The effects of country wealth on the energy mix

by

D.H.B. Konings

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on Friday August 28, 2020 at 14:00.

Student number:	4247159	
Project duration:	December 1st, 2019 – August 28th, 2020	
Thesis committee:	Prof. dr. J.J.C. Geerlings,	TU Delft, supervisor
	Prof. dr. ir. S. van Cranenburgh ,	TU Delft, 2^{nd} read
	Prof. dr. ir. P.M. Herder ,	TU Delft, 3^{rd} read
	Dr. T. S. Akkerhuis,	Royal Dutch Shell, supervisor

This thesis is confidential and cannot be made public until August 28, 2022.

TU Delft

Acknowledgment

First of all I would like to thank everyone who has supported me in working on this project. Thesis committee members, Hans Geerlings, Sander van Cranenburgh, Paulien Herder, and Thomas Akkerhuis, thank you for the feedback that helped in with the initial spelling and grammar errors, and especially in sharping the thinking that went in to the thesis and interpretation of the results. Shell's Scenarios Team, thank you for granting me the opportunity to work on my thesis with your team, and thank you for all the support and opportunities to reflect on the outcomes of the results and what this means for scenario development. A personal thank you to Chief Energy Advisor Wim Thomas and Martin Haigh, who granted me the opportunity to work with key parts of Shell's World Energy Model in order to support my research. Thank you to Tashi Erdmann for helping me out with dusting off my statistics. And a special thank you to Thomas Akkerhuis for supporting and mentoring me throughout the entire project. I have enjoyed working on this project and feel like a lot has been learned, this was not possible without your fantastic guidance.

Abstract

The total investments in the global energy system amounted to 1.85 trillion in 2018 [1]. Globally over 400 exajoule [2] was provided to energy consumers. The energy consumption is closely related to the emission of green house gasses [3] [4] [5]. In the Kyoto and Paris agreements the international community formalized the intention to reduce the emission of green house gasses in an effort to mitigate global warming [6] [7]. Economic progress has long been linked to increased energy consumption [8] [9]. Mitigation of green house gasses whilst facilitating economic growth poses one of the major challenge of the twenty first century [10] [11]. Development of climate policy and business strategy that facilitates both the mitigation of global warming and economic growth requires understanding of the energy market. In an effort to expand the understanding of the energy market we examine the historic effects of a country's wealth on the country's energy mix by applying a multinomial logit choice model for energy carriers to various sectors on a global scale. We find evidence for ordered wealth effects in sector categories Heavy Industry, Agriculture & Other Industry, Passenger Transport Rail, Freight Transport Rail, Residential Heating & Cooking, and Services.

List of Figures

List of Tables

Contents

1
Introduction

The global total final energy consumption was in excess of 400 exajoule in 2018 [12], and consumption has been growing by 1.6 % on an annual basis between 1973 and 2018 [12]. The annual investments associated with the energy industry amounted to 1.85 trillion USD in 2018 [1]. The energy industry is vast and an important enabler of economic growth, supporting a better quality of life [13] [5] [10] [11].

However, the global energy consumption is correlated with emission of green house gases (GHG) [3] [4] [5] . These green house gases in the atmosphere are correlated with the rising temperatures across the globe [14]. In the Kyoto, and Paris agreements the international community has committed to reducing GHG emission [6]. In the Paris agreement 196 parties have agreed to put forward their best efforts to reduce GHG emissions through nationally determined contributions (NDCs) [7]. The reduction of GHG emission and maintaining economic growth are at odds with each other so, in order to deliver on the environmental goals, solid strategy and policy development are essential.

Effective development of strategy and policy is facilitated by projections on their potential impact. Some of the better known energy system models are those of the International Energy Agency (IEA) [15], Imperical College London (Imperial) [16], Massachusets Institute of Technology (MIT) [17], The United Nations (UN) [18], McKinsey & Company [19], BP [20], ExxonMobil [21], and Shell's World Energy Model (WEM) [22].

For the development of strategy and climate policy both total energy consumption and types of energy carriers that are consumed are important, examples of energy carriers are gasoline and electricity. These carriers differ vastly in their processing and their environmental impact [3] [23]. From a business perspective, energy carrier demand information gives steer to investments. From a climate policy perspective it can reveal impact on green house gas emissions. This thesis aims to contribute to the understanding of what drives end-users to consume particular types of energy.

These energy models cover two facets, quantifying the total energy consumption, total primary energy (TPE) consumption and total final energy consumption (TFC), and projecting the market shares of energy carriers. The modeling approaches can be divided into three categories; bottom-up models, top-down models, and hybrids of the two [24] [25] [26].

Bottom-up approaches compute total energy consumption and energy carrier market shares based on individual market participant data. The market is modelled through gathering data on each end-user and creating projections for each end-user, who collectively create the market. The end-user refers to the consumer of the product. Consecutively, projections on the energy end-users are used for the total market projections. Increasingly more complex models require increasing quantities of data, and more complex data [27].

Top-down models can use general markers such as the development of prices, growth domestic product (GDP), and more, to provide projections on how the market will develop in terms of size and market shares. A top-down model can be scaled easily due to the less restrictive data requirements and therefore facilitate a consistent and transparent projection for various markets when consistent data is scarce [28]. Hybrid models facilitate the use of both bottom up and top-down approaches, such a model could use bottom-up modelling for one of the sectors and top-down for the others.

The energy system is comprised of several local and global markets with local and global stakeholders. Additionally, the problem of global warming is global. This drives the necessity for models with the ability to provide projections on the development of the energy system on a global scale. A consistent approach is easily restricted by data requirements [27], due to the scarcity of available data that is consistent on a global scale. Therefore the top-down approach is popular for projections on a potential future outcome of the energy system on the global scale.

This thesis expands the knowledge about top-down modeling of energy carrier market share modelling. When using realized data for the variables, Shell's WEM displays a trend breach between the energy carrier market shares it projects and the realized energy carrier market shares. Common causes for such a trend breach are the use of non-representative data, illegitimate model assumptions, or omitted variable bias [29]. This thesis focuses on the latter.

It is well known that energy consumption is correlated with GDP [30] [5] [10] [11]. Locally, it has been found that wealth is correlated with changes in the consumption of energy carriers [31] [32] [33] [34] [35]. We are researching if such a correlation can also be found on a global scale, and in multiple sectors. Wealth is projected to grow significantly over the coming couple of decades [13]. If there is evidence to be found of a relation between wealth and choice for energy carriers this may provide incentives to alter investment decisions by companies and policy development by governments and non-governmental organizations. Such shifts for companies can entail a shift from hydrocarbons to renewables for instance, and for governments a stronger support for a steep GDP growth trajectory.

We hypothesize that a country's wealth affects the energy mix. Which is to say that increasing wealth changes demand for types of energy carriers. To develop evidence concerning this hypothesis we try to answer the question; "Is the choice for energy carriers related to a country's wealth?". If this question can be answered with yes, the next question is; "Is there an order to the relation of a country's wealth and the choice for energy carriers?". In case there is evidence for a regional wealth effect and there is evidence for an ordered effect in country wealth effect on energy carrier choice, then the last question would be: "How much is end-user choice for energy carrier affected by increasing country wealth?"

In order to test the hypothesis of a relation between country wealth and end-user choice for energy carriers, we make use of revealed preference theory [36] [37] [38] [39] [40]. Market shares in a market are the result of consumer purchase decisions. Such a decision can be seen as an expression of preference for one of the alternatives available in the market. Revealed preference theory asserts that the way a market takes shape is the result of consumers choosing products that they prefer the most.

Applying the revealed preference framework requires data on the consumption of energy carriers. Through the IEA there are data available on end-user consumption of energy in absolute quantities for 198 countries, divided among 66 energy consuming sectors and 26 carriers. The consumption data from the IEA forms the basis of the revealed preferences of the end-users.

In order to compare the alternatives available to the end-user the concept of utility is applied. Utility has been developed as a way of expressing consumer preference in a single number. Utility is the result of attributes, such as price, of an alternative and characteristics, such as wealth, of the decision maker [41]. This approach allows for the decomposition of the utility in to it's constituent parts in order to facilitate the estimation of what attributes affect end-user choice. Shell's WEM formalizes a utility function that is linear in it's variables. The variables in 2019 were the price of energy carriers and a mode constant assumed to summarize the historic predisposition of energy carriers due to developments relevant to the energy system. It should be noted that using a constant to absorb unobserved variables does not mean that the effects of these variables are actually constant.

Estimating the effects of the relative price of energy carriers requires data on the prices of the energy carriers. The price data are available trough Enerdata for 198 countries, in 57 sector-energy carrier combinations. In order to investigate the effects of a country's wealth we add a proxy for wealth for the utility function. As a proxy for wealth we use real growth domestic product (GDP) per capita [42] [43] [44], where data are provided by the World Bank and the population data are provided by the United Nations.

The attributes that are considered, price and wealth, can be expressed on a continuous scale. However, a choice for an energy carrier is a discrete outcome. For the translation of the utilities of different energy carriers to their respective choice probabilities the multinomial logit (MNL) model is used, more theory behind the MNL is presented in chapter 2. The MNL can be used for estimating the effects of independent variables, continuous and discrete, on dependent variables that are discrete. The MNL has been developed based on the assumption that consumers choose the alternative that maximizes utility at the moment of choosing [45] [46] [47] [48] [39]. Maximum likelihood estimation (MLE) [48] [39] is used to estimate parameters for the mode constant, the marginal price and the wealth proxy.

The top-down approach to market projections is chosen due to the consistent approach it enables [9], even with scarce data. This is important from a coherent strategy and policy development perspective [10][11] [26]. In the research at hand we have to deal with incomplete data on energy consumption and pricing. This means that the estimates resulting from the data are potentially biased as missing data may lead to a non-representative sample. This is especially relevant for the research at hand, since wealthier countries tend to have better administration providing more complete data.

Additionally, there is a lack of knowledge about the true population size and applying an end-user choice framework infers that decisions are made on an end-user level. The consumption data is on an aggregate level, without knowledge on the number of end-users. A lack of knowledge out the number of end-users and the quantity of energy they consume means that we have no exact knowledge about the sample size we use. The approach that is used to deal with a lack of knowledge about the sample size is discussed in chapter 3.

The rest of this thesis report is structured as follows; chapter two elaborates on the theory. The third chapter discusses the methodology of the research. The results are presented in chapter four of the report. These results are discussed in the fifth chapter. Chapter 6 presents the conclusions and recommendations for further research.

2
Theory

This chapter presents the necessary theory for understanding the framework that we use to understand if country wealth is related to end-user choice for energy carriers. The theory chapter is divided into three sections. In the first section the multinomial logit model with the utility function that we use in our research is developed. Different approaches to market modeling are discussed in the second section of this chapter. In the third section, the theory involved in the uncertainty quantification for the parameter estimates is discussed.

2.1. Functional form

This section is divided into three subsections. The first subsection elaborates on the research problem, after which the second subsection develops the utility function that is applied in the research. The last subsection provides an overview of the multinomial logit function that maps the utilities onto probability space. Collectively, the utility function and the multinomial logit function form the functional form for the computation of the end-user choice probabilities for the energy carriers.

2.1.1. Energy mix modelling

This section provides further introduction of the research question. Any market takes shape by the purchase decisions made by the consumers in the market, the market for energy carriers is no different. The nature of the energy industry causes long supply chains with multiple intertwined markets for energy carriers, this research focuses on the end-users of the energy products. Therefore the term end-user is used to describe a consumer in the market for energy carriers, this means that end-users can be individuals as well as institutions such as companies.

The alternatives available to a consumer at a given time collectively represent a choice set. Therefore the term choice set refers to the set of energy carriers available to the end-users to choose from.

Our research aims to expand the knowledge about the markets for energy carriers. In order to do so, we are trying to infer what attributes of the alternatives in the choice set and characteristics of the end-user are of influence on the decision the the end-user makes. Mathematically we can describe this as follows:

$$f : I \rightarrow D \tag{2.1}$$

f is a function maps attributes of alternatives and characteristics of end-users provided through variables in *I* onto decision space *D*. The variables in *I* are the result of the data processing as described in chapter 3 and include, the attributes of the energy carriers, the characteristics of the end-users. Decision space *D* represents the consumption of energy carriers.

The data available are that of the aggregate energy consumption by carrier, sector, country, and year. In order to understand why the market for energy carriers is shaped the way it is, the the end-user choice framework is applied, which makes use of the concept of utility. Utility is not directly observed, instead it is inferred. This means that this is a latent variable problem. Based on variations in the distribution over choices in *D* and a set of variables *I* consisting of attributes such as the relative price and end-user characteristics such as regional wealth, a utility function is inferred. Based on our hypothesis, knowledge about the choice set and the end-users, we develop the mathematical framework in section 2.1.2.

2.1.2. The Utility Function

The end-user choice framework for energy carrier choice modeling is based on random utility maximization [36] [46]. Utility is a concept that functions as a building block of the mathematical framework used for the ordering of a set of alternatives available to the end-user. Consumers are assumed to always choose the alternatives that provides the largest utility, i.e. they are utility maximizing. Therefore, a utility function should assign a number to an alternative in the choice set based on it's attributes and the characteristics of the decision maker [41], such that:

$$C_n = i \iff U_i > U_j \qquad \forall j \neq i. \tag{2.2}$$

In eq. (2.2) C_n is the choice by end-user n for and alternative from the choice set, and U_i and U_j are the utilities associated with alternative i and alternative j respectively.

The utility function used for our research is linear in the variables. By acknowledging that there are attributes unknown and unobservable to us we recognize the need for an error term. This brings us to the formulation of the utility for alternative i to decision maker n:

$$U_{n,i} = \boldsymbol{\beta}_{n,i} \cdot \mathbf{x}_{n.i} + \epsilon_{n,i}. \tag{2.3}$$

In eq. (2.3) $U_{n,i}$ is the utility for decision maker n of alternative i in the choice set, $\boldsymbol{\beta}_{n,1}$ is a vector of parameters corresponding to the variables in vector $\mathbf{x}_{n,i}$, and $\epsilon_{n,i}$ is the error term representing additional unobserved attributes of the alternatives and characteristics of the end-users on the utility.

The utility function that is used has three variables; mode constant, the energy carrier price, and the wealth proxy indicating the average wealth in the decision maker's market. The specific data used for these variables and how this data is processed is elaborated on in chapter 3. The utility function we use for all sectors is formalized as follows:

$$U_{n,i} = \beta_{0,i} x_{0,i} + \beta_{1,i} \cdot x_{1,n,i} + \beta_{2,i} \cdot x_{2,n} + \epsilon_{n,i}. \tag{2.4}$$

Here $U_{n,i}$ is the utility that energy carrier i brings to end-user n. It should be noted that this is a stochastic value, this is where the name random utility maximization (RUM) is derived from [46]. $\beta_{0,i}$ is the parameter associated with the constant $x_{0,i}$, this parameter is the mode constant. The mode constants absorbs the effects of any variables not included in the model, such as a predisposition of an energy carrier because of natural resources. The correction is required because, among other things, the energy system does not start out with a clean slate, the historic developments reflected in the data would introduce a bias if not corrected for. The mode constant can be interpreted as the utility the energy carrier enjoys for the end-user due to historic developments in the energy system in our data set. $\beta_{1,i}$ is the parameter for price. $x_{1,n,i}$ is the price for of the energy carrier. $\beta_{2,i}$ is the parameter for the wealth proxy. $x_{2,n}$ is the wealth proxy for all sectors in the region where end-user n consumes the energy carrier. $\epsilon_{n,i}$ is the error term.

2.1.3. The Multinomial Logit Function

To develop a function that computes the choice probabilities for the energy carriers we revisit eq. (2.2) and the utility in eq. (2.4) which allows us to formulate the probability that alternative i is chosen:

$$P_{r,s,n,i} = Prob(U_{r,s,n,i} - U_{r,s,n,j} > 0, \qquad \forall j \neq i). \tag{2.5}$$

In eq. (2.5) $P_{r,s,n,i}$ is the probability that in sector s in region r the end-user n chooses energy carrier i. Equation (2.5) provides an important property of the applied framework when combined with the independence of irrelevant alternatives of the alternatives (IIA) [48] [39] in the choice set. The IIA property means that the results of the comparison of two alternatives is not influenced by an irrelevant alternative. If the alternatives in the choice set are IIA, then only relative utility of the alternatives in the choice set is relevant to the decision.

$$\begin{aligned}\Delta U_{r,s,n,i-j} &= U_{r,s,n,i} - U_{r,s,n,j} \\ &= \beta_{0,r,s,n,i-j} x_{0,r,s,n,i-j} + \beta_{1,r,s,n,i-j} \cdot x_{1,r,s,n,j} + \beta_{2,r,s,n,i-j} \cdot x_{2,r} + \epsilon_{r,s,n,i-j}\end{aligned} \tag{2.6}$$

In order to answer the question if a country's wealth is related to the energy mix we aim to estimate $\beta_{2,r,s,n,i-j}$ in eq. (2.6). This parameter represents the utility of energy carrier i with respect to energy carrier j in relation with the wealth of country r. Equation (2.6) provides a completely formalized equation for comparing the alternatives in the choice set. However, it does not provide the choice probabilities necessary to estimate the parameters.

To develop the functional form that computes the choice probabilities of the energy carriers based on their utility eq. (2.4), needs to be split into two parts; the parameters and variables, and the error term $\epsilon_{r,s,n,i}$:

$$V_{r,s,n,i} = \beta_{0,r,s,n,i} x_{0,r,s,n,i} + \beta_{1,r,s,n,i} \cdot x_{1,r,s,n,i} + \beta_{2,r,s,n,i} \cdot x_{2,r} \tag{2.7}$$

$$error_{r,s,n,i} = \epsilon_{r,s,n,i} \tag{2.8}$$

Such that:

$$U_{r,s,n,i} = V_{r,s,n,i} + \epsilon_{r,s,n,i}. \tag{2.9}$$

When the error term $\epsilon_{r,s,n,i}$ is independent and identically distributed (i.i.d.) with a Generalized Extreme Value (GEV) type 1 distribution, then $\epsilon_{r,s,n,i-j} = \epsilon_{r,s,n,i} - \epsilon_{r,s,n,j}$ ineq. (2.6) is logistically distributed [39] [48]. If $\epsilon_{r,s,n,i-j}$ in eq. (2.6) is logistically distributed it can be shown that eq. (2.6) leads to the closed form probability function multinomial logit:

$$P_{r,s,n,i} = \frac{e^{V_{r,s,n,i}}}{\sum_{j=1}^{J} e^{V_{r,s,n,j}}}. \tag{2.10}$$

In eq. (2.10) $P_{r,s,n,i}$ is the probability that energy carrier i is chosen by end-user n in sector s and region r and $V_{r,s,n,i}$ is the result from eq. (2.7). The derivation of the multinomial logit model can be found in appendix C.

2.2. The Market Model

The probability of an end-user's decision to choose for a specific product affects the market. Under the theory that has been developed so far we can state that the end-user choice probabilities resulting from eq. (2.10) can be equated to the energy carrier market shares [47], or energy mix on an annual basis, given that the parameter estimations are executed with the annual total final energy consumption that is free for choice.

The equivalence is not entirely trivial, to see this we derive the equivalence statement. The market in year t is the aggregate of all end-user decisions in year t:

$$M_t = \sum_{n=1}^{N} \sum_{i=1}^{J} C_{n,i} \cdot Q_n. \tag{2.11}$$

In eq. (2.11) M_t is the market in year t and C_n is the choice for energy carrier i by end-user n. This consumer's decision consists of a volume Q_n. The J in eq. (2.11) refers to the choice set. The choice probability in eq. (2.10) is developed for all end-users in the market [48] [39]. If all decisions $C_{n,i}$ in eq. (2.11) are based on eq. (2.10) then with a large enough N the the market size is $\sum_{n=1}^{N} Q_n$ and the ratio of energy carriers in market M_t is equal to the choice probabilities as described by eq. (2.10), such that:

$$M_{t,i} = M_t \cdot P_{r,s,n,i} = \sum_{n=1}^{N} C_{n,i} \cdot Q_n. \tag{2.12}$$

In eq. (2.12) $M_{t,i}$ is the market for energy carrier i in year t, M_t is the total market for energy carriers in year t, $P_{n,i}$ is the probability that end-user n chooses carrier i as in eq. (2.10), $C_{n,i}$ is the choice of end-user n for energy carrier i, and Q_n is the corresponding volume of the choice of end-user n. The equivalence of the market shares and the choice probabilities is used to estimate the β's in eq. (2.4), such that the parameter estimates for the model can be interpreted as the effect of the variables on the energy carrier market shares.

2.2.1. The Full Market Model

There are two ways to approach the market, the full market model and a variable market model. When applying a choice model, the full market model is equivalent to assuming that on an annual basis the entire market is up for choice. This means that under the full market model the annual energy carrier market is equal to the total final consumption of energy that year. From a market perspective this may be correct. However the choice modelling framework requires that all end-users can choose freely what energy carriers they want to use to produce the energy service they want, in order to be able to interpret the parameter estimates as effects on the end-user choice probabilities.

There are two distinct advantages to the full market approach, the simplicity of the interpretation of the estimation results and the fact that it does not require additional information to estimate parameters. There are also pitfalls. The full market model neglects the existence of market momentum. Market momentum refers to the slow change of market shares as a result of an external shock, this shock can be price changes or other factors relevant to market shares. For the energy system, investments in production assets, that can only use one type of energy carrier, which are utilized for multiple years is one of the major sources of market momentum.

From the perspective of the wealth effects on the end-user choice probabilities, the market momentum may distort the parameter estimates. When applying the choice modelling framework the market shares are interpreted as the choice probabilities. If the changes in the market shares as a result of changes in the variables are dampened by the market momentum, then the effect of the variables on the end-user choice may be under estimated. The annual market in full market model can be formalized as follows:

$$M_{r,s,t} = (1 + r_{growth}) \cdot TC_{r,s,t-1}. \tag{2.13}$$

Here $M_{r,s,t}$ is the market in year t, $TC_{r,s,t-1}$ is the total consumption in year $t-1$, and r_{growth} is the growth rate going from year $t-1$ to year t.

2.2.2. The New and Churn Market Model

The variable market model considered for our research is referred to as a new and churn market model (NCMM). In the new and churn model the annual market is defined as the growth of the total market (new) and a percentage of the total market in the previous year (churn). Churn is usually a term used when referring to market attrition, in the context of this model it represents the part of the annual energy consumption that has no predetermined energy carrier. Another way to look at this is to say that $100 \cdot (1 - churnrate)\%$ of the annual energy market will use the same type of energy carriers it did the previous year. The new and churn model can be formalized as follows:

$$M_{r,s,t} = TC_{r,s,t-1} \cdot r_{growth,r,s,t-1} + TC_{r,s,t-1} \cdot r_{churn,r,s} \tag{2.14}$$

Here $M_{r,s,t}$ is the market for year t in sector s in region r and $TC_{r,s,t-1}$ is the total consumption in year $t-1$, r_{growth} is the growth rate going from year $t-1$ to year t, and r_{churn} is the churn rate.

The churn rate usually associated with asset replacement or contract expiry. The new and churn market model considers market momentum created by for instance investments into energy consuming production assets planned to last for multiple years by carrying a percentage of each market segment to the next market year.

The most important advantage of the new and churn model is that it incorporates market momentum. This changes how we can interpret the parameter estimates. Under the new and churn market model we only consider the part of the market where end-users are realistically free to choose among energy carriers. The relation of variables with the NCMM market can therefore be interpreted as the relation of these variables with the choice probabilities of the end-users. The pitfall is that there is an additional unknown in the model, the churn rate. The interpretation of parameter estimations with a new and churn model are only valid for the churn rate used during estimation, because the churn rate affects the magnitude of the changes in the market shares.

The new and churn market model focuses on that section of the market where the changes in the energy mix are more pronounced. Provided that the assumed churn rate is a good fit, these more pronounced changes in the energy mix potentially facilitate parameter estimation because the magnitude of the changes in the market shares are enlarged.

2.3. Uncertainty quantification

One of the purposes of top-down energy system modeling is overcoming the scarcity is consistent data, whilst providing a consistent modelling approach. The historic data used for our research is no exception and has gaps too. The gaps in the data introduce uncertainty in the parameter estimates, because the data are a sample of a true population. Therefore, we need to determine the uncertainty of our estimates.

In addition to uncertainty, the data gaps may introduce a bias in the estimates. If specific parts of the data are missing then the data set may be non-representative for the intended purposes. This means that the parameter estimates cannot be generalized for the entire population, instead it would only be suitable to model the sub-population properly represented in the sample.

From an end-user perspective the sample size is not available in the data. The consumption data provided by the IEA is the aggregate of all end-user choices, however the number of market participants or end-user choices is not quantified. The unknown sample size in each market introduces a challenge to the research. Parameter estimates can be performed, however the metrics generally used to assess uncertainty of the estimates: standard error, likelihood ratio, z- and p-values, are computed using the sample size. These can therefore not be relied upon. This leaves us with two additional challenges; examining the influence of the hypothetical size Q_n in eq. (2.11) and the quantification of the uncertainty of our parameter estimates.

To overcome the challenge of the unknown population size N we disaggregate the total final consumption data. This is equivalent to assuming a fixed Q_n, such that:

$$N = \frac{TC}{Q_n}. \tag{2.15}$$

In eq. (2.15) TC is the total consumption, and Q_n is the size of a choice batch. A choice batch refers to the consumption volume of end-user n similar to Q_n in eq. (2.11). Executing estimations for varying Q_n should reveal the impact Q_n has on the parameter estimates. When consecutive disaggregation scales align their parameter estimates sufficiently, then this can be interpreted as evidence for a suitable disaggregation scale that does not bias the parameter estimates. The magnitude of Q_n can give insight in the how influential the data accuracy of the IEA is.

To overcome the challenge of the invalid uncertainty quantification, bootstrapping is introduced. By repetitively estimating the parameters with different, same size, samples with data points picked with uniform probability from the original data set we create a distribution for the parameter estimates. The uncertainty quantification is based on the t-test. The only difference is that we notice that there is no N available, therefore it is not possible to compute a standard error. In an effort to resolve the issue of the standard error we introduce the relative distance defined as the mean of the estimate divided by it's standard deviation:

$$RD = \frac{\bar{\beta}}{\sigma_\beta} \tag{2.16}$$

In eq. (2.16) RD is the relative distance, $\bar{\beta}$ is the mean parameter estimate for β, and σ_β is the standard deviation of the distribution of the parameter estimate for β. The relative distance is subsequently used to assess the parameter estimates and their distance to zero. Their distance to zero is used similarly as it is when testing a null hypothesis, if the magnitude of the parameter estimate is large enough with respect to the standard deviation then the parameter is concluded to be sufficiently different from zero to suggest an effect of the variable associated with the parameter. Chapter 3 provides more on the assessment of the uncertainty of the parameter estimates presented in chapter 4.

3 Methodology

This chapter builds on the theory developed in the theory chapter. We proceed by discussing how we executed the theory developed in chapter 2. The first section of this chapter will discuss the characteristics of the data that has been used during the research. The second section discusses the execution of the parameter estimations. We will give special attention to our strategy for dealing with uncertainty in our estimations.

3.1. Data

This section will discuss the properties of the data that has been used for the research. After providing a full summary of the raw data available, we will continue with the processing of the data to the feature space that has been used to estimate the model parameters.

3.1.1. Data Description

The dependent variable in the model is the the total final consumption (TEFC) of the energy carriers. The raw data is provided by the International Energy Agency (IEA). The IEA provides annual TFC data on a country level, from 1960 up to and including 2016. The IEA discerns 26 sectors and 66 energy products. The unit for the provided TFC data is terajoule (TJ).

The raw price data comes from Enerdata. Enerdata provides annual pricing from 1978 up to and including 2016. This data is provided as the in real term 2015 United States Dollars (USD) price per ton of oil equivalent (toe). This means that prices are inflation corrected and portray the same price unit for all carriers. Enerdata provides prices on a country level and discerns 17 different combinations of sectors and carriers.

There is no direct quantification of wealth and therefore we use a widely used metric for the regional wealth; growth domestic product (GDP) per capita [42] [43] [44]. This requires two types of data, GDP data and the population data. The GDP data is provided by the World bank, from 1978 up to and including 2016. The GDP data is provided in real 2016 USD, at purchasing power parity exchange rates. This means that the data is inflation corrected and reflects the standard of living in different countries. The population data is provided by the UN, from 1950 up to and including 2016.

From a data availability perspective Enerdata limits the data set from 1978 up to and including 2016.

3.1.2. Data Processing

This sub section will address three issues with the data. Firstly we elaborate on the data segmentation, there is a clear mismatch of definitions between the IEA TFC data and the Enerdata pricing data which needs a mapping. Secondly, in order to compare the quantities of energy carriers that have been consumed the concept of Energy Service (ES) is introduced. This concept is addressed after the segmentation of the data. Thirdly, the application of the market models is discussed.

The energy consumption data is categorized into specific types of energy consumption, the carriers, and different sectors where the energy is consumed to produce services. The categorization facilitates the estimation of the effects of various variables on the energy mix. Without the categorization, the heterogeneity in the end-users and varying choice sets would drastically impair the ability to estimate how variables affect the energy mix. [39] [48]

The TFC data and the pricing data show a mismatch in the categorization of energy consumption into sectors and carriers. In order to estimate the effects of price changes and wealth changes the segmentation needs to align. To facilitate the alignment of the region-sector-energy carrier combinations the segmentation of Shell's World Energy Model (WEM) is applied [22].

The WEM identifies one hundred geographical regions. From an energy consumption perspective, the 82 largest countries are the first 82 regions. From here on the word region can be interpreted as country or group of countries. The remaining countries are combined into 18 regions. They can be found in appendix A. There are fourteen different energy consuming sectors in the WEM, these are broadly categorized as industry (2 sectors), transport (8 sectors), buildings (3 sectors), and non-energy use. This division is a result of modelling, energy consumption and business considerations. There are ten different energy carriers available in the WEM for the energy consumption presented in table 3.1.

The fact that IEA, Enerdata, World Bank, the UN and the WEM use different country sets, different carrier structures, and different sectoral structures, requires a set of mapping rules. The exact mapping is proprietary to Shell, and is internally called the “balances process”, referring to the IEA's yearly energy balances.

To give an idea of the mapping we provide some examples; for the energy carriers the mapping of charcoal and petroleum coke, two different carriers in the IEA data, are mapped onto the carrier solid hydrocarbon fuels in the WEM. Similar is the IEA sector air transport, which is divided into passenger transport air and freight transport air in the WEM based on information from external sources.

For the country to region mapping the US remains the US, as it is one of the largest 82 energy consuming countries. On the other hand, Bermuda is separately discerned in UN country data and part of the region Rest of North America in the WEM, as it is not in the list of the 82 largest energy consumers. The result of this process is a set of 14.000 unique region-sector-carrier combinations(100×10×14). These 14.000 unique region-sector-carrier combinations are on an annual basis.

Table 3.1: The energy carriers identified as distinct energy carriers in Shell's World Energy Model.

Energy Carriers
Solid Hydrocarbon Fuels (SHCF)
Liquid Hydrocarbon Fuels (LHCF)
Gaseous Hydrocarbon Fuels (GHCF)
Electricity - Commercial
Electricity - Solar PV Distributed
Hydrogen
Heat - Commercial
Heat - Solar Thermal Distributed
Biomass - Commercial
Biomass - Traditional

The categorization of energy carriers presented in table 3.1 for the WEM is created with one foot in history and one foot in the future. Therefore, some of the discerned energy carriers cannot realistically be included into the choice set. The technologies have not been or are still not available on a scale such that they are realistically an alternative to the end-users in the period 1978 to 2016. Therefore the energy carriers, hydrogen, decentral photovoltaic electricity, decentral solar thermal, and commercial heat are dropped from the data set. Heat - Commercial refers to district heating, which is historically only used in specific countries and is therefore discarded from the choice set.

The pricing data from Enerdata provides a scarce data set on the marginal energy prices. The marginal price is the price for consuming an additional quantity of energy carrier, it does not take into account any capital expenditure. In order to compare alternatives in the choice set, the data on the attributes for all the alternatives needs to be complete. This means that missing pricing data for either one of the alternatives requires us to drop the data point. The raw marginal pricing data from Enerdata has numerous gaps, such that a restrictive number of region-sector-year data points have at least one missing energy carrier price. The price gaps are filled using a proprietary algorithm from Shell's scenarios team.

This algorithm has an ordered set of approaches to compute the missing pricing data. Assuming a price is missing, the first approach finds two prices for the energy carrier for the last year before the data gap and the first year after the data gap, consecutively it computes a cumulative average growth rate (CAGR) base on the price difference and based on the CAGR the missing price is computed. Whenever there is no more data on either side of the data gap, pricing data from within the surrounding region is used, i.e. the prices changes from the other carriers in the sector are used as an indication for the annual price changes in the price for the energy carrier with the missing data. Whenever the other energy carriers are also missing data points to use as a steer, the steer is sought in the surrounding regions. The prices and their growth from the surrounding regions are averaged out and used to compute the prices for the region with the missing pricing data.

Energy carriers have different forms and attributes, among which the extractable energy content. This influences the amount of work that an end-user can extract form the energy carrier. If you choose to travel by car for 20 km on electricity you will need a different quantity of energy than when you choose petrol. Since the consumption data is given in terajoule there is need for a conversion factor that allows for comparison of the consumed quantities from an end-user perspective. These units are referred to as energy service in the IEA model and Shell's WEM [15] [22], ExxonMobil and BP use a conversion to tonne of oil equivalent [21] [20]. The sectors identified as separate market within a region in Shell's World Energy Model and their respective Energy Service units are displayed in table 3.2.

Table 3.2: The sectors identified as unique markets within a region in Shell's World Energy Model and their respective Energy Service units.

Sector	**Energy Service Unit**
Heavy Industry	Tonne of Steel equivalent
Agriculture & Other Industry	MJ Heating Requirement in Buildings
Services	MJ Heating Requirement in Buildings
Passenger Transport - Ship	Passenger Kilometre
Passenger Transport - Rail	Passenger Kilometre
Passenger Transport - Road	Passenger Kilometre
Passenger Transport - Air	Passenger Kilometre
Freight Transport - Ship	Tonne Kilometre
Freight Transport - Rail	Tonne Kilometre
Freight Transport - Road	Tonne Kilometre
Freight Transport - Air	Tonne Kilometre
Residential Heating & Cooking	MJ Heating Requirement in Buildings
Residential Lighting & Appliances	Electricity Need
Non Energy Use	Oil equivalent for output

The efficiency with which a terajoule of a carrier is converted into an energy service is called the energy service efficiency (ESE), such that:

$$ESE = \frac{ES}{EC} \tag{3.1}$$

In eq. (3.1) ESE is energy service efficiency, ES is energy service, and EC is the quantity of energy consumption required for the production of the energy service. Using ESEs, the TFC data in terajoules can be converted to energy service (ES) data in sector-specific measurement units as presented in table 3.2. The ESEs are provided by Shell and are year-country-sector-carrier specific. This allows for energy carrier market shares to be compared based on the energy service produced with the carriers. An example is the kilometres travelled on liquid hydrocarbon fuels compared to gaseous hydrocarbon fuels in the passenger transport sector, instead of the terajoules of liquid hydrocarbon fuels and gaseous hydrocarbon fuels.

Energy carriers' prices are converted to energy service prices accordingly. The pricing data from Enerdata is in the unit USD per tonne of oil equivalent. After a unit conversion from USD per toe to USD per terajoule the same energy service efficiencies that are used for the conversion of TFC to total final energy service consumption can be used to convert the price from USD per terajoule to USD per energy service. This conversion enables the comparison of the price of the energy carriers in the choice set on an equal footing.

This leaves us with the task of determining a suitable churn rate. To determine the churn rate, parameter estimates have been executed under the new and churn market model with churn rates ranging from 5 % up to and including 100%. These experiments provide us with varying estimates, variances, and feature spaces. Where the parameter estimates are stable, the variances are acceptable and the features spaces drop crisis years, then the churn rates are assessed to be suitable.

Ideally we would estimate parameters for each unique region-sector. Due to the incomplete data on each region, parameter estimations for each individual region are not feasible. Therefore, we are required to assume that each region can be modeled with the same parameters. Such that we can pool the region-sector feature spaces into one sector feature space.

Differences in the energy carrier market shares should now be the result of relative price differences within the region and the wealth proxy of the region. Such wealth effects would create an additional effect of wealth on the energy system, in addition to the concept of energy ladders [12]. This means that in addition to the total energy consumption, also the energy mix in a country would be affected by wealth.

Following the steps set out in this section resulted in the features spaces that have been used to estimate the parameters in eq. (2.7). The summary statistics for the features spaces can be found in appendix E. The summary statistics for the features spaces can be found in appendix E. The distribution of the data with respect to wealth can be found in appendix G, these suggest that there is sufficient data for less wealthy regions.

In order to facilitate the interpretation of the parameter estimates, one of the energy carriers in the choice set is chosen to function as the pivot. This means that for each year-region-sector data point one of the carriers is used as a reference point for the relative variables. To do so all parameters associated with one of the carriers are set equal to zero. For the mode constants and the wealth parameter simply defining one of the carriers' parameter estimates to be zero is sufficient. For the price parameter we set the price for one of the carriers in each year-region-sector data point equal to zero by subtracting the energy service price for that carriers from all energy service prices. This leaves us only with the price differences for each year-region-sector data point.

3.2. Parameter Estimation

The first part of this section will discuss how we go about the disaggregation of the data. The second part of this section elaborates on the choice for maximum likelihood estimation for the parameter estimations. The third part of this section will focus on the approach that we have taken to assess the uncertainty of the estimations.

3.2.1. Data disaggregation

The total final energy service consumption data is the aggregation of decisions for specific carriers. This poses three problems for our analysis. Firstly, the absence of sample size prohibits formalizing the estimation error. Secondly, parameter estimations based on the market shares alone would give a distorted image of the energy system as it grants equal weights to the US energy market and the Dutch energy market unless weighing is applied. Thirdly, the preferred multinomial logit maximum likelihood estimation software package does not work with market shares as the dependent variable. To overcome all problems we apply disaggregation to the individual data points.

Each single year-country level data point is disaggregated into a new set of data points. Such that it generates N data points for each year-country level data point:

$$N = \frac{TFEC}{Q_n} \tag{3.2}$$

In eq. (3.2) $TFEC$ is the total final energy consumption for a year-country level data point, and Q_n is the size of a choice batch associated with end-user n (for instance 1 PJ). The disaggregated data points receive labels indicating the decision for an energy carrier corresponding to the market shares based on energy service consumption in the original year-country level data point, such that:

$$MS_i = \frac{N_i}{N} \tag{3.3}$$

In eq. (3.3) MS_i is the market share of energy carrier i, N_i is the number of choice batches for energy carrier i, and N is the total number of choice batches for the year-country data point. So, we go from 1 year-country level data point to several. If these new data points were to be aggregated, they would result in the exact single year-country level data point that we started out with.

This disaggregation approach resolves the problems introduced in the first paragraph of this subsection. The population size is still unknown, however this approach allows for analyzing the effects on the parameter estimations of assumptions about the population size. This is not only useful for the quantification of the errors in our parameter estimates, it also reveals at what scale an error in the data from a data bank will potentially bias the parameter estimates. This can be interpreted as a minimum required accuracy for the data to fit a model such as the one at hand.

Since the data has been pooled into global sectors, the disaggregation of the data based on $TFEC$ is equivalent to giving each region a weight equal to the size of the energy market:

$$\frac{N_a}{N_b} \approx \frac{TFEC_a}{TFEC_b}. \tag{3.4}$$

In eq. (3.4) N_a and N_b are the number of choice batches for region a and b, and $TFEC_a$ and $TFEC_b$ are the total final energy consumption for region a and b. Reducing Q_n will lead to a larger N which is equivalent to a larger number of data points, leading to a closer fit of the market shares.

3.2.2. Maximum Likelihood estimation

For the parameter estimations maximum likelihood estimation (MLE) is used. This method estimates parameters by maximizing the likelihood of the occurrence of the data set as a function of the parameters. The method is proven to be consistent under trivial assumptions, most importantly the assumption that the model is a good reflection of reality. It is especially popular for estimating discrete choice models such as the multinomial logit [39] [46]. For computational convenience, the monotone logarithm-transformation is usually applied and MLE amounts to maximizing the log-likelihood::

$$\log L(\boldsymbol{\beta}_S) = \sum_t \sum_r \sum_n \sum_i log\left(\frac{e^{V_{r,S,n,i}}}{\sum_j e^{V_{r,S,n,j}}}\right). \tag{3.5}$$

In eq. (3.5) the subscript r is for region, n is for a choice batch, i is for an energy carrier, S is for the sector the parameters are estimated for, j indicates a sum of all the energy carriers in the choice set, and $V_{r,S,n,i}$ is the utility as formulated in eq. (2.4). Which due to the disaggregation process is equivalent to:

$$\log L(\boldsymbol{\beta}_S) = \sum_t \sum_r \sum_n \sum_i \left(\frac{TFC_{r,S,n,i}}{\sum_t \sum_r \sum_n \sum_j TFC_{r,S,n,j}} \cdot log\left(\frac{e^{V_{r,S,n,i}}}{\sum_j e^{V_{r,S,n,j}}} \right) \right). \tag{3.6}$$

It should be noted that when Q_n in eq. (3.2) is small enough, then eq. (3.6) is equivalent to a MLE regression of the variable against the market shares with the energy carrier market size as weight.

3.2.3. Uncertainty quantification

The disaggregation does not solve the problem of not knowing our population size. Formal uncertainty measures rely on the population size. Therefore, we devised an approach resulting in 'relative uncertainty'. We apply the principle of the t-statistic to quantify the relative uncertainty:

$$RD = \frac{\bar{\beta}}{\hat{\sigma}_\beta} \tag{3.7}$$

In eq. (3.7) RD is the relative distance, $\bar{\beta}$ is the mean estimate for β, and σ_β is the the standard deviation for the distribution of $\hat{\beta}$, the estimate of β. The relative distance refers to the distance between the parameter estimate and a reference point. An actual t-test assumes that $\bar{\beta}$ follows a normal distribution with variance σ^2/n, for the relative distance neither need to be true. If the relative distance to zero is large enough for a parameter estimate β associated with independent variable x, then we interpret this as evidence for an effect of the independent variable x on the dependent variable. Whatever we interpret as being "large enough" will become apparent within the context of other estimates in the chapter 4.

To approximate a parameter distribution, we apply bootstrapping before the disaggregation for the parameter estimations. According to a uniform distribution a sample of the data set is draw with replacement. The sample size is equal to the size of the original data set. Successively the sample is disaggregated for the parameter estimation. A repetition of this process will result in a distribution of the maximum likelihood estimator. To be able to analyse the influence of the disaggregation scale, we perform bootstrapping for 0.1, 1, 10 and 100 PJ scale disaggregation.

3.2.4. Computational toolkit

All data processing and computational work has been executed using Python 3.7. The python packages pandas and numpy have been used have been used to process the data. In order to create the model, the python api statsmodels was used. From stats models the MNLogit python method was used in combination with a Newton Raphson maximum likelihood estimation method for fitting the model to the data. All other computations have been programmed by hand.

4

Results

This chapter presents the outcomes of the parameter estimations and uncertainty quantification. The parameter estimates are presented as the mean value of the bootstrap results, the value between the square brackets is the relative distance defined in eq. (2.16). In the first section we present the findings on disaggregation scale. The second section presents the sensitivity on churn rate assumptions. The parameter estimates are presented in the third section of this chapter.

It is important to note that there are no results for the sectors Passenger Transport Ship, Passenger Transport Air, Freight Transport Ship, Freight Transport Road, Freight Transport Air, and Residential Lighting and Appliances, because there is no choice between multiple carriers in these sectors. The Non Energy Use sector represents all energy carriers used in chemical processes, not for their energy content, and as such, they are not interchangeable. Therefore, we exclude the Non Energy Use sector.

4.1. Disaggregation Scale Sensitivity

In table 4.1 the parameter estimates and the relative distances to zero are presented for the full market model. At this point the interpretation of the estimation outcomes themselves are not yet the point of interest. The purpose of this table is assessing the influence of the disaggregation scale on the parameter estimates.

In Table 4.1 we start by analyzing the difference between a disaggregation scale of 1 PJ and 10 PJ by analyzing the difference between columns (1) and (2), (4) and (5), and (7) and (8). A noticeable change in the parameter estimates between the columns in the column pairs would provide evidence of an effect of the disaggregation scale on the parameter estimates. The results in the column pairs do not provide evidence of effects on the parameter estimates by a change in disaggregation scale from 1 PJ to 10 PJ. It should be noted that most of the estimates do present evidence of a slightly increased standard deviation.

The second difference that we analyze is the difference in the parameter estimates for the disaggregation scales of 10 PJ and 100 PJ, these can be found in table 4.1 in columns (2) and (3), (5) and (6), and (8) and (9). Similar to the 1 PJ and 10 PJ scale an effect of the change in disaggregation scale presents itself as a noticeable change in parameter estimates. A change in disaggregation scale from 10 PJ to 100 PJ has noticeable effects on the parameter estimates. Either the relative distances drop drastically or parameter estimates increase with disproportional changes in relative distance. 100 PJ disaggregation is not deemed suitable for executing the disaggregation for the parameter estimations, because the resolution of the data becomes too low, resulting in erratic parameter estimates.

The slight increase in standard deviation for the change in disaggregation scale from 1 PJ to 10 PJ does no provide any information about the precision of the estimates. As mentioned in chapter 2 and chapter 3 the change in disaggregation scale does not provide information about the actual population size underlying the data set. However the points estimates show no noticeable change, especially compared to the change in disaggregation scale from 10 to 100 PJ. The lack of evidence for an effect on the parameter estimates between 1 PJ and 10 PJ is interpreted as evidence for a stable parameter estimation for a disaggregation scale below 10 PJ.

In addition to the 1 PJ, 10 PJ, and 100 PJ scale disaggregation experiments under the full market model regime, we created bootstraps for the new and churn model with disaggregation scales of 0.1 PJ, 1 PJ, and 10 PJ. The estimation results are presented in table 4.2. The parameter estimates portray similar results compared to the full market model described in the previous paragraphs. The 0.1 PJ disaggregation and the 1 PJ disaggregation parameter estimates in columns (1) and (2), (4) and (5), and (7) and (8) show no noticeable differences. Similar to the full market model there are minute changes in relative distance. However the step from 1 to 10 PJ disaggregation, from columns (2), (5), and (8) to (3), (6) and (9) respectively, introduces noticeable differences. This is likely the result of the lower granularity in the data. The lack of evidence for an effect on the parameter estimates between 0.1 PJ and 1 PJ is interpreted as evidence for a stable parameter estimation for a disaggregation scale below 1 PJ

The new and churn market model considers the market growth and the market churn, this is a significant reduction in the considered annual energy consumption. The different scale of consumed energy quantity results in two different optimal disaggregation scales. Consistency in disaggregation scale across market models and the increasing computational time at higher granularity resulted in the choice for a 1 PJ disaggregation scale for the parameter estimates presented in section 4.3.

Table 4.1: Estimation results for the estimates of β_0, β_1, and β_2, i.e. the *mode constants*, the *relative energy service price parameters*, and the *wealth proxy effect* respectively, for a disaggregation scale of 1, 10 and 100 PJ under the Full Market Model; (1) Heavy Industry 1 PJ, (2) Heavy Industry 10 PJ, (3) Heavy Industry 100 PJ, (4) Passenger Transport Road 1 PJ, (5) Passenger Transport Road 10 PJ, (6) Passenger Transport Road 100 PJ, (7) Residential Heating & Cooking 1 PJ, (8) Residential Heating & Cooking 10 PJ, and (9) Residential Heating & Cooking 100 PJ.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\beta_{0,GHCF}$	−3.607	−3.702	−4.147	0.328	−0.122	−4.211	0.604	0.709	1.288
	(−20.37)	(−19.10)	(−16.42)	(0.667)	(−0.271)	(−4.000)	(5.591)	(5.932)	(7.939)
$\beta_{1,GHCF}$	−0.032	−0.036	−0.054	104.6	110.7	190.43	−111.6	−113.1	−114.7
	(−22.09)	(−21.58)	(−13.64)	(14.78)	(11.60)	(6.140)	(−12.41)	(−11.33)	(−11.64)
$\beta_{2,GHCF}$	4.514	4.522	4.229	0.069	1.627	17.746	−2.055	−2.317	−3.376
	(12.01)	(11.12)	(6.154)	(0.046)	(0.897)	(2.344)	(−9.695)	(−8.117)	(−8.124)
$\beta_{0,LHCF}$	−1.645	−1.728	−2.179	1.685	1.324	−1.262	0.346	0.411	0.842
	(−13.63)	(−15.42)	(−14.01)	(3.038)	(2.594)	(−1.109)	(3.048)	(3.453)	(4.613)
$\beta_{1,LHCF}$	−0.024	−0.027	−0.044	134.6	140.4	202.5	−93.53	−95.33	−97.92
	(−16.34)	(15.74)	(−10.55)	(19.62)	(15.37)	(6.403)	(−11.19)	(10.03)	(4.613)
$\beta_{2,LHCF}$	−1.196	−1.338	−1.895	2.941	4.697	22.808	−3.514	−3.729	−4.788
	(−3.307)	(−3.990)	(−3.097)	(1.970)	(2.550)	(3.011)	(−12.87)	(−12.16)	(−9.971)
$\beta_{0,SHCF}$	−3.254	−3.326	−3.945				0.402	0.509	−4.788
	(−13.31)	(−11.01)	(−8.941)				(3.685)	(4.624)	(−9.249)
$\beta_{1,SHCF}$	−0.036	−0.038	−0.058				−149.6	−153.5	1.069
	(−19.86)	(−20.62)	(−13.99)				(−17.76)	(15.85)	(5.853)
$\beta_{2,SHCF}$	−0.323	−0.476	−1.397				−13.97	−14.787	−18.953
	(−0.667)	(−0.794)	(−1.370)				(−20.11)	(−21.93)	(−19.05)

Table 4.2: Estimation results for the estimates of β_0, β_1, and β_2, i.e. the *mode constants*, the *relative energy service price parameters*, and the *wealth proxy effect* respectively, for a disaggregation scale of 0.1, 1 and 10 PJ under the New and Churn Market Model; (1) Heavy Industry 0.1 PJ disaggregation, (2) Heavy Industry 1 PJ disaggregation, (3) Heavy Industry 10 PJ disaggregation, (4) Passenger Transport Road 0.1 PJ disaggregation, (5) Passenger Transport Road 1 PJ disaggregation, (6) Passenger Transport Road 10 PJ disaggregation, (7) Residential Heating & Cooking 0.1 PJ disaggregation, (8) Residential Heating & Cooking 1 PJ disaggregation, and (9) Residential Heating & Cooking 10 PJ disaggregation.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\beta_{0,GHCF}$	−3.907	−3.965	−4.180	0.003	−0.155	−1.337	0.530	0.570	0.910
	(−11.10)	(−9.870)	(−10.31)	(0.007)	(−0.373)	(−1.582)	(3.305)	(3.708)	(5.159)
$\beta_{1,GHCF}$	−0.036	−0.038	−0.046	113.3	117.0	143.3	−96.36	−98.95	−94.26
	(9.216)	(−8.155)	(−5.777)	(12.92)	(12.622)	(7.092)	(−4.580)	(−5.314)	(−4.246)
$\beta_{2,GHCF}$	5.252	5.231	5.073	0.124	0.628	4.932	−1.473	−1.608	−2.088
	(6.201)	(5.441)	(4.948)	(0.067)	(0.343)	(0.958)	(−4.055)	(−4.294)	(−4.065)
$\beta_{0,LHCF}$	−2.166	−2.246	−2.496	1.043	0.924	0.033	0.181	0.224	0.450
	(−9.080)	(−8.731)	(−6.881)	(2.006)	(1.874)	(0.034)	(1.230)	(1.868)	(2.740)
$\beta_{1,LHCF}$	−0.029	−0.031	−0.039	142.6	146.0	171.2	−82.71	−85.65	−81.95
	(−7.560)	(−6.969)	(−4.813)	(16.86)	(15.51)	(8.385)	(−4.358)	(−5.014)	(−4.168)
$\beta_{2,LHCF}$	−0.354	−0.405	−0.914	3.540	4.228	9.613	−3.217	−3.405	−3.880
	(−0.554)	(−0.437)	(−0.971)	(1.971)	(2.353)	(1.818)	(−6.880)	(−7.716)	(−7.024)
$\beta_{0,SHCF}$	−2.956	−2.932	−3.120				0.062	0.118	0.387
	(−6.781)	(6.232)	(−4.822)				(0.434)	(0.844)	(2.256)
$\beta_{1,SHCF}$	−0.038	−0.039	−0.048				−133.9	−138.3	−144.48
	(−9.352)	(−7.995)	(−6.040)				(−6.686)	(−8.052)	(−6.675)
$\beta_{2,SHCF}$	−1.663	−2.073	−2.766				−12.47	−12.96	−15.30
	(−1.919)	(−2.134)	(−2.107)				(−14.71)	(−15.94)	(−15.30)

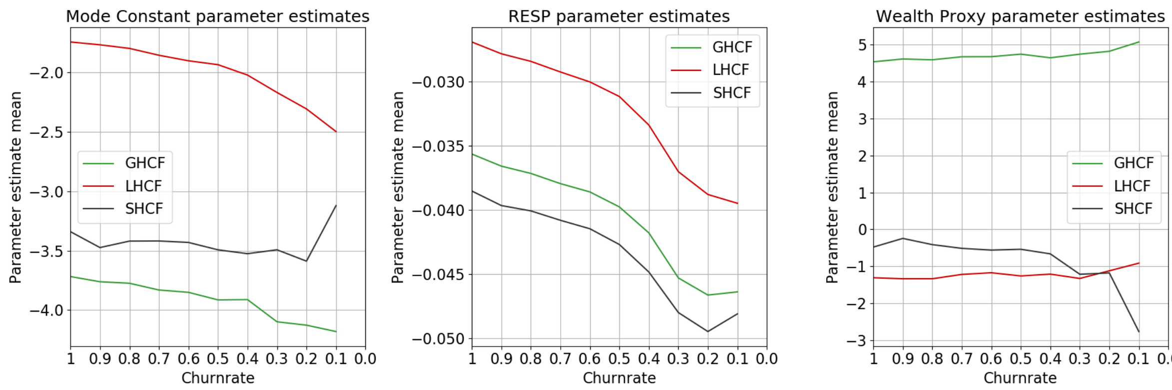

Figure 4.1: Churn rate sensitivity plot for Heavy Industry with parameter estimates for churn rates ranging from 10 to 100% at a disaggregation scale of 10 PJ. GHCF is Gaseous Hydrocarbon Fuels, LHCF is Liquid Hydrocarbon Fuels, and SHCF is Solid Hydrocarbon Fuels, RESP is the relative energy service price.

4.2. Churn Rate Sensitivity

Under the new and churn market model the market considered by the model consists of the annual growth and the churned part of the market from the previous year. This is an approach to model the inertia in the annual market shares for energy carriers.

In order to find a suitable churn rate we need to analyze the impact of different churn rates on the parameter estimates. In order to analyze the effects of the churn rates on the parameter estimates we use churn rate sensitivity plots. These plots consist of parameter estimates and standard deviations for the estimates for churn rates ranging from 10 to 100 %. Based on the results in the churn rate sensitivity plots we choose churn rates for the new and churn market model used to compute the parameter estimates presented in section 4.3.

4.2.1. Sector Group: Industry

Figure 4.1 presents the parameter estimates for Industry. For Industry the mode constants are slightly declining, wealth is rather stable except for Solid Hydrocarbon Fuels (SHCF), and price parameters stabilize below a churn rate of 20%. Based on these results we find that the churn rate range 5% to 20% requires additional analysis. Therefore an additional churn rate sensitivity plot is presented in fig. 4.2 for 5% churn rate to 20% churn rate in steps of 5% with a disaggregation scale of 1 PJ.

In fig. 4.2 it can be observed that Gaseous Hydrocarbon Fuels (GHCF) has a stable mode constant, where Liquid Hydrocarbon Fuels (LHCF) and SHCF show an increasing magnitude and a decreasing magnitude respectively. Both appear to be the result of a change in the estimates for the regional wealth effect presented in the right-hand figure. The change in the parameter estimates for the regional wealth proxy effect for SHCF is especially of interest for the research, since the magnitude for the parameter estimate increases and as such appears to present evidence for a a wealth effect at lower churn rates.

In order to understand what drives the change in the SHCF wealth parameter the size of the feature space is plotted in fig. 4.3. This figure presents the data points in the features space on the y-axis set out against the decreasing churn rate on the x-axis. The reduction of data points may seem like a lot, however the data set spans a period with 4 major crises (1979 energy Crisis, 1987 stock market crash, early 2000’s dotcom bubble, and the financial crisis 2007-2008). The new and churn market model requires 2 years of data to create data points for 1 year due to the annual differences required for the computation. A crisis year is therefore likely to reduce the data set by 2 year-country level data points, this could add up to 200 data points. The data reducing effects of any other year-country level data gaps may be aggravated by the crisis year, therefore the observed reduction in the feature space is reasonable.

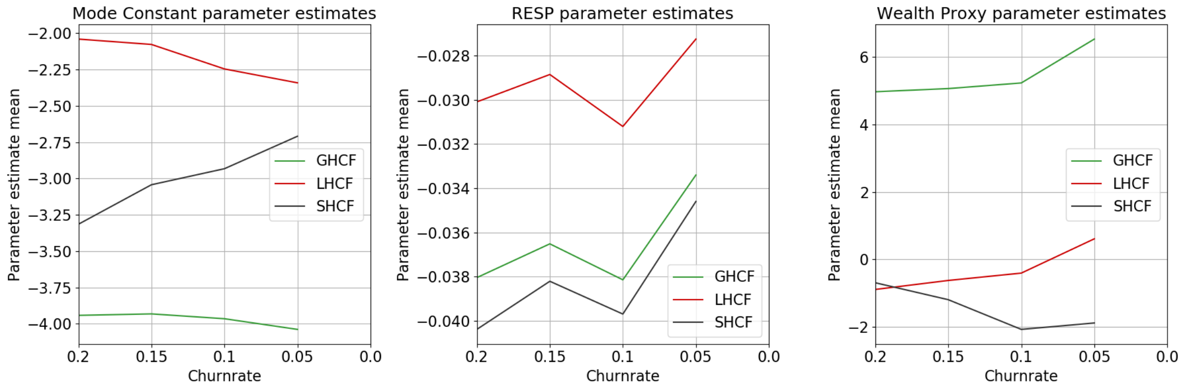

Figure 4.2: Churn rate sensitivity plot for Heavy Industry with parameter estimates for churn rates ranging from 5 to 20% at a disaggregation scale of 1 PJ. GHCF is Gaseous Hydrocarbon Fuels, LHCF is Liquid Hydrocarbon Fuels, and SHCF is Solid Hydrocarbon Fuels, RESP is the relative energy service price.

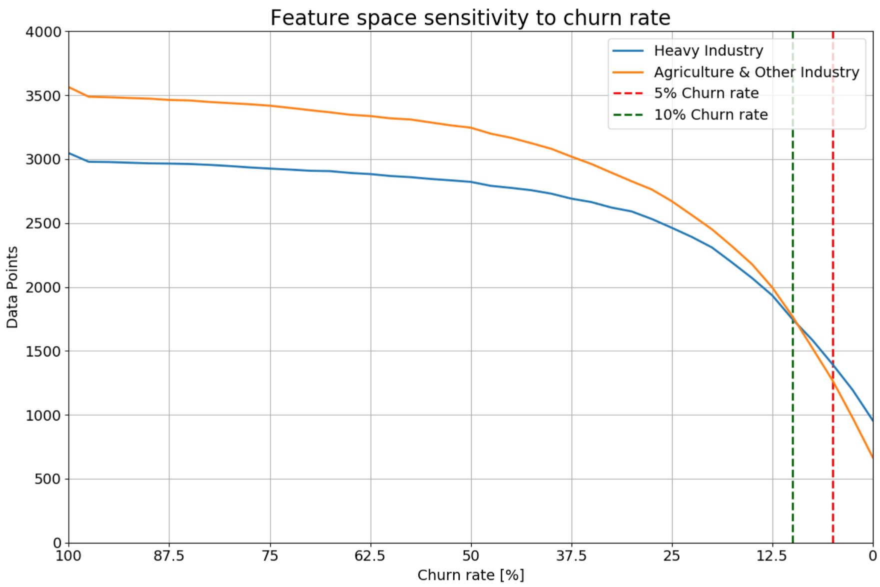

Figure 4.3: Visual representation of sensitivity of the size of the data set to the churn rate under the new and churn market model for the sector group Industry.

In crisis year the global steel consumption drops, SHCF consumption is strongly related to steel consumption. Including such years in the data set is likely to affect the parameter estimates suggesting a positive relation between SHCF consumption and GDP, because in crisis years SHCF consumption and GDP both stagnate or fall. In addition to the crisis years another cause is the market momentum that is included in the new and churn model. Production assets running on SHCF tend to have long asset lifetimes, a high churn rate would therefore likely overestimate the actually churned SHCF market share. This is equivalent to underestimating the changes in the market shares, resulting in parameter estimates with lower magnitudes.

Based on the rather stable relative energy service price parameter (RESP) estimates presented in fig. 4.2, the relatively wealth proxy parameter estimates, and the additional data reduction from 10% to 5% churn rate shown in fig. 4.3, the churn rate is kept conservative at 10 % for the sector group Industry.

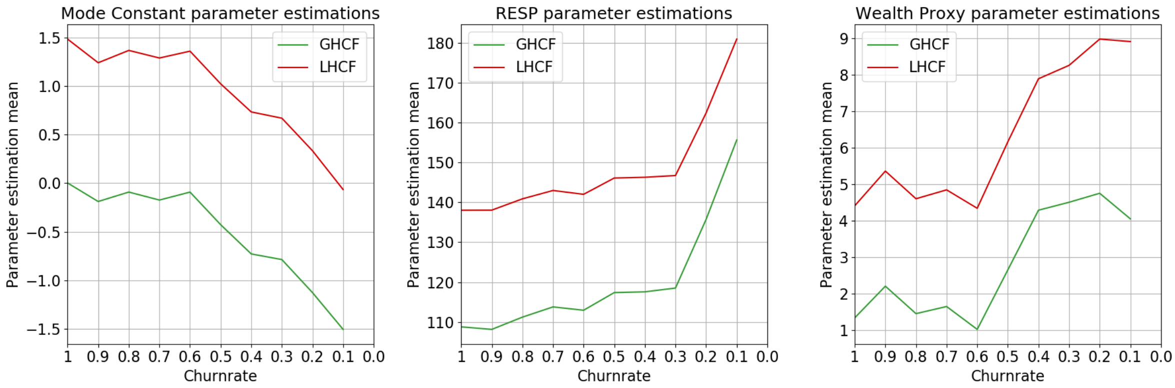

Figure 4.4: Churn rate sensitivity plot for Passenger Transport Road with parameter estimates for churn rates ranging from 10 to 100% at a disaggregation scale of 10 PJ. GHCF is Gaseous Hydrocarbon Fuels, LHCF is Liquid Hydrocarbon Fuels, and SHCF is Solid Hydrocarbon Fuels, RESP is the relative energy service price.

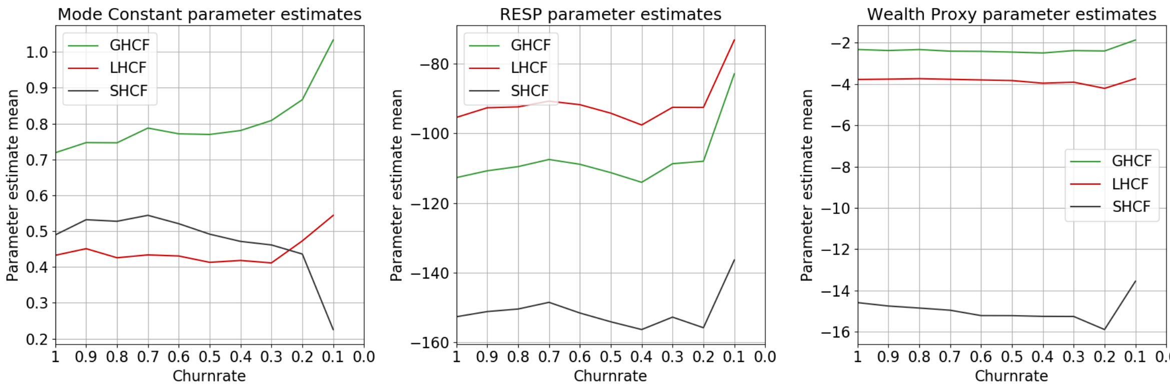

Figure 4.5: Churn rate sensitivity plot for Residential Heating & Cooking with parameter estimates for churn rates ranging from 10 to 100% at a disaggregation scale of 10 PJ. GHCF is Gaseous Hydrocarbon Fuels, LHCF is Liquid Hydrocarbon Fuels, and SHCF is Solid Hydrocarbon Fuels, RESP is the relative energy service price.

4.2.2. Sector Group: Transport

Figure 4.4 presents the parameter estimates for Passenger Transport Road for churn rates of 10 up to a 100 %. Looking at the parameter estimates for the relative energy service price parameter estimates in the middle figure, the parameter estimates suggest that there is a positive relation with price. This should be interpreted as rising relative prices are related to a relative rising consumption. This makes no sense if both alternatives provide the same service, therefore we can say that these estimates make no sense. The reason is discusses later in chapter 5. For now we base the churn rate for the sector group Transport on the time a production asset for the sector Passenger Transport Road, a car, belongs to one end-user. This is approximately 6-7 years [49] [50].

4.2.3. Sector Group: Buildings

Figure 4.5 presents the parameter estimates for Residential Heating & Cooking for churn rates of 10 up to a 100 %. Residential Heating & Cooking shows remarkably stable parameter estimates. There is noticeable change in the parameter estimates for the relative energy service price displayed in the middle figure. To further investigate the influence of the churn rate on the parameter estimates a set of estimates at a higher granularity is presented in fig. 4.6. The parameter estimates in the figure have a churn rate ranging from 5% to 20% and a disaggregation scale of 1PJ.

In fig. 4.6 it is visible that the estimates for 20% and 15% are approximately equal, however reducing the churn rate to 10% reduces the magnitude of the relative energy service price parameter and increases the mode constant estimates. This is contrary to what is to be expected for the new and churn model, a lower churn rate should increase the observed changes in the market. This should in turn

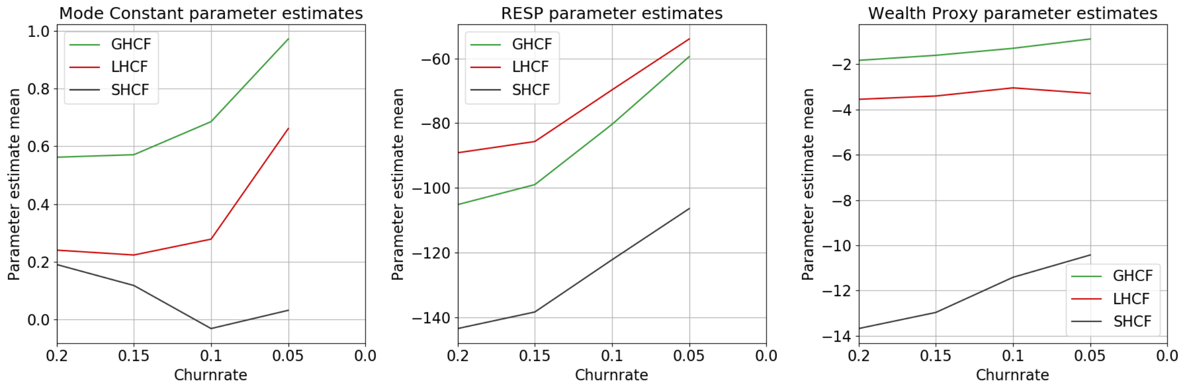

Figure 4.6: Churn rate sensitivity plot for Residential Heating & Cooking with parameter estimates for churn rates ranging from 5 to 20% at a disaggregation scale of 1 PJ. GHCF is Gaseous Hydrocarbon Fuels, LHCF is Liquid Hydrocarbon Fuels, and SHCF is Solid Hydrocarbon Fuels, RESP is the relative energy service price.

increase the magnitudes of the relevant variables.

The comparatively stable parameter estimates and decreasing magnitude of the relative energy service price parameter estimates beyond 15% in combination with the comparatively stable number of data points leads us to choose 15% as the applied churn rate for the new and churn market model in the sector group Buildings.

4.3. Parameter Estimation Results

This section presents the results of the parameter estimations. The section has four subsections, the first three each covering one of the sector groups Industry, Transport, and Buildings, and the final summarizing the results. For each sector the parameter estimates for the Full Market Model (FMM), the New and Churn Market Model (NCMM), and the NCMM without a wealth proxy variable are presented. These three different approaches are presented because they each have a different interpretation. This is because the two different market models have different definitions of the market, and the NCMM without the wealth proxy serves as a sensitivity check. Each of the models contributes to our understanding of the historic relation between the wealth proxy and energy carrier consumption, especially because estimating all three enables us to interpret the differences.

The FMM results can be interpreted as a historic relation between the variables and the energy mix. The NCMM parameter estimates can be interpreted as estimates of the historic relation between the variables and the end-user choice for energy carriers. The difference is subtle yet crucial. The relation with the historic energy mix does not model the market momentum and is therefore a relation between the variables and historic trends in the energy mix. The NCMM parameter estimates more closely reflect the relation with end-user choice, because the NCMM accounts for market momentum.

4.3.1. Sector Group: Industry

Table 4.3 presents the results for Heavy Industry in columns (1), (2), and (3) and Agriculture & Other Industry in columns (4), (5), and (6). All estimates are with respect to the energy carrier electricity, the last paragraph of the section on data processing in chapter 3 describes the concept of a pivot in more detail.

Column (1) in table 4.3 presents the parameter estimates for Heavy Industry under the Full Market Model. The β_0's represent the *mode constants*. Their relative distances (eq. (2.16)) are among the highest in the table, which suggests that the population parameters are different from zero. A constant effect should be interpreted as the historic predisposition of a carrier, which comes on top of the effects of price and/or wealth. The historic predisposition needs to be corrected, because the other estimates would otherwise be susceptible for a spurious regression. The β_1's are all negative as is to be expected [39], these parameters are similar to price elasticity's in the sense that they reflect the effect of price changes on choice probability for energy carriers.

The β_2's are of most interest for the research into the effects of the wealth proxy on the regional energy mix. Considering their relative distances the results for Heavy Industry under the FMM infers that increasing wealth is related to an increasing choice probability for Gaseous Hydrocarbon Fuels (GHCF), and a decreasing choice probability for Liquid Hydrocarbon Fuels (LHCF). For Solid Hydrocarbon Fuels (SHCF) the parameter estimate infers that there is no noticeable effect from wealth under the Full Market Model. It should be noted that the relative distance for $\beta_{2,LHCF}$ is low compared to the relative distance for $\beta_{2,GHCF}$.

Column (2) presents the parameter estimates for Heavy Industry under the New and Churn Market Model with a churn rate of 10%. The first thing that can be noticed is the drop in relative distance for most parameter estimates. The change in relative distance is the result of an increased standard deviation, as is clear due to the similar magnitudes of the estimates with the lower relative distances. The increased standard deviation is as expected because the NCMM focuses on a smaller market, only market churn and growth. The smaller market reduces the number of data points. Still all the change in energy carrier consumption happens within the market, effectively increasing the fluctuations in market shares. The increased fluctuations in the market shares and the reduced number of data points both increase the standard deviation.

The second noteworthy result is the similar relative energy service price parameter estimates. The β_1's in column (2) should be interpreted as the relative energy service price parameter estimates for the new and churn market, i.e. 10% plus the market growth. Even thought their interpretation is slightly different the magnitude cannot be distinguished based on the relative distances. The means that for both models the relative energy service price effect on the market shares is equal, which is contrary to the expectation of higher parameter estimates for the NCMM. The approximately equal parameter estimates for the price effects means that under the NCMM the market would react slower to price shocks.

The most noteworthy change is the shift in the estimates for the β_2's. GHCF choice probabilities are still positively related to the wealth proxy, however LHCF and SHCF have switched places. The smaller magnitude for the parameter estimate for $\beta_{2,LHCF}$ and the reduced relative distance in column (2) no longer provide evidence for a relation of the market share of LHCF with the wealth proxy. The estimate for the relation of regional wealth and the SHCF market share is now negative, this shows in the increased magnitude of the parameter estimate for $\beta_{2,SHCF}$ and the increased relative distance. This means that compared to electricity SHCF will have a reduced choice probability when the wealth proxy increases.

Column (3) presents the parameter estimates for Heavy Industry under the NCMM with a churn rate of 10% without the wealth proxy. This run is executed to validate that the estimates for the price parameter and the wealth parameter are not affected by the potential multicolinearity, because wealth and prices may be correlated. Comparing columns (2) and (3) shows that the price parameter estimates, the β_1's appear to be unaffected by the introduction of the wealth variable. The values for β_0 do change because baseline wealth is now absorbed in the β_2 parameter estimates.

Column (4) in table 4.3 presents the parameter estimates for Agriculture & Other Industry under the Full Market Model. Similar to the Heavy Industry we find β_0's that are different from zero, based on the relative distance values that are comparatively high.

The most noticeable parameter estimate is that of $\beta_{1,LHCF}$, which is positive. This suggests a positive relation between the relative energy service price and the market share for LHCF, i.e. increasing prices drives increasing consumption. This is a spurious result. The other relative energy service price parameters are negative, which is as expected. The difference from the Heavy Industry relative energy service price parameter estimates is in part because the parameter estimates reflect the different measurement unit of the energy service price.

The wealth proxy parameter estimates, β_2's, and their relative distances provide evidence suggesting

Table 4.3: Estimation results for the estimates of β_0, β_1, and β_2, i.e. the *mode constants*, the *relative energy service price* parameters, and the *wealth proxy* parameters respectively, for (1) Heavy Industry - Full Market Model, (2) Heavy Industry - New and Churn Market Model - Churn rate 10%, (3) Heavy Industry - New and Churn Market Model - Churn rate 10% without the wealth variable, (4) Agriculture & Other Industry - Full Market Model, (5) Agriculture & Other Industry - New and Churn Market Model - Churn rate 10%, and (6) Agriculture & Other Industry - New and Churn Market Model - Churn rate 10% without the wealth variable, all at 1 PJ disaggregation.

	(1)	(2)	(3)	(4)	(5)	(6)
$\beta_{0,GHCF}$	-3.607	-3.965	-2.489	-2.843	-2.661	-1.527
	(-20.36)	(-9.889)	(-11.06)	(-19.17)	(-8.727)	(9.056)
$\beta_{1,GHCF}$	-0.032	-0.038	-0.038	-167.2	-121.0	-110.45
	(-22.01)	(-8.229)	(-8.915)	(-13.13)	(-4.336)	(-4.626)
$\beta_{2,GHCF}$	4.514	5.231		3.760	4.645	
	(12.09)	(5.525)		(11.77)	(7.178)	
$\beta_{0,LHCF}$	-1.645	-2.246	-2.410	-0.602	-0.707	-0.654
	(-13.63)	(-8.679)	(-14.16)	(-9.11)	(-6.801)	(-6.603)
$\beta_{1,LHCF}$	-0.024	-0.031	-0.032	70.74	23.31	20.90
	(-16.34)	(-7.027)	(-7.820)	(9.045)	(1.202)	(1.049)
$\beta_{2,LHCF}$	-1.196	-0.405		-1.115	0.194	
	(-3.307)	(-0.521)		(-4.291)	(0.467)	
$\beta_{0,SHCF}$	-3.254	-2.932	-3.421	-1.654	-1.337	-2.148
	(-13.32)	(-6.419)	(-6.650)	(-5.639)	(-3.162)	(-6.590)
$\beta_{1,SHCF}$	-0.036	-0.039	-0.041	-189.5	-137.6	-135.0
	(-19.86)	(-8.004)	(-8.386)	(-11.97)	(-4.855)	(-4.372)
$\beta_{2,SHCF}$	-0.32	-2.073		-7.68	-7.784	
	(-0.667)	(-2.206)		(-9.830)	(-4.852)	

a positive relation with wealth for GHCF market shares, and a negative relation for LHCF and SHCF. These results show different a different relation between wealth and the carrier market shares than found in column (1), especially the SHCF and the relative distance for the wealth parameter estimate.

Column (5) presents parameter estimates for Agriculture & Other Industry under the new and churn market model with a churn rate of 10%. There are two noteworthy difference with the full market model estimates are the relative energy service price parameter estimate and the wealth proxy parameter estimate for liquid hydrocarbon fuels. These are both not convincingly different from zero based on the relative distances.

For the relative energy price parameter estimate this was to be expected. The estimate in column (5) infers no price effects. This may be the situation whenever there is no realistic substitute for the energy carrier, such as with diesel powered agricultural machinery.

Under the NCMM the β_2 estimates and their relative distances for Agriculture & Other Industry suggest a relation between wealth and the energy carrier market shares. In particular the results suggest the wealth proxy to be positively related to GHCF market shares, to not be related to LHCF market shares, and negatively related to SHCF.

Again columns (6) functions as a validation of the estimation results and the relevance of the wealth

Table 4.4: Estimation results for the estimates of β_0, β_1, and β_2, i.e. the *mode constants*, the *relative energy service price* parameters, and the *wealth proxy* parameters respectively, for (1) Passenger Transport Road - Full Market Model, (2) Passenger Transport Road - New and Churn Market Model - Churn rate 15%, and (3) Passenger Transport Road - New and Churn Market Model - Churn rate 15% without the wealth variable, all at 1 PJ disaggregation.

	(1)	(2)	(3)
$\beta_{0,GHCF}$	0.328	−0.155	−0.172
	(0.666)	(−0.372)	(−0.564)
$\beta_{1,GHCF}$	104.6	117.0	119.1
	(14.78)	(12.62)	(10.75)
$\beta_{2,GHCF}$	0.069	0.627	
	(0.045)	(0.343)	
$\beta_{0,LHCF}$	1.685	0.924	1.684
	(3.038)	(1.874)	(5.021)
$\beta_{1,LHCF}$	134.6	146.0	151.5
	(19.62)	(15.51)	(14.67)
$\beta_{2,LHCF}$	2.941	4.228	
	(1.970)	(2.353)	

proxy. The column provides evidence similar to that for the Heavy Industry. That is, the relative price parameter estimates are fairly stable suggesting not a disturbing effect from multicollinearity.

For the sector group Industry the results as presented in table 4.3 provide evidence for a relation between the wealth proxy, and the market shares for energy carriers. In particular, the share of GHCF increases with wealth while the share of SHCF decreases with wealth. There is no convincing evidence that the share of LHCF is affected by wealth.

4.3.2. Sector Group: Transport

Table 4.4 and table 4.5 present the parameter estimation results for Passenger Transport Road, Passenger Transport Rail and Freight Transport Rail.

Column (1) in table 4.4 presents the parameter estimates for the sector Passenger Transport Road under the Full Market Model. The first thing that should be noted are the positive relative energy service price parameters with large relative distances in a similar range to those in the previous subsection, this suggests a positive correlation between price and market shares. This is in contradiction with theory on price elasticity. Columns (2) and (3) show similar results.

These estimates are the result of a violation of requirements of the choice set. The data run from 1978 to 2016, a period where road transport was feasible with LHCF and for some regions to some extent with GHCF. The pivot, Electricity, has not been a feasible energy carrier for passenger transport on the road until lately, and only in wealthy countries. Executing the parameter estimation based on this data set will show no market share for electricity no matter what happens to price or wealth.

Table 4.5 presents the parameter estimates for the sectors Passenger Transport Rail and Freight Transport Rail. For the Passenger Transport Rail the parameter estimates for the full market model in columns (1) show positive price parameters. This is likely the result of the violation of the ability for end-users the choose the energy carriers. The wealth parameter estimates suggest a phase out of hydrocarbon fuels for this sector. The estimates for the wealth are credible, based on their relative distances. Their magnitude suggests a quick phase out of SHCF and a slower phase out of LHCF in the Passenger Transport Rail sector.

The relative energy service price parameters show no evidence for an effect on the energy carrier market shares under the new and churn market model in column (2). This suggests that the energy carrier market shares in the sector Passenger Transport Rail are not affected by the prices of the energy carriers. The parameter estimates for the wealth parameters show similar results under the NCMM as they do under the FMM. These parameter estimates suggest an increasing share of electricity for increasing wealth, whilst LHCF and SHCF will have decreasing market shares.

It should be noted that the results in column (3) confirm the instability of the relative energy service price parameters. Where the price parameter estimates lost any evidence for relevance at the switch from FMM in column (1) to NCMM in column (2), the drop of the wealth variable shows erratic results for the parameter estimates in column (3).

The parameter estimates in column (4) are the results for the sector Freight Transport Rail. The estimates show mode constants, β_0's, that are small in magnitude yet of influence based on the relative distances associated with them. For the price parameter estimates, β_1's, the results combined with the relative distances show that LHCF market shares are negatively related to relative price, whilst the relative distance of the SHCF price parameter show no evidence for a price effect. The wealth parameter estimates, β_2's, in column (4) combined with their relative distances show that the market share of LHCF will increase in increasing wealth, where SHCF will be phased out in increasing wealth.

The estimates for the NCMM in column (5) show results that lead to the same effects. The only noticeable difference is the magnitude of the wealth parameter estimate for SHCF, suggesting a stronger wealth effect. The results in column (6) show results that suggest stability for estimates in columns (4) and (5). That is, the relative price parameter estimates are fairly stable suggesting not a disturbing effect from multicollinearity.

The parameter estimates for the sector group Transport portray some spurious results, most likely as a result of clear violations of model assumptions, in particular freedom of choice. Therefore the estimates for the sector Passenger Transport Road are deemed incredible. The parameter estimates for the rail sectors suggest an effect of the wealth proxy and no effect of relative energy service price. For Passenger Transport Rail wealth is positively related with Electricity compared to LHCF and SHCF. For Freight Transport Rail wealth is positively related with LHCF compared to Electricity and negatively with SHCF compared to Electricity.

Table 4.5: Estimation results for the estimates of β_0, β_1, and β_2, i.e. the *mode constants*, the *relative energy service price* parameters, and the *wealth proxy* parameters respectively, for (1) Passenger Transport Rail - Full Market Model, (2) Passenger Transport Rail - New and Churn Market Model - Churnrate 15%, (3) Passenger Transport Rail - New and Churn Market Model - Churnrate 15% without the wealth variable, for (4) Freight Transport Rail - Full Market Model, (5) Freight Transport Rail - New and Churn Market Model - Churnrate 15%, and (6) Freight Transport Rail - New and Churn Market Model - Churnrate 15% without the wealth variable, all at 1 PJ disaggregation.

	(1)	(2)	(3)	(4)	(5)	(6)
$\beta_{0,LHCF}$	1.152	1.036	−0.872	−0.389	−0.899	1.120
	(11.51)	(5.282)	(−8.904)	(−3.580)	(−5.215)	(6.441)
$\beta_{1,LHCF}$	182.2	138.0	48.37	−157.7	−146.3	−145.9
	(9.921)	(1.486)	(0.950)	(−4.684)	(−2.195)	(−2.494)
$\beta_{2,LHCF}$	−9.181	−9.206		5.417	7.851	
	(−27.92)	(−15.43)		(11.31)	(11.93)	
$\beta_{0,SHCF}$	3.748	3.108	−4.228	1.325	2.695	−0.935
	(7.783)	(3.576)	(−16.14)	(6.666)	(4.065)	(−2.893)
$\beta_{1,SHCF}$	412.9	334.9	117.8	27.15	90.06	0.637
	(9.296)	(0.950)	(5.093)	(0.741)	(0.930)	(0.025)
$\beta_{2,SHCF}$	−250.5	−264.0		−44.70	−156.14	
	(−8.254)	(−4.154)		(−6.134)	(−3.363)	

4.3.3. Sector Group: Buildings

Table 4.6 presents the parameter estimation results for Services and Residential Heating & Cooking.

Columns (1), (2), and (3) present the parameter estimates for the sector Services under the Full Market Model (FMM) , the New and Churn Market Model (NCMM) with a churn rate of 15%, and the New and Churn Market Model without the wealth proxy variable, respectively. The parameter estimates for the relative energy service price parameters, β_1's, are all negative, as expected. The parameter estimates for the wealth proxy, β_2's, suggest an ordered wealth proxy relation for both the FMM and the NCMM, with respect to Electricity the parameter estimates suggest a negative relation of the energy carrier market share for the other energy carriers in descending order; GHCF, LHCF, and lastly SHCF. The relative distances for all parameter estimates in columns (1), (2), and (3), are comperatively large, with the exception for the β_0 esimate for LHCF. These relative distances support the suggestion of effects on the energy carriers' market shares. The relative price parameter estimates are stable between column (2) and column (3) suggesting no disturbing effect from multicollinearity.

Columns (4), (5), and (7) present the parameter estimates for the sector Residential Heating & Cooking under the Full Market Model (FMM) , the New and Churn Market Model (NCMM) with a churn rate of 15%, and the New and Churn Market Model without the wealth proxy variable, respectively. The parameter estimates for the relative energy service price, β_1's, are all negative as expected. Their relative distances suggest a relation between price and the market shares for the energy carriers. The parameter estimates for the wealth proxy, β_2's, and their relative distances suggest an ordered wealth proxy relation for both the FMM and the NCMM. With respect to Electricity the parameter estimates infer a negative relation of the energy carrier market share for the other energy carriers in descending order; GHCF, LHCF, and lastly SHCF.

Column (6) presents the parameter estimates for the sector Residential Heating & Cooking under the New and Churn Market Model (NCMM) with a churn rate of 15% with an additional energy carrier; Traditional Biomass. Traditional Biomass is defined as the non-commercial biomass used in the sector Residential Heating & Cooking. This is biomass that is hand gathered, this means that there is no price data available. It is still a main source of energy for many people, and therefore we wanted to include it in our estimations. In order to include Traditional Biomass in the choice set a shadow price had to be estimated, this price has been set at roughly 80% of the cost price of Solid Hydrocarbon Fuels in 1978. The price is fixed across the entire data set.

This additional element in the choice set offers a suitable test for the independence of irrelevant alternatives property, the other parameter estimates should remain unaffected [39]. Comparing the parameter estimates in columns (5) and (6) parameter estimates are nearly equal in magnitude, in addition the relative distances are nearly unaffected. The magnitude and the relative distances associated with the parameter estimates for Traditional Biomass support the notion of an ordered wealth effect placing Traditional Biomass in the least favoured category with a negative relation with the wealth proxy with the largest magnitude.

For the sector group Buildings the parameter estimates as presented in table 4.6 in columns (3) and (6) suggest an ordered relation of the wealth proxy and the market shares for energy carriers. In this order Electricity is ranked first, followed by GHCF, then LHCF, and SHCF comes in last for Services, whilst the addition of Traditional Biomass shows that Traditional Biomass is the least favoured energy carrier in Residential Heating & Cooking. It should be noted that the magnitude of the relative distances is comparatively large in comparison with previous sector groups.

Table 4.6: Estimation results for the estimates of β_0, β_1, and β_2, i.e. the *mode constants*, the *relative energy service price* parameters, and the *wealth proxy* parameters respectively, for (1) Services - Full Market Model, (2) Services - New and Churn Market Model - Churnrate 15%, (3) Services - New and Churn Market Model - Churnrate 15% without the wealth variable, for (4) Residential Heating & Cooking - Full Market Model, (5) Residential Heating & Cooking - New and Churn Market Model - Churnrate 15%, (6) Residential Heating & Cooking - New and Churn Market Model - Churnrate 15% including Traditional Biomass and (7) Residential Heating & Cooking - New and Churn Market Model - Churnrate 15% without the wealth variable, all at 1 PJ disaggregation.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\beta_{0,GHCF}$	−2.203	−1.921	−2.290	0.604	0.571	0.523	0.283
	(−16.96)	(−9.820)	(−14.66)	(5.591)	(3.777)	(2.510)	(1.834)
$\beta_{1,GHCF}$	−282.0	−237.7	−225.4	−111.6	−98.95	−100.7	−85.66
	(−14.71)	(−7.630)	(−7.317)	(−12.41)	(−5.284)	(−4.385)	(−5.053)
$\beta_{2,GHCF}$	−1.261	−1.490		−2.055	−1.608	−1.626	
	(−4.043)	(−3.802)		(−9.695)	(−4.364)	(−4.588)	
$\beta_{0,LHCF}$	−1.379	−1.448	−2.740	0.346	0.223	0.119	−0.447
	(−13.03)	(−8.281)	(−14.49)	(3.048)	(1.733)	(0.851)	(−3.752)
$\beta_{1,LHCF}$	−277.0	−232.8	−220.5	−93.53	−85.65	−87.49	−71.13
	(−14.74)	(−7.336)	(−7.330)	(−11.19)	(−5.079)	(−4.390)	(−4.966)
$\beta_{2,LHCF}$	−4.943	−4.704		−3.514	−3.405	−3.187	
	(−16.04)	(−9.895)		(−12.87)	(−7.645)	(−7.992)	
$\beta_{0,SHCF}$	−1.430	−1.397	−4.012	−0.402	0.117	−0.079	−1.907
	(−8.284)	(−7.017)	(−13.39)	(3.685)	(0.821)	(−0.488)	(−8.952)
$\beta_{1,SHCF}$	−311.9	−264.7	−240.5	−149.6	−138.2	−146.3	−125.8
	(−15.37)	(−8.930)	(−8.045)	(−17.76)	(−8.038)	(−6.678)	(−7.369)
$\beta_{2,SHCF}$	−14.85	−13.43		−13.96	−12.96	−13.32	
	(−22.87)	(−9.914)		(−20.11)	(−16.05)	(−13.76)	
$\beta_{0,Biom}$						1.537	
						(9.178)	
$\beta_{1,Biom}$						−128.7	
						(−6.221)	
$\beta_{2,Biom}$						−17.00	
						(−22.63)	

4.3.4. Summary of the Results

This subsection provides a clear overview of the qualitative implications of the parameter estimates presented in the previous subsections. The results of are summarized in table 4.7. This table presents the energy carriers ordered by the magnitude of the wealth effect on their respective end-user choice probabilities under the new and churn market model.

In order to facilitate an interpretation of what these effects mean quantitatively we have created a table with an estimation of how the wealth effects hold up against prices changes. In order to do so a comparison between a 10% price increase of the relative energy service price is multiplied by the new and churn market model relative energy service price parameter to compute the effect on the utility of

Table 4.7: Tabular presentation of the regional wealth effect on the energy mix, shifting market shares from lower ranked energy carriers towards higher ranked energy carriers. GCHF is Gaseous Hydrocarbon Fuels, LHCF is Liquid Hydrocarbon Fuels, SHCF is Solid Hydrocarbon Fuels, and Trad. Biomass is Traditional Biomass.

	Industry		Transport			Buildings	
Wealth order	Heavy Industry	Agriculture & Other Industry	Passenger Transport Road	Passenger Transport Rail	Freight Transport Rail	Residential Heating & Cooking	Services
1	GHCF	GHCF		Electricity	LHCF	Electricity	Electricity
2	Electricity (2/3)	Electricity		LHCF	Electricity	GHCF	GHCF
3	LHCF (2/3)	LHCF		SHCF	SHCF	LHCF	LHCF
4	SHCF	SHCF				SHCF	SHCF
5						Trad. Biomass	

a 10% price change of the relative energy service price. Please note that this is the relative price, so the price of the energy carrier with respect to the pivot (in our case electricity). To compute the 10% price change, the median relative energy service price has been used. The resulting change in utility is divided by the wealth parameter estimate in order to compute the wealth change required to undo the effect of the price change on the utility.

Another way to phrase this is computing for what value of wealth changes the following holds:

$$\beta_{1,i} \cdot \Delta x_{1,n,i} = \beta_{2,i} \cdot \Delta x_{2,n} \tag{4.1}$$

In eq. (4.1) $\beta_{1,i}$ is the price parameter estimate, similar to the $\beta_{1,i}$ in eq. (2.4), $\beta_{2,i}$ is the estimated wealth parameter, similar to the $\beta_{2,i}$ in eq. (2.4), and $\Delta x_{1,n,i}$ and $\Delta x_{2,n}$ are the changes in relative energy service price and wealth. The resulting estimates for the values for the wealth change equal to a 10% change in the median relative energy service price are presented in table 4.8. Note that a higher magnitudes means less effect of wealth on the carrier market share with respect to electricity. Negative values mean that this carrier has a higher choice probability in increasing wealth, therefore an increase in price would be countered by a decrease in wealth.

Table 4.8: Wealth changes that are comparable to a 10% increase in the relative energy service price based on the median relative energy service price. GCHF is Gaseous Hydrocarbon Fuels, LHCF is Liquid Hydrocarbon Fuels, SHCF is Solid Hydrocarbon Fuels, and Trad. Biom. is Traditional Biomass.

	Industry		Transport			Buildings	
Carrier	Heavy Industry	Agriculture & Other Industry	Passenger Transport Road	Passenger Transport Rail	Freight Transport Rail	Residential Heating & Cooking	Services
GHCF	≈-18.000	≈-4.000				≈14.000	≈35.000
LHCF	No Wealth Effect	No Wealth Effect		No Price Effect	≈-2.000	≈3.000	≈10.000
SHCF	≈50.000	≈3.500		No Price Effect	No Price Effect	≈2.000	≈3.000
Trad. Biom.						≈1.000	

Table 4.8 provides a better understanding of the magnitude of the wealth effects compared to price changes. The table shows that, especially for the LHCF, the SHCF, and Traditional Biomass, the wealth parameter estimates suggest a significant effect.

5 Discussion

This chapter discusses the results presented in chapter 4 and the applied framework to determine the wealth effects on the energy mix. The chapter has two sections, the first discusses most important results presented in chapter 4 and what can be derived from them. The second section will discuss the limitations to the model and how these may be improved.

5.1. Discussion of the results

Before the discussion of the results for the specific sector groups is presented it is important to revisit the two different market models and how they affect the interpretation of the parameter estimates. For the full market model (FMM) the market momentum is not taken into account. Without taking the market momentum into account a maximum likelihood estimation of the model parameters on the annual energy markets' market shares is estimating parameters for a relation with the historic market shares. Therefore, the parameter estimates are an estimation of the historic relation between the associated variables and the energy carrier market shares in the total annual market for energy carriers.

The new and churn market model (NCMM) considers market momentum and estimate based on a market that is open for end-user choice, therefore the parameter estimates can be interpreted as estimates of the relation between the variables and end-user choice probabilities for the energy carriers. It should be noted that the interpretation for the NCMM can also be phrased as the relation between the variables and the market shares, with the important distinction that the considered market is the new and churn market as defined in eq. (2.14).

5.1.1. Sector group: Industry

In addition to the interpretation of the results for the sector group Industry as given in chapter 4, there are two noticeable results in the parameter estimates. The first, is the slightly lower relative energy service price parameter estimate for Heavy Industry in columns (1) and (2) of table 4.3, and the non-negative and indiscernible from zero estimate for Agriculture & Other Industry as presented in columns (4) and (5) of table 4.3. The second, is the changing order of wealth effects in the switch from the FMM to the NCMM.

The estimate for the relative energy service price parameter can be interpreted as the effect of the relative energy service price on the market share of the energy carrier. A lower magnitude of the parameter estimate for one carrier suggests replacement is less easy than for other carriers. This explains the slightly lower magnitude of the LHCF price parameter estimate for Heavy Industry in columns (1) and (2) in table 4.3. LHCF is widely used in Heavy Industry for processes where electricity is not yet technologically feasible, such as operating heavy machinery.

The magnitude for the parameter estimate for Agriculture & Other industry suggests no effect of price. Under the full market model, column (4) in table 4.3, the estimate is even positive, this means that higher prices and higher market shares go hand in hand. If both prices and market shares rise it can

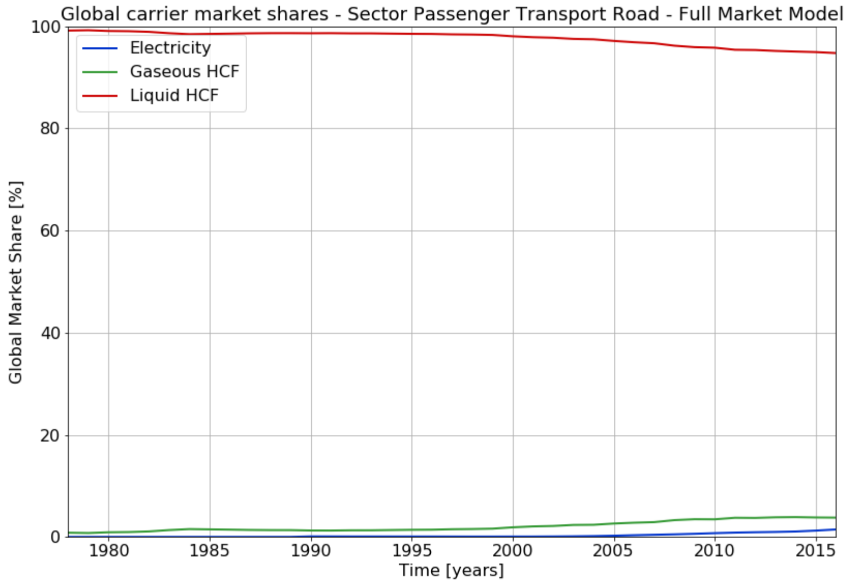

Figure 5.1: Global carrier market shares - sector Passenger Transport Road - Full Market Model.

be interpreted as a lack of substitution alternatives, i.e. there is a lack of competition in the market. For the sector Agriculture & Other Industry this may be true because of the use of LHCF for agricultural machinery used to work the land.

The order of the wealth effects for Heavy Industry on liquid hydrocarbon fuels and solid hydrocarbon fuels flips in the transition from the full market model to the new and churn market model. This is a clear example of a situation where the market momentum and exceptional data, in the from of crisis years, play a role. The slowly declining share of LHCF combined with the SHCF's sensitivity to crisis years leads to an over estimation of the effect of regional wealth on the LHCF market shares and an under estimation of effects on the SHCF market shares.

The market for SHCF in Heavy Industry is closely tied to the steel market, which is closely tied to economic growth. In the data set there are a couple of of crisis years where economic growth slowed or reverted to recession and the market share of solid hydrocarbon fuels dropped. The combination of stagnating wealth and market shares for SHCF invokes parameter estimates with a positive sign. This opposes the negative regional wealth effect on the solid hydrocarbon fuels market in other years across the data set. This problem was solved in the new and churn market model, because it ensured the drop of crisis years.

5.1.2. Sector group: Transport

The transport sector provided some peculiar results for the Passenger Transport Road sector presented in table 4.4. Especially, the price parameter estimates that are positive with large relative distances. Such that they suggest a positive relation between price and energy carrier market shares for LHCF and GHCF compared to electricity.

This is largely attributable to the violation of requirements to the choice set, meaning that electricity was not a realistic alternative for most of the data points. The electricity was not realistic due to a lack of electric vehicles and charging infrastructure for most of the data period. Theoretically this should not necessarily be a problem as the mode constant should absorb any of the constant differences between the carriers. However, as of the last 5 to 10 years the electrical power vehicles are a realistic element of the choice set. This means that there has been a change in the alternatives within the data. Therefore, the mode constants is not entirely able to absorb the limitations to choosing electricity.

Over time the liquid hydrocarbon fuels and gaseous hydrocarbon fuels market shares remained com-

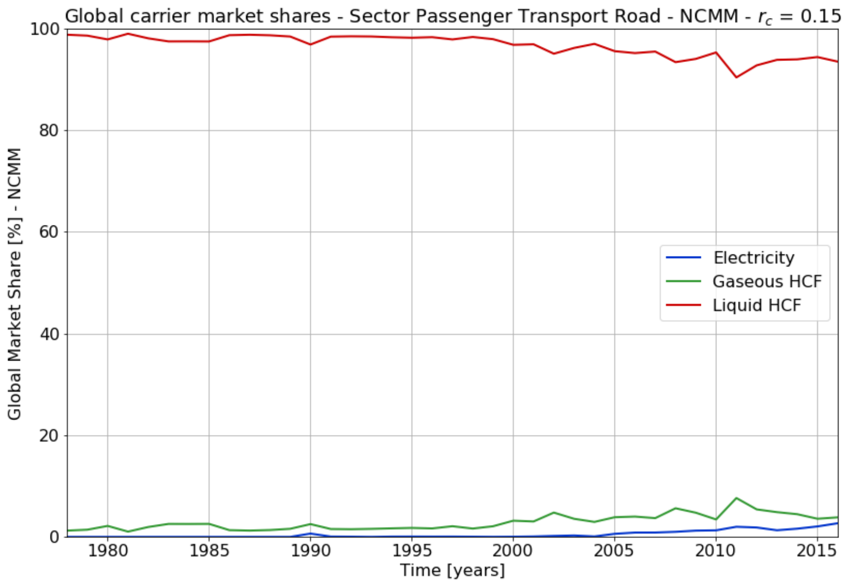

Figure 5.2: Global carrier market shares - sector Passenger Transport Road - New and Churn Market Model - Churn rate = 0.15.

paratively constant, as shown in fig. 5.1 and fig. 5.2. Such that it appeared as a landslide win for the hydrocarbon fuels, and LHCF in particular, from an end-user choice perspective. Combined with rising prices the result this positive price parameter estimates. Repeating the estimations without electricity should provide more knowledge about the development of the market shares for LHCF and GHCF. In order to include Electricity in the choice set, data is required from countries where electricity is a feasible alternative to LHCF and GHCF. Such data may become available in the near future. An alternative solution could be the introduction of a dummy for the availability of electricity to the choice set, however it is difficult to determine what a suitable metric for ”available” would be. So the feasibility of the dummy option is difficult to assess.

For the rail transport sectors the full market model results for the sector Passenger Transport Rail in column (1) in table 4.5 show similar results for the price parameter estimates as observed for the Passenger Transport Road sector. However for the Passenger Transport Rail sector the positive parameter estimates become indiscernible from zero under the new and churn market model, presented in column (2) in table 4.5. This is likely the result of the inability of end-users to choose energy carriers in the sector Passenger Transport Road. The results portray the historic development in the sector, meaning that the sector in more wealthy countries is shifting from SHCF and LHCF towards electricity, independent of energy carriers prices. This shows most clearly in the results for the NCMM in column (2), where the price parameters show to have no discernible effect and all the change in market shares for the energy carriers appears to be related to changes in wealth.

For the sector Freight Transport Rail both columns (4) and (5) in table 4.5, there are estimate suggesting no discernible effect of the energy service price of SHCF on the market share, similar to the Passenger Transport Rail sector. The relative energy service price parameter estimates for LHCF show the expected negative relation.

Compared to the Passenger Transport Rail sector the most noticeable result is the positive parameter estimate for the wealth proxy. This can be explained by increasing international trade for countries experiencing economic growth. Whenever the infrastructure is not yet suitable for trains running on electricity, a diesel powered locomotive is still the solution. Especially for long haul cross border trade this is still a preferred mode of transportation. There are no issue with grid compatibility for diesel powered locomotives, which there are for electrically powered trains, additionally there is no need for creating an electricity network for the sole purpose of rail transport.

5.1.3. Sector group: Buildings

Both sectors Services and Buildings portray estimation results in table 4.6 that are in line with current expectations about high quality carriers and low quality carriers. This means that the parameter estimates suggest that cleaner and more easy to use energy carriers such as Electricity and GHCF show increasing market shares, where LHCF and SHCF show decreasing market shares in relation to increasing wealth. Energy service price effects are comparable for both sectors as their energy service units are equal (MJ heating requirement), suggesting similar price effects.

The most important result is the order in the estimated of the wealth effects. These show a transition away from energy carriers that have the stigma of being of low quality, such as traditional biomass and solid hydrocarbon fuels, towards higher quality carriers such as electricity and gaseous hydrocarbon fuels, as people get more wealthy in line with the research on the local drivers of energy choice in these sectors [31] [32] [33] [34]. From an end-user perspective this makes sense. There are distinct health and practicality issues for heating buildings and cooking based on traditional biomass and SHCF compared to GHCF and electricity.

5.1.4. Drivers of the Wealth Effect

In the previous subsections numerous examples for the peculiar results have been discussed. This shows the complexity of the relation between wealth and the end-user choice probabilities, which indicate the consumption pattern of the countries and regions considered during the research. The wealth parameters can therefore not be interpreted as the effect that the wealth of an individual has on the end-user choice probabilities. Instead the wealth parameters estimated in this piece of research are the aggregate effects, on the end-user choice probabilities for energy carriers within a country or region, of everything wealth enables, such as: policy reform, different industry, different food consumption, different transport, etc. [12].

The mapping discussed in chapter 3 influences the data aggregation of various sectors and the carriers they use. For countries the mapping rules are straightforward, since country borders have remained comparatively stable world wide. However the mapping of the multiple carriers identified by the IEA onto the carriers identified by Shell's WEM, in order to be able to compare prices, reduces granularity. In addition the pricing is determined per sector. This means that mapping decisions, on for instance a sub-sector belonging in Heavy Industry or Agriculture & Other Industry, may have an impact on the parameter estimates. Both carrier mapping and sector mapping may therefore influence the data such that the effects result in higher end-user heterogeneity.

A clear example can be found in the IEA World Energy Outlook [12]. In the sector Heavy Industry processing of different materials such as Iron and Steel, Chemical and Petrochemicals, Non-Ferrous Metals, Non-Metallic Minerals, Paper, Pulp and Print, show different energy carrier consumption patterns. The concentration of these types of industry vary from country to country. Concentration of a specific material processing industry in wealthy countries, will therefore affect the parameter estimates for the wealth effects.

The realization that the relation is the result of the aggregate effects associated with wealth requires additional reflection on the causality of the wealth effects. The research focuses on the effects of wealth on the energy mix:

$$\text{Wealth} \leftrightarrow \text{Energy Mix} \tag{5.1}$$

From our research it cannot be concluded if the arrow is pointing left to right or right to left. Previous research provides evidence causality both ways [5]. Meaning that higher GDP per capita impacts energy consumption, however energy consumption does not directly impact GDP per capita. Additional research will be required to determine the causality of the wealth effects.

The results of the research suggest wealth effects on end-user choice probabilities for the sectors categories Heavy Industry, Agriculture & Other Industry, Passenger Transport Rail, Freight Transport Rail, Residential Heating & Cooking, and Services. Figure 5.3 displays the total final consumption of energy in the world from 1960 up to and including 2018. The energy consumption in fig. 5.3 is segmented in the 14 sector categories as identified by Shell's WEM, such that the figure shows that the

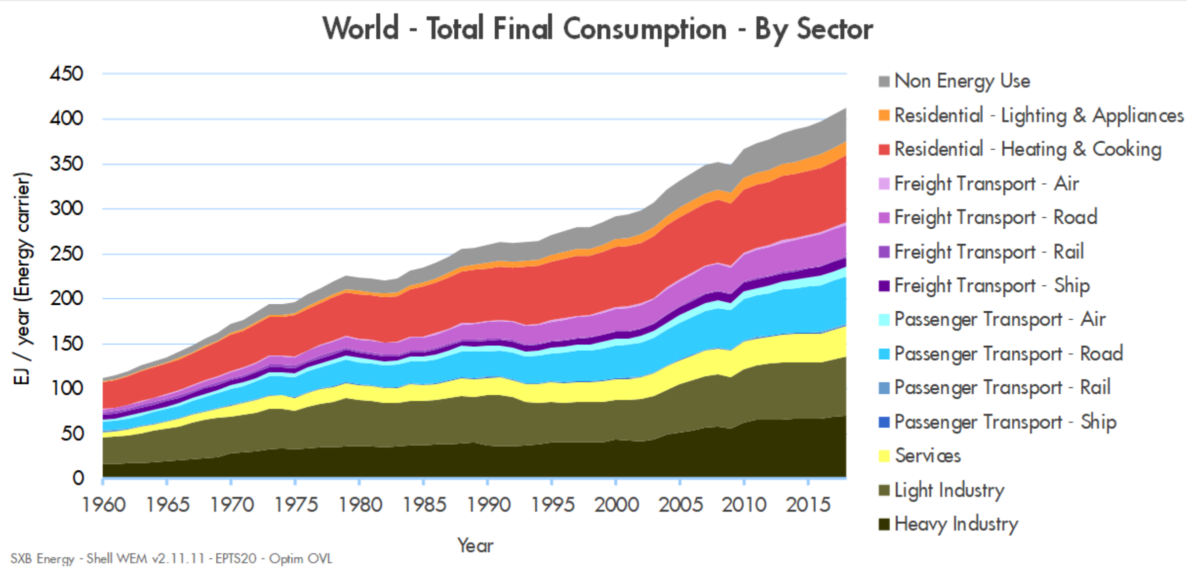

Figure 5.3: The Total Final Consumption of the World segmented in the 14 sector categories as identified by Shell's WEM.

results from the research are relevant for roughly two thirds of the total final consumption of energy in the world. It should be noted that Light Industry and Agriculture & Other Industry represent the same sector.

5.2. End-user choice modelling for the energy sector

In chapter 1 the scale of the energy sector and the challenge that this sector faces has been introduced. The purpose of the research is to expand the knowledge about the energy market and how it takes shape over time in order to facilitate improved modeling for future projections. The first subsections (section 5.2.1, section 5.2.2) discuss conceptual limitations, where later subsections (section 5.2.3, section 5.2.4, section 5.2.5) discuss limitations specific to the applied framework.

5.2.1. Top-down vs. Bottom-up

The first point of discussion is the difference between the top-down and the bottom-up approach. The main difference between the approaches is that the top-down approach tries to model the market as a whole, where a bottom-up model uses data on each individual consumer and models the market by simply summing over the individual consumers. Top down modelling makes assumptions and generalizations in order to model the market with a limited set of markers, reducing the required amount of data. Where the bottom up approach requires consumption data for each individual consumer.

The top-down framework that has been applied to execute the research on the wealth effects resulted in a model that relies heavily on assumptions. These assumptions inevitably result in a modeling bias. When modeling energy carrier market shares, a top-down model is therefore most likely less accurate then a bottom-up model where actual end-user data is used to project future energy carrier market shares from the bottom up. However in absence of sufficient data to coherently compute energy consumption with a bottom-up model, a top-down model can be an excellent solution.

A third option is to create a hybrid model that is capable of processing data in a bottom-up approach of regions with abundant data. In absence of sufficient data the model would revert to the top-down modelling approach. Such an approach allows for the optimal use of the available data. However the approach will not be as consistent as possible, which is a desired property for strategy and policy development that crosses borders.

5.2.2. Revealed Preference

The framework that has been set up to investigate the effects of wealth on the energy mix is based on the concept of revealed preference in a consumer market. The concept of revealed preference relies on the assumption that end-users will choose the alternatives they prefer the most whenever they are presented with a choice. A consequence is that the market presents a reflection of end-user preference for the available alternatives.

The most important drawback of this method is the absence of alternatives in the choice set that may become relevant in the future. Considering an energy transition this is a major drawback for energy system modelling. This limitation is also what caused the peculiar estimates for the sector Passenger Transport Road found in table 4.4. The currently applied framework requires us to either include or exclude an alternative from the choice set. Electricity was not realistically part of the choice set until the last 5 to 10 years of the data set. Including electricity for the entire time span in the choice set suggests for the associated variables electricity is not chosen, resulting in parameter estimates corresponding to not choosing electricity. However electricity cannot realistically be chosen, the spurious results showed that our model was not able to cope with this.

The absence of these choice alternatives may be partially solved by introducing stated preference data [51]. There are models available that incorporate both revealed and stated preference data [39]. This may prove to enable scenario development [11] with an improved foundation. Gathering stated preference data for the entire energy system may however prove to be restrictively cumbersome. If gathering stated preference data proves to be too cumbersome, then modelers will have to resort to inferences based on historic data.

5.2.3. Utility function

In the applied model the utility function is linear in the variables. Linearity is the simplest assumption that can be made. The utility function consists of three variables: the constant, the relative energy service price, and the wealth proxy. The constant term is essentially a term that absorbs relevant yet unobserved variables. In the absence of knowledge about these unobserved variables it is difficult to make and argument for or against linearity for the constant. However for the relative energy service price parameter and the wealth proxy parameter the assumption of linearity may introduce an additional bias because it could be too simplistic.

For the relative energy service price a linear effect on utility is a reasonable assumption for smaller differences. However, larger price differences may have a non linear impact on the utility. At some point additional price difference is not likely to have an impact. In the features spaces presented in appendix E there are some extreme values for the relative energy service price that could be receiving disproportionate weights in the parameter estimates due to the linearity assumption. An example is are the relative energy service prices in the sector Heavy Industry where the 100th percentile value is a factor 10 times that of the median value.

Similarly, for the regional wealth proxy a linear effect on utility is not necessarily an assumption that needs to be true. Solid Hydrocarbon Fuels are likely to remain an essential ingredient for for instance cement, and LHCF are likely to remain a part of rail transport for long-haul distances where infrastructure required for electricity is simply not feasible.

In addition to how utility may depend on the currently assessed variables, there may be other variables of interest. An example would be including capital expenditure. The issue electric vehicles discussed earlier in the sector Passenger Transport Road is a good example. Based on marginal cost of energy, the cost used in the currently applied model, is electricity a cheaper energy carrier for transport than LHCF. However, driving an electric car is still seen as a luxury good by most people because of the initial investment associated with the car. Controlling capital expenditure may therefore improve the modelling capability.

5.2.4. Multinomial Logit

The applied model is a multinomial logit model. Applying a multinomial logit model means making two important assumptions, the first is the existence of a single true population parameter for the variables and the second is the absence of correlation in the error terms.

Assuming the existence of a single true population parameter in our model means that each country reacts in exactly the same way to changes in variables. If this is assumption is not true then it will introduce a bias in the model.

No correlation in the error term means that compared to each other the alternatives in the choice set cannot be grouped based on specific attributes. Whenever there are specific attributes that are distinctly different for alternatives such that they can be grouped within the choice set, then groups of end-users may react to the attribute effectively correlating the error terms that should absorb the variation in the individual preferences of the decision makers. An example for energy carriers would be combustible carriers compared to non-combustibles.

The mixed logit model is able to handle stochastic population parameters and correlated error terms and could therefore be applied to reduce the potential model bias. However, for creating projections the mixed logit model is less favourable then the simpler multinomial logit. The two most important reasons are the additional efforts required for parameter estimation, and the added stochastics in the forecasting process [39].

A logit model by definition assumes that the error term follows a Generalized Extreme Value type 1 error distribution. There are a multitude of other potential distributions. These generally lack the closed form property that the logit model offers. However, these may be more realistic. Examples are probit models which assume normal error term distributions or semi parametric distributions [Paper on semi parametric distributions], this approach does not require additional assumptions about an error term distribution, potentially offering an opportunity to reduce the models bias due to less assumptions.

5.2.5. Disaggregation

The lack of knowledge about the population size required an alternative strategy for the uncertainty quantification. Because p-values and z-scores depend on the knowledge of the sample size in their computation. To that end we used a relative distance measure, defined as the mean and standard deviation of bootstrap results in eq. (3.7).

It should be mentioned that another approach would be possible, one could perform the similar parameter estimations based solely on the market shares whilst using the market sizes as weights in the maximum likelihood estimation without disaggregation:

$$\log L(\boldsymbol{\beta}_S) = \sum_t \sum_r \sum_n \sum_i \left(\frac{TFC_{r,S,n,i}}{\sum_t \sum_r \sum_n \sum_j TFC_{r,S,n,j}} \cdot log \left(\frac{e^{V_{r,S,n,i}}}{\sum_j e^{V_{r,S,n,j}}} \right) \right). \tag{5.2}$$

Such an approach would require an other software package, or writing a maximum likelihood estimator for this function by hand. It could solve the problem of including different weights for the different market sizes and the incompatibility of the preferred package with market shares, mentioned in chapter 3.

However, this approach would not solve the unknown real population size and sample size, and the interpretation of the results would be slightly different. A maximum likelihood estimation of the population parameters based on the market shares, FMM or NCMM, should be interpreted as the relation between the variable and the market shares for these respective to the market. Our approach under the NCMM can be interpreted as the estimation of the relation between a variable and the historic end-user choice probabilities.

5.2.6. Market Model

The data that we are dealing with is annual market data. This is a time series, for market time series generally display auto-correlation or market momentum. The multinomial logit model is not equipped to deal with the market momentum by itself. Therefore the new and churn market model is added to the modelling, this approach should effectively factor in market momentum.

Under the full market model the data points are used as one batch. This is the result of the wealth hypothesis, in order to test this hypothesis we assume that countries follow a similar trajectory in their energy mix as their wealth grows. By using the data points as one big batch we implicitly assume that the error terms are independently and identically distributed. This would only be true if the error terms show no auto-correlation, or market momentum. This means that the model requires accurate churn rate estimations. Determining the optimal churn rates for each sector may be a research topic by itself.

6

Conclusions and Recommendations

In this chapter we summarize and conclude. The chapter has two sections, the first provides the conclusions to the research and the second provides recommendations for further research.

One of the models used to quantify projections on the potential outcomes a future energy system is Shell's World Energy Model (WEM). The WEM quantifies potential outcomes of a future global energy system based on input variables corresponding to a scenario developed by the scenarios team in the strategy department at Shell. For the demand side modeling in the WEM the End-User Choice module incorporates end-user choice modelling concerning energy carriers. Over the last couple of years, a discrepancy can be identified between reality and the model output given realized input variables. We developed the hypothesis that this discrepancy may be the result of omitted variable bias. As such the introduction of a relevant variable should reduce the discrepancy. In particular, we expected that increasing wealth influences not only quantity of energy that is consumed, but also how it is consumed by end-users. This means that wealth should be related to changes in the energy mix.

6.1. Conclusions

Through applying the end-user choice framework developed in chapter 2 and chapter 3 we find effects of wealth on the energy mix. We found an ordered effect of wealth on the energy mix in the sector groups: Industry, Rail Transport, and Buildings.

These findings are based on the parameter estimates for the wealth proxy variable, relating wealth to energy carrier market shares. The parameter estimates presented are in table 4.3, table 4.5, and table 4.6. These parameter estimates are estimations of the population parameters for the utility function in the end-user choice framework. Whenever the parameter estimates showed sufficient indication of an effect in the utility function this was interpreted as a suggestion for an effect of the associated variable on the energy mix.

The results of the research are summarized in table 6.1, which presents the relevant energy carriers, ordered by the magnitude of the wealth effect on their respective end-user choice probabilities under the new and churn market model. As we expected, cleaner energy solutions such as electricity and gaseous fuels are on top and solid fuels and traditional biomass are located at the bottom.

Table 6.1: Tabular presentation of the regional wealth effect on the energy mix, shifting market shares from lower ranked energy carriers towards higher ranked energy carriers. GCHF is Gaseous Hydrocarbon Fuels, LHCF is Liquid Hydrocarbon Fuels, SHCF is Solid Hydrocarbon Fuels, and Trad. Biomass is Traditional Biomass.

	Industry		Rail Transport		Buildings	
Wealth order	Heavy Industry	Agriculture & Other Industry	Passenger Transport Rail	Freight Transport Rail	Residential Heating & Cooking	Services
1	GHCF	GHCF	Electricity	LHCF	Electricity	Electricity
2	Electricity (2/3)	Electricity	LHCF	Electricity	GHCF	GHCF
3	LHCF (2/3)	LHCF	SHCF	SHCF	LHCF	LHCF
4	SHCF	SHCF			SHCF	SHCF
5					Trad. Biomass	

6.2. Recommendations

In chapter 5 several limitations of the applied model and some of their effects observed in the results have been discussed. These limitations lead to the following recommendations for improving the experiment model improvement.

The first is variables in the utility function; testing for additional variables and non-linear relations may provide an area of further model improvement.

The second is the functional form, a critical assumption is the one about the distribution of the error term in the utility function. The assumption of the GEV type 1 distribution is convenient, but other distribution may provide a better fit. Alternative functional forms such as the multinomial probit are less convenient, however they may prove to be a better fit for the modelling of the energy sector.

In addition to the assumption about the shape of error term distribution the multinomial logit model assumes no correlation among the error terms and deterministic population parameters for the variables. The mixed logit model may provide a better suited alternative to the multinomial logit model, since is it able to cope with the correlation of the error terms and with stochastic population parameters.

The third area for further research is churn rates. The new and churn market model is able to mimic market momentum, as long as the correct churn rate is implemented. For that, literature research or a survey study is advised.

The last recommendation for further research delves into the underlying causes for the relation of wealth and the energy mix. Our research displays a relation between wealth and end-user choice probabilities for energy carriers, however the question of what the drivers underlying this relation are remains unanswered. Better understanding of where this relation comes from may help in the decoupling of global GDP growth and green house gas emissions. A potential start is higher granularity modeling, by creating sub-sectors for the distinct end-user groups within the sectors used for this research.

Bibliography

[1] IEA. *World Energy Investment 2019*. International Energy Agency, 2019.

[2] IEA. *Key World Energy Statistics (KWES) 2019*. International Energy Agency, 2019.

[3] IEA. *Global Energy & CO2 Status Report 2019*. International Energy Agency, 2019.

[4] Muhammad Azhar Khan, Muhammad Zahir Khan, Khalid Zaman, and Lubna Naz. Global estimates of energy consumption and greenhouse gas emissions. *Renewable and Sustainable Energy Reviews*, 29:336 – 344, 2014.

[5] Sectoral trends and driving forces of global energy use and greenhouse gas emissions. *Mitigation and Adaptation Strategies for Global Change*, 3:263–319, 1998.

[6] United Nations. Kyoto protocol to the united nations framework convention on climate change, 1998.

[7] United Nations. Paris agreement, 2015.

[8] The environmental kuznets curve: an empirical analysis. *Environment and Development Economics*, 2(4):401–416, 1997.

[9] Edward B. Barbier David I. Stern, Michael S. Common. Economic growth and environmental degradation: The environmental kuznets curve and sustainable development. *World Development*, 24(7):1151–1160, 1996.

[10] Peter Lopion, Peter Markewitz, Martin Robinius, and Detlef Stolten. A review of current challenges and trends in energy systems modeling. *Renewable and Sustainable Energy Reviews*, 96:156 – 166, 2018.

[11] Stefan Pfenninger, Adam Hawkes, and James Keirstead. Energy systems modeling for twenty-first century energy challenges. *Renewable and Sustainable Energy Reviews*, 33:74 – 86, 2014.

[12] IEA. *World Energy Outlook 2019*. International Energy Agency, 2019.

[13] United Nations. World economic situation and prospects, 2020.

[14] Intergovernmental Panel on Climate Change. *Climate Change 2007: The Physical Science Basis*. Cambridge University Press, 2007.

[15] International Energy Agency. World energy model - scenario analysis of future energy trends, 2019.

[16] Imperial College London & Sustainable Gas Institute. Modular energy systems simulation environment (muse), 2020.

[17] Massachusetts Institute of Technology. Sustainable energy system analysis modelling environment (sesame), 2020.

[18] United Nations Statistics Division World Bank Group International Energy Agency, International Renewable Energy Agency and World Health Organization. 2019 tracking sdg7 - the energy progress report, 2019.

[19] McKinsey & Company. Global energy perspective, 2019.

[20] BP Energy. World energy outlook, 2019.

[21] ExxonMobil. World energy outlook, 2020.

[22] Royal Dutch Shell. World energy model, 2020.

[23] IEA. *World Energy Balances 2019*. International Energy Agency, 2019.

[24] C.O. Wene. Energy-economy analysis: Linking the macroeconomic and systems engineering approaches. *Energy*, 21(9):809 – 824, 1996.

[25] Chris Bataille Jean-Charles Hourcade, Mark Jaccard and Frederic Ghersi. Hybrid modeling: New answers to old challenges introduction to the special issue of the energy journal. *The Energy Journal*, Hybrid Modeling(Special Issue 2):1–12, 2006.

[26] Nic Rivers and Mark Jaccard. Combining top-down and bottom-up approaches to energy-economy modeling using discrete choice methods. *The Energy Journal*, 26(1):83–106, 2005.

[27] P. Geoffrey Allen and Robert Fildes. *Econometric Forecasting*, pages 303–362. Springer US, 2001.

[28] Ruud Weijermars, Peter Taylor, Olivier Bahn, Subir Ranjan Das, and Yi-Ming Wei. Review of models and actors in energy mix optimization - can leader visions and decisions align with optimum model strategies for our future energy systems? *Energy Strategy Reviews*, 1(1):5 – 18, 2012. Energy System Models Natural Gas Strategies.

[29] James H Stock and Mark W Watson. *Introduction to econometrics*. Pearson Education Limited, 2015.

[30] Ugur Soytas and Ramazan Sari. Energy consumption and gdp: causality relationship in g-7 countries and emerging markets. *Energy Economics*, 25(1):33 – 37, 2003.

[31] Dil Bahadur Rahut, Akhter Ali, Khondoker Abdul Mottaleb, and Jeetendra Prakash Aryal. Wealth, education and cooking-fuel choices among rural households in pakistan. *Energy Strategy Reviews*, 24:236 – 243, 2019.

[32] M. Narasimha Rao and B. Sudhakara Reddy. Variations in energy use by indian households: An analysis of micro level data. *Energy*, 32(2):143 – 153, 2007.

[33] Justice Tei Mensah and George Adu. An empirical analysis of household energy choice in ghana. *Renewable and Sustainable Energy Reviews*, 51:1402 – 1411, 2015.

[34] Bing-Dong Hou, Xin Tang, Chunbo Ma, Li Liu, Yi-Ming Wei, and Hua Liao. Cooking fuel choice in rural china: results from microdata. *Journal of Cleaner Production*, 142:538 – 547, 2017.

[35] Empirical link between economic growth, energy consumption and co2 emission in australia. *The Journal of Developing Areas*, 47(2):81–92, 2013.

[36] P.A. Samuelson. A note on the pure theory of consumer's behaviour. *Economica*, 5:61–71, 1938.

[37] John Muellbauer. Household production theory, quality, and the "hedonic technique.". *American Economic Review*, 64(6):977–94, 1974.

[38] Sherwin Rosen. Hedonic prices and implicit markets: Product differentiation in pure competition. *Journal of Political Economy*, 82(1):34–55, 1974.

[39] K. E. Train. *Discrete Choice Methods with Simulation*. Cambridge University Press, 2003.

[40] Steven Berry, James Levinsohn, and Ariel Pakes. Automobile prices in market equilibrium. *Econometrica*, 63(4):841–890, 1995.

[41] Kelvin J. Lancaster. A new approach to consumer theory. *Journal of Political Economy*, 74(2):132–157, 1966.

[42] Carol S. Carson. The history of the united states national income and product accounts: The development of an analytical tool. *Review of Income and Wealth*, 21(2):153–181, 1975.

[43] Martin Feldstein. Underestimating the real growth of gdp, personal income, and productivity. *Journal of Economic Perspectives*, 31(2):145–64, May 2017.

[44] Sheiner Lousie Dynan, Karen. Gdp as a measure of economic well-being, 1975.

[45] R.D. Luce. *Individual Choice Behavior: A Theoretical Analysis*. Wiley, 1959.

[46] Daniel McFadden. Modelling the choice of residential location. Cowles Foundation Discussion Papers 477, Cowles Foundation for Research in Economics, Yale University, 1977.

[47] D. McFadden. The choice theory approach to market research. *Marketing Science*, 5:275–297, 1986.

[48] Lerman S. R. Ben-Akiva, M. *Discrete Choice Analysis*. MIT Press, 1985.

[49] Statista. Average length of vehicle ownership in the united states in 2006 and 2016, by vehicle type, 2020.

[50] IHS Markit. Average length of new vehicle ownership in the united, 2020.

[51] Mark Wardman. A comparison of revealed preference and stated preference models of travel behaviour. *Journal of Transport Economics and Policy*, 22(1):71–91, 1988.

A

World Energy Model Countries & Regions

Table A.1: The regions identified as unique geographical areas within a region in Shell's World Energy Model.

USA	Netherlands	Hungary	Peru
China	Pakistan	Norway	Mozambique
Russia	Argentina	Portugal	Ecuador
India	Malaysia	(Kuwait)	Angola
Japan	Egypt	Bangladesh	Yemen
Germany	Venezuela	Ethiopia	Luxembourg
France	Kazakhstan	North Korea	International marine bunkers
Canada	Belgium	Israel	Rest of Europe West Other
United Kingdom	Vietnam	Denmark	Rest of Europe East Other
Brazil	Sweden	Tanzania	Rest of EU New 13
South Korea	Uzbekistan	Bulgaria	Baltic States
Italy	(United Arab Emirates)	Syria	Rest of Central Asia
Indonesia	Czech Republic	Slovakia	Rest of East Asia
Mexico	Philippines	(Qatar)	Rest of SE Asia
Iran	Romania	Kenya	Rest of South Asia
Saudi Arabia	Finland	Libya	Rest of Middle East
Spain	Algeria	Sudan	Rest of Arabian Peninsula
Ukraine	Austria	New Zealand	Rest of North Africa
South Africa	Iraq	DR Congo	Rest of East Africa
Australia	Greece	Turkmenistan	Rest of Southern Africa
Taiwan	Singapore	Myanmar	Rest of West Africa
Nigeria	Colombia	Ireland	Rest of North America
Thailand	Chile	Oman	Rest of Central America & Caribbean
Poland	Belarus	Azerbaijan	Rest of South America
Turkey	Switzerland	Morocco	Rest of Oceania

B

World Energy Model Sectors & Energy Carriers

Table B.1: The sectors identified as unique markets within a region in Shell's World Energy Model.

Sectors
Heavy Industry
Agriculture & Other Industry
Services
Passenger Transport - Ship
Passenger Transport - Rail
Passenger Transport - Road
Passenger Transport - Air
Freight Transport - Ship
Freight Transport - Rail
Freight Transport - Road
Freight Transport - Air
Residential Heating & Cooking
Residential Lighting & Appliances
Non Energy Use

Table B.2: The energy carriers identified as distinct energy carriers in Shell's World Energy Model.

Energy Carriers
Solid Hydrocarbon Fuels (SHCF)
Liquid Hydrocarbon Fuels (LHCF)
Gaseous Hydrocarbon Fuels (GHCF)
Electricity - Commercial
Electricity - Solar PV Distributed
Hydrogen
Heat - Commercial
Heat - Solar Thermal Distributed
Biomass - Commercial
Biomass - Traditional

C

Multinomial Logit Derivation

The multinomial logit derivation is based on the assumption that the error term of the utility has a Generalized Extreme Value type 1 distribution.

The density for the error term:

$$f(\epsilon_{n,j}) = e^{-\epsilon_{n,j}} e^{-e^{-\epsilon_{n,j}}}, \tag{C.1}$$

Such that the cumulative distribution is:

$$F(\epsilon_{n,j}) = e^{-e^{-\epsilon_{n,j}}}. \tag{C.2}$$

Assuming independence of irrelevant alternatives and identical independent draws from the eq. (C.1) then $\epsilon_{n,j,i} = \epsilon_{n,j} - \epsilon_{n,i}$ is logistically distributed. Such that:

$$F(\epsilon_{n,j,i}) = \frac{e^{n,j,i}}{1 + e^{n,j,i}}. \tag{C.3}$$

The probability alternative i is chosen by decision maker n is:

$$\begin{aligned} P_{n,i} &= Prob(V_{n,i} + \epsilon_{n,i} > V_{n,j} + \epsilon_{n,i} \qquad \forall j \neq i) \\ &= Prob(\epsilon_{n,j} > V_{n,i} - V_{n,j} + \epsilon_{n,i} \qquad \forall j \neq i) \end{aligned} \tag{C.4}$$

Assuming the independence of ϵ the probabilities are given by integrating over $\epsilon_{n,i}$.

$$\begin{aligned} P_{n,i} &= \int_{\epsilon_{n,i}=-\infty}^{\infty} \left(\prod_{j \neq i} e^{-e^{-(V_{n,i}-V_{n,j}+\epsilon_{n,i})}} \right) e^{-\epsilon_{n,i}} e^{-e^{-\epsilon_{n,i}}} d\epsilon_{n,i} \\ &= \int_{\epsilon_{n,i}=-\infty}^{\infty} \left(\prod_{j} e^{-e^{-(V_{n,i}-V_{n,j}+\epsilon_{n,i})}} \right) e^{-\epsilon_{n,i}} d\epsilon_{n,i} \\ &= \int_{\epsilon_{n,i}=-\infty}^{\infty} \exp\left(-\sum_{j} e^{-(V_{n,i}-V_{n,j}+\epsilon_{n,i})} \right) e^{-\epsilon_{n,i}} d\epsilon_{n,i} \\ &= \int_{\epsilon_{n,i}=-\infty}^{\infty} \exp\left(-e^{-\epsilon_{n,i}} \sum_{j} e^{-(V_{n,i}-V_{n,j})} \right) e^{-\epsilon_{n,i}} d\epsilon_{n,i} \end{aligned} \tag{C.5}$$

As $\epsilon_{n,i}$ approaches ∞, $e^{-\epsilon_{n,i}}$ approaches 0, and as $\epsilon_{n,i}$ approaches 0, $e^{-\epsilon_{n,i}}$ approaches $-\infty$, using this:

$$\begin{aligned}
P_{n,i} &= \int_{\epsilon_{n,i}=-\infty}^{\infty} \exp\left(-e^{-\epsilon_{n,i}} \sum_j e^{-(V_{n,i}-V_{n,j})}\right) e^{-\epsilon_{n,i}} d\epsilon_{n,i} \\
&= \int_{\infty}^{0} \exp\left(-e^{-\epsilon_{n,i}} \sum_j e^{-(V_{n,i}-V_{n,j})}\right) e^{-\epsilon_{n,i}} d\epsilon_{n,i} \\
&= \int_{0}^{\infty} \exp\left(-e^{-\epsilon_{n,i}} \sum_j e^{-(V_{n,i}-V_{n,j})}\right) - e^{-\epsilon_{n,i}} d\epsilon_{n,i} \\
&= \frac{\exp(-e^{-\epsilon_{n,i}} \sum_j e^{-(V_{n,i}-V_{n,j})})}{-\sum_j e^{-(V_{n,i}-V_{n,j})}} \Bigg|_0^{\infty} \\
&= \frac{1}{\sum_j e^{-(V_{n,i}-V_{n,j})}} \\
&= \frac{e^{V_{n,i}}}{\sum_j e^{V_{n,j}}}
\end{aligned} \tag{C.6}$$

D
Parameter Estimate Distribution Plots

For all figures in this appendix the following abbreviations have been used:

- GHCF - Gaseous Hydrocarbon Fuels
- LHCF - Liquid Hydrocarbon Fuels
- SHCF - Solid Hydrocarbon Fuels

The following attachments can be seen in the figures:

- Const - Referring to the mode constant parameter estimate β_0
- ESP - Referring to the relative energy service price parameter estimate β_1
- Wealth - Referring to the wealth parameter estimate β_2

Combinations of the above mentioned abbreviations are used in the titles and x-labels for the figures. The y-labels show the ”occurrence in bootstrap” as label, this refers to the density.

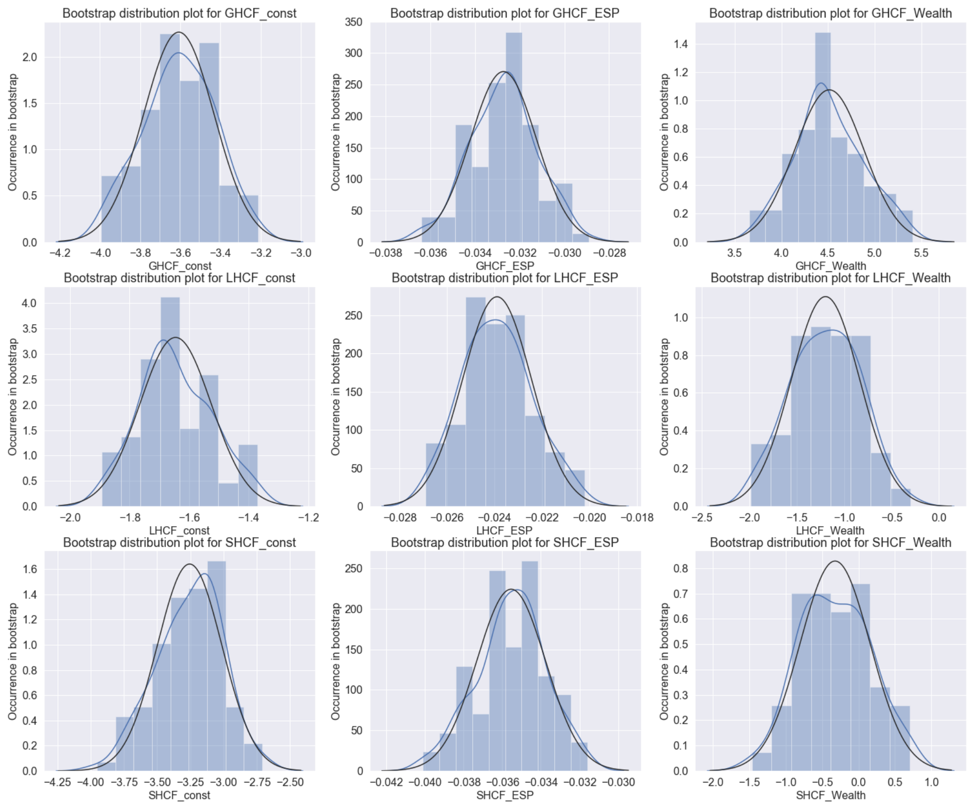

Figure D.1: Parameter Estimate Distribution Plots Heavy Industry Full Market Model

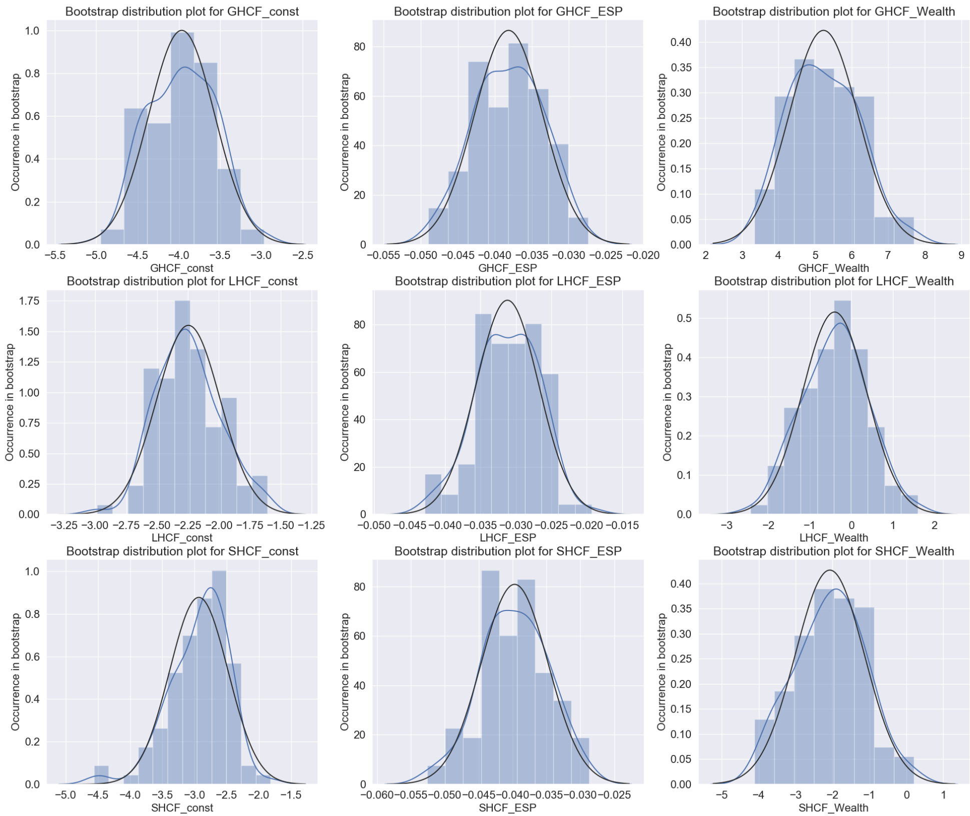

Figure D.2: Parameter Estimate Distribution Plots Heavy Industry New and Churn Market Model

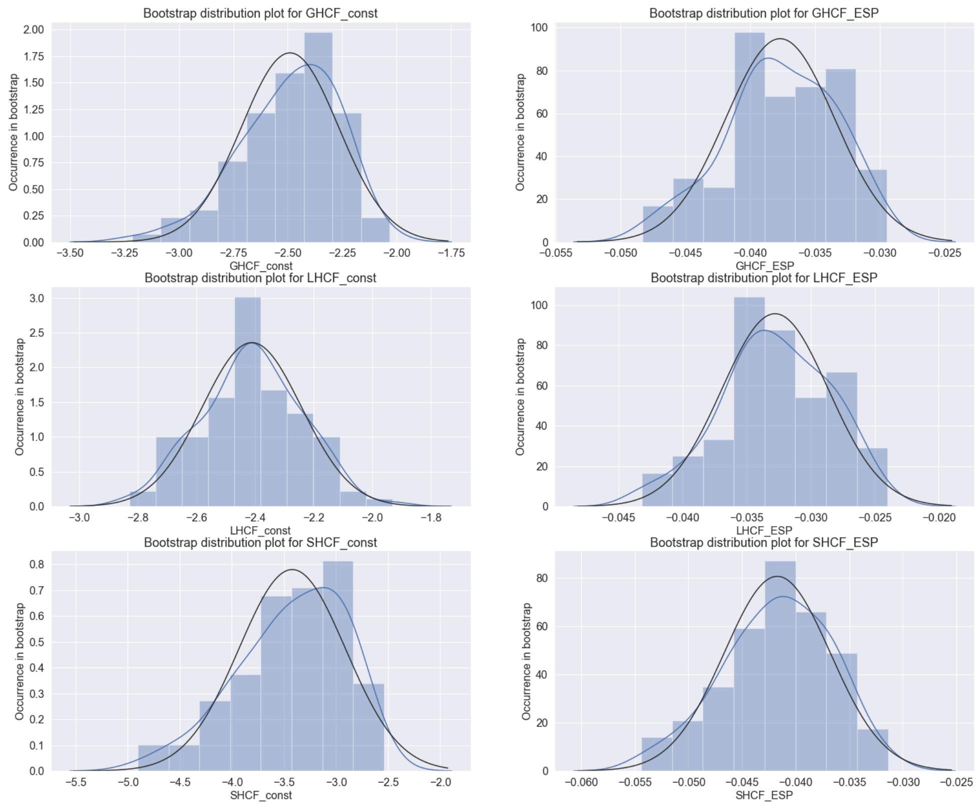

Figure D.3: Parameter Estimate Distribution Plots Heavy Industry New and Churn Market Model without the wealth variable

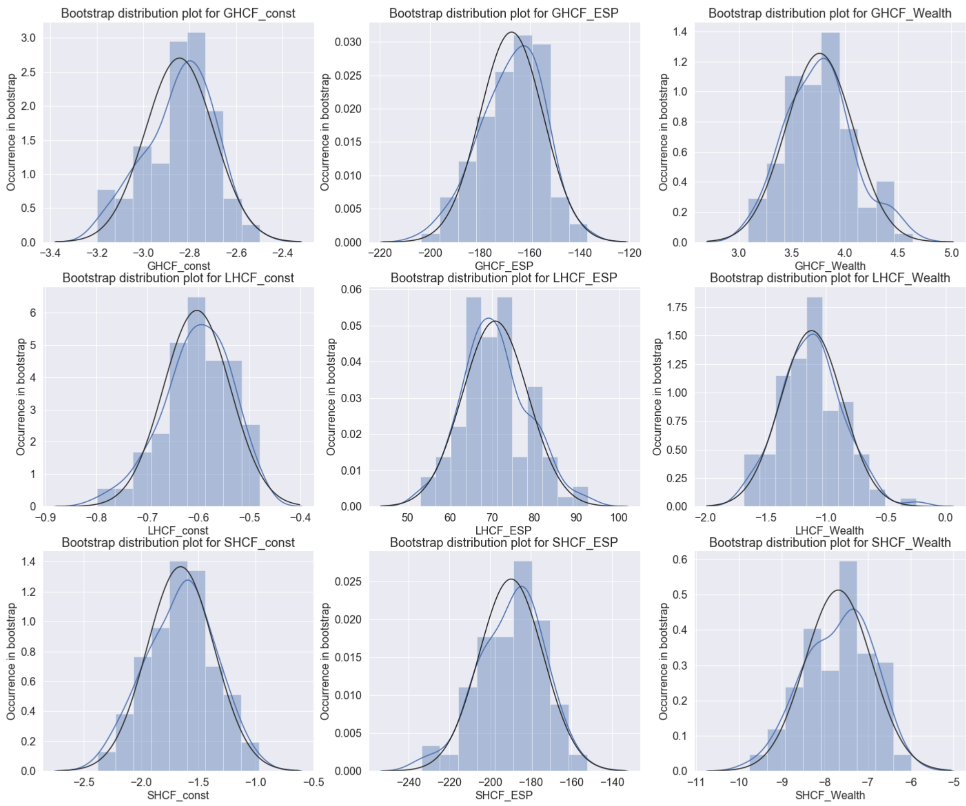

Figure D.4: Parameter Estimate Distribution Plots Agriculture & Other Industry Full Market Model

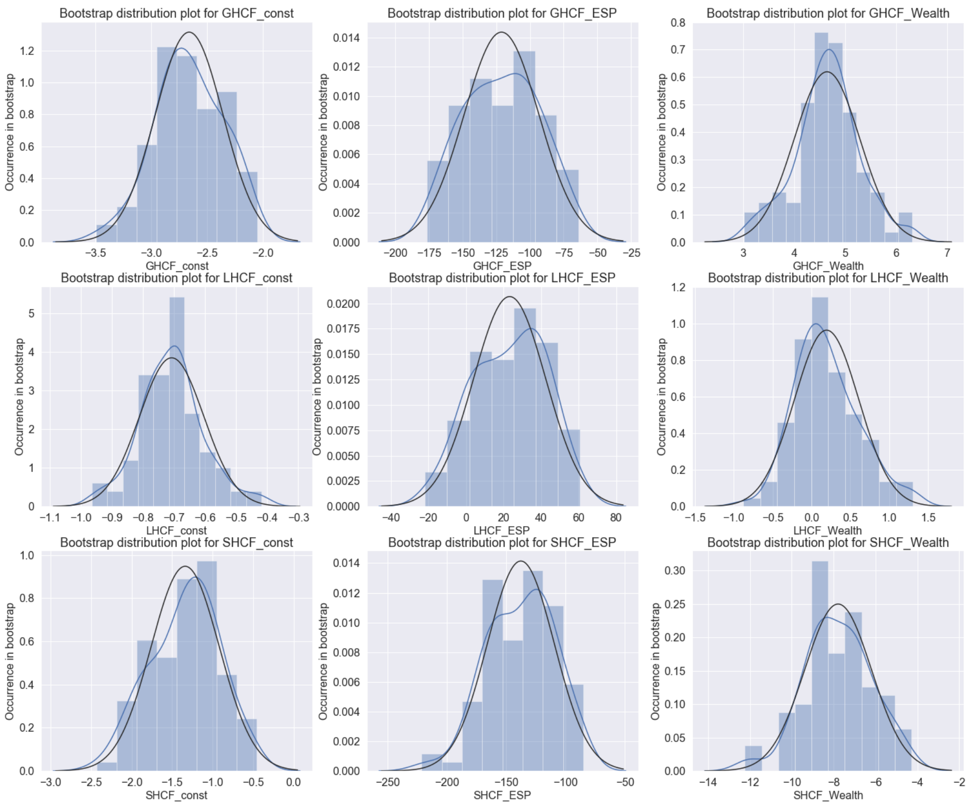

Figure D.5: Parameter Estimate Distribution Plots Agriculture & Other Industry New and Churn Market Model

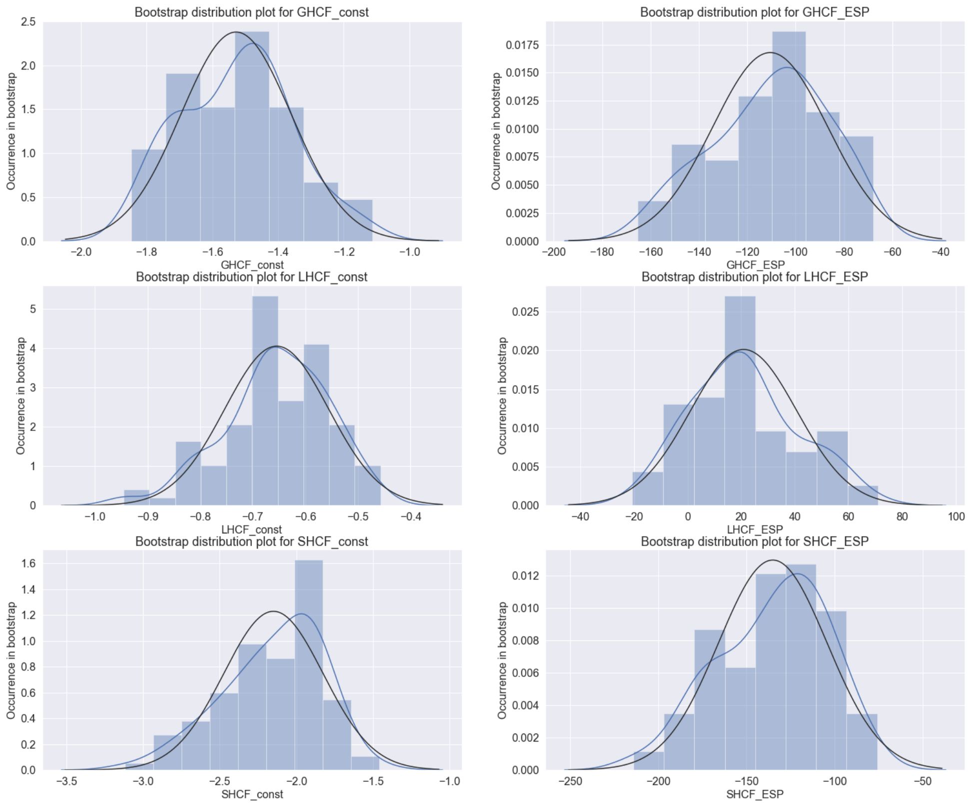

Figure D.6: Parameter Estimate Distribution Plots Agriculture & Other Industry New and Churn Market Model without the wealth variable

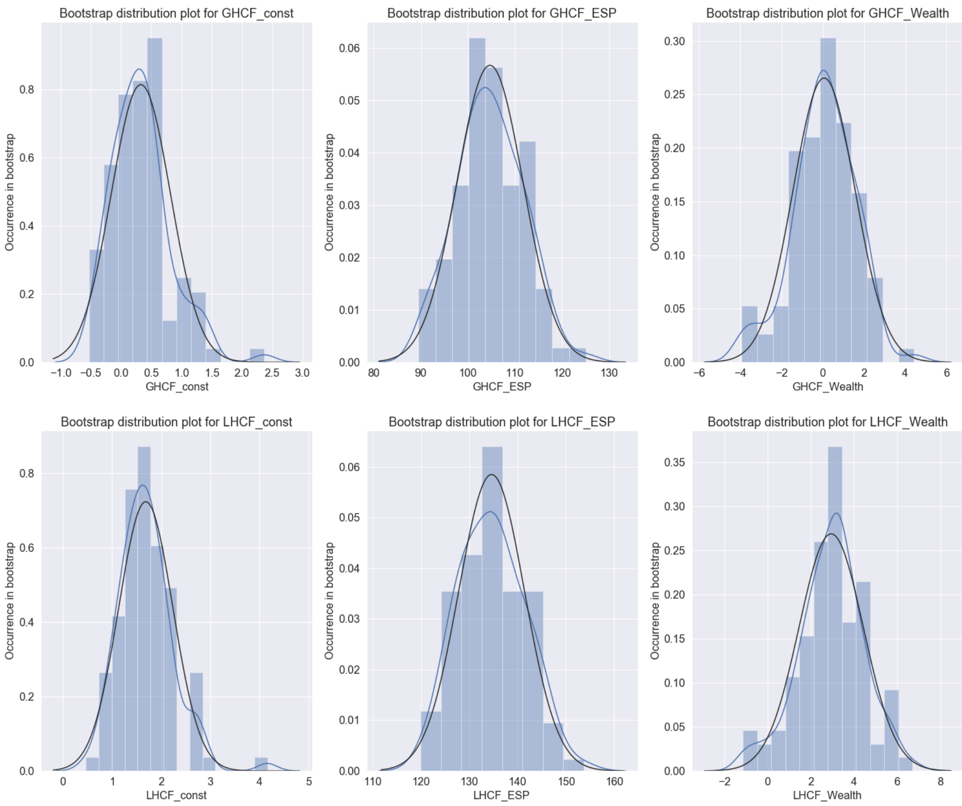

Figure D.7: Parameter Estimate Distribution Plots Passenger Transport Road Full Market Model

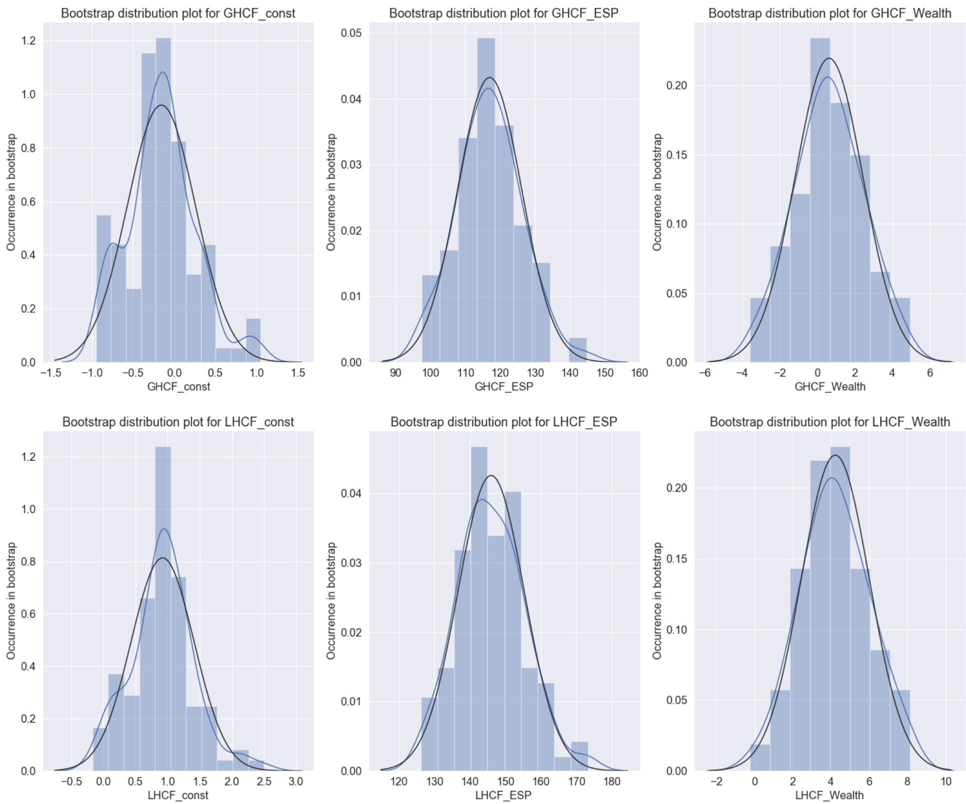

Figure D.8: Parameter Estimate Distribution Plots Passenger Transport Road New and Churn Market Model

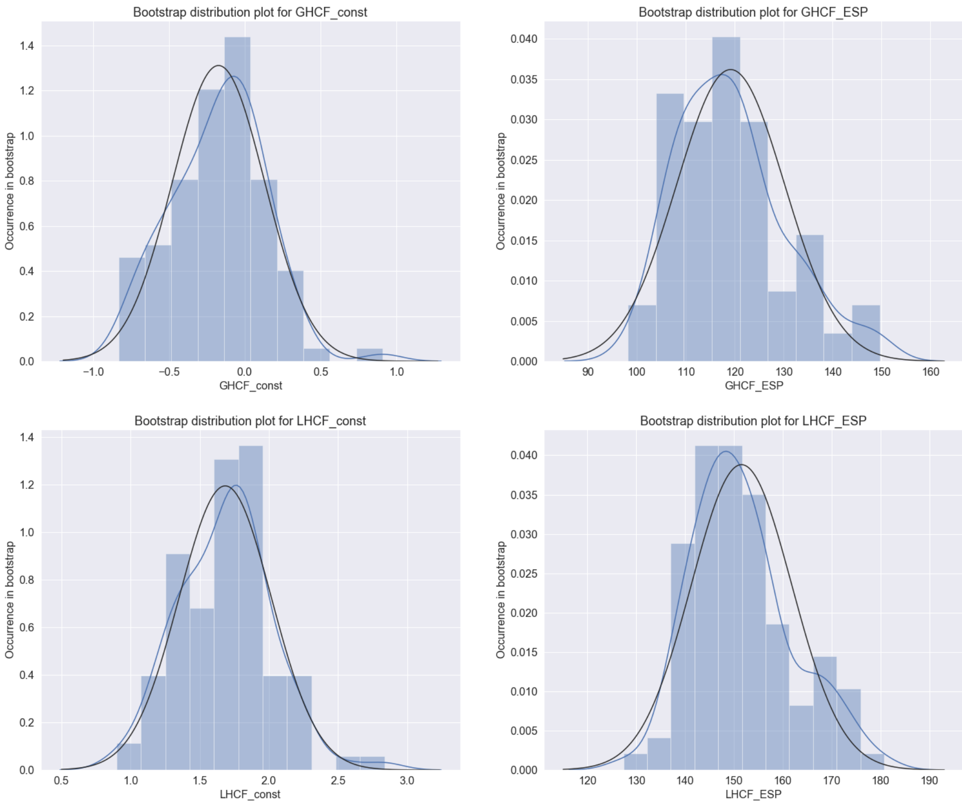

Figure D.9: Parameter Estimate Distribution Plots Passenger Transport Road New and Churn Market Model without the wealth variable

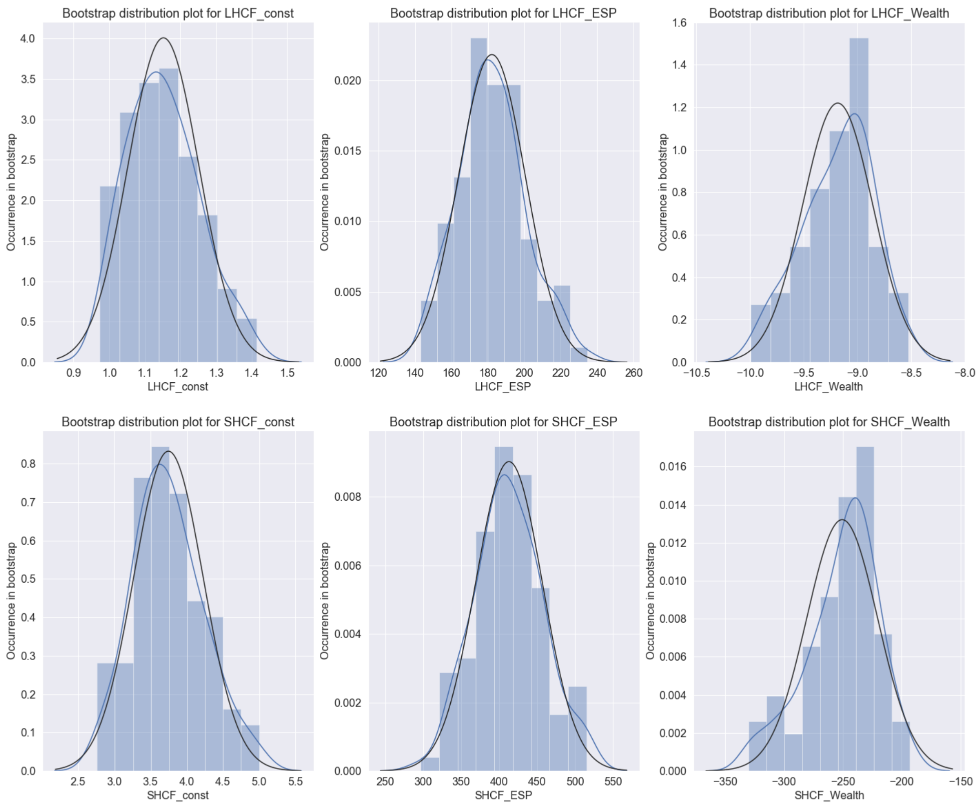

Figure D.10: Parameter Estimate Distribution Plots Passenger Transport Rail Full Market Model

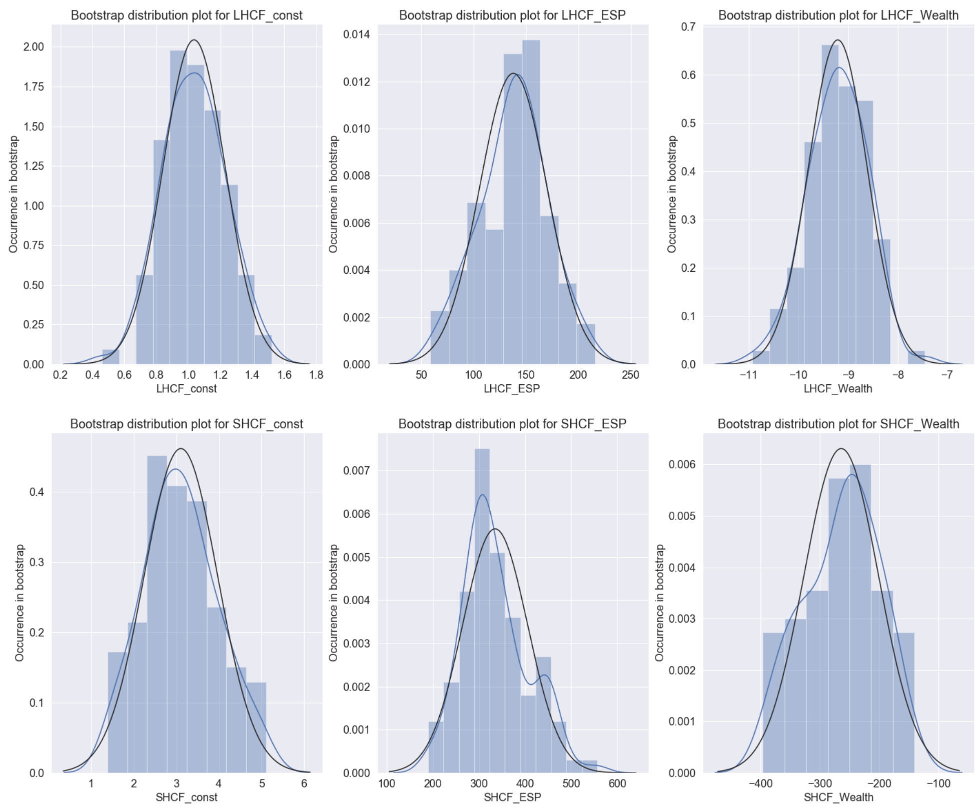

Figure D.11: Parameter Estimate Distribution Plots Passenger Transport Rail New and Churn Market Model

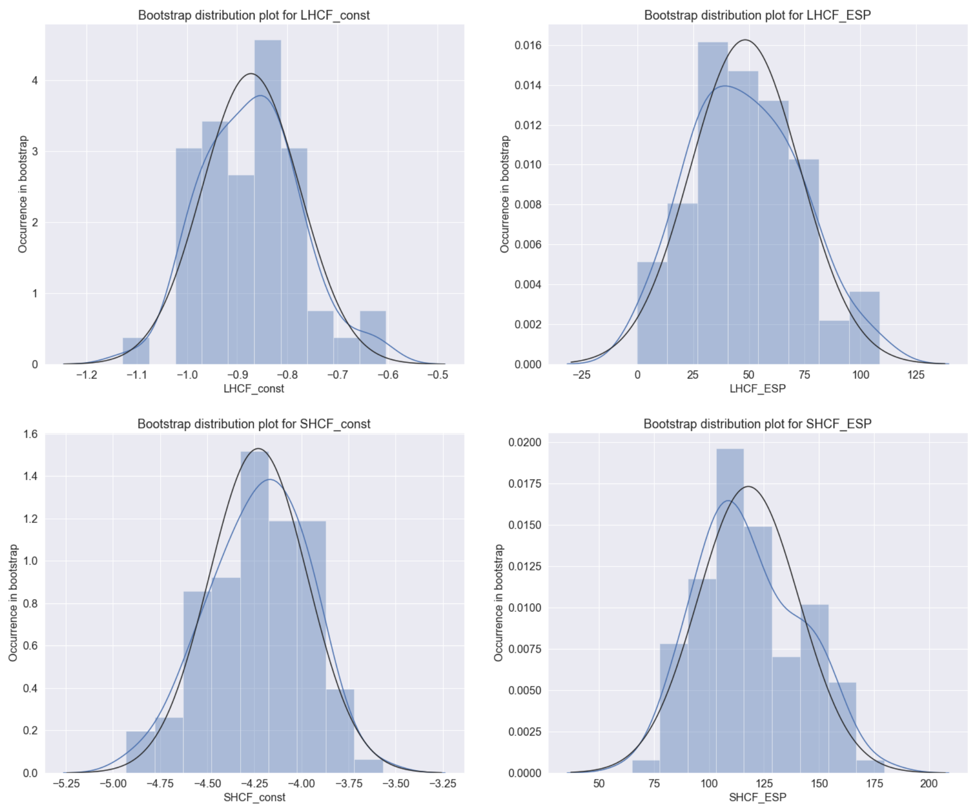

Figure D.12: Parameter Estimate Distribution Plots Passenger Transport Rail New and Churn Market Model without the wealth variable

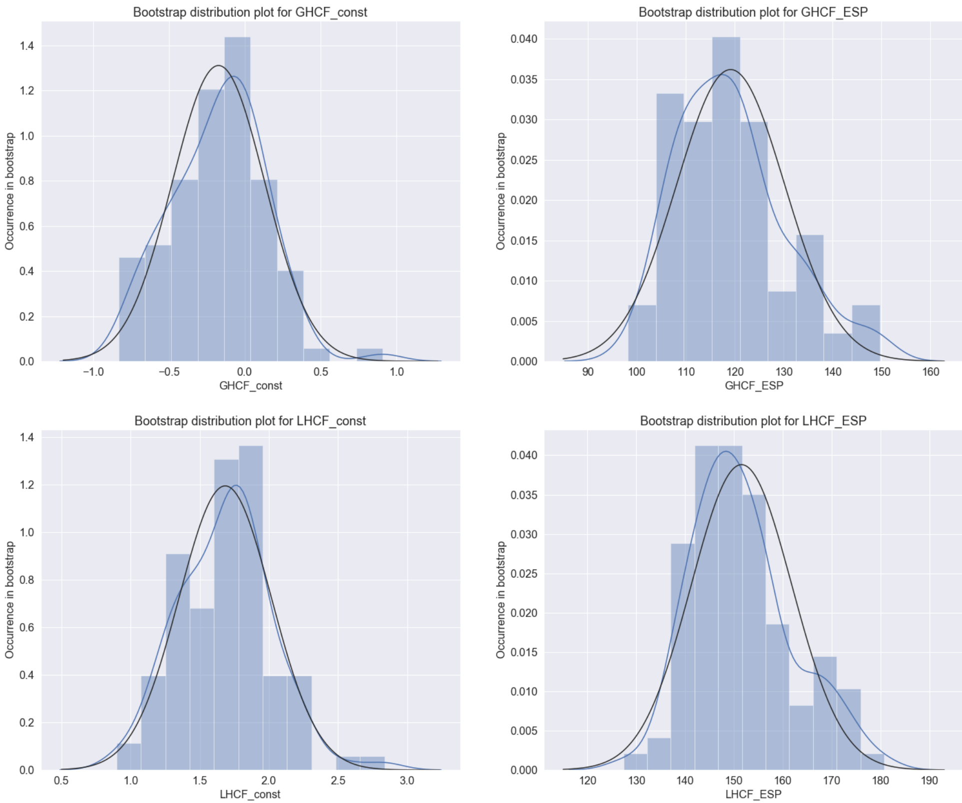

Figure D.13: Parameter Estimate Distribution Plots Passenger Transport Road New and Churn Market Model without the wealth variable

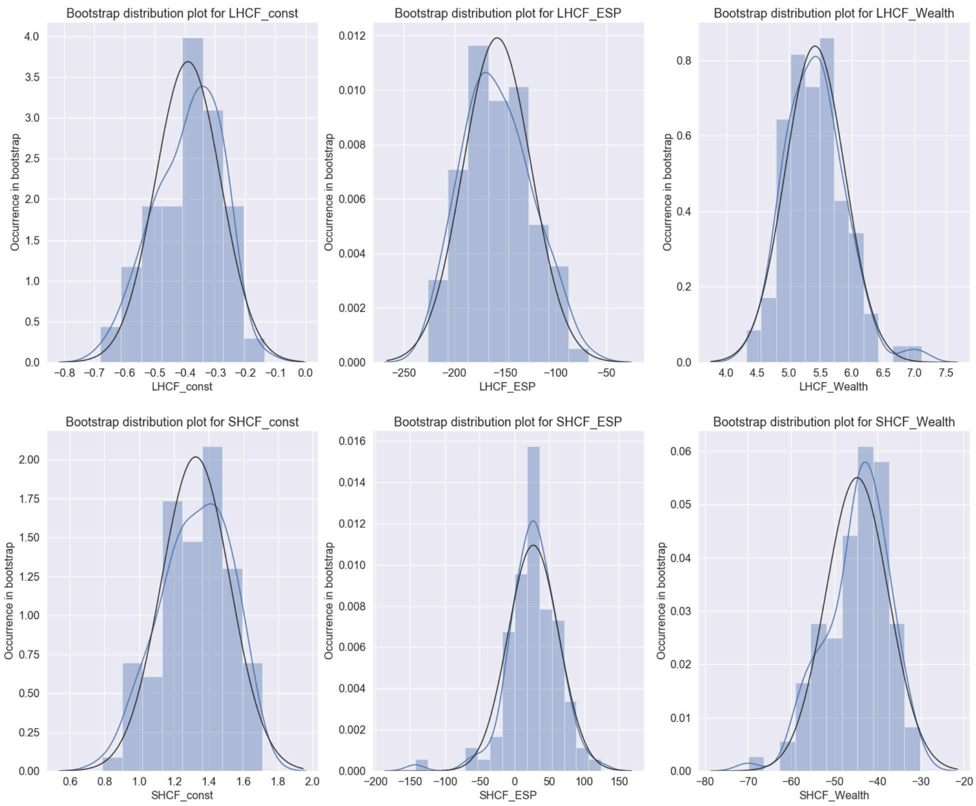

Figure D.14: Parameter Estimate Distribution Plots Freight Transport Rail Full Market Model

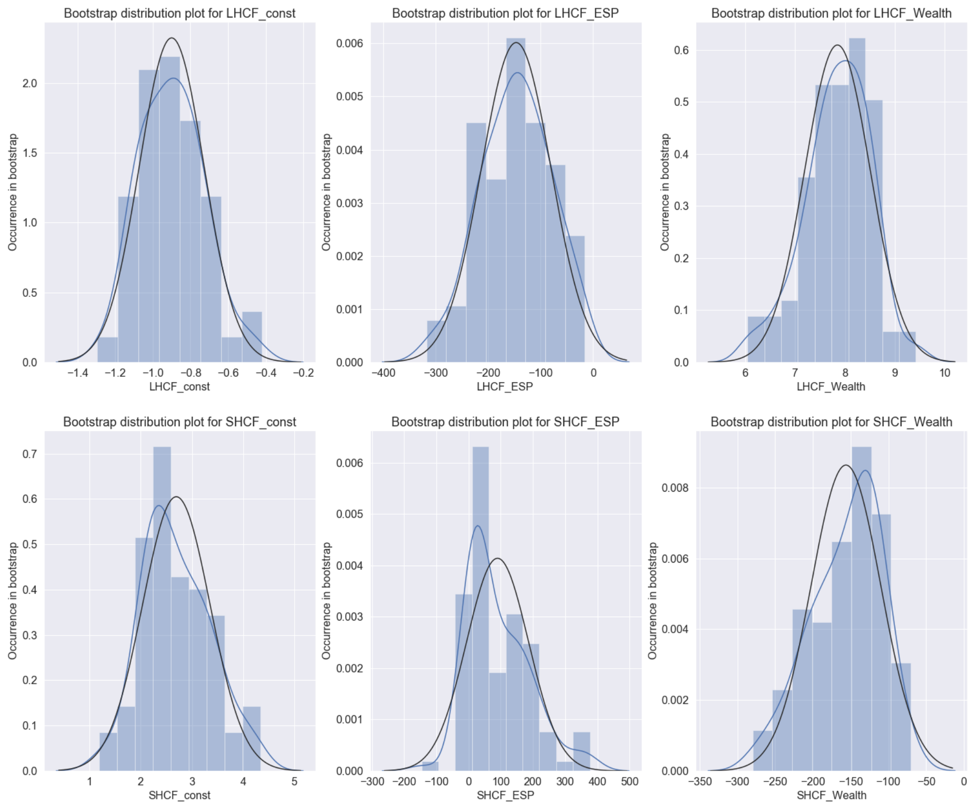

Figure D.15: Parameter Estimate Distribution Plots Freight Transport Rail New and Churn Market Model

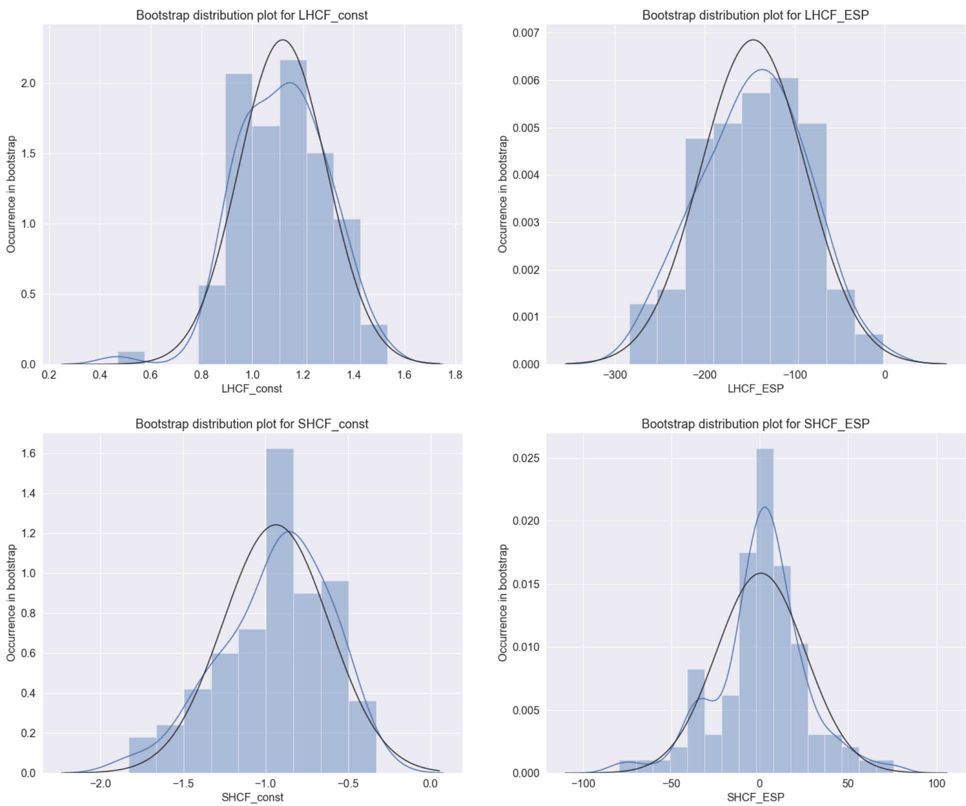

Figure D.16: Parameter Estimate Distribution Plots Freight Transport Rail New and Churn Market Model without the wealth variable

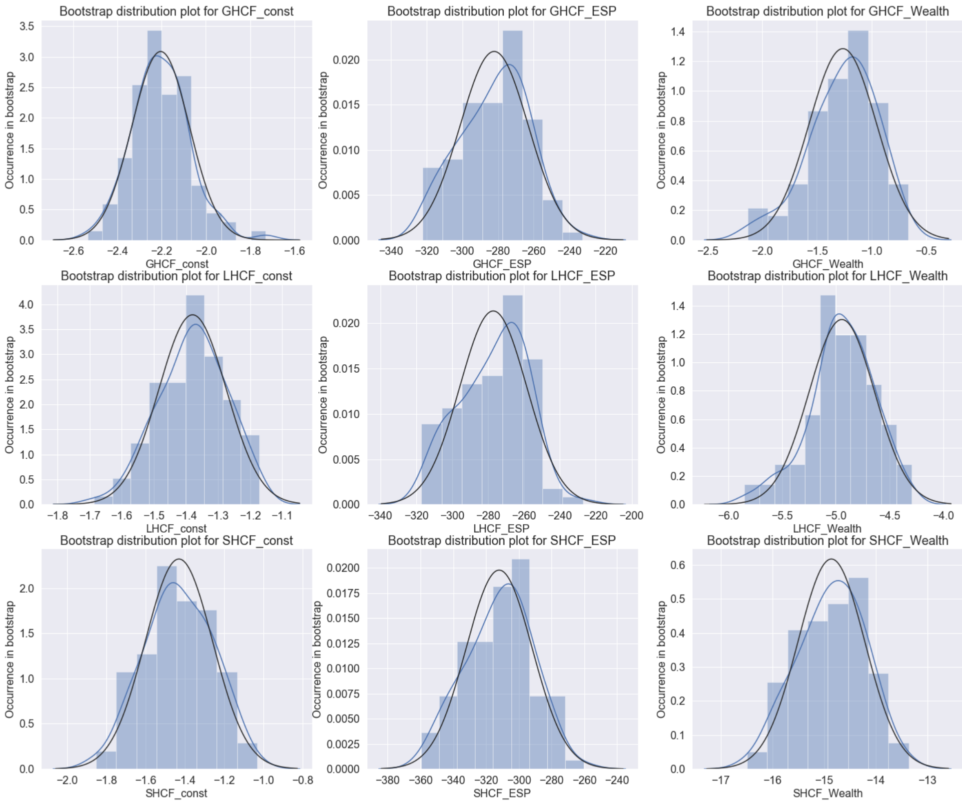

Figure D.17: Parameter Estimate Distribution Plots Services Full Market Model

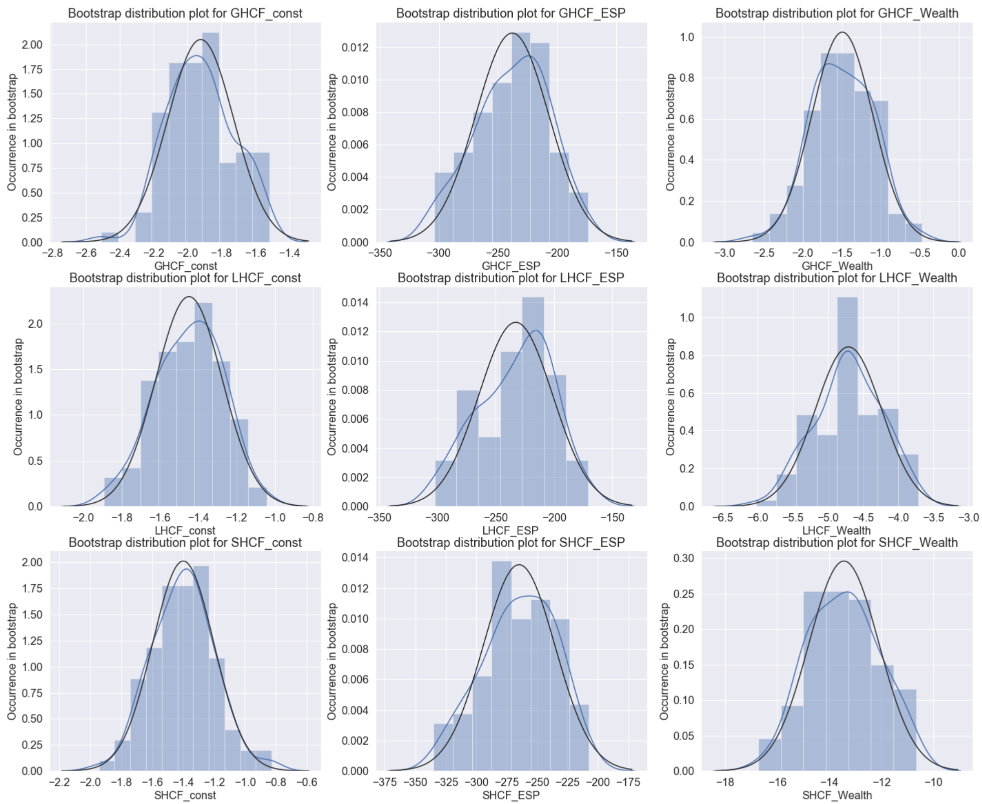

Figure D.18: Parameter Estimate Distribution Plots Services New and Churn Market Model

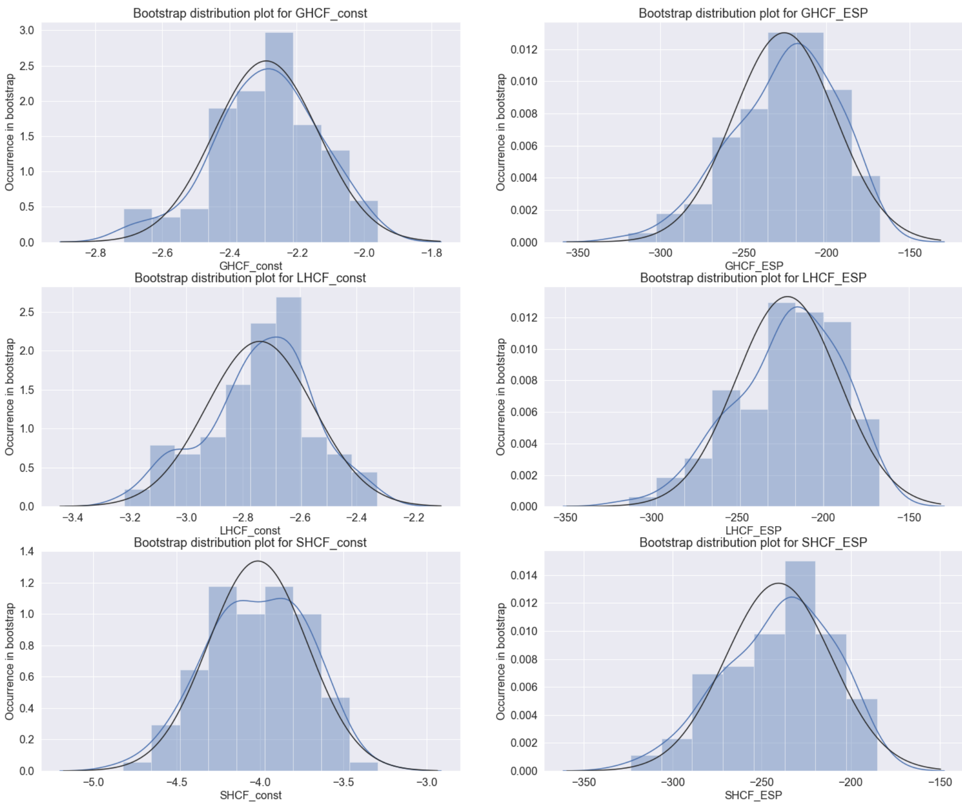

Figure D.19: Parameter Estimate Distribution Plots Services New and Churn Market Model without the wealth variable

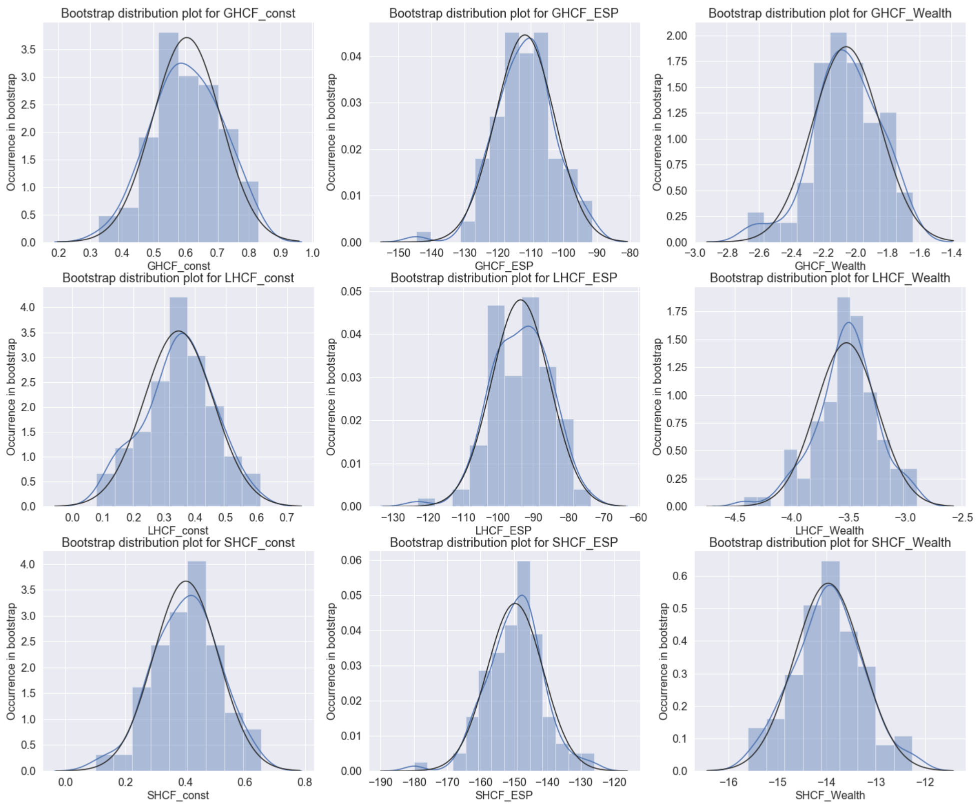

Figure D.20: Parameter Estimate Distribution Plots Residential Heating & Cooking Full Market Model

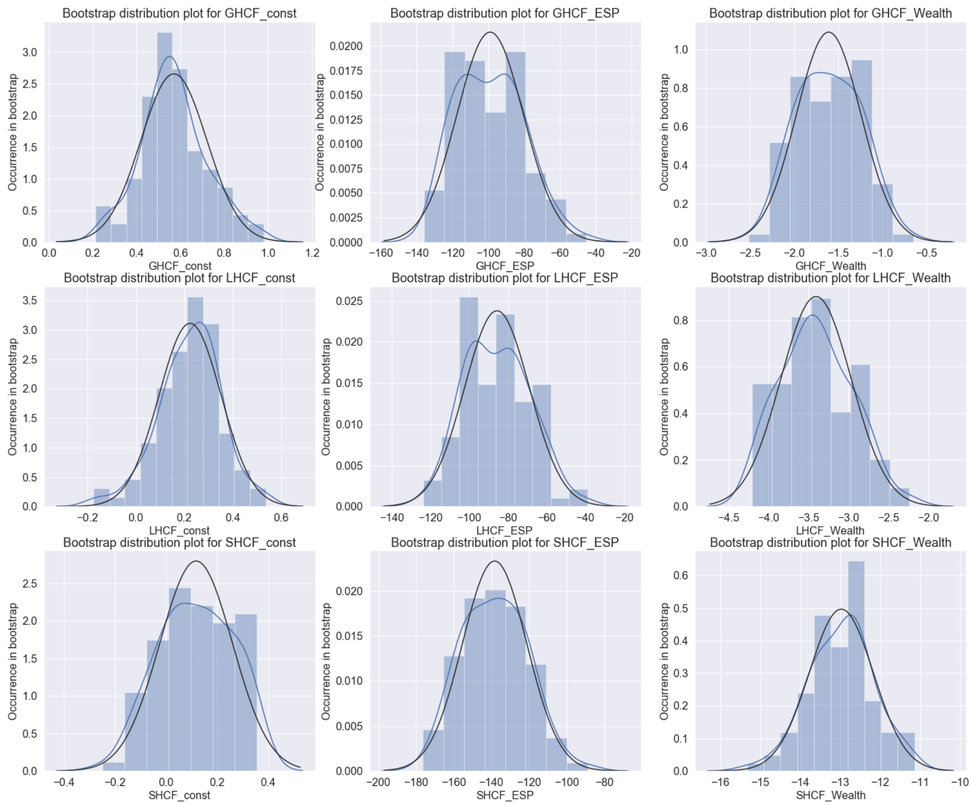

Figure D.21: Parameter Estimate Distribution Plots Residential Heating & Cooking New and Churn Market Model

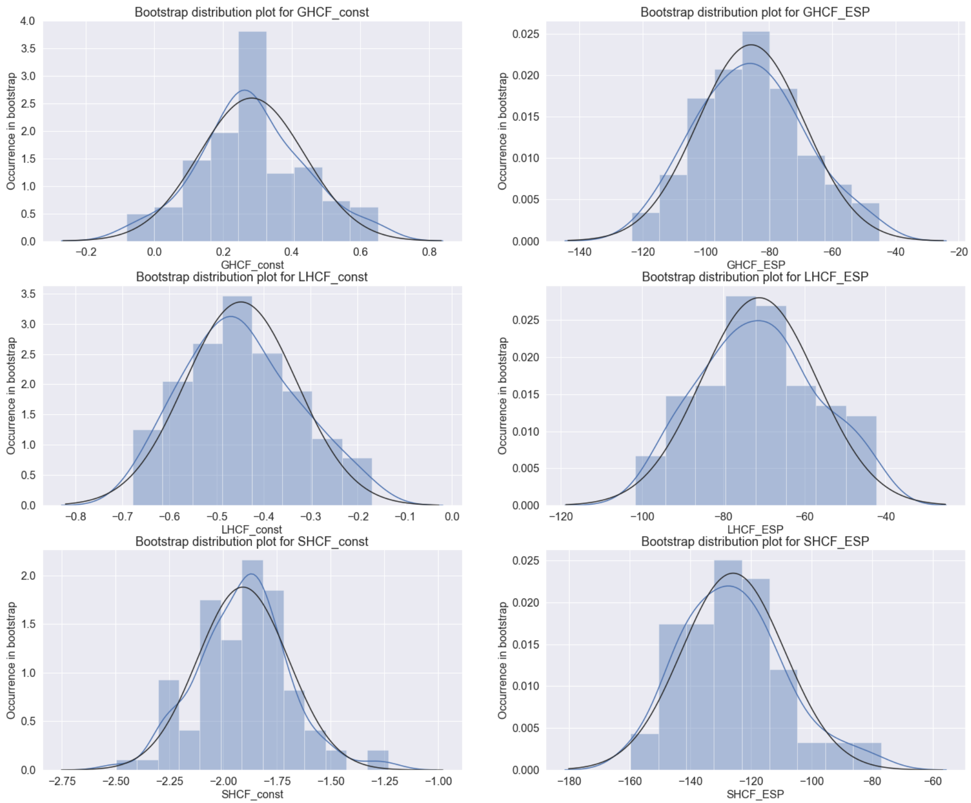

Figure D.22: Parameter Estimate Distribution Plots Residential Heating & Cooking New and Churn Market Model without the wealth variable

E

Summary Statistics for the Feature Spaces

Table E.1: Summary Statistics for the sector Heavy Industry under the Full Market Model.

	Year	Electricity - Commercial ESP	Electricity - Commercial MS	Gaseous Hydrocarbon Fuels ESP	Gaseous Hydrocarbon Fuels MS	Liquid Hydrocarbon Fuels ESP	Liquid Hydrocarbon Fuels MS	Solid Hydrocarbon Fuels ESP	Solid Hydrocarbon Fuels MS	TFC	Wealth
count	3047	3047	3047	3047	3047	3047	3047	3047	3047	3047	3047
mean	1998	0.000	0.275	-288.9	0.301	-68.45	0.129	-319.5	0.239	5.36E+05	0.196
std	11.2	0.000	0.265	248.8	0.285	380.4	0.199	274.3	0.266	1.96E+06	0.167
min	1978	0.000	0.000	-2123	0.000	-1814	0.000	-2065	0.000	4.19E+01	0.000
25%	1988	0.000	0.000	-375.2	0.048	-192.9	0.000	-419.6	0.043	2.01E+04	0.062
50%	1998	0.000	0.241	-230.0	0.240	-44.42	0.051	-260.4	0.160	9.11E+04	0.149
75%	2008	0.000	0.355	-126.9	0.431	65.07	0.172	-148.1	0.326	3.93E+05	0.295
max	2016	0.000	1.000	161.3	1.000	5291	1.000	1977	1.000	3.11E+07	1.054

Table E.2: Summary Statistics for the sector Agriculture & Other Industry under the Full Market Model.

	Year	Electricity - Commercial ESP	Electricity - Commercial MS	Gaseous Hydrocarbon Fuels ESP	Gaseous Hydrocarbon Fuels MS	Liquid Hydrocarbon Fuels ESP	Liquid Hydrocarbon Fuels MS	Solid Hydrocarbon Fuels ESP	Solid Hydrocarbon Fuels MS	TFC	Wealth
count	3563	3563	3563	3563	3563	3563	3563	3563	3563	3563	3563
mean	1997	0.000	0.358	-0.016	0.141	0.017	0.291	-0.025	0.059	5.35E+05	0.183
std	11.16	0.000	0.174	0.023	0.151	0.047	0.185	0.023	0.120	1.43E+06	0.167
min	1978	0.000	0.003	-0.622	0.000	-0.549	0.000	-0.625	0.000	1.17E+03	0.000
25%	1988	0.000	0.234	-0.023	0.016	0.001	0.154	-0.032	0.000	6.78E+04	0.048
50%	1998	0.000	0.352	-0.014	0.094	0.013	0.250	-0.022	0.011	1.44E+05	0.132
75%	2007	0.000	0.469	-0.007	0.221	0.029	0.395	-0.013	0.054	5.01E+05	0.283
max	2016	0.000	0.996	0.078	0.849	0.866	0.991	0.029	0.838	1.60E+07	1.054

Table E.3: Summary Statistics for the sector Passenger Transport Road under the Full Market Model.

	Year	Electricity - Commercial ESP	Electricity - Commercial MS	Gaseous Hydrocarbon Fuels ESP	Gaseous Hydrocarbon Fuels MS	Liquid Hydrocarbon Fuels ESP	Liquid Hydrocarbon Fuels MS	TFC	Wealth
count	3540	3540	3540	3540	3540	3540	3540	3540	3540
mean	1997	0.000	0.001	0.028	0.026	0.077	0.973	3.74E+05	0.183
std	11.16	0.000	0.009	0.029	0.065	0.053	0.066	1.44E+06	0.167
min	1978	0.000	0.000	-0.130	0.000	-0.041	0.316	5.67E+02	0.000
25%	1988	0.000	0.000	0.007	0.000	0.045	0.984	3.19E+04	0.047
50%	1998	0.000	0.000	0.026	0.000	0.071	1.000	7.92E+04	0.132
75%	2007	0.000	0.000	0.046	0.014	0.098	1.000	2.23E+05	0.282
max	2016	0.000	0.133	0.149	0.684	0.441	1.000	1.63E+07	1.054

Table E.4: Summary Statistics for the sector Passenger Transport Rail under the Full Market Model.

	Year	Electricity - Commercial ESP	Electricity - Commercial MS	Liquid Hydrocarbon Fuels ESP	Liquid Hydrocarbon Fuels MS	Solid Hydrocarbon Fuels ESP	Solid Hydrocarbon Fuels MS	TFC	Wealth
count	2589	2589	2589	2589	2589	2589	2589	2589	2589
mean	1998	0.000	0.556	0.003	0.428	0.020	0.016	9.73E+03	0.204
std	11.07	0.000	0.373	0.008	0.371	0.020	0.112	1.91E+04	0.168
min	1978	0.000	0.000	-0.032	0.000	-0.015	0.000	3.86E-01	0.000
25%	1989	0.000	0.171	-0.002	0.079	0.006	0.000	8.79E+02	0.069
50%	1999	0.000	0.645	0.001	0.329	0.015	0.000	2.48E+03	0.161
75%	2008	0.000	0.909	0.006	0.788	0.027	0.000	7.86E+03	0.302
max	2016	0.000	1.000	0.075	1.000	0.220	1.000	2.32E+05	1.054

Table E.5: Summary Statistics for the sector Freight Transport Rail under the Full Market Model.

	Year	Electricity - Commercial ESP	Electricity - Commercial MS	Liquid Hydrocarbon Fuels ESP	Liquid Hydrocarbon Fuels MS	Solid Hydrocarbon Fuels ESP	Solid Hydrocarbon Fuels MS	TFC	Wealth
count	2580	2580	2580	2580	2580	2580	2580	2580	2580
mean	1998	0.000	0.414	0.002	0.560	0.006	0.025	2.02E+04	0.204
std	11.07	0.000	0.373	0.007	0.378	0.011	0.126	6.75E+04	0.168
min	1978	0.000	0.000	-0.025	0.000	-0.023	0.000	4.45E-03	0.000
25%	1989	0.000	0.004	-0.001	0.211	0.000	0.000	5.97E+02	0.069
50%	1999	0.000	0.379	0.001	0.578	0.003	0.000	2.20E+03	0.161
75%	2008	0.000	0.759	0.004	0.981	0.011	0.000	6.49E+03	0.302
max	2016	0.000	1.000	0.069	1.000	0.087	1.000	6.05E+05	1.054

Table E.6: Summary Statistics for the sector Services under the Full Market Model.

	Year	Electricity - Commercial ESP	Electricity - Commercial MS	Gaseous Hydrocarbon Fuels ESP	Gaseous Hydrocarbon Fuels MS	Liquid Hydrocarbon Fuels ESP	Liquid Hydrocarbon Fuels MS	Solid Hydrocarbon Fuels ESP	Solid Hydrocarbon Fuels MS	TFC	Wealth
count	3362	3362	3362	3362	3362	3362	3362	3362	3362	3362	3362
mean	1998	0.000	0.562	-0.026	0.135	-0.012	0.149	-0.039	0.036	2.56E+05	0.188
std	11.10	0.000	0.296	0.029	0.180	0.038	0.222	0.029	0.110	8.60E+05	0.167
min	1978	0.000	0.000	-0.302	0.000	-0.253	0.000	-0.307	0.000	8.37E+01	0.000
25%	1988	0.000	0.357	-0.037	0.000	-0.027	0.000	-0.051	0.000	1.47E+04	0.051
50%	1998	0.000	0.548	-0.022	0.069	-0.011	0.056	-0.035	0.000	6.03E+04	0.139
75%	2007	0.000	0.802	-0.010	0.200	0.002	0.196	-0.021	0.004	1.39E+05	0.289
max	2016	0.000	1.000	0.280	1.000	0.518	1.000	0.005	0.964	8.85E+06	1.054

Table E.7: Summary Statistics for the sector Residential Heating & Cooking under the Full Market Model.

	Year	Electricity - Commercial ESP	Electricity - Commercial MS	Gaseous Hydrocarbon Fuels ESP	Gaseous Hydrocarbon Fuels MS	Liquid Hydrocarbon Fuels ESP	Liquid Hydrocarbon Fuels MS	Solid Hydrocarbon Fuels ESP	Solid Hydrocarbon Fuels MS	TFC	Wealth
count	2364	2364	2364	2364	2364	2364	2364	2364.00	2364	2364	2364
mean	1997	0.000	0.174	-0.028	0.303	-0.016	0.151	-0.028	0.045	1.15E+05	0.176
std	11.24	0.000	0.195	0.029	0.263	0.030	0.192	0.029	0.123	2.75E+05	0.168
min	1978	0.000	0.000	-0.306	0.000	-0.232	0.000	-0.297	0.000	0.00E+00	0.000
25%	1987	0.000	0.044	-0.040	0.052	-0.029	0.010	-0.040	0.000	1.16E+04	0.043
50%	1997	0.000	0.115	-0.023	0.251	-0.011	0.073	-0.024	0.000	3.44E+04	0.119
75%	2007	0.000	0.220	-0.011	0.478	0.002	0.221	-0.011	0.014	8.55E+04	0.276
max	2016	0.000	1.000	0.172	0.993	0.142	0.960	0.036	0.916	3.19E+06	1.054

Table E.8: Summary Statistics for the sector Heavy Industry under the New and Churn Market Model with a churn rate of 10%.

	Year	Electricity - Commercial ESP	Electricity - Commercial MS	Gaseous Hydrocarbon Fuels ESP	Gaseous Hydrocarbon Fuels MS	Liquid Hydrocarbon Fuels ESP	Liquid Hydrocarbon Fuels MS	Solid Hydrocarbon Fuels ESP	Solid Hydrocarbon Fuels MS	TFC	Wealth
count	1749	1749	1749	1749	1749	1749	1749	1749	1749	1749	3047
mean	1997	0.262	-310	0.297	-108	0.118	-342	0.274	8.63E+04	0.174	0.196
std	11.22	0.304	258	0.318	323	0.213	280	0.309	3.60E+05	0.157	0.167
min	1978	0.000	-2123	0.000	-1814	0.000	-2065	0.000	0.00E+00	0.000	0.000
25%	1987	0.000	-412	0.014	-230	0.000	-457	0.013	1.52E+03	0.048	0.062
50%	1997	0.193	-245	0.181	-82	0.016	-275	0.165	1.00E+04	0.126	0.149
75%	2006	0.356	-137	0.49	36	0.139	-159	0.399	4.98E+04	0.269	0.295
max	2016	1.000	161	1.000	1878	1.000	1977	1.000	4.96E+06	1.006	1.054

Table E.9: Summary Statistics for the sector Agriculture & Other Industry under the New and Churn Market Model with a churn rate of 10%.

	Year	Electricity - Commercial ESP	Electricity - Commercial MS	Gaseous Hydrocarbon Fuels ESP	Gaseous Hydrocarbon Fuels MS	Liquid Hydrocarbon Fuels ESP	Liquid Hydrocarbon Fuels MS	Solid Hydrocarbon Fuels ESP	Solid Hydrocarbon Fuels MS	TFC	Wealth
count	1771	1771	1771	1771	1771	1771	1771	1771	1771	1771	1771
mean	1997	0.000	0.350	-0.017	0.149	0.013	0.278	-0.025	0.07	9.25E+04	0.173
std	11.22	0.000	0.204	0.026	0.174	0.036	0.216	0.026	0.14	3.71E+05	0.165
min	1978	0.000	0.000	-0.622	0.000	-0.549	0.000	-0.625	0.00	1.00E+02	0.000
25%	1987	0.000	0.190	-0.023	0.011	0.000	0.114	-0.032	0.00	8.86E+03	0.039
50%	1997	0.000	0.339	-0.015	0.083	0.011	0.226	-0.022	0.01	2.19E+04	0.117
75%	2006	0.000	0.481	-0.007	0.233	0.026	0.394	-0.013	0.07	7.13E+04	0.276
max	2016	0.000	0.995	0.078	0.884	0.615	0.997	0.018	0.85	1.17E+07	1.006

Table E.10: Summary Statistics for the sector Passenger Transport Road under the New and Churn Market Model with a churn rate of 15%

	Year	Electricity - Commercial ESP	Electricity - Commercial MS	Gaseous Hydrocarbon Fuels ESP	Gaseous Hydrocarbon Fuels MS	Liquid Hydrocarbon Fuels ESP	Liquid Hydrocarbon Fuels MS	TFC	Wealth
count	3226	3226	3226	3226	3226	3226	3226	3226	3226
mean	1997	0.000	0.002	0.028	0.035	0.076	0.964	6.22E+04	0.181
std	11.19	0.000	0.013	0.029	0.093	0.052	0.095	2.29E+05	0.164
min	1978	0.000	0.000	-0.130	0.000	-0.041	0.012	0.00E+00	0.000
25%	1988	0.000	0.000	0.007	0.000	0.044	0.978	5.03E+03	0.045
50%	1997	0.000	0.000	0.026	0.000	0.070	1.000	1.41E+04	0.130
75%	2007	0.000	0.000	0.046	0.020	0.098	1.000	4.04E+04	0.281
max	2016	0.000	0.319	0.149	0.988	0.441	1.000	2.97E+06	1.054

Table E.11: Summary Statistics for the sector Passenger Transport Rail under the New and Churn Market Model with a churn rate of 15%

	Year	Electricity - Commercial ESP	Electricity - Commercial MS	Liquid Hydrocarbon Fuels ESP	Liquid Hydrocarbon Fuels MS	Solid Hydrocarbon Fuels ESP	Solid Hydrocarbon Fuels MS	TFC	Wealth
count	1956	1956	1956	1956	1956	1956	1956	1.96E+03	1956
mean	1999	0.000	0.572	0.003	0.417	0.020	0.011	1.88E+03	0.214
std	11.13	0.000	0.387	0.008	0.385	0.021	0.086	6.40E+03	0.170
min	1978	0.000	0.000	-0.032	0.000	-0.015	0.000	0.00E+00	0.000
25%	1989	0.000	0.152	-0.002	0.054	0.006	0.000	1.58E+02	0.069
50%	1999	0.000	0.698	0.001	0.289	0.016	0.000	4.32E+02	0.174
75%	2008	0.000	0.941	0.007	0.822	0.027	0.000	1.33E+03	0.324
max	2016	0.000	1.000	0.064	1.000	0.220	1.000	2.32E+05	1.054

Table E.12: Summary Statistics for the sector Freight Transport Rail under the New and Churn Market Model with a churn rate of 15%

	Year	Electricity - Commercial ESP	Electricity - Commercial MS	Liquid Hydrocarbon Fuels ESP	Liquid Hydrocarbon Fuels MS	Solid Hydrocarbon Fuels ESP	Solid Hydrocarbon Fuels MS	TFC	Wealth
count	1989	1989	1989	1989	1989	1989	1989	1989	1989
mean	1999	0.000	0.421	0.002	0.559	0.006	0.02	3.39E+03	0.212
std	11.05	0.000	0.388	0.007	0.392	0.011	0.11	1.32E+04	0.170
min	1978	0.000	0.000	-0.025	0.000	-0.022	0.00	0.00E+00	0.000
25%	1990	0.000	0.001	-0.001	0.161	0.000	0.00	9.72E+01	0.069
50%	2000	0.000	0.349	0.001	0.605	0.003	0.00	3.61E+02	0.170
75%	2008	0.000	0.807	0.004	0.995	0.011	0.00	1.06E+03	0.324
max	2016	0.000	1.000	0.061	1.000	0.087	1.00	2.65E+05	1.054

Table E.13: Summary Statistics for the sector Services under the New and Churn Market Model with a churn rate of 15%

	Year	Electricity - Commercial ESP	Electricity - Commercial MS	Gaseous Hydrocarbon Fuels ESP	Gaseous Hydrocarbon Fuels MS	Liquid Hydrocarbon Fuels ESP	Liquid Hydrocarbon Fuels MS	Solid Hydrocarbon Fuels ESP	Solid Hydrocarbon Fuels MS	TFC	Wealth
count	2362	2362	2362	2362	2362	2362	2362	2362	2362	2362	2362
mean	1997	0.000	0.584	-0.027	0.130	-0.014	0.142	-0.039	0.034	4.33E+04	0.180
std	11.29	0.000	0.319	0.028	0.194	0.033	0.230	0.029	0.118	1.41E+05	0.167
min	1978	0.000	0.000	-0.222	0.000	-0.211	0.000	-0.236	0.000	0.00E+00	0.000
25%	1987	0.000	0.347	-0.039	0.000	-0.028	0.000	-0.052	0.000	2.23E+03	0.042
50%	1997	0.000	0.586	-0.023	0.051	-0.012	0.033	-0.035	0.000	9.38E+03	0.122
75%	2007	0.000	0.885	-0.010	0.184	0.001	0.183	-0.021	0.000	2.52E+04	0.282
max	2016	0.000	1.000	0.098	1.000	0.366	1.000	0.005	0.979	1.67E+06	1.006

Table E.14: Summary Statistics for the sector Residential Heating & Cooking under the New and Churn Market Model with a churn rate of 15%

	Year	Electricity - Commercial ESP	Electricity - Commercial MS	Gaseous Hydrocarbon Fuels ESP	Gaseous Hydrocarbon Fuels MS	Liquid Hydrocarbon Fuels ESP	Liquid Hydrocarbon Fuels MS	Solid Hydrocarbon Fuels ESP	Solid Hydrocarbon Fuels MS	TFC	Wealth
count	2364	2364	2364	2364	2364	2364	2364	2364.00	2364	2364	2364
mean	1997	0.000	0.174	-0.028	0.303	-0.016	0.151	-0.028	0.045	1.15E+05	0.176
std	11.24	0.000	0.195	0.029	0.263	0.030	0.192	0.029	0.123	2.75E+05	0.168
min	1978	0.000	0.000	-0.306	0.000	-0.232	0.000	-0.297	0.000	0.00E+00	0.000
25%	1987	0.000	0.044	-0.040	0.052	-0.029	0.010	-0.040	0.000	1.16E+04	0.043
50%	1997	0.000	0.115	-0.023	0.251	-0.011	0.073	-0.024	0.000	3.44E+04	0.119
75%	2007	0.000	0.220	-0.011	0.478	0.002	0.221	-0.011	0.014	8.55E+04	0.276
max	2016	0.000	1.000	0.172	0.993	0.142	0.960	0.036	0.916	3.19E+06	1.054

F

Feature Space Plots

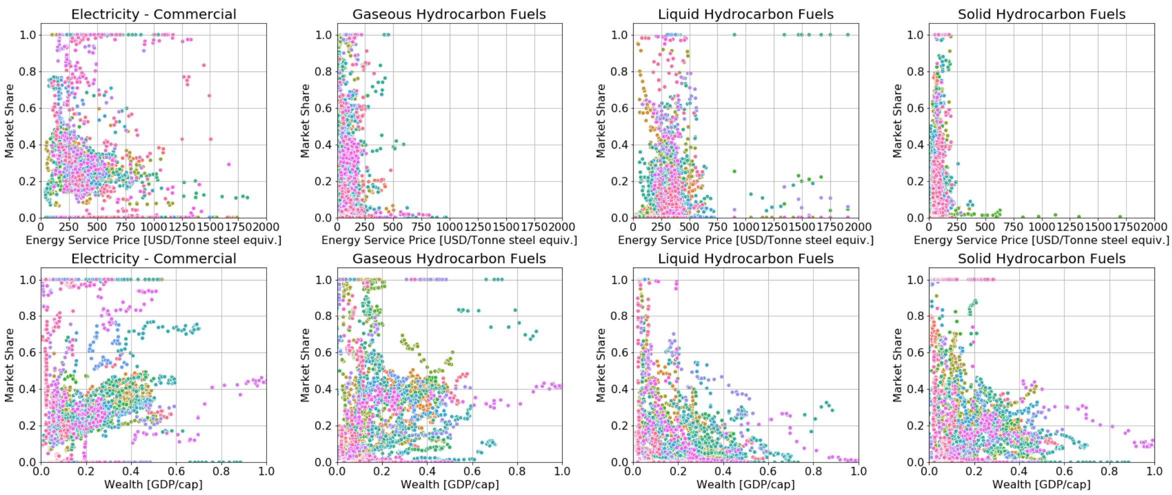

Figure F.1: Visual representation of the feature space for Heavy Industry depicting energy carrier market shares plotted against the energy service price, the wealth proxy, and urbanization for the full market model.

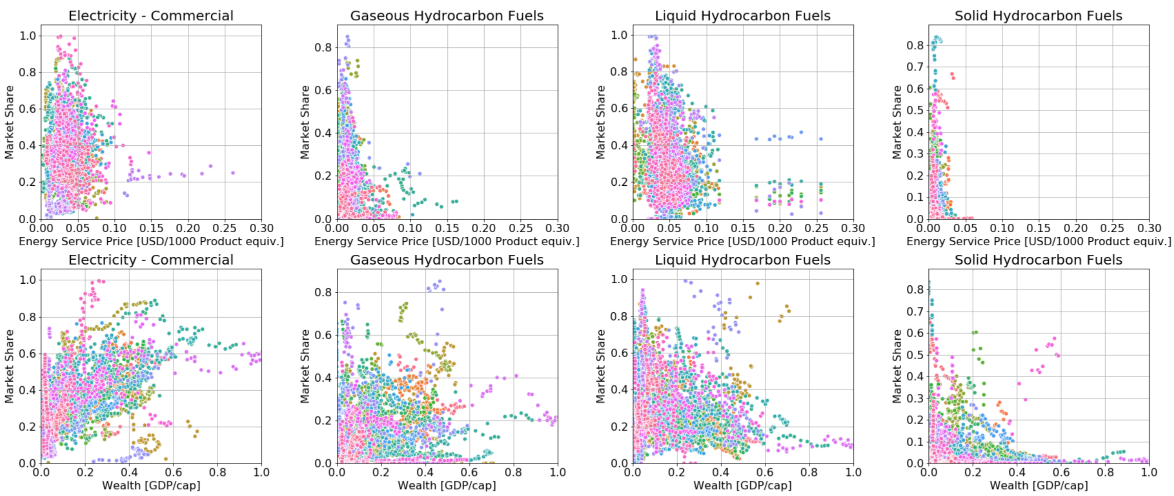

Figure F.2: Visual representation of the feature space for Agriculture & Other Industry depicting energy carrier market shares plotted against the energy service price, the wealth proxy, and urbanization for the full market model.

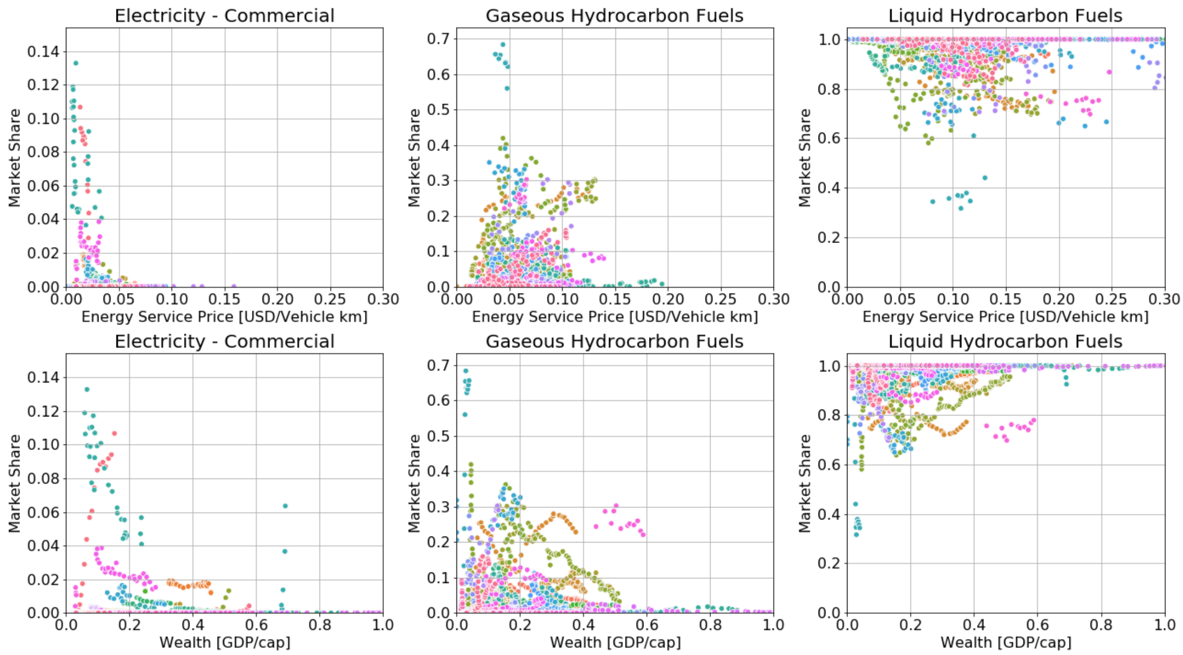

Figure F.3: Visual representation of the feature space for Passenger Transport Road depicting energy carrier market shares plotted against the energy service price, the wealth proxy, and urbanization for the full market model.

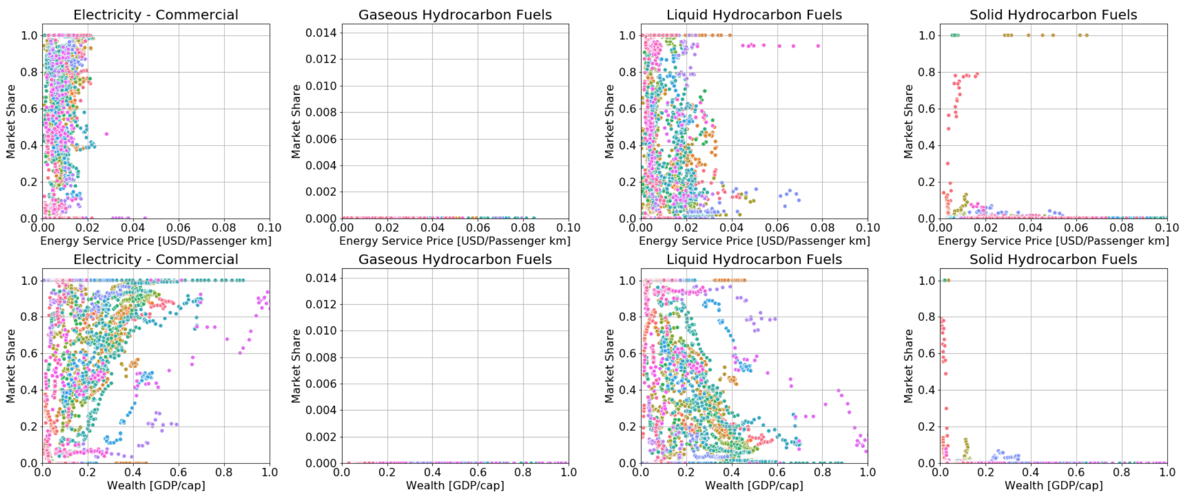

Figure F.4: Visual representation of the feature space for Passenger Transport Rail depicting energy carrier market shares plotted against the energy service price, the wealth proxy, and urbanization for the full market model.

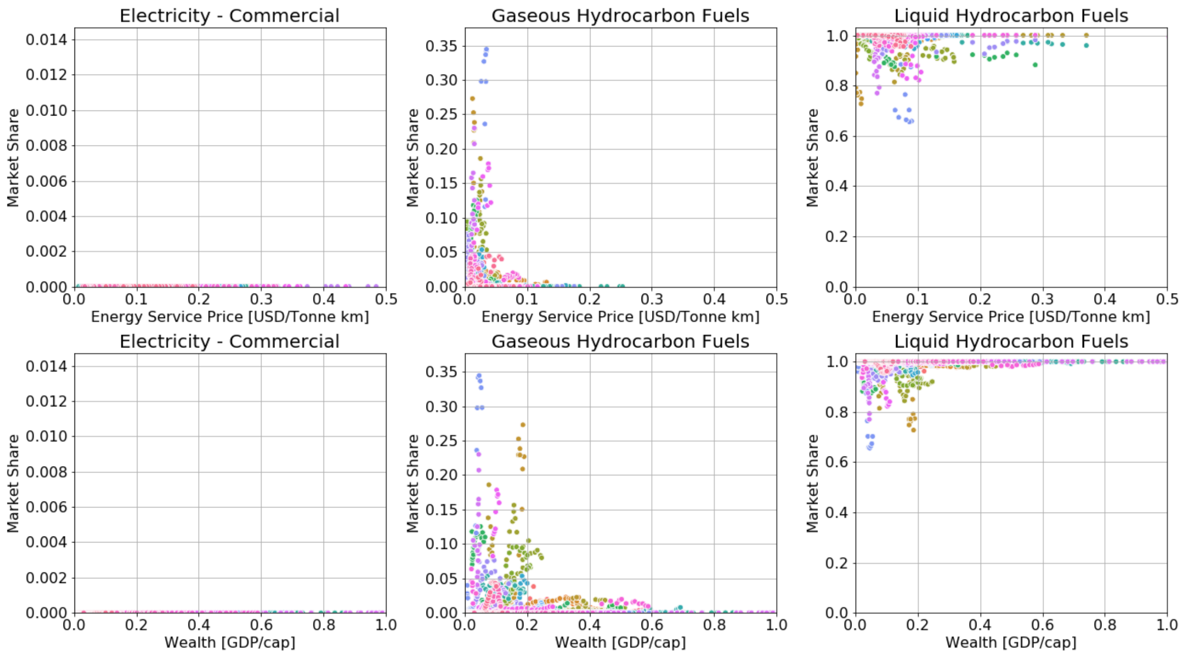

Figure F.5: Visual representation of the feature space for Freight Transport Road depicting energy carrier market shares plotted against the energy service price, the wealth proxy, and urbanization for the full market model.

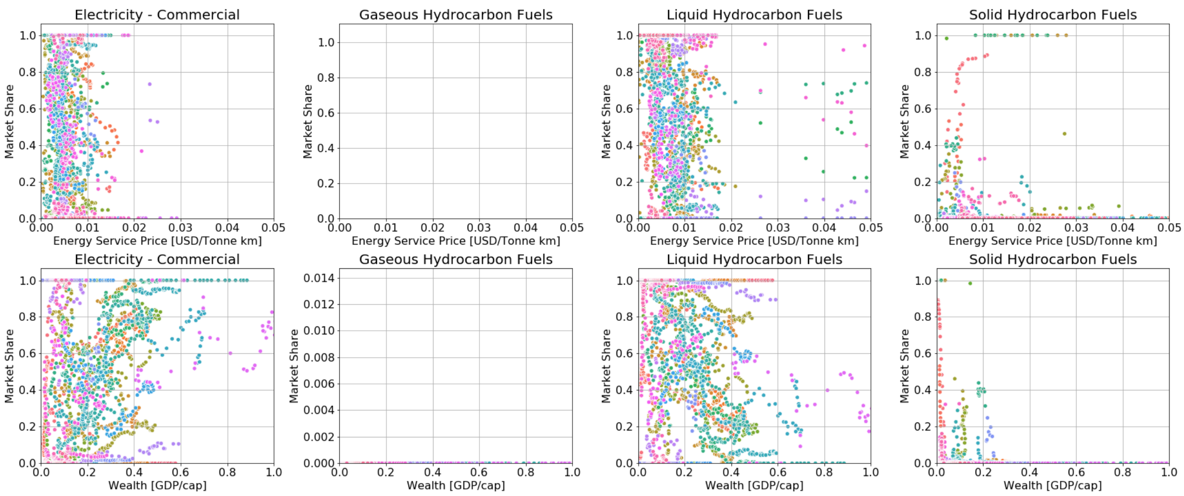

Figure F.6: Visual representation of the feature space for Freight Transport Rail depicting energy carrier market shares plotted against the energy service price, the wealth proxy, and urbanization for the full market model.

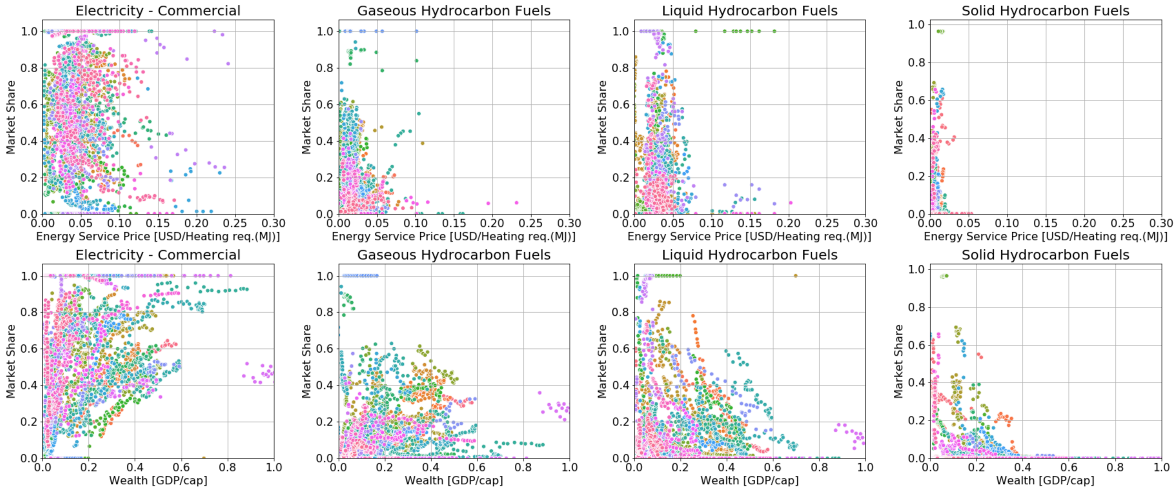

Figure F.7: Visual representation of the feature space for Services depicting energy carrier market shares plotted against the energy service price, the wealth proxy, and urbanization for the full market model.

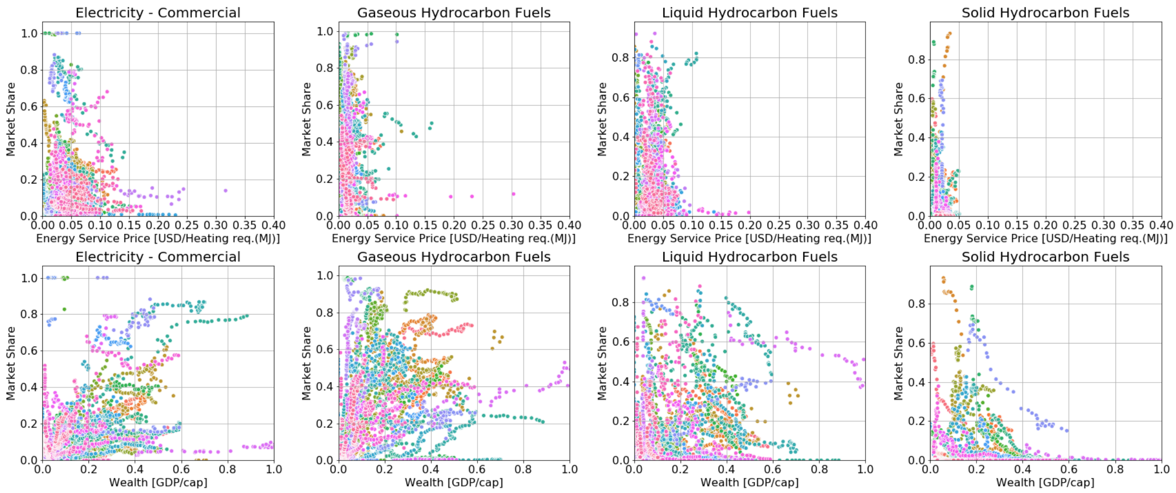

Figure F.8: Visual representation of the feature space for Residential Heating & Cooking depicting energy carrier market shares plotted against the energy service price, the wealth proxy, and urbanization for the full market model.

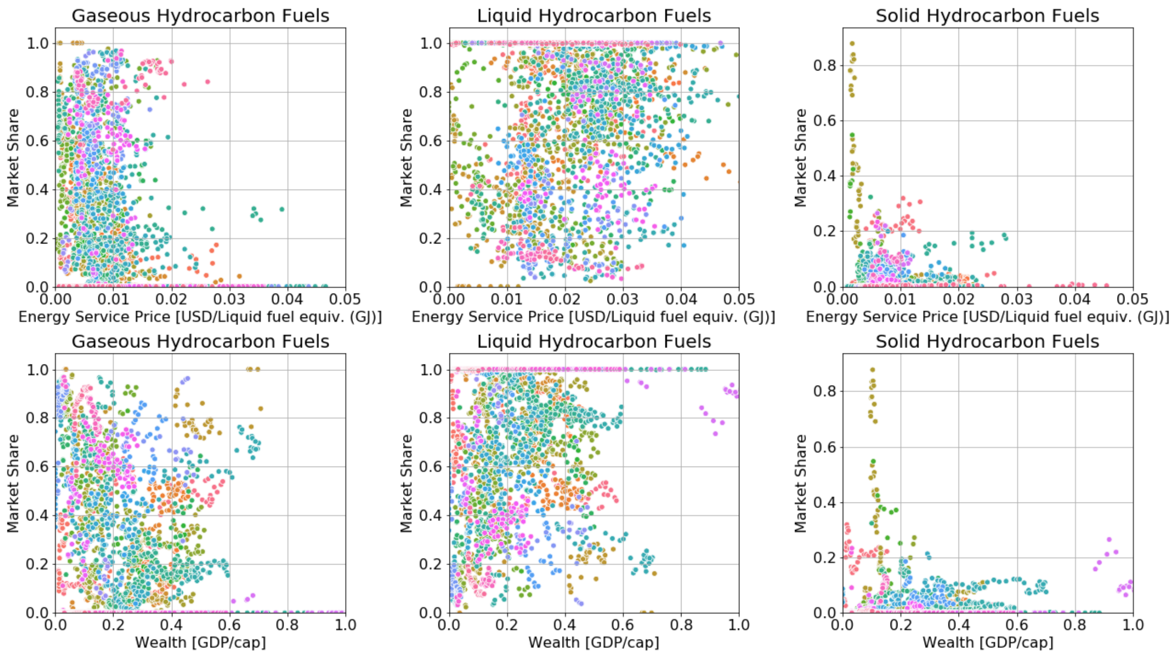

Figure F.9: Visual representation of the feature space for Non Energy Use depicting energy carrier market shares plotted against the energy service price, the wealth proxy, and urbanization for the full market model.

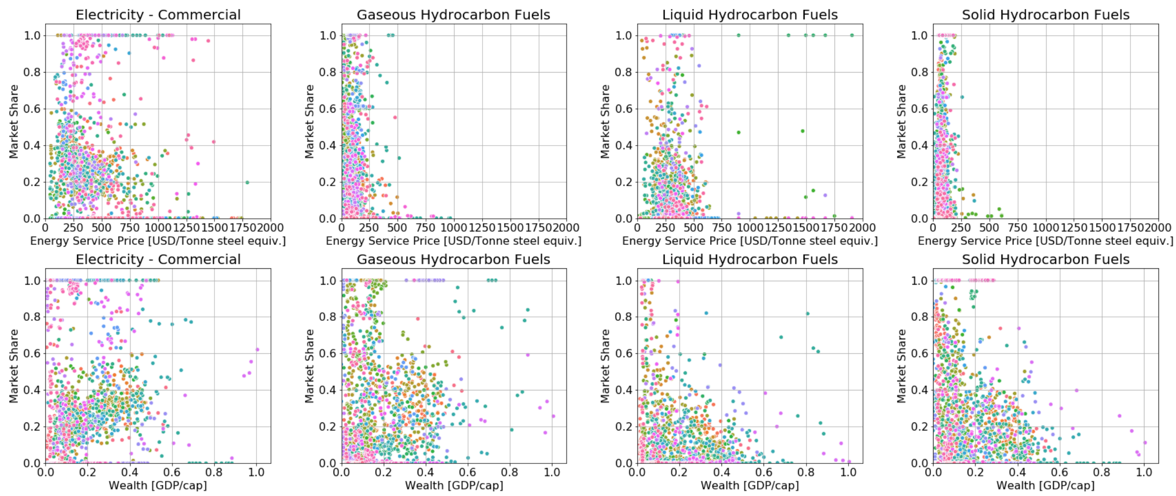

Figure F.10: Visual representation of the feature space for Heavy Industry depicting energy carrier market shares plotted against the energy service price, the wealth proxy, and urbanization for the new and churn market model with a churn rate of 10%.

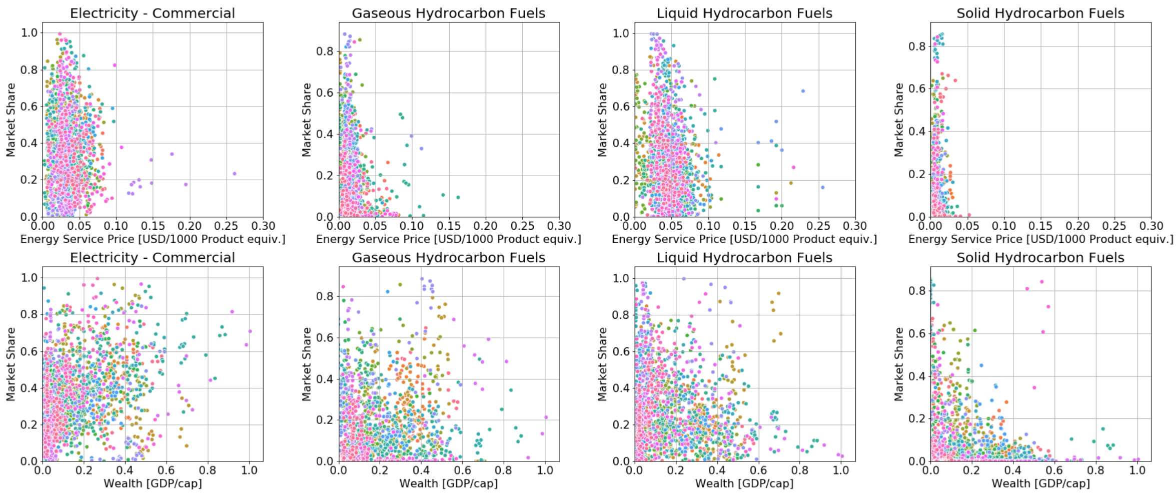

Figure F.11: Visual representation of the feature space for Agriculture & Other Industry depicting energy carrier market shares plotted against the energy service price, the wealth proxy, and urbanization for the new and churn market model with a churn rate of 10%.

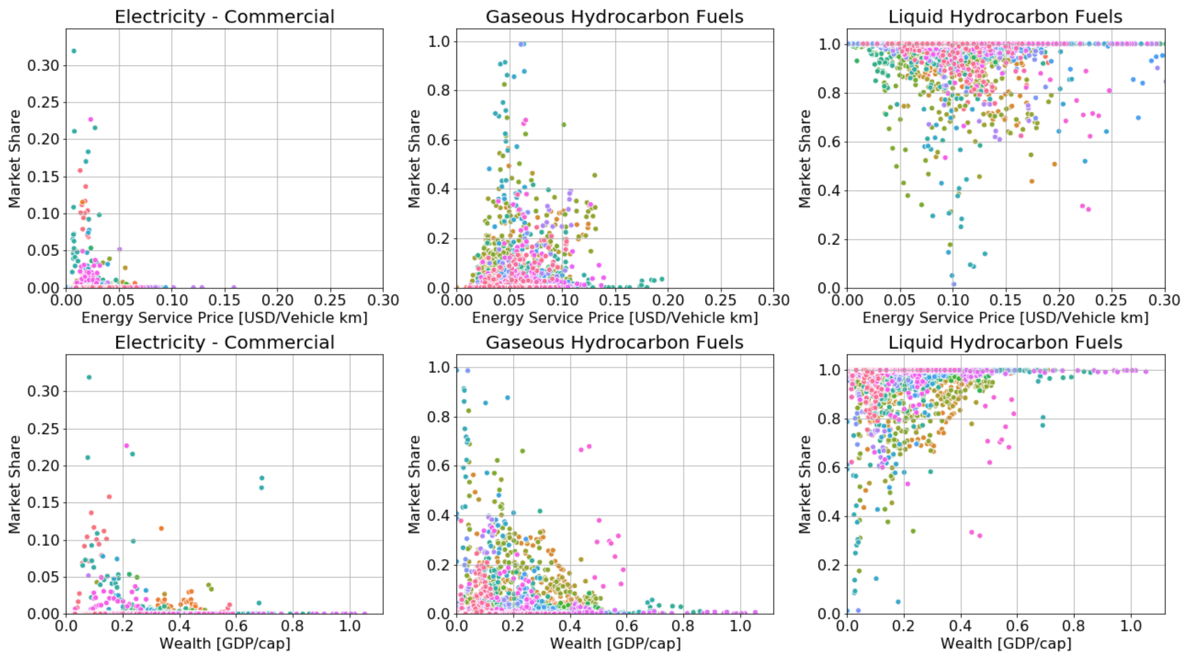

Figure F.12: Visual representation of the feature space for Passenger Transport Road depicting energy carrier market shares plotted against the energy service price, the wealth proxy, and urbanization for the new and churn market model with a churn rate of 15%.

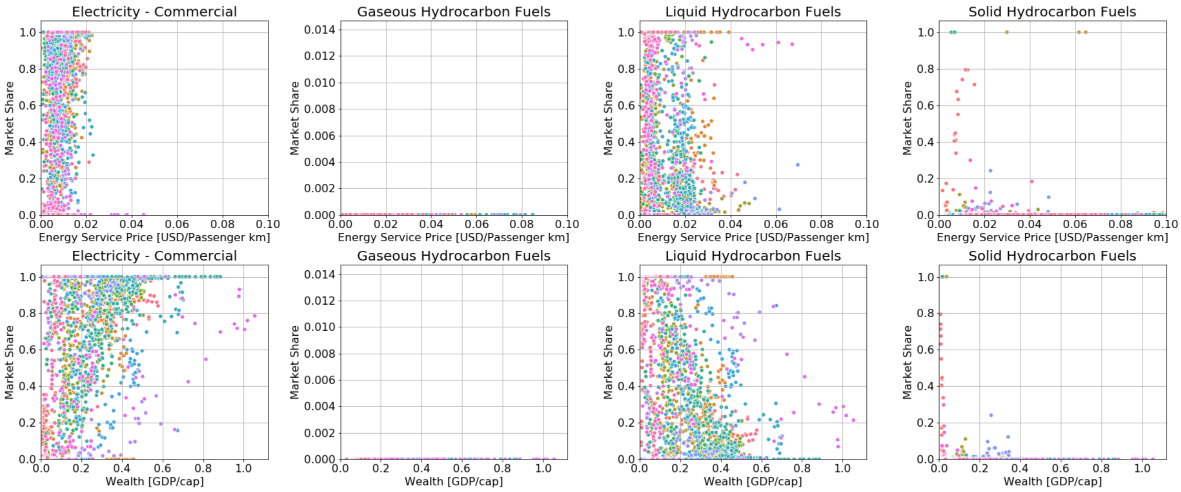

Figure F.13: Visual representation of the feature space for Passenger Transport Rail depicting energy carrier market shares plotted against the energy service price, the wealth proxy, and urbanization for the new and churn market model with a churn rate of 15%.

Figure F.14: Visual representation of the feature space for Passenger Transport Ship depicting energy carrier market shares plotted against the energy service price, the wealth proxy, and urbanization for the new and churn market model with a churn rate of 15%.

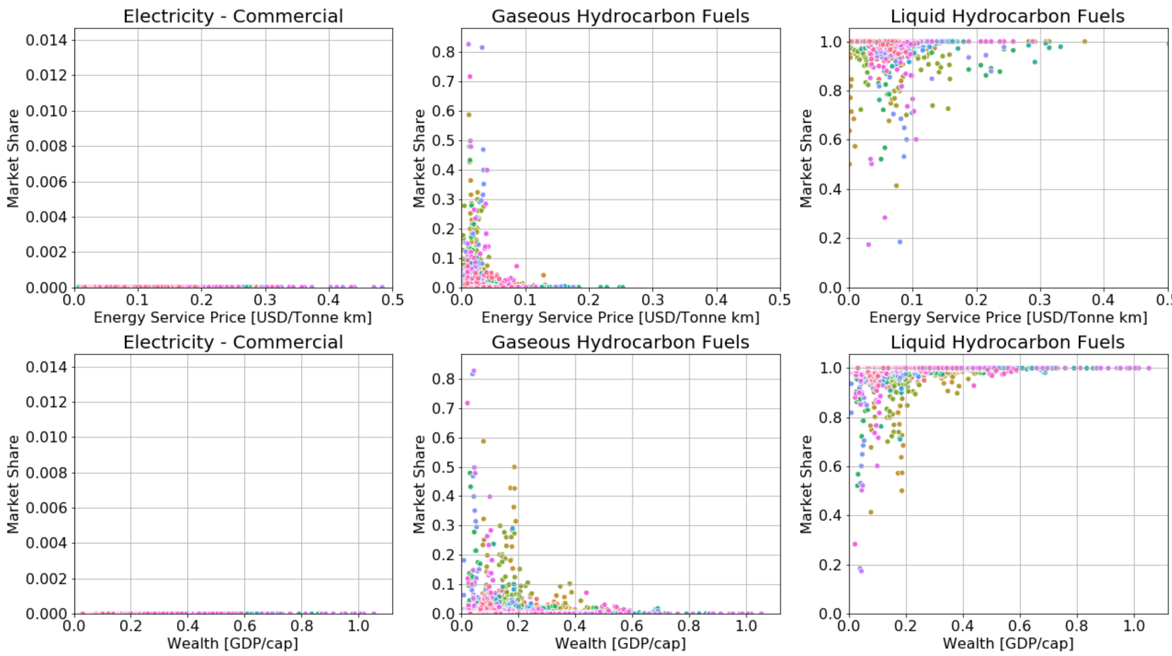

Figure F.15: Visual representation of the feature space for Freight Transport Road depicting energy carrier market shares plotted against the energy service price, the wealth proxy, and urbanization for the new and churn market model with a churn rate of 15%.

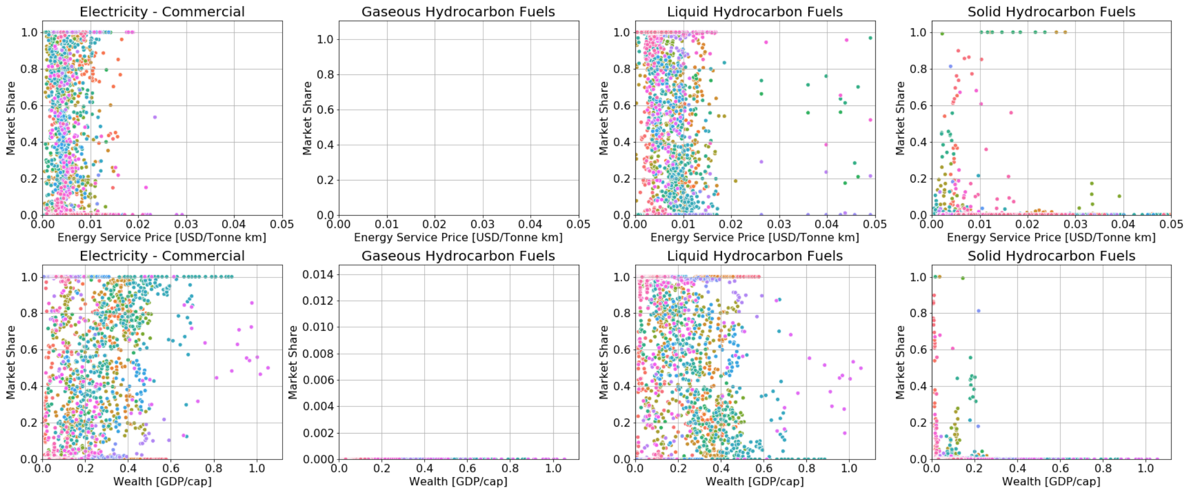

Figure F.16: Visual representation of the feature space for Freight Transport Rail depicting energy carrier market shares plotted against the energy service price, the wealth proxy, and urbanization for the new and churn market model with a churn rate of 15%.

Figure F.17: Visual representation of the feature space for Freight Transport Ship depicting energy carrier market shares plotted against the energy service price, the wealth proxy, and urbanization for the new and churn market model with a churn rate of 15%.

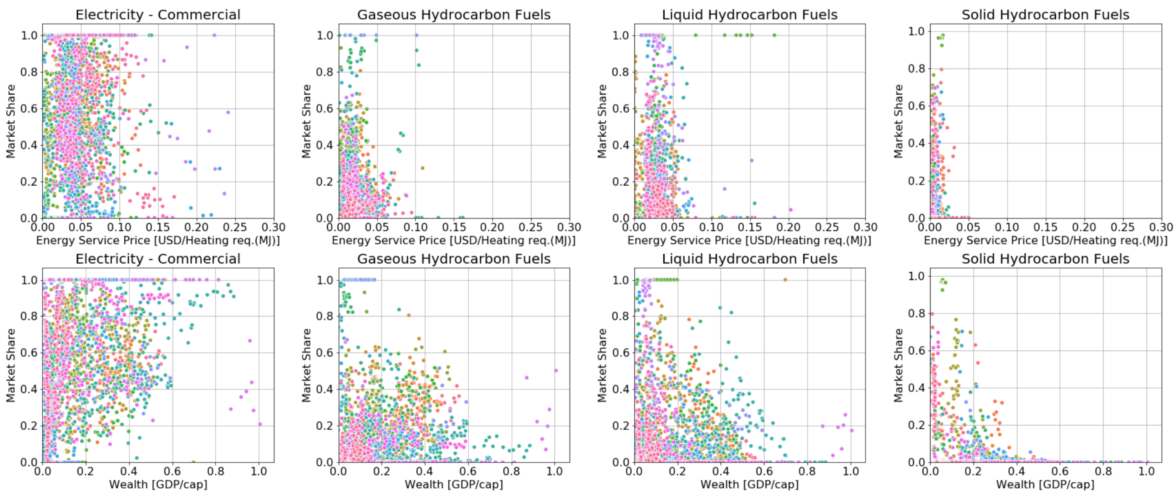

Figure F.18: Visual representation of the feature space for Services depicting energy carrier market shares plotted against the energy service price, the wealth proxy, and urbanization for the new and churn market model with a churn rate of 15%.

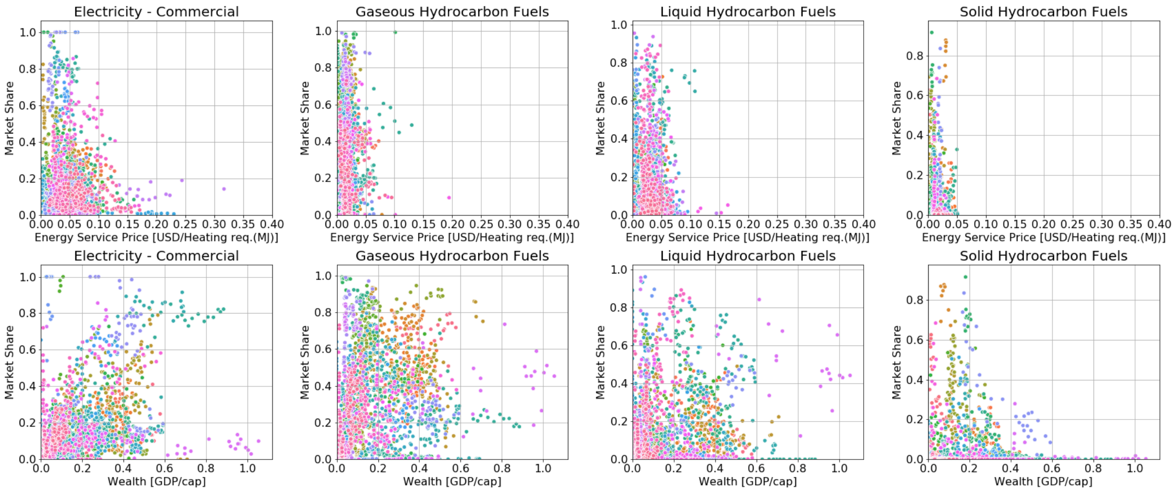

Figure F.19: Visual representation of the feature space for Residential Heating & Cooking depicting energy carrier market shares plotted against the energy service price, the wealth proxy, and urbanization for the new and churn market model with a churn rate of 15%.

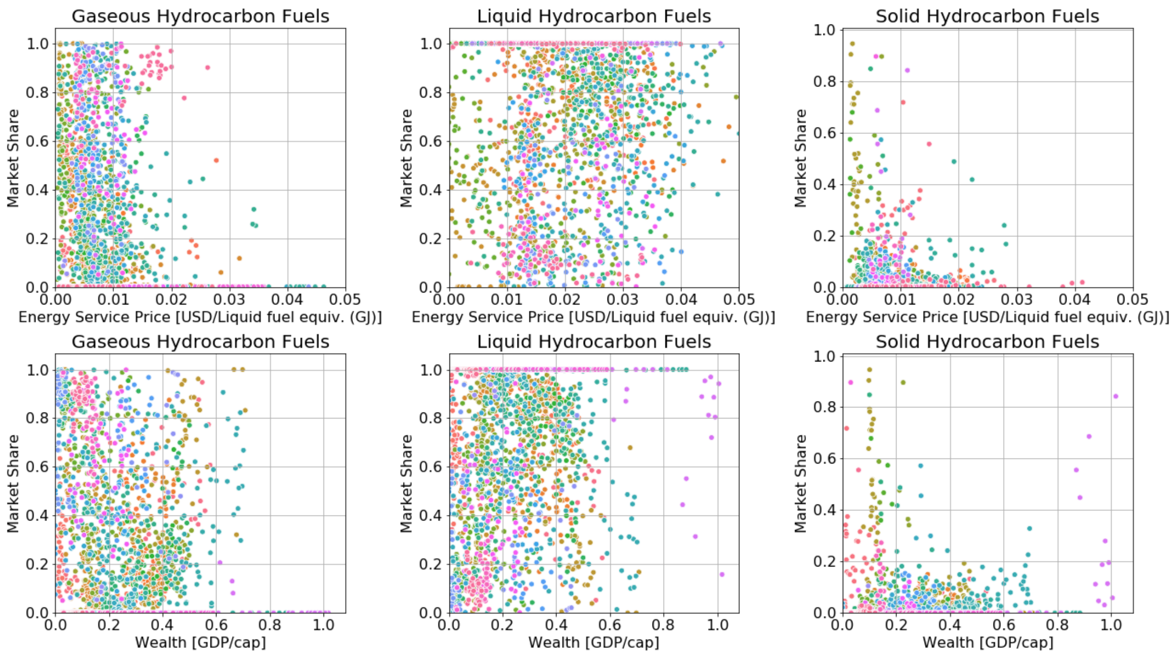

Figure F.20: Visual representation of the feature space for Non Energy Use depicting energy carrier market shares plotted against the energy service price, the wealth proxy, and urbanization for the new and churn market model with a churn rate of 15%.

G

Data distribution plots: Wealth

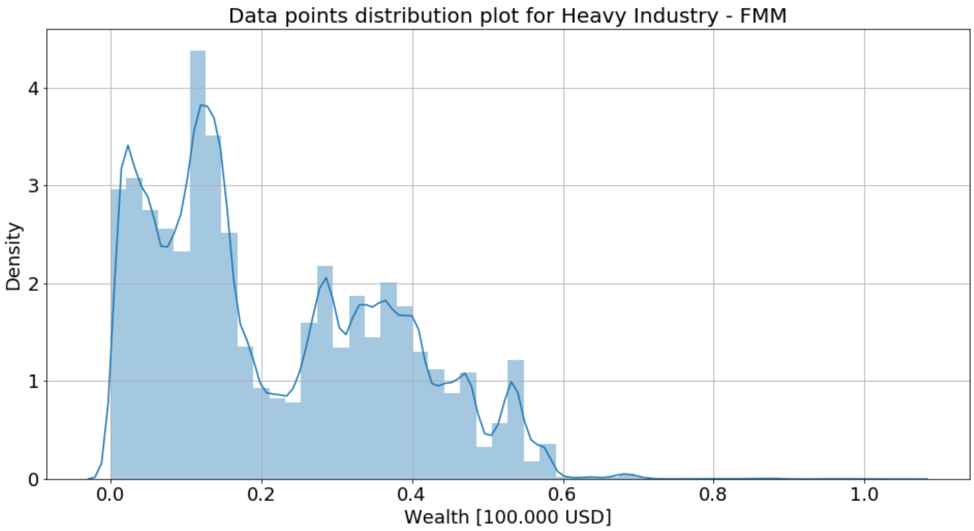

Figure G.1: Visual representation of the distribution of the data points with respect to the Wealth variable for the sector Heavy Industry under the full market model.

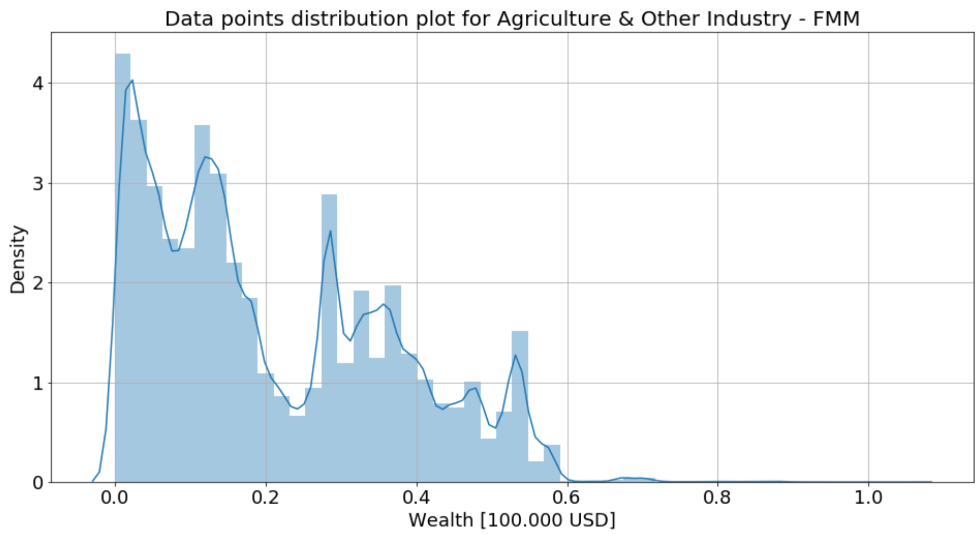

Figure G.2: Visual representation of the distribution of the data points with respect to the Wealth variable for the sector Agriculture & Other Industry under the full market model.

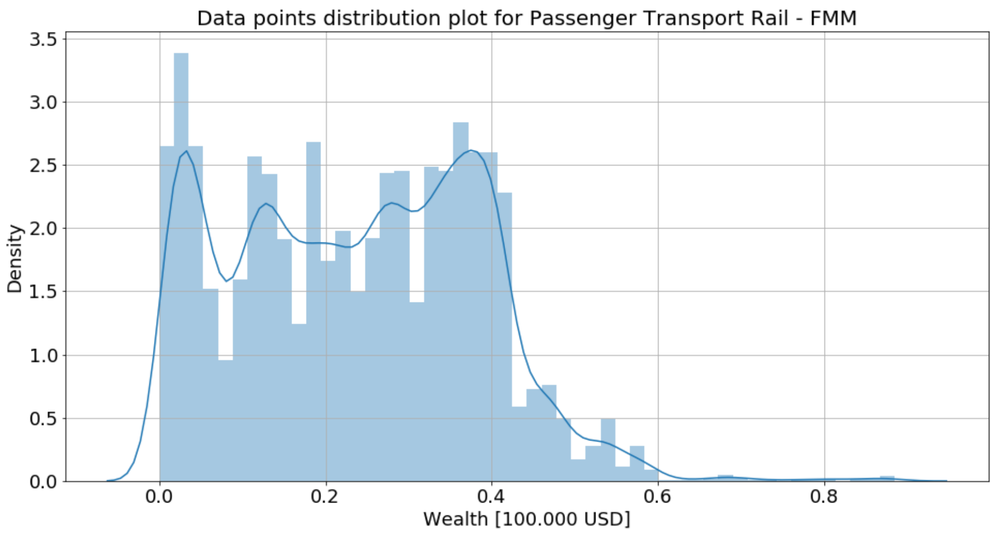

Figure G.3: Visual representation of the distribution of the data points with respect to the Wealth variable for the sector Passenger Transport Rail under the full market model.

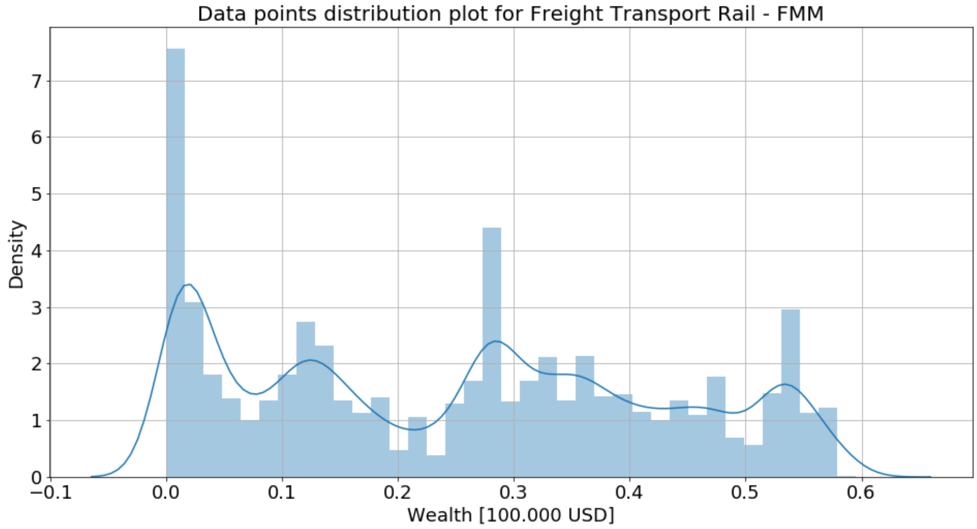

Figure G.4: Visual representation of the distribution of the data points with respect to the Wealth variable for the sector Freight Transport Rail under the full market model.

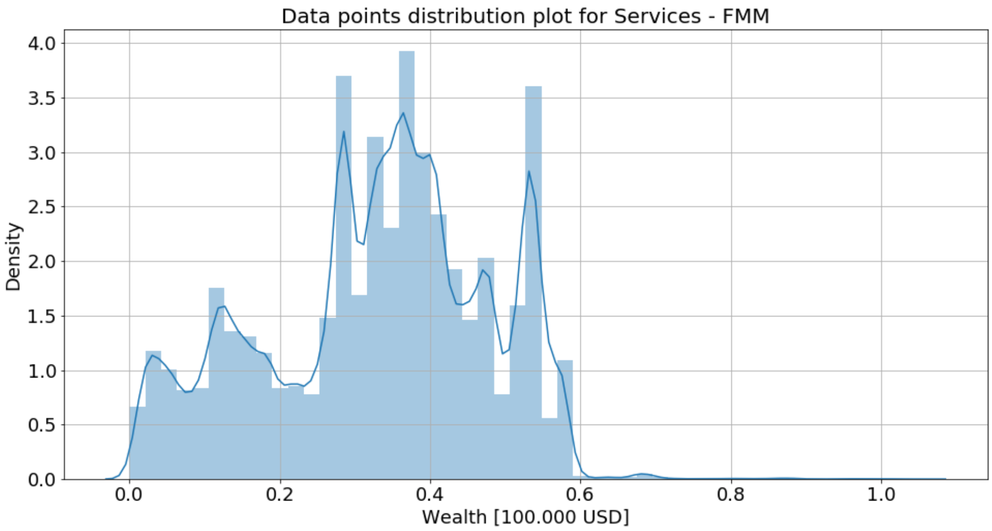

Figure G.5: Visual representation of the distribution of the data points with respect to the Wealth variable for the sector Services under the full market model.

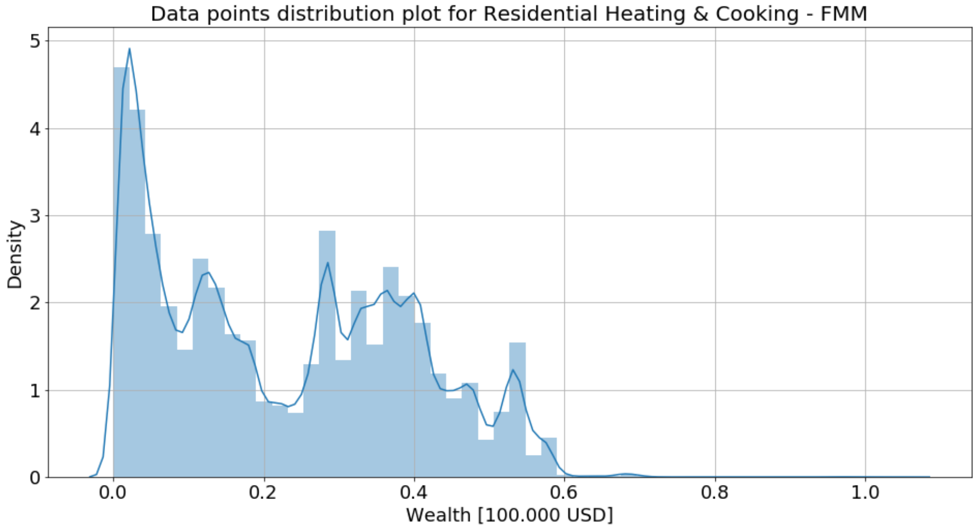

Figure G.6: Visual representation of the distribution of the data points with respect to the Wealth variable for the sector Residential Heating & Cooking under the full market model.

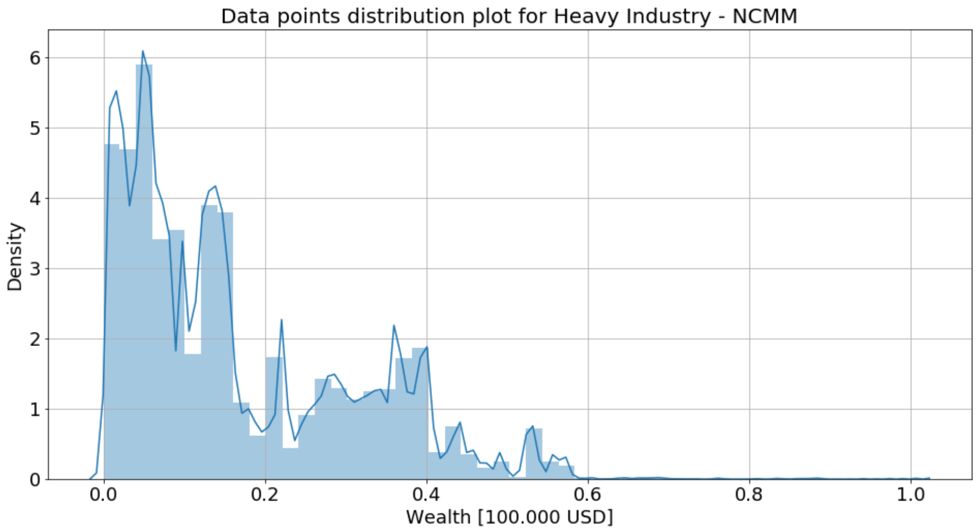

Figure G.7: Visual representation of the distribution of the data points with respect to the Wealth variable for the sector Heavy Industry under the new and churn market model with a churn rate of 0.10.

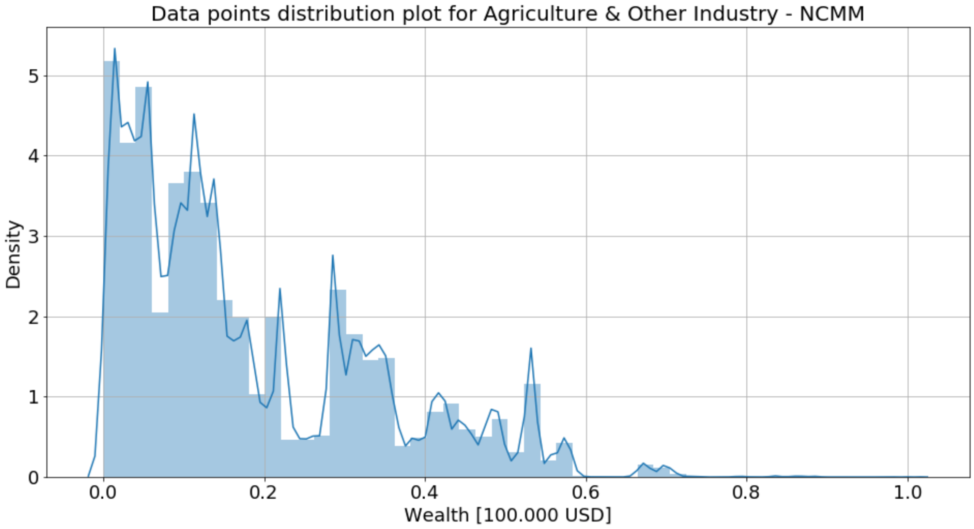

Figure G.8: Visual representation of the distribution of the data points with respect to the Wealth variable for the sector Agriculture & Other Industry under the new and churn market model with a churn rate of 0.10.

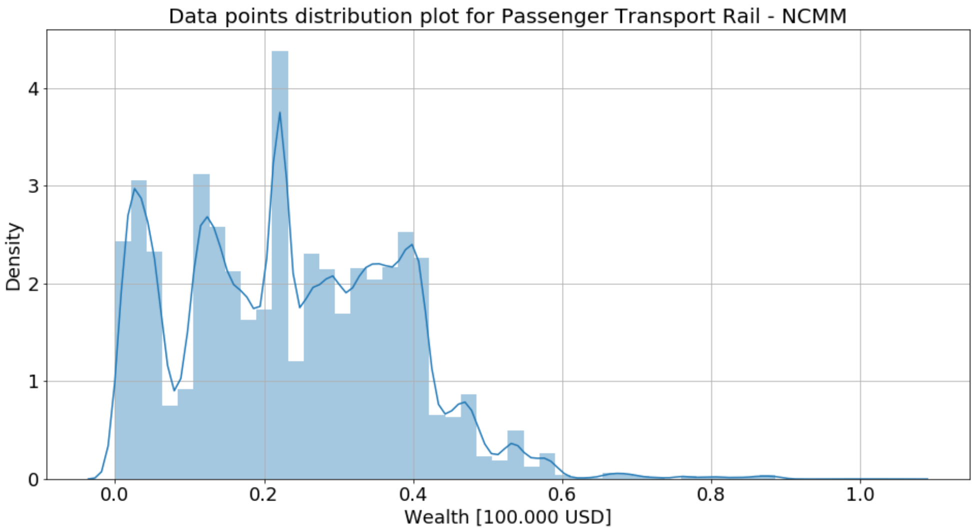

Figure G.9: Visual representation of the distribution of the data points with respect to the Wealth variable for the sector Passenger Transport Rail under the new and churn market model with a churn rate of 0.15.

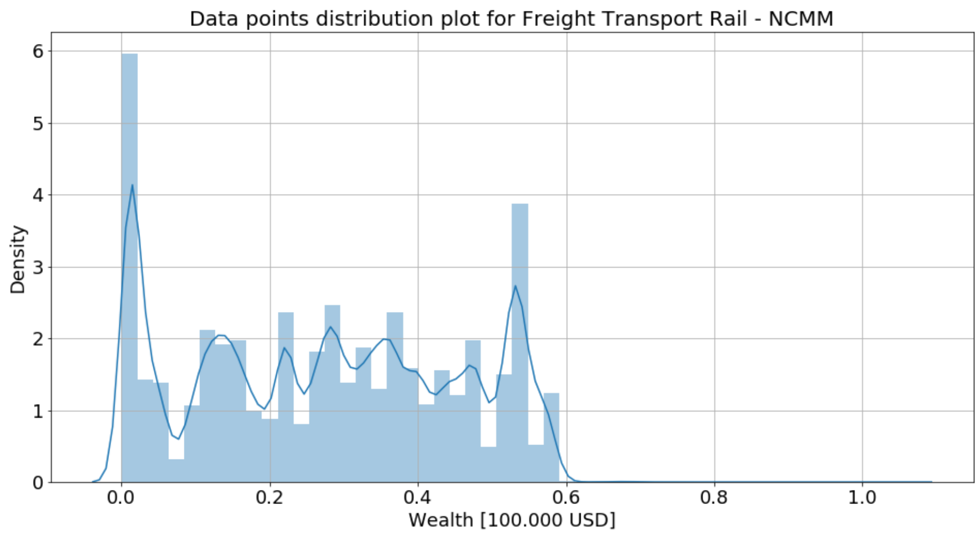

Figure G.10: Visual representation of the distribution of the data points with respect to the Wealth variable for the sector Freight Transport Rail under the new and churn market model with a churn rate of 0.15.

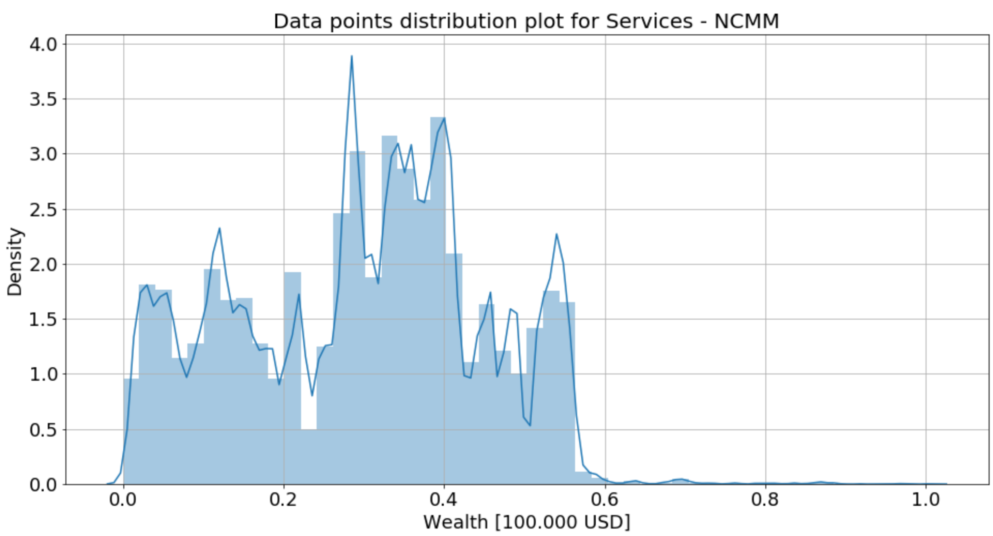

Figure G.11: Visual representation of the distribution of the data points with respect to the Wealth variable for the sector Services under the new and churn market model with a churn rate of 0.15.

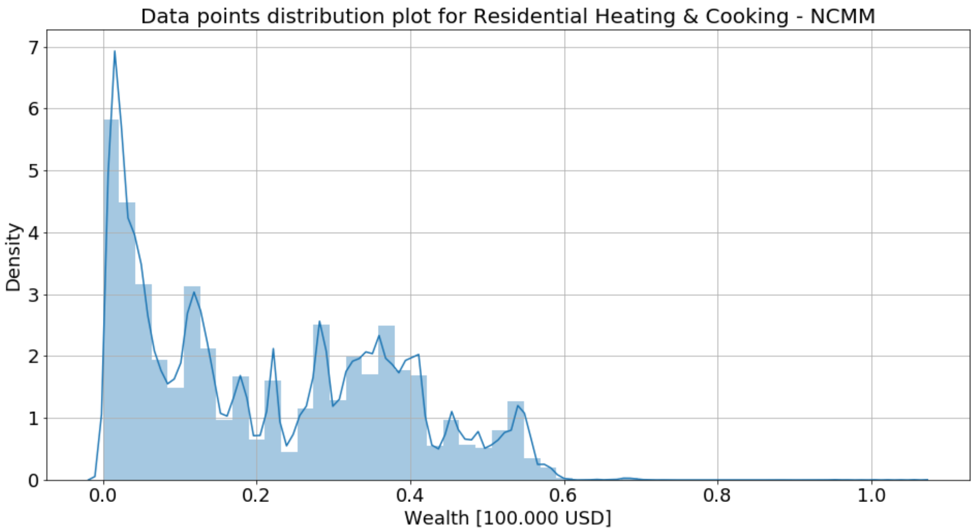

Figure G.12: Visual representation of the distribution of the data points with respect to the Wealth variable for the sector Residential Heating & Cooking under the new and churn market model with a churn rate of 0.15.