cost model to data on recreational use of Centennial Park in Sydney and Martinez-Cruz and Sainz-Santamaria's (2015) application of a latent class count data model to data on recreational use

[23] Although the contingent valuation method and the hedonic pricing method have been used to value the benefits from urban open spaces, they fail to capture the direct benefits of users, by not accounting for endogeneity of time spent on site.

of two parks in Mexico City. However, neither Lockwood and Tracy (1995) nor Martinez-Cruz and Sainz-Santamaria (2015) explored the endogeneity of time spent on-site.

Historically, the travel cost recreational demand literature has focused on estimating demand for national parks (Trice and Wood 1958, Clawson 1959, Martínez-Espineira and Amoako-Tuffour 2008), hunting sites (Creel and Loomis 1990), beaches (Hynes and Greene 2013) and lake and river based recreational fishing sites (Grogger and Carson 1991, Egan and Herriges 2006). These sites usually differ from urban parks when it comes to the size and the variation of the users' travel costs. The lack of variation in the travel costs of users of urban parks can make it extremely hard to estimate and trace out a demand curve for trips based on those costs, which can result in a failure to provide welfare estimates.[24] When travel cost cannot explain the variation in trips between individuals, other factors, including time spent on-site, must be dominant factors behind recreational demand.[25]

McConnell (1992) and Larson (1993) derived recreational demand models with endogenously determined on-site time. McConnell (1992) shows that by modelling the demand for trips and the length of stay for each trip with a basic model[26] the standard welfare estimates hold under endogenous on-site time. Thus, implying that only the quantity of trips and not the quantity of time spent on-site is relevant for welfare estimation. Larson (1993) accounts for time spent on-site by assuming recreationists jointly choose the number of trips and total duration of recreation.[27] However, Larson (1993) does not consider welfare calculation under his duration specification, but focuses on analyzing the scarcity value of

[24] Conversely, Smith and Kopp (1980) found that with increasing distance from the recreational site, the spatial limitations of the method can impact welfare estimates substantilly.

[25] Empirical estimation of welfare estimates is further complicated when urban parks in addition are tourist attractions such as Central Park. Not only is it difficult to allot travel cost for a multipurpose trip, but there are at least two latent demand functions behind the recreational demand that fundamentally differ from each other, one for locals and another for out of town visitors. The locals' demand curve is likely to be relatively flat in travel cost while the visitors' demand curve will be much steeper and more in line with what is seen for national parks.

[26] McConnell (1992) assumes utility is a function of trips, x, time spent on-site, t, and a Hicksian bundle, z: $u(x, t, z)$.

[27] Larson (1993) defines the utility function as $u(d, r, z)$ where d is total days spent on-site, r is trips and z a Hicksian bundle.

time. More recently, Hellström (2006) studied the joint choice of the number of leisure trips and nights spent on-site by Swedish households, by estimating a bivariate count data model, and found evidence of positive cross-price elasticities, implying that trips and nights are substitutes.

This paper adds to the literature on time spent on-site in three ways: Firstly, we define the recreational good as a function of two components; trips taken to the recreational site and time spent on-site. We do this because intrinsically time spent on-site must be a fundamental part of the recreational experience and therefore the quantity measure of recreation, just as trips to the grocery store are not a complete measure of the quantity of milk an individual buys. Our model essentially combines McConnell's (1992) single site model with endogenous on-site time and nonlinear budget constraint with Larson's (1993) duration demand specification in which the individual jointly chooses the number of trips to a given site and the time to spend on-site. As a result, we provide an estimate for willingness to pay (WTP) for access that is consistent with utility theory and accounts for changes in the price of time spent on-site. Our WTP estimate can vary greatly from standard single site WTP estimate, depending on e.g. the substitution pattern between trips and time spent on-site, but will under certain conditions provide the same measure. Thus, it provides a more general estimate for WTP for access which encompasses the standard single site WTP for access.

Secondly, we provide an intuitively appealing and theory-consistent method of estimating the joint model of trips taken and time spent on-site by estimating a hurdle model, where time spent on site is only observed if a trip is observed. A stochastic process that depends on the structure of the underlying demand functions as well as the sampling procedure is assumed to have generated the data for this joint recreational demand decision. Since the frequency part of recreational demand takes the form of non-negative discrete integers, it is modelled by a count data model as is customary in the single site recreational demand literature (e.g., Shaw 1988, Creel and Loomis 1990, Grogger and Carson 1991, Shonkwiler and Shaw 1996, Egan and Herriges 2006, Hynes and Greene 2013). Time spent on-site can only take positive non-zero values and although it is observed discretely, it is generated continuously and is modeled with a gamma model.[28] In addition, time spent on-site is likely to be right skewed due to heterogeneity in the user group with the existence of at least a few heavy users that spend significantly more time on-site than is indicated by the group's overall average. To allow for non-independence of the two parts of the model, normally distributed random effects with a non-zero covariance are introduced to each part. Due to computational challenges, there are only a few articles, which apply hurdle models that account for non-zero correlation between stages in the econometric count data literature (e.g., Winkelmann 2004, Min and Agresti 2005). The two-part model is approximated by Gauss-Hermite integration and estimated with a maximum likelihood procedure, using a Dual Quasi-Newton (DQN) method for the optimization.

Thirdly, to demonstrate the model's usefulness, we apply the model to data gathered on-site at an urban park in Iceland, Heiðmörk. We provide an estimate of WTP for access that accounts for the opportunity cost of time spent on-site. This estimate provides policymakers with valuable information on recreational demand and the economic value of the area, which has faced substantial development pressure through the years. We further compare the estimates form our model with the standard methods and evaluate the consequences of omitting time on welfare estimates.

The remainder of the paper is organized as follows. Section 2 presents the theoretical model and a WTP for access measure that accounts for endogenously determined on-site time. Section 3 presents the econometric model used to estimate the demand for recreation. In

[28] The assumption that the DGP of time spent on site is continuous and not discrete as in Hellström (2006) makes our statistical modelling approach significantly different.

section 4, the background to the study and the data are described. Section 5 provides the results and in section 6 the results and their implications are discussed.

2. Theoretical Model

Assume that the individual obtains utility by using a flow of recreational services, x, and a composite commodity bundle, z, which is normalized at a price $\bar{p}_z = 1$. Contrary to the standard single site travel cost model, time spent on-site is determined endogenously. To experience the recreational services, the individual needs to take a trip to the recreational site and to spend time on-site. Let n be the number of trips and t be the time spent on site. For simplicity, t is assumed fixed between trips for the same individual.[29] Hence, to maximize utility the individual simultaneously chooses the number of trips and how much time to spend on-site. Recreation is therefore a function of trips and time spent on-site, as is shown by the following identity:

$$x \equiv x(n, t). \tag{1}$$

Following Larson (1993), trips can either be a source of utility or disutility to the individual, and the utility function is specified as $u(x(n,t), n, z)$. It is assumed to be quasi-concave and exhibit joint weak complementarity across the number of trips and time spent on-site, i.e. without a trip there is no utility derived from time spent on-site, $\partial u(x(0,t), 0, z)/\partial t = 0$, and without time spent on-site there is no utility derived from a trip, $\partial u(x(n,0), n, z)/\partial n = 0$.

The opportunity cost of travel time and time spent on-site can differ from each other and are not necessarily bound by or equal to the wage rate, as is the case when individuals cannot substitute labor and leisure freely (Bockstael, Strand and Hanemann 1987). Assuming

[29] This assumption is consistent with our data, but might not hold true in all cases.

the opportunity costs of travel time and time spent on-site are exogenous prices, the individual faces the nonlinear budget constraint:

$$Y = z + p_n n + p_t t n, \qquad (2)$$

where Y is income, p_n is the total travel cost per trip and p_t is the price per unit (hour) of time spent on-site. The price, p_n, includes all out-of-pocket costs incurred by a trip, such as the marginal cost of driving or the cost of subway/bus fares as well as the opportunity cost of time spent travelling.[30] For each individual, the travel time and prices in equation (2) are fixed but vary between individuals.

The individual's maximization problem is:

$$\max_{\{n,t,z\}} \quad u(x(n,t),n,z) \quad \text{s.t.} \quad Y = z + p_n n + p_t t n. \qquad (3)$$

The solution to this maximization problem is the indirect utility function, $V(p_n, p_t, Y)$. Although, we define recreation as function of trips and time spent on-site, Roy's identity provides the same results as in McConnell (1992). As shown in section A1 of the Appendix, Roy's identity provides the Marshallian demand for trips:

$$n(p_n, p_t, Y) = -\frac{\partial V/\partial p_n}{\partial V/\partial Y} = n, \qquad (4)$$

but does not directly provide the demand for time spent on-site since:

$$t(p_n, p_t, Y) \neq -\frac{\partial V/\partial p_t}{\partial V/\partial Y} = nt. \qquad (5)$$

The optimal value for time spent on-site is given by (McConnell 1992):

$$t(p_n, p_t, Y) = \frac{\partial V/\partial p_t}{\partial V/\partial p_n} = t. \qquad (6)$$

[30] Total travel cost per trip can be broken down as follows: $p_n = p_c + p_\tau \tau$, where p_c is all out of-pocket costs incurred by a trip, τ is travel time and p_τ is the price per unit (hour) of travel time.

In terms of total duration, which is how Larson (1993) defined the recreational good x, equation (5) provides an intuitive specification for the demand of recreation in identity (1) as:

$$x \equiv nt. \tag{7}$$

Thereby, Roy's identity provides the demand for the total duration of recreation $x(p_n, p_t, Y) = nt$ in equation (5).

The interpretation of the first-order conditions (FOCs) of equation (3) are derived in section A2 of the Appendix[31]. Corresponding results for the associated dual cost minimization problem are derived in section A3 of the Appendix.

2.1 Willingness to Pay for Access

The welfare measure of the standard single site travel cost model fails to account for the underlying consumer surplus associated with spending time on-site, which is intuitively implausible and likely to result in downwards biased WTP estimates. McConnell's (1992) findings that the standard single site WTP for access provides an unbiased estimate of welfare despite time spent on-site is endogenously determined is merely a result of how he defines the recreational good and the utility function and not a general rule that holds irrespective. The systematic bias of welfare estimates is likely to be substantial for open access urban parks where potentially a large share of the users has negligible travel costs but substantial welfare benefits associated with spending time on site.

The compensating and equivalent variations provide exact measures of welfare gains/losses associated with price changes. Consumer surplus, which is obtained by

[31] As shown in the Appendix, it follows from the first-order conditions (FOCs) of the maximization problem (3) that the solution satisfies: $\frac{\partial u/\partial n}{\partial u/\partial z} = p_n$, $\frac{\partial u/\partial x}{\partial u/\partial z} = p_t$ and $\frac{\partial u/\partial n}{\partial u/\partial x} = \frac{p_n}{p_t}$. The FOCs imply that the marginal cost of recreation is the price of time spent on-site while the marginal cost of trips is the price of trips.

integrating the Marshallian demand curve over a price change, can serve as an approximation of compensating and equivalent variations with known bounds as shown by Willig (1976). An approximation of WTP for access is given by the Marshallian demand integrated over a price change from the current price level, p_{x_0}, to the choke price, $\bar{p}_x$, at which there is no demand for recreation:

$$WTP = \int_{p_{x_0}}^{\bar{p}_x} x dp_x. \tag{8}$$

Obtaining WTP for access in the model presented in equation (3) with the demand specification in equation (7) is not as straightforward as it is in the standard travel cost model, where time spent on-site is treated as an exogenous variable. Because of the nonlinearity of the budget constraint and the definition of the recreational good, the price of the duration of recreation, p_x, depends on the price per trip, p_n, the price per unit of time spent on-site, p_t, as well as on the time spent on-site itself, t. This can be seen from the budget constraint in equation (2), where the last two parts make up the total cost of the duration of recreation:

$$p_x x = p_n n + p_t t n. \tag{9}$$

By utilizing the definition in equation (7), p_x becomes:

$$p_x = \frac{p_n}{t} + p_t. \tag{10}$$

Since p_x is a function of p_n, p_t and t, the total differential, dp_x, in equation (8) depends on the partial derivatives and the differentials of those variables in the following manner:

$$dp_x = \tfrac{1}{t} dp_n + dp_t - \tfrac{1}{t^2} p_n dt. \tag{11}$$

The WTP for access in equation (8) can thus be written as:

$$\int x dp_x = \int n dp_n + \int x dp_t - \int \frac{n p_n}{t} dt. \tag{12}$$

(i) (ii) (iii)

Equation (12) is therefore a Riemann-Stieltjes integral, since p_x is a function and not a variable. Thus, integrating over *x* w.r.t. p_x and disregarding this fact will lead to incorrect WTP estimates. The first part (i) of equation (12) is the area under the Marshallian demand curve for the number of trips, which is the conventional measure of WTP for access in single site travel cost models and represents the recreationists' consumer surplus from taking trips to a recreational area. The second part (ii) is the area under the Marshallian demand curve for the duration of recreation with respect to the price of time spent on-site. Area (ii) represents the recreationists' consumer surplus from spending time on-site for trips taken over the entire season. The final part (iii) of equation (12) is the total travel cost per time unit (hour) integrated over a change in time spent on-site from the current price level to the choke price. The negative sign in front of part (iii) indicates that even though trips and time spent on-site are complementary goods in the sense that one cannot consume the recreational good without consuming them both, they are in fact substitutes for each other once the individual takes part in the recreational experience. Parts (ii) and (iii) in the WTP for access have been ignored in the recreational demand literature.

Equation (12) contains a total differential for t, dt, which is an endogenous variable in the model. The total differential for $t(p_n, p_t, Y)$ is given by:

$$dt = t_{p_n} dp_n + t_{p_t} dp_t + t_Y dY, \tag{13}$$

where $t_{p_n} = \partial t / \partial p_n$ is the marginal demand for time spent on-site when the price of trips increases, $t_{p_t} = \partial t / \partial p_t$ is the marginal demand for time spent on-site when the price of time spent on-site increases and $t_Y = \partial t / \partial Y$ is the marginal demand for time spent on-site when income increases. Therefore, the total differential for p_x in equation (11) can be rewritten in terms of partial differentials of the exogenous variables, with t acting as a scaling factor:

$$dp_x = \frac{1}{t} dp_n + dp_t - \frac{1}{t^2} p_n t_{p_n} dp_n - \frac{1}{t^2} p_n t_{p_t} dp_t - \frac{1}{t^2} p_n t_Y dY. \tag{14}$$

Part (iii) in the WTP for access in equation (12) can thus be written as:

$$-\int \frac{np_n}{t} dt = -\int \frac{np_n}{t} t_{p_n} dp_n - \int \frac{np_n}{t} t_{p_t} dp_t - \int \frac{np_n}{t} t_Y dY. \qquad (15)$$

In section A.4 of the appendix, it is shown that by using equation (15) and the definitions of own-price, cross-price and income elasticities of demand for time spent on-site as provided in table 1, equation (12) can be rewritten as:

$$\int x dp_x = \underset{\text{(a)}}{\int n(1-\epsilon_{t,n}) dp_n} + \underset{\text{(b)}}{\int x\left(1-\frac{p_n}{tp_t}\epsilon_t\right) dp_t} - \underset{\text{(c)}}{\int \frac{np_n}{Y}\epsilon_{t,Y} dY}. \qquad (16)$$

(Table 1 about here)

The WTP for access measure in equation (16) is presented as a function of the elasticities of demand for time spent on-site to highlight its welfare implications.[32] It consists of three distinct parts. The first part (a) is a measure of the consumer surplus associated with taking trips to the site. This measure accounts for the relationship between the demand for time spent on-site and the demand for trips through the cross-price elasticity, $\epsilon_{t,n}$. The more closely related the demand for time spent on-site is to the demand for trips, the larger the consumer surplus with respect to trips will be. Conversely, the less related they are, the smaller the consumer surplus with respect to the number of trips will be. Table 2 provides an overview of the effects of the cross-price elasticity on the size of the consumer surplus.

(Table 2 about here)

Since integration of the demand curve for the number of trips inherently involves changes in its price, p_n, it is useful to examine how those price changes affect the demand for time spent on-site. Figure 1, which has indifference curves, U, and budget lines, Y, in the n, t plane, the demand for trips, d_n, in the n, p_n plane, the demand for time spent on-site, d_t,[33] in

[32] Table 1 provides the means to present equation (16) in a number of ways, e.g. as a function of the income and price elasticities for recreation, x, and in terms diversion ratios.
[33] The demand curves are linear in Figure 1 for demonstrational purposes.

the t, p_t, plane and iso-demand curves for recreation, x, in the p_n, p_t plane, breaks the effects of an increase in p_n down into substitution and income effects. Figure 1 shows that as the price of trips rises the demand curve for t shifts outward because the individual substitutes time spent on-site for trips, a shift from d_t^0 to $d_t^{0\text{-}1}$. The shift is represented by the cross-price elasticity in part (a). The iso-demand curve for x shows the recreationist' opportunities to acquire recreation at different prices and it shifts outwards from x^0 to x^1 as the price for trips increases, which represents a reduction in the level of recreation obtainable. Figure 1 shows the iso-demand curve for x being upward sloping which can be confusing. The lower left quadrant the iso-demand curve is located in has the p_t axis moving inwards rather than outwards which causes the upwards slope. If the iso-demand curve was located in the lower right quadrant, which is the customary presentation of demand curves, it would be downward sloping.

(Figure 1 about here)

The second part (b) in equation (16) is a measure of the consumer surplus associated with spending time on-site for all the trips taken over the entire season. It accounts for the relationship between the demand for the number of trips and the demand for time spent on-site for each trip through a scaling factor, $\left(1 - \frac{p_n}{tp_t}\epsilon_t\right)$, which includes the relative price of a trip, $\frac{p_n}{tp_t}$, and the own-price elasticity of time spent on-site, ϵ_t. Together the relative price of a trip and the own-price elasticity cause a shift of d_n, the demand for trips, that is parallel to the shift from d_t^0 to $d_t^{0\text{-}1}$ in figure 1. The relative price of a trip is measure of the opportunity cost of trips in terms of time spent on-site. The size of part (b) increases, ceteris paribus, as the opportunity cost of a trip measured in time spent on-site increases since recreationists will substitute this time with the number of trips. Furthermore, ceteris paribus, the more elastic the demand for time spent on-site is, the larger the scaling factor for the consumer surplus will be. However, if it is perfectly inelastic, i.e. $\epsilon_t = 0$, or the price of a trip is dwarfed by the opportunity cost of time spent on-site per trip, tp_t, then part (b) collapses into part (ii) in equation (12), which is the consumer surplus with respect to time spent on-site that does not account for the relationship between the number of trips and time spent on-site. There are two reasons why this can occur. Firstly, as is often the case with urban parks, the price of travel is negligible for users who live near the area. For example, the transportation costs are negligible for many local users of Central Park in New York, where a large share of Manhattan's residents live within a mile radius from the park. Secondly, it can occur if the opportunity cost of total time spent on-site is high. For the opportunity cost of total time spent on-site to be able to dwarf the price of the trip, the price of time spent on-site must be relatively high since the daily time spent on-site is bounded by the individual's available time each day. A relatively high price of time spent on-site is not farfetched in affluent metropolitan areas where individuals often need to work long hours six or seven days a week to keep up with their employers' demands. Table 2 provides an overview of the theoretical effects of opportunity cost of a trip and the own-price elasticity of time spent on-site on the size of the consumer surplus.

The third part (c) accounts for the income effects corresponding to the substitution effects in parts (a) and (b). Together the share of income spent on trips, $\frac{np_n}{Y}$, and the income elasticity of demand for time spent on-site, $\epsilon_{t,Y}$, control how far the demand curves shift following price changes. Such a shift in the demand for time spent on-site is shown as the move from $d_t^{0\text{-}1}$ to d_t^1 in figure 1. Other things constant, the more income that is spent on trips or the higher the income elasticity is, assuming time spent on-site is a normal good, the greater will the income effect and hence the lower part (c) be. As is shown in table 2, part (c) becomes zero if the demand for time spent on-site is unaffected by changes in income. In the case of urban parks, the share of income spent on trips is likely to be small for most individuals and therefore part (c) might be insignificant in many cases.

The WTP measure in equation (16) collapses into the WTP for access measure of the standard single site travel cost model, part (i) in equation (12) when the time spent on-site is not a variable in the model. If an individual cannot choose the amount of time spent on-site freely, i.e., it is constant and exogenously determined, the own-price elasticity, ϵ_t, and the income elasticity of time spent on-site, $\epsilon_{t,Y}$, as well as the cross-price elasticity of time spent on-site and trips, $\epsilon_{t,n}$, all become zero. In this case, equation (16) collapses into the first two parts of equation (12), i.e. $\int x dp_x = \int n dp_n + t \int n dp_t$. The presented model will not provide the same estimate for WTP for access as the standard single site WTP measure does[34] unless the price of time spent on-site, p_t, is a constant. It should be noted that this result is not limited to recreational goods, but extends to all flow goods. How much our measure deviates from the conventional measure depends on the substitution pattern between trips and time spent on-site, their relative costs and responsiveness to income changes.

3. Econometric Model

The joint recreational demand decision to be modeled is how many trips, *n*, the individual takes to a recreational area and the length of each visit on-site, *t*. If the number of trips is assumed to be generated from a discrete distribution belonging to the Katz family of distributions (Katz 1965), $f_N(n|B)$ for $n = 0,1,2, \ldots$ where B is a vector of exogenous variables, then the probability of observing $n = 0$ is given by $\Pr(N = 0|B)$. Time spent on-site is only observed if the individual takes a trip to the site, i.e., when $n > 0$. Thus, the data generating process (DGP) for $t|n > 0$ is assumed to follow a continuous distribution defined only over positive real values, $g_{T|n>0}(t|n > 0, B)$ for $t > 0$. Time spent on-site is defined as the following two-part model:

[34] This will still hold true even if the opportunity cost of time spent on-site, tp_t, is added to the travel cost variable in the standard single site model.

$$g_T(t|B) = \begin{pmatrix} \Pr(N = 0|B) & \text{if } t = 0 \\ \Pr(N > 0|B) g_{T|n>0}(t|n > 0, B) & \text{if } t > 0 \end{pmatrix}. \tag{17}$$

As time spent on site is only realized when a trip takes place, $f_N(N = n|B)$ and $g_{T|n>0}(t|n > 0, B)$ are likely to be stochastically correlated. Introducing random effects, ε_i, $i = 1,2$ to each part of the model can accommodate this stochastic correlation. This will result in the two mixture distributions $f_N(N = n|B, \varepsilon_1)$ and $g_{T|n>0}(t|n > 0, B, \varepsilon_2)$. Assuming, the random effects are jointly normal, we have:

$$\varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix} \sim MVN \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix} \right], \tag{18}$$

the conditional mean of *n* and *t* can be specified as:

$$\mathrm{E}(n|B, \varepsilon_1) \tag{19}$$

and

$$\mathrm{E}(t|B, \varepsilon_2) = \Pr(N > 0|B) \mathrm{E}(t|n > 0, B, \varepsilon_2). \tag{20}$$

The marginal effects of $\mathrm{E}(t|B, \varepsilon_2)$ are given by:

$$\frac{\partial \mathrm{E}(t|B, \varepsilon_2)}{\partial B_i} = \frac{\partial \mathrm{E}(t|n > 0, B, \varepsilon_2)}{\partial B_i} \Pr(N > 0|B). \tag{21}$$

Then, the likelihood of the model is given by:

$$L = \int \left[\prod_{i=1}^{M} f_N(n_i|B, \varepsilon_1)\, g_{T|n>0}(t_i|n > 0, B, \varepsilon_2) \right] \vartheta(\varepsilon) d\varepsilon, \tag{22}$$

where $\vartheta(\varepsilon)$ is the normal density function of ε. The probabilities of interest, $\Pr(N = 0|B, \varepsilon_1)$ and $\Pr(N > 0|B, \varepsilon_1)$, can be calculated from the maximum likelihood estimates.

To fit the model in equation (22), the marginal likelihood must be found by integrating out the random effects. However, there does not exist a closed form solution to this integral and numerical maximization methods must be applied. In our case, there is a single random effects term in each equation, and the Gauss-Hermite integral approximation is preferred due to its computational ease. The joint density of our distributions are given by the following expression:

$$\xi(n,t|B,\varepsilon) = \prod_{i=1}^{M} f_N(n_i|B,\varepsilon_1)\, g_{T|n>0}(t_i|n>0,B,\varepsilon_2). \tag{23}$$

The variance covariance matrix of ε is given by $\Sigma = LL'$, where L is the lower triangular Cholesky factor of Σ.[35] We transform ε such that $\varepsilon = \sqrt{2}Lc$, where $c' = (c_1, c_2)$ and $c_i \sim N(0,1)$, $i = 1,2$. Then, as shown in section A6 in the Appendix, using equations (22) and (23) the likelihood can be written as:

$$L = \int \xi(n,t|B,\varepsilon) \frac{1}{|2\pi\Sigma|^{1/2}} \exp\left(-\frac{1}{2}\varepsilon'\Sigma^{-1}\varepsilon\right) d\varepsilon \tag{24}$$

$$= \frac{1}{\pi}\int \xi\left(n,t\middle|B,\sqrt{2}Lc\right)\exp(-c'c)\, dc.$$

Then, the Gauss-Hermite approximation of the likelihood using *m* quadrature points for each dimension of ε, and following the notation in Min and Agresti (2005), is given by:

$$L \approx \frac{1}{\pi}\sum_{l_1=1}^{m} w_{l_1}^{(1)} \sum_{l_2=1}^{m} w_{l_2}^{(2)}\, \xi\left(n,t\middle|B,\sqrt{2}Lc\right), \tag{25}$$

where $c_{l_k}^{(k)}$ are the nodes (i.e., points of evaluation) and $w_{l_k}^{(k)}$ are the weights $(k = 1,2)$ of the Gauss-Hermite integration of order *m*. The nodes *c* and weights *w* have been calculated for the Gauss-Hermite integration, and can be found in tables in Abramovitz and Stegun (1971).

The approximation in equation (25) is then optimized using the Dual Quasi-Newton (DQN) method. The DQN method computes the gradient, but does not need to calculate the second derivatives, since the Hessian is approximated. The DQN updates the Cholesky factor of the approximated Hessian, instead of updating the approximate inverse Hessian as the standard Quasi-Newton method, and this method can save computation time.[36]

---

[35] The Cholesky decomposition of the variance-covariance matrix is given by the following expression: $\Sigma = \begin{bmatrix} L_{11} & 0 \\ L_{21} & L_{22} \end{bmatrix}\begin{bmatrix} L_{11} & L_{21} \\ 0 & L_{22} \end{bmatrix} = \begin{bmatrix} L_{11}^2 & L_{21}L_{11} \\ L_{21}L_{11} & L_{21}^2 + L_{22}^2 \end{bmatrix}$, see for example Cameron and Trivedi (2005).

[36] For a detailed overview of the quasi-Newton methods see for example Fletcher (1987).

3.1 Distribution Assumptions

In order to estimate equation (25), the number of trips, n, is assumed to be generated by a Poisson log-normal mixture distribution $f_N(n|B,\varepsilon_1)$, where the conditional expectation and variance of n is given by $E(n|B,\varepsilon_1) = r(B)\exp(\varepsilon_1) = \lambda$ and $\mathrm{var}(n|B,\varepsilon_1) = \lambda$, where $r(B)$ is defined over $\mathbb{R}^+ \cup \{0\}$.[37] Furthermore, time spent on-site, $t|n > 0$, is assumed to be generated by a gamma log-normal mixture distribution, where the conditional expectation and variance of $t|n > 0$ is given by $E(t|n > 0, B, \varepsilon_2) = k\phi = h(B,\varepsilon_2)$ and $\mathrm{var}(t|n > 0, B, \varepsilon_2) = k\phi^2$, respectively, and $h(B,\varepsilon_1)$ is defined over $\mathbb{R}^+ \cup \{0\}$.[38] Finally, ε is assumed to be distributed jointly normal as in equation (18).[39]

The unconditional distribution of n, with respect to ε_1 is then given by:

$$f_N(n|B) = \int \frac{\exp(-r(B)\exp(\varepsilon_1))(r(B)\exp(\varepsilon_1))^n \exp(-\varepsilon_1^2/2\sigma_{11})}{n!\sqrt{2\pi\sigma_{11}}} d\varepsilon_1, \text{for } n = 0,1,2,\ldots \tag{26}$$

To be able to evaluate equations (20) and (21), the probability of observing $n > 0$, $\Pr(N > 0|B,\varepsilon_1) = 1 - \Pr(N = 0|B)$, is needed. The probability of observing $n = 0$ is given by:

$$\Pr(N = 0|B) = \int \frac{\exp(-r(B)\exp(\varepsilon_1))(r(B)\exp(\varepsilon_1))^0 \exp(-\varepsilon_1^2/2\sigma_{11})}{0!\sqrt{2\pi\sigma_{11}}} d\varepsilon_1 \tag{27}$$

$$= \int \exp(-r(B)\exp(\varepsilon_1)) \frac{1}{\sqrt{2\pi\sigma_{11}}} \exp(-\varepsilon_1^2/2\sigma_{11})\, d\varepsilon_1.$$

However, there does not exist a closed form solution to the integral in equation (27), and it is approximated by using Gauss-Hermite integration. We do the following transformation, $\varepsilon_1 = \sqrt{2\sigma_{11}}c$, where $c \sim N(0,1)$. Then, equation (27) can be rewritten as:

[37] For overviews of discrete distributions and their relationships see for example Cameron and Trivedi (2013) and Johnson et al. (2005).

[38] For a detailed overview of the gamma distributions see for example Johnson et al. (1994).

[39] A possible extension would be to account for endogenous stratification, but our estimation is merely for demonstrational purposes so this additional compliction is left out.

$$\Pr(N = 0|B) = \int \exp\left(-r(B)\exp\left(\sqrt{2\sigma_{11}}c\right)\right)\frac{1}{\sqrt{2\pi}}\exp(-c^2)\,dc \tag{28}$$

$$\approx \frac{1}{\sqrt{2\pi}}\sum_{l=1}^{m} w_l \exp\left(-r(B)\exp\left(\sqrt{2\sigma_{11}}c_l\right)\right),$$

where m is the number of quadrature points, and w_l and c_l are the weights and the nodes of a Gauss-Hermite integration. Usually it is sufficient to use 4-6 quadrature points, since by then the approximation is not changing significantly, however, adding more points of evaluation only increases accuracy. In the following empirical analysis, we use 8 quadrature points. As was stated in the previous subsection, values for w and c can be found in tables in Abramovitz and Stegun (1971).

4. Data and Empirical Specification

Data was gathered on-site in Heiðmörk, an open access urban park on the fringe of the capital area, from the beginning of July 2008 until the end of September 2009. Heiðmörk was chosen as a study site due to its popularity amongst recreationists' in Iceland's capital area and because of the wide range of ecosystem services it provides to the residents of Reykjavík and its neighboring communities (Eiriksdottir et al. 2016). Although nature is in abundance in Iceland, nature areas intended for recreation are relatively scarce within the capital area. Furthermore, in the last few decades the capital area has sprawled in every possible direction making open wilderness increasingly further away from the center of the area. Heiðmörk is by far the largest recreational area in the vicinity of the capital area covering around 3,000 hectares of vegetated areas, lava fields, two lakes, caves and a water basin as well as offering picnic areas, playgrounds and over 40 kilometers of trails for pedestrians and horseback riders. Heiðmörk offers a wide variety of recreational opportunities to its users, including hiking, horseback riding, cross-country skiing, fishing and picking berries and mushrooms.

The park is used extensively year around and its users are heterogeneous with respect to socio-demographics and recreational activities (Eiriksdottir et al. 2016).

The data was gathered on a per group basis and the process resulted in a sample of 2,392 observations with a 67% participation rate, thereof only 1,525 observations were complete without missing values. A comprehensive discussion on the sampling methodology and the survey design as well as the reasons for missing data and how missing data can be handled is provided in Eiriksdottir et al. (2016). The dataset contains variables for the number of trips taken in the previous calendar month, how much time was spent on-site on the sampled occasion, travel mode, group size, round trip distance, an allotted relative importance of the trip in cases of multipurpose trips, the recreational activity undertaken on the sampled occasion as well as socio-economic variables. The socio-economic variables include gender, age, size of household, the number of children in the household, marital status, education, job market participation and annual disposable household income. Data on substitute sites was not gathered due to a lack in alternatives.

Our empirical application is included solely to demonstrate the impact on the WTP estimates from incorrectly specifying the travel cost model, by not allowing for endogenous time spent on site. We therefore estimate a rather simple empirical model, excluding such factors as socio-economic variables. The Marshallian demand functions, $n(p_n, p_t, Y)$ and $t(p_n, p_t, Y)$, are assumed to take the semi-log functional form as is standard in the single site count data recreational demand literature. Income, Y, is omitted from the analysis since the price of time, both time spent travelling and time spent on-site, is represented with a scaling of Y and the inclusion of income would probably cause multicollinearity problems in estimation.[40] The number of trips and time spent on-site are specified as:

[40] A detailed discussion on the theoretical issues of how the opportunity cost of time should be estimated is beyond the scope of this paper.

$$n = \exp(\alpha_0 + \alpha_1 p_n + \alpha_2 p_t + \varepsilon_1) \quad (29)$$

and

$$t = \exp(\beta_0 + \beta_1 p_n + \beta_2 p_t + \varepsilon_2). \quad (30)$$

The demand for recreation is given by:

$$x = nt = \exp\big((\alpha_0 + \beta_0) + (\alpha_1 + \beta_1)p_n + (\alpha_2 + \beta_2)p_t + (\varepsilon_1 + \varepsilon_2)\big). \quad (31)$$

Following Eiriksdottir et al. (2016), the price of time, p_t, whether it was spent travelling or on-site, was based on one third[41] of the hourly wage rate given 1,800 working hours per year. On a group basis, it was computed as:

$$p_{t_i} = (1/3 \cdot income_i / 1800 \cdot adults_i), \quad (32)$$

where $income_i$ was estimated with in interval regression method based on the respondent's reported yearly income category and $adults_i$ is the number of adults in the group travelling together. The price of travel, p_n, on a group basis was calculated as:

$$p_{n_i} = \left(dist_i \cdot cost/km + ttime_i \cdot p_{t_i}\right) \cdot \%ofpurpose_i, \quad (33)$$

where $dist_i$ is the roundtrip distance from the respondent's home, $cost/km$ is the marginal cost of travel based on travel mode, $ttime_i$ is the roundtrip travel time based on travel mode and $\%ofpurpose_i$ is the relative importance of the trip in the case of a multipurpose trip and 1 otherwise. [42]

Based on the 1,525 complete observations in the dataset, 77% of respondents took 5 trips or fewer in the last calendar month and fewer than 2.5% took 20 trips or more, with the average being 4.20 trips. This indicates the presence of heavy users, i.e., users that visit the park substantially more often than the average. The average time spent on-site was 80

[41] A 1/3 of the hourly wage rate is widely accepted as a lower bound of the opportunity cost of time (Parsons 2003).

[42] Although using a self-reported percentage of trip is an imperfect measure of the travel cost it is better than dropping observations that report a multipurpose trip. Dropping the observations would lead to a bias in the WTP measure. It can also be argued that it is nomore flawed than the acceped way to measure the opportunity cost of travel time.

minutes, ranging from 3 minutes to 7.5 hours. On average, there were 1.56 adults in each group travelling together, the average travel cost on a group basis was 620 ISK and the average time cost on a group basis was 880 ISK. Further summary statistics are provided in table 3 and in Eiriksdottir et al. (2016).

(Table 3 about here)

Under the demand specifications given in equations (29) and (30) and given that α_1, $(\alpha_2 + \beta_2)$ and α_2 are all < 0 and $\beta_1 > 0$,[43] the WTP estimate in equation (16) takes the closed form:

$$\int_{p_{x_0}}^{\infty} x dp_x = -\frac{1}{\alpha_1} n - \frac{1}{\alpha_2 + \beta_2} x + \frac{\beta_1}{\alpha_1} n p_n \left(1 - \frac{1}{\alpha_1 p_n}\right) + \frac{\beta_2}{\alpha_2} n p_n, \tag{34}$$

where α_1 is the half-price elasticity of the demand for trips, $(\alpha_2 + \beta_2)$ is the half-price elasticity[44] of the demand for the duration of recreation with respect to the price of time spent on-site, $p_n \left(1 - \frac{1}{\alpha_1 p_n}\right) = p_n \left(1 + \frac{1}{|\epsilon_n|}\right)$ is the same measure as marginal revenue in profit maximization, $\frac{\beta_1}{\alpha_1}$ is the share of marginal consumption diverted to time spent on-site when the price of trips increases and $\frac{\beta_2}{\alpha_2}$ is the lost demand for time spent on-site as a share of marginal consumption resulting from an increase in the price of on-site time.

Homogeneity of degree zero is imposed by dividing both prices with the Icelandic price index. As is shown in A5, there are no symmetry conditions on price effects between *n* and *t*. Although the data was gathered on-site, respondents were asked about trips taken in the previous calendar month. For simplification, we do not correct for endogenous stratification

---

[43] If however, α_1, $(\alpha_2 + \beta_2)$ and α_2 are not < 0 and β_1 is not > 0, then there does not exist a closed form equation for the WTP for access in equation (16).

[44] The half-price elasticity shows a percentages change in the demand of the underlying variable for a unit change in its price.

of the sampling procedure[45] and the data is not truncated from zero since it includes zeros for those who did not take a trip in the previous calendar month.

5. Empirical Results

To highlight the difference between the welfare estimate of our model and the conventional single site travel cost model, where time spent on-site is exogenously determined and assumed to be constant across users, we provide estimation results and WTP estimates for this model. We estimated the standard single site travel cost model under the Poisson assumption and the parameter estimates are shown in table 4. The associated WTP estimates are shown in table 5. The estimated average monthly WTP is significant and around ISK 6,400 per group.[46]

(Table 4 about here)

(Table 5 about here)

Table 6 shows the parameter estimates and associated *t*-values from the Poisson gamma log-normal mixture model. The Poisson log-normal mixture (PLNM) part shows a significant and negative relationship between travel cost and the number of trips taken. Furthermore, there is a negative relationship between the opportunity cost of time and the number of trips taken. These results indicate that when the price of travel increases the users take fewer trips and with increasing cost of time the users take fewer trips. The estimation results from the gamma log-normal mixture (GLNM) part show a significant and positive relationship between travel cost and the average time spent on site. Thus, as travel cost increases individuals take fewer trips, but spend more time on site on each trip. Furthermore, as the opportunity cost of time increases, the time spent on site decreases.

[45] However, Moeltner and Shonkwiler (2010) found that on-site sampling issues carry over across seasons.
[46] The average exchange rate for the year 2015 was 130 ISK/1 USD.

(Table 6 about here)

Table 7 shows the estimates of the elements of the Cholesky matrix and the covariance between the number of trips, *n*, and the time spent on-site, *t*. All estimates are statistically significant, demonstrating the importance of allowing for non-zero covariance between the two stages of the model.

(Table 7 about here)

Table 8 shows the estimated elasticities and associated *t*-values for each part of the model and the combined effect. The own-price elasticity of the number of trips is significant at the 1% level, and it shows that a 1% increase in the cost of travel reduces the frequency of trips by about 31%. The own-price elasticity of time spent on-site is also significant at the 1% level and indicates that a 1% increase in its price reduces the average time spent on-site by about 0.1%. The cross-price elasticity between the number of trips and the cost of time spent on-site is significant at the 1% level and indicates that a 1% increase in the cost of time spent on-site decreases the frequency of trips by about 0.14%, i.e., time spent on-site is a complementary to the number of trips. However, the cross-price elasticity between time spent on-site and the cost of trips indicates that the number of trips is a substitute to average time spent on site. For a 1% increase in the cost of trips, the average time spent on-site increases about 0.26%. Hence, the substitution pattern between trips and time spent on-site is not symmetric.

The combined elasticities show the total effects on the duration of recreation, *x* of changes in travel costs and time cost. The price elasticity with respect to the cost of trips is not significant. The price elasticity for the duration of recreation with respect to the cost of time spent on-site is significant at the 1% level, and shows that a 1% increase in the price of time reduces the total time spent in Heiðmörk by about 0.25%. Thus, the duration of recreation is a normal good with respect to changes in prices.

(Table 8 about here)

In table 9, we provide welfare estimates based on equation (34) These estimates account for the effects of changes in the price of time spent on-site on the demand for total duration of recreation. The estimated average monthly WTP per group is approximately 15,200 ISK, the average monthly WTP per person is approximately 9,800 ISK, and the average WTP per person per hour per trip is approximately 3,100 ISK. The average monthly WTP per group is substantially higher than the 6,400 ISK for the Poisson model in table 5. These results indicate that the standard single site travel cost model underestimates welfare substantially.

(Table 9 about here)

An average monthly WTP per group of 15,200 ISK corresponds to 117 USD and an average WTP per person per hour per trip of 3,100 ISK corresponds to around 24 USD. However, the dollar amounts should be treated with some caution since environmental benefits and WTP estimates are likely to vary from context to context (Kristofersson and Navrud 2005)[47]. Nevertheless, the WTP estimates are similar to values that other researchers have found in recent years. Hynes and Greene (2013) estimated a per trip per person WTP for access to Silverstrand beach outside of Galway, Ireland. They used a panel negative binomial endogenously stratified and truncated count model, and found a WTP of 30.54 EUR, which corresponds to 4,459 ISK.[48] They also estimated a per trip per person WTP for access to Silverstrand beach using a panel negative binomial endogenously stratified and truncated latent class count model, and found a weighted WTP of 16.93 EUR, which corresponds to 2,472 ISK. Martinez-Cruz and Sainz-Santamaria (2015) estimated a per trip per person WTP for access to Desierto de los Leones natural park outside of Mexico City using a negative

---

[47] There is no meaningful way to transfer benefits between societies with different underlying populations and preferences. Furthermore, variations in exchange rates and the currency controls that have been in effect in Iceland since 2008 complicate the tranferability of benefits to other countries.
[48] The average exchange rate for the year 2015 was 146 ISK/1 EUR.

binomial endogenously stratified and truncated latent class count model, and found a WTP of 33 USD for 82% of the users and a WTP of 12 USD for 18% of the users. For the present dataset, Eiriksdottir et al. (2015) found that by including the opportunity cost of time in the travel cost variable and by including time spent on-site as an exogenous variable[49] in a NB2 count model, the estimated average monthly WTP per group was approximately 20,600 ISK. However, none of these papers accounted for endogeneity of time spent on-site.

In our case, the simpler model presented in table 4 succeeded in providing estimates of the WTP values based on the travel costs. However, for a park that is located centrally rather than on the fringe of a city, it may be impossible to provide estimates based on travel costs because they approach zero for many visitors.

6. Conclusions

The travel cost method has limitations for estimating the demand and WTP for access to urban parks. This is especially true when the parks are centrally located and there is a lack of variation in the users' travel costs. In the single site model, the basic assumption is that all users spend the same amount of time on-site, which is a great simplification of behavior. This simplification is less likely to hold true for neighborhood parks or other local parks than for large national or state parks, where an intrinsic amount of time is needed to experience what the park has to offer in one trip. We add to the travel cost literature by allowing time spent on-site to be endogenously determined and reflected in the welfare estimate.

Based on a duration model for recreational demand, we present a WTP measure for access that fully accounts for time spent on-site. This WTP measure differs from the WTP measure of the standard single site travel cost model where time spent on-site is an exogenous variable. We show that when time spent on-site is endogenously determined and

[49] This approach was advocated in McConnell (1992) if time spent on-site is believed to vary across individuals.

the demand for recreation is defined in terms of duration, the magnitude of WTP depends on the substitution pattern between the number of trips and time spent on-site as well as their relative income effects. Furthermore, our WTP measure collapses into the WTP measure of the standard single site model when time spent on-site is not a variable in the model.

Our econometric model is capable of estimating the demand for duration as a two-part model that allows for correlation between the two underlying parts; the decision of how many trips to take and the decision of how much time to spend on-site on each trip. The frequency part is modeled with a Poisson log-normal count model and the length of stay part is modeled with a gamma log-normal model that only allows non-negative values. The likelihood function of this model does not have a closed form solution and is therefore approximated using a Gauss-Hermite integration, and it is optimized with the numerical DQN method. We compare this model's results to the standard single site count model, and find that our model provides substantially higher estimates of WTP than the standard single site model. In the standard single site model, the estimated average monthly WTP per group is approximately 6,400 ISK, however, this estimate increases to 15,200 ISK in our model. This WTP value suggests that Heiðmörk is an integral part of social welfare in Iceland's capital area.

In reality, consumers face a multitude of different recreational choices. Our analysis should therefore be extended to choices between recreational sites to estimate how price changes affect the demand for alternatives and the interaction e.g. between site specific properties, trip cost and time spent on site.

Appendix: A.1 The Relationship between Our Theoretical Model and McConnell's (1992) Model

The individual's maximization problem is:

$$\max_{\{n,t,z\}} \quad u(x(n,t),n,z) \quad \text{s.t.} \quad Y = z + p_n n + p_t t n \tag{1}$$

with the associated Lagrangian function:

$$\mathcal{L} = u(x(n,t),n,z) + \lambda(Y - z - p_n n - p_t t n).$$

The first-order conditions (FOCs) to (3) are given by:

$$\frac{\partial \mathcal{L}}{\partial n} = \frac{\partial u}{\partial x}\frac{\partial x}{\partial n} + \frac{\partial u}{\partial n} - \lambda(p_n + p_t t) = 0, \tag{2}$$

$$\frac{\partial \mathcal{L}}{\partial t} = \frac{\partial u}{\partial x}\frac{\partial x}{\partial t} - \lambda p_t n = 0, \tag{3}$$

$$\frac{\partial \mathcal{L}}{\partial z} = \frac{\partial u}{\partial z} - \lambda = 0, \tag{4}$$

And

$$\frac{\partial \mathcal{L}}{\partial \lambda} = Y - z - p_n n - p_t t n = 0. \tag{5}$$

From equation (4), we have:

$$\frac{\partial u}{\partial z} = \lambda. \tag{6}$$

Therefore, equations (2) and (3) become:

$$\left(\frac{\partial u}{\partial x}\frac{\partial x}{\partial n} + \frac{\partial u}{\partial n}\right)\frac{1}{(p_n + p_t t)} = \frac{\partial u}{\partial z} \tag{7}$$

and

$$\frac{\partial u}{\partial x}\frac{\partial x}{\partial t}\frac{1}{(p_t n)} = \frac{\partial u}{\partial z}. \tag{8}$$

Combining equations (7) and (8) results in the marginal rate of substitution:

$$-\frac{\left(\frac{\partial u}{\partial x}\frac{\partial x}{\partial n} + \frac{\partial u}{\partial n}\right)}{\frac{\partial u}{\partial x}\frac{\partial x}{\partial t}} = -\frac{p_n + p_t t}{p_t n}. \tag{9}$$

The solution to (1) is the indirect utility function:

$$V(p_n, p_t, Y) = u\Big(x\big(n(p_n, p_t, Y), t(p_n, p_t, Y)\big), n(p_n, p_t, Y), z(p_n, p_t, Y)\Big). \tag{10}$$

For simplicity, let us define:

$$V_1 \equiv \partial V/\partial p_n,$$

$$V_2 \equiv \partial V/\partial p_t \tag{11}$$

and

$$V_Y \equiv \partial V/\partial Y.$$

Furthermore, the budget constraint $Y = z + p_n n + p_t tn$ at the optimum can be written as:

$$Y = z(p_n, p_t, Y) + p_n n(p_n, p_t, Y) + p_t n(p_n, p_t, Y) t(p_n, p_t, Y). \tag{12}$$

Roy's Identity for $n(p_n, p_t, Y)$

Differentiate the indirect utility function (10) with respect to the price of a trip to obtain:

$$V_1 = \frac{\partial u}{\partial x}\frac{\partial x}{\partial n}\frac{\partial n}{\partial p_n} + \frac{\partial u}{\partial x}\frac{\partial x}{\partial t}\frac{\partial t}{\partial p_n} + \frac{\partial u}{\partial n}\frac{\partial n}{\partial p_n} + \frac{\partial u}{\partial z}\frac{\partial z}{\partial p_n}. \tag{13}$$

From the FOCs in equations (2)-(4), we have:

$$V_1 = \lambda(p_n + p_t t)\frac{\partial n}{\partial p_n} + \lambda p_t n \frac{\partial t}{\partial p_n} + \lambda \frac{\partial z}{\partial p_n}. \tag{14}$$

The budget constraint differentiated with respect to the price of a trip is:

$$0 = \frac{\partial z}{\partial p_n} + n + p_n \frac{\partial n}{\partial p_n} + p_t t \frac{\partial n}{\partial p_n} + p_t n \frac{\partial t}{\partial p_n}$$

$$-n = (p_n + p_t t)\frac{\partial n}{\partial p_n} + p_t n \frac{\partial t}{\partial p_n} + \frac{\partial z}{\partial p_n}. \tag{15}$$

Therefore, from equations (14) and (15) we have:

$$V_1 = -\lambda n. \tag{16}$$

The indirect utility function (10) differentiated with respect to income gives:

$$V_Y = \frac{\partial u}{\partial x}\frac{\partial x}{\partial n}\frac{\partial n}{\partial Y} + \frac{\partial u}{\partial x}\frac{\partial x}{\partial t}\frac{\partial t}{\partial Y} + \frac{\partial u}{\partial n}\frac{\partial n}{\partial Y} + \frac{\partial u}{\partial z}\frac{\partial z}{\partial Y}. \tag{17}$$

From the FOCs in equations (2)-(4), we have:

$$V_Y = \lambda(p_n + p_t t)\frac{\partial n}{\partial Y} + \lambda p_t n \frac{\partial t}{\partial Y} + \lambda \frac{\partial z}{\partial Y}. \tag{18}$$

The budget constraint differentiated with respect to income is:

$$1 = \frac{\partial z}{\partial Y} + p_n \frac{\partial n}{\partial Y} + p_t t \frac{\partial n}{\partial Y} + p_t n \frac{\partial t}{\partial Y}$$
$$1 = (p_n + p_t t)\frac{\partial n}{\partial Y} + p_t n \frac{\partial t}{\partial Y} + \frac{\partial z}{\partial Y}. \tag{19}$$

From equations (18) and (19), the marginal utility of income is:

$$V_Y = \lambda. \tag{20}$$

By using Roy's identity, the Marshallian demand for the number of trips is:

$$n(p_n, p_t, Y) = -\frac{V_1}{V_Y} = \frac{\lambda n}{\lambda} = n. \tag{21}$$

Roy's Identity for $t(p_n, p_t, Y)$

The indirect utility function differentiated with respect to the price of time spent on-site is:

$$V_2 = \frac{\partial u}{\partial x}\frac{\partial x}{\partial n}\frac{\partial n}{\partial p_t} + \frac{\partial u}{\partial x}\frac{\partial x}{\partial t}\frac{\partial t}{\partial p_t} + \frac{\partial u}{\partial n}\frac{\partial n}{\partial p_t} + \frac{\partial u}{\partial z}\frac{\partial z}{\partial p_t}. \tag{22}$$

From the FOCs in equations (2)-(4), we have:

$$V_2 = \lambda(p_n + p_t t)\frac{\partial n}{\partial p_t} + \lambda p_t n \frac{\partial t}{\partial p_t} + \lambda \frac{\partial z}{\partial p_t}. \tag{23}$$

The budget constraint differentiated with respect to the price of time spent on-site results in:

$$0 = \frac{\partial z}{\partial p_t} + p_n \frac{\partial n}{\partial p_t} + p_t t \frac{\partial n}{\partial p_t} + p_t n \frac{\partial t}{\partial p_t} + nt$$
$$-nt = (p_n + p_t t)\frac{\partial n}{\partial p_t} + p_t n \frac{\partial t}{\partial p_t} + \frac{\partial z}{\partial p_t}. \tag{25}$$

From equations (23) and (24) we have:

$$V_2 = -\lambda n t. \tag{25}$$

Roy's identity is unable to provide the demand for time spent on-site since:

$$t(p_n, p_t, Y) \neq -\frac{V_2}{V_Y} = \frac{\lambda n t}{\lambda} = nt. \tag{26}$$

The optimal value for time spent on-site is given by:

$$t(p_n, p_t, Y) = \frac{V_2}{V_1} = \frac{\lambda nt}{\lambda n} = t. \tag{27}$$

A.2 The Interpretation of the FOCs of the Theoretical Model

Larson (1993) defined recreational experience as:

$$x \equiv nt, \tag{28}$$

which can be interpreted as the total duration of recreation. The maximization problem in equation (1) becomes:

$$\begin{array}{ccc} \max & u(nt, n, z) & \\ \{n, t, z\} & \text{s.t.} & Y = z + p_n n + p_t tn \end{array} \tag{29}$$

with the associated Lagrangian function:

$$\mathcal{L} = u(nt, n, z) + \lambda(Y - z - p_n n - p_t tn).$$

The FOCs to (29) are given by:

$$\frac{\partial \mathcal{L}}{\partial n} = \frac{\partial u}{\partial x} t + \frac{\partial u}{\partial n} - \lambda(p_n + p_t t) = 0, \tag{30}$$

$$\frac{\partial \mathcal{L}}{\partial t} = \frac{\partial u}{\partial x} n - \lambda p_t n = 0, \tag{31}$$

$$\frac{\partial \mathcal{L}}{\partial z} = \frac{\partial u}{\partial z} - \lambda = 0, \tag{32}$$

And

$$\frac{\partial \mathcal{L}}{\partial \lambda} = Y - z - p_n n - p_t tn = 0. \tag{33}$$

From equations (31) and (32), we have:

$$\frac{\partial u}{\partial x} = \lambda p_t \tag{34}$$

and

$$\frac{\partial u}{\partial z} = \lambda. \tag{35}$$

By inserting equation (34) into equation (30), we obtain:

$$\lambda p_t t + \frac{\partial u}{\partial n} - \lambda p_n - \lambda p_t t = 0$$

$$\frac{\partial u}{\partial n} = \lambda p_n. \tag{36}$$

It follows from equations (34), (35) and (36) that the necessary and sufficient conditions for maximum are given by:

$$-\frac{\partial u/\partial n}{\partial u/\partial z} = -p_n, \tag{37}$$

$$-\frac{\partial u/\partial x}{\partial u/\partial z} = -p_t \tag{38}$$

And

$$-\frac{\partial u/\partial n}{\partial u/\partial x} = -\frac{p_n}{p_t}. \tag{39}$$

A.3 The Dual Problem to the Theoretical Model

The individual's expenditure minimization problem is to minimize expenditure given a certain utility level, $\bar{u}$, or:

$$\min_{\{n,t,z\}} \quad z + p_n n + p_t tn \quad \text{s.t.} \quad u(nt, n, z) = \bar{u} \tag{40}$$

with the associated Lagrangian function:

$$\mathcal{L} = z + p_n n + p_t tn - \lambda(u(nt, n, z) - \bar{u}).$$

The FOCs to (40) are given by:

$$\frac{\partial \mathcal{L}}{\partial n} = p_n + p_t t - \lambda\left(\frac{\partial u}{\partial x}t + \frac{\partial u}{\partial n}\right) = 0, \tag{41}$$

$$\frac{\partial \mathcal{L}}{\partial t} = p_t n - \lambda \frac{\partial u}{\partial x} n = 0, \tag{42}$$

$$\frac{\partial \mathcal{L}}{\partial z} = 1 - \lambda \frac{\partial u}{\partial z} = 0 \tag{43}$$

And

$$\frac{\partial \mathcal{L}}{\partial \lambda} = u(nt, n, z) - \bar{u} = 0. \tag{44}$$

From equations (42) and (43), we have:

$$\frac{p_t}{\partial u/\partial x} = \lambda \tag{45}$$

and

$$\frac{1}{\partial u/\partial z} = \lambda. \tag{46}$$

By inserting equation (45) into equation (41), we have:

$$p_n + p_t t - \frac{p_t}{\frac{\partial u}{\partial x}}\left(\frac{\partial u}{\partial x} t + \frac{\partial u}{\partial n}\right) = 0,$$

$$p_n + p_t t - p_t t - p_t \frac{\partial u/\partial n}{\partial u/\partial x} = 0 \tag{47}$$

and

$$-\frac{\partial u/\partial n}{\partial u/\partial x} = -\frac{p_n}{p_t}.$$

It follows from equations (45)-(47) that expenditure minimization provides identical results as utility maximization. The solution to problem (40) is the expenditure function:

$$e(p_n, p_t, \bar{u}) = z(p_n, p_t, \bar{u}) + p_n n(p_n, p_t, \bar{u}) + p_t t(p_n, p_t, \bar{u}) n(p_n, p_t, \bar{u}). \tag{48}$$

Hicksian, or compensated, demand functions are derived by differentiating the expenditure function with respect to prices (Shephard's lemma). The utility function with compensated demand functions is given by:

$$u(n(p_n, p_t, \bar{u}) t(p_n, p_t, \bar{u}), n(p_n, p_t, \bar{u}), z)(p_n, p_t, \bar{u}) = \bar{u}. \tag{49}$$

Shephard's Lemma for $n(p_n, p_t, \bar{u})$

Differentiate the expenditure function with respect to the price of a trip to obtain:

$$\frac{\partial e(p_n, p_t, \bar{u})}{\partial p_n} = (p_n + p_t t)\frac{\partial n}{\partial p_n} + p_t n\frac{\partial t}{\partial p_n} + \frac{\partial z}{\partial p_n} + n. \tag{50}$$

From the FOCs given by equations (41)-(43), we have:

$$\frac{\partial e(p_n, p_t, \bar{u})}{\partial p_n} = \lambda\left(\frac{\partial u}{\partial x}t + \frac{\partial u}{\partial n}\right)\frac{\partial n}{\partial p_n} + \lambda\frac{\partial u}{\partial x}n\frac{\partial t}{\partial p_n} + \lambda\frac{\partial u}{\partial z}\frac{\partial z}{\partial p_n} + n. \tag{51}$$

Differentiate the utility function (49) with respect to the price of a trip, to obtain:

$$\frac{\partial u}{\partial x}t\frac{\partial n}{\partial p_n} + \frac{\partial u}{\partial x}n\frac{\partial t}{\partial p_n} + \frac{\partial u}{\partial n}\frac{\partial n}{\partial p_n} + \frac{\partial u}{\partial z}\frac{\partial z}{\partial p_n} = 0$$

and (52)

$$\left(\frac{\partial u}{\partial x}t + \frac{\partial u}{\partial n}\right)\frac{\partial n}{\partial p_n} + \frac{\partial u}{\partial x}n\frac{\partial t}{\partial p_n} + \frac{\partial u}{\partial z}\frac{\partial z}{\partial p_n} = 0.$$

From equations (51) and (52), Shephard's lemma provides the compensated demand functions for the number of trips:

$$n(p_n, p_t, \bar{u}) = \frac{\partial e(p_n, p_t, \bar{u})}{\partial p_n} = \lambda \cdot 0 + n. \tag{53}$$

Shephard's Lemma for $t(p_n, p_t, \bar{u})$

Differentiate the expenditure function with respect to the price of time spent on-site to obtain:

$$\frac{\partial e(p_n, p_t, \bar{u})}{\partial p_t} = (p_n + p_t t)\frac{\partial n}{\partial p_t} + p_t n\frac{\partial t}{\partial p_t} + \frac{\partial z}{\partial p_t} + nt. \tag{54}$$

From the FOCs in equations (41)-(43), we have:

$$\frac{\partial e(p_n, p_t, \bar{u})}{\partial p_n} = \lambda\left(\frac{\partial u}{\partial x}t + \frac{\partial u}{\partial n}\right)\frac{\partial n}{\partial p_t} + \lambda\frac{\partial u}{\partial x}n\frac{\partial t}{\partial p_t} + \lambda\frac{\partial u}{\partial z}\frac{\partial z}{\partial p_t} + nt. \tag{55}$$

Differentiate the indirect utility function, given by equation (49) with respect to the price of time spent on-site to obtain:

$$\frac{\partial u}{\partial x}t\frac{\partial n}{\partial p_t} + \frac{\partial u}{\partial x}n\frac{\partial t}{\partial p_t} + \frac{\partial u}{\partial n}\frac{\partial n}{\partial p_t} + \frac{\partial u}{\partial z}\frac{\partial z}{\partial p_t} = 0 \tag{56}$$

and

$$\left(\frac{\partial u}{\partial x}t + \frac{\partial u}{\partial n}\right)\frac{\partial n}{\partial p_t} + \frac{\partial u}{\partial x}n\frac{\partial t}{\partial p_t} + \frac{\partial u}{\partial z}\frac{\partial z}{\partial p_t} = 0.$$

By combining equations (55) and (56), we see that Shephard's lemma is unable to provide the compensated demand functions for time spent on-site since:

$$t(p_n, p_t, \bar{u}) \neq \frac{\partial e(p_n, p_t, \bar{u})}{\partial p_n} = \lambda \cdot 0 + nt. \tag{57}$$

However, the derivative of the expenditure function with respect to the price of on-site time provides the compensated demand for the duration of recreation, $x = nt$. Hence, expenditure minimization provides the same results as utility maximization on the indirect utility function in equation (10) and Roy's identity in equations (21) and (26).

A.4 Welfare Estimation

Willingness to pay (WTP) for access is obtained by integrating the Marshallian demand curve over a price change from the current price, p_{x_0}, to the choke price, $\bar{p}_x$:

$$WTP = \int_{p_{x_0}}^{\bar{p}_x} x dp_x. \tag{58}$$

The price of the duration of recreation depends on the price of trips, the price of time spent on-site as well as on time itself. The budget constraint $Y = z + p_n n + p_t tn$ shows that the total cost of recreation is:

$$p_x x = p_n n + p_t tn. \tag{59}$$

By utilizing the definition of demand in equation (28), equation (59) becomes:

$$p_x x = p_n n + p_t x \tag{60}$$

$$p_x = \frac{p_n}{t} + p_t.$$

The total differential of equation (60) is:

$$dp_x = \frac{1}{t}dp_n + dp_t - \frac{1}{t^2}p_n dt, \tag{61}$$

and the WTP for access in equation (58) can be rewritten as:

$$\int_{p_{x_0}}^{\bar{p}_x} x dp_x = \int_{p_{x_0}}^{\bar{p}_x} x\frac{1}{t}dp_n + \int_{p_{x_0}}^{\bar{p}_x} x dp_t - \int_{p_{x_0}}^{\bar{p}_x} x\frac{1}{t^2}p_n dt \tag{62}$$

$$\int_{p_{x_0}}^{\bar{p}_x} x dp_x = \int_{p_{x_0}}^{\bar{p}_x} n dp_n + \int_{p_{x_0}}^{\bar{p}_x} x dp_t - \int_{p_{x_0}}^{\bar{p}_x} \frac{np_n}{t} dt.$$

However, equation (62) still contains a total differential for t, dt, which is an endogenous variable in the model. The total differential for $t(p_n, p_t, Y)$ is given by:

$$dt = t_{p_n}dp_n + t_{p_t}dp_t + t_Y dY. \tag{63}$$

By plugging equation (63) into equation (61), we get:

$$dp_x = \frac{1}{t}dp_n + dp_t - \frac{1}{t^2}p_n t_{p_n} dp_n - \frac{1}{t^2}p_n t_{p_t} dp_t - \frac{1}{t^2}p_n t_Y dY. \tag{64}$$

Therefore, the WTP for access in equation (62) becomes:

$$\int_{p_{x_0}}^{\bar{p}_x} x dp_x = \int_{p_{x_0}}^{\bar{p}_x} n dp_n + \int_{p_{x_0}}^{\bar{p}_x} x dp_t - \int_{p_{x_0}}^{\bar{p}_x} \frac{np_n}{t} t_{p_n} dp_n - \int_{p_{x_0}}^{\bar{p}_x} \frac{np_n}{t} t_{p_t} dp_t$$

$$- \int_{p_{x_0}}^{\bar{p}_x} \frac{np_n}{t} t_Y dY. \tag{65}$$

(a) (b) (c) (d) (e)

Parts (c), (d) and (e) of equation (65) can be rewritten using the following demand elasticities:

	Own-price elasticity	Cross-price elasticity	Income elasticity	(66)
n	$\epsilon_n = \frac{\partial n}{\partial p_n}\frac{p_n}{n}$	$\epsilon_{n,t} = \frac{\partial n}{\partial p_t}\frac{p_t}{n}$	$\epsilon_{n,Y} = \frac{\partial n}{\partial Y}\frac{Y}{n}$	

t	$\epsilon_t = \frac{\partial t}{\partial p_t}\frac{p_t}{t}$	$\epsilon_{t,n} = \frac{\partial t}{\partial p_n}\frac{p_n}{t}$	$\epsilon_{t,Y} = \frac{\partial t}{\partial Y}\frac{Y}{t}$

Thus, part (c) of equation (65) becomes:

$$\int_{p_{x_0}}^{\bar{p}_x} \frac{n}{t} p_n t_{p_n} dp_n = \int_{p_{x_0}}^{\bar{p}_x} n\epsilon_{t,n} dp_n, \tag{67}$$

part (d) becomes

$$\int_{p_{x_0}}^{\bar{p}_x} \frac{n}{t} p_n t_{p_t} dp_t = \int_{p_{x_0}}^{\bar{p}_x} np_n \frac{\epsilon_t}{p_t} dp_t \tag{68}$$

and part (e) becomes

$$\int_{p_{x_0}}^{\bar{p}_x} \frac{n}{t} p_n t_Y dY = \int_{p_{x_0}}^{\bar{p}_x} \frac{np_n}{Y} \epsilon_{t,Y} dY. \tag{69}$$

Therefore, equation (65) can be written as:

$$\int_{p_{x_0}}^{\bar{p}_x} x dp_x = \int_{p_{x_0}}^{\bar{p}_x} n dp_n + \int_{p_{x_0}}^{\bar{p}_x} x dp_t - \int_{p_{x_0}}^{\bar{p}_x} n\epsilon_{t,n} dp_n - \int_{p_{x_0}}^{\bar{p}_x} np_n \frac{\epsilon_t}{p_t} dp_t - \int_{p_{x_0}}^{\bar{p}_x} \frac{np_n}{Y} \epsilon_{t,Y} dY. \tag{70}$$

By collecting the parts with the same total differentials, equation (70) can be rewritten as:

$$\int_{p_{x_0}}^{\bar{p}_x} x dp_x = \int_{p_{x_0}}^{\bar{p}_x} n(1 - \epsilon_{t,n}) dp_n + \int_{p_{x_0}}^{\bar{p}_x} x\left(1 - \frac{p_n}{tp_t}\epsilon_t\right) dp_t - \int_{p_{x_0}}^{\bar{p}_x} \frac{np_n}{Y} \epsilon_{t,Y} dY. \tag{71}$$

If the demand functions $n(p_n, p_t, Y)$ and $t(p_n, p_t, Y)$ have a semi-log functional form, as is typically assumed in the recreational demand literature, and the price of time spent on-site is assumed to be proportional to income, i.e. $p_t = \delta Y$, the demand functions are:

$$n = \exp(\alpha_0 + \alpha_1 p_n + \alpha_2 p_t) \tag{72}$$

and

$$t = \exp(\beta_0 + \beta_1 p_n + \beta_2 p_t).$$

The associated demand for duration of recreation, x, is:

$$\begin{aligned} x = nt &= \exp(\alpha_0 + \alpha_1 p_n + \alpha_2 p_t)\exp(\beta_0 + \beta_1 p_n + \beta_2 p_t) \\ &= \exp((\alpha_0 + \alpha_1 p_n + \alpha_2 p_t) + (\beta_0 + \beta_1 p_n + \beta_2 p_t)) \\ &= \exp((\alpha_0 + \beta_0) + (\alpha_1 + \beta_1) p_n + (\alpha_2 + \beta_2) p_t). \end{aligned} \tag{73}$$

The associated own-price and cross-price elasticities are:

	Own-price elasticity	Cross-price elasticity
n	$\epsilon_n = \alpha_1 p_n$	$\epsilon_{n,t} = \alpha_2 p_t$
t	$\epsilon_t = \beta_2 p_t$	$\epsilon_{t,n} = \beta_1 p_n$

(74)

With the semi-log functional form for demand, equation (72), the choke price $\bar{p}_x$ becomes infinity and equation (70) becomes:

$$\begin{aligned} \int_{p_{x_0}}^{\infty} x\,dp_x = &\int_{p_{x_0}}^{\infty} \exp(\alpha_0 + \alpha_1 p_n + \alpha_2 p_t)\,dp_n \\ &+ \int_{p_{x_0}}^{\infty} \exp((\alpha_0 + \beta_0) + (\alpha_1 + \beta_1) p_n + (\alpha_2 + \beta_2) p_t)\,dp_t \\ &- \beta_1 \int_{p_{x_0}}^{\infty} \exp(\alpha_0 + \alpha_1 p_n + \alpha_2 p_t)\,p_n dp_n \\ &- \beta_2 p_n \int_{p_{x_0}}^{\infty} \exp(\alpha_0 + \alpha_1 p_n + \alpha_2 p_t)\,dp_t. \end{aligned} \tag{75}$$

Integration by parts gives: $\int \exp(\alpha_0 + \alpha_1 p_n + \alpha_2 p_t)\,p_n dp_n = p_n \frac{\exp(\alpha_0+\alpha_1 p_n+\alpha_2 p_t)}{\alpha_1} - \frac{1}{\alpha_1}\int \exp(\alpha_0 + \alpha_1 p_n + \alpha_2 p_t) dp_n = p_n \frac{\exp(\alpha_0+\alpha_1 p_n+\alpha_2 p_t)}{\alpha_1} - \frac{\exp(\alpha_0+\alpha_1 p_n+\alpha_2 p_t)}{{\alpha_1}^2}$ and equation (75) becomes:

$$\int_{p_{x_0}}^{\infty} x dp_x = \frac{1}{\alpha_1}\exp(\alpha_0 + \alpha_1 p_n + \alpha_2 p_t)\Big|_{p_n = p_{x_0}}^{p_n \to \infty} \tag{76}$$

$$+ \frac{1}{\alpha_2 + \beta_2}\exp((\alpha_0 + \beta_0) + (\alpha_1 + \beta_1)p_n + (\alpha_2 + \beta_2)p_t)\Big|_{p_t = p_{x_0}}^{p_t \to \infty}$$

$$- \beta_1\left(\frac{\exp(\alpha_0 + \alpha_1 p_n + \alpha_2 p_t)}{\alpha_1}(p_n - 1/\alpha_1)\Big|_{p_n = p_{x_0}}^{p_n \to \infty}\right)$$

$$- \beta_2 p_n\left(\frac{\exp(\alpha_0 + \alpha_1 p_n + \alpha_2 p_t)}{\alpha_2}\Big|_{p_t = p_{x_0}}^{p_t \to \infty}\right).$$

Finally, if α_1, $(\alpha_2 + \beta_2)$ and α_2 are all < 0 and $\beta_1 > 0$, then equation (76) becomes:

$$\int_{p_{x_0}}^{\infty} x dp_x = -\frac{1}{\alpha_1}n - \frac{1}{\alpha_2 + \beta_2}x + \frac{\beta_1}{\alpha_1}n(p_n - 1/\alpha_1) + \frac{\beta_2}{\alpha_2}np_n$$

$$\int_{p_{x_0}}^{\infty} x dp_x = -\frac{1}{\alpha_1}n - \frac{1}{\alpha_2 + \beta_2}x + \frac{\beta_1}{\alpha_1}np_n\left(1 - \frac{1}{\alpha_1 p_n}\right) + \frac{\beta_2}{\alpha_2}np_n. \tag{77}$$

A5. Symmetry Conditions

It can be shown that the matrix of compensated substitution effects for $x(p_n, p_t, \bar{u})$ and $n(p_n, p_t, \bar{u})$ are symmetric and negative semidefinite, but the same does not hold for the matrix of compensated substitution effect for $t(p_n, p_t, \bar{u})$, nor between $t(p_n, p_t, \bar{u})$ and $n(p_n, p_t, \bar{u})$. First we need to show that $e(p_n, p_t, \bar{u})$ is concave in p_n and p_t. For concavity, fix the utility level at $\bar{u}$, and let $p_n'' = \rho p_n + (1-\rho)p_n'$ and $p_t'' = \rho p_t + (1-\rho)p_t'$ for $\rho \in [0,1]$. Suppose t'', n'', $x'' = t''n''$, and z'' are optimal solutions to the expenditure problem, in A3 equation (40), when prices are p_n'' and p_t''. If so,

$$\begin{aligned} e(p_n'', p_t'', \bar{u}) &= z'' + p_n''n'' + p_t''x'' \\ &= z'' + \rho p_n n'' + (1-\rho)p_n'n'' + \rho p_t x'' + (1-\rho)p_t'x'' \end{aligned} \tag{78}$$

$$= z'' + \rho[p_n n'' + p_t x''] + (1-\rho)[p_n' n'' + p_t' x'']$$

$$\geq \rho e(p_n, p_t, \bar{u}) + (1-\rho) e(p_n', p_t', \bar{u}),$$

where the inequality follows from $u(x'', n'', z'') \geq \bar{u}$, and the definition of the expenditure function, which implies that

$$z'' + p_n n'' + p_t x'' \geq e(p_n, p_t, \bar{u}), \tag{79}$$

and

$$z' + p_n n' + p_t x' \geq e(p_n', p_t', \bar{u}). \tag{80}$$

Next we need to show that for all p_t and $\bar{u}$, the compensated demand $x(p_n, p_t, \bar{u})$ is the derivative vector of the expenditure function w.r.t. p_t. That is $x(p_n, p_t, \bar{u}) = \partial e(p_n, p_t, \bar{u})/\partial p_t$. This follows directly from the Envelope theorem. Thus,

$$\frac{\partial e(p_n, p_t, \bar{u})}{\partial p_t} = \frac{\partial \mathcal{L}(x, n, z)}{\partial p_t} = x(p_n, p_t, \bar{u}). \tag{81}$$

The second derivative of the expenditure function gives

$$\frac{\partial^2 e(p_n, p_t, \bar{u})}{\partial p_t^2} = \frac{\partial x(p_n, p_t, \bar{u})}{\partial p_t}. \tag{82}$$

Since the expenditure function is concave the matrix of compensated substitution effects for $x(p_n, p_t, \bar{u})$ is symmetric and negative semidefinite. It then follows that the matrix of compensated substitution effects for $t(p_n, p_t, \bar{u})$ is not symmetric and negative semidefinite, nor the substitution effects between $t(p_n, p_t, \bar{u})$ and $n(p_n, p_t, \bar{u})$, but are jointly symmetric through the identity $x \equiv nt$ and equation (82).

Restrictions on the matrix of compensated substitution effects for $n(p_n, p_t, \bar{u})$ can be found using the same argument as above. The compensated demand $n(p_n, p_t, \bar{u})$ is the derivative vector of the expenditure function w.r.t. p_n. That is $n(p_n, p_t, \bar{u}) = \partial e(p_n, p_t, \bar{u})/\partial p_n$. This follows directly from the Envelope theorem. Thus,

$$\frac{\partial e(p_n, p_t, \bar{u})}{\partial p_n} = \frac{\partial \mathcal{L}(x, n, z)}{\partial p_n} = n(p_n, p_t, \bar{u}). \tag{83}$$

The second derivative of the expenditure function gives

$$\frac{\partial^2 e(p_n, p_t, \bar{u})}{\partial p_t^2} = \frac{\partial n(p_n, p_t, \bar{u})}{\partial p_n}. \tag{84}$$

Since the expenditure function is concave the matrix of substitution effects for $n(p_n, p_t, \bar{u})$ is symmetric and negative semidefinite. Q.E.D.

It then follows from the derivation above that in a single site model there are no symmetry restriction imposed.

A6. Econometric Model

Econometric Model ML Optimization

Simplification of the log-likelihood using Cholesky factorization:

$$\begin{aligned} L &= \int \xi(n, t|B, \varepsilon) \frac{1}{|2\pi\Sigma|^{1/2}} \exp\left(-\frac{1}{2}\varepsilon^{\mathrm{T}}\Sigma^{-1}\varepsilon\right) d\varepsilon \\ &= \frac{1}{\pi}\int \xi\left(n, t\middle|B, \sqrt{2}Lc\right)\exp\left(-\frac{1}{2}\left(\sqrt{2}Lc\right)^{\mathrm{T}}\Sigma^{-1}\sqrt{2}Lc\right) dc \\ &= \frac{1}{\pi}\int \xi\left(n, t\middle|B, \sqrt{2}Lc\right)\exp(-c^T L^T \Sigma^{-1} Lc)\, dc \\ &= \frac{1}{\pi}\int \xi\left(n, t\middle|B, \sqrt{2}Lc\right)\exp(-c'c)\, dc. \end{aligned} \tag{85}$$

Econometric Model Empirical Calculations

To calculate the variance for the WTP estimate, we differentiate equation (77) by the price parameters:

$$\begin{aligned} \frac{\partial WTP_i}{\partial \alpha_1} &= \frac{n_i}{\alpha_1^2} - \frac{\beta_1 n_i p_n}{\alpha_1^2} + \frac{2\beta_1 n_i}{\alpha_1^3} = W_{\alpha_1}, \\ \frac{\partial WTP_i}{\partial \alpha_2} &= \frac{x_i}{(\alpha_2 + \beta_2)^2} - \frac{\beta_2 n_i p_n}{\alpha_2^2} = W_{\alpha_2}, \end{aligned} \tag{86}$$

$$\frac{\partial WTP_i}{\partial \beta_1} = \frac{n_i p_n}{\alpha_1} - \frac{n_i}{\alpha_1^2} = W_{\beta_1}$$

and

$$\frac{\partial WTP_i}{\partial \beta_2} = \frac{x_i}{(\alpha_2 + \beta_2)^2} + \frac{n_i p_n}{\alpha_2} = W_{\beta_2}.$$

The vector of averages of these derivatives over all groups is $\begin{pmatrix}\widehat{W}_{\alpha_1} & \widehat{W}_{\alpha_2} & \widehat{W}_{\beta_1} & \widehat{W}_{\beta_2}\end{pmatrix}$. By using the delta method, the variance of the WTP estimate is given by:

$$\begin{pmatrix}\widehat{W}_{\alpha_1} & \widehat{W}_{\alpha_2} & \widehat{W}_{\beta_1} & \widehat{W}_{\beta_2}\end{pmatrix}\begin{pmatrix}\theta_{\alpha_1\alpha_1} & \theta_{\alpha_1\alpha_2} & \theta_{\alpha_1\beta_1} & \theta_{\alpha_1\beta_2}\\ \theta_{\alpha_2\alpha_1} & \theta_{\alpha_2\alpha_2} & \theta_{\alpha_2\beta_1} & \theta_{\alpha_2\beta_2}\\ \theta_{\beta_1\alpha_1} & \theta_{\beta_1\alpha_2} & \theta_{\beta_1\beta_1} & \theta_{\beta_1\beta_2}\\ \theta_{\beta_2\alpha_1} & \theta_{\beta_2\alpha_2} & \theta_{\beta_2\beta_1} & \theta_{\beta_2\beta_2}\end{pmatrix}\begin{pmatrix}\widehat{W}_{\alpha_1}\\ \widehat{W}_{\alpha_2}\\ \widehat{W}_{\beta_1}\\ \widehat{W}_{\beta_2}\end{pmatrix}. \tag{87}$$

References

Abramowitz, Milton, and Irene A. Stegun. (1971). *Handbook of Mathematical Functions*. New York: Dover Press.

Bockstael, Nancy E., Strand, Ivar E., and W. Michael Hanemann. (1987). Time and recreational demand. *American Journal of Agricultural Economics* 69 (2): 293-302.

Cameron, A. Colin, and Pravin K. (2013). *Regression Analysis of Count Data*. Cambridge: Cambridge University Press.

Cameron, A. Colin, and Pravin K. Trivedi. (2005). *Microeconometrics: Methods and Applications*. New York: Cambridge University Press.

Clawson, Marion. (1959). *Methods of Measuring the Demand for and Value of Outdoor Recreation.* Washington: Resources for the Future.

Creel, Michael D., and John B. Loomis. (1990). Theoretical and empirical advantages of truncated count data estimators for analysis of deer hunting in California. *American Journal of Agricultural Economics* 72 (2): 434–441.

Egan, Kevin, and Joseph Herriges. (2006). Multivariate count data regression models with individual panel data from an on-site sample. *Journal of Environmental Economics and Management* 52 (2): 567-581.

Eiriksdottir, Kristin, Arnason, Ragnar, Davidsdottir, Brynhildur, and Dadi Kristofersson. (2017). Handling missing values in travel cost models: An application of multiple imputation and a latent class count data model. *Forthcoming as working paper from the Institute of Economic Studies, University of Iceland*.

Fletcher, R. (1987), *Practical Methods of Optimization*, Second Edition, Chichester: John Wiley & Sons, Inc.

Grogger, J.T., and R. T. Carson. (1991). Models for truncated counts. *Journal of Applied Econometrics* 6 (3): 225-238.

Hellström, Jörgen. (2006). A bivariate count data model for household tourism demand. *Journal of Applied Econometrics* 21 (2): 213–226.

Hynes, Stephen, and William Greene. 2013. A panel travel cost model accounting for endogenous stratification and truncation: A latent class approach. *Land Economics* 89 (1): 177-192.

Johnson, Norman L., Kemp, Adrienne W., and Samuel Kotz. (2005). *Univariate Discrete Distributions*, 3rd ed. New York: Wiley.

Johnson, Norman L., Kotz, Samuel, and N. Balakrishnan. (1994). *Continuous Univariate Distributions*, volume I, 2nd ed. New York: Wiley.

Katz, Leonard. (1965). Unified treatment of a broad class of discrete probability distributions. *Proceedings of the International Symposium on Discrete Distributions. Montreal:* 175 – 182.

Konijnendijk, Cecil C., Annerstedt, Matilda, Nielsen, Anders B., and Sreetheran Maruthaveeran. (2013). Benefits of urban parks: A systematic review. A report for IFPRA.

http://ign.ku.dk/english/employees/landscape-architecture-planning/?pure=files%2F44944034%2FIfpra_park_benefits_review_final_version.pdf Accessed December 1, 2015.

Kristofersson, Dadi., and Ståle Navrud. (2005). Validity tests of benefit transfer – are we performing the wrong tests? *Environmental and Resource Economics* 30 (3): 279-286.

Larson, Douglas M. (1993). Joint recreation choices and implied values of time. *Land Economics* 69 (3): 270-286.

Lockwood, Michael, and Kathy Tracy. (1995). Nonmarket economic valuation of an urban recreation park. *Journal of Leisure Research* 27 (2): 155-167.

Martinez-Cruz, Adan L., and Jaime Sainz-Santamaria. (2015). Recreational value of two peri-urban forests in Mexico City. http://www.researchgate.net/publication/282132523_Recreational_value_of_two_peri-urban_forests_in_Mexico_City. Accessed December 1, 2015.

Martínez-Espineira, Roberto, and Joe Amoako-Tuffour. (2008). Recreation demand analysis under truncation, overdispersion, and endogenous stratification: An application to Gros Morne National Park. *Journal of Environmental Management* 88 (4): 1320-1332.

McConnell, Kenneth. E. (1992). On-site time in the demand for recreation. *American Journal of Agricultural Economics* 74 (4): 918-925.

McCormack, Gavin R., Rock, Melanie, Toohey, Ann M., and Danica Hignell. (2010). Characteristics of urban parks associated with park use and physical activity: A review of qualitative research. *Health & Place* 16 (4): 712-726.

Min, Yongyi, and, Alan Agresti. (2005). Random effect models for repeated measures of zero-inflated count data. *Statistical Modelling* 5 (1): 1–19.

Moeltner, Klaus, and J. Scott Shonkwiler. (2010). Intercept and recall: Examining avidity carryover in on-site collected travel data. *Transportation Research Part D: Transport and*

Environment, 15 (7): 418-427.

More, Thomas A., Stevens, Thomas, and Geoffrey P. Allen. (1988). Valuation of urban parks. *Landscape and Urban Planning* 15 (1-2): 139-152.

Parsons, George R. (2003). The travel cost model. In *A Primer on Nonmarket Valuation* volume 3, eds. P. A. Champ, K. J. Boyle and T. C. Brown. Amsterdam, Netherlands: Kluwer Academic Publishers.

Shonkwiler, J. Scott, and W. Douglass Shaw. (1996). Hurdle count-data models in recreation demand analysis. *Journal of Agricultural and Resource Economics* 21 (2): 210-219.

Shaw, Daigee. (1988). On-site samples' regression: problems of nonnegative integers, truncation, and endogenous stratification. *Journal of Econometrics* 37 (2): 211-223.

Smith, V. Kerry, and Raymond Kopp. (1980). The spatial limits of the travel cost recreational demand model. *Land Economics* 56 (1): 64-72.

Trice, Andrew H., and Samuel E Wood. (1958). Measurement of recreation benefits. *Land Economics* 34 (3): 195–207.

Winkelmann, Rainer. (2004). Health care reform and the number of doctor visits - an econometric analysis. *Journal of Applied Econometrics* 19 (4): 455-472.

Willig, Robert D. (1976). Consumer’s surplus without apology. *American Economic Review* 66 (4): 589-597.

Tables and Figures

Table 1. Definitions of Uncompensated Demand Elasticities

	Own-Price Elasticity	Cross-Price Elasticity	Income Elasticity
Trips: n	$\epsilon_n = \frac{\partial n}{\partial p_n}\frac{p_n}{n}$	$\epsilon_{n,t} = \frac{\partial n}{\partial p_t}\frac{p_t}{n}$	$\epsilon_{n,Y} = \frac{\partial n}{\partial Y}\frac{Y}{n}$
Time spent on-site: t	$\epsilon_t = \frac{\partial t}{\partial p_t}\frac{p_t}{t}$	$\epsilon_{t,n} = \frac{\partial t}{\partial p_n}\frac{p_n}{t}$	$\epsilon_{t,Y} = \frac{\partial t}{\partial Y}\frac{Y}{t}$

	Trip-Price Elasticity	Time-Price Elasticity	Income Elasticity
Recreation: x	$\epsilon_{x,n} = \epsilon_n + \epsilon_{t,n}$	$\epsilon_{x,t} = \epsilon_{n,t} + \epsilon_t$	$\epsilon_{x,Y} = \epsilon_{n,Y} + \epsilon_{t,Y}$

Table 2. The Effects of Uncompensated Demand Elasticities on WTP

Equation (16)	Demand Elasticity	Effect on WTP
Part (a)	$\epsilon_{t,n} = 0$	$\int n(1-\epsilon_{t,n})dp_n = \int ndp_n$
	$\epsilon_{t,n} > 0$	$\int n(1-\epsilon_{t,n})dp_n < \int ndp_n$
	$\epsilon_{t,n} < 0$	$\int n(1-\epsilon_{t,n})dp_n > \int ndp_n$
Part (b)	$\epsilon_t = 0$ $\frac{p_n}{tp_t} \rightarrow 0$	$\int x\left(1-\frac{p_n}{tp_t}\epsilon_t\right)dp_t = \int xdp_t$
Part (c)	$\epsilon_{t,Y} = 0$	$\int \frac{np_n}{Y}\epsilon_{t,Y}dY = 0$

Table 3. Summary Statistics of Variables

Variable	Description	Mean	Std. Dev.	Min	Max
Trips	Number of trips taken last calendar month	4.20	5.95	0.00	31.00
Hours	Time spent on-site	1.30	0.90	0.05	7.50
Travel cost	Travel cost on a group basis scaled by a factor 1/1000 in ISK	0.62	0.42	0.00	6.38
Time cost	Time cost on a group basis scaled by a factor of 1/1000 in ISK	0.88	0.59	0.15	6.15
Adults	Number of adult visitors in a group	1.56	0.65	1.00	6.00

Table 4. Estimation Results from the Poisson Model

	Poisson	
	Estimate	t-value
Constant	1.75***	73.16
Travel cost	-231.46***	-14.72
Log likelihood	-6669.00	

Note: Significance codes:
*** Significance at the 1% level;
** significance at the 5% level; * significance at the 10% level.

Table 5. Willingness to Pay Estimates in ISK

	Poisson		
	Estimate	95% Confidence Interval	
Per group	6339.53***	6271.12	6407.94
Per person	4063.80***	4019.95	4107.65
Per person per hour per trip	968.48***	958.03	978.93

Note: The first row shows the average monthly WTP per group. The second row shows the average monthly WTP per person and the third row shows the average WTP per person per hour per trip. The confidence intervals are calculated using the delta method. Significance codes: *** Significance at the 1% level; ** significance at the 5% level; * significance at the 10% level.

Table 6. Estimation Results from the Poisson Gamma Log-Normal Mixture Model

	Trips		Hours	
	Estimate	t-value	Estimate	t-value
Constant	1.09^{***}	12.06	0.07^{*}	1.79
Travel cost	-211.18^{***}	-5.36	172.57^{***}	6.36
Time cost	-65.47^{**}	-2.30	-51.69^{***}	-3.82
Log likelihood	-5456.50			

Note: The standard deviations are estimated using a sandwich estimator. The travel cost and time cost parameters are scaled by 421.1. Significance codes: *** Significance at the 1% level; ** significance at the 5% level; * significance at the 10% level.

Table 7. Cholesky Factors and Covariance Estimates

	Estimate	t-value
L1	-1.27***	-44.69
L12	0.20***	6.91
L2	-0.18***	-6.23
Cov(n,t)	-0.26***	-6.83

Note: L1, L2, and L12 are the Cholesky factors from the lower triangular Cholesky matrix. Standard deviations are estimated using a sandwich estimator. The *t*-value of Cov(*n,t*) is calculated using the delta method. Significance codes: *** Significance at the 1% level; ** significance at the 5% level; * significance at the 10% level.

Table 8. Estimated Elasticities

	Trips		Hours		Combined	
	Estimate	t-value	Estimate	t-value	Estimate	t-value
Travel cost	-0.31***	-5.36	0.26***	6.36	-0.06	-0.81
Time cost	-0.14**	-2.30	-0.11***	-3.82	-0.25***	-4.21

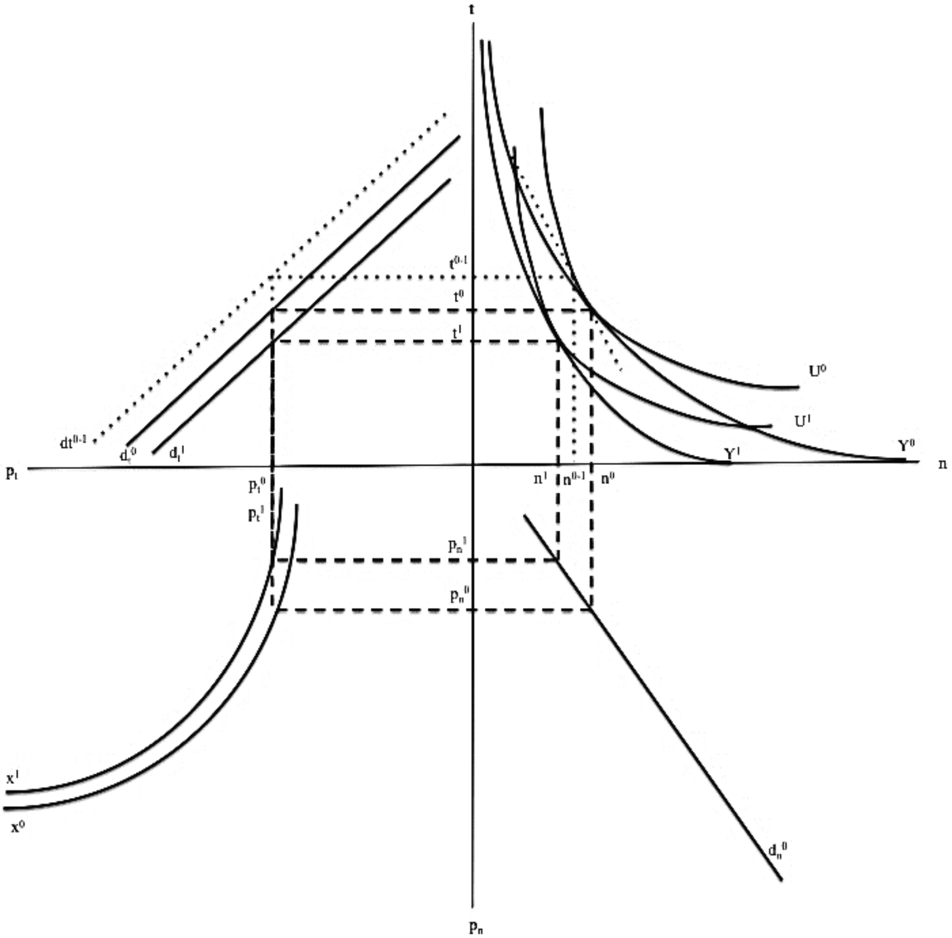
Note: The *t*-values are calculated using the delta method. Significance codes:
*** Significance at the 1% level; ** significance at the 5% level; * significance at the 10% level.

Table 9. Willingness to Pay Estimates in ISK

	Estimate	95% Confidence Interval	
Per group	$15{,}243^{***}$	13,618	16,868
Per person	$9{,}771^{***}$	8,729	10,813
Per person per hour per trip	$3{,}128^{***}$	2,794	3,461

Note: The first row shows the average monthly WTP per group. The second row shows the average monthly WTP per person and the third row shows the average WTP per person per hour per trip. The confidence intervals are calculated using the delta method. Significance codes: *** Significance at the 1% level; ** significance at the 5% level; * significance at the 10% level.

Figure 1. Utility Maximization, Demand Functions and Iso-Demand Curves for a Change in the Price of Trips



Paper 3

Habits in Frequency of Purchase Models

Arnar Buason[1]

[1] School of Economics and Business, Norwegian University of Life Sciences, P.O. Box 5003, 1432 Ås, Norway. Corresponding author is Arnar Buason: arnar.buason@nmbu.no

Abstract

We introduce a demand system, which incorporates habits to explain the dynamics of consumption of everyday consumer goods. In this model, the quantities purchased are modelled as a result of two decisions; how often to purchase each good and how much to purchase on each shopping occasion. An econometric model is developed for these decisions, and it is estimated by Bayesian methods. The data generating processes of the frequency and average quantity purchased of each good is, respectively, assumed to follow a multivariate Poisson log-normal and a multivariate gamma log-normal distribution. The mean and variance of the marginal distribution of the multivariate gamma log-normal distribution is also derived. As an illustration of the usefulness of the model, it is applied to model purchases of fresh fish in France. The results suggest that habits are most important for shopping frequencies, while price effects are more important for the average quantities purchased. These results are then used to produce an example of a profitable pricing strategy.
Key Words: Frequency of purchase, multivariate log-normal Poisson distribution, multivariate log-normal gamma distribution, habit formation, fish, France.

1. Introduction

The importance of habits in consumers purchase decisions has been well documented in the economic literature (e.g., Pollak 1970, Alessie and Kapteyn 1991, and Dynan 2000). A household's consumption over time is determined by multiple factors, and two of the key factors for explaining repeated purchases of everyday consumer goods are habits and duration[50]. Habits are negatively related to consumer utility, since the consumer does not respond optimally to changes in relative prices when influenced by habits, if compared to the classical utility maximization problem. Habit can be used by marketing and price strategies of

[50] We refere to duration as the combination of durability of the good and the consumers personal preference of time between purchases.

sellers to increase profits. Thus, understanding the dynamics of consumption and habits can provide essential information for policy and marketing decisions, see for example Holt and Goodwin (1997), Rickertsen et al. (1995), Heaton (1995), Rickertsen (1998), Fuhrer (2000), Arnade and Gopinath (2006), and Zhen et al. (2011).

Recent research on habit formation has focused on habits of product choice and purchased quantities, see for example Adamowicz and Swait (2012) and Zhen et al. (2010). However, the relationships between purchase frequency, habits, and product duration have not yet been analyzed. We divide a household's decision on how much to buy of everyday consumer goods over a period of time into two decisions; how much to buy each time and how often to go shopping. The knowledge of these relationships can provide information that can be useful for formulating marketing strategies. For example, if habits dominate durability in the shopping frequency, a marketing strategy that increases the frequency of shopping could be introduced. Even though the average consumer may purchase less in each trip, the profits over time could increase due to increased sales of other goods. Especially when the price of fresh fish is lowered the loss from the price reduction is minimal since the product is purchased very infrequently. However, the frequency of purchase increase due to the price reduction can lead to gains much higher than the loss, even though the increase is small, since consumers purchase a range of other products in each shopping trip.[51]

Purchase frequencies are rarely incorporated into consumer demand analysis. Exceptions include some studies that use infrequency of purchase models (IPMs), see for example Meghir and Robin (1992) and Robin (1993). These articles introduce purchase frequencies instead of a binary choice of whether to buy or not. Robin (1993) shows that when a demand system is adjusted by the predictions from such a count data model, it explains more of the variation in the data than a conventional double-hurdle model. Count

[51] This example is demonstrated with an empirical example in section 5.3.

data models have been used within health and environmental economics and marketing for analyzing topics like brand success, brand loyalty, and store choice.[52]

In applied demand analysis, one is typically faced with estimating a system of stochastically related equations. However, to estimate a multivariate count data model with unrestricted covariance structures for more than two goods is more challenging. For example, Meghir and Robin (1992) and Robin (1993) estimated their count data demand equations without allowing for any covariance structure between the equations. A few studies have estimated count data demand systems with unrestricted covariance structure. Two examples are Egan and Herriges (2006) who estimated a travel cost model using the multivariate Poisson log-normal model,[53] and Chib and Winkelmann (2001) who used the same model but applied Bayesian methods. Such simulation based methods do not require an approximation of the likelihood function but only sampling from the posterior and do not suffer so greatly from the curse of dimensionality as the Gaussian quadrature.

This article adds to the existing literature in three ways. Firstly, we extend the models developed in Spinnewyn (1981), Pashardes (1986), and Muellbauer and Pashardes (1992). In their models, dynamics were introduced by assuming that consumers derive utility from the flow of services provided by the consumption of goods. Our model extends these models by allowing the choice of the service stock to be derived from two decisions: (i) how often to purchase each good and (ii) how much to purchase on each occasion. Furthermore, we formulate the dynamic structure of habit formation and duration so it can be applied to semi-

---

[52] The dependent variable in these studies are typically the number of visits to doctors or the number of trips to a recreational site. Count data models used for the estimation of recreational demand include Smith (1988), Creel and Loomis (1990), Hellerstein (1991) and Egan and Herriges (2006). Count data models in health economics include Deb and Trivedi (2002), Munkin and Trivedi (1999) and Wang (2003). Examples of applications in marketing include Kau and Ehrenberg (1984), Uncles *et al*. (1995), Bhattacharya (1997) and Uncles and Lee (2006).

[53] Egan and Herriges (2006) approximated the log-likelihood using Gaussian-quadrature, which is a good choice for small problems but becomes a difficult when faced with any significant number of dimensions.

logarithmic demand specifications, which also are the standard specifications of the expected value in a count data model.

Secondly, we develop a theory-consistent empirical model, which allows for the joint estimation of purchase frequencies and the average purchased quantities on each shopping occasion by using Bayesian methods. For the purchase frequencies, a multivariate Poisson log-normal mixture model[54] is estimated, and for the average quantities a multivariate gamma log-normal mixture model is estimated. This empirical model allows for a non-restricted covariance structure between the equations in each subsystem and also between the two subsystems of the model. However, for ease of computation in our empirical application, we assume that the two subsystems are stochastically unrelated but let the covariance structure between equations in each system be unrestricted.[55] The mean and variance of the marginal distribution of a multivariate gamma log-normal distribution have previously not been derived, and we derive them by using the methods of iterative expectations.

Thirdly, the model is applied to French scanner data for fresh fish purchases for the period 2005-2008. Our empirical system consists of three categories of fresh fish: wild, farmed, and fish that the consumer does not know whether is wild or farmed. We refer to this category as other fish in the reminder of this paper. By using this commodity specification, we can calculate the premiums paid for wild as compared to farmed fish. Furthermore, the model shows how households respond when they have limited information regarding an important product attribute.[56] To avoid the large share of zeros in the data set, we aggregate the daily purchases to monthly aggregates. The two systems of demand equations are estimated using a semi-logarithmic functional form, and we combine them to estimate the

[54] This model was first introduced by Aitchison and Ho (1981).
[55] This simplifying assumption could lead to an incorrect specification of the variance if the two systems are stochastically correlated, but even so this will not lead to inconsistent or biased estimates of the parameters.
[56] Discussion on wild versus farmed fish can be found, for example, in Herrmann et al. (1993), Asche et al. (2005), Asche and Guttormsen (2014).

effects on total quantities purchased. The model shows how habit effects can be divided into habits in purchase frequencies and habits in average purchased quantities.

The rest of the article is organized as follows. In section two, the microeconomic model is described. The econometric model is developed in section three. In section four, we illustrate the usefulness of the model by applying French scanner data for purchases of fresh fish. Section five contains the results and section six concludes.

2. Microeconomic Model

Following Spinnewyn (1981), and Muellbauer and Pashardes (1992), let Z_{it} be services in period *t* provided by the flow of good x_i purchased in period *t* or earlier periods. Let Z_{it} be defined as the weighted sum of the logarithm of current and past purchases as follows:

$$Z_{it} = \sum_{\tau=0}^{\infty} d_i^{\tau} \ln x_{it-\tau} = \ln x_{it} + d_i Z_{it-1}. \tag{1}$$

The degree of durability of good x_i is determined by a simple parameter d_i, where $0 \leq d_i < 1$, see for example Muellbauer and Pashardes (1992) and Zhen et al. (2011).[57] The parameter d_i does not just reflect the biological durability of the good, but also personal preference of time between purchases (Zhen et al. 2011). For example, fresh fish is not a durable good, however, this lack of durability does not imply that a purchase of fresh fish today has no impact on purchases of fish in the future. A consumer might wish to purchase fish once every three days to maintain a healthy diet.

Next, we extend this standard model by assuming that the decision of purchased quantity x_{it} of good *i* in a time period *t* is determined by two decisions: (i) how often to purchase good *i* in period *t* and (ii) how much to purchase on average of good *i* on each

---

[57] The durability of good x_{it} could be determined more generally by a decay function which could be more sophisticated than a single parameter, but to keep the analysis simple we choose to go with the convention.

shopping occasion in period t. These decisions are expressed by the identity $x_{it} \equiv n_{it}q_{it}$, where n_{it} is purchase frequency and q_{it} is average quantity.

Then, we follow Muellbauer and Pashardes (1992) and assume that habits are developed over time, where the desired level of the utility generating service stock Z_{it}^{*} is defined as follows:

$$Z_{it}^{*} = \exp(Z_{it} - \phi_i Z_{it-1}). \tag{2}$$

Habits are thus introduced by a single parameter ϕ_i for each good where $0 \leq \phi_i < 1$. Habits are therefore treated as the opposite of durability. The grater the habit formation parameter is the larger service stock Z_{it} needs to be maintained to reach the desired level of utility generating stock Z_{it}^{*}. In Pashardes (1986), the effects of durability and habits are treated jointly as a single parameter reflecting the total effect. However, we are interested in how the individual components of durability and habits relate to purchase frequency and quantities purchased, respectively. Substituting equation (1) into equation (2) gives the following desired level of the utility generating service stock:

$$\begin{aligned} Z_{it}^{*} &= \exp(\ln x_{it} + d_i Z_{it-1} - \phi_i Z_{it-1}) \\ &= x_{it}\exp\big((d_i - \phi_i)Z_{it-1}\big). \end{aligned} \tag{3}$$

Thus, if durability, d_i, dominates habits, ϕ_i, then the total effect is positive and the utility generating stock is greater than the quantity of good *i* purchased in period *t*. The opposite is true when habits dominate durability. Then, the net effect is negative and the utility generating stock is less than the quantity of the good *i* purchased in period *t*.

Following Zhen et al. (2011) consumer's life time utility is assumed to be weakly separable over time such that:

$$U = v[v_t(Z_{0t}^{*}, \dots, Z_{mt}^{*}), v_{t+1}(Z_{0t+1}^{*}, \dots, Z_{mt+1}^{*}), \dots, v_T(Z_{0T}^{*}, \dots, Z_{mT}^{*})], \tag{4}$$

and the present value of the lifetime budget constraint is:

$$W_t = \sum_{\tau=t}^{T} \sum_{i=0}^{m} \hat{p}_{i\tau} Z_{i\tau}^{*}, \tag{5}$$

where $\hat{p}_{i\tau}$ is the user's cost in period τ of service stock $Z_{i\tau}^{*}$. This user's cost can be thought of as either rational or myopic. However, in the following analysis the myopic assumption will be used for simplicity, since neither assumption has proven to be consistently more accurate, see for example Zhen et al. (2010).[58]

Due to the weakly separable utility function (4), the consumer can allocate period to period budget $y_t = \sum_i \hat{p}_{it} Z_{it}^{*}$ and then maximizes the utility $v_t(Z_{0t}^{*}, \dots, Z_{mt}^{*})$ in each period individually. This gives a *m* dimensional system of Marshallian demand functions of the form, $Z_{it}^{*} = g(\hat{p}_{it}, y_t)$ in each period. Substituting equation (3) into the demand equation gives the following expressions:

$$\begin{gathered} x_{it}\exp\big((d_i - \phi_i)Z_{it-1}\big) = g(\hat{p}_{it}, y_t) \\ x_{it} = g(\hat{p}_{it}, y_t)\exp\big((\phi_i - d_i)Z_{it-1}\big). \end{gathered} \tag{6}$$

Equation (6) gives the demand for good x_{it} where both habits and durability have been incorporated. Inserting the identity $x_{it} \equiv n_{it}q_{it}$ into equation (6) and further assuming that $g(\hat{p}_{it}, y_t) \equiv n(\hat{p}_{it}, y_t)q(\hat{p}_{it}, y_t)$, gives:

$$n_{it}q_{it} = n(\hat{p}_{it}, y_t)q(\hat{p}_{it}, y_t)\exp\big((\phi_i - d_i)Z_{it-1}\big). \tag{7}$$

Equation (7) summarizes our theoretical contribution to the model. The multiplicative form of equation (7) allows us to analyze two systems of demand equations in each period, $n_{it}(\cdot)$ and $q_{it}(\cdot)$, which provide a more detailed and useful information regarding consumer behavior. To investigate how the purchase frequencies and average quantities purchased contribute to habit and durability, we assume that the habit and duration parameters can be additively separated, $\phi_i = \psi_i + \omega_i$ and $d_i = \varphi_i + \zeta_i$, where ψ_i and φ_i originate from n_{it} and ω_i and ζ_i originate from q_{it}. Where ψ_i and φ_i are the habit and duration parameters of the

[58] The myopic assumption is a simplifying assumption under which the consumer does not account for user costs of stocks.

frequency part, respectively. The parameters ω_i and ζ_i have the same interpretation but for the average quantity part of the model. For example, if ψ_i dominates φ_i but ζ_i dominates ω_i, it would be more profitable to influence a larger increase in n_{it} relative to q_{it}. Equation (7) can be rewritten as:

$$n_{it}q_{it} = n(\hat{p}_{it}, y_t)\exp\big((\psi_i - \varphi_i)Z_{it-1}\big)q(\hat{p}_{it}, y_t)\exp\big((\omega_i - \zeta_i)Z_{it-1}\big), \qquad (8)$$

i.e., total purchases of x_{it} is found by multiplying the frequency of purchase, $n_{it} = n(\hat{p}_{it}, y_t)\exp\big((\psi_i - \varphi_i)Z_{it-1}\big)$ with the average quantity purchased $q_{it} = q(\hat{p}_{it}, y_t)\exp\big((\omega_i - \zeta_i)Z_{it-1}\big)$.

3. Econometric Model

The frequency of shopping $n_i = (n_{i11}, n_{i12}, \dots, n_{iKT})$ is assumed to follow a discrete distribution $f_{Ni}(n_i|\beta_i, C)$, for $n_{ikt} = 0,1,2, \dots$ where β_i is a vector of parameters, and $k = 1,2, \dots, K$ and $t = 1,2, \dots, T$, where *k* and *t* refer to households and time period, respectively. *N* denotes a random variable and n_{ikt} refers to an observed value of *N*, and *C* is a matrix of explanatory variables. The average purchases $q_i = (q_{i11}, q_{i12}, \dots, q_{iKT})$ are only observed when a trip to the shop takes place. Thus, the variable $q_i|n_i > 0$ is assumed to follow a continuous distribution $f_{Qi|n_i>0}(q_i|\alpha_i, C, n_i > 0)$, defined only over positive values, where α_i is a vector of parameters, the interpretation of *Q* and q_{ikt} is that q_{ikt} is an observed value of the random variable *Q*.[59] The data generating process (DGP) for average quantity purchased is therefore represented by the two-part model:

[59] The conditional mean of n_i and q_i are given as follows: $\mathrm{E}(n_i|\beta_i, C, b_{Ni})$, $\mathrm{E}\big(q_i\big|\alpha_i, C, b_{Qi}\big) = \Pr(N > 0|\beta_i, C)\mathrm{E}\big(q_i\big|\alpha_i, C, n_i > 0, b_{Qi}\big)$, The marginal effects of $\mathrm{E}\big(q_i\big|\alpha_i, C, b_{Qi}\big)$ are given by: $\frac{\partial \mathrm{E}(q_i|\alpha_i, C, b_{Qi})}{\partial c_i} = \frac{\partial \mathrm{E}(q_i|\alpha_i, C, n_i > 0, b_{Qi})}{\partial c_i}\Pr(N > 0|\beta_i, C)$.

$$f_Q(q_i|\alpha_i, C) = \begin{pmatrix} \Pr(N = 0|\beta_i, C) & \text{if } q_i = 0 \\ \Pr(N > 0|\beta_i, C) f_{Qi|n>0}(q_i|\alpha_i, C, n_i > 0) & \text{if } q_i > 0 \end{pmatrix}. \tag{9}$$

The decisions to purchase a good and how much to purchase of the good in each trip is likely to be related decisions, and it is desirable to model them as stochastically correlated. Furthermore, the demand for one good is directly related to the demand to the other, and it desirable to allow for correlation between the equations within each of the two systems. To allow for these correlations, random effects are introduced to both densities $f_{Ni}(n_i|\beta_i, C, b_{Ni})$ and $f_{Qi|n>0}\left(q_i\middle|\alpha_i, C, n_i > 0, b_{Qi}\right)$, where b_{Ni} and b_{Qi} are random effects for frequencies and average quantities respectively, and are assumed to follow a multivariate normal distribution:

$$\begin{bmatrix} b_N \\ b_Q \end{bmatrix} \Bigg| D \sim MVN\left(\begin{bmatrix} 0^m \\ 0^m \end{bmatrix}, \begin{bmatrix} D_N & D_{NQ} \\ D_{NQ} & D_Q \end{bmatrix}\right), \tag{9}$$

where $b_N = (b_{N1}, \dots, b_{Nm})$, $b_Q = \left(b_{Q1}, \dots, b_{Qm}\right)$, and D is the unrestricted block covariance matrix. The joint probability density function for n_i and q_i is then given as follows:

$$p(n_i, q_i|\beta_i, \alpha_i, D, C) = \int \prod_{t=1}^{T} f_{Ni}(n_{ikt}|\beta_i, C, b_{Nikt})\, f_{Qi|n>0}(q_i|\alpha_i, C, n_i > 0)\phi(b_{ik}|0, D)db_{ik}. \tag{10}$$

The product operator is inside the integral since b_N and b_Q each have one draw for the T random variables $n_{ik1}, n_{ik2}, \dots, n_{ikT}$ and $q_{ik1}, q_{ik2}, \dots, q_{ikT}$, respectively. Thus, there is a new draw for each cluster, but not for each time period within a cluster. The likelihood is then given by:

$$L = \prod_{k=1}^{K} \prod_{i=1}^{M} p(n_i, q_i|\beta_i, \alpha_i, D, C). \tag{11}$$

Since the joint density $p(n_i, q_i|\beta_i, \alpha_i, D, C)$ does not have a closed form solution, the likelihood L can not be optimized with conventional Newton methods, so we turn to simulation methods.[60]

[60] For a simple problem, it is also feasible to use the Gaussian-quadrature.

3.1 Distributional Assumptions

To be able to estimate the parameters of our model, the conditional distribution assumptions of $n_i|\beta_i, C$ and $q_i|\alpha_i, C, n_i > 0$ need to be determined. The number of shopping trips is assumed to follow a Poisson distribution, $n_i|\beta_i, C, \sim$ Poisson(μ_i) and the average quantity purchased in each trip is assumed to follow a gamma distribution $q_i|\alpha_i, C, n_i > 0 \sim$ Gamma(κ_i, η_i). The parameter of the Poisson distribution is specified as $\lambda_i = \exp(C_i\beta_i)$. The mean of the gamma distribution is specified as $\kappa_i\eta_i = \exp(C_i\alpha_i)$. Now let $v_{io} = \exp(b_{io})$, where $v_{io} = (v_{i1o}, v_{i2o}, \dots, v_{iJo})$, and $v_{io} \sim LN(\mu_o, \Sigma_o)$, where $\mu_o = \exp(0.5\text{diag}(D_o))$, and $o = N, Q$. The variance covariance matrix is then $\Sigma_o = (\text{diag}(\mu_o))[\exp(D_o) - 11'](\text{diag}(\mu_o))$. Multiplying both means with this random effect gives, $\lambda_i v_{Ni} = \exp(C_i\beta_i + b_{Ni})$ and $\kappa_i\eta_i v_{Qi} = \exp(C_i\alpha_i + b_{Qi})$. To simplify the analysis, the covariance between n_i and q_i is assumed to be zero. This could lead to incorrect variance specification, but it will not lead to inconsistent parameter estimates. However, a non-restricted covariance matrix is assumed between clusters within both parts of the model. Equation (9) is then reduced to:

$$\begin{bmatrix} b_N \\ b_Q \end{bmatrix} \Big| D \sim MVN\left(\begin{bmatrix} 0^m \\ 0^m \end{bmatrix}, \begin{bmatrix} D_N & 0_{NQ} \\ 0_{NQ} & D_Q \end{bmatrix}\right). \quad (12)$$

Under this specification, the Poisson part of the model is the Log-Normal Poisson model in Atchinson and Ho (1989). It is thus possible to derive the mean and variance of the marginal distribution of n_i without integration. Let $\tilde{\lambda}_i = \lambda_i\mu$, where $\tilde{\lambda}_i = (\tilde{\lambda}_{i11}, \tilde{\lambda}_{i12}, \dots, \tilde{\lambda}_{iKT})$ and $\tilde{\Lambda}_i = \text{diag}(\tilde{\lambda}_i)$. Applying the law of iterated expectations, one obtains $E(n_i|\beta_i, C, D_N) = \tilde{\lambda}_i$ and $\text{var}(n_i|\beta_i, C, D_N) = \tilde{\Lambda}_i + \tilde{\Lambda}_i[\exp(D_N) - 11']\tilde{\Lambda}_i$. Then, the covariance between n_{ikt} and n_{ift} is calculated as: $\text{cov}(n_{ikt}, n_{ift}) = \tilde{\lambda}_{ikt}(\exp(d_{Nikt,ift}) - 1)\tilde{\lambda}_{ift} =$

$\lambda_{ikt}\exp\left(0.5\left(d_{Nikt,ikt}\right)\right)\left(\exp\left(d_{Nikt,ikf}\right)-1\right)\lambda_{jkt}\exp\left(0.5\left(d_{Nift,ift}\right)\right)$. For a more detailed derivation of this result see Atchinson and Ho (1989).

The mean and variance of the marginal distribution of q_i have, as far as we know, not been derived in the literature but they can also be derived without integration. We start by defining $\delta_{ikt} \equiv \kappa_{ikt}\eta_{ikt}$, and the mean of the marginal distribution of q_i is:

$$\mathrm{E}\left[q_{ikt}|\beta_i, C, D_Q\right] = \mathrm{E}_{v|\delta}\left[\mathrm{E}_{q|\delta,v}\left[q_{ikt}|\delta_{ikt}, v_{ikt}, D_Q\right]\right] \tag{13}$$

$$= \mathrm{E}_{v|\delta}[\delta_{ikt}v_{ikt}] = \delta_{ikt}\mathrm{E}_{v|\delta}[v_{ikt}] = \delta_{ikt}\mu \equiv \tilde{\delta}_{ikt}.$$

The variance of the marginal distribution of q_i is:

$$\mathrm{var}\left[q_{ikt}|\beta_i, C, D_Q\right] \tag{14}$$

$$= \mathrm{E}_{v|\delta}\left[\mathrm{var}_{q|\delta,v}\left[q_{ikt}|\delta_{ikt}, v_{ikt}, D_Q\right]\right] + \mathrm{var}_{v|\delta}\left[\mathrm{E}_{q|\delta,v}\left[q_{ikt}|\delta_{ikt}, v_{ikt}, D_Q\right]\right]$$

$$= \mathrm{E}_{v|\delta}[v_{ikt}\kappa_{ikt}\eta_{ikt}^2] + \mathrm{var}_{v|\delta}[v_{ikt}\delta_{ikt}]$$

$$= \kappa_{ikt}\eta_{ikt}^2\mathrm{E}_{v|\delta}[v_{ikt}] + \delta_{ikt}^2\mathrm{var}_{v|\delta}[v_{ikt}]$$

$$= \kappa_{ikt}\eta_{ikt}^2\mu + \delta_{ikt}^2\mu^2\left(\exp\left(d_{Qkk}\right)-1\right)$$

$$\eta_{ikt}\tilde{\delta}_{ikt} + \tilde{\delta}_{ikt}^2\left(\exp\left(d_{Qkk}\right)-1\right).$$

Let $\tilde{\delta}_i = \left(\tilde{\delta}_{i11}, \tilde{\delta}_{i12}, \dots, \tilde{\delta}_{iKT}\right)$, $\eta_i = (\eta_{i11}, \eta_{i12}, \dots, \eta_{iKT})$, $\widetilde{\Delta}_i = \mathrm{diag}\left(\tilde{\delta}_i\right)$, and $\mathrm{H}_i = \mathrm{diag}(\eta_i)$. Then, we can rewrite equation (14) in matrix form as:

$$\mathrm{var}\left[q_i|\beta_i, C, D_Q\right] = \mathrm{H}_i\widetilde{\Delta}_i + \widetilde{\Delta}_i\left(\exp\left(D_Q\right) - 11'\right)\widetilde{\Delta}_i. \tag{15}$$

The covariance between q_{ikt} and q_{ift} becomes:

$$\mathrm{cov}\left(q_{ikt}, q_{ift}\right) = \tilde{\delta}_{ikt}\left(\exp\left(d_{Qikt,ift}\right)-1\right)\tilde{\delta}_{ift} \tag{16}$$

$$= \delta_{ikt}\exp\left(0.5\left(d_{Qikt,ikt}\right)\right)\left(\exp\left(d_{Qikt,ift}\right)-1\right)\exp\left(0.5\left(d_{Qift,ift}\right)\right)\delta_{ift},$$

$$k \neq f.$$

3.2 Priors and MCMC Sampling

To estimate our model using Bayesian methods, we choose to use uninformative priors, see for example Chib and Winkelmann (2001) for a discussion. Let $\beta \sim \mathrm{N}(\beta_0, B_0^{-1})$, $\alpha \sim \mathrm{N}(\alpha_0, A_0^{-1})$, $\kappa \sim \mathrm{Gamma}(k_0, s_0)$, $D_N^{-1} \sim \mathrm{Wishart}(v_{N0}, R_{N0})$, and $D_Q^{-1} \sim \mathrm{Wishart}(v_{Q0}, R_{Q0})$, where $\beta_0, B_0, \alpha_0, A_0, k_0, s_0, v_{N0}, R_{N0}, v_{Q0},$ and R_{Q0} are known hyperparameters and $\mathrm{Wishart}(\cdot,\cdot)$ is the Wishart distribution with v_{o0} degrees of freedom and scale matrix R_{o0}, where $o = N, Q$. By the Bayes theorem, the posterior density of the two parts of the model are proportional to the following expressions:

$$\phi(\beta|\beta_0, B_0^{-1}) f_W(D_N^{-1}|v_{N0}, R_{N0}) \prod_{k=1}^{K}\prod_{i=1}^{M} f_N(n_{ik}|\beta, b_{Nik})\, \phi(b_{Nik}|0, D_N), \tag{17}$$

$$\phi(\alpha|\alpha_0, A_0^{-1}) f_W\left(D_Q^{-1}\middle|v_{Q0}, R_{Q0}\right) \prod_{k=1}^{K}\prod_{i=1}^{M} f_{Q|n>0}\left(q_{ik}\middle|\alpha, b_{Qik}, n_{ik} > 0\right) \phi\left(b_{Qik}\middle|0, D_Q\right), \tag{18}$$

where f_W is the Wishart distribution. We then construct Markov chains using the blocks of parameters b_N, b_Q, β, α, D_N and D_Q, and their full conditional distributions:

$$[b_N|n, \beta, D_N]; \quad [\beta|n, b_N]; \quad [D_N|b_N], \tag{19}$$

$$[b_Q|q, \alpha, D_Q]; \quad [\alpha|q, b_Q]; \quad [D_Q|b_Q]. \tag{20}$$

The simulation output is generated by recursively simulating these distributions, using the most recent values of the conditioning variables in each step. The sampling of b_N and b_Q starts with specifying the target densities:

$$\pi(b_N|n, \beta, D_N) = \prod_{k=1}^{K}\prod_{i=1}^{M} \pi(b_{Nik}|n_{ik}, \beta, D_N), \tag{21}$$

$$\pi\left(b_Q\middle|q, \alpha, D_Q\right) = \prod_{k=1}^{K}\prod_{i=1}^{M} \pi\left(b_{Qik}\middle|q_{ik}, \alpha, D_Q\right) \tag{22}$$

To sample the density of the *k*th household of the *i*th cluster of the target densities we specify:

$$\pi(b_{Nik}|n_{ik},\beta,D_N) = c_{ik}\phi(b_{Nik}|0,D_N)\prod_{t=1}^{T}\exp[-\exp(C_{ikt}\beta_t + b_{Nikt})] \tag{23}$$

$$\times[-\exp(C_{ikt}\beta_t + b_{Nikt})]^{n_{ikt}}$$

$$\equiv c_{ik}\pi^+(b_{Nik}|n_{ik},\beta,D_N),$$

$$\pi\left(b_{Qik}\middle|q_{ik},\alpha,\kappa,D_Q\right) = i_{ik}\phi\left(b_{Qik}\middle|0,D_Q\right)\prod_{t=1}^{T}\frac{q_{ikt}^{C_{ikt}\alpha_t+b_{Nikt}}}{\Gamma(C_{ikt}\alpha_t + b_{Nikt})\kappa^{C_{ikt}\alpha_t+b_{Qikt}}} \tag{24}$$

$$\times\ \exp[-q_{ikt}/\kappa]$$

$$\equiv i\pi^+\left(b_{Qik}\middle|q_{ik},\alpha,\kappa,D_Q\right).$$

The target distributions are a Poisson log-normal mixture and a Gamma log-normal mixture, respectively. We utilize a random walk Metropolis algorithm, for a discussion see for example Roberts et al. (1997). The proposal density is found by approximating the target density around the modal value by a multivariate *t*-distribution. Let $\hat{b}_{Nik} =$ argmax $\ln\pi^+(b_{Nik}|n_{ik},\beta,D_N)$ and $V_{b_{Nik}} = \left(H_{b_{Nik}}\right)^{-1}$ be the inverse of the Hessian of $\ln\pi^+(b_{Nik}|n_{ik},\beta,D_N)$ at the mode $\hat{b}_{Nik}$. To find these estimates, we used the Newton-Raphson algorithm. Then, our proposal density is $b_{Nik}^{(s)}\left|b_{Nik}^{(s-1)}\right. \sim t\left(b_{Nik}^{(s-1)}, V_{b_{Nik}}, v\right)$, where *v* is the degrees of freedom and *s* indicates the draw number. We then make a random draw e_{ik} from $t\left(0, V_{b_{Nik}}, v\right)$, where $b_{Nik}^{(s)} = b_{Nik}^{(s-1)} + e_{ik}$ and we move from $b_{Nik}^{(s-1)}$ to $b_{Nik}^{(s)}$ with probability

$$r = \min\left\{\frac{\pi^+\left(b_{Nik}^{(s)}\middle|n_{ik},\beta,D_N\right)}{\pi^+\left(b_{Nik}^{(s-1)}\middle|n_{ik},\beta,D_N\right)}, 1\right\}. \tag{25}$$

We then sample *u* from a uniform distribution $\mathrm{U}(0,1)$ and if $u < r$ then $b_{Nik}^{(s)} = b_{Nik}^*$ otherwise $b_{Nik}^{(s-1)} = b_{Nik}^*$. We use the exact same steps for the sampling of b_Q.

The sampling of β and α follows the approach described above. The respective target distributions are given as follows:

$$\pi(\beta|n, b_N, D_N) = \phi(\beta|\beta_0, B_0^{-1}) \prod_{k=1}^{K}\prod_{i=1}^{M}\prod_{t=1}^{T} \exp[-\exp(C_{ikt}\beta_t + b_{Nikt})] \tag{26}$$

$$\times [-\exp(C_{ikt}\beta_t + b_{Nikt})]^{n_{ikt}}$$

$$\pi\big(\alpha|q, b_Q, \kappa, D_Q\big) = \phi(\alpha|\alpha_0, A_0^{-1}) \prod_{k=1}^{K}\prod_{i=1}^{M}\prod_{t=1}^{T} \frac{q_{ikt}^{C_{ikt}\alpha_t + b_{Nikt}}}{\Gamma(C_{ikt}\alpha_t + b_{Nikt})\kappa^{C_{ikt}\alpha_t + b_{Qikt}}} \tag{27}$$

$$\times\ \exp[-q_{ikt}/\kappa].$$

Sampling D_N^{-1} and D_Q^{-1} is a simpler process than the other two blocks of parameters, since we specified a hyperprior, which resulted in a Wishart distribution. We sample D_o^{-1}, $o = N, Q$, from a distribution proportional to:

$$f_W(\,D_o^{-1}|v_{o0}, R_{o0}^{-1}) \prod_{k=1}^{K}\prod_{i=1}^{M} \phi(b_{oik}|0, D_o). \tag{28}$$

The distribution of $D_o^{-1}|b_0$ then results in aWishart distribution:

$$D_o^{-1}|b_0 \sim \text{Wishart}\left(M + v_{o0}, \left[R_{o0}^{-1} + \sum_{k=1}^{K}\sum_{i=1}^{M}(b'_{oik}b_{oik})\right]^{-1}\right), \tag{29}$$

with $M + v_{o0}$ degrees of freedom and a scale matrix $[R_{o0}^{-1} + \sum_{k=1}^{K}\sum_{i=1}^{M}(b'_{oik}b_{oik})]^{-1}$.

4. Empirical Illustration

We use a scanner data set, which is collected by TNS Worldpanel and includes weekly purchases of fresh fish for about 6,000 French households for the period 2005-2008. The data set includes a detailed description of social, geographical and other characteristics of the participating households, and includes variables as diverse as shoe size and socioeconomic class. However, we do not include these household characteristics in our analysis because the main point of our example is to demonstrate the usefulness of the method. A large proportion of zero observations is a standard problem when working with micro-data, and we reduced the number of zero observations by aggregating the data over months.[61]

4.1. Specification of Variables

To estimate equation (8), the total service stock of each type of fish Z_{ikt} needs to be predicted. The first three months of the first year, 2005, are used to predict Z_{ikt}. From equation (1) we note that the relationship between the first two periods can be written as:

$$Z_{ik1} = \ln x_{ik1} + \left(\frac{d_i}{1-d_i}\right)\ln x_{ik0}. \tag{30}$$

The functional form for $n(p_{it}, y_t)$ and $q(p_{it}, y_t)$ is assumed to be semi-logarithmic. Thus, $g(p_{it}, y_t)$ is also semi-log. Equation (6) can then be written as:

$$x_{ik2} = \exp(C_{ikt}\gamma_i + (\phi_i - d_i)Z_{ik1}) \tag{31}$$

$$\ln x_{ik2} = C_{ikt}\gamma_i + (\phi_i - d_i)\ln x_{ik1} + (\phi_i - d_i)\left(\frac{d_i}{1-d_i}\right)\ln x_{ik0}$$

Equation (31) is estimated to obtain estimates of ϕ_i and d_i, where d_i is used to predict Z_{ikt}, and Equation (8) is estimated by the following two systems of equations:

$$E(n_{ik}|C_i) = \exp\left(\beta_{ik} + \sum_{s=1}^{S}\beta_{is}\left(\frac{p_{sk}}{CPI}\right) + \theta_i\left(\frac{y_k}{CPI}\right) + (\psi_i - \varphi_i)\hat{Z}_{ikt-1} + b_{Nik}\right) \tag{32}$$

$$E(q_{ik}|C_i) = \exp\left(\alpha_{ik} + \sum_{s=1}^{S}\alpha_{is}\left(\frac{p_{sk}}{CPI}\right) + \varpi_i\left(\frac{y_k}{CPI}\right) + (\omega_i - \zeta_i)\hat{Z}_{ikt-1} + b_{Qik}\right), \tag{33}$$

where the price of fish category *i* for household *k* is denoted by p_{ik} and the total fish expenditure of household *k* is given by y_k, and CPI is the average French consumer price index. The semi-logarithmic demand equations in Equations (32) and (33) are integrable

[61] The problem of zero observations is also a smaller problem when the data generating process is discrete as it is with the Poisson distribution, since the probability of observing a zero is positive.

when the restrictions $\beta_{is} = 0 \; \forall \; i \neq s$ and $\theta_i = \theta \; \forall \; i$ are imposed (LaFrance and Hanemann, 1989).

4.2. Descriptive Statistics

The variables used in our example are given in table 1. The dependent variables in the count data system are the frequency of purchase of wild fish, farmed fish, and fish produced with unknown production technology, for short referred to as other fish in the remaining of the article. The category of wild fish is mainly made up of cod and other white fish, but the farmed fish category is mainly salmon. The average frequency of purchase of wild fish is around 1.5 over the four years sample period. This frequency is the highest among the three groups, but it is still very low. The main reason for this low average value is the large number of zero observations and the maximum observed number of purchases is 27. The dependent variables in the continuous part of the model are the truncated average quantities of purchase of wild fish, farmed fish, and other fish. The average quantities are ranging between 650 and 700 grams over the sample period. However, there are large variations. For example, the minimum purchase of wild fish is 15 grams while the maximum quantity is more than 15 kilograms. The data set does not contain any information regarding prices of the products. Prices are therefore estimated by dividing expenditures by quantities purchased of each good in each shopping trip to create unit values. When zero purchases are recorded, there is no price available so the average price is used. This approach is used in other demand studies such as Allais et al. 2010 and Bertail and Caillavet (2008). There is however a problem with this approach since prices will be influenced by choice of quality. These issues are discussed, for example, by Deaton (1997). The exogenous variables are real prices (unit values divided by the average French CPI), lagged service stocks of the different types of fish and households' total expenditures on fresh fish. There is relatively little variation in the average prices for the different fish categories, due to aggregation. The calculation of these prices are discussed in the next section. The calculation of the service stock of wild fish is discussed in the previous section. We can observe substantial variation in the average service stocks of the different fish types. The largest average service stock is for wild fish with a stock of almost 1.7 kilograms.

(Table 1 about here)

5.1 Empirical Results

Tables 2 – 4 presents the results of the estimation of the MPLN and MGLN systems for each type of category. The parameter estimates, associated *t*-values and Geweke *Z*-values, for a test of stationarity of the Markov chains, for the various fish types are shown.[62] For the chains to be stationary the Geweke-Z should be below 2. Table 2 presents results for wild fish. From Geweke *Z* – values for the MPLN part, we see that all Markov Chains are stationary, and our parameter estimates can be interpreted reliably. However, two of the of the Geweke *Z* estimates in the MGLN are above 2 and should thus be interpreted with caution. The lag of stock parameter is positive in both equations. As discussed in Section 2, a positive stock parameter implies that habits dominate durability in the equations for purchase frequency and average quantity. Thus, the average consumer purchase wild fish more frequently and in higher quantity than would be optimal in the absence of habits. Due to habits sellers can sell more than they would be able to sell to consumers without habits. All three time dummies are negative, which indicates that the purchase frequency, has been declining during 2006, 2007, and 2008 as compared with 2005. However, the time dummies are all positive for the average quantity part, indicating an increase in quantity purchased relative to 2005. The monthly time trend shows the exact same pattern as the annual dummies.

[62] For a discussion of the Geweke convergence test see for example Nylander et al. (2008).

(Table 2 about here)

Table 3 presents the results for farmed fish. The Geweke *Z*-values indicate that the Markov Chains have converged, and our estimates are reliable. The lag of stock parameter is positive in both equations, indicating that habits dominate duration for both frequency and average quantity, i.e., habits results in more frequent and higher purchases of farmed fish. The time dummies indicate a declining purchase frequency relative to 2005, but average quantity has increased relative to 2005. The time trend shows a decline in purchase frequency over time, but an increase in average purchased quantity.

(Table 3 about here)

Table 4 presents the results for other fish. The Geweke Z-values show that all Markov chains are stationary. The results are similar to the results for wild and farmed fish. This is not surprising given that this category consists of a mixture of fishes produced with unknown production technology. Habit dominates duration for purchase frequency and average quantity. The time dummies show that purchase frequency has been declining relative to 2005. The time trend has been declining for purchase frequency, but increasing for average quantity.

(Table 4 about here)

Table 5 shows the parameter estimates from the cross-equation covariance matrix for the count data part of the model and the average frequency part. Sigma11 Sigma22, and Sigma33 refer to the variance of the random effects of demand equation 1-3, respectively, in each system. Sigma12 and Sigma21 show the covariance between random effects of equation one and two, and Sigma13 and Sigma31 show the covariance between equation one and three. Finally, Sigma23 and Sigma32 are the show the covariance between equation two and three. There is a positive covariance, and thus positive correlation, between the purchases of wild and farmed fish. This is not does not indicate that the groups are compliments in price, only that those who are consumers of wild fish also consume farmed fish. There is negative correlations between wild and other fish and farmed and other fish. These results indicate that the buyers of wild fish and farmed fish are not the same as those who buy other fish, on average.

(Table 5 about here)

5.2 Elasticities

Table 6 shows the calculated elasticities for the frequencies of purchase, average purchased quantities, and total purchased quantities as calculated from equations (32) and (33). Net habits are constructed from the habit components of frequencies and average quantities. When the lag of service stock of wild fish increases by 10%, the purchase frequency increases by 0.66% and the average quantity increases by 0.05%. The effect on total purchases is 0.71%. These results indicate that most of the effect of habit formation on total purchases originates from habits in purchase frequency. The results for farmed and other fish are similar, and we conclude that changes due to habits in total purchases of fresh fish mainly are related to habits in the frequency of purchase and only to a less extent from habits in average purchases in each shopping trip.

Price changes have more effect on average quantities than on purchase frequencies. When the price of wild fish is reduced by 1%, the purchase frequency increases by 0.12%, the average quantity increases by 0.32%, and total purchases increases by 0.43%.

As discussed in Section 2, the total expenditure elasticity for the frequency of purchase is constrained to be the equal for the three fish types to fulfill the symmetry condition in a semi-logarithmic demand system. However, the total expenditure elasticity for average purchases differs between the fish types. For wild fish the total expenditure elasticity for purchase frequency is 0.15, for average purchases 0.19 and for total purchases 0.34.

(Table 6 about here)

The differences in absolute values between the habit, own-price and total expenditure elasticities for the purchase frequencies, average purchases and total purchases for the three types of fish are shown in Table 7. The *t*-values for tests of no difference are also shown. There is no statistically significant difference between any of the stock variables. Thus, the habits for purchasing the three types of fish are very similar. However, there are several significant difference between the own-price elasticities for purchasing frequencies, average purchases and total purchases among the three types of fish. For example, the purchase frequency of wild fish is more sensitive to price changes than the purchase frequency of farmed or other fish. The total expenditure elasticities for average purchases are statistically different at the 5% level for all the three types of fish. Farmed fish is less elastic with respect to total expenditure than wild fish that is less elastic than other fish.

(Table 7 about here)

5.3 Example

Let's assume that a supermarket wishes to change the prices of fish to increase overall profit, using the results above. What we would like to acknowledge is that a price change of one or two fish categories will have a relatively small effect on revenue, but the sales of other products will make up for most of the revenue through changes in frequency and thus habits, since frequencies are the driving force of habits.

First, we assume that the supermarket owner disregards the importance of frequencies and increases the price of wild fish from 9.4 EUR/Kg to 10.3 EUR/Kg for two periods (i.e. two months). This will lead to a revenue increase of 0.08 EUR per customer on average. However, this is a supermarket and customers do not only buy fresh fish. Let us assume that on average consumers purchase other products, for example, for 150 ERU in each trip. Then if we note that the frequency change from the price change of wild fish is -0.00147 then we will decrease revenue by 0.22 EUR, leading to a net loss of 0.138 EUR per customer during the two months.

Now if we assume that the supermarket owner knows the importance of habits and frequencies he can turn this previous loss to a gain. Now what owner wants is to produce as high frequency as possible by lowering price in such a way that revenue loss from the price reduction is minimal. Now let us assume that we lower the price of wild fish from 9.4 EUR/Kg to 8.5 EUR/Kg, and lower the price of farmed fish from 8.5 EUR/Kg to 7.6 EUR/Kg. Then the revenue loss from the price reduction is -0.15 EUR over these two months. However, the frequency change is positive and equal to 0.00195 and if we assume the 150 EUR expenditure on other goods in each trip we will have a revenue gain of 0.58 EUR per individual over the two months. This will lead to a net gain of 0.43 EUR per customer on average during the two months.

Thus, by taking account of habits and that their driving force is frequencies it is possible to make a profit from simple price changes. The revenue gains above are quite small for each individual during these two months, but the reason for that is the extremely low frequency of purchase for fresh fish in France. If we would look at customers who purchase once a month and not once a year the numbers would be even more in favor of the pricing strategy.

6. Conclusions

This paper extends the theoretical model of Spinnewyn (1981) and Muellbauer and Pashardes (1992) in two ways. First, we specify the dynamic structure of habit formation and duration such that it can be applied to semi-logarithmic demand specifications. Second, we allow the service stock to be derived from the purchase frequency of a good and the average purchased quantity, instead of the total purchased quantity. This separation can be useful for marketing purposes and allows us to calculate habit formation elasticities for purchase frequency, average quantity, and total quantity.

The paper also contributes to the econometric literature in two ways. First, a Bayesian framework for the joint estimation of a demand system for the purchasing frequencies of different goods and a second demand system for average purchased quantities of these goods is introduced. We use the multivariate Poisson log-normal and the multivariate gamma log-normal to model the data generating processes of purchase frequencies and average quantities. The Bayesian framework allows for unrestricted covariance structures within each demand system. Second, as far as we know, the mean and variance of the marginal distribution of the multivariate gamma log-normal distribution has not been previously derived, and we derive the mean and the variance by using the methods of iterative expectations.

We provide an empirical illustration of the econometric model by using French scanner data for fish purchases. We include fresh wild fish, fresh farmed fish, and other fresh fish in our demand model. We find that habits in total purchases almost solely originate from habits in purchase frequencies, while habits in average purchased quantities are of minor importance. Contrary to the effects of habits, changes in total purchases in response to price changes is mainly determined by the changes in average quantities purchased, and not frequencies.

We then provide an example of a pricing strategy which utilizes the aforementioned results to increase revenue. We show that a supermarket owner could increase revenue from increasing the price of wild fish, but will then lose money on the sales of other products due to the decrease in shopping frequency. We then show that if the average per kilo price of wild and farmed fish is decreased in a particular way it is possible to make a minor loss from the price reduction and a profit from sales of other products, due to increased purchase frequency, which leads to a net increase in revenue.

Possibilities for future research. First, allowing for unrestricted covariance structure between systems would be desirable. Second, applying the model to a more disaggregate group of products. Thirdly, estimating a larger system with fish types along with meat for example.

References

Adamowicz, W.L., and Swait, J. D. (2012). Are Food Choices Really Habitual? Integrating Habits, Variety-Seeking, and Compensatory Choice in a Utility-Maximizing Framework. *American Journal of Agricultural Economics* 95, 17-41.

Aitchison, J., and Ho, C. H. (1989). The Multivariate Poisson-log Normal Distribution. *Biometrika*, 76, 643-653.

Alessie, R., and Kapteyn, A. (1991). Habit Formation, Interdependent Preferences and Demographic Effects in the Almost Ideal Demand System. *The Economic Journal* 101, 404-419.

Allais, O., Bertail, P., and, Nichele, V. (2010). The Effect of a Fat Tax on French Households' Purchases: A Nutrition Approach. *American Journal of Agricultural Economics* 92, 228-245.

Arnade, C, and M. Gopinath. (2006). The Dynamics of Individuals' Fat Consumption. *American Journal of Agricultural Economics* 88. 836-850.

Asche, F., Guttormsen, A.G., Sebulonsen, T., and Sissener, E. H. (2005). Competition between farmed and wild salmon: the Japanese salmon market. *Agricultural Economics* 33, 333-340.

Asche, F., Guttormsen, A.G. (2014). Seafood Markets and Aquaculture Production: Special Issue Introduction. *Marine Resource Economics* 29, 301-304.

Bertail, P., and, Caillavet, F. (2008). Fruit and Vegetable Consumption Patterns: A Segmentation Approach. *American Journal of Agricultural Economics* 90, 827-842.

Bhattacharya, C.B. (1997). Is Your Brand's Loyalty too Much, too Little, or Just Right? Explaining Deviations in Loyalty from the Dirichlet Norm. *International Journal of Research in Marketing* 14, 421-435.

Chib, S. and Winkelmann, R. (2001). Markov Chain Monte Carlo Analysis of Correlated Count Data. *Journal of Business and Economic Statistics* 19, 428-435.

Creel, M.D., and Loomis, J.B. (1990). Theoretical and Empirical Advantages of Truncated Count Data Estimators for Analysis of Deer Hunting in California. *American Journal of Agricultural Economics* 72, 434–441.

Deaton, A. (1997). The Analysis of Household Surveys. Baltimore: Johns Hopkins University.

Deb, P., and Trivedi, P.K. (2002). The Structure of Demand for Health Care: Latent Class versus Two-Part Models. *Journal of Health Economics* 21, 601-625.

Dynan, K.E. (2000). Habit Formation in Consumer Preferences: Evidence from Panel Data. *The American Economic Review* 90, 391-406.

Egan, K., and Herriges, J. (2006). Multivariate Count Data Regression Models with Individual Panel Data from an On-Site Sample. *Journal of Environmental Economics and Management* 52, 567-581.

Fuhrer, J. C (2000). Habit Formation in Consumption and Its Implications for Monetary-Policy Models. *Journal of Political Economy* 90, 367-3

Heaton, J. (1995). An Empirical Investigation of Asset Pricing with Temporally Dependent Preference Specifications. *Econometrica* 63, 681–717.

Hellerstein, D.M. (1991). Using Count Data Models in Travel Cost Analysis with Aggregate Data. *American Journal of Agricultural Economics* 73, 860–866.

Herrmann, M.L. Mittelhammer, R.C., and Lin, B. (1993). Import Demands for Norwegian Farmed Atlantic Salmon and Wild Pacific Salmon in North America, Japan and the EC. *Canadian Journal of Agricultural Economics* 41, 111-125.

Holt, M. T., and B. K. Goodwin. (1997). Generalized Habit Formation in an Inverse Almost Ideal Demand System: An Application to Meat Expenditures in the U.S. Empirical Economics 22, 2

Kau, A.K., and Ehrenberg, A.S.C. (1984). Patterns of Store Choice. *Journal of Marketing Research* 21, 399-409.

Meghir, C., and Robin, J.M. (1992). Frequency of Purchase and the Estimation of Demand Systems. *Journal of Econometrics* 53, 53–85.

Muellbauer, J., and Pashardes, P. (1992). Tests of Dynamic Specification and Homogeneity in a Demand System. In *Aggregation, Consumption, and Trade*, ed. L. Philips and L.S. Taylor, 55-98. Boston: Kluwer Academic Publishers.

Munkin, M.K., and Trivedi, P.K. (1999). Simulated Maximum Likelihood Estimation of Multivariate Mixed-Poisson Regression Models, with Application. *Econometrics Journal* 2, 29-48.

Nylander, J.A.A., Wilgenbusch, J.C., Warren, D.L., and Swofford, D.L. (2008). AWTY (are we there yet?): a system for graphical exploration of MCMC convergence in Bayesian phylogenetics. *Bioinformatics* 24, 581-583.

Pashardes, P. (1986). Myopic and forward looking behavior in a dynamic demand system. International Economic Review 27, 287-297.

Pollak, A.R. (1970). Habit Formation and Dynamic Demand Functions. *Journal of Political Economy* 78, 745-763.

Rickertsen K., J.A. Chalfant, and M. Steen (1995). "The Effects of Advertising on the Demand for Vegetables," *European Review of Agricultural Economics*, 22, 481-494.

Rickertsen, K. (1998). "The Demand for Food and Beverages in Norway," *Agricultural Economics*, 18:89-100.

Roberts, G.O., Gelman, A., and Gilks, W.R. (1997). Weak Convergence and Optimal Scaling of Random Walk Metropolis Algorithms. *The Annals of Applied Probability* 7, 110-120.

Robin, J.M. (1993). Analysis of the Short-Run Fluctuations of Households' Purchases. *The Review of Economic Studies* 60, 923-934.

Smith, V.K. (1988). Selection and Recreation Demand. *American Journal of Agricultural Economics* 70, 29-36.

Spinnewyn, F. (1981). Rational Habit Formation. *European Economic Review* 15: 91–109.

Uncles, M., and Lee, D. (2006). Brand Purchasing by Older Consumers: An Investigation Using the Juster Scale and the Dirichlet Model. *Marketing Letters* 17, 17-29.

Wang, P. (2003). A Bivariate Zero-Inflated Negative Binomial Regression Model for Count Data with Excess Zeros. *Economics Letters* 78, 373-378.

Zhen, C., Wholgenant, M.K., Karns, S., and Kaufman, P. (2011). Habit Formation and Demand for Sugar-Sweetened Beverages. *American Journal of Agricultural Economics* 93, 175-193.

Table 1: Descriptive Statistics

Variable	Description	Mean	Std. Dev	Min	Max
FWild	Number of trips to buy wild fish	1.51	1.65	0.00	27.00
FFarmed	Number of trips to buy farmed fish	0.94	0.88	0.00	14.00
FOther	Number of trips to buy other fish	1.31	1.20	0.00	16.00
QWild	Truncated average purchased quantities of wild fish	669.75	631.80	14.66	15172.20
QFarmed	Truncated average purchased quantities of farmed fish	701.00	642.89	20.00	11175.00
QOther	Truncated average purchased quantities of other fish	658.87	549.70	11.80	10475.40
WLag of Stock	The lagged service stock of wild fish	1657.54	2204.19	0.00	34650.19
FLag of Stock	The lagged service stock of farmed fish	504.33	837.95	0.00	19012.51
OLag of Stock	The lagged service stock of other fish	924.20	1208.82	0.00	21364.52
Expendit ures	Real household expenditures on fresh fish	0.14	0.14	10·0.01	2.52
WPrice	Real price wild fish (per kilo)	0.10	0.03	10·0.03	0.59
FPrice	Real price of farmed fish (per kilo)	0.09	0.02	10·0.02	0.58
OPrice	Real per of other fish (per kilo)	0.10	0.03	10·0.07	0.58

Table 2: Posterior Summary for Wild Fish Based on the MPLN and MGLN.

	Frequencies			Average Purchases		
Variable	**Mean**	**t-value**	**Geweke Z**	**Mean**	**t-value**	**Geweke Z**
Constant	0.06	3.45	-0.72	6.47	501.39	2.28
WLag of Stock·10,000	0.40	22.40	-	0.05	3.08	-
WPrice	-1.12	-10.56	0.15	-4.82	-68.81	-3.61
Expenditure	1.06	75.99	-0.06	2.20	140.40	0.18
Dummy06·10	-0.84	-9.20	-0.57	0.07	8.42	-0.53
Dummy07	-0.11	-12.21	-1.50	0.09	10.14	-0.37
Dummy08	-0.15	-15.31	0.56	0.10	11.58	0.11
Months·10	-0.11	-11.76	0.31	0.03	3.70	-0.74

Note: Notes: Time06, Time07 and Time08 are annual dummy variables, which takes the value of 1 in the indicated years. Months is a monthly time trend. Geweke *Z* provides the *Z*-value for a test of stationarity of the Markov chains. The Geweke *Z* could not be calculated for the lag of stock variable, due to computational issues. The multiplication in front of variable names indicate scaling.

Table 3: Posterior Summary for Farmed Fish Based on the MPLN and MGLN.

	Frequencies			Average Purchases		
Variable	**Mean**	**t-value**	**Geweke Z**	**Mean**	**t-value**	**Geweke Z**
Constant	-0.48	-18.22	1.84	6.57	373.48	1.22
Lag of						
Stock·1000	0.13	28.20	-	0.02	4.74	-
Price	-0.64	-3.35	-0.85	-5.94	-53.27	-0.64
Expenditure	1.06	75.99	-0.06	2.20	140.40	0.18
Dummy06	-0.09	-7.37	-0.19	0.05	4.60	-1.05
Dummy07	-0.03	-2.46	-1.21	0.06	4.81	-1.62
Dummy08	-0.05	-4.06	-1.48	0.09	7.56	-1.06
Months·10	-0.04	-2.72	-0.15	0.05	4.46	-0.27

Note: Time06, Time07 and Time08 are annual dummy variables, which takes the value of 1 in the indicated years. Months is a monthly time trend. Geweke *Z* provides the *Z*-value for a test of stationarity of the Markov chains. The Geweke *Z* could not be calculated for the lag of stock variable, due to computational issues. The multiplication in front of variable names indicate scaling.

Table 4: Posterior Summary for Fish with Unknown Origin Based on the MPLN and MGLN.

	Frequencies			Average Purchases		
Variable	**Mean**	**t-value**	**Geweke Z**	**Mean**	**t-value**	**Geweke Z**
Constant	0.03	1.55	0.68	6.45	500.29	0.74
Lag of Stock·1000	0.08	24.75	-	0.01	4.57	-
Price	-0.71	-6.19	0.13	-4.76	-63.77	0.19
Expenditure	1.06	75.99	-0.06	2.20	140.40	0.18
Dummy06	-0.09	-9.56	-0.06	0.08	9.65	-0.80
Dummy07	-0.12	-11.76	-1.36	0.09	10.34	-1.32
Dummy08	-0.14	-13.44	-0.43	0.10	12.66	-1.42
Months·10	-0.07	-7.02	-1.08	0.05	5.66	-0.07

Note: Time06, Time07 and Time08 are annual dummy variables, which takes the value of 1 in the indicated years. Months is a monthly time trend. Geweke *Z* provides the *Z*-value for a test of stationarity of the Markov chains. The Geweke *Z* could not be calculated for the lag of stock variable, due to computational issues. The multiplication in front of variable names indicate scaling.

Table 5: Posterior Summary Cross-Equation Covariance Matrix from MPLN and MGLN.

	Frequencies			Average Purchases		
Variable	**Mean**	**t-value**	**Geweke Z**	**Mean**	**t-value**	**Geweke Z**
Sigma11	0.55	36.15	0.45	0.13	36.17	-0.03
Sigma12	0.08	8.52	0.32	0.10	32.21	1.27
Sigma13	-0.15	-20.81	-0.10	0.10	34.16	0.11
Sigma21	0.08	8.52	0.32	0.10	32.21	1.27
Sigma22	0.60	33.55	-1.29	0.13	30.63	0.30
Sigma23·10	-0.05	-0.64	-0.93	0.92	32.31	1.28
Sigma31	-0.15	-20.81	-0.10	0.10	34.16	0.11
Sigma32·10	-0.05	-0.64	-0.93	0.92	32.31	1.28
Sigma33	0.39	36.06	-0.80	0.11	34.16	-0.44

Note: Sigma11 Sigma22, and Sigma33 represent the variance of the random effects of demand equation 1-3, respectively. The other Sigma estimates are covariance parameters of the random effects between equations. Geweke *Z* provides the *Z*-value for a test of stationarity of the Markov chains. The multiplication in front of variable names indicate scaling.

Table 6: Elasticities Purchase Frequencies, Average Purchases, and Total Purchases

	Purchase Frequencies					
	Wild Fish		Farmed Fish		Other Fish	
	Estimate	t-value	Estimate	t-value	Estimate	t-value
Lag of Stock·10	0.66	22.40	0.65	28.20	0.70	24.75
Price	-0.12	-10.56	-0.06	-3.35	-0.07	-6.19
Expenditure	0.15	75.99	0.15	75.99	0.15	75.99
	Average Quantities					
	Wild Fish		Farmed Fish		Other Fish	
	Est.	t-val.	Est.	t-val.	Est.	t-val.
Lag of Stock·100	0.50	3.08	0.47	4.74	0.65	4.57
Price	-0.32	-68.81	-0.27	-53.27	-0.31	-63.77
Expenditure	0.19	140.40	0.15	140.40	0.20	140.40
	Total Quantities					
	Wild Fish		Farmed Fish		Other Fish	
	Est.	t-val.	Est.	t-val.	Est.	t-val.
Lag of Stock·10	0.71	21.20	0.70	27.77	0.77	24.16
Price	-0.43	-36.17	-0.33	-17.78	-0.38	-30.33
Expenditure	0.34	143.05	0.30	134.27	0.35	143.73

The multiplication in front of variable names indicate scaling.

Table 7: Elasticity Differences

	MPLN		MGLN		Combined Model	
	Difference	t-value	Abs. Diff.	t-value	Abs. Diff.	t-value
(Lag Wild - Lag Farmed)·100	0.12	0.33	0.02	0.10	0.14	0.34
(Lag Wild - Lag Other)·100	0.39	-0.96	0.16	-0.77	0.56	-1.21
(Lag Farmed - Lag Other)·100	0.52	-1.42	0.18	-1.05	4.66	-1.73
Price Wild - Price Farmed	0.06	-2.66	0.04	-6.25	0.10	-4.46
Price Wild - Price Other	0.04	-2.81	0.01	-1.05	0.05	-3.00
Price Farmed - Price Other	0.01	0.53	0.04	5.11	0.05	2.10
Exp. Wild – Exp. Farmed	-	-	0.04	24.40	0.04	13.05
(Exp. Wild – Exp. Other)·10	-	-	0.04	-2.06	0.04	-1.19
Exp. Farmed - Exp. Other	-	-	0.05	-26.37	0.05	-14.23

The multiplication in front of variable names indicate scaling.

Paper 4

Fond of fish? A Count Data Analysis of How Frequently French Consumers Purchase Seafood

Arnar Buason[1] and Sveinn Agnarsson[2]

[1] School of Economics and Business, Norwegian University of Life Sciences, P.O. Box 5003, 1432 Ås, Norway. Corresponding author is Arnar Buason: arnar.buason@nmbu.no.
[2] Department of Economics, University of Iceland, Reykjavik, Iceland.

Abstract

France is one of the largest markets for fish in the Europe. This article estimates a theory-consistent demand system for frequencies of purchases fish in France. We include five types of fish and use scanner data. The results show that consumers have very heterogeneous preferences for purchasing different types of fish. As an example, the average consumer of fresh salmon differs substantially from the consumer of frozen white fish. The typical consumer of fresh salmon is a healthy upper-class individual with university education who comes from a small household in Paris or the north of France, but the average frozen white fish consumer is an older, lower, middle to lower class individual who comes from a large household in the south France.

Key Words: Frequency of purchase, negative binomial, demand, EU, France.

1. Introduction

After China, Indonesia, and India, the European Union (EU) is the fourth largest producer of sea food in the world in value terms. EU produced more than 3% of the global production in 2013 and is the largest fish trader[63] in the world (EUMOFA, 2015). The demand for sea food in the EU has been growing in recent years, and EU is a large net importer of seafood. In 2014 the value of EU seafood import was four times higher than the value of total meat import (EUMOFA, 2015). The high seafood import was mainly due to increased imports of salmon and shrimp, and most of the imports are either frozen or prepared products.

Among the EU countries France is one of the largest consumer markets for seafood and the biggest market for salmon (Xie and Myrland, 2011). Even though historically France is a great fishing nation, today it produces a relatively small fraction of its seafood

[63] Measured as all EU exports plus imports.

consumption, meaning that the majority of French seafood consumption is imported (EUMOFA, 2015).

Only 37% of French consumers eat fish two times a week, as is recommended by health authorities. 34% of French consumers only eat fish 2-3 times a month or less. Whereas the frequent consumers are older individuals with higher income (Norwegian Seafood Council, 2016).

A great range of seafood is available in the French fish market such as; salmon, cod, shrimp, saithe, trout, whiting, sea bream, sea bass, nile perch, sardine, pangasius, sole, mackrel, skate, etc. The most popular is actually canned tuna, but after that comes salmon and cod, both in terms of value and frequency of purchases (Norwegian Seafood Council, 2016). It is therefore that out of all the possible fish products available in the French market this study focuses on the two most important; salmon and cod. More specifically, fresh salmon, frozen Salmonidae, fresh cod and frozen white fish.[64]

The purpose of this research is to study the French fish market and specifically the characteristics of consumers who purchase different types of fish. We approach the problem from the angle of purchasing behavior, that is, we estimate how frequently households purchase various types of fish by estimating a demand system for purchase frequencies rather than quantities. Our model combines aspects of demand system analysis with the count data approach typically used in the marketing literature.

Demand analysis aims at understanding how demand relates to prices, income, and socioeconomic variables. Using this framework, the demand for fish in France has been thoroughly studied by, for example, Asche et al. (2011) who analyzed the demand growth of Atlantic Salmon in the EU and France, and Gobillon and Wolff (2015) who investigated

[64] It would be desirable to include more types of fish, but due to computational issues this is difficult and we aggregate the frozen products into salmonide and white fish. Fresh salmon and cod are the main focus, due to their importance, and are therefore specified as their own category.

spatial variations in product prices in French fish markets. Onozaka et al. (2014) analyzed the relationship between consumer perception and salmon consumption frequencies using a latent class model. Xie and Myrland (2011) applied an empirical test for the aggregation levels of French household demand for salmon.

The marketing literature has focused more on count data models, which have been widely applied for evaluating brand success, brand loyalty, and store choice. Kau and Ehrenberg (1984) used the negative binomial Dirichlet[65] model to predict store choice. Uncles et al. (1995) provide a review on buyer regularities based on predictions from the negative binomial (NB) Dirchlet. Bhattacharya (1997) estimated deviations from brand loyalty and compared these deviations with the predictions from the Dirchlet model, and Uncles and Lee (2006) estimated the purchase frequency of different age groups using predictions from the NB Dirichlet model.[66]

Some studies have applied count data models on demand systems (Meghir and Robin, 1992; Robin, 1993). However, these studies only used the estimated probabilities to adjust conventional demand models to account for the actual purchase frequency as an alternative to estimate the choice of whether to purchase or not.[67] By adjusting the demand system with the probabilities calculated from the choice of how often to goes to the store the model utilizes more of the information in the data compared to when probabilities are calculated from the choice of purchasing or not. We follow the basic setup of Meghir and Robin (1992) but add a budget constraint where the consumer's total income depends on income and other transfers.

[65] The negative binomial Dirichlet model has two stages; the first is based on the multivariate beta distribution known as the Dirichlet distribution, and the second is a Poisson gamma mixture which produces a variant of the negative binomial model. The Dirichlet is assumed to be the data generating process (DGP) of some choice and the negative binomial is assumed to be the DGP of the frequency of the corresponding choice.

[66] The Dirichlet distribution is generally not used in economics, but one example is Shonkwiler and Anglin (2005) who used it to estimate the willingness to pay for removing grazing land from hiking trails.

[67] Count data models have also been used to estimate recreational demand and the demand for health care. The Poisson and negative binomial have been applied, for example, by Creel and Loomis (1990), and Hellerstein (19991) to estimate recreational demand. Munkin and Trivedi (1999), Deb and Trivedi (2002), and Wang (2003) estimated the demand for health care.

To demonstrate the potential usefulness of this model, we include an empirical example where we use French scanner data to estimate a system of demand equations in terms of purchase frequencies.

The rest of the paper is organized as follows. Section two presents the microeconomic model. Section three describes the statistical model and distribution assumptions. In section four, we provide a description of our data, as well as the empirical specification of the microeconomic model. Section five presents our results and section six concludes.

2. Microeconomic Model

The consumer faces three decision variables *l*, $c = (c_1, c_2, \dots, c_M)'$, and $n = (n_1, n_2, \dots, n_M)'$, where *l* is leisure, *c* is a vector denoting total purchases on each shopping occasion of *M* different goods, and *n* is a vector of the corresponding purchase frequencies. We assume that the consumer's utility function, $u(l, c, n)$, is weakly separable in *c* and *n*, and quasi-concave in *l*, *c*, and *n*. The consumer has wage income $y = wh$, where *h* is the hours spent working and *w* is the hourly wage rate, and receives other transfers, *R*. The total available time is *T*, and it is divided between total hours worked *h*, time spent on purchasing goods $L(n)$, and leisure time *l*. $L(n)$ is assumed to be increasing in *n*, and the consumer's optimization problem is:

$$\max_{l,c,n}\{u(l,c,n): wT + R = p'c + wl + wL(n), l > 0, c > 0, n > 0\}, \tag{1}$$

where *p* is a vector of prices corresponding to *c*. Here we assume that the opportunity cost of time is *w* and assumed to be the same for any activity, work shopping and leisure. This is a simplifying assumption which is not deemed to pose any problems for the analysis, but should be explored in future research. The solution to this optimization problem are three sets of Marshallian demand equations, $l(p, w, R)$, $n_i(p, w, R)$, and $c_i(p, w, R)$, where $i = 1,2, \dots,$.

3. Statistical Model

In the microeconomic model above, the consumer is faced with three decisions; how much time to spend on leisure, how much to purchase in terms of quantities and how often to purchase goods. In this article, we only estimate the purchasing frequency decision. Let us assume that the number of shopping trips is generated by a discrete distribution with a probability mass function $f_N(n|Z_i)$ for $n = 0,1,2,...$ where Z_i is a matrix of exogenous variables, then the probability of observing $n = 0$ is given by $\Pr(N = 0|Z_i)$.

To be able to estimate the model, we specify a probability mass function of *n*. In count data estimation usually a Poisson distribution is assumed as in Meghir and Robin (1992). This is a valid choice since quasi-maximum likelihood will lead to an unbiased estimation even though the distribution assumptions are incorrect as long as the mean is correctly specified. However, due to the Poisson limitations of equidispersion we assume the negative binomial distribution instead:

$$f(n_i|Z_i) = \frac{\Gamma(\theta + n_i) r^{\theta} (1-r)^{n_i}}{\Gamma(1+n_i)\Gamma(\theta)}, r = \frac{\theta}{\theta + \lambda}, 0,1,2, ... \qquad (2)$$

The conditional mean of n_i is then $\mathrm{E}(n_i|Z_i) = \lambda_i$ and the conditional variance is $\mathrm{V}(n_i|Z_i) = \lambda\left(1 + \frac{1}{\theta}\lambda\right) = \lambda(1 + \kappa\lambda) = (\lambda + \kappa\lambda^2)$. This specification of the negative binomial model is known as the NB2, due to the square of the lambda parameter in the variance specification.[68] As in Meghir and Robin (1992) we do not assume a stochastic relationship between different $\mathrm{E}(n_i|Z_i) = \lambda_i$. To be able to estimate a count data system with a dimension greater than two or three and an unrestricted covariance matrix one must use simulation based methods, see for example Chib and Winkelmann (2001).

4. Data and Empirical Specification

[68] See Cameron and Trivedi (2013) for details.

The data is a rotating consumer panel[69] of fish purchases in France from 2010-2013, collected by Kantar Worldpanel[70]. The data consists of weekly data on fish purchases of French households and their sociodemographic information. The recorded sociodemographic information is extremely detailed and includes everything from age to the number of cats. During each year, there are around 20,000 households in the panel with an annual rotation of one third of the participants. Each household then remains in the panel for an average of all four years. In 2010, 2011, 2012 and 2013 the respective number of participating households are 18,494, 18,651, 19,729 and 20,623. Households are selected by stratification according to a few socioeconomic variables. The total number of household in the survey during the four years is 43,127. The annual total frequency of fish purchases in 2010, 2011, 2012 and 2013 respectively is 171,013, 165,143, 171,909 and 168,856. Thus, total annual purchase frequency is rather stable during the sample period. The total number of observations for all four years is 96,917.

The Kantar purchase data is collected with the use of bar codes. Kantar provides all surveyed households with a hand-held scanner and other relevant equipment to register purchases at home. Due to this form of data collection, this type of data has become known as scanner data. To register purchases without a bar code, households are assigned to two groups in order to reduce their workload. Each group is then required to register their purchases for specific types of food. In our study, we only analyze the fish purchase data and therefore assume weak separability among different food categories in the consumer's utility function. Thus, assuming that the utility maximizer chooses her consumption of different food categories separately with respect to the share of income used to spend on each product group. This weak spearability assumption is common in applied food demand analysis, see

[69] Here, a rotating consumer panel refers to a panel of served households where new households are added when others are inactive, to keep a stable number of households in the data collection process.
[70] Kantar Worldpanel, former TNS Worldpanel, is an international company which focuses on data collection and consultancy in consumer markets.

for example Rickertsen et al. (2003) and Gustavsen and Rickertsen (2003). Scanner data have frequently been used in food demand analysis (e.g., Allais et al., 2010).

A common problem encountered in scanner data sets is the large share of zero observations. The data is aggregated over years to reduce the share of zero purchases and also the likelihood of autocorrelation. We have specified a demand system consisting of fresh salmon, fresh cod, frozen Salmonidae, frozen white fish, and all other fish. Table 1 shows the empirical purchase frequency of these types of fish. The first column shows the frequencies, the second column shows the corresponding frequencies observed in the data over the four years, and finally column three denotes the percentage of observations with the observed frequencies. Even though the data has been aggregated over time they there is still significant share of zero observations. We see that most consumers only purchase from these categories once during the four years. However, we see that a significant share of consumers purchase from the other category more than twenty times over the sample period, namely we observe 4321 households who purchase more than 20 times over the four years.

(Table 1 about here)

The data set does not contain any information regarding prices of the products. Prices are therefore estimated by dividing expenditures by quantities purchased of each good in each shopping trip to create unit values. When zero purchases are recorded, there is no price available so the yearly average price is used. This approach is used in other demand studies such as Allais et al. (2010)[71] and Bertail and Caillavet (2008). However, these constructed prices will be influenced by the consumer's choices of quality, as has been pointed out by Deaton (1997). Thus, higher income households will tend to purchase higher quality fish and

[71] They use scanner data collected by TNS Worldpanel. TNS is the former name of Kantar which collected the data used in our study.

therefore the unit values will be positively related to income and expenditure, which will cause the price effects to be biased upwards.

The constructed unit values include a few outliers, which we believe are due to errors in the data recording process. We dropped the price observations which were more than 15 standard deviations from the mean, in total 22 observations. However, the dropped do not have any significant effect on the estimates.

Table 2 provides descriptive statistics of all variables used in the analysis, except for time dummies. The first five variables in the table are the dependent variables in the demand system, mentioned above, showing a fairly low average frequency during the four years. However, it should be noted that none of these five mean estimates are statistically different from zero, and should therefore not receive too much attention. The estimated prices and total fish expenditures are normalized by dividing each price with the average French consumer price index. This normalization gives real prices and a demand system that is homogenous of degree zero in prices and expenditures.

(Table 2 about here)

We include many sociodemographic variables in the analysis to analyze the characteristics of consumers who purchase what type of fish. The choice of variables is determined mainly by convention established in the literature, see for example Allais et al. (2010). The average family size is 2.6. The average age of the household head is about 46 years[72]. Age is a standard variable to include in a demand study, but is of even greater interest when analyzing fish demand. Bourre and Paquotte (2008) show that older people in France are not consuming enough fish, and according to the French recommended dietary allowance (RDA) for people who are 65 years or older, seafood provides many of the necessary

[72] The household head is the one who makes most of the purchasing decisions.

nutrients[73]. Low intakes of these nutrients in older consumer's diets can thus be solved by increasing fish consumption. It is thus a good marketing opportunity to advertise the health benefits of fish in general and especially for the elderly. We include the body mass index of the household head (BMI), since weight is a good health indicator. The average BMI is around 24.8, which indicates that the average household head is at a healthy weight.

We include four dummy variables for social class: upper class, upper middle class, lower middle class, and lower class.[74] We include two non-conventional variables, the number of cats and the number of motor vehicles. Cats can potentially provide a significant contribution to fish consumption, and a motor vehicle may facilitate a low shopping frequency in general due to the possibility to purchase a lot of food on each occasion. Dummy variables for different levels of education are included: Secondary education or less, high school education or equivalent, and university education. The sample is well educated with over 50% with high school education or equivalent, and close to 40% with university education. Furthermore, we include dummy variables for different regions in France: South, north, east, west, central, and Paris metropolitan area. Finally, we follow Allais et al. (2010) and include dummy variables for housing status: Housing owner, tenant, and free accommodation.

As is conventional when estimating count data models the conditional expectation is defined as a semi-logarithmic function.[75] Our demand function is thus given by the following expression:

$$\mathrm{E}\left(n_{ij}\middle|Z_i\right) = \exp\left(\beta_{ij} + \sum_{s=1}^{M} \beta_{is}\left(p_{sj}/CPI\right) + \theta_i\left(y_j/CPI\right)\right) \tag{3}$$

[73] Fish, for example, provides 25% of vitamin RDA, 56% of the vitamin B12 RDA, 28% of iodine RDA, 23% of selenium RDA, and 203% of DHA RDA (Bourre and Paquotte, 2008).
[74] Class variables are specified by Kantar.
[75] For an application and theoretical restrictions of a semi-logarithmic functional form using count data see for example Shonkwiler and Englin 2005.

for $i = 1,2, \dots, M$ fish categories and $j = 1,2, \dots, J$ households. The price of fish category i for household j is denoted by p_{ij}. The total fish expenditure of household j is given by y_j, and *CPI* is the French consumer price index. For the demand system to be consistent with economic theory we need to impose a number of restrictions on the parameters. LaFrance and Hanemann (1989) and LaFrance (1990) derived the restrictions for this specific functional form, among others. In our case, we must have $\beta_{is} = 0 \; \forall \; i \neq s$ and $\theta_i = \theta \; \forall \; i$ to impose symmetry. As discussed above, we imposed homogeneity of degree zero by dividing prices and total expenditure by the French CPI. We then model household heterogeneity in equation (3) by specifying:

$$\beta_{ij} = \eta_{0i} + \sum_{k=1}^{K} z_{kj}\eta_{ki} \tag{4}$$

where z_{kj} represents household characteristics and η_{ki} are parameters to be estimated. Finally, it should be noted that even though cross-price effects are zero, $\beta_{is} = 0$, compensated cross price effects are not. From the Slutsky equation we have:

$$e_{isj} = n_{ij}\frac{\partial n_{sj}}{\partial y_j} = \theta n_{ij} n_{sj} \tag{5}$$

where e_{isj} is the compensated substitution effect between products *i* and *s* for household j, and the *n*'s are purchase frequencies of different product groups by household j.

5. Results

Table 3 shows the results from the negative binomial estimation. The five demand equations were estimated as a system.

(Table 3 about here)

From the table, we see that there is no effect the same for all types, but there are still some consumer characteristics that are generally more common among fish consumers than others. In genera fish consumers are older individuals who are healthy, in terms of weight, The typical French fresh salmon consumer is a healthy upper-class individual with university education who comes from a small household. Furthermore, this individual owns his own housing and has an above average number of motor vehicles, a further indicator that she is well-off. This consumer comes from the Paris metropolitan area or the north of France. The time dummies are all positive and significant, showing that the demand for fresh salmon has increased over the sample period. These time dummies capture effects on fresh salmon demand which are not otherwise accounted for in the model. One of the factors which could be affecting these variables, other than positive publicity, are cross-price effects, which are not accounted for in the model. These effects should be positive indicating that fresh salmon and the other four categories are substitute goods.

The frozen Salmonidae consumer is a healthy, younger, lower middle class individual with university education. Furthermore, this individual owns an above average number of cats. The consumer comes from the east, the south, or the Paris metropolitan area. Only the 2011 time dummy is statistically significant. The time dummy has a negative sign showing a reduction in demand between 2010 and 2011. This might be an indicator of a transition from frozen to fresh salmon.

The average fresh cod consumer comes from either of two groups of people: both are healthy, older, upper middle class individuals. The former one has only secondary education while the latter one has university education. Both groups come from a small household and do not own their own housing. These consumers are found in the east and west parts of France. The time dummies for 2011 and 2013 are significant, but have alternate signs. The demand for fresh cod in 2011 decreased relative to 2010, but then increased in 2013 relative to 2010. It is difficult to say exactly what causes these effects, but these variables might be picking up substitution effects and preference changes.

The average frozen white fish consumer is an older, lower middle to lower class individual from the south of France who comes from a large household. The BMI parameter is positive and significant, at the 10% level, which could indicate that consumers in this group are not as healthy as in the other categories. Several of the product forms in this category, e.g. fish fingers, are less healthy than those for fresh salmon and cod. The household head comes from a large family and is from a lower class than the fresh fish consumers and is thus more likely to buy cheaper less healthy food. The time dummies for 2012 and 2013 are significant and show that demand increased in 2012 but decreased in 2013, relative to 2010.

The consumer of other fish is a healthy, older, lower middle class individual with secondary education who comes from a large household in the center of France. This category includes a great variety of different fish products, but two of the largest ones are canned tuna and shrimp. The time dummy for 2011 was significant and shows that demand decreased that year relative to 2010.

Table 4 presents elasticities with respect to own-price, total expenditure, age, and BMI. The results indicate that the own-price elasticities of different product groups varies significantly between frozen and fresh fish. The own-price elasticities of fresh salmon and fresh cod are quite low demonstrating that price changes have little effect on these products purchase frequency. Frozen fish on the other hand has a significantly higher own-price elasticity, which is further evidence that price is a deciding factor for the consumer purchasing decision. This price sensitivity is consistent with consumers who are less well-off and therefore need to think more about the price before she feeds her family. Thus, campaigns reducing the price of frozen fish is a good way to increase store traffic. Furthermore, increasing the price of fresh salmon and cod at the same time as the price of frozen fish is lowered and advertised could make up a good short-term pricing strategy.

The age elasticities[76] indicate that sellers should target their sales of fresh cod and frozen white fish towards older individuals, but sales of frozen Salmonidae towards younger consumers. And as was mentioned previously, an advertising strategy, focused at older individuals, could be focused on health effects, namely the large concentration of vitamin D, B12, iodine and selenium, which are important to prevent potential health problems at older age.

(Table 4 about here)

Table 5 shows the compensated substitution effect for the five fish categories, calculated by using equation (5), where the substitution effects are assumed to be symmetric. The results suggest that if the price of fresh salmon increases by one euro per kilo then the purchase frequency of frozen Salmonidae increases by 0.36, fresh cod increases by 0.33, frozen white fish increases by 0.18, and other fish increases by 3.8, over the whole period. Thus, reducing the price of frozen white fish by one euro per kilo reduces the purchase frequency of salmon the least, and increases purchase frequency of frozen white fish by 0.38%, as we can see from table 4 and 5. Moreover, the compensated substitution effects are largest for the group frozen white suggesting that frozen white fish is the best product group to put on sale to increase store traffic. To counter the loss associated with the price reduction of white fish, the price of fresh salmon that has the lowest own-price elasticity -0.06% could be increased.

(Table 5 about here)

Given our results, an example of a good marketing strategy would be to reduce the price of frozen white fish and advertise it as healthy for all age groups and especially elderly. Then, the seller should discretely increase the price of fresh salmon. This strategy could work particularly well on households who purchase fish infrequently and are less likely to be up to

[76] The age elasticity is interpreted as how frequencies change in response of a 1% change in age. Even though it is quite peculiar to talk about percentage changes in age, it is obvious that age is in itself continuous.

date on prices in different stores (Chen and Rey, 2012). A more precise development of these strategies will however be left to future research.

6. Summary and Conclusions

In this study, we derived consumer demand functions for the number of shopping trips by adding a budget constraint to the model of Meghir and Robin (1992). We then suggested a semi-logarithmic functional form for these demand functions, and estimated the system for five fish types. We used French scanner data of fish purchases for the years 2010-2013 and assumed a negative binomial DGP for the purchase frequencies.

Our empirical results show a significant heterogeneity between groups who purchase different types of fish, such as the average consumer of fresh salmon differs substantially from the consumer of frozen white fish. The typical consumer of fresh salmon is a healthy upper-class individual with university education who comes from a small household in Paris or the north of France, but the average frozen white fish consumer is an older, lower, middle to lower class individual who comes from a large household in the south France. From the calculated frequencies, we showed that most fish consumers purchase fish rarely, but a relatively small group of consumers purchase fish frequently. It is therefore even more important to locate which groups are infrequent purchasers and which groups are high frequency. Then fish retailers can use detailed description of consumers to know which groups to focus on in their marketing campaigns.

We see three areas of future research. Firstly, the demand for average quantities purchased could also be estimated. This type of extension could potentially provide a framework that allows for marketing strategies with more complex interactions between frequency of purchase and average purchases on each occasion. Secondly, the statistical model could be estimated with simulation based methods, which would allow us to incorporate an unrestricted covariance matrix between the demand equations. Thirdly, we have very many zero frequencies of purchase and the statistical model could be extended to a zero-inflated framework to account this specifically.

References

Allais, O., Bertail, P., and, Nichele, V. (2010). The Effect of a Fat Tax on French Households' Purchases: A Nutrition Approach. *American Journal of Agricultural Economics* 92, 228-245.

Asche, F., Dahl, R.E., Gordon, D.V., Trollvik, T., and. Aandahl, P. (2011). Demand Growth for Atlantic Salmon: The EU and French Markets. *Marine Resource Economics*, 26, 255-265.

Bhattacharya, C.B. (1997). Is Your Brand' s Loyalty too Much, too Little, or Just Right? Explaining Deviations in Loyalty from the Dirichlet Norm. *International Journal of Research in Marketing* 14, 421-435.

Bourre, J.M., and, Paquotte, P. (2008). Seafood (Wild and Farmed) For the Elderly: Contribution to the Dietary Intakes of Iodine, Selenium, DHA, and Vitamins B12 and D. *The Journal of Nutrition, Health and Aging* 12, 186 – 192.

Bertail, P., and, Caillavet, F. (2008). Fruit and Vegetable Consumption Patterns: A Segmentation Approach. *American Journal of Agricultural Economics* 90, 827-842.

Chen, Z., and Rey, P. (2012). Loss Leading as an Exploitative Practice. *The American Economic Review* 102, 3462-3482.

Chib, S. and Winkelmann, R. (2001). Markov Chain Monte Carlo Analysis of Correlated Count Data. Journal of Business and Economic Statistics 19, 428-435.

Creel, M.D., and Loomis, J.B. (1990). Theoretical and Empirical Advantages of Truncated Count Data Estimators for Analysis of Deer Hunting in California. *American Journal of Agricultural Economics* 72, 434–441.

Deaton, A. (1997). The Analysis of Household Surveys. Baltimore: Johns Hopkins University.

Deb, P., and Trivedi, P.K. (2002). The Structure of Demand for Health Care: Latent Class versus Two-Part Models. *Journal of Health Economics* 21, 601-625.

European Market Observatory for Fisheries and Aquaculture Products (EUMOFA). (2015). Retrieved from: https://www.eumofa.eu/documents/20178/66003/EN_The+EU+fish+market_Ed+2015.pdf/4cbd01f2-cd49-4bd1-adae-8dbb773d8519

Gobillon, L., and, Wolff, F.C. (2015). Evaluating the Law of One Price Using Micro Panel Data: The Case of The French Fish Market. *American Journal of Agricultural Economics* 98, 134-153.

Gustavsen, G.W. and K. Rickertsen (2003). Forecasting Ability of Theory-Constrained Two-Stage Demand Systems. *European Review of Agricultural Economics*, 30:539-558.

Hellerstein, D.M. (1991). Using Count Data Models in Travel Cost Analysis with Aggregate Data. *American Journal of Agricultural Economics* 73, 860–866.

Kau, A.K., and Ehrenberg, A.S.C. 1984. Patterns of Store Choice. *Journal of Marketing Research* 21, 399-409.

LaFrance, J.T., and Hanemann, W.M. (1989). The Dual Structure of Incomplete Demand Systems. *American Journal of Agricultural Economics* 71, 262-274.

LaFrance, J.T. (1990). Incomplete Demand Systems and Semilogarithmic Demand Models. *Australian Journal of Agricultural Economics* 34, 118-131.

Meghir, C., and Robin, J.M. (1992). Frequency of Purchase and the Estimation of Demand Systems. *Journal of Econometrics* 53, 53–85.

Munkin, M.K., and Trivedi, P.K. (1999). Simulated Maximum Likelihood Estimation of Multivariate Mixed-Poisson Regression Models, with Application. *Econometrics Journal* 2, 29-48.

The Norwegian Seafood Council (2016). Retrieved from: https://seafood.no/contentassets/6b8fa7b9742b4b0d9b2cc4a8d87af7a5/french-seafood-study.pdf

Onozaka, Y., Hansen, H., and, Sørvig, A. (2014). Consumer Product Perceptions and Salmon Consumption Frequency: The Role of Heterogeneity Based on Food Lifestyle Segments. *Marine Resource Economics* 29, 351-374.

Rickertsen, K., D. Kristofersson, and S. Lothe (2003). Effects of Health Information on Nordic Meat and Fish Demand. *Empirical Economics* 28, 249-273.

Robin, J.M. 1993. Analysis of the Short-Run Fluctuations of Households' Purchases. The *Review of Economic Studies* 60, 923-934.

Shonkwiler, J.S., and Englin, J. (2005). Welfare Losses Due to Livestock Grazing on Public Lands: A Count Data Systemwide Treatment. *American Journal of Agricultural Economics* 87, 302-313.

Uncles, M., Ehrenberg, A.S.C., and Hammond, K. (1995). Patterns of Buyer Behavior: Regularities, Models, and Extensions. *Marketing Science* 14, 71-78.

Uncles, M., and Lee, D. (2006). Brand Purchasing by Older Consumers: An Investigation Using the Juster Scale and the Dirichlet Model. *Marketing Letters* 17, 17-29.

Wang, P. (2003). A Bivariate Zero-Inflated Negative Binomial Regression Model for Count Data with Excess Zeros. *Economics Letters* 78, 373-378.

Xie, J., and Myrland, Ø. (2011). Consistent Aggregation in Fish Demand: A Study of French Salmon Demand. *Marine Resource Economics* 26, 267-280.

Table 1: Empirical Purchase Frequency

	Fresh Salmon		Frozen Salmonidae		Fresh Cod		Frozen White Fish		Other Fish	
	Freq.	%	Freq.	%	Freq.	%	Freq.	%	Freq.	%
0	70646	72.910	77669	80.158	83548	86.225	84461	87.168	31631	32.644
1	11206	11.565	10138	10.463	5632	5.812	8078	8.337	10956	11.307
2	5300	5.470	3778	3.899	2550	2.632	2295	2.369	8587	8.862
3	2971	3.066	1865	1.925	1552	1.602	943	0.973	6725	6.941
4	1945	2.007	1123	1.159	972	1.003	429	0.443	5600	5.779
5	1242	1.282	698	0.720	667	0.688	227	0.234	4685	4.835
6	835	0.862	419	0.432	428	0.442	152	0.157	3898	4.023
7	612	0.632	315	0.325	340	0.351	94	0.097	3283	3.388
8	502	0.518	209	0.216	226	0.233	73	0.075	2828	2.919
9	340	0.351	153	0.158	198	0.204	31	0.032	2346	2.421
10	232	0.239	120	0.124	145	0.150	33	0.034	2083	2.150
11	193	0.199	69	0.071	141	0.146	19	0.020	1775	1.832
12	147	0.152	62	0.064	100	0.103	14	0.014	1548	1.598
13	135	0.139	67	0.069	68	0.070	12	0.012	1323	1.365
14	89	0.092	45	0.046	61	0.063	8	0.008	1166	1.203
15	66	0.068	31	0.032	41	0.042	5	0.005	976	1.007
16	85	0.088	23	0.024	30	0.031	5	0.005	814	0.840
17	50	0.052	17	0.018	31	0.032	3	0.003	706	0.729
18	40	0.041	22	0.023	26	0.027	2	0.002	626	0.646
19	37	0.038	8	0.008	25	0.026	2	0.002	552	0.570
20	31	0.032	15	0.015	16	0.017	1	0.001	466	0.481
> 20	191	0.197	49	0.051	98	0.101	8	0.008	4321	4.459

Note: The first column shows frequencies. The second column shows the corresponding frequencies over the four years. Column three then denotes the percentage of observations with the observed frequencies.

Table 2: Descriptive Statistics

Variable	Mean	Std. Dev.	Min	Max
Fresh Salmon (Freq.)	0.836	2.371	0	87
Frozen Salmonidae (Freq.)	0.479	1.532	0	44
Fresh Cod (Freq.)	0.440	1.765	0	49
Frozen White Fish (Freq.)	0.230	0.870	0	35
Other Fish (Freq.)	4.998	7.823	0	132
Real Price of Fresh Salmon	0.161	0.039	0.017	0.716
Real Price of Frozen Salmonidae	0.157	0.036	0.001	0.698
Real Price of Fresh Cod	0.150	0.021	0.015	0.429
Real Price of Frozen White Fish	0.099	0.018	0.006	0.405
Real Price of Other Fish	0.108	0.053	0.008	0.918
Real Expenditures on Fish	0.641	0.861	0.003	16.576
Family Size	2.617	1.381	1	9
Age (Head of Household)	46.490	15.364	17	95
BMI (Head of Household)	24.873	4.920	11.019	59.285
Upper Class	0.138	0.345	0	1
Upper Middle Class	0.290	0.454	0	1
Lower Middle Class	0.411	0.492	0	1
Lower Class	0.161	0.367	0	1
Number of Cats	0.524	0.951	0	8
Number of Vehicles	1.491	0.809	0	8
Secondary Edu or Less	0.092	0.289	0	1
High School Level Education	0.515	0.500	0	1
University Education	0.393	0.488	0	1
South	0.202	0.402	0	1
North	0.099	0.299	0	1
East	0.091	0.288	0	1
West	0.196	0.397	0	1
Central	0.219	0.414	0	1
Paris Metropolitan Area	0.193	0.394	0	1
Housing Owner	0.590	0.492	0	1
Tenant	0.374	0.484	0	1
Free Accommodation	0.036	0.186	0	1

Table 3: Estimation Results from Negative Binomial Model

	Fresh Salmon		Frozen Salmonidae		Fresh Cod		Frozen White Fish		Other Fish	
	Est.	t-val.	Est.	t-val.	Est.	t-val.	Est.	t-val.	Est.	t-val.
Const.	-0.58	-9.59	-0.35	-4.84	-2.28	-24.32	-1.98	-21.57	1.32	56.51
Price	-0.37	-2.67	-2.44	-14.66	-0.66	-2.26	-3.86	-11.18	-3.90	-62.28
Exp.	0.91	242.79	0.91	242.79	0.91	242.79	0.91	242.79	0.91	242.79
Family Size · 10	-0.50	-7.14	0.03	0.30	-0.63	-6.28	0.83	8.24	0.29	10.79
Age · 1000	0.04	0.07	-6.04	-8.9	27.89	35.25	1.9	2.21	5.16	22.72
Upper Class	0.05	2.31	-0.02	-0.83	-0.05	-1.63	-0.22	-5.98	-0.05	-4.93
Lower M. Class	-0.12	-6.81	0.07	3.33	-0.18	-7.25	0.16	5.82	0.03	3.74
Lower Class	-0.31	-12.23	-0.04	-1.22	-0.43	-11.46	0.21	5.57	0.02	1.68
Nr. of Cats · 10	-0.27	-3.51	0.22	2.37	-0.66	-6.00	0.18	1.58	-0.02	-0.69
Nr. of Motor V.	0.03	2.82	-0.02	-1.82	0.07	4.79	-0.02	-0.99	-0.02	-5.01
D2011	0.06	2.83	-0.11	-4.45	-0.07	-2.51	0.05	1.7	-0.02	-2.08
D2012	0.23	11.78	-0.02	-0.98	-0.02	-0.71	0.11	3.78	-0.02	-1.92
D2013	0.13	6.47	0.01	0.21	0.10	3.49	-0.09	-2.89	-0.01	-1.03
Sec. Edu.	0.03	1.22	-0.15	-4.74	0.09	2.67	0.05	1.26	0.02	2.00
High. Edu.	0.14	9.10	0.06	3.27	0.06	2.72	-0.05	-2.15	-0.02	-3.43
House owner · 10	-0.78	-4.75	0.30	1.48	-1.07	-4.46	0.22	0.85	0.05	0.72
Accom. for Free	0.06	1.47	-0.03	-0.52	0.05	0.82	-0.16	-2.6	-0.01	-0.46
BMI · 10	-0.03	-2.34	-0.07	-4.08	-0.18	-8.37	0.04	1.83	-0.02	-3.04
South	-0.19	-9.04	0.07	2.84	-0.39	-12.48	0.18	5.76	-0.02	-2.17
North	0.12	4.55	0.05	1.6	0.06	1.67	-0.23	-5.52	-0.13	-11.66
East	-0.13	-4.62	0.08	2.38	0.10	2.69	-0.07	-1.68	-0.08	-7.61
West	0.04	1.92	-0.13	-4.62	0.16	5.33	-0.10	-3.07	-0.04	-4.07
Paris	0.13	6.01	0.17	6.01	-0.21	-6.57	-0.07	-1.93	-0.08	-8.86
Dispersion Par.	0.69	63.22	1.15	93.96	1.25	83.42	1.33	81.40	-1.08	-134.06

Table 4: Elasticities

	Fresh Salmon		Frozen Salmonidae		Fresh Cod		Frozen White Fish		Other Fish	
	Est.	t-val.	Est.	t-val.	Est.	t-val.	Est.	t-val.	Est.	t-val.
Price	-0.06	-2.67	-0.38	-14.66	-0.10	-2.26	-0.38	-11.18	-0.42	-62.28
Exp.	0.58	242.79	0.58	242.79	0.58	242.79	0.58	242.79	0.58	242.79
Age · 10	0.01	0.07	-2.81	-8.9	12.96	35.25	0.85	2.21	2.40	22.72
BMI	-0.08	-2.34	-0.18	-4.08	-0.44	-8.37	0.10	1.83	-0.04	-3.04

Table 5: Compensated Substitution Effects

	Estimate
Fresh Salmon VS Frozen Salmonidae	0.36
Fresh Salmon VS Fresh Code	0.33
Fresh Salmon VS Frozen White Fish	0.18
Fresh Salmon VS Other Fish	3.80
Frozen Salmonidae VS Fresh Cod	0.19
Frozen Salmonidae VS Frozen White Fish	0.10
Frozen Salmonidae VS Other Fish	2.18
Fresh Cod VS Frozen White Fish	0.09
Fresh Cod VS Other Fish	2.00
Frozen White Fish VS Other Fish	1.05

Note: All t-values are 242.79 and identical according to equation (5).

Paper 5

How often, how much?
Analysis of consumption of Label Rouge salmon in France

Arnar Buason[1], Audur Hermannsdottir[2], and Sveinn Agnarsson[2]

[1] School of Economics and Business, Norwegian University of Life Sciences, P.O. Box 5003, 1432 Ås, Norway. Corresponding author is Arnar Buason: arnar.buason@nmbu.no.
[2] Department of Economics, University of Iceland, Reykjavik, Iceland.

Abstract

The paper introduces a theory-consistent way of estimating a flexible infrequency of purchase model which accounts for the actual purchase frequency of consumers and not just whether a consumer buys a good or not. The model is applied to the consumption of salmon sold under the Label Rouge, other salmon and all other fish in the French fish market. Using French household scanner data, it is shown that consumers´ perception and loyalty differs substantially between fresh salmon bearing the Label Rouge label and non-labelled salmon, demonstrating that the Label Rouge is able to produce the desired effect of product differentiation.
Keywords: France, consumption, salmon, frequency, Label Rouge

1. Introduction

France is one of the most important seafood markets in Europe. In 2015, the country ranked third in the EU, with total household expenditure for farmed and wild seafood totalling around € 8.5 billion (EUMOFA, 2017). Fresh seafood made up around a quarter of that value, with salmon and cod the most important species. Most of the salmon is farmed while most of cod is wild. As noted by Chen, Alfnes and Rickertsen (2015), no ecolabelling program for farmed fish has so far gained wide international acceptance, but in France organic labels such as Agriculture Biologique are widely used for food. The quality label Label Rouge is, however, well known in France since its introduction in 1965 and open to all food products, including seafood. At first though it was only applied to agricultural products, with Scottish salmon becoming the first fish – and indeed non-French – product to be awarded the label. Today, Norwegian and Irish salmon is also sold under this label.

In this paper we employ count data methods to analyse what determines how often French consumers buy fresh Label Rouge salmon as well as salmon not bearing this label, and then use that information to adjust the demand system for the frequency of consumer purchases and not just whether a purchase takes place or not. The classical econometric models by Tobin (1958), Cragg (1971), Heckman (1974), Lee and Pitt (1987), Wales and Woodland (1983), and Phaneuf et al. (2000) assume that zero purchases are due to corner solutions from utility maximizing behaviour. However, this is not the only reason for observed zero purchases in micro data, and infrequency of purchase models (IPM's) assume that zero observations can also be generated by short term purchase fluctuations (i.e. infrequency of purchases) (Deaton and Irish, 1984; Kay Keen and Morris, 1984; Pudney, 1985; and Pudney, 1986). Blundell and Meghir (1987) introduce a two regime IPM which assumes that all observed zeros in a data set are due to infrequency of purchases, i.e., there are no households which are non-consumers. Such a specification is reasonable in many cases including purchases of clothes, cars, housing or other durable goods, or when the data set contain a subset of households which all consume a specific product group. Blundell and Meghir (1987) also present a double hurdle model, based on Cragg (1971), which nests both the Tobit and the IPM and is more general than its two alternatives. A generalization of the IPM framework is introduced by Meghir and Robin (1992) and Robin (1993), where the models are extended to account for the actual purchase frequency of households over a given survey period, and not just whether a purchase took place or not, as in Blundell and Meghir (1987). This approach therefore uses more of the information in the data set and is thus able to produce more precise estimates than conventional models, as is shown in Robin (1993).

This paper extends the literature in three ways. First, by introducing a theory consistent way of estimating the flexible infrequency of purchase model introduced by Robin (1993). Previously Robin (1993) had log-linearized the model and estimated it without any theoretical constraints. We however estimate the model in two stages. (i) We estimate a count data model for frequencies of purchase, (ii) we estimate the demand system, adjusted with the probabilities calculated in the first step, as linear almost ideal demand system (LAIDS) introducing additivity, homogeneity, and symmetry. Second, by applying the model to the demand for fresh salmon in France, using scanner data on 20,000 households during the years 2011-2013. The method used makes it possible to take account of zero observations in a sophisticated manner by accounting for the actual frequency of purchases and not just whether a consumer buys the product or not. This is a significant problem when analyzing household purchases of a product like Label Rouge salmon, as the product group is such a small fraction of all seafood products purchases by consumers that its fraction of the market is next to none.

2. Labelling and Label Rouge

Consumers are slowly but steadily starting to put more emphasis on quality products and are increasingly more interested in knowing where products come from and how they are produced (Dimara and Skuras, 2005; Grunert, 2005; Whitmarsh and Palmieri, 2011). This might, among other things, be the result of increased globalization of food trade that has created concerns among consumers regarding the quality of the products they purchase (Mariojouls and Wessells, 2002). The emphasis on quality varies between countries, but in France the demand for high quality has been increasing and has become more solid then ever (Monfort, 2006).

When forming attitudes and food quality expectations, consumers look for quality cues (Brunsø, Verbeke, Olsen and Jebbesen, 2009). Quality cues in form of informational labelling can reduce consumers' risk of buying food that might not satisfy their needs or negatively affect them in some way (Dufeu, Ferrandi, Gabriel, and Gall-Ely, 2014). Quality labelling can therefore be a powerful signal and assist consumers in their purchase decisions (Dimara and Skuras, 2005; Hocquette et al., 2013). Such labels provide information about attributes that some segments of consumers find important; for example that the product is safe and that it is produced in an environmentally friendly way through a socially acceptable process (Monfort, 2006). Nevertheless, the efficiency of labelling as an informational source has been questioned by some, who argue that the efficiency of labels as assisting in purchase decisions depends on how important consumers deem the labelled information (Dimara and Skuras, 2005). Labels can also cause confusion among consumers since there are many available signs regarding quality and origin, governmentally certified or not (Mariojouls and Wessells, 2002).

Even though consumers' concerns for quality may drive quality labelling, they are also a means to address competition by differentiating similar products by production process (Mariojouls and Wessells, 2002). Within food production there is an increased awareness that competing on price alone is not necessarily the best strategy (Grunert, 2005). Many firms are becoming more customer-focused and emphasize added value for customers (Grunert, 2005) who generally speaking have a favourable perception of products marked with official quality labels (Hocquette et al., 2012)

For a long time, France has been a world leader in food labelling programs (Mariojouls and Wessells, 2002). One of their well-known food labelling trademarks is Label Rouge, set up in 1960 by the Ministry of Agriculture and is managed by a national commission for labels and certifications. Label Rouge is a signal of superior quality, representing intrinsic quality, food safety and environmentally sound production practices (Dufeu et al., 2014; Westgren, 1999). The label was designed to differentiate high quality food products from standard products (Monfort, 2006), by certifying that the products marked with Label Rouge possess specific characteristics that make them of higher quality than standard products (Hocquette et al., 2012). The quality label was developed as a tool for improving the price mechanisms for products from capture fisheries (Charles and Boude, 2001) and this has indeed been the case; products marketed with the Label Rouge label are sold at huge premium (Mariojouls and Wessells, 2002; Westgren, 1999).

Possible benefits of using Label Rouge are not just financial. Using the label conveys a positive image and the demand for products marked with the Label Rouge is more stable so the customers seem to be loyal to the label (Monfort, 2006). According to Monfort (2006) the label fosters sales, especially in specific upper grade niches. Among French consumers, Label Rouge is the most widely recognized product quality predictor (Hocquette et al., 2013) with a good reputation (Dufeu et al., 2014; Monfort, 2006), and is thought of as being highly trustworthy (Dufeu et al., 2014; Hocquette et al., 2013).

3. Theoretical and statistical model

Consider a consumer who faces the following optimization problem:

$$\max_{l,c,n}\{U(l,c,n): wT + R = p'c + wl + wL(n), l > 0, c > 0, n > 0\} \tag{1}$$

where U denotes utility, l leisure, $c = (c_1, c_2, \dots, c_M)'$ is a vector of consumption goods and $n = (n_1, n_2, \dots, n_M)'$ represents the corresponding purchase frequency. The consumers budget constraint is $y + R = p'c$, where $p = (p_1, p_2, \dots, p_M)$. The consumer has wage income $y = hw$, where h is the hours spent working and w is the hourly wage rate, and other income through transfers and undeclared activities, R. Total available time is T, which is split into total hours worked h, time spent purchasing goods which is given by the function $L(n)$, which is increasing in n, and other non-market hours, l. The time constraint is then $T = l + k + L(n)$. The solution to the optimization problem are three sets of Marshallian demand equations, $l(p, w, R)$, $n_i(p, w, R)$, and $c_i(p, w, R)$, where $i = 1,2, \dots, M$.

The consumer faces three decisions; how much time to spend on non-market activities, how much to purchase and how often to purchase goods. This paper deals with the purchase frequency decision and how much a consumer purchases of goods, which is an approximation for consumption demanded.

It is assumed that the number of shopping trips, n, is generated by a discrete distribution with a probability mass function $f_N(n|Z_i)$ for $n = 0,1,2, ...$ where z_i is a matrix of exogenous variables. The analysis builds on the flexible infrequency of purchase model introduced by Robin (1993), where the model accounts for the actual purchase frequency and not just whether a purchase takes place or not. This type of model therefore utilizes more of the information in the data than a standard hurdle model would, and takes account of zeros in the data set as conventional hurdle models do. The model can be represented as follows:

$$\begin{cases} if\ n_j = 0: w_j = 0 \\ if\ n_j > 0: b_j P(n_j > 0|z_j) = f(z_j'\gamma)\left[\frac{n_j P(z_j'\beta)}{z_j'\beta}\right]^{\alpha} + u_j \end{cases} \tag{2}$$

where n_j is frequency of purchase of good j, b_j are budget shares, z_j are exogenous variables, $P(\cdot)$ is the probability of a purchase, and β, γ, and α are parameters to be estimated. The parameter α is a key parameter which determines the relationship between purchase frequency and budget shares. If $\alpha = 0$ the model reduces to a standard type hurdle model which only accounts for whether a purchase takes place or not. If on the other hand $\alpha > 0$, the model accounts for the actual purchase frequency. To keep things simple, but still allowing for $\alpha > 0$, it is assumed that that this relationship is given by $\alpha = 1$. Thus, if α is significant we show that it is actually important to account for frequencies when adjusting a demand system with probabilities of purchase. However, since we restrict $\alpha = 1$ it will not be possible to see exactly what the relationship is. Allowing α to take any value will be left for future research.

The data generating process of n_j is assumed to follow the truncated negative binomial distribution. The truncation makes it possible to account for the large share of zeros in the data set, and the negative binomial provides a more flexible variance specification, namely the possibility of over- and underdispersion.

The utility function given by (1) is assumed to be weakly separable, i.e. the consumer maximizes her utility with respect to different food groups separately both in terms of quantities and frequencies, subject to the budget share allocated to the consumption of each group. This makes it possible to estimate a complete demand system in terms of fish, conditional on the aforementioned assumption of weak separablity. The demand system consists of demand equations for three different types of fish products; fresh salmon, Label Rouge salmon which is also fresh, and all other fish products.

The model in equation (2) is estimated in two stages. The first part of the model consists of the purchase frequencies, which are used to adjust the demand system for how often a consumer purchases. As is conventional when estimating count data models, the conditional expectation is defined as a semi-logarithmic function, see for example Shonkwiler and Englin (2005). These demand functions, which relate to purchases frequencies rather than quantity demanded, are thus given by the following expression:

$$\mathrm{E}\left(n_{ij}\middle|Z_i\right) = \exp\left(\beta_{ij} + \sum_{s=1}^{M} \beta_{is}\left(p_{sj}/CPI\right) + \theta_i\left(y_j/CPI\right)\right) \tag{3}$$

where $i = 1,2,3$ denotes the three different fish categories and $j = 1,2,\ldots,J$ households. The price of fish category *i* for household *j* is denoted by p_{ij}. The total fish expenditure of household *j* is given by y_j, and *CPI* is the average French consumer price index over the three years. In order for the demand system to be consistent with economic theory a number of restrictions must be imposed on the parameters. Following LaFrance and Hanemann (1989) and LaFrance (1990) the symmetry restrictions may be derived as $\beta_{is} = 0\ \forall\ i \neq s$ and $\theta_i = \theta\ \forall\ i$.

Homogeneity of degree zero is then imposed on the Marshallian demand system by dividing both prices and expenditure by the French consumer price index (*CPI*). Household heterogeneity may then be model by defining the constant term in (1) as:

$$\beta_{ij} = \eta_{0i} + \sum_{k=1}^{K} z_{kj}\eta_{ki} \tag{4}$$

where z_{kj} represents household characteristics and η_{ki} are parameters to be estimated. The frequency system described in equations (3) and (4) are then used to predict the probability of observing a positive frequency $P(n_j > 0|z_j)$ in equation (2). Furthermore, $\left[\frac{n_j P(z_j'\beta)}{z_j'\beta}\right]^{\alpha}$ is also predicted using the same results. Finally, it should be noted that even though cross-price effects are restricted to be zero, i.e., $\beta_{is} = 0$, compensated cross-price effects are not. The Slutsky equation yields:

$$e_{isj} = n_{ij}\frac{\partial n_{sj}}{\partial y_j} = \theta n_{ij} n_{sj} \tag{5}$$

where e_{isj} is the compensated substitution effect between products *i* and *s* for household *j*, and the *n*'s are purchase frequencies of different product groups by household *j*.

The second part of the model specifies $f(z_j'\gamma)$ in equation (2). Here, this function consists of a system of demand equations given by the linear approximated almost ideal demand system (LA/AIDS)[77]:

$$b_{ij} = \alpha_{ij} + \sum_{s=1}^{M} \gamma_{is} \ln p_{sj} + \varphi_i(\ln y_j - \ln P_j) \tag{6}$$

[77] See Deaton and Muelbauer for a representation and discussion of the AIDS/LAIDS model.

where b_{ij} is the budget share of household j for good i, p_{sj} is the unit value price of household j for good s, y_j is total expenditure of fish for household j, and finally P_j is a price index, approximated by Stone's price index, $\ln P_j = \sum_s^M b_{sj} \ln p_{sj}$. Household heterogeneity is specified by specifying the constant term in equation (6) as:

$$\alpha_{ij} = \tau_{0i} + \sum_{k=1}^{K} z_{kj}\tau_{ki} \tag{7}$$

where z_{kj} represents household characteristics and τ_{ki} are parameters to be estimated.

To be in accordance with economic theory, the parameters of the demand equations must satisfy the following restriction:

Adding up: $\sum_i \tau_{ki} = 1, \sum_i \varphi_i = 0$, and $\sum_i \gamma_{is} = 0$

Homogeneity: $\sum_s \gamma_{is} = 0$ and

Symmetry: $\gamma_{is} = \gamma_{si}, i \neq s$

4. Data

The data used in this paper comes from Kantar Worldpanel and consist of scanner data of weekly purchases of fish by 20,000 French households during the three-year period 2011-2013. However, there is rotation of households, since a share of households drop out each year. In all there are 43,127 households included in the study. As indicated by the data type, the data is collected using hand held bar code scanners which Kantar provides to all households, along with other relevant equipment for the data collection. This method makes the data extremely detailed since the bar code contains significant fraction of the important information regarding the product. As well as observations on purchases, the data set also contains detailed information on sociodemographic- and geographical household characteristics. The households are selected by stratification according to a few socioeconomic variables. As is common with micro data, the fraction of zero observations is significant, as can be seen from Table 1 which shows that 73% of the observations on purchases of fresh salmon are zero, i.e., no salmon was purchased. The corresponding numbers are 98% for Label Rouge salmon and 30% for other fish products. These numbers demonstrate the importance of the econometric methods used in the paper.

(Table 1 about here)

Descriptive statistics of all the variables used in the analysis, except for annual dummies, are given in Table 2. The mean frequency of fresh salmon purchases is 0.81 indicating that on average consumers bought fresh salmon almost once during the time period under observation, 2011-2013. By contrast, the mean frequency of fresh Label Rouge purchases is only 0.03, implying that consumers almost never bought fresh Label Rouge salmon. Other fish products are purchased more frequently, or six times on average. The difference between the purchase frequencies is therefore quite large, which is not surprising, given that the category Label Rouge salmon refers to a very narrowly defined product while the category other products refers to a wide range of different seafood products.

The price variables in Table 2 refer to log standardized prices, but it should be noted that these prices are unit values and not real prices. The prices are calculated from the amount of fish purchased in kilos and the expenditures in euros. This is a common practice, see for example Allais et al. (2010) and Bertail and Caillavet (2008)[78]. Most of the social and geographical variables used in this study have been used in other studies (see for instance Allais et al. (2010)), except for the body mass index (BMI) variable which is here used as an indicator of good health. Individuals with a BMI in the range 18.5-24.9 may be considered as healthy, while

[78] These constructed unit values might be endogenous due to the choice of product quality by the consumer.

those with a lower BMI are regarded as underweight and those with a higher BMI as either overweight or obese (BMI > 30) (Eurostat, 2017).

(Table 2 about here)

5. Results and discussion

The model described in (2) was estimated in two stages; first the frequency system outlined in equations (3) and (4) and then the almost ideal demand system set out in equations (6) and (7), after adjusting for the actual purchase frequencies described in equation (2). For each equation, the frequencies were assumed to be a function of price, total expenditure, family size, number of children under the age of 16, age, education, class status[79], education, geographical location, body mass index (BMI), type of accommodation and dummy variables for the years 2012 and 2013. The model includes three education variables; secondary education or less, high school, and university, as well as four class variables; lower class, lower and upper middle class, and upper class. There are six geographical areas; north, south, east, west, central and Paris, and the model also takes into consideration whether the consumer is a tenant, does not pay rent or owns his own house or flat. The base version of the model, i.e. when all dummy variables except year dummies take a value of zero, assumes a consumer who is of upper middle class standing, has completed university and lives in her own house or flat in central France.

The estimation results are presented in Table 3. The family size parameter is significant and negative in both salmon frequency equations indicating that individuals from smaller families purchase fresh salmon more often. The education parameter estimates reveal that there are two types of French fresh salmon consumers, those with secondary education or less are keener on regular fresh salmon, while those with university education prefer Label Rouge salmon. Salmon consumers are mostly from the upper middle class, but purchases are not limited to a

[79] The class is measured by work status, such as unemployed, routine occupation, manager etc. See Rose and Harrison (2010) for a detailed discussion.

certain age group. The number of children under 16 has no effect on the demand for non-Label Rouge fresh salmon, but is almost significant at the 5% level for the Label Rouge salmon. Those consuming regular salmon live in housing where they pay no rent, but the Label Rouge salmon consumer are better off and own their own house or flat. The BMI parameter is significant and negative for those buying Label Rouge salmon. Individuals that most frequently purchase Label Rouge salmon are healthy university-educated upper middle class citizen who live in their own house or flat in the center of France, but not Paris.

The frequency of purchase of fresh salmon increased from 2011 to 2012, which is consistent with the growing salmon market in France and the marketing of farmed salmon from Norway. The sign of the 2013 year dummy is negative but not significant. Both time dummies are insignificant in the Label Rouge equation.

The average buyer of other fish is a healthy older individual with secondary education who comes from a large family in the centre of France.

(Table 3 about here)

Table 4 shows elasticities from the truncated negative binomial model. The own-price elasticity of fresh salmon and other fish is the same, thus a 1% increase in price will reduce the corresponding purchase frequency by 0.054%, which is quite low. It should be noted however that these averages are conditioned on frequencies being positive. The own-price elasticity of Label Rouge salmon is even lower than the two other categories, when the price of Label Rouge salmon increases by 1% the purchase frequency decreases by 0.001%. That is changes in price of label Rouge has little effect on frequencies. These differences in elasticities between fresh salmon and Label Rouge labelled salmon open up some interesting opportunities for pricing strategies. One strategy could, for instance, be to marginally decrease the price of fresh salmon and instead raise the price of Label Rouge salmon, as this should increase purchases of salmon while having very limited effect on the frequency of Label Rouge salmon purchases, thus resulting in increases in overall store traffic.

The BMI elasticities for Label Rouge and other fish are significant and negative, where the effect on LR purchases are quite small. The effect on other fish is much greater; a 1% increase in BMI is associated with a 2.05% reduction in other fish purchases.

(Table 4 about here)

The compensated substitution effects from the truncated negative binomial model are shown in Table 5. Increasing the price of fresh non-labelled salmon by one euro per kilo will only increase purchase frequency of Label Rouge salmon by 0.01, which is a very small substitution effect. By contrast the purchase frequency of other fish will increase by 1.71 more trips if the price of salmon goes up by one euro per kg. The substitution effects from an increase in the price of Label Rouge salmon are close to zero. Thus, a marginal increase in the price of Label Rouge salmon has very small effect on the purchase frequency of both other salmon and other seafood products. These results yields a picture of a very loyal Label Rouge salmon consumer who does neither reduce her consumption of Label Rouge salmon when the prices of that product goes up, as is evident from Table 4, nor switch to other seafood products. Increases in the price of other fish have very different effects on purchases of fresh salmon and Label Rouge salmon. The compensated substitution effect is quite strong in the former case, as the purchase frequency of fresh salmon rises by 1.66 when the price of other fish is increased by one euro per. kg., but the price rise has hardly any effect on how often consumers buy Label Rouge salmon.

(Table 5 about here)

The estimation results from the purchase frequency adjusted LA/AIDS model are presented in Table 6. The average consumer of fresh salmon is a younger upper class individual with higher education, who comes from a small family in Paris or the North of France. Moreover, this type of consumer not only purchases fresh salmon in more quantity than other consumers but also more frequently. This group of customers is thus the perfect group to target in order to increase sales of fresh salmon.

The average consumer of Label Rouge salmon is a healthy older upper middle class individual from Paris. Thus, this group of consumers is similar to those that buy fresh salmon, but there is nevertheless a distinct difference between the two groups of consumers; the quality label attracts older consumers whereas younger individuals favour cheaper fresh salmon without any quality label.

(Table 6 about here)

The calculated elasticities for fresh salmon, Label Rouge salmon and other fish are shown in Table 7 where the elasticities for other fish are calculated using the adding up condition. The price elasticities show that the three groups are substitutes. The expenditure elasticities of all groups are close to 1, thus a 1% increase in fish expenditure is associated with a 1% increase in the budget share of each of the three groups.

(Table 7 about here)

Finally, Table 8 shows the results from t-tests of differences between elasticities from the LA/AIDS model. That is the null hypothesis is no difference between parameters. The results reveal that the own-price elasticity of fresh salmon is significantly more elastic than for Label Rouge salmon. Older consumers who buy Label Rouge salmon are thus statistically more loyal, in the sense of how sensitive they are to price changes, than younger individuals who buy regular fresh salmon. Furthermore, the cross-price elasticities between fresh salmon and Label Rouge are significantly different. Therefore, the effect of an increase in salmon price on demand for Label Rouge salmon is larger than the increase in the price of Label Rouge salmon on demand for fresh salmon. The own-price elasticity of fresh salmon is larger than that of other fish. Thus, fresh salmon price is more elastic than fish products on average.

(Table 8 about here)

The Label Rouge label has a significant effect on consumers' perception of fresh salmon. Firstly, while younger adults purchase fresh salmon, older individuals prefer Label Rouge salmon. Moreover, the average young fresh salmon consumer is an upper-class citizen from Paris or the north of France whereas the average Label Rouge salmon consumer is an upper middle class individual from Paris. Thus, the two groups are distinctly different and marketing strategies should emphasize that these are two as separate consumer groups. Secondly, the own-price elasticity of fresh salmon is greater than that of Label Rouge salmon, indicating that older, upper middle class consumers are more loyal towards Label Rouge salmon than younger, upper class individuals are towards non-labelled fresh salmon. The label thus makes it possible for consumers to significantly differentiate between fresh salmon and fresh Label Rouge salmon. Label Rouge does not have such a strong foothold among younger consumers and thus their willingness to pay for products bearing this label is lower than that of the older consumers.

6. Conclusions

The article introduces a theory consistent almost ideal demand representation of the model due to Robin (1993). The model is a two-step procedure. In the first step, a truncated negative binomial model is estimated and the predictions and probabilities calculated in that step are then used to adjust the almost ideal demand system, which is estimated in the second step. This procedure makes it possible to account for actual purchase frequencies and not just the probability of positive purchase. Thus, we utilize more of the information in the data than conventional hurdle models and thus produce more accurate estimates, as shown by Robin (1993).

The model is applied to French scanner data of fish purchases from 2011-2013 to analyse the demand for Label Rouge salmon in the French consumer market. The results show that average consumers of fresh salmon and fresh salmon bearing the Label Rouge are significantly different. Furthermore, consumers who purchase Label Rouge salmon are more loyal than those who buy non-label fresh salmon. The label is thus able to reach its goal of differentiating between fresh salmon and fresh Label Rouge salmon, both in terms of consumer perception and loyalty. This study therefore confirms the results obtained by Monfort (2006), that Label Rouge customers are loyal to the label and that the label is well suited to foster sales among high-end consumers.

References

Allais, O., P. Bertail, and Nichele, V. (2010). The effects of a fat tax on French households' purchases: A nutritional approach. *American Journal of Agricultural Economics* 92, 228–245.

Bertail, P., and, Caillavet, F. (2008). Fruit and Vegetable Consumption Patterns: A Segmentation Approach. *American Journal of Agricultural Economics* 90, 827-842.

Blundell, R. W. and Meghir, C. (1987). Bivariate Alternatives to the Tobit Model. *Journal of Econometrics* 34, 179-200.

Brunsø, K., Verbeke, W., Olsen, S. O., and Jeppesen, L. F. (2009). Motives, barriers and quality evaluation in fish consumption situations: Exploring and comparing heavy and light users in Spain and Belgium. *British Food Journal* 111, 699-716.

Charles E., and Boude, J.P. (2001). *Enhancement strategy, artisanal fishing products and the theory of conventions*. XIII EAFE Conference. Available from: http://www.eafe-fish.org/conferences/salerno/papers/paper14_erwancharles.doc.

Chen, X., Alfnes, F., and Rickertsen, K. (2015). Consumer preferences, ecolabels, and effects of negative environmental information. *AgBioForum*, 18, 327-226.

Cragg, (1971). Some Statistical Models for Limited Dependent Variables with Application to Durable Goods. *Econometrica*, 39, 829-844.

Deaton, A. S. and Irish, M. (1984). A Statistical Model for Zero Expenditure in Household Budget. *Journal of Public Economics*, 23 59-80.

Deaton, A., and Muellbauer, J., (1980). An Almost Ideal Demand System. *The American Economic Review* 70, 312-326.

Dimara, E. and Skuras, D. (2005). Consumer demand for informative labelling of quality food and drink products: A European Union case study. *Journal of Consumer Marketing* 22, 90-100.

Dufeu, I., Ferrandi, J. M., Gabriel, P., and Le Gall-Ely, M. (2014). Socio-environmental multi-labelling and consumer willingness to pay. *Recherche et Applications en Marketing (English Edition) 29,* 35-56.

European Market Observatory for Fisheries and Aquaculture Products (EUMOFA). (2017). Retrieved from http://www.eumofa.eu/predefined-queries2.

EUROSTAT. (2017). Statistics Explained. Overweight and obesity - BMI statistics. Retrieved from http://ec.europa.eu/eurostat/statistics-explained/index.php/Overweight_and_obesity_-_BMI_statistics.

Gruner, K. G. (2005). Food quality and safety: consumer perception and demand. *European Review of Agricultural Economics*, 2005 32, 369-391

Heckman, J. (1974). Shadow Prices, Market Wages, and Labor Supply. *Econometrica* 42, 679-694.

Hocquette, J.-F., Botreau, R., Picard, B., Jacquet, A., Pethick, D. W., and Scollan, N. D. (2012). Opportunities for predicting and manipulating beef quality. *Meat Science 92*, 197-209.

Hocquette, J. F., Jacquet, A., Giraud, G., Legrand, I., Sans, P., Mainsant, P., and Verbeke, W. (2013). Quality of food products and consumer attitudes in France. In *Consumer attitudes to food quality products* (pp. 67-82). Wageningen Academic Publishers.

Kay, J. A., Keen, M. J. and Morris, C. N. (1984). Estimating Consumption from Expenditure Data. *Journal of Public Economics* 23, 169-181.

Lee, L. F. and Pitt, M. M. (1986). Microeconomic Demand Systems with Binding Nonnegativity Constraints: The Dual Approach. *Econometrica* 54, 1237-1242.

Mariojouls, C. and Rohem-Wessells, C. (2002). Certification and quality signals in the aquaculture sector in France. *Marine Resource Economics* 17, 175-180.

Meghir, C., and Robin, J.M. (1992). Frequency of Purchase and the Estimation of Demand Systems. *Journal of Econometrics* 53, 53–85.

Monfort, C.M. (2006). *Adding value to salmon helps capturing market shares*. IIFET 2006 Portsmouth Proceedings. Corvallis, Oregon: International Institute of Fisheries Economics & Trade.

Phaneuf, D.J. Kling, C.L, and, Herriges, J.A, (2000). Estimation and Welfare Calculations in a Generalized Corner Solution Model with an Application to Recreation Demand. *The Review of Economics and Statistics* 82, 83-92.

Pudney, S. (1985). *Frequency of Purchase and Engel Curve Estimation*. Discussion Paper A56, LSE Econometrics Programme. London: LSE.

Pudney, S. (1986) Modelling Individual 5 Choice: The Econometrics of Corners, Kinks and Holes (Oxford: Basil Blackwell).

Robin, J.M. (1993). Analysis of the Short-Run Fluctuations of Households' Purchases. *The Review of Economic Studies* 60, 923-934.

Rose, D. and Harrison, E. (2010). Social Class in Europe An Introduction to the European Socio-Economic Classification. Abington: Routledge.

Tobin, J. (1958). Estimation of Relationships for Limited Dependent Variables. *Econometrica* 26, 24-36.

Wales, T. J. and Woodland, A. D. (1983). Estimation of Consumer Demand Systems with Binding Nonnegativity Constraints. *Journal of Econometrics* 21, 437-468.

Shonkwiler, J.S., and Englin, J. (2005). Welfare Losses Due to Livestock Grazing on Public Lands: A Count Data Systemwide Treatment. *American Journal of Agricultural Economics* 87, 302-313.

Westgren, R E. (1999). Delivering food safety, food quality, and sustainable production practices: The Label Rouge Poultry System in France. *American Journal of Agricultural Economics* 81, 1107-1111.

Whitmarsh, D. and Palmieri, M. G. (2011). Consumer behaviour and environmental preferences: A case study of Scottish salmon aquaculture. *Aquaculture Research*, 42, 142-147.

Table 1: Empirical purchase frequency

	Fresh salmon		Label Rouge salmon		Other fish	
	Freq.	%	Freq.	%	Freq.	%
0	53682	73.036	72212	98.246	22570	30.707
1	8536	11.613	863	1.174	7401	10.069
2	4007	5.452	196	0.267	5905	8.034
3	2301	3.131	91	0.124	4629	6.298
4	1452	1.975	36	0.049	4001	5.443
5	920	1.252	27	0.037	3476	4.729
6	638	0.868	19	0.026	2958	4.024
7	465	0.633	11	0.015	2674	3.638
8	372	0.506	7	0.010	2328	3.167
9	230	0.313	7	0.010	2031	2.763
10	169	0.230	9	0.012	1705	2.320
> 10	729	0.992	23	0.031	13823	18.807

Table 2: Descriptive statistics

Variable	Mean	Std. dev.	Min	Max
Fresh salmon frequency	0.81	2.27	0	56
Label Rouge salmon frequency	0.03	0.42	0	32
Other fish frequency	6.04	9.19	0	162
Fresh salmon budged share	0.13	0.23	0	1
Label Rouge budget share·10	0.03	0.03	0	1
Other fish budget share	0.87	0.24	0	1
Log standardized price of fresh salmon	-0.03	0.24	-2.24	2.21
Log standardized price of Label Rouge salmon·10	-0.04	0.09	-1.96	1.14
Log standardized price other fish	-0.11	0.47	-2.65	2.71
Log expenditures / Stone Price Index	3.56	1.16	-0.92	7.76
Family size	2.62	1.39	1	9
Number of children under 16	1.68	1.00	1	9
Age of head of household	46.69	15.22	17	94
BMI of head of household	24.91	4.93	11.02	59.29
Upper class	0.14	0.35	0	1
Upper middle class	0.28	0.45	0	1
Lower middle class	0.41	0.49	0	1
Lower class	0.16	0.37	0	1
Secondary education or less	0.09	0.29	0	1
High school education	0.52	0.50	0	1

University education	0.39	0.49	0	1
South	0.20	0.40	0	1
Center	0.22	0.41	0	1
North	0.10	0.30	0	1
East	0.09	0.29	0	1
West	0.20	0.40	0	1
Paris metropolitan area	0.19	0.40	0	1
Housing owner	0.59	0.49	0	1
Tenant	0.37	0.48	0	1
Free accommodation	0.04	0.19	0	1

Table 3: Estimation results from the truncated negative binomial model

	Fresh salmon		Label Rouge salmon		Other fish	
	Est.	t-val.	Est.	t-val.	Est.	t-val.
Constant	-0.19	-1.95	-2.77	-3.36	1.73	71.44
Price	-0.73	-3.71	-0.08	-0.08	-0.49	-12.27
Expenditure	0.75	166.58	0.75	166.58	0.75	166.58
Family size·10	-0.63	-4.79	-3.43	-4.23	0.07	2.7
Nr. of children under 16	-0.01	-0.75	0.22	1.78	0.01	3.17
Age·10	-0.01	-1.48	-0.05	-0.82	0.02	7.04
Upper class	0.02	0.58	-0.72	-3.88	-0.04	-6.03
Lower middle class	-0.04	-1.52	-0.18	-1.06	0.03	5.53
Lower class	-0.07	-1.74	-0.45	-1.77	0.04	4.98
D2012·10	1.17	4.36	2.64	1.54	0.02	0.43
D2013·10	-0.02	-0.07	2.51	1.52	-0.08	-1.65
Sec. edu.	0.16	4.05	0.19	0.82	0.02	2.3
High. edu.	0.14	5.77	0.14	0.9	-0.04	-8.37
Tenant·10	-0.04	-0.13	-2.55	-1.49	0.20	4.21
Accom. for free·10	1.25	2.07	0.05	0.01	0.18	1.51
BMI·10	-0.04	-1.54	-0.38	-2.47	-0.85	-19.81
South	-0.20	-5.73	0.13	0.59	-1.32	-16.72

North	0.02	0.48	-0.95	-3.12	-1.78	-21.3
East	-0.09	-2.12	-0.59	-1.95	-1.70	-20.34
West	0.04	1.24	-0.02	-0.12	-1.47	-19.34
Paris	0.06	1.71	0.12	0.62	-1.30	-16.14
Dispersion par.	0.76	17.06	3.72	6.86	-1.21	-108.86

Table 4: Selected elasticities from the truncated negative binomial model

	Fresh salmon		Label Rouge salmon		Other fish	
	Est.	t-val.	Est.	t-val.	Est.	t-val.
Price·10	-0.54	-3.71	-0.01	-0.08	-0.54	-12.27
Expenditure	0.22	166.58	0.05	166.58	0.47	166.58
Age	-0.03	-1.48	-0.02	-0.82	0.09	7.04
BMI	-0.04	-1.54	-0.09	-2.47	-2.05	-19.81

Table 5: Compensated substitution effects from the truncated negative binomial model

	Estimate
Fresh salmon vs. Label Rouge·10	0.10
Fresh salmon vs. other	1.71
Rouge vs. fresh salmon·100	0.09
Rouge vs. other	0.02
Other vs. fresh salmon	1.66
Other vs. Label Rouge	0.01

Note: The t-value for all estimates are 166.58 due to equation (5).

Table 6: Estimation results from the frequency adjusted AIDS model

	Fresh salmon		Label Rouge salmon	
	Est.	t-val.	Est.	t-val.
Constant·10	2.29	19.72	-0.04	-4.81
Fresh salmon price·10	-0.45	-14.64	0.03	6.41
Label Rouge price·10	0.03	6.41	-0.05	-11.02
Other fish price·10	0.43	13.89	0.02	10.26
Expenditure·10	0.10	7.31	0.01	16.7
Family size·1000	2.04	1.11	-0.07	-0.62
Nr. of children under 16·100	-1.48	-5.9	0.03	1.7
Age·1000	-1.40	-11.57	0.03	3.62
Upper class·100	2.22	4.62	-0.03	-0.88
Lower midle class·100	-2.17	-5.9	-0.04	-1.73
Lower class·100	-4.96	-9.6	-0.03	-0.82
D2012·100	1.59	4.42	0.06	2.75
D2013·10	0.05	0.00357	0.02	6.55
Sec. edu. ·100	-0.33	-0.62	-0.03	-0.76
High. edu. ·1000	27.60	8.31	-0.02	-0.11
Tenant·1000	-7.80	-2.27	0.02	0.09
Accom. for free·100	1.233	1.49	-0.06	-1.1
BMI·1000	0.19	0.63	-0.05	-2.59
South·100	-2.44	-5.45	-0.03	-1.02

North·1000	38.88	7.01	-0.07	-0.2
East·100	-1.58	-2.77	0.02	0.48
West·100	0.50	1.1	0.04	1.24
Paris·10	0.28	6.1	0.01	3.36

Table 7: Elasticities from the adjusted AIDS model

	Fresh salmon		Label Rouge salmon		Other fish	
	Est.	t-val.	Est.	t-val.	Est.	t-val.
Fresh salmon price	-1.37	-110.93	0.02	50.98	0.27	18.95
Label Rouge price	0.89	6.87	-0.84	-5.21	0.36	7.99
Other price·10	4.21	211.97	0.03	19.39	-10.40	-4593.16
Total Expenditure	1.08	141.62	1.03	62.26	0.99	554.63

Note: t-values are calculated by boot strapping.

Table 8: Price and expenditure elasticity differences

	Abs. diff.	t-value
P11 – P22	0.54	-3.32
P11 – P33	0.33	-26.87
P22 – P33	0.20	1.26
P12 – P21	0.87	-6.71
P13 – P31	0.15	-10.06
P23 – P32	0.35	7.93
E1 – E2	0.05	2.51
E1 – E3	0.09	11.67
E2 – E3	0.05	2.72

Note: t-values are calculated by boot strapping.
P11 is the price of salmon in the equation for salmon.
P12 is the price of salmon in Label Rouge salmon etc.

Arnar Mar Buason

Arnar Mar Buason was born in Reykjavik Iceland 1987. He holds a BS degree in Economics from the University of Iceland (2011), and a MSc degree in Economics from the Norwegian University of Life Sciences (2013).

The thesis consists an introduction and five research papers. The papers are independent, but are all related to the integration of purchase or travel frequency in a demand analysis framework.

The motivation for the thesis is to point out the importance of purchase frequencies in demand analysis. Furthermore, to develop its theoretical foundations within microeconomics and econometrics, as well as providing a range of applications to demonstrate the value of the methods introduced.

Paper 1 introduces a microeconomic which incorporates how often consumers purchase goods. The paper also develops a corresponding econometric model. Finally, an empirical example is provided to show how the method can be used to form profitable loss-leader strategies.

Paper 2 introduces a travel cost model where the consumer jointly chooses the number of visits to a recreational site and how much time to spend at the site. The paper also develops a corresponding econometric model, which is applied to an Icelandic, stated preference, data set.

Paper 3 develops a microeconomic model which incorporates habits to explain the dynamics of consumption of everyday consumer goods and storable goods. It is assumed that habits can be separated into how often to shop for goods and how much to purchase each time. Finally, an empirical example is provided to demonstrate how the method can be used to formulate profitable marketing strategies.

Paper 4 is an empirical paper analyzing the French fish market using purchase frequency methods, with the aim of studying the markets consumer heterogeneity.

Paper 5 is an empirical paper which compares demand for fresh Label Rouge salmon and other fresh salmon in France to investigate if this label can produce the desired product differentiation and increased loyalty.

Supervisors: Kyrre Rickertsen and Dadi Kristofersson

School of Economics
and Business
Norwegian University of
Life Sciences (NMBU)
P.O Box 5003
N-1432 Ås, Norway

Contact:arnarmar@hi.is

Thesis number 2017:54
ISSN 1894-6402
ISBN 978-82-575-1453-2