Department of Economics

Master Essay I

Have We Built Too Large?

A study of the optimal size of student housing in Lund

Author:
Pontus Josefsson

Supervisor:
Peter Jochumzen

Abstract

This essay investigates what housing size is the optimal for the average student in Lund. The optimal housing size for the average student is defined as the size that maximizes the average student's net benefit from size in monetary terms. It is found by equating the marginal willingness to pay for size and the marginal cost for size. The marginal willingness to pay is derived from estimates of choice models (multinomial logit, nested logit, and random parameters logit) estimated on data sampled from students at Lund University. The marginal cost is derived from estimates of a hedonic regression model estimated on data from AF Bostäder. The optimal housing size estimates for the average student in Lund is in the range 19.0-24.8 square meters. The result can be used when planning new student housing in Lund. The method can be applied to any housing market as long as the estimates from the hedonic regression model can be interpreted as the marginal cost. For this to hold, as in this case, the data must be generated by a housing company that reinvests all profit into the company. Moreover, the method is not only applicable to housing size, but to virtually any housing service there is.

Keywords: housing size; student housing; discrete choice, logit, hedonic regression

Acknowledgements

I would like to thank Peter Jochumzen for his guidance, Fredrik Svensson Rosell for all stimulating discussions, and my family for their support.

Contents

List of Figures

List of Tables

1 Introduction

In Sweden there is a large shortage of housing (Hansson 2014). The market for student housing is not an exception. Between 40-50 percent of the cities with universities/colleges have had a shortage of student housing in the last five years. In some cities it would be sufficient with hundreds of new student housing units (i.e. corridor rooms and/or apartments), while other are in need of thousands of new student housing units (Studentbostadsförtgen 2015). But what kind of housing should be built? That is, what kind of housing do students demand? Considering the range of all possible aspects of housing and the fact that student housing markets are different in different cities, this is a very broad question. In this essay, I therefore focus on one aspect of housing, namely the size, and on one student housing market, namely that in Lund, where at least 1000 new student housing units need to be built (Studentbostadsförtgen 2015). More specifically, I investigate what housing size the average student in Lund prefer. That is, what housing size is the optimal for the average student in Lund? I also contrast the optimal housing size for the average student to the average size of the current housing stock. That is, is the average housing size of the current housing stock the optimal housing size for the average student?

I define the optimal housing size for the average student to be the size that maximizes the average student's net benefit from size in monetary terms. It is found by equating the marginal willingness to pay for size and the marginal cost for size. The marginal willingness to pay is derived from estimates of choice models (multinomial logit, nested logit, and random parameters logit) estimated on data sampled from students at Lund University. The marginal cost is derived from estimates of a hedonic regression model estimated on data from AF Bostäder. I have found optimal housing size estimates for the average student in Lund in the range 19.0-24.8 square meters. The optimal size estimates vary with respect to the age and location of the housing unit. The optimal size is lower for central housing units than for non-central housing units, and also lower for newer housing units than for older housing units. This is due to the fact that the marginal cost estimates are higher for central and newer housing units compared to non-central and older housing units, all other things being equal. The optimal size estimates also differ because the different choice models give different marginal willingness to pay estimates. To examine if the average housing size of the current housing stock is the optimal housing size for the average student, I have carried out Monte Carlo simulations. Based on the results from the simulations I cannot conclude that the average housing size of the current housing stock is not the optimal housing size for the average student.

The remainder of this essay is structured in the following way. Section 2 describes related previous research. Section 3 first describes what is meant by optimal housing size and how it is found by equating the marginal willingness to pay and the marginal cost, and then it describes how the willingness to pay and marginal cost can be derived from the parameters in choice models and hedonic regression models, respectively. Section 4 describes the data used to estimate the choice models and the hedonic regression model. Section 5 describes the choice models. Section 6 presents the results from the choice models and the hedonic regression model, and also the marginal willingness to pay, the marginal cost, and the optimal size estimates. Section 7 summarizes the findings and discusses policy implications and ideas for future research.

2 Previous Research

Student housing size has previously been studied by Nilsson (2015). Nilsson examines attitudes for student housing in Luleå with direct questions such as "would you consider living in an apartment..." and "how much would you be willing to pay for...". However, in contrast to this study, Nilsson does not arrive at an estimate of the optimal size.

To get an estimate of the optimal size I combine two methods that separately have been used in analyzes of the housing market: choice modeling and hedonic price analysis. However, no study that I know of has combined a willingness to pay estimate, derived from a choice model, and a marginal cost estimate, derived from a hedonic regression model, to get an estimate of the optimal housing size on any housing market.

Choice modeling can be used to derive estimates of the marginal willingness to pay. An example in the context of a housing market is a study by Torres, Greene and Ortúzar (2013). Torres, Greene and Ortúzar used choice modeling to estimate potential new residents' willingness to pay for housing and neighbourhood attributes in Santiago, Chile, so that government and private developers could gain knowledge about the potential new residents' preferences. That is, their goal was similar to mine. However, Torres, Greene and Ortúzar did not arrive at an estimate of the optimal level of the housing services.

Hedonic price analysis is used to estimate implicit prices of attributes. It was popularized by Griliches (1961) and the theory was further developed by Rosen (1974). However, the pioneering hedonic price analysis – coining the term "hedonic" – was made by Court (1939) (Griliches 1991, Goodman 1998). Both Court (1939) and Griliches (1961) studied automobile prices, but much of the work have later focused on housing markets (see, for example, Kain and Quigley 1970, and Wilhelmsson 2000). The estimates of implicit prices are usually interpreted as marginal willingness to pay for attributes. For example, Wilhelmsson (2000) estimates the marginal willingness to pay for a reduction of noise. However, in this essay I interpret the estimates of implicit prices as the marginal cost for attributes. I do this because the data is from the students' own housing company in Lund, AF Bostäder, which is a foundation that that reinvests all profit into the company. Since all profit is reinvested into the company, I do not expect there to be a discrepancy between the rent tenants pay and the housing company's cost.

Choice modeling is used to model decision processes. An essential toolkit for this is the logit familiy of models (Hensher and Greene 2003). The logit models used in this essay are the multinomial logit model, the nested logit model and the random parameters logit model. The multinomial logit model is the most commonly used discrete choice model. The reason for this is that its formula has a closed form solution (Train 2002). It was first derived by Luce (1959) from the assumption of "independence from irrelevant alternatives," which implies proportional substitution across alternatives. Marschak (1960) proved that it is consistent with utility maximization. Luce and Suppes (1965) then showed that it also can be derived by assuming extreme value distributed unobserved utility. The analysis was completed by McFadden (1974) who proved the converse: the logit model for choice probabilities necessarily implies that the unobserved utility follow an extreme value distribution. However, although the multinomial logit model is popular due to its closed form solution, it is inappropriate in situations where "independence from irrelevant alternatives" is violated. This was pointed out by Chipman (1960) and Debreu (1960).

When "independence from irrelevant alternatives" is violated more general models are needed, such as the nested logit model and the random paramers logit model. The nested logit model has long been the main modeling tool when a more sophisticated analysis is required (Hensher and Greene 2003). Like the multinomial logit model, it is consistent with utility maximization as McFadden (1978) showed. The random parameters logit model is an alternative to the nested logit model. McFadden and Train (2000) showed that it is so general that it, to any degree of accuracy, can approximate any discrete choice model derived from utility maximization. However, the random parameters logit model does not have a closed form solution, and it thus relies on simulation methods. The full power of the random parameters logit could therefore not be utilized until the late 1990's (see, for example, Bhat 1998, and Brownstone and Train 1999), after the development of simulated maximum likelihood estimation by Börsch-Supan and Hajvassiliou (1993), and Hajivassiliou and Ruud (1994), among others, that enabled estimation of open-form models (see, for example, Hensher and Greene 2003).

3 Theory

3.1 Optimal Housing Size

In order to determine what housing size is optimal for the average student, it is first necessary to define what is meant by optimal housing size. For the average student, I define the optimal housing size to be the size that maximizes the individual's net benefit in monetary terms. The net benefit is calculated by subtracting the total cost for size from the total willingness to pay for size. I assume that cost, as a function of size, is linear over the interval that I consider in this essay. I also assume that individuals' total willingness to pay for size is increasing at a diminishing rate, so that the total willingness to pay, as a function of size, is concave. Figure 3.1 depicts a total cost curve and a total willingness to pay curve, representing these assumptions.

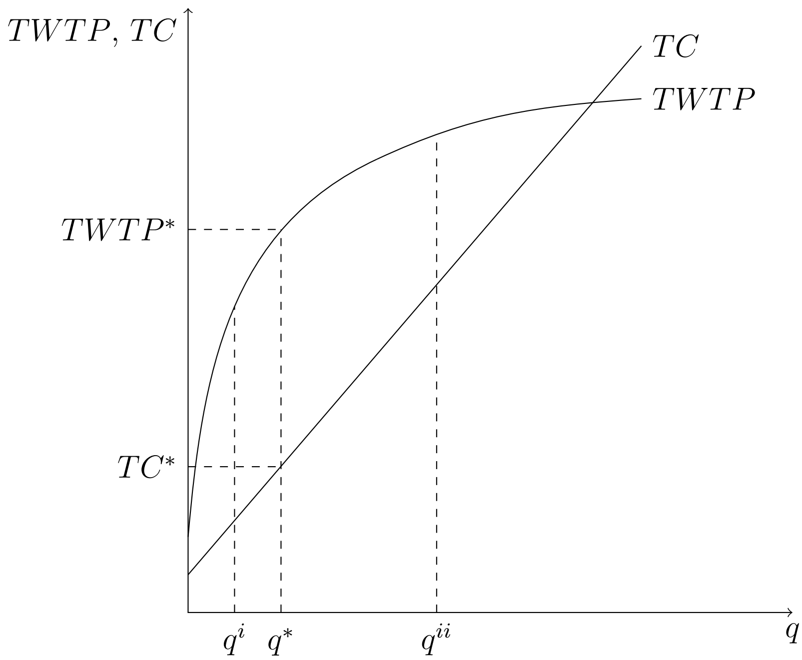

Figure 3.1: The optimal size, denoted q^*, maximizes the net benefit. It is found by equating the marginal willingness to pay and the marginal cost.

The optimal size, denoted q^*, is the size that maximizes the net benefit. In the figure, the net benefit is represented by the distance between the total willingness to pay curve (TWTP) and the total cost curve (TC). This distance is maximized at q^*, where the marginal willingness to pay is equal to the marginal cost. The marginal willingness to pay is represented by the slope of the total willingness to pay curve and the marginal cost is represented by the slope of the total cost curve. At q^* these curves have the same slope. At q^i, on the other hand, the slope of the total willingness to pay curve is steeper than the slope of the total cost curve. Thus, the marginal willingness to pay is higher than the marginal cost. This implies that the individual can get better off by enjoying a larger size at the price of a higher cost. Similarly, at q^{ii} the slope of the total willingness to pay curve is less steep than the slope of the total cost curve. Thus, the marginal willingness to pay is lower than the marginal cost. This implies that the individual can get better off by enjoying a smaller size in exchange for a lower cost.

Thus, to find an estimate of the optimal housing size for the average student I simply equate estimates of the average marginal willingness to pay for size and and the average marginal cost for size. Then, to find out whether or not the average housing size of the current housing stock is the optimal housing size for the average student I compare the estimate of the optimal size for the average student to the average housing size of the current housing stock.

In the next section, I describe the theoretical meaning of marginal willingness to pay and how I will use choice modeling to estimate it. Thereafter, in Section 3.3, I describe how I will estimate the marginal cost by taking a hedonic pricing approach.

3.2 Marginal Willingness to Pay

One feature of housing that makes it different from other goods is that the housing stock is heterogenous. Each housing unit offers a bundle of housing services, i.e. attributes, that makes it different from other housing units. For example, housing units differ in size, layout, style, quality of the interior and the exterior, utilities (heating and electrical), and neighborhood (accessibility to jobs, education- and social opportunities; local public goods and taxes; environmental quality) (O'Sullivan 2012). To me, it therefore seems implausible to assume that all housing units give an individual the same level of utility. Instead, I will assume that the level of utility an individual gets from a housing unit, and consequently the individual's willingness to pay for a housing unit, depends on the quality of the housing services it offers.

An econometric method for modeling individuals' choices between alternatives, based on this way of viewing goods, is choice modeling. In choice modeling, goods are treated as the embodiment of a bundle of attributes, and individuals are assumed to choose the bundle of attributes that gives them the highest level of utility. The effect the different attributes have on utility is captured by the estimated parameters, and from them it is possible to derive estimates of the average marginal willingness to pay for the attributes. That is, I can use choice modeling to derive estimates of the average marginal willingness to pay for size. However, before I describe how choice modeling can be used to derive estimates of the average marginal willingness to pay for a size, I describe the theoretical meaning of willingness to pay for quality.

Compensating variation, which is the conventional welfare measure for price changes (Hicks 1939. See, for example, Mas-Colell, Whinston and Green 1995), can be extended to also measure the willingness to pay for quality changes (Mäler 1974). Following Hanemann (1991), this can be shown assuming an individual who has preferences over a finite number L of goods. The individual's consumption level of the goods are represented by a consumption vector

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_L \end{bmatrix},$$

which lists the amount consumed of each of the L goods. The individual also has preferences over the quality of the goods. For simplicity it is assumed that the quality level for all goods but one remain constant throughout the analysis. Thus, the quality level for the good of focus can be denoted by a scalar q. The quality level is assumed to be fixed exogenously, in contrast to the consumption levels which the individual chooses freely subject to a budget constraint.

The budget constraint is simply that consumption is restricted to those consumption vectors that the individual can afford. This is determined by prices and the wealth of the individual. The prices of the L commodities are represented by a price vector

$$p = \begin{bmatrix} p_1 \\ \vdots \\ p_L \end{bmatrix},$$

which lists the unit cost for each of the L commodities. Here it is assumed that the individual cannot influence the price level for any of the commodities. The individual's wealth is denoted by w. The individual can afford a consumption vector x if the wealth w is not exceeded by the total cost of that consumption vector

$$p \cdot x = p_1 x_1 + \cdots + p_L x_L \leq w.$$

Thus, the problem that the individual faces is to choose the most preferred consumption vector x subject to this budget constraint.

What consumption vector is the most preferred depends on the individual's preferences. Assume that the individual's preferences can be represented by a continuous, differentiable utility function $u(x, q)$. More specifically, assume that the individuals preferences are monotone and convex. Monotone preferences means that the individual prefers larger amounts of commodities to smaller amounts, and a higher quality level to a lower quality level. This implies that the utility function is increasing. Convex preferences means that the individual require increasingly larger amounts of one commodity in compensation for giving up successive units of another commodity. This implies that the utility function is quasiconcave. With this representation of the individual's utility, and given prices $p \gg 0$ (i.e. $p_\ell > 0$ for all $\ell = 1, \ldots, L$) and wealth $w > 0$, the problem of choosing the most preferred consumption vector x can now be expressed as the following utility maximization problem

$$\begin{aligned} \max_x \quad & u(x, q) \\ \text{s.t.} \quad & p \cdot x \leq w. \end{aligned}$$

The solution to the utility maximization problem is the individual's set of optimal consumption vectors, denoted by $x(p, q, w)$. The resulting utility value is given by the indirect utility function $v(p, q, w) = u[x(p, q, w), q]$, for each q and $(p, w) \gg 0$. Thus, the indirect utility is equal to $u(x^*, q)$ for any optimal consumption vector $x^* \in x(p, q, w)$.

Now, suppose that the quality increases from q^0 to $q^1 > q^0$, while prices and wealth remain constant at $(p, w) \gg 0$. The quality increase results in a utility increase from utility level $u^0 = v(p, q^0, w)$ to $u^1 = v(p, q^1, w)$. Thus, the welfare change from the quality increase in terms of utility is given by $v(p, q^1, w) - v(p, q^0, w)$. The measure of the welfare change in monetary terms is given by the compensating surplus, which is defined by

$$v(p, q^1, w - CS) = v(p, q^0, w).$$

The compensating surplus, here denoted CS, is the amount of money that the individual can give up and still maintain the initial utility level, if at the same time the quality is increased from q^0 to q^1. This can also be shown by a graphical analysis.

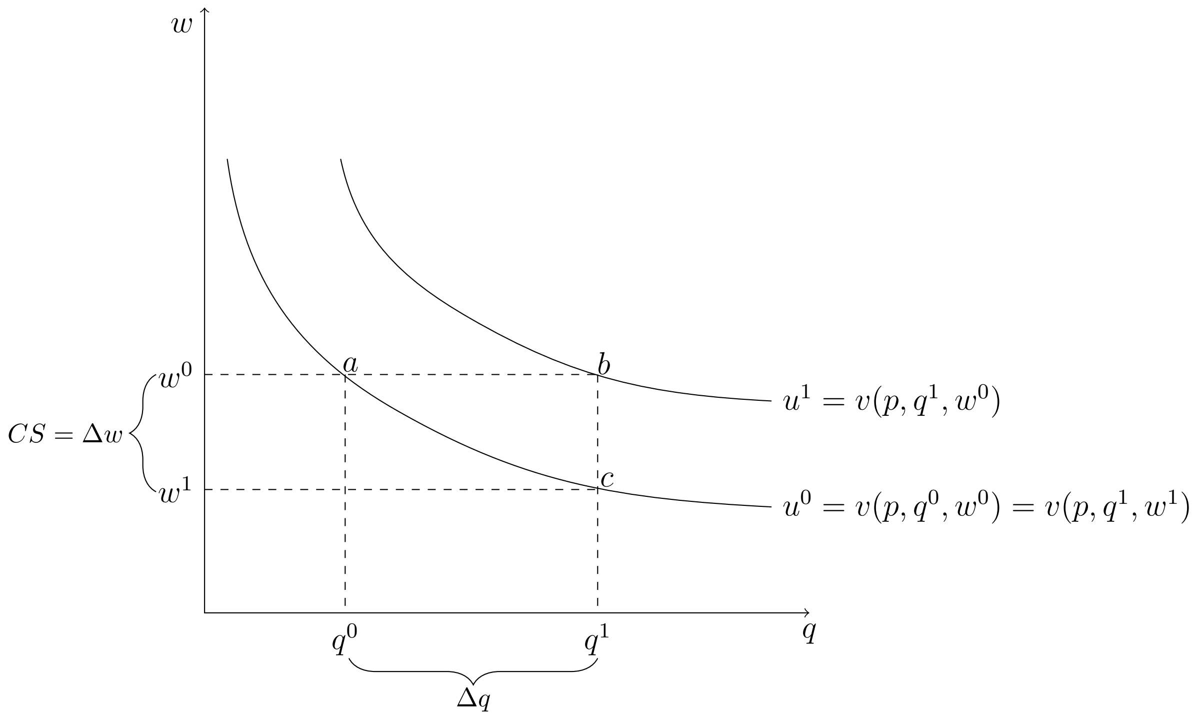

Figure 3.2: The willingness to pay for a quality increase from q^0 to q^1 is equal to $w^0 - w^1$.

Initially the price vector is p, wealth[1] is w^0, the quality is q^0, and the individual's resulting utility level is $u^0 = v(p, q^0, w^0)$. This is represented by point a in Figure 3.2. When the quality increases from q^0 to q^1, while prices and wealth remain constant at (p, w^0), the individual's utility increases and reaches utility level $u^1 = v(p, q^1, w^0)$. This is represented by point b. If instead the wealth were reduced by $w^0 - w^1$ at the same time as the quality increased from q^0 to q^1, the individual would maintain the initial utility level. This is represented by point c. Thus, $w^0 - w^1$ is the amount of money that the individual can give up and still maintain the initial utility level, if at the same time the quality is increased from q^0 to q^1. This is the compensating surplus, which in the case of a quality increase is the individual's willingness to pay. Of course, the individual would be willing to pay an amount less than $w^0 - w^1$ to enjoy the quality increase, as this would result in a utility net increase. However, the individual would not be willing to pay an amount higher than $w^0 - w^1$ to enjoy the the quality increase, as this would result in a utility net decrease. Thus, $w^0 - w^1$ is the individual's (maximum) willingness to pay for the quality increase from q^0 to q^1.

Until now, nothing have been said about the magnitude of the quality change. The magnitude of the increase has simply been assumed to be arbitrary and equal to Δq. However, remember from Section 3.1 that the optimal size is found by equating the

[1] Here the initial wealth is denoted w^0 instead of w.

marginal willingness to pay and the marginal cost. Thus, for the purpose of this essay, it is necessary to consider the marginal willingness to pay, that is, the willingness to pay for arbitrarily small quality increases.

The marginal willingness to pay is given by the amount of money dw that the individual can give up and still maintain the initial utility level, if at the same time the quality is increased by an arbitrarily small amount equal to dq. This can be written as

$$\frac{\partial v(p,q,w)}{\partial q}dq + \frac{\partial v(p,q,w)}{\partial w}dw = 0,$$

which in turn can be rewritten as

$$\frac{dw}{dq} = -\frac{\partial v(p,q,w)/\partial q}{\partial v(p,q,w)/\partial w}. \tag{3.1}$$

This is negative since both the marginal utility of quality and marginal utility of wealth is positive: $dv/dq > 0$, $dv/dw > 0$. Thus, the individual can give up money and still maintain the initial level of utility. This amount is the marginal willingness to pay. I now turn to describe how choice modeling can be used to derive estimates of the average marginal willingness to pay for a quality change.

Choice models can be derived from utility maximizing behavior by using a random utility model. This will be described later on in Section 5.1. In a random utility model the utility U an individual gets from an alternative is given by

$$U = V + \varepsilon,$$

where V is a component that depends on observed factors and ε is a component that depends on unobserved factors. The component that depends on observed factors can be specified to be either linear or non-linear in parameters. However, throughout this essay the parameters will enter linearly: $V = \boldsymbol{x}'\boldsymbol{\beta}$, where $\boldsymbol{x}$ is a vector of observed factors related to the alternative for the individual, and $\boldsymbol{\beta}$ is a vector of parameters. It is from the estimates of these parameters that the estimate of the marginal willingness to pay is derived (see, for example, Cameron and Trivedi 2005, and Train 2003).

As an example, consider the following observed utility component

$$V = \beta_0 p + \beta_1 q + \beta_2 q^2,$$

where q is a quality attribute, and p is the price. I assume that the utility decreases as the price increases, and increases as the quality attribute increases, but at a diminishing rate: $\beta_0 < 0,\ \beta_1 > 0,\ \beta_2 < 0$. The sum of ratios $-\beta_1/\beta_0 - 2(\beta_2/\beta_0)q$ represents the individual's marginal willingness to pay for an increase in the quality attribute. To see this, first remember that the marginal willingness to pay is given by the amount of money that the individual can give up and still maintain the initial utility level, if at the same time the quality is increased by an arbitrarily small amount. Then, consider an increase in the quality attribute q and in the price p such that the observed utility does not change

$$\frac{\partial V}{\partial p}dp + \frac{\partial V}{\partial q}dq = 0.$$

This can be rewritten as

$$\frac{dp}{dq} = -\frac{\beta_1}{\beta_0} - 2\frac{\beta_2}{\beta_0}q. \tag{3.2}$$

This is positive and increasing at a diminishing rate over the relevant quality interval: the first term is positive ($\beta_0 < 0,\ \beta_1 > 0$), and the second term is negative ($\beta_0 < 0,\ \beta_2 < 0$) and a function of the quality (q).

This result is equivalent to the one in (3.1). However, the expression in (3.1) is negative because it expresses the willingness to pay as a decrease in wealth. The expression in (3.2), on the other hand, is positive since it expresses the willingness to pay as an increase in price (a higher price means less money for consumption of other goods). Note also that the expression for the willingness to pay in (3.2) represents the marginal willingness to pay curve in Figure 3.1.

3.3 Marginal Cost

To estimate the cost I take a hedonic pricing approach. As was pointed out before, a housing unit consists of a bundle of attributes, i.e. housing services, such as size, layout, style, quality of the interior and the exterior, and neighborhood etc (see, for example, O'Sullivan 2012). Hedonic pricing can be used to estimate the relationship between the attributes and the price. The estimates give implicit prices of attributes, which usually are interpreted as the marginal willingness to pay for attributes. However, instead of interpreting the estimates as the marginal willingness to pay, I will interpret them as the marginal cost. I do this because the data is generated under other circumstances than what is usually assumed in the standard hedonic pricing theoretical framework. For example, Rosen (1974) considers the case of pure competition. The data I use, however, is from the students' own housing company in Lund, AF Bostäder, a foundation which "according to its regulations has a function to own and manage buildings intended as inexpensive and suitable housing for active students who are members of Akademiska Föreningen (the Academic Society) in Lund" (AF Bostäder 2015b). AF Bostäder reinvests all profit into the company (Studentlund 2015). I therefore interpret the estimates of the implicit prices as the marginal cost.

The implicit prices are estimated using a hedonic regression model

$$p_i = z_i'\beta + \varepsilon_i, \quad i = 1, 2, \ldots, n.$$

Here p_i is the observed price, $z_i = (z_{i1}, \ldots, z_{ik})'$ is a vector of observed housing services, $\beta = (\beta_1, \ldots, \beta_k)'$ is a vector of coefficients, and ε_i is an error term for the ith of the n observations. The implicit prices are given by the coefficients β since

$$\frac{\partial E[p_i|z_i]}{\partial z_{ij}} = \beta_j.$$

Holding everything else constant, the coefficient β_j is the expected value of the change in the price from a marginal increase in the jth housing attribute. Of course, this requires that the explanatory variables are exogenous, that is, the expected value of the error term ε_i given all observed housing services in z_i is zero: $E[\varepsilon_i|z_i] = 0$.

4 Data

The data I use to estimate the marginal cost with the hedonic regression model is collected from AF Bostäder. That data is described in section 4.1. The data I use to estimate the marginal willingness to pay with the choice models is collected from students at Lund University. That data is described in section 4.2

4.1 Hedonic Regression Model Data

The data I use to estimate the marginal cost in the hedonic regression model is collected from AF Bostäder. AF Bostäder is the students' own housing company in Lund. At their website all currently available housing units are presented (AF Bostäder 2015a). I have collected a sample of 246 housing units. Among these 246 housing units, there are 176 apartments and 70 corridor rooms. The most frequent housing type is the single-room apartment and the second most frequent is the corridor room. The different housing types and their frequencies in the sample are shown in a bar chart in Figure 4.1.

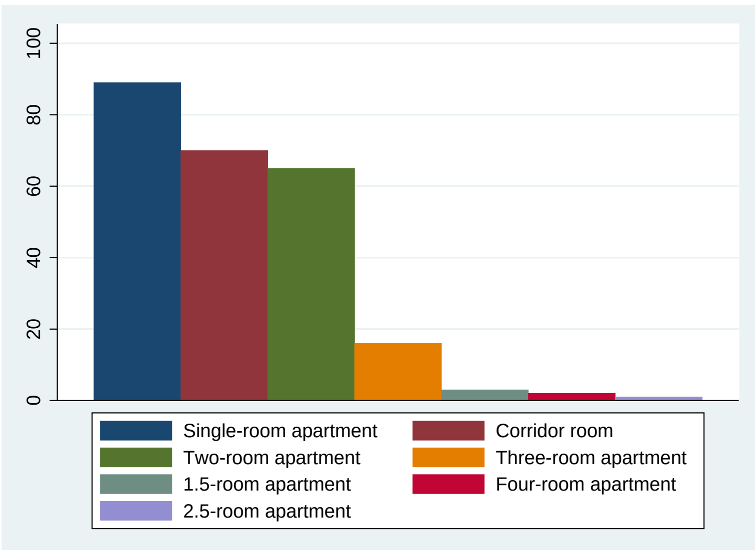

Figure 4.1: The collected sample contain 89 single-room apartments, 70 corridor rooms, 65 two-room apartments, 16 three-room apartments, 3 1.5-room apartments, 2 four-room apartments and 1 2.5-room apartment.

The housing units are described by a set of attributes. In the following, I describe the attribute monthly rent and attributes that I have found affect monthly rent. Since my goal with the hedonic regression model is to estimate the marginal cost of size, the emphasis is on the description of monthly rent and size. The other attributes are used as control variables and in interaction terms with size. They are therefore given a more brief treatment. I first describe monthly rent and I then move on to describe size, and the attributes used as control variables and in interaction terms.

The monthly rent[1], in SEK, varies between about 2100 to 9100. The median is 3654, which is almost 500 less than the mean 4140. These statistics are presented in Table 4.1, and the distribution is shown in a histogram in Figure 4.2. Clearly, lower rents are more common than higher rents. This reflect the frequencies of different housing types in the sample. Table 4.1 also shows that, for single-room and 2-room apartments, the mean monthly rent is more than 900 higher for newer housing units than for older housing units. Here older housing units are defined as housing units built in the years 1960-1997 and newer housing units are defined as housing units built in the years 2004–2015. The effect of the age of housing units on monthly rent is discussed further after the description of the attribute size, to which I now turn.

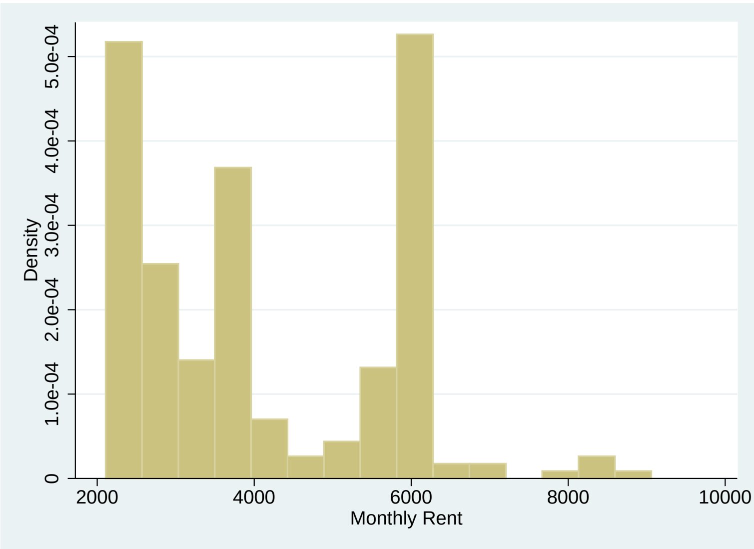

Figure 4.2: The monthly rent, in SEK, varies between about 2100 and 9100. The median is 3654, the mean is 4140 and the standard deviation is 1657.

[1]For some apartments and corridor rooms the payment period is not 12 but 9 months. That is, there is no rent during June, July and August. I have therefore transformed the monthly rent for the apartments and corridor rooms with a 9 months payment period into a 12 months payment period. See Table 4.8 for more details.

Table 4.1: Monthly Rent

Housing Type	Older Housing Units				Newer Housing Units				Full Sample			
	N	$\bar{x}$	s	md	N	$\bar{x}$	s	md	N	$\bar{x}$	s	md
Single-room apartment	35	2976	416	2879	54	3896	572	3654	89	3534	684	3654
1.5-room apartment	3	4022	729	4083					3	4022	729	4083
2-room apartment	12	5305	314	5333	53	6217	76	6224	65	6049	386	6217
2.5-room apartment	1	6791	0	6791					1	6791	0	6791
3-room apartment	15	6575	1046	6039	1	9059	0	9059	16	6730	1186	6039
4-room apartment	2	6395	210	6395					2	6395	210	6395
Corridor room	70	2449	279	2362					70	2449	279	2362

Notes to Table 4.1: The table shows the number of observations (N), mean ($\bar{x}$), standard deviation (s) and median (md). The statistics are given for older and newer housing units, and for the full sample. The monthly rent is measured in SEK. For the full sample, the mean is 4140, the median is 3654, and the standard deviation is 1657.

The size, in square meters, varies between about 17 to 94. The median is 22.2, which is about 10 less than the mean 30.4. These statistics are presented in Table 4.3, and the distribution is shown in a histogram in Figure 4.3. Clearly, smaller housing units are more common than larger housing units. Again, this reflect the frequencies of different housing types in the sample. Table 4.3 also show that for 2-room apartments the mean size is more than 13 higher for older housing units than for newer housing units. For single-room apartments, on the other hand, the pattern is reversed. However, in that case, the difference is only 0.5.

I have found that the effect of size on monthly rent varies with the location and age of the housing unit. I therefore use variables describing the location and year of construction in interaction terms with size. AF Bostäder provides information on the distance to the city centre from the different residential areas. According to AF Bostäder, the distance to the city centre from their residential areas is in the range 0.4–2.6 km. In the analysis, I use a dummy variable for location in an interaction term with size. I categorize the residential areas as central if the distance to the city centre is in the range 0.4–1.0 km.

AF Bostäder also provides information on when the different residential areas were built. The year of construction varies in the range 1960-2015. In the analysis, I use a dummy variable for year of construction in an interaction term with size. I categorize the residential areas as newer if the year of construction is in the range 2004–2015. Table 4.2 show the frequencies of central and newer housing units in the sample. Older housing units are more frequent than newer, and central housing units are more frequent than non-central. However, among the older housing units, non-central housing units are more frequent than central housing units. Among the newer housing units, on the other hand, all but one is central. I now move on to describe the control variables.

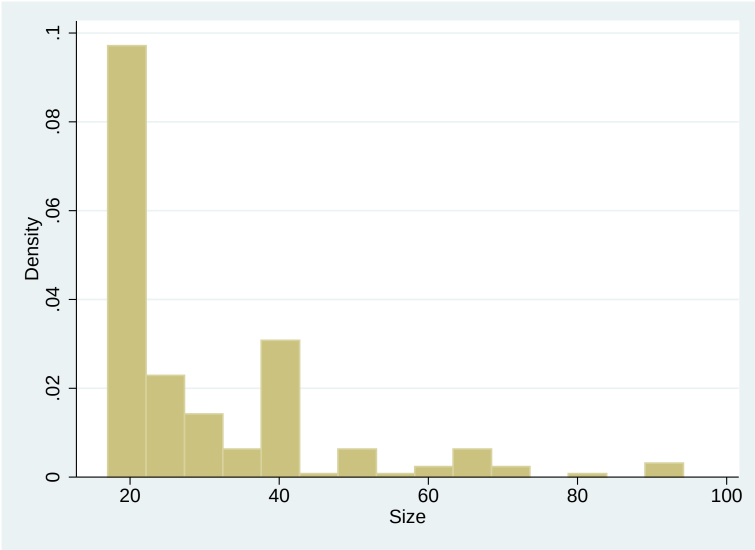

Figure 4.3: The size, in square meters, varies between about 17 and 94. The median is 22.2, the mean is 30.4 and the standard deviation is 15.2.

Table 4.2: Location and Age

Location	Year of construction 1960-1997		2004-2015		Total	
Non-central	110	44.7%	1	0.4%	111	45.1%
Central	28	11.4%	107	43.5%	135	54.9%
Total	138	56.1%	108	43.9%	246	100%

Notes to Table 4.2: The table shows the frequencies and relative frequencies of the housing units by location and year of construction. The housing units I refer to as central are situated within a distance of 0.4-1.0 km to the city centre. The housing units I refer to as newer are built in 2004-2015.

Table 4.3: Housing Size

Housing Type	Older Housing Units				Newer Housing Units				Full Sample			
	N	$\bar{x}$	s	md	N	$\bar{x}$	s	md	N	$\bar{x}$	s	md
Single-room apartment	35	23.7	3.7	21.9	54	24.2	4.4	22.1	89	24.0	4.1	22.1
1.5-room apartment	3	34.1	3.3	36.0					3	34.1	3.3	36
2-room apartment	12	50.1	4.3	51.8	53	36.6	3.0	38.6	65	39.1	6.2	38.6
2.5-room apartment	1	82.5	0	82.5					1	82.5	0	82.5
3-room apartment	15	71.6	14.4	64.5	1	63.7	0	63.7	16	71.1	14.0	64.5
4-room apartment	2	71.1	3.2	71.1					2	71.1	3.2	71.05
Corridor room	70	19.2	1.1	19.2					70	19.2	1.1	19.2

Notes to Table 4.3: The table shows the number of observations (N), mean ($\bar{x}$), standard deviation (s) and median (md). The statistics are given for older and newer housing units, and for the full sample. The size is measured in square meters. For the full sample, the mean is 30.4, the median is 22.2, and the standard deviation is 15.2.

One of the control variables I use is housing type. I divide the sample into three housing type groups: apartment, corridor room in a large corridor, and corridor room in a small corridor. Table 4.4 shows the frequencies of the three groups. Apartment is the most frequent housing type, corridor room in a large corridor is the second most frequent, and corridor room in a small corridor is least frequent. The group of apartments simply consists of all apartments in the sample. The group of large corridor rooms consists of all corridor rooms in corridors with 12 or more corridor rooms. The group of small corridor rooms consists of all corridor rooms in corridors with up to 11 corridor rooms. The cut off point 12 is used because the median and mean number of rooms in corridors are 12. The rationale for this division into three housing type groups is that the costs for shower and kitchen are split among more tenants in corridors compared to apartments, and also split among more tenants in large corridors compared to small corridors.

However, some corridor rooms have private showers instead of shared showers. I therefore also divide the corridor rooms into two groups: corridor rooms with shared shower and corridor rooms with private shower. Table 4.5 shows the frequencies for corridor rooms with shared shower and corridor rooms with private shower. Private shower is more common than shared shower.

Also the apartments are divided into groups. In some apartments the cost of electricity is included in the rent. I therefore divide the apartments into two groups: apartments with the cost of electricity included in the rent and apartments without the cost of electricity included in the rent. Table 4.6 shows the frequencies for apartments with and without the cost of electricity included in the rent. It is more common to not have the cost of electricity included in the rent. Some apartments also have a storage room. I therefore also divide apartments into another two groups: apartments with a storage room and apartments with no storage room. Table 4.7 shows the frequencies for apartments with and without a storage room. It is more common to have a storage room.

Table 4.4: Housing Types

Housing Type	Older Housing Units		Newer Housing Units		Full Sample	
Apartment	68	27.6%	108	43.9%	176	71.5%
Room in a large corridor	47	19.1%	0	0%	47	19.1%
Room in a small corridor	23	9.3%	0	0%	23	9.3%

Notes to Table 4.4: The table shows the frequencies and relative frequencies of the different housing types for older and newer housing units, and for the full sample.

Table 4.5: Shower

Shower	Small Corridor		Large Corridor		Total	
Shared	13	18.6%	6	8.6%	19	27.1%
Private	10	14.3%	41	58.6%	51	72.9%

Notes to Table 4.5: The table shows the frequencies and relative frequencies of shared and private shower for small and large corridors.

Table 4.6: Electricity

Electricity	Older Apartment		Newer Apartment		Total	
Not included	59	33.5%	107	60.8%	166	94.3%
Included	9	5.1%	1	0.6%	10	5.7%

Notes to Table 4.6: The table shows the frequencies and relative frequencies of apartments with and without the cost of electricity included in the rent.

Table 4.7: Storage

Storage Room	Older Apartment		Newer Apartment		Total	
Not included	26	14.8%	0	0.0%	26	14.8%
Included	42	23.9%	108	61.4%	150	85.2%

Notes to Table 4.7: The table shows the frequencies and relative frequencies of apartments with and without a storage room.

The variables are summarized in Table 4.8. In the table I have written short notes on the meaning of the variables and how I expect the explanatory variables to affect the dependent variable monthly rent. I expect size to have a positive effect on the monthly rent. Moreover, I also expect the effect to vary with the location and age of the housing unit. Size is probably more costly near the city centre, due to higher land prices, and also more costly for newer housing units than for older ones. Size therefore enter the hedonic regression model in three different terms: $Size$, $Size \times Central$ and $Size \times Newer$. $Size$ is a continuous variable, measuring the the size in square meters. $Size \times Central$ and $Size \times Newer$ are interaction terms of $Size$ and dummy variables for centrally located and newer housing units, respectively, that account for differences in the effect of size due to the location and age of the housing unit.

I also expect the housing type to have an effect on the monthly rent. As I pointed out before, the rationale for this is that the costs for shower and kitchen are split among more people for corridor rooms compared to apartments, and also split among more people for large corridors compared to small corridors. I therefore expect large corridors to be cheaper than small corridors, and small corridors to be cheaper than apartments, all other things being equal. I control for the effect of housing type on monthly rent with the variables *Large_Corridor*, *Small_Corridor* and *Apartment*. These are dummy variables for corridor rooms in large corridors, corridor rooms in small corridors and apartments, respectively. *Small_Corridor* is the base alternative in the hedonic regression model.

However, as I also pointed out before, some corridor rooms have private showers instead of shared showers. All other things being equal, a private shower should be more costly than a shared shower, since the cost for a shared shower is split among more tenants. I control for this with the variable *Shower*, which is a dummy variable for corridor rooms with a private shower.

Some apartments have the cost of electricity included in the rent and some have a storage room. I expect both having the cost of electricity included in the rent and having a storage room to have a positive effect on the monthly rent. I control for this with *Electricity* and *Storage*. These are dummy variables for apartments with the cost of electricity included in the rent and apartments with a storage room, respectively.

Table 4.8: Hedonic Regression Model Variables

Variable	Meaning	Type	Expected Effect On Monthly Rent
$Monthly\_Rent$	The monthly rent in SEK, converted into a 12 month payment period.	Continuous	Dependent
$Size$	The size in square meters.	Continuous	Positive
$Size \times Central$	The size in square meters, if centrally located.	Continuous	Positive
$Size \times Newer$	The size in square meters, if newer.	Continuous	Positive
$Large\_Corridor$	Whether or not it is a large corridor room.	Dummy	Negative
$Small\_Corridor$	Whether or not it is a large corridor room.	Dummy	Base
$Apartment$	Whether or not it is an apartment.	Dummy	Positive
$Shower$	Whether or not the shower is private, if it is a corridor room.	Dummy	Positive
$Electricity$	Whether or not electricity is included, if it is an apartment.	Dummy	Positive
$Storage$	Whether or not a storage room is included, if it is an apartment.	Dummy	Positive

Notes to Table 4.8: The table summarizes the variables used in the hedonic regression model. $Monthly\_Rent =$ (Monthly rent) $\times$ (Payment period)/12, where the payment period is either 9 or 12 months.

4.2 Choice Model Data

The data I use to estimate the marginal willingness to pay with the choice models is collected from 92 students at Lund University by means of a questionnaire. The questionnaire consists of two parts. In the first part the students make choices among different housing alternatives, and in the second part the students answer a couple of follow-up question about themselves. The answers in the first part gives me data on the students' preferences for housing attributes, and the answers in the second part gives me data on the students' individual attributes, which I use as control variables. An example of the questionnaire is shown in Appendix A.

In the first part the students face three separate choice scenarios. Each choice scenario consist of a choice between five different housing alternatives:

- Shared room in a small corridor
- Room in a small corridor
- Room in a large corridor
- Single-room apartment
- Shared 2-room apartment

The housing alternatives are the same in all choice scenarios. The alternatives are described by a set of alternative attributes that vary across alternatives within each choice scenario and over choice scenarios, and also over questionnaires. These are monthly rent, size, and location. That is, each choice scenario have the same alternatives but the monthly rent, size and location of the alternatives vary across alternatives within each choice scenario and over choice scenarios, and also over questionnaires.

The idea is that respondents have preferences over alternatives and also preferences over the alternative attributes, and that they reveal those preferences when they chose among the alternatives. The variables that describe the alternative attributes and my expectations of their effect on utility are summarized in Table 4.9.

Monthly rent is defined as the monthly rent per tenant per month in SEK. I expect the utility to be decreasing with monthly rent. The variable for monthly rent is labeled $Price$, since I use the respondents actual monthly rent as a control variable. $Price$ is a continuous variable.

Size is defined as the housing size in square meters. I expect the utility of size to be increasing at a diminishing rate. I therefore use two continuous variables for size, labeled $Size\_Ind$ and $Size\_Ind\_2$. $Size\_Ind$ is the size of the housing unit per individual. I use this transformation of the housing size since two students are assumed to share the housing in two of the alternatives (shared room in a small corridor and shared 2-room apartment). See Table 4.9 for more details. $Size\_Ind$ is expected to be positive and capture the increase in utility from size, and $Size\_Ind\_2$, which is the square of $Size\_Ind$, is expected to be negative and capture the diminishing rate of the increase in utility from size.

The location is defined to be central if the distance to the city centre is in the range 0.4-1.0 km, and non-central if the distance to the city centre in the range 1.0-2.6 km. I expect a higher utility for a central location compared to a non-central location, all other things being equal, since I expect proximity to services and entertainment to be desired. The variable for location is a dummy variable, labeled *Central*.

Table 4.9: Choice Model Alternative Attribute Variables

Attribute	Meaning	Type	Expected Effect On Utility
Price	The monthly rent per person in SEK.	Continuous	Negative
Size_Ind	The size of the apartment/corridor room in square meters per individual.	Continuous	Positive
Size_Ind_2	The square of *Size_Ind*.	Continuous	Negative
Central	The location of the apartment/corridor room.	Dummy	Positive

Notes to Table 4.9: The table summarizes the alternative attributes used in the choice models. *Size_Ind* = (Size of the housing alternative)/(Number of tenants), where the number of tenants is the number of tenants living in the housing alternative, that is, either 1 or 2.

In the second part the respondents answer a couple of follow-up question about themselves. The answers to the questions are summarized in Table 4.10 and Table 4.11. 4% of the respondents are exchange students, and 66% study at the bachelor's level. The most common home faculty is the School of Economics and Management, and the least common are the Faculty of Science and the Faculty of Medicine. 40% of the respondents share an apartment, 35% of the students live alone in an apartment, and 25% live in a corridor room. 37% of the respondents are male, and 40% are in a relationship. The average monthly rent is 3400 SEK, and the average age is 23.5. The variables that describe the the individual attributes are summarized in Table 4.12.

Table 4.10: Individual Attributes

Individual Attribute		N	%
Exchange student	Yes	4	4.3
	No	88	95.7
Bachelor's level	Yes	61	66.3
	No	31	33.7
Home faculty	Engineering	8	8.7
	Science	3	3.3
	Law	11	12.0
	Social Sciences	23	25.0
	Medicine	3	3.3
	Humanities and Theology	6	6.5
	Economics and Management	38	41.3
Accommodation	Corridor room	23	25.0
	Alone in an apartment	32	34.8
	Shared apartment	37	40.2
Male	Yes	34	37.0
	No	58	63.0
Relationship	Yes	37	40.2
	No	55	59.8

Notes to Table 4.10: The table shows the frequencies and relative frequencies of the answers to follow-up questions on individual attributes. Here, relationship consists of the marital statuses relationship, engaged and married.

Table 4.11: Individual Attributes

Individual Attribute	$\bar{x}$	s	md
Monthly Rent	3474.4	894.4	3400.0
Age	23.5	2.8	23.0

Notes to Table 4.11: The table shows the mean ($\bar{x}$), standard deviation (s) and median (md). The monthly rent is measured in SEK, and the age in years.

Table 4.12: Choice Model Individual Attribute Variables

Attribute	Meaning	Type
Exchange	Whether or not the individual is an exchange student.	Dummy
Bachelor	Whether or not the individual study at the bachelor's level.	Dummy
Engineering	Whether or not the individual's home faculty is the Faculty of Engineering.	Dummy
Science	Whether or not the individual's home faculty is the Faculty of Science.	Dummy
Law	Whether or not the individual's home faculty is the Faculty of Law.	Dummy
Social	Whether or not the individual's home faculty is the Faculty of Social Sciences.	Dummy
Medicine	Whether or not the individual's home faculty is the Faculty of Medicine.	Dummy
Humanities	Whether or not the individual's home faculty is the Faculty of Humanities and Theology.	Dummy
Economics	Whether or not the individual's home faculty is the School of Economics and Management.	Dummy
Corridor	Whether or not the individual live in a corridor.	Dummy
Apt_Alone	Whether or not the individual live alone in an apartment.	Dummy
Apt_Share	Whether or not the individual live in a shared apartment.	Dummy
Rent	The monthly rent paid by the individual in SEK.	Continuous
Age	The individual's age.	Continuous
Male	Whether or not the individual is male.	Dummy
Relationship	Whether or not the individual is in a relationship, engaged or married, as opposed to single.	Dummy

Notes to Table 4.12: The table summarizes the individual attribute variables used in the choice models.

5 Choice Models

To estimate the marginal willingness to pay, I use discrete choice multinomial models. Discrete choice multinomial models describe individual's choices among several different alternatives. In this case, the different alternatives are the five housing types that the respondents choose between in the questionnaire. The students' choices are assumed to be determined by both observed and unobserved factors. The observed factors are the attributes describing the alternatives and individuals in the questionnaire. For example, the alternatives are described by monthly rent, size and location, and the individuals are described by level of study, home faculty and age, etc. Since the individual's choices are assumed to be determined by both observed and unobserved factors, the choices cannot be predicted exactly. Instead, the probabilities of the alternatives are derived, conditional on the observed factors. The marginal willingness to pay estimates are then derived from the estimated parameters of monthly rent and size.

I consider three different discrete choice models: the multinomial logit model, the nested logit model and the random parameters logit model. The simplest of the models is the multinomial logit model. However, it may be inappropriate due to its restrictive assumptions. For example, it cannot capture taste variation that is associated with unobserved factors. I therefore also consider the nested logit and the random parameters logit models. They are derived from less restrictive assumptions and hence more general. However, one thing that the three models all have in common is that they can be derived from utility maximizing behavior. They are therefore called random utility models (see, for example, Cameron and Trivedi 2005, and Train 2003). In section 5.1 I describe the random utility model, from which the different models can be derived by making different assumptions about the unobserved factors. In sections 5.2-5.4 I then move on to describe the different models and how they are estimated. In section 5.5 I describe how goodness of fit can be evaluated and how hypotheses can be tested. This section draws on Cameron and Trivedi (2005), and Train (2003).

5.1 Random Utility Models

The selection probabilities of the alternatives in a discrete choice multinomial model are called choice probabilities. They can be derived from utility maximizing behavior by using a random utility model. In a random utility model the individual, labled i, chooses among several alternatives. The alternative the individual chooses is the alternative that gives the highest level of utility. The utility U_{ij} individual i gets from alternative j is assumed to be determined by both observed and unobserved factors. The observed factors are collected in a component denoted V_{ij} and the unobserved factors are collected in a component denoted ε_{ij}. The utility U_{ij} is specified as a sum of the observed factors V_{ij} and the unobserved factors ε_{ij}

$$U_{ij} = V_{ij} + \varepsilon_{ij}. \tag{5.1}$$

Thus, individual i chooses alternative j if

$$\begin{aligned} U_{ij} &> U_{ik} \ \forall k \neq j \\ \Leftrightarrow V_{ij} + \varepsilon_{ij} &> V_{ik} + \varepsilon_{ik} \ \forall k \neq j. \end{aligned}$$

Since the ε_i's are unobserved, the choice cannot be predicted exactly. However, by treating the ε_i's as random it is possible to calculate the probabilities of the alternatives. The probability that individual i chooses alternative j is given by

$$\begin{aligned} P_{ij} &= \text{Prob}(U_{ij} > U_{ik} \ \forall k \neq j) \\ &= \text{Prob}(V_{ij} + \varepsilon_{ij} > V_{ik} + \varepsilon_{ik} \ \forall k \neq j). \end{aligned}$$

Different discrete choice multinomial models can be generated from different specifications of the ε_i's joint distribution. For example: the multinomial logit and the random parameters logit models are derived under the assumption that the ε_i's are independent and identically distributed random terms that follow the extreme value distribution; the nested logit model is derived under the assumption that the marginal distribution of the ε_i's is univariate extreme value and that their joint cumulative distribution is the generalized extreme value distribution.

It is worth pointing out that the distribution of the ε_i's depend on the specification of the V_i's. The reason for this is that the ε_i's in (5.1) are defined as the difference between the U_i's and the V_i's. Thus, if the V_i's are specified sufficiently, with all relevant explanatory variables included, the ε_i's are essentially white noise. However, if relevant explanatory variables are omitted, there might be correlation among the ε_i's.

5.2 Multinomial Logit

5.2.1 Multinomial Logit – Choice Probabilities

The choice probabilities for the multinomial logit model is derived under the assumption that the ε_i's are independent and identically distributed random terms that follow the extreme value distribution, with probability density function

$$f(\varepsilon_{ij}) = e^{-\varepsilon_{ij}} e^{-e^{\varepsilon_{ij}}}, \quad j = 1, 2, \ldots, m, \tag{5.2}$$

and cumulative distribution function

$$F(\varepsilon_{ij}) = e^{-e^{-\varepsilon_{ij}}}, \quad j = 1, 2, \ldots, m. \tag{5.3}$$

The probability that individual i chooses alternative j is given by

$$\begin{aligned} P_{ij} &= \text{Prob}(U_{ij} > U_{ik} \ \forall k \neq j) \\ &= \text{Prob}(V_{ij} + \varepsilon_{ij} > V_{ik} + \varepsilon_{ik} \ \forall k \neq j) \\ &= \text{Prob}(\varepsilon_{ik} < \varepsilon_{ij} + V_{ij} - V_{ik} \ \forall k \neq j). \end{aligned}$$

If ε_{ij} where not random but given, this would simply be the cumulative distribution for each ε_{ik} evaluated at $\varepsilon_{ij} + V_{ij} - V_{ik}$. According to (5.3) this is $e^{-e^{-\varepsilon_{ij}+V_{ij}-V_{ik}}}$. Since the ε_i's are assumed to be independent, this cumulative distribution over all $k \neq j$ is the product of the individual cumulative distributions

$$P_{ij}|\varepsilon_{ij} = \prod_{k \neq j} e^{-e^{-(\varepsilon_{ij}+V_{ij}-V_{ik})}}.$$

The choice probability is the integral of $P_{ij|\varepsilon_{ij}}$ over all values of ε_{ij}, weighted by its probability density function in (5.2).

$$P_{ij} = \int \left(\prod_{k \neq j} e^{-e^{-(\varepsilon_{ij}+V_{ij}-V_{ik})}} \right) e^{-\varepsilon_{ij}} e^{-e^{\varepsilon_{ij}}} \ d\varepsilon_{ij}.$$

This integral has the closed form solution

$$P_{ij} = \frac{e^{V_{ij}}}{\sum_{k=1}^{m} e^{V_{ik}}}.$$

With a linear specification of V_{ij} the choice probabilities take the following form

$$P_{ij} = \frac{e^{\boldsymbol{x}'_{ij}\boldsymbol{\beta}}}{\sum_{k=1}^{m} e^{\boldsymbol{x}'_{ik}\boldsymbol{\beta}}},$$

where $\boldsymbol{x}_{ij}$ is a vector of observed variables associated with alternative j for individual i, and $\boldsymbol{\beta}$ is a vector of parameters. The linear specification is not as restrictive as it first might seem. In fact, functions that are linear in parameters can approximate any function arbitrarily close under fairly general conditions.

The critical assumption of the multinomial logit model is the independence assumption. As was pointed out in section 5.2.1, the distribution of the ε_i's depend on the specification of the V_i's, and if relevant explanatory variables are omitted there might be correlation among the ε_i's. In that case the independence assumption is violated and, consequently, the model is misspecified. Therefore, I also consider the nested logit model and the random parameters model.

5.2.2 Multinomial Logit – Estimation

Since the multinomial logit choice probabilities have a closed form solution and the data necessarily is multinomial distributed, maximum likelihood can be applied when estimating the parameters of the multinomial logit model. Consider a random sample of N individuals. For each individual in the sample, introduce m binary variables for the dependent variable y_i

$$y_{ij} = \begin{cases} 1 & \text{if individual } i \text{ chose alternative } j \\ 0 & \text{otherwise.} \end{cases}$$

Thus, for each individual, exactly one of $y_{i1}, y_{i2}, \ldots, y_{im}$ will be non-zero. With this notation, since P_{ij} raised to the power of zero is equal to one, the probability that the individual chooses the observed choice can be written as

$$f(y_i) = P_{i1}^{y_{i1}} \times \cdots \times P_{im}^{y_{im}} = \prod_{j=1}^{m} P_{ij}^{y_{ij}}.$$

It follows that, if each individual's decision is independent of the other individuals decisions, the probability that each individual chooses the observed choice can be written as

$$L(\boldsymbol{\beta}) = \prod_{i=1}^{N} \prod_{j=1}^{m} P_{ij}^{y_{ij}}.$$

This is the likelihood function for a sample of N independent observations. The log-likelihood function is

$$\ln L(\boldsymbol{\beta}) = \sum_{i=1}^{N} \sum_{j=1}^{m} y_{ij} \ln P_{ij}.$$

The maximum likelihood estimator $\hat{\boldsymbol{\beta}}$ is the value of $\boldsymbol{\beta}$ that maximizes this log-likelihood function. Furthermore, the covariance matrix is minus the inverse of the information matrix.

5.3 Random Parameters Logit

5.3.1 Random Parameters Logit – Choice Probabilities

The choice probabilities for the random parameters logit model (also called the mixed logit model) is derived under the assumption that the ε_i's follow the extreme value distribution, with probability density function

$$f(\varepsilon_{ij}) = e^{-\varepsilon_{ij}} e^{-e^{\varepsilon_{ij}}}, \quad j = 1, 2, \ldots, m,$$

and cumulative distribution function

$$F(\varepsilon_{ij}) = e^{-e^{-\varepsilon_{ij}}}, \quad j = 1, 2, \ldots, m.$$

That is, the assumptions about the ε_i's are the same as for multinomial logit model. The assumption that distinguishes random parameters logit model from the multinomial logit model is that the parameters are allowed to be random in the random parameters logit model. The utility individual i gets from alternative j is given by

$$\begin{aligned} U_{ij} &= V_{ij} + \varepsilon_{ij}, \\ &= \boldsymbol{x}'_{ij}\boldsymbol{\beta}_i + \varepsilon_{ij}, \quad j = 1, 2, \ldots, m, \end{aligned} \tag{5.4}$$

where $\boldsymbol{\beta}_i$ follow some distribution with probability density function $f(\boldsymbol{\beta}_i|\boldsymbol{\theta})$. Here $\boldsymbol{\theta}$ represent a set of parameters that describe the probability density function of $\boldsymbol{\beta}_i$. For example, the probability density function of $\boldsymbol{\beta}_i$ can be described by its mean $\boldsymbol{\beta}$ and its covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{\beta}}$: $\boldsymbol{\theta} = \{\boldsymbol{\beta}, \boldsymbol{\Sigma}_{\boldsymbol{\beta}}\}$. If $\boldsymbol{\beta}_i$ where not random but a known, the choice probabilities would simply be the choice probabilities for the multinomial logit model. That is, the choice probabilities conditional on $\boldsymbol{\beta}_i$ is

$$P_{ij}(\boldsymbol{\beta}_i) = \frac{e^{\boldsymbol{x}'_{ij}\boldsymbol{\beta}_i}}{\sum_{k=1}^{m} e^{\boldsymbol{x}'_{ik}\boldsymbol{\beta}_i}}.$$

However, $\boldsymbol{\beta}_i$ is not known but random. Therefore it is not possible to condition on $\boldsymbol{\beta}_i$. Hence, the unconditional choice probabilities need to be derived. The unconditional choice probabilities are obtained by integrating out the randomness

$$P_{ij} = \int \frac{e^{\boldsymbol{x}'_{ij}\boldsymbol{\beta}_i}}{\sum_{k=1}^{m} e^{\boldsymbol{x}'_{ik}\boldsymbol{\beta}_i}} f(\boldsymbol{\beta}_i|\boldsymbol{\theta})\ d\boldsymbol{\beta}_i. \tag{5.5}$$

Note that this is a multidimensional integral. It is common (see, for example, Revelt and Train 1998) to assume that $\boldsymbol{\beta}_i$ follow a normal distribution

$$\boldsymbol{\beta}_i \sim \mathcal{N}[\boldsymbol{\beta}, \boldsymbol{\Sigma}_{\boldsymbol{\beta}}], \tag{5.6}$$

or a log-normal distribution.

$$\ln \boldsymbol{\beta}_i \sim \mathcal{N}[\boldsymbol{\beta}, \boldsymbol{\Sigma}_{\boldsymbol{\beta}}].$$

The log-normal distribution is common to use when the parameters are known beforehand to have the same sign for all individuals. If $\boldsymbol{\beta}_i$ is specified to follow a normal distribution as described in (5.6), the expression in (5.5) becomes

$$P_{ij} = \int \frac{e^{\boldsymbol{x}'_{ij}\boldsymbol{\beta}_i}}{\sum_{k=1}^{m} e^{\boldsymbol{x}'_{ik}\boldsymbol{\beta}_i}} \Phi(\boldsymbol{\beta}_i|\boldsymbol{\beta}, \boldsymbol{\Sigma}_{\boldsymbol{\beta}})\ d\boldsymbol{\beta}_i.$$

where $\Phi(\boldsymbol{\beta}_i|\boldsymbol{\beta}, \boldsymbol{\Sigma}_{\boldsymbol{\beta}})$ is the multivariate normal probability density function for $\boldsymbol{\beta}_i$, with parameters $\boldsymbol{\beta}$ and $\boldsymbol{\Sigma}_{\boldsymbol{\beta}}$.

So what is the purpose of allowing the parameters to be random? By allowing the parameters to be random the multinomial logit model is generalized to allow the utilities of different alternatives to be correlated. It thus solves the multinomial logit model's problem with the independence assumption, discussed in section 5.2.1. To see this, the utility in (5.4) can be rewritten by decomposing the coefficients $\boldsymbol{\beta}_i$ into their mean $\boldsymbol{\beta}$ and deviations $\boldsymbol{u}_i$

$$
\begin{aligned}
U_{ij} &= \boldsymbol{x}'_{\boldsymbol{ij}}\boldsymbol{\beta} + v_{ij},\\
v_{ij} &= \boldsymbol{x}'_{ij}\boldsymbol{u}_i + \varepsilon_{ij},
\end{aligned}
$$

where $\boldsymbol{u}_i \sim \mathcal{N}[\mathbf{0}, \boldsymbol{\Sigma}_{\boldsymbol{\beta}}]$. The covariance in utility among alternatives become $\mathrm{Cov}[v_{ij}, v_{ik}] = E[(\boldsymbol{x}'_{ij}\boldsymbol{u}_i+\varepsilon_{ij})(\boldsymbol{x}'_{ik}\boldsymbol{u}_i+\varepsilon_{ik})] = \boldsymbol{x}'_{ij}\boldsymbol{\Sigma}_{\boldsymbol{\beta}}\boldsymbol{x}_{ik}, j \neq k$. The random parameters logit model takes the multinomial logit model as a special case when all $\boldsymbol{u}_i$ are non-stochastic.

5.3.2 Random Parameters Logit – Estimation

When estimating the parameters in the random parameters logit model, the idea is the same as when estimating the parameters in the multinomial logit model. The maximum likelihood estimator $\hat{\boldsymbol{\beta}}$ is the value of $\boldsymbol{\beta}$ that maximizes the log-likelihood function

$$
\ln L(\boldsymbol{\beta}) = \sum_{i=1}^{N}\sum_{j=1}^{m} y_{ij} \ln P_{ij}.
$$

Unfortunately, there is no closed-form solution to the integral in the formula for the choice probabilities P_{ij} in (5.5). Simulation methods are therefore needed. By replacing the integral with the average of S evaluations of the integrand at random drawn values of $\boldsymbol{\beta}_i$ from the probability density function $f(\boldsymbol{\beta}_i|\boldsymbol{\theta})$, the simulated log-likelihood function is obtained

$$
\ln \hat{L}(\boldsymbol{\theta}) = \sum_{i=1}^{N}\sum_{j=1}^{m} y_{ij} \ln \left[\frac{1}{S} \sum_{s=1}^{S} \frac{e^{\boldsymbol{x}'_{ij}\boldsymbol{\beta}_i^{(s)}}}{\sum_{k=1}^{m} e^{\boldsymbol{x}'_{ik}\boldsymbol{\beta}_i^{(s)}}} \right].
$$

Here, $\boldsymbol{\beta}_i^{(s)}$, $s = 1, \ldots, S$, are the random drawn values of $\boldsymbol{\beta}_i$ from the probability density function $f(\boldsymbol{\beta}_i|\boldsymbol{\theta})$. Since the parameters in $\boldsymbol{\theta}$ are unknown the simulation procedure need to be iterative. Consistent estimates requires that $N \to \infty$ and that $S \to \infty$ as well as $\sqrt{N}/S \to \infty$.

5.4 Nested Logit

5.4.1 Nested Logit – Choice Probabilities

The nested logit model partition the alternatives into subsets, called nests. It is convenient to picture this with a tree diagram, where each branch represents a nest (i.e subset) of the alternatives and every leaf on each branch represents an alternative. Figure 5.1 and Figure 5.2 depict the two nesting schemes that I consider. More generally, the alternatives are partitioned into J non-overlapping subsets. The jth subset has K_j alternatives, labeled $j1, \ldots, jk, \ldots, jK_j$.

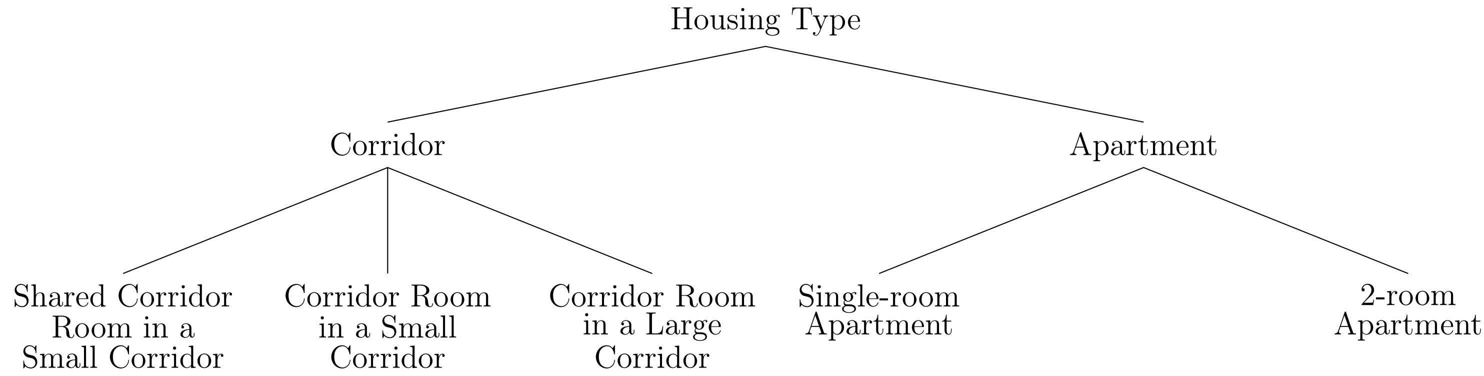

Figure 5.1: The first nesting scheme portion the alternatives into nests with regard to if they are corridor rooms or apartments.

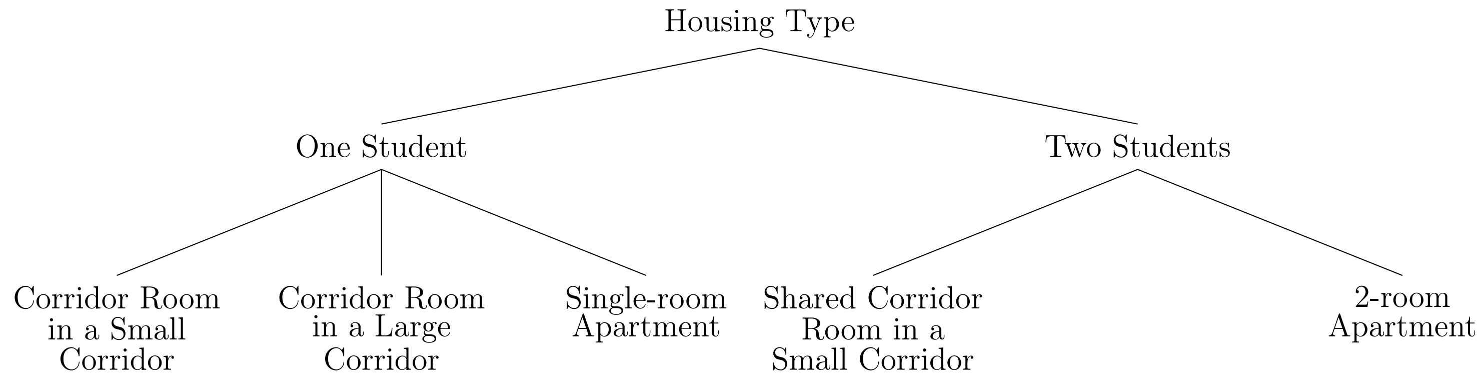

Figure 5.2: The second nesting scheme portion the alternatives into nests with regard to if they accommodate one or two students.

The utility from the kth of K_j alternatives in the jth of J subsets is

$$U_{jk} = V_{jk} + \varepsilon_{jk}, \quad k = 1, 2, \ldots, K_j, \quad j = 1, 2, \ldots, J.$$

The choice probabilities for the nested logit model is derived under the assumption that the ε_i's ($\varepsilon_i = \varepsilon_{i11}, \dots \varepsilon_{i1K_1}, \dots, \varepsilon_{iJK_j}$) has the cumulative distribution function

$$F(\varepsilon_{jk}) = \exp\left(-\sum_{j=1}^{J}\left(\sum_{k=1}^{K_j} e^{-\varepsilon_{jk}/\lambda_j}\right)^{\lambda_j}\right).$$

This is a type of the generalized extreme value distribution. For this generalized extreme value distribution, the ε_i's are correlated within nests and uncorrelated across nests. The parameter λ_j is a function of the correlation in unobserved factors among the alternatives in nest j. However, it is not equal to the correlation, but inversely related to it: $\lambda_j = \sqrt{1 - \mathrm{Corr}[\varepsilon_{jk}, \varepsilon_{j\ell}]}$. Thus, a higher value of λ_j means greater independence.

The probability that alternative k in subset j is chosen is given by

$$\begin{aligned} P_{jk} &= \Pr[U_{jk} \geq U_{lm} \ \forall l, m], \\ &= \Pr[V_{jk} + \varepsilon_{jk} \geq V_{lm} + \varepsilon_{lm} \ \forall l, m], \\ &= \Pr[\varepsilon_{lm} \leq \varepsilon_{jk} + V_{jk} - V_{lm} \ \forall l, m], \end{aligned}$$

which has the closed form solution

$$P_{jk} = \frac{e^{V_{jk}/\lambda_j}\left(\sum_{l=1}^{K_l} e^{V_{jl}/\lambda_j}\right)^{\lambda_j - 1}}{\sum_{m=1}^{J}\left(\sum_{l=1}^{K_l} e^{V_{ml}/\lambda_m}\right)^{\lambda_m}}, \tag{5.7}$$

where

$$V_{jk} = \boldsymbol{z}_j'\boldsymbol{\alpha} + \boldsymbol{x}_{jk}'\boldsymbol{\beta_j}, \quad k = 1, 2, \dots, K_j, \quad j = 1, 2, \dots, J.$$

Here $\boldsymbol{z}_j$ is a vector of observed variables associated with subset j, and $\boldsymbol{x}_{jk}$ is a vector of observed variables associated with alternative k in subset j. Moreover, $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}_j$ are vectors of parameters. The formula for the choice probabilities in (5.7) can be rewritten to make it simpler to interpret. This is done by factoring the probability P_{jk} as P_j, the probability that an alternative in subset j is chosen, times $P_{k|j}$, the probability that alternative k is chosen given that an alternative in subset j is chosen

$$\begin{aligned} P_{jk} &= P_j P_{k|j} \\ &= \frac{e^{\boldsymbol{z}_j'\boldsymbol{\alpha} + \lambda_j I_j}}{\sum_{m=1}^{J} e^{\boldsymbol{z}_m'\boldsymbol{\alpha} + \lambda_m I_m}} \frac{e^{\boldsymbol{x}_{jk}'\boldsymbol{\beta}_j/\lambda_j}}{\sum_{l=1}^{K_j} e^{\boldsymbol{x}_{jl}'\boldsymbol{\beta}_j/\lambda_j}}, \end{aligned} \tag{5.8}$$

where

$$I_j = \ln\left(\sum_{l=1}^{K_j} e^{\boldsymbol{x}'_{jl}\boldsymbol{\beta}_j/\lambda_j}\right).$$

Thus, P_j, the probability of choosing an alternative in nest j, can be written as a logit formula for a choice among nests, and $P_{k|j}$, the probability of choosing alternative k given that an alternative in nest j is chosen, can be written as a logit formula for a choice among the alternatives in the nest. The quantity I_j in the formula for P_j is the log of the denominator of $P_{k|j}$. It hence bring information about the alternatives in the nests to the choice among nests.

The model is consistent with utility-maximizing behavior for all possible values of the explanatory variables if $\lambda_j\ \forall j$ is between zero and one. For λ_j greater than one the model is only consistent with utility maximizing behavior for some range of the explanatory variables, and for negative λ_j the model is inconsistent with utility maximizing behavior. In the latter case an improvement in the attributes of an alternative can decrease its probability.

The nested logit model successfully solves the multinomial logit model's problem with the independence assumption if all correlation among the ε_i's is accommodated within the nests. If there is no correlation among the ε_i's, then $\lambda_j = 1$ for all j and the choice probabilities in (5.7) become simply the multinomial logit model. That is, the nested logit model takes the multinomial logit model as a special case when the multinomial logit model's independence assumption is not violated.

5.4.2 Nested Logit – Estimation

Since the nested logit choice probabilities have a closed form solution and the data necessarily is multinomial distributed, maximum likelihood can be applied when estimating the parameters of the nested logit model. Consider a random sample with N individuals. For each individual in the sample, introduce $K_1 + \cdots + K_J$ binary variables for the dependent variable y_i.

$$y_{ijk} = \begin{cases} 1 & \text{if individual } i \text{ chose alternative } k \text{ in nest } j \\ 0 & \text{otherwise.} \end{cases}$$

Thus, for each individual, exactly one of $y_{i11}, \ldots, y_{i1K_1}, \ldots, y_{iJ1} \ldots, y_{iJK_J}$ will be non-zero. With this notation, since P_{ijk} raised to the power of zero is equal to one, the probability that the individual chooses the observed choice can be written as

$$\begin{aligned} f(y_i) &= P_{i11}^{y_{i11}} \times \cdots \times P_{i1K_1}^{y_{i1K_1}} \times \cdots \times P_{iJ1}^{y_{iJ1}} \times \cdots P_{iJK_J}^{y_{iJK_J}} \\ &= \prod_{j=1}^{J} \prod_{k=1}^{K_j} P_{ijk}^{y_{ijk}} = \prod_{j=1}^{J} \prod_{k=1}^{K_j} \left[P_{ij} P_{ik|j} \right]^{y_{ijk}} = \prod_{j=1}^{J} \left(P_{ij}^{y_{ij}} \prod_{k=1}^{K_j} P_{ik|j}^{y_{ijk}} \right), \end{aligned}$$

where y_{ij} equals one if an alternative in subset j is chosen, and zero otherwise. It follows that, if each individual's decision is independent of the other individuals decisions, the probability that each individual chooses the observed choice can be written as

$$L(\boldsymbol{\beta}) = \prod_{i=1}^{N} \prod_{j=1}^{J} \left(P_{ij}^{y_{ij}} \prod_{k=1}^{K_j} P_{ik|j}^{y_{ijk}} \right).$$

This is the likelihood function for a sample of N independent observations. The log-likelihood function is

$$\ln L(\boldsymbol{\beta}) = \sum_{i=1}^{N} \sum_{j=1}^{J} y_{ij} \ln P_{ij} + \sum_{i=1}^{N} \sum_{j=1}^{J} \sum_{k=1}^{K_j} y_{ijk} \ln P_{ik|j}.$$

The maximum likelihood estimators $\hat{\boldsymbol{\alpha}}$, $\hat{\boldsymbol{\beta}}_j$ and $\hat{\lambda}_j$ are the values of $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ and λ_j that maximizes this log-likelihood function. The estimators are consistent and efficient under fairly general conditions (Brownstone and Small 1989). Furthermore, the covariance matrix is minus the inverse of the information matrix.

5.5 Goodness of Fit and Hypothesis Testing

The goodness of fit statistic used with multinomial models is the likelihood ratio index. It measures how well the model performs compared to a model without any explanatory variables. The likelihood ratio index is defined as

$$LRI = 1 - \frac{\ln L(\hat{\boldsymbol{\beta}})}{\ln L(\mathbf{0})},$$

where $\ln L(\hat{\boldsymbol{\beta}})$ is the log-likelihood function evaluated with the estimated parameters, and $\ln L(\mathbf{0})$ is the log-likelihood function evaluated with all parameters set equal to zero. The value of the likelihood ratio index is between zero and one. When the model performs no better than a model without any explanatory variables, the likelihood ratio index is zero, and when the model perfectly predicts choices, the likelihood ratio index is one. Thus, in a comparison between two models, the one with the higher log-likelihood ratio index fits the data better.

Hypotheses about individual parameters are tested using t-tests. More complex hypotheses are tested using likelihood ratio tests.

6 Result

6.1 Hedonic Regression Model Result

I estimate a hedonic regression model on the data described in Section 4.1. The dependent variable is monthly rent and the explanatory variable of interest is size. Size enters the regression model on its own and in interaction with dummy variables for centrally located and newer housing units. I also use a set of control variables. The variables where described in Section 4.1 and summarized in Table 4.8. The model is specified below

$$\begin{aligned} Monthly\_Rent = &\beta_0 + \beta_1 Size + \beta_2 Size \times Central + \beta_3 Size \times Newer \\ &+ \beta_4 Large\_Corridor + \beta_5 Apartment + \beta_6 Shower \\ &+ \beta_7 Electricity + \beta_8 Storage. \end{aligned}$$

The parameters are estimated with OLS and the result is reported in the first result column of Table 6.1, where I refer to it as Hedonic Regression Model 1 (HRM1). A Breusch-Pagan test finds strong evidence for heteroskedasticity (p-value=0.001). I therefore report heteroskedasticity consistent standard errors for this model. The R^2-value is 0.96. That is, the model captures most of the variation in the data. However, a RESET test for omission of variables and/or inappropriate functional form is rejected at the 1%-level (p-value=0.000). Moreover, the sign of the estimated parameters are the expected for all explanatory variables, and 6 out of 8 of these are individually significant at the 1%-level. The other two are not even individually significant at the 10%-level. An F-test for joint significance of the two individually insignificant explanatory variables fails to reject the null hypotheses of no joint significance. This indicates that a model with only the 6 individually significant explanatory variables has no worse explanatory power than that with the 6 individually significant explanatory variables and the two individually insignificant explanatory variables.

To decide whether or not I should drop the two insignificant variables, I compare the AIC value for HRM1 to the AIC value from an alternative model without the two insignificant variables. The estimation result for the alternative model is reported in the second result column of Table 6.1, where I refer to it as Hedonic Regression Model 2 (HRM2). The reported standard errors are heteroskedasticity consistent, since a Breusch-

Pagan test finds strong evidence for heteroskedasticity in HRM2 (p-value=0.018). The R^2-value is still 0.96. The sign of the estimated parameters are the expected for all explanatory variables and they are individually significant at the 1%-level. A RESET test is still rejected at the 1%-level (p-value=0.000). The AIC is lower for HRM2 than for HRM1. Hence, I conclude that HRM2 is preferred to HRM1.

Remember that the data is sampled from AF Bostäder that reinvests all profit into the company. I therefore interpret the parameter estimates as the marginal cost. Thus, the estimated marginal cost for size in terms of monthly rent is 79 SEK for older housing units in a non-central location. For central housing units the estimate is 8 SEK higher, and for newer housing units the estimate is 48 SEK higher. That is, the marginal cost estimate of size for an older housing unit in a central location is 87 SEK, and for a newer housing unit in a non-central location it is 127 SEK, and for a newer housing unit in a central location it is 135 SEK. The marginal cost estimates are summarized in Table 6.2.

Moreover, all other things being equal, the monthly rent for a corridor room in a large corridor is 292 SEK lower than the monthly rent for a corridor room in a small corridor. Also, all other things being equal, the monthly rent for an apartment is 185 SEK higher than the monthly rent for a corridor room in a small corridor. Furthermore, all other things being equal, the monthly rent for a corridor room with a private shower is 396 SEK higher than the monthly rent for a corridor room with shared shower.

The result that electricity and storage room does not affect the monthly rent is not in line with what I expected. To check the robustness of this result, I also estimate the model on only the older housing units in the data set. The estimation result for this model is reported in the third result column of Table 6.1, where I refer to it as Hedonic Regression Model 3 (HRM3). In this case the estimated parameters does not only have the expected sign. They are now also individually significant at the 1%-level. For older housing units, all other things being equal, the monthly rent for apartments is 186 SEK higher if the cost of electricity is included in the rent, and 549 SEK higher for apartments with a storage room.

I also estimate the model on only the newer housing units in the data set. The estimation result for this model is reported in the fourth result column of Table 6.1, where I refer to it as Hedonic Regression Model 4 (HRM4).

Now that I have estimates of the marginal cost of size, I move on to the result of the choice models which I use to derive estimates for the marginal willingness to pay for housing size.

Table 6.1: Hedonic Regression Model Estimates

Dependent variable: Monthly Rent				
Variables	HRM1[a]	HRM2[a]	HRM3[b]	HRM4[c]
Size	79.39	79.18	66.54	144.47
	(1.83)***	(1.46)***	(1.77)***	(7.98)***
Size × *Central*	8.31	8.22	6.39	10.50
	(1.81)***	(1.69)***	(1.35)***	(2.34)***
Size × *Newer*	47.99	47.79		
	(2.29)***	(1.90)***		
Large_Corridor	−291.51	−291.96	−307.53	
	(75.28)***	(75.04)***	(78.81)***	
Apartment	165.13	184.90	150.31	
	(49.81)***	(55.68)***	(34.84)***	
Shower	397.11	395.89	384.85	
	(75.52)***	(74.76)***	(76.26)***	
Electricity	93.59		186.06	
	(78.70)		(62.62)***	
Storage	8.45		548.59	
	(86.68)		(93.93)***	
Constant	788.53	794.185	1065.06	332.75
	(55.53)***	(44.99)***	(52.63)***	(175.61)*
R-squared	0.96	0.96	0.98	0.91
F-statistic	1260***	1657***	1812***	1034***
AIC	3588.769	3585.339	1890.977	1598.972
N	246	246	138	108

Notes to Table 6.1: Robust standard errors are in parentheses. *Significant at 10%, **Significant at 5%, ***Significant at 1%. Data are sampled from AF Bostäder. [a]Full sample, [b]Sample of older housing units, [c]Sample of newer housing units. *Size* is a continuous variable, measuring the the size in square meters. *Size* × *Central* and *Size* × *Newer* are interaction terms of *Size* and dummy variables for centrally located and newer housing units, respectively. *Large_Corridor* and *Apartment* are dummy variables for corridor rooms in large corridors and apartments, respectively. The base alternative is a corridor room in a small corridor. *Shower* is a dummy variable for corridor rooms with a private shower. *Electricity* and *Storage* are dummy variables for apartments with the cost of electricity included in the rent, and apartments with a storage room, respectively.

Table 6.2: Marginal Cost Estimates

Housing, Location	Marginal Cost
Older, Non-central	79.18
Older, Central	87.40
Newer, Non-central	126.97
Newer, Central	135.19

Notes to Table 6.2: Estimates of marginal cost for size in terms of monthly rent are reported in SEK.

6.2 Choice Models Result

I estimate multinomial logit, nested logit, and random parameters logit models on the data described in Section 4.2. The dependent variable is housing type and the explanatory variables of interest is monthly rent and size. The variables where described in Section 4.2 and summarized in Table 4.9 and Table 4.12.

The first model I estimate is the multinomial logit model. I begin by not controlling for the individual attributes. A summary of the result is reported in the first result column of Table 6.3, where I refer to it as Multinomial Logit 1 (MNL1). The complete result is shown in Table B.1 in Appendix B. The sign of the estimated parameters are the expected for all alternative attribute variables, and they are all also individually significant at the 1%-level.

I continue by estimating a multinomial logit model with individual attributes entering the observed utility component to see if that result in a model with a better fit to the data. A summary of the result is reported in the second result column of Table 6.3, where I refer to it as Multinomial Logit 2 (MNL2). The complete result is shown in Table B.2 in Appendix B. Again, the sign of the estimated parameters are the expected for all alternative attribute variables, and they are all significant at the 1%-level. Moreover, the likelihood ratio index is higher compared to that of MNL1. This indicate that the controlling for individual attributes improves the fit to the data. A likelihood ratio test with the null hypothesis that the individual attribute control variables are not jointly significant gives a p-value of 0.000. Hence, I conclude that controlling for individual attributes significantly improves the fit to the data.

The next model I estimate is the nested logit model. I consider two alternative nesting schemes. In the first nesting scheme I portion the alternatives into nests with regard to if they are corridor rooms or apartments, and in the second I portion the alternatives into nests with regard to if they accommodate one or two students. The nested logit models are estimated without the individual attributes as control variables. The reason for this is that it simply is not possible to estimate a nested logit model with that many parameters on my data.

A summary of the result from the first nested logit model is reported in the third result column of Table 6.3, where I refer to it as Nested Logit 1 (NL1). The complete result is shown in Table B.3 in Appendix B. Again, the sign of the estimated parameters are the expected for all alternative attribute variables, and they are all significant at the 1%-level. Moreover, the λ_j parameter is 0.87, with a confidence interval in the range $(0.22, 1.52)$, for the nest of apartments, and 1.86, with a confidence interval in the range $(0.12, 3.59)$, for the nest of corridor rooms (see Table B.3 in Appendix B). That is, neither of the λ_j parameters are significantly smaller than zero or significantly larger than one. This means that I cannot reject the assumption of utility maximizing individuals (remember from Section 5.4 that the model is consistent with utility-maximizing behavior for all possible values of the explanatory variables if λ_j $\forall j$ is between zero and one). Moreover, the likelihood ratio index is higher compared to that of MNL1. This indicate that letting the ε_i's be correlated within the nests improves the fit to the data. A likelihood ratio test with the null hypothesis $\lambda_1 = \lambda_2 = 1$ gives a p-value of 0.193. Thus, the null hypothesis cannot be rejected. Hence, I cannot conclude that letting the ε_i's be correlated within the nests significantly improves the fit to the data.

A summary of the result from the second nested logit model is reported in the fourth result column of Table 6.3, where I refer to it as Nested Logit 2 (NL2). The complete result is shown in Table B.4 in Appendix B. For this model, only $Size\_Ind$ and $Size\_Ind\_2$ are significant at the 1%-level. $Price$ is significant at the 5%-level and $Central$ is not even significant at the 10%-level. Moreover, the λ_j parameter is 0.91, with a confidence interval in the range $(0.17, 1.65)$, for the nest of housing units accommodating one student, and 0.58, with a confidence interval in the range $(-0.16, 1.30)$, for the nest of housing units accommodating two students (see Table B.4 in Appendix B). That is, again, neither of the λ_j parameters are significantly smaller than zero or significantly larger than one. This means that I cannot reject the assumption of utility maximizing individuals. Moreover, the likelihood ratio index is higher compared to that of MNL1. This indicate that letting the ε_i's be correlated within the nests improves the fit to the data. A likelihood ratio test with the null hypothesis $\lambda_1 = \lambda_2 = 1$ gives a p-value of 0.735. Thus, the null hypothesis cannot be rejected. Hence, I cannot conclude that letting the ε_i's be correlated within the nests significantly improves the fit to the data.

The last model I estimate is the random parameters logit model. The random parameters logit model is estimated without the individual attributes as control variables. The reason for this is that it simply is not possible to estimate a random parameters logit model with that many parameters on my data. Moreover, I specify the parameters to follow a normal distribution. A summary of the result is reported in the fifth result column of Table 6.3, where I refer to it as Random Parameters Logit (RPL). The complete result is shown in Table B.5 in Appendix B. The sign of the estimated parameters are the expected for all alternative attribute variables, and they are all also individually significant at the 1%-level. Moreover, a likelihood ratio test with the null hypothesis that parameters are not random gives a p-value of 0.000. Thus, the null hypothesis is rejected. Hence, I conclude that the multinomial logit model's independence assumption is violated and that the random parameters logit model without control variables (RPL) is more appropriate and gives a significantly better fit to the data compared to the multinomial logit model without control variables (MNL1). Thus, in an ideal situation with enough data, a random parameters logit model would probably have been the best model.

Table 6.3: Choice Model Estimates

Dependent Variable: Housing Type					
Variable	MNL1	MNL2	NL1	NL2	RPL
Price	−0.00065	−0.00078	−0.00086	−0.00057	−0.00220
	(0.00020)***	(0.00022)***	(0.00030)***	(0.00025)**	(0.00101)*
Size_Ind	0.24402	0.27458	0.30642	0.23039	1.94415
	(0.04811)***	(0.05363)***	(0.09224)***	(0.07381)***	(0.50523)***
Size_Ind_2	−0.00402	−0.00446	−0.00480	−0.00380	0.03670
	(0.00089)***	(0.00099)***	(0.00155)***	(0.00127)***	(0.01050)***
Central	0.50518	0.51530	0.56322	0.41963	2.66485
	(0.15576)***	(0.17529)***	(0.21406)***	(0.21838)	(1.08647)*
$\ln L(\hat{\boldsymbol{\beta}})$	−370.976	−307.898	−369.330	−370.669	−333.107
LRI	0.052	0.213	0.056	0.053	0.149
N	1380	1380	1380	1380	1380

Notes to Table 6.3: Standard error are in parentheses. *Significant at 10%, **Significant at 5%, *Significant at 1%. Data are sampled from students at Lund University. Table B.1-B.5 in Appendix B are extensions of this table, including complete results. *Price* is a continuous variable, measuring the monthly rent per person in SEK. *Size_Ind* is a continuous variable, measuring the size of an apartment/corridor room per person in square meters. *Size_Ind_2* is the square of *Size_Ind*. *Central* is a dummy variable for centrally located housing units.

From the parameter estimates of *Price*, *Size_Ind* and *Size_Ind_2*, I can now derive estimates of the marginal willingness to pay for size. As described in Section 3.2, the marginal willingness to pay estimates are given by

$$MWTP = -\frac{\beta_1}{\beta_0} - 2\frac{\beta_2}{\beta_0}q,$$

where β_0 is the parameter estimate of $Price$, β_1 is the parameter estimate of $Size\_Ind$, β_2 is the parameter estimate of $Size\_Ind\_2$, and q is the size. Table 6.4 summarizes the marginal willingness to pay estimates derived from the parameter estimates of MNL1, MNL2, NL1, NL2, and RPL.

Table 6.4: Marginal Willingness to Pay Estimates

Model	Marginal Willingness to Pay
MNL1	$375.42 - 12.37q$
MNL2	$352.03 - 11.44q$
NL1	$356.30 - 11.16q$
NL2	$404.19 - 13.33q$
RPL	$883.70 - 33.36q$

Notes to Table 6.4: Estimates of willingness to pay for size, as a function of size, in terms of monthly rent, are reported in SEK.

6.3 Optimal Housing Size Result

As described in Section 3.1, the optimal size is found by equating the marginal willingness to pay and the marginal cost for size. Thus, to get estimates of the optimal size, I equate the marginal willingness to pay estimates in Table 6.4 and the marginal cost estimates in Table 6.2. The result is shown in Table 6.5. The estimates of the optimal size is in the range 18.95-24.83 square meters.

Different estimates of the marginal willingness to pay and the marginal cost result in different estimates of the optimal size, within the range 18.95-24.83. The optimal size is lower for central housing units than for non-central housing units, and also lower for newer housing units than for older housing units. This is due to the fact that the marginal cost estimates are higher for central and newer housing units compared to non-central and older housing units, all other things being equal.

Table 6.5: Optimal Housing Size Estimates

	Older		Newer	
Model	Non-central	Central	Non-central	Central
MNL1	23.95	23.28	20.08	19.42
MNL2	23.85	23.13	19.67	18.95
NL1	24.83	24.09	20.55	19.81
NL2	24.38	23.77	20.80	20.18
RPL	24.12	23.87	22.68	22.44

Notes to Table 6.5: The table summarizes the optimal housing size estimates for older, newer, central and non-central housing units. The estimates are reported in square meters.

The optimal size is also lower when the marginal willingness to pay estimates are derived from the multinomial logit model without individual attributes as control variables (MNL1), compared to to when the marginal willingness to pay estimates are derived from the nested logit and random parameters logit models without individual attributes as control variables (NL1, NL2 and RPL). Also, remember that NL1 and NL2 give a better fit to the data, and RPL gives a significantly better fit to the data, compared to MNL1. Altogether this indicates that the restrictive assumptions of the multinomial logit model causes a downward bias on the optimal size.

The optimal size is also lower when the marginal willingness to pay estimates are derived from the multinomial logit model with individual attributes as control variables (MNL2) compared to when the marginal willingness to pay estimates are derived from the multinomial logit model without individual attributes as control variables (MNL1). This indicates that omission of individual attributes as control variables causes an upward bias on the optimal size. Thus, the optimal size derived from the willingness to pay estimates of the nested logit and random parameters logit models (NL1, NL2 and RPL) would probably have been lower if these models were estimated with individual attributes as control variables.

All in all, the optimal size is probably less than 0.5 square meter less than the estimates of RPL (0.47 square meters is the maximum difference between MNL1 and MNL2, due to omitted variables bias). That is, the optimal size is probably in the range 21.94-24.12 square meters, depending on the age and location of the housing unit.

To examine if the average housing size of the current housing stock is the optimal housing size for the average student I have carried out Monte Carlo simulations. I have found that the average housing size of the current housing stock is within a distance of two standard deviations from the optimal housing size for the average student. Hence, I cannot conclude that the average housing size of the current housing stock is not the optimal housing size for the average student.

7 Conclusion

The optimal housing size estimates for the average student in Lund is in the range 19.0-24.8 square meters. It is found by equating the marginal willingness to pay for size and the marginal cost for size. The marginal willingness to pay is derived from estimates of choice models (multinomial logit, nested logit, and random parameters logit) estimated on data sampled from students at Lund University. The marginal cost is derived from estimates of a hedonic regression model estimated on data from AF Bostäder. Different estimates of the marginal willingness to pay and the marginal cost result in different estimates of the optimal size, within the range 19.0-24.8. However, all in all, when considering potential biases, the optimal size is probably in the range 21.9-24.1, depending on the age and location of the housing unit. The optimal size is lower for central housing units than for non-central housing units, and also lower for newer housing units than for older housing units. This is due to the fact that the marginal cost estimates are higher for central and newer housing units compared to non-central and older housing units, all other things being equal.

To examine if the average housing size of the current housing stock is the optimal housing size for the average student, I have carried out Monte Carlo simulations. Based on the results from the simulations I cannot conclude that the average housing size of the current housing stock is not the optimal housing size for the average student.

The result can be used when planning new student housing in Lund. Since there is a large shortage of student housing, there is a need to build new, and when doing so it is of course desirable to match the new housing with the students' preferences. For this purpose my findings in this essay can be of help. However, I have only focused on the optimal size for the average student. But it is natural to assume that there is a variation in the students' preferences for housing size. Thus, an important issue for future research is to find an estimate of the distribution of the optimal size.

Of course, the method can also be applied to student housing markets in other cities and to other parts of the housing market as long as the estimates from the hedonic regression model can be interpreted as the marginal cost. That is, the data must be generated by a housing company that reinvests all profit into the company. This is the crucial assumption. However, it may hold for, for example, municipal housing companies as well. In that case, considering the large shortage of housing in general, the method can be used to match new housing with consumers' preferences also on other parts of the housing market. Moreover, the method is not only applicable to housing size, but to virtually any housing service there is.

References

AF Bostäder. 2015a. *Available housings and storages.* https://www.afbostader.se/en/available-housings/ (Accessed 2015-11-08).

AF Bostäder. 2015b. *We focus on students!.* https://www.afbostader.se/en/about-us/ (Accessed 2015-10-14).

Bhat, C. 1998. Accommodating variations in responsiveness to level-of-service variables in travel mode choice models. *Transportation Research Part A: Policy and Practice*, 32, 495-507.

Brownstone, D. & Small, K. 1989. Efficient Estimation of Nested Logit Models. *Journal of Business & Economic Statistics*, 7, 67-74.

Brownstone, D. & Train, K. 1999. Forecasting new product penetration with flexible substitution patterns. *Journal of Econometrics*, 89, 109-129.

Börsch-Supan, A. & Hajivassiliou, V.A. 1993. Smooth unbiased multivariate probability simulators for maximum likelihood estimation of limited dependent variable models. *Journal of Econometrics.* North-Holland, 58, 347-368.

Cameron, C.A. & Trivedi, P.K. 2005. *Microeconometrics: Methods and Applications.* Cambridge University Press.

Chipman, J. 1960. The foundations of utility. *Econometrica*, 28, 193-224.

Court, A. T. 1939. Hedonic price indexes with automotive examples. In *The Dynamics of Automobile Demand.* General Motors, 98-119.

Debreu, G. 1960. Review of individual choice behavior by R.D.Luce. *American Economic Review*, 50, 186-188.

Goodman, A.C. 1998. Andrew Court and the Invention of Hedonic Price Analysis. *Journal of Urban Economics*, 44, 291-298.

Griliches, Z. 1961. Hedonic prices for automobiles: An econometric analysis of quality change. In *The Price Statistics of the Federal Government.* National Bureau of Economic Research, 173-196.

Griliches, Z. 1991. Hedonic price indexes and the measurement of capital and productivity: Some historical reflections. In E.R.Berndt & J.E.Triplett eds., *Fifty Years of Economic Measurement: The Jubilee of the Conference on Research in Income and Wealth.* University of Chicago Press, 185-206.

Hajivassiliou, V. & Ruud, P. 1994. Classical estimation methods for ldv models using simulation. In R.Engle & D.McFadden, eds., *Handbook of Econometrics.* North-Holland, 4, 383–441.

Hanemann, W.M. 1991. Willingness to Pay and Willingness to Accept: How Much Can They Differ? *The American Economic Review*, 80, 635-647.

Hansson, B. 2014. *Låst läge på bostadsmarknaden, Marknadsrapport, maj 2014.* Boverket. http://www.boverket.se/globalassets/publikationer/dokument/2014/marknadsrapport-maj-2014.pdf (Accessed 2015-10-12).

Hensher, D.A. & Greene, W.H. 2003. The Mixed Logit model: The state of practice. *Transportation*, 30, 133-176.

Hicks, J.R. 1946. *Value and capital: an inquiry into some fundamental principles of economic theory.* 2nd ed. Oxford University Press.

Kain, J.F. & Quigley, J.M. 1970. Measuring the Value of Housing Quality. *Journal of the American Statistical Association*, 65, 532-548.

Luce, R.D. 1959. *Individual Choice Behavior*, John Wiley and Sons.

Luce, R.D. & Suppes, P. 1965. Preferences, utility and subjective probability. In

R.D.Luce, R.Bush & E.Galanter, eds., *Handbook of Mathematical Psychology.* John Wiley and Sons, 249-410.

Marschak, J. 1960. Binary choice constraints on random utility indications. In K.Arrow, S.Karlin & P.Suppes, eds., *Mathematical methods in the social sciences, 1959 : proceedings of the First Stanford Symposium.* Stanford University Press, 312-329.

Mas-Colell, A., Whinston, M.D. & Green, J.R. 1995. *Microeconomic Theory.* Oxford University Press.

McFadden, D. 1974. Conditional logit analysis of qualitative choice behavior. In P.Zarembka, ed., *Frontiers in Econometrics.* Academic Press, 105-142.

McFadden, D. 1978. Modeling the choice of residential location. In A.Karlqvist, L.Lundqvist, F.Snickars & J.Weibull, eds., *Spatial Interaction Theory and Planning Models.* North-Holland, 75-96.

McFadden, D. & Train, K. 2000. Mixed mnl models of discrete response. *Journal of Applied Econometrics*, 15, 447-470.

Mäler, K.G. 1974. *Environmental economics: a theoretical inquiry.* Johns Hopkins University Press.

Nilsson, I. 2014. *Minirevolutionen: En konsekvensanalys av ändrade byggregler för utformning av studentbostäder.* Thesis. Luleå tekniska universitet. http://pure.ltu.se/portal/sv/studentthesis/minirevolutionen(675c4e27-e002-4056-8c28-5619dd281889).html (Accessed 2015-10-12).

O'Sullivan, A. 2012. *Urban Economics.* 8th ed. McGraw-Hill.

Revelt, D. & Train, K. 1998. Mixed logit with repeated choices: Households' Choices of Appliance Efficiency Level. *The Review of Economics and Statistics*, 80, 647-657.

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

Studentbostadsföretagen. 2015. *STUDBOGUIDEN 2015: En genomgång av student-bostadssituationen på landets studiorter*. http://studentbostadsforetagen.se/wp-content/uploads/2015/08/Studboguiden-2015.pdf (Accessed 2015-10-14).

Studentlund. 2015. *AFB manages over 5900 homes and is southern Sweden's largest student housing company*. http://www.studentlund.se/bostad/af-bostader/ (Accessed 2015-10-14).

Torres, I., Greene, M. & Ortúzar, J. 2013. Valuation of housing and neighbourhood attributes for city centre location: A case study in Santiago. *Habitat International*, 39, 62-74.

Train, K. 2002. *Discrete Choice Methods with Simulation*. Cambridge University Press. http://eml.berkeley.edu/books/choice2.html (Accessed 2015-09-28).

Wilhelmsson, M. 2000. The Impact of Traffic Noise on the Values of Single-family Houses. *Journal of Environmental Planning and Management*, 43, 799-815.

A Questionnaire

(See next three pages for an example of the questionnaire)

Student housing questionnaire

- In answering this questionnaire you are contributing with valuable information to my master essay.
- I will use the information to analyze the housing situation for students to find out if the housing alternatives available matches the students preferences and, consequently, how the alternatives available can be altered to better fit the students preferences.
- Please choose your most preferred alternative from the following hypothetical alternatives.
- Assume that all alternatives are student apartments/corridor rooms.
- Assume the following for all apartments and corridor rooms:
  - Electricity included
  - Balcony not available
  - Heating/Water included
  - Furniture not included
  - Internet access available
  - Payment period 12 months
  - It is *your* monthly rent that is stated for the various alternatives
- Assume the following for corridor rooms:
  - Common kitchen in corridor
  - Private shower
  - Small corridors accommodate 8 students
  - Large corridors accommodate 16 students
  - Assume the following for the shared corridor room alternative
    * Assume that you and another student jointly decide to share a corridor room
  - All features not shown for corridor rooms are identical for all corridor rooms
- Assume the following for all single-room apartments:
  - Private kitchen
  - Private shower
  - All features not shown for single-room apartments are identical for all single-room apartments
- Assume the following for all shared 2-room apartments:
  - Private kitchen
  - Private shower
  - Assume that you and another student jointly decide to share a 2-room apartment
  - All features not shown for shared 2-room apartments are identical for all shared 2-room apartments

Table 1: Please choose your most preferred alternative from the following hypothetical alternatives.

Your monthly rent	Housing type	Size	Distance to city centre	Choose one alternative
2700 kr	Shared room in a small corridor	32 m^2	0.4–1.0 km	□
3600 kr	Room in a small corridor	20 m^2	0.4–1.0 km	□
4500 kr	Room in a large corridor	36 m^2	1.0–2.6 km	□
3100 kr	Single-room apartment	14 m^2	1.0–2.6 km	□
3100 kr	Shared 2-room apartment	44 m^2	0.4–1.0 km	□

Table 2: Please choose your most preferred alternative from the following hypothetical alternatives.

Your monthly rent	Housing type	Size	Distance to city centre	Choose one alternative
4700 kr	Shared room in a small corridor	64 m^2	1.0–2.6 km	☐
2400 kr	Room in a small corridor	12 m^2	0.4–1.0 km	☐
2300 kr	Room in a large corridor	14 m^2	1.0–2.6 km	☐
3400 kr	Single-room apartment	22 m^2	1.0–2.6 km	☐
2900 kr	Shared 2-room apartment	36 m^2	0.4–1.0 km	☐

Table 3: Please choose your most preferred alternative from the following hypothetical alternatives.

Your monthly rent	Housing type	Size	Distance to city centre	Choose one alternative
1600 kr	Shared room in a small corridor	20 m^2	0.4–1.0 km	☐
6000 kr	Room in a small corridor	38 m^2	0.4–1.0 km	☐
2100 kr	Room in a large corridor	10 m^2	1.0–2.6 km	☐
6300 kr	Single-room apartment	38 m^2	1.0–2.6 km	☐
5600 kr	Shared 2-room apartment	68 m^2	0.4–1.0 km	☐

About you

1. **Is Lund University your home university?**
   - ☐ Yes
   - ☐ No, I am an exchange student
2. **What is you primary level of study?**
   - ☐ Bachelor's level
   - ☐ Master's level
   - ☐ Doctoral level
3. **What faculty is your home faculty at Lund University?**
   - ☐ Faculty of Engineering
   - ☐ Faculty of Science
   - ☐ Faculty of Law
   - ☐ Faculty of Social Sciences
   - ☐ Faculty of Medicine
   - ☐ Faculties of Humanities and Theology
   - ☐ School of Economics and Management
4. **How do you live?**
   - ☐ I live in a corridor room
   - ☐ I live alone in an apartment/house
   - ☐ I share an apartment/house
5. **What is your monthly rent?**
   My monthly rent is: ________ kr
6. **What is your approximate average monthly budget?**
   My monthly budget is: ________ kr
7. **How old are you?**
   I am ________ years old
8. **What is your marital status?**
   - ☐ Single
   - ☐ In a relationship
   - ☐ Engaged
   - ☐ Married
   - ☐ Other

9. What is your gender?

- ☐ Male
- ☐ Female
- ☐ Other

B Complete Results

Table B.1: Complete Result: MNL1

Alternative Attributes		$\ln L(\hat{\boldsymbol{\beta}})$	LRI	N
$Price$	-0.00065	-370.976	0.052	1380
	(0.00020)***			
$Size\_Ind$	0.24402			
	(0.04811)***			
$Size\_Ind\_2$	-0.00402			
	(0.00089)***			
$Central$	0.50518			
	(0.15576)***			
		Alternatives		
Individual Attributes	CRSC	CRLC	SRA	STRA
$Constant$	1.55236	1.33147	2.62456	2.20715
	(0.36010)***	(0.34939)***	(0.36269)***	(0.32455)***

Notes to Table B.1: Standard error are in parentheses. *Significant at 10%, **Significant at 5%, *Significant at 1%. CRSC: Corridor room in a small corridor. CRLC: Corridor room in a large corridor. SRA: Single-room apartment. STRA: Shared two-room apartment.

Table B.2: Complete Result: MNL2

Alternative Attributes		$\ln L(\hat{\boldsymbol{\beta}})$	LRI	N
Price	−0.00078	−307.898	0.213	1380
	(0..00022)***			
Size_Ind	0.27458			
	(0.05363)***			
Size_Ind_2	−0.00446			
	(0.00099)***			
Central	0.51530			
	(0.17529)***			
		Alternatives		
Individual Attributes	CRSC	CRLC	SRA	STRA
Exchange	15.34	15.03	14.78	14.57
	(1484.19)	(1484.19)	(1484.19)	(1484.19)
Bachelor	0.62	−0.22	0.25	0.39
	(1.01)	(0.94)	(0.88)	(0.89)
Engineering	−19.49	−17.29	−15.41	−17.32
	(1516.75)	(1516.75)	(1516.75)	(1516.75)
Science	−1.48	−0.98	−0.63	−1.76
	(2079.47)	(2079.47)	(2079.47)	(2079.47)
Law	−5.25	−2.38	−0.40	−3.23
	(1626.19)	(1626.19)	(1626.19)	(1626.19)
Social	−17.08	−15.96	−13.31	−16.21
	(1516.75)	(1516.75)	(1516.75)	(1516.75)
Humanities	−2.37	−3.16	−0.39	−3.42
	(1957.67)	(1957.67)	(1957.67)	(1957.67)
Economics	−17.00	−15.26	−13.56	−15.34
	(1516.75)	(1516.75)	(1516.75)	(1516.75)
Corridor	1.61	1.10	0.78	−0.11
	(1.39)	(1.38)	(1.33)	(1.32)
Alone	−0.85	−0.66	−0.13	−2.21
	(1.20)	(1.14)	(1.08)	(1.07)**
Rent_1000	0.65	0.25	0.62	0.85
	(0.51)	(0.49)	(0.45)	(0.45*)
Age	−0.24	0.05	−0.01	−0.08
	(0.19)	(0.17)	(0.16)	(0.16)
Relationship	−0.49	−1.27	0.17	0.16
	(0.99)	(0.92)	(0.84)	(0.84)
Male	2.60	1.63	1.28	1.79
	(1.07)**	(1.03)	(0.97)	(0.98)*
Constant	20.56	14.77	13.60	16.93
	(1516.76)**	(1516.75)	(1516.75)	(1516.75)

Notes to Table B.2: Standard error are in parentheses. *Significant at 10%, **Significant at 5%, *Significant at 1%. CRSC: Corridor room in a small corridor. CRLC: Corridor room in a large corridor. SRA: Single-room apartment. STRA: Shared two-room apartment.

Table B.3: Complete Result: NL1

Alternative Attributes		$\ln L(\hat{\boldsymbol{\beta}})$	LRI	N
Price	-0.00086	-369.330	0.056	1380
	(0.00030)***			
Size_Ind	0.30642			
	(0.09224)***		λ-parameters	
Size_Ind_2	-0.00480		Corridor	0.86815
	(0.00155)***			(0.33152)
Central	0.56322		Apartment	1.85577
	(0.21406)***			(0.88515)
		Alternatives		
Individual Attributes	CRSC	SCRSC	SRA	STRA
Constant	0.26384	-1.25821	0.68895	0.25949
	(0.23659)	(0.49690)**	(0.61245)	(0.54181)

Notes to Table B.3: Standard error are in parentheses. *Significant at 10%, **Significant at 5%, *Significant at 1%. CRSC: Corridor room in a small corridor. SCRLC: Shared corridor room in a small corridor. SRA: Single-room apartment. STRA: Shared two-room apartment.

Table B.4: Complete Result: NL2

Alternative Attributes		$\ln L(\hat{\boldsymbol{\beta}})$	LRI	N
Price	-0.00057	-370.669	0.053	1380
	(0.00025)**			
Size_Ind	0.23039			
	(0.07381)***		λ-parameters	
Size_Ind_2	-0.00380		One student	0.91044
	(0.00127)***			(0.37977)
Central	0.41963		Two students	0.56821
	(0.21838)*			(0.37179)
		Alternatives		
Individual Attributes	CRSC	SCRSC	SRA	STRA
Constant	0.18801	-0.55504	1.14511	0.79013
	(0.22547)	(0.63715)	(0.48180)**	(0.53700)

Notes to Table B.4: Standard error are in parentheses. *Significant at 10%, **Significant at 5%, *Significant at 1%. CRSC: Corridor room in a small corridor. SCRLC: Shared corridor room in a small corridor. SRA: Single-room apartment. STRA: Shared two-room apartment.

Table B.5: Complete Result: RPL

Alternative Attributes		$\ln L(\hat{\boldsymbol{\beta}})$	LRI	N
$Price$	-0.00220	-333.107	0.149	1380
	(0.00102)**			
$Size\_Ind$	1.94415			
	(0.50523)***			
$Size\_Ind\_2$	-0.03670			
	(0.01050)***			
$Central$	2.66485			
	(1.08647)**			
		Alternatives		
Individual Attributes	CRSC	CRLC	SRA	STRA
$Constant$	6.13054	2.00398	6.95347	6.00910
	(1.52847)***	(1.25141)	(1.69955)***	(1.58074)***
		$\boldsymbol{L}$		
l11 l32 l63 l65:	4.1546	6.2701	0.6414	0.2112
	(1.4425)***	(2.2480)***	(0.2691)**	(0.4027)
l21 l42 l73 l75:	5.2928	-6.2524	-0.0019	-0.0059
	(1.9840)***	(2.0143)***	(0.0044)	(0.0071)
l31 l52 l83 l85:	-3.7798	-0.0015	4.5547	0.2595
	(1.6883)**	(0.0010)	(1.4064)***	(0.8538)
l41 l62 l44 l66:	1.7153	0.0581	-7.3952	-0.5161
	(1.2536)	(0.3637)	(2.3496)***	(0.2252)**
l51 l72 l54 l76:	-0.0042	-0.0054	-0.001	0.0100
	(0.0011)***	(0.0069)	(0.0008)	(0.0044)**
l61 l82 l64 l86:	0.4335	0.7448	0.3438	-2.3001
	(0.2217)*	(0.7384)	(0.2631)	(0.8252)***
l71 l33 l74 l77:	-0.0005	-1.6443	-0.0068	-0.0015
	(0.0039)	(1.2345)	(0.0047)	(0.0011)
l81 l43 l84 l87:	3.5486	-9.1848	-2.0485	1.2767
	(1.0696)***	(2.4860)***	(1.0900)*	(0.6646)*
l22 l53 l55 l88:	6.9701	-0.0012	0.0013	0.5139
	(2.1968)***	(0.0012)	(0.0008)*	(0.5465)

Notes to Table B.5: Standard error are in parentheses. *Significant at 10%, **Significant at 5%, *Significant at 1%. CRSC: Corridor room in a small corridor. CRLC: Corridor room in a large corridor. SRA: Single-room apartment. STRA: Shared two-room apartment. The the covariance matrix for the random coefficients is given by $\boldsymbol{V} = \boldsymbol{L}\boldsymbol{L}'$