Abstract

In Norway, the average house size has increased with 8.4 m^2 between 2001 and 2013. During the same time, households have experienced a huge increase in disposable income, house prices have reach unprecedentedly high levels and the size of households have been reduced. This paper examines which forces determine the owner-occupiers' demand for house size in Norway and why the average house size has increased between 2001 and 2013. Using household level data from Statistics Norway's Survey of Living Conditions, I estimate an identical cross-sectional regression for each of the five survey years, regressing the number of square meter house size on income, price, household characteristics, geographical and structure related variables. Inserting average values for all explanatory variables, I find given the explanatory variables, higher predicted house size each survey year. The increase in predicted house size exceeds the increase in the average square meters indicating changes inherent in my explanatory variables combined reduce demand for house size. Specifically, a trend towards smaller households drives down the size of dwellings, all else equal. Instead, macroeconomic variables including the interest rate, credit availability and price expectations can help explain why we live bigger.

Preface

This thesis is the final requirement for the Master of Science in Economics at the Norwegian University of Science and Technology (NTNU).

First, I would like to thank my supervisor, Professor Fredrik Carlsen for trusting me with this project and providing helpful guidance along the way. Second, I would express my appreciation to those who have proofread this paper. Finally, a special thanks to Anne Larsen Viken at the Department of Economics, whom in a devoted manner has helped me facilitating studies abroad and all other practical issues occurring during my time as a student.

The data applied in the present thesis are retrieved from Statistics Norway's (SSB) Survey of Living Conditions 2001, 2004, 2007, 2012 and 2013. The data sets have been anonymized and made available by the Norwegian Social Science Data Services (NSD). Neither NSD nor SSB are responsible for the analyses or interpretations presented in this paper.

Trondheim, August, 2015
Tor Midtbø

Table of Contents

List of Figures

List of Tables

1 Introduction

Everyone needs a place to stay and a house is an essential commodity as it gives shelter. On the other hand, once the basic needs are met, housing can be a way of showing off status. Housing represents a special and complex commodity quite different from a standard homogeneous good. Households are different in their characteristics and preferences which make the households demand different sizes and styles of housing. Not only is the structure itself of importance to the demand but also the land on which it sits. This is because a home is spatial immobile and households care about neighborhood characteristics as well as distance to commercial areas. Further, a dwelling is highly durable and is physically modifiable. All these features separate the housing good from the standard textbook homogeneous commodity. Consequently, houses represent imperfect substitutes with a varying degree of substitutability between the housing markets. In addition, owner-occupied housing is for most households their largest investment and housing as a capital asset substantially affects demand. An analysis taking all these features of the housing good into consideration is highly problematic because of the complex nature of the housing good. This paper tries to overcome some of these issues by only studying one part of the housing demand, namely the demand for house size. A way of studying housing demand is by regarding housing as a service. In this setting, housing is appreciate as a flow and defined as the amount of housing consumed over a period. House size can be seen as a simple type of housing service measuring the most important physical component of a house.

The first main goal of this thesis is to study which forces determine the demand for house size for owner-occupiers in Norway. More specific, I use the cross sectional Survey on Living Conditions from Statistics Norway for 2001, 2004, 2007, 2012 and 2013. The variable of interest, house size measured in square meters is regressed on a span of explanatory variables including income, house price, household characteristics, regional and house specific variables. Identical regressions are run each survey year and estimated coefficients and income- and price elasticity are discussed both for the separate years and changes over time.

The second main concern of the paper is to explain the increased house sizes seen over a 13 year period. Inserting for the same values of all the explanatory variables all years let us study the effect on demand for house size given all the right side variables. In the time period of interest, households have experienced an increase in their disposable incomes. House prices increased – especially in the cities, while households move away from the traditional family structure (SSB, 2015b). In particular, how these changes affect demand for house size is studied through a shift analysis. Other factors important for the Norwegian housing demand are discussed including the interest rate, credit availability, expectations and the tax system. Later, the analysis is carried out exclusively for households which feel they live in a right-sized dwelling in order to get a closer estimate of households' ideal demand and to test the robustness of my main results. Finally, a chapter on econometric challenges and limitations reveals some challenges and drawbacks of the study with particular concern of omitted variable bias and measurement error. But first, I look into the existing literature on the demand for housing services.

2 Background and literature

The concept of housing demand has been ambiguously understood in the economic literature. However, Rothenberg (1991) categorized the vast literature on housing demand into four groups. That is the demand for housing services, the individual housing attributes of the composite housing good, the demand for home ownership versus renting (tenure choice) and the spatial allocation of the household. Only the first group will be discussed in this paper.

2.1 Literature on the demand for housing services

Looking at housing as a service, let us appreciate housing as a flow changing over time similar to other economic variables such as money. The term is defined as the amount of housing consumed over a period. Demand for housing services is determined by the willingness to pay of consumers, which is made up by the ability to pay, housing needs and preferences or "taste" for housing. Especially the last one is hard to measure but one way to do so is to look at housing expenditure as a proxy for housing services. The willingness to pay is highly correlated with housing expenditures since persons with high willingness to pay also on average pay more and have larger houses which increase mortgage and maintenance costs, payments included in the housing expenditures. Those with "taste" for better housing also demand more from the dwelling, increasing housing expenditures. A normal way of estimating the demand in this setting is by using housing expenditure as a proxy for services and regressing it on the price of housing, price of non-housing expenditure and socioeconomic variables including income (Zabel, 2004).

2.1.1 Income and price elasticity of demand

Studying housing services is a way to uncover the income and price elasticity of demand. Considering elasticity, let us understand consumers' behavior of the housing good. Specifically, income elasticity gives the percentage demand change from a percentage increase in income. Price elasticity has the same interpretation although from a house price increase.

The income elasticity largely depends on the type of income we are using. With perfect capital markets, demanders will only care about their permanent income as they can freely save or borrow money if their current income differs from the permanent. Following the argument, consumers do not change their housing demand when changes in income are transitory, but rather look over the time horizon making decisions on the permanent income, thereby smoothing their consumption (Friedman, 1957; Goodman & Kawai, 1982). Assuming the permanent income hypothesis holds, an analysis using the current income instead of the permanent will under some assumptions give a downward bias for the income elasticity because of the measurement error of including the transitory income. An illustration of bias follows:

$$E(\hat{\eta}_y) = \eta_y P_y$$

where $\hat{\eta}_y$ is an estimator of the true permanent income elasticity η_y. P_y is the ratio of the variance of permanent and current income given by:

$$P_y = \frac{\sigma_{yp}^2}{\sigma_y^2} = \frac{\sigma_{yp}^2}{\sigma_{yp}^2 + \sigma_{yt}^2} < 1$$

Where $\sigma_{yp}^2$is the variance of the permanent income, σ_{yt}^2 of the transitory and σ_y^2 of the current income. The variance of the current income equals the sum of the variances of the permanent and transitory income. Here we assume that the permanent and transitory incomes are uncorrelated. The size of the downward bias increases with the relative size of the transitory income variance compared to the one of the permanent income (Mayo, 1981). Unfortunately, the permanent income is often not measurable. Different approaches have been used including using aggregated data which hopefully average out the transitory income, using lagged income as instrumental variable or using average of income over several years (Mayo, 1981). A more general discussion on measurement error and the relevance for my thesis is found in Chapter 7.2.

There is, on the other side, also a reason to include the transitory income in the housing demand. First, the capital markets are not perfect and households cannot freely borrow on their lifetime earnings. Another reason is that buying a house is a large part of a household's investment portfolio – often the largest they ever make. A positive transitory income would imply higher savings that would raise the housing demand since the household wants to invest more in housing since it is a part of their investment portfolio. In addition, higher consumption also follows from a transitory income increase. To separate this effect, the permanent and transitory effects should be separated to avoid estimation biases (Goodman & Kawai, 1982).

In an influential paper, Mayo (1981) sums up the existing findings from the U.S. housing market and shows large discrepancies in the estimated income and price elasticity of demand. Looking at the income elasticity reveals a substantial difference between aggregated and micro data. Some of the aggregated studies which use averages within metropolitan areas instead of individual household data show income elasticity above unity. The studies using micro data and permanent income for owner-occupiers show elasticity in the range from 0.36 to 0.87 with most in the range from 0.5 to 0.7. With regard to price elasticity, the results depend on how the price variable is defined, however, most sensible estimates lies within the range -0.3 to -0.9. Interestingly both price and income elasticity increase in absolute value with income. The higher the income the larger is the negative effect of a percentage increase in house price on the housing demand. Moreover, when the household gets richer, a larger portion of a wage increase goes towards housing but still less than the increase in income as the elasticity lies below unity.

In a relatively newer article, Zabel (2004) uses data from 1993 and 2001 for 36 and 38 metropolitan areas in the United States. He first estimates the housing price using hedonic regression in order to find the implicit price for housing services. The dependent variable is the natural log of owner's price valuation of his/her house and property. Explanatory variables include the number of rooms, bathrooms, and bedrooms, the age of the house and whether a garage is present. In addition, other variables for owner characteristics including the logarithm of permanent income, age and dummy variables for male, white and married. The hedonic method allows for estimation of the marginal price of each characteristic, all important for the price valuation. Thus, the value of a constant quality house can be computed. This hedonic regression is performed on each of the metropolitan areas and a price index is created. Then, the demand for housing services is estimated with the price valuation over the price index as a depended variable. The model is estimated both with current and permanent income. In 2001 the elasticity is 0.36 when permanent income is used and 0.166 for current income. The elasticity, here looking at permanent income, also increases with the level of income as observed in the paper by Mayo (1981). Individual income of owners in the 10 to 20^{th} income percentile has an elasticity of 0.16 while for the richer group of the 80 to 90^{th} percentile has an income elasticity of 0.64 (Zabel, 2004).

Looking at Norwegian data, Larsen (2014) estimates the income elasticity for owner occupied housing consumption. The house owner's consumption of housing services is, unlike the renter, not observable since he does not pay rent. To overcome this challenge, the author estimates the rental value an owner would pay if the owner rented his/her own house. This procedure is based on the rental equivalence principle that assumes owning and renting are perfect substitutes with no self-selection in tenure caused by household income, family size and house size. The fact that the Norwegian rental market is small and partly regulated speaks in disfavor of the rental equivalence principle, which is found not to hold (Nesbakken, 2008). Still this method is used by many statistical agencies including Statistics Norway (SSB) to estimate the owner-occupiers' housing consumption used in the Consumer Price Index. Values of rent from SSB's rental survey of 2007 by is estimated with a hedonic regression onto housing attributes like house size, geographical location and municipality size giving a partial price for each housing and spatial component. Information on owner-occupiers, taken from the Consumer Expenditure Survey, is inserted for tenants with similar housing attributes, thus assigning what households would pay in rent if they had rented their own house, also called “imputed rent”. The income elasticity is found by regressing the income share of imputed rent on gross income and household characteristics. Evaluated at the mean income, yields an elasticity of 0.32 which is close to the estimate by Zabel (2004). Also here the elasticity is clearly heterogeneous with the level of income. Income levels in the 10 to 20th percentile give an elasticity of demand of 0.10 while it is 0.59 for the 80 to 90th income percentile, which also is comparable to the estimates found by Zabel (2004) (Larsen, 2014).

2.1.2 Life cycle and homeownership

The demand for housing services can be affected by the age of the decision maker. Artle and Varaiya (1978) present a life cycle model for homeownership under a set of borrowing constraints. According to their theory the owner rate should first increase with age as income and savings rise. But to keep a good standard of living after retirement, individuals should release home equity either by downsizing or taking up a mortgage. A large debate started when Mankiw and Weil (1989) showed that the housing demand would move in a hump-shaped curve with age in coherence with the life cycle model. Specifically, they expected that when the baby-boomer cohorts in the United States reached retirement, it would put a downward pressure on the demand and house prices. R. Green and Hendershott (1996) agreed the demand would fall with age if characteristics like income, education and household size could vary with age. The reason for lower demand was that older households in 1980 had substantially lower levels of education than the younger households at the time. When income and education are held constant, however, then the demand would slightly increase from the late 1950s indicating, in spite of empty nests and lower income, that individuals maintain or even increase their housing demand as they age.

Looking at newer European data, Angelini and Laferrere (2011) found low residential mobility for European elderly compared to other age groups. More specifically, they estimated an annual mobility rate of 2 percent correcting for non-random drop outs from the sample. The rate differed from 4.4 percent in Denmark and Sweden to 1.0 percent in Austria and 0.3 percent in Greece.

Reasons for not moving include monetary and especially psychological transaction costs. For instance only a small portion of the population is actively looking for new accommodation, and those who do will only acquire information about a few housing options, not the whole market. A reason why is the time consuming process of attaining sufficient information about the complex and heterogeneous good a house represents (R. K. Green & Malpezzi, 2003; Rothenberg, 1991). In particular we can think of being out on the market actively searching for new dwellings is more costly for elderly than the population in general. This reflects the argument that elderly may to a lesser degree than others act according to their underlying demand, contributing to market inefficiencies. The elderly may regard their house as a secure asset in case of need and their home equity as long-term care insurance. Another argument against the life cycle model is the motive of bequest, making elderly remain homeowners. For among these reasons, high proportions of elderly own their homes, with homeownership rates higher than 70 percent for those aged between 50 to 79 years in most European countries (Angelini & Laferrere, 2011). The fact that elderly stay in large single homes long after their offspring leave the nest is also a tendency found in my dataset (Sandlie & Grødem, 2013).

The few elderly that move tend to downsize their housing, either by moving to smaller dwellings or nursing homes/institutions. Income halts this process in two ways: First, high income induces less need to capitalize on housing value by moving to smaller accommodations or going from homeowners to renters. Lower income individuals often have to stay in their dwelling, since they lack the funds to move. Hence, mobility is twice income constrained; some are prevented to move, while others would like to stay in their homes longer (Angelini & Laferrere, 2011).

2.1.3 Veblen and neighborhood effects

Broadly speaking, owner-occupied housing represents a composite commodity fulfilling several human needs. First, housing is a basic consumption good as every individual needs shelter, and in that perspective it constitutes a necessity. Once shelter is preserved, then the consumer wants larger and better housing to attain comfort, and later, status. The Norwegian-American economist Thorstein Veblen introduced this concept formally in his famous book "Theory of the leisure class" ([1899]1994). The book lays out a framework where preferences are determined socially in contrast to the neoclassical theory of consumption, where utility is maximized with exogenous preferences and the demand decision only depends on the individual decision maker. More directly, he introduced the term "conspicuous consumption" which is defined as spending money on luxurious goods and services to publicly display economic power and status. According to Veblen, all classes of society engage in this behavior as they try to mimic the consumption patterns of the class above theirs. As a consequence, consumers pay more for a good than the intrinsic utility suggests making the whole society worse off by engaging in overconsumption (Trigg, 2001).

One of the goods subject for conspicuous consumption is housing. In the United States, the trend has been towards larger homes with an increase in average house size of new detached houses from 100m^2 in the 1950's to 217m^2 in 2002 (Wilson & Boehland, 2005). The term "the bigger the better" has been an important force in the American housing market, making people buy the largest houses they can afford. Further, buying larger homes go hand in hand with other forms of conspicuous consumption since a large home is necessary to store all the other things we "need", including multiple forms of entertainment *et cetera* (Wilson & Boehland, 2005). Critique of the conspicuous consumption theory has questioned the one-way tickle down of preferences from the upper to the lower classes. His theory is also criticized for lacking subtlety and sophistication in assuming that buying large and expensive is the main way of showing status. Instead, critics focus on the postmodern consumer freeing himself from the traditional social structures by individually giving meaning to commodities (Trigg, 2001).

There is evidence for a strong social element in the housing demand. The American authors Ioannides and Zabel (2003) have shown that the individual housing demand in United States has largely affected the mean housing demand of neighbors. This effect, also called "neighborhood effects", say that neighbors' decisions to fix, renovate or make additions to their homes induce increased housing consumption. The authors also find evidence of a contextual neighborhood effect where housing demand is influenced by the neighbors' characteristics. This effect can be interpreted that housing preferences not only is depended on the individual, but also being affected by its neighbors.

3 Theory and data

In this section a theoretical model is presented that let us determine equilibrium house size by maximizing the household's utility function under a budget constraint. Next, I look into the data material used for the empirical analysis, where each included variable is studied separately. Finally, descriptive statistics is presented with particular concern on how the average values change over the time period.

3.1 Theory on house size

The household utility function can be written as:

$$U(Size, X, \alpha) \tag{1.0}$$

where

Size = number of square meter house size

X = consumption of all other goods

α = a vector of parameters which characterize the household preferences for house size.

The household budget constraint is given by:

$$P_A \cdot Size + P_x \cdot X = INC \tag{1.1}$$

where

P_X = price of other consumption goods

INC = Household income

P_A = the user cost of house size per square meter for owner-occupiers. Similar to the price of housing services discussed in the literature, although here only referring to house size and is given per square meter. The user cost can as a simplification be described as a function of square meter sales price and the real interest rate.

$$P_A = P(P_B, r)$$

where

P_B = market price per square meter house size for owner-occupiers.

r = real interest rate

A household maximizes its utility given by (1.1) with respect to house size and other consumption goods subject to the budget constraint given by (1.2).

Max $U(Size, X, \alpha)$

subject to

$P(P_B, r) \cdot Size + P_x \cdot X = INC$

inserting X from the budget constraint (1.2) into the optimization problem gives

$$\underset{Size,X}{\text{Max}}\ U\left(Size, \left[\frac{INC}{P_x} - \frac{P(P_B,r)\cdot Size}{P_x}\right], \alpha\right)$$

then maximizing with respect to size gives the first order condition

$$U_{size}(Size, INC, \alpha) - U_X(Size, INC)\frac{P(P_B,r)}{P_x} = 0 \qquad (1.3)$$

As a simplification, I assume the price of other consumption goods is normalized to one, that is, $P_X = 1$. Further, the marginal utility of other consumption goods does not depend on house size assuming only a limited share of the household budget is used on house size. U_X now only depends on household income. The first order condition can now be written as:

$$U_{size}(Size, INC, \alpha) - U_X(INC) \cdot P(P_B, r) = 0 \qquad (1.4)$$

With price normalized to one, the marginal utility of other consumption goods is equal to the marginal utility of income given as U_{INC}, since the marginal utility of other goods now only depends on household income. The simplified first order condition is given by (1.5).

$$U_{size}(Size, INC, \alpha) - U_{INC}(INC) \cdot P(P_B, r) = 0 \qquad (1.5)$$

Further, I assume that the marginal utility of house size no longer depends on the household income. In other words, a square meter house size increase yields the same gain in utility regardless of the household income level. The final simplified first order condition is given by (1.6).

$$U_{Size}(Size, \alpha) = P(P_B, r) \cdot U_{INC}(INC) \tag{1.6}$$

The optimal number of square meter is found where the marginal utility of house size equals the marginal utility of income times a price function depending on price per square meter and the real interest rate.

$$\frac{\partial U}{\partial Size} = U_{size} > 0 \,, \; \frac{\partial U}{\partial X} = U_X > 0, \; \frac{\partial^2 U}{\partial Size^2} < 0, \quad \frac{\partial^2 U}{\partial X^2} < 0$$

There is a positive marginal utility of house size and of other consumption goods (or equivalent of income). Negative second derivatives assure us a maximum point to the optimization problem is found. Alpha is a shift parameter describing the household preferences for house size. A high alpha indicates the household highly prefer a large home, thus the parameter is defined so that the marginal utility of house size increases with alpha.

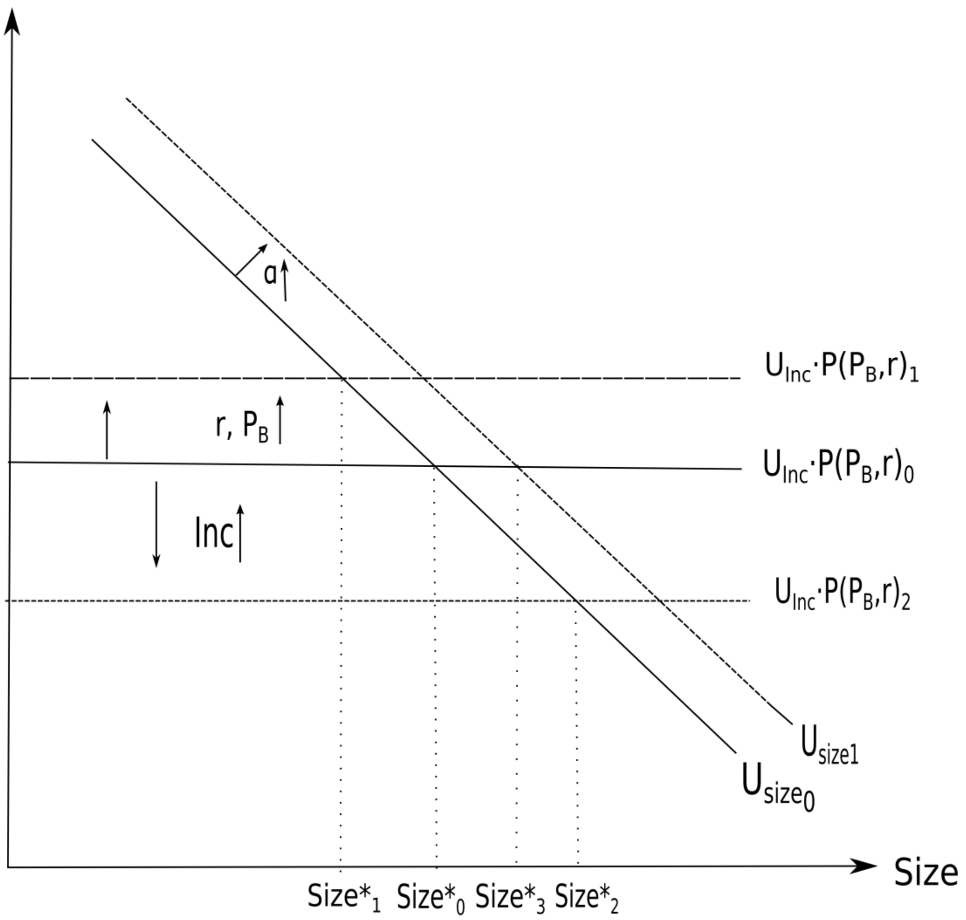
$$\frac{\partial U_{Size}}{\partial \alpha} > 0$$

Finally, equilibrium house size can be presented in a graph shown in Figure 3.1 with house size and its marginal utility on the axes. $P(P_B, r) \cdot U_{INC}(Inc)$ does not depend on house size, hence given as a horizontal line in the diagram. As a consequence of the negative second derivative of the utility function with respect to house size, the marginal utility curve will fall with house size. Equilibrium house size is given where the two curves intercept.

$$Size^* = Size(\overset{\div}{P_B}, \overset{\div}{r}, \overset{+}{\alpha}, \overset{+}{INC})$$

The optimal house size falls with higher square meter price. A higher real interest rate reduces the ability to pay and increase the opportunity cost of housing, thereby lowering the optimal number of square meter house size. On the other hand, the optimal house size will increase with household income since a house is a normal good. Intuitively, the optimal number of square meters raises with higher preferences for house size.

Figure 3.1 Equilibrium house size



From an initial equilibrium of Size*$_0$, an increase in the interest rate or square meter price will shift the horizontal curve up to $U_{INC}\cdot P(P_B, r)_1$, lowering the optimal house size to Size*$_1$. A positive income shock (INC↑) gives a downwards shift for the marginal income curve, increasing optimal house size to Size*$_2$. Higher preferences for house size (α↑), for instance following an increase in the number of household members, provides an outwards shift in the marginal utility of house size curve from U_{size0} to U_{size1}. New equilibrium is found at Size*$_3$.

This section has shown how the dwelling's square meter price, household income, real interest rate and preferences may theoretical affect the optimal number of square meter demanded. Next section, a look into the data material used for the empirical analysis is taken, highlighting the variables discussed here and others of importance to the demand for house size.

3.2 Data material

In the following analysis, I use data from the Survey of Living Conditions collected by SSB. A representative sample of Norwegians over the age of 16 is randomly selected to participate in the survey. The data is collected mainly by telephone interviews, with some house visits. Since 1996, the survey has been carried out annually with a set of three rotating topics. The survey is a representative cross sectional survey, but the chosen individuals are asked to join several times making it a rotating panel of old and new participants. The data files are anonymized and made available for students and researchers by the Norwegian Social Science Data Service (NSD) (Thorsen & Revold, 2014).

Housing conditions are a specific topic in the surveys of 1997, 2001, 2004, 2007 and 2012. As a part of the EEA agreement, the survey has since 2011 been coordinated with the Eurostat's Statistics on Income and Living Conditions, EU-SILC, making it possible to compare living conditions across Europe. Included in the annual EU-SILC survey are some housing related questions previous only available every third year, making it possible to incorporate data from 2013 in the analysis as well. The rotating panel was reduced from eight to four years in 2012, inviting many new individuals to participate in the survey. As a result, the number of households increased from 2 350 in 2007 to 4 351 in 2012. Conversely, the respondent rate dropped from 66% to 56% in the same period. A lower participation rate is not a problem if individuals drop out on random. Non-random drop outs can be compensated for with so called frequency weights, taking into consideration what group or characteristics that drop out (Thorsen & Revold, 2014). However, since the concern of my analysis is on household level, this weighting is not regarded as of major importance and therefore ignored (Sandlie & Grødem, 2013). Unfortunately, I was unable to include the data set from 1997 in my analysis due to dissimilarities in some survey questions and answer categories. Moreover, because of an omitted income variable is the data set from 2011 excluded. On the whole, data from five years from the Survey of Living Conditions are used in this thesis, namely in 2001, 2004, 2007, 2012 and 2013.

All individuals in Norway above the age of 16 should have the same probability to be invited to participate in the survey. My analysis focuses on the household demand for house size, and therefore every household should have the same probability to be elected. Households with more members above the age of 16 are more likely to be invited into the survey than others. To correct for this selection bias, every observation is weighted with the inverse of the number of household members above the age of 16. Introducing this probability weight let us interpret the data on household instead of on individual level (Thorsen & Revold, 2014).

$$Household\ weight = \frac{1}{Number\ of\ household\ members\ over\ the\ age\ of\ 16}$$

The total number of households account to 16 032 for all years. The household weight has been applied to most of the analysis including descriptive statistics.

3.2.1 Square meter house size

In the following analysis, the self-reported number of square meters housing structure is the variable of interest. The interviewed are asked to quantify the total area within the outer walls, but only the used living area of basement and attic. Areas rented out to other households should not be included. Studying house size definitely has certain clear advantages. First, it is a clear objective measurement limiting subjective variation. Moreover, the unit is objectively the same and interesting to study over time. Also between different housing types, square meter is a comparable measurement as size describes an important aspect of every housing type. Last, the number of square meters variable is available for all years and reported for almost all households. All these effects speak in favor of using square meter house size as the variable of interest (Sandlie & Grødem, 2013).

On the other side, there are also some clear drawbacks only looking house size. As discussed before, is the demand for housing determined of many different characteristics that together make up the housing good. Only looking at house size, will ignore the quality of the dwelling structure and neighborhood characteristics not caught up by the explanatory variables. However, I will stress that this analysis looks at the demand for house size, not the general housing demand. As seen in the literature, most studies use a monetary measurement for housings services. My analysis only has a physical measurement, looking at house size as a simplified version of housing services.

The tenure choice between owner-occupancy and renting can also affect the housing demand. A renter makes the decision on house size only as a consumption good while the owner also take it into consideration as an investment. Higher prices can indicate higher capital gains if we regard housing as a pure financial asset (Dusansky, Koç, & Onur, 2010). This positive effect on demand from a price increase is not seen in the rental market, thus showing a difference in demand with respect to tenure choice. Hence, renters are shown to have higher price elasticity (in absolute value) than owners, and an analysis ignoring tenure choice could give seriously biased estimates (Gillingham & Hagemann, 1983). Also, with respect to Norwegian data, we see that renters on average are younger, live smaller and are more frequently situated in apartments in the larger cities compared to owners. In addition, income is also lower (Nesbakken, 2008). As a consequence, the following analysis look only at owner-occupiers, that is, households owning their own house, through a housing cooperative (*Borettslag)* or through a cooperative housing corporation *(boligaksjeselskap)*. A problem occurs if the ratio between renters and homeowners change over the time period, since then the sample composition no longer stays the same. In other words, a survey year with higher percentage of owners contains more households with an investment motivation, making over time comparison less accurate. This is not a major concern for my analysis as the homeownership ratio of all households remains fairly stable throughout the time period, ranging from the minimum of 81.7 % homeowners in 2004 to the maximum of 83.6 % in 2013.

3.2.2 Household income

Information on household income is taken from Statistic Norway's income register, thus not a survey question. Sometimes the number of household members asked in the interview does not match with the reported number of members of that specific household in the income register. To solve this problem, individual incomes from the household members that coincide with the survey are summarized into the household income. Further, interviews are mostly held in the first half of the year and many questions refer to the last year. As a result, the income from the year before is connected to the survey. Data from 2013 is matched with income information from 2012, 2012 from 2011 and so on (Thorsen & Revold, 2014).

The chosen income variable is the real after tax household income. In order to separate out the real effects, the income for all years has been adjusted for inflation using SSB's Consumer Price Index with 2012 as the reference year (SSB, 2015e). For some years, the household income is truncated of privacy concerns. As a consequence, the average values might be affected. Still, sample distribution should not have been altered to a substantial degree (Thorsen & Revold, 2014). Income is truncated in the data sets for 2007, 2012 and 2013, and these values will not be comparable with income in 2001 and 2004, where there is no upper limit. The lowest capped income is 900 000 NOK in 2013. In order for an identically defined and comparable income variable, all inflation adjusted incomes exceeding the lowest truncated value were set equal to it.

3.2.3 Square meter sale price

The price variable of interest is the price per square meter. In the interview, the respondent is asked to report the expected current sale price of their dwelling. A price per square meter variable is created by dividing the self-reported expected sale price on the number of square meter house size of the current dwelling. In order to keep the analysis in real terms, the variable has been adjusted for inflation in the same fashion as income. Finally, extreme values below 1 500 NOK and above 100 000 NOK per square meter house size have been dropped from the sample.

3.2.4 Household composition

Information on the size of each household is taken from SSB's population register, but the information is double-checked in the interview. The age of each household member is taken from the population register and included in the data set. In order to describe age structure within the household as well as their size, I have created three variables, namely the household number of children (0-17), adults (18-66) and elderly (67 and older).

3.2.5 Geographical location and size of populated area

In the data sets, the geographical location of the household is only found on regional level, where the regions are given by the Nomenclature Units for Territorial Statistics (NUTS) established by Eurostat. These regions are Oslo and Akershus, Hedmark and Oppland, Sør-Østlandet, Agder and Rogaland, Vestlandet, Trøndelag and Nord Norge[1]. Regional information of the household is taken from the population register (Thorsen & Revold, 2014). Each region is modeled as a dummy variable which equals 1 if the household live in that region and 0 otherwise.

In the data set there is a variable giving the size of the populated area where the household lives. The variable is given as one of seven groups, but I have separated them further into four categories and modeled each of them as dummy variable. The first group, sparsely populated area (*spredtbygd strøk)*, contains areas with less than 200 inhabitants. The next category, urban settlement (*tettbygd strøk)*, contains areas with inhabitants from 200 up to 20 000, where the distance between the dwellings cannot exceed 50 meters. Further, densely populated area (*tettsted)*, includes populated areas between 20 000 and 100 000 residents. Last category, includes populated areas and cites (*storby)* above 100 000 inhabitants. For the three first survey years the information is gathered as a part of the interview, whereas for 2012 and 2013 the information is taken directly from SSB's population register (Thorsen & Revold, 2014).

3.2.6 Educational attainment

Also, the respondent level of educational attainment is included, taken from the education register. Unfortunately, some survey years only give information about the interview object – not all members of the household. Thus, in order to make a comparable analysis, only education of the interview object is incorporated. The selected standard is the NUS2000 which divide attained education into nine groups. I have subsequently compressed educational attainment into six categories, each modeled as a dummy variable which equals 1 if the interview object has maximum attained that level of education and 0 otherwise. First group combine all respondents with maximum attained education of grade ten (*ungdomskole*). Individuals with more but highest attained education of upper secondary (*videregående*) are grouped together. Higher education is separated into two categories. The first group includes lower level up to a bachelor degree from universities and colleges. Whereas the other group includes individuals exceeding four years of higher education, a group which contain master degrees and vocational studies (*profesjonssutdanninger) (SSB, 2000).* Also a category

[1] Here Sør-Østlandet constitutes the counties of Østfold, Vestfold, Buskerud and Telemark. The category Vestlandet contains the counties of Hordaland, Sogn og Fjordane and Møre og Romsdal.

comprised of persons with doctorates is included. Last, a missing or unknown education group is included since it, on average for all survey years, constitutes 1.75% of the observations. This is done to avoid losing observations and to leave the reference group unaffected.

3.2.7 Type of housing

Arguably a variable of major importance to the house size demand is the type of housing. One of the survey questions asks about the housing type of the household's primary residence. A households that wants to live in a detached free standing (*frittstående*) residential buildings most likely differ in the size demanded compared to a household living in an apartment or multidwelling house, all else equal. There is also a category for linked houses (*rekkehus)* and one for semi-detached houses that includes housing structures shared with up to four households (*to-, tre- og firemannsbolig*). In addition, there is an auxiliary category of other housing types which include households living in commercial buildings, in boats and others (Thorsen & Revold, 2014). For each response alternative in the survey, I have created a dummy variable which equals 1 if the household have that housing type and 0 otherwise.

3.3 Descriptive statistics

3.3.1 Descriptive statistics of included variables

In order to get an overview over the included variables of my analysis and how they change over time, the average values for the five surveys are presented in Table 3.1. In addition, the table includes a column for the overall average values. With the probability weight applied we find estimates of the household mean, and not of the individual which is given by the sample mean. The standard deviation of the sample is here not of particular interest as it relates to the individual. Thus, instead standard errors of the estimate of houshold averages are reported (Sribney, 1997).

Table 3.1 Summary statistics of included variables

	2001		2004		2007		2012		2013		Full set			
Variable	**Mean**	**St. Error**	**Mean**	**St. Error**	**Mean**	**St. Error**	**Mean**	**St. Error**	**Mean**	**St. Error**	**Mean**	**St. Error**	**Min**	**Max**
Number of square meter housing	130.6	1.356	136.8	1.378	134.5	1.4558	137.9	1.0879	139.0	1.0766	135.8	1.2709	13	780
After tax household income	428493	4551	464385	4530	517011	5022	541409	3846	554180	3743	501096	4338	0	900000
Price per square meter	13752	171.12	15514	181.48	21127	262.85	23592	207.88	24267	210.24	19650	206.71	1511	100000
Truncated income dummy	0.0307	0.0033	0.0462	0.0039	0.0774	0.005	0.1142	0.0044	0.1217	0.0045	0.0780	0.0042	0	1
Household characteristics														
Number of children	0.672	0.0217	0.709	0.0210	0.658	0.0207	0.610	0.0149	0.585	0.0140	0.647	0.0185	0	7
Number of adults	1.496	0.0199	1.540	0.0185	1.512	0.0188	1.459	0.0141	1.435	0.0135	1.488	0.0170	0	6
Number of elderly	0.265	0.0124	0.249	0.0119	0.245	0.0120	0.288	0.0093	0.299	0.0092	0.269	0.0110	0	3
Region														
Oslo_Akershus	0.2395	0.0099	0.2384	0.0094	0.2477	0.0098	0.2623	0.0073	0.2589	0.0070	0.2493	0.0087	0	1
Hedemark_Oppland	0.0742	0.0059	0.0804	0.0059	0.0771	0.0060	0.0697	0.0042	0.0685	0.0041	0.0740	0.0052	0	1
Sør-Østlandet	0.1960	0.0090	0.1897	0.0086	0.1878	0.0087	0.1854	0.0064	0.1882	0.0063	0.1894	0.0078	0	1
Agder_Rogaland	0.1360	0.0077	0.1504	0.0077	0.1306	0.0073	0.1383	0.0056	0.1385	0.0054	0.1388	0.0067	0	1
Vestlandet	0.1549	0.0080	0.1607	0.0079	0.1638	0.0082	0.1603	0.0060	0.1655	0.0059	0.1611	0.0072	0	1
Trøndelag	0.0907	0.0064	0.0815	0.0059	0.0928	0.0064	0.0916	0.0047	0.0903	0.0045	0.0894	0.0056	0	1
Nord-Norge	0.1088	0.0071	0.0989	0.0064	0.1001	0.0067	0.0924	0.0048	0.0900	0.0046	0.0980	0.0059	0	1
Size of populated area														
Sparsely populated area (< 200)	0.1910	0.0087	0.1717	0.0082	0.1611	0.0080	0.1736	0.0061	0.1675	0.0059	0.1730	0.0074	0	1
Urban settlement (200 ≤ > 20 000)	0.3562	0.0108	0.3574	0.0103	0.3574	0.0106	0.3045	0.0075	0.3134	0.0074	0.3378	0.0093	0	1
Densely populated area (20 000 ≤ > 100 000)	0.2048	0.0092	0.2238	0.0091	0.2356	0.0095	0.1402	0.0058	0.1393	0.0056	0.1887	0.0078	0	1
Cities (> 100 000)	0.2471	0.0100	0.2469	0.0096	0.2445	0.0098	0.3818	0.0080	0.3792	0.0078	0.2999	0.0091	0	1
Level of attained education														
Primary and lower secondary	0.1554	0.0085	0.1539	0.0081	0.1961	0.0089	0.1658	0.0061	0.1627	0.0058	0.1668	0.0075	0	1
Upper secondary education	0.5361	0.0113	0.5325	0.0109	0.4359	0.0111	0.4331	0.0082	0.4256	0.0079	0.4726	0.0099	0	1
Higher education lower level (≤ 4 years)	0.2279	0.0095	0.2214	0.0090	0.2505	0.0097	0.2874	0.0074	0.2906	0.0072	0.2556	0.0086	0	1
Higher education higher level (> 4 years)	0.0551	0.0052	0.0636	0.0052	0.0825	0.0062	0.1000	0.0049	0.1070	0.0049	0.0816	0.0053	0	1
Doctorate	0.0086	0.0019	0.0039	0.0012	0.0097	0.0021	0.0091	0.0016	0.0099	0.0016	0.0083	0.0017	0	1
Missing education	0.0169	0.0026	0.0247	0.0030	0.0252	0.0031	0.0046	0.0010	0.0043	0.0011	0.0151	0.0022	0	1
Type of housing														
Apartment	0.1904	0.0097	0.1862	0.3894	0.2176	0.0100	0.2258	0.0074	0.2207	0.0072	0.2081	0.0847	0	1
Detached house	0.6152	0.0113	0.6071	0.4886	0.5687	0.0112	0.5713	0.0083	0.5701	0.0080	0.5865	0.1055	0	1
Linked house	0.1121	0.0072	0.1085	0.3111	0.1097	0.0069	0.1017	0.0049	0.1010	0.0047	0.1066	0.0670	0	1
Semi-detached house	0.0773	0.0062	0.0917	0.2886	0.0999	0.0070	0.0973	0.0051	0.1045	0.0051	0.0941	0.0624	0	1
Other housetypes	0.0050	0.0018	0.0066	0.0808	0.0042	0.0017	0.0040	0.0011	0.0037	0.0010	0.0047	0.0173	0	1
Number observations	**2282**		**2455**		**2350**		**4351**		**4594**		**16032**			

Income and price are reported in NOK. The truncated income dummy equals 1 when income is 900 000 and zero otherwise. Household characteristics are given by the number of children, adults and elderly in the household. Sparsely populated area has less than 200 inhabitants. Urban settlement has between 200 and 20 000. Densely populated area has between 20 000 and 100 000 whereas Cities has more than 100 000 inhabitants. Region, size of populated area, education and type of dwelling are all given as a dummy variable set equal to 1 if the household/interview object fits that category and 0 otherwise. Mean values of all dummy variables gives the share of all households fitting in the category. Minimum and maximum values are reported for the full set only. The household weight is applied.

Looking at the average values, reveals interesting information on how these variables change over the 13 year period. To measure the aggregated effect on the demand from households with higher income levels than the capped value, a binary variable is introduced that equals 1 if income is 900 000 NOK and 0 otherwise. Interestingly, the percentage of households exceeding this income level rises every year from 3.1 % in 2001 to 12.2 % in 2013. Moreover, mean income increases with 125 687 NOK in the same time period, showing a steady increase in net real household income. Both of these variables picture a stable growth in real income over time giving implications for the house size demand analyzed more later.

The estimated sale price per square meter also increases every year. Even adjusted for inflation, the average price increases from 13 752 NOK in 2001 to 24 267 NOK in 2013. This accounts for a remarkable 76.5 % increase over the 13 year period. SSB's own house price index states a 72.3 % house price increase in the same period once inflation is adjusted for, similar in magnitude to the growth in the price per square meter given by Table 3.1 (SSB, 2015b).

Looking at size and age structure of households, show changes over the time period. The number of kids follows a hump-shaped curve over time, with the highest value in 2004 when each household has 0.71 children on average, a number reduced to 0.59 children per household in the 2013 survey. The same tendency for the number of adults is found, falling from 1.54 in 2004 to 1.44 in 2013, indicating more single households or that children above the age of 18 leave their nests earlier. On the other side, a small increase in the average number of old household members from 0.25 in 2007 to 0.30 in 2013 is seen. Still, looking at household size on the whole, the tendency moves towards smaller households with average members falling from 2.50 members in 2004 to 2.32 in 2013.

The ongoing trend of urbanization and immigration attracts more people to the largest cities in Norway. According to SSB, there has been a population growth of 22.2 % in Oslo from 2005 to 2015 (SSB, 2015c). However, looking at my data set, the regional dummies reveal no clear pattern of regional movement. In fact, the percentage of respondents in each region does not change much over the 13 year period. The only exception is Oslo and Akershus where 24.0% of the interviewed lived in 2001, a number increasing to 26.2% in 2012.

Larger changes are found looking at the size of the populated area where households live. Starting with sparsely populated areas, there is a steady decline from 19.1% of all households in 2001 to 16.1% in 2007. Then the percentage increases for 2012 before it falls to 16.8 % in 2013. Further, looking only at the first three sample years, the percentage of respondents who say they live in a city or urban settlement stay constant. Most of the decline in sparsely populated areas is followed by a movement into densely populated areas which experience an increase from 20.5% to

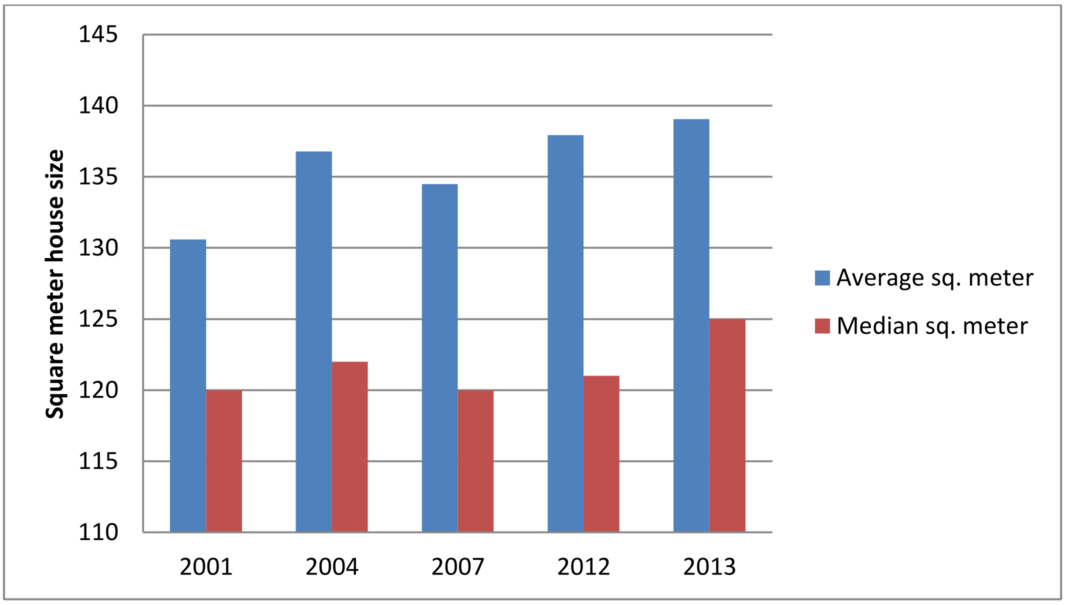
23.6%. Surprisingly, the allocation of inhabitants changes substantially for the two last sample years. Most striking, the percentage living in densely populated areas falls from 23.6% in 2007 to 14.0% in 2012 - a way larger fall than between any of the other survey years. On the other hand, cities experience an increase from 24.4% to 38.0% during the same time period - a category that remains somewhat stable between the other sample years. Same goes for the percentage living in urban settlements where there is a roughly five percentage point decrease between 2007 and 2012, but stay stable between the other survey years. The variable determining the size of the populated area where the household live is defined the same way for all survey years. Still, there is a crucial difference as from 2001 to 2007 the respondents are asked in the interview about the size of their populated area, whereas for the two last years this information is taken directly from the population register, thus no longer self-reported. Unfortunately, the reason why self-reported and register information differ is to me unknown. As a consequence of the deviation, comparison between the average values over time might not be valid.

Moving on to educational attainment, we also find some changes between the samples. The sample percentage with maximum of upper secondary education shrinks steadily from 53.6% in 2001 to 42.6% in 2013. Stable growth for higher education, both lower and higher level is found. Here, 22.8% of the sample is in the lower level higher education group in 2001, a number that increases to 29.1% in 2013. Looking at higher level, we see a remarkable doubling in percentage points from 5.5 % to 10.7%. Since the other dummies do not change substantially, people who previously only had a high school degree now seem to a greater extent take higher education.

Last, looking at housing types also reveal deviations over the time period. In 2001, 19.0 % of all households report living in an apartment whereas 22.1% answer the same in 2012. The increase seems largely to come from detached houses experiencing a drop from 61.5 % in 2001 to 56.9 % in 2007 and constant thereafter. To summarize, there seem to be some kind of movement in housing type from detached houses to apartments. An explanation could be immigration into the larger cities, where households of economical and spatial reasons decide to reside in multidwelling houses.

3.3.2 Descriptive statistics on house size

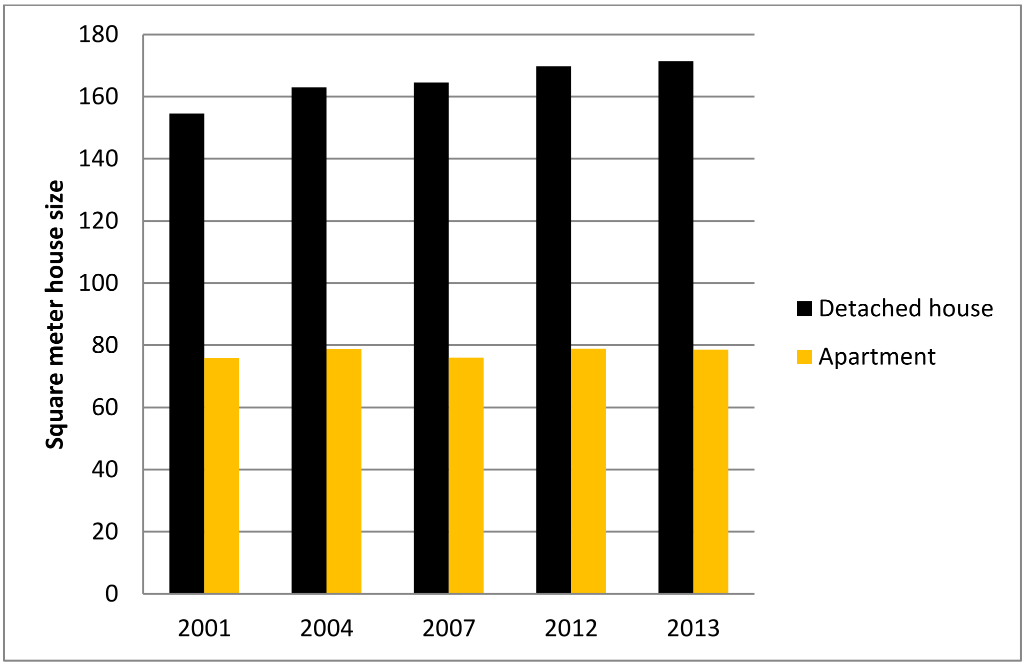
Figure 3.2 Average and median square meter house size



The mean number of square meters house size rises from 130.6m^2 in 2001 to 139.0m^2 in 2013 - an increase of 8.4m^2 in 12 years. The average house size increase for all years except for 2007, where mean value is 2.3m^2 lower than 2004. The same general pattern is seen looking at the median values[2], but there are some differences. By looking at the Figure 3.2, we see that the mean exceeds the median for all years. A reason for divergence could be that a few households have much larger houses than the rest, thus increasing the mean values. We can also see that the total increase over the time period only is 5m^2 studying median values. A possible explanation for the higher increase in mean compared to median values could be that the increase in house sizes, seen in the data, to some extent is driven by households with larger homes.

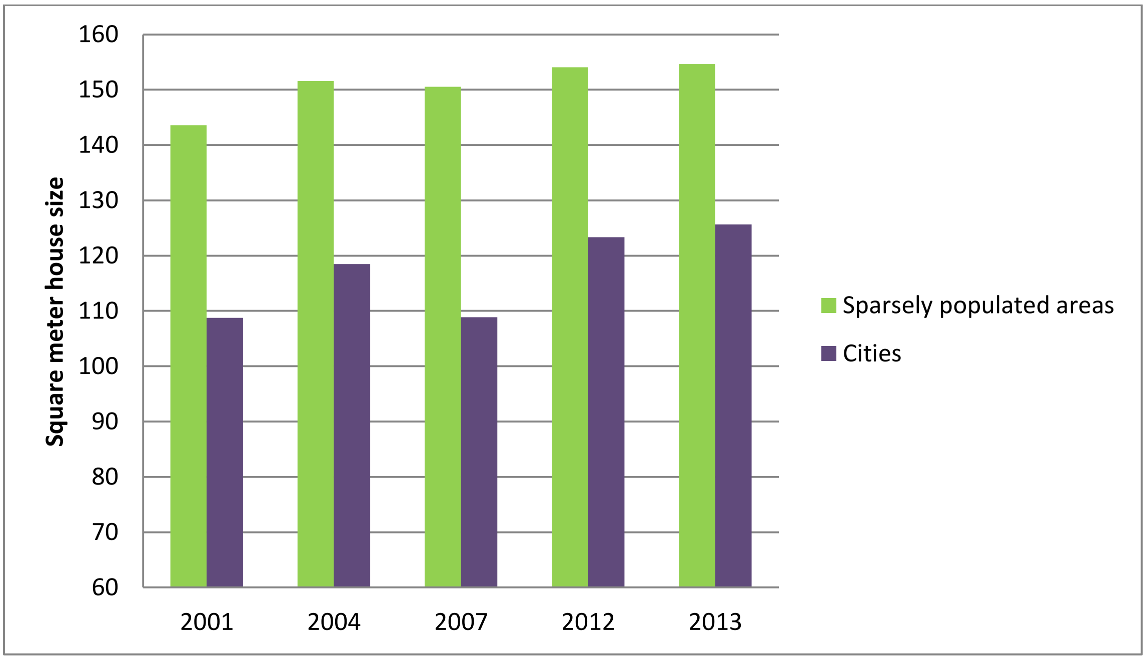
[2] By ranking houses after size, the median number is found in the middle of the distribution.

Figure 3.3 Average square meter house sizes for detached houses and apartments



In order to find more information on where the growth in house sizes have taken place, I have computed the mean values for each house type. Graphed average number of square meters for detached houses and apartments are shown in Figure 3.3. Interestingly, almost no increase in the size of multidwelling houses throughout the time period is found, but experience a small fall in 2007 consistent with the full sample. On the other hand, detached houses experience growth every sample year with a total increase of about 17m^2 from 2001 to 2013. Since detached houses constitute for roughly 60% of the full sample, a large portion of the total growth in house size is seen for households with detached houses.

Figure 3.4 Average square meter house size in sparsely populated areas and cities



Looking at average house size for households living in the largest and smallest populated areas, given by Figure 3.4, shows that households living in sparsely populated areas have larger dwellings. More surprising, cities experience a mean increase of $17m^2$ from 2001 to 2013, equal to the increase seen for detached houses. On the other hand, house sizes in sparsely populated areas increase with $10m^2$. On the whole, these findings indicate, although simplified, that much of the growth in house sizes happen in the cities, but in other housing types than apartments.

4 Empirical specification

4.1 Regression analysis

Regression analysis is a statistical method which can be used to study the relationship between two or more variables. A regression model predicts the value of the variable of interest based on the explanatory variables. A common estimation method named Ordinary Least Squares (OLS) seeks to minimize the sum of the squared error terms. That is, every vertical distance from an actual observation to the estimated regression line is squared, and then sum of these squares is minimized (Wooldridge, 2012). OLS is the chosen estimator for this analysis. One crucial assumption for OLS to be an unbiased estimator is that the error term has a zero conditional mean:

$$E(u|x_1, x_2, \ldots, x_k) = 0$$

where u is the error term and $x_1, x_2, \ldots, x_k$ are explanatory variables. The error term of the regression cannot be correlated with any of the explanatory variables of the regression. Potential reasons for violation include omitted variables and measurement error, which both are valid concerns of my analysis, thus elaborated on in more detail later in the thesis.

4.2 Main empirical specification

The main empirical analysis is carried out by regressing the number of square meter house size on income, price, household characteristics, geographical and structure related variables. A major goal of the analysis is to compare the results over time. Thus identical regressions given by equation (2.1) are carried out for each of the five samples.

$$ln(Size_i) = \beta_0 + \beta_1 ln(Income_i) + \beta_2 ln(Sq.meterPrice_i) + \alpha_1 INCdummy_i + \beta_3 Children_i + \beta_4 Adults_i + \beta_5 Elderly_i + \alpha_r + \alpha_p + \alpha_e + \alpha_h + u_i \qquad (2.1)$$

where
i= household.
r= 1, 2, ..., 7 denotes region of household i.
p= 1, 2, …., 4 gives the category of populated area where household i is situated.
e= 1, 2, …., 6 denotes the level of attained education of the interview object household i.
h= 1, 2, …., 5 denotes the housing type of household i’s primary residence.

$ln(Size_i)$ = natural log of square meter house size.

$ln(Sq.meterPrice_i)$ = natural log of the square meter house price.

$$INCdummy_i = \begin{cases} \dfrac{1\ if\ household\ income\ is\ equal\ 900\ 000}{0\ otherwise} \end{cases}$$

$Children_i, Adults_i\ and\ Elderly_i$ gives the household number of children, adults and elderly respectively.

$\alpha_r, \alpha_p, \alpha_p, \alpha_e, \alpha_h$ gives the fixed effect of region, size of populated area, education and house type respectively.

u_i is the error term for household i. u_i represent all other factors important to the household demand for house size other than my explanatory variables. The error term should be white noise, that is, uncorrelated with any of the included explanatory variables to avoid omitted variable bias.

After trying out different functional forms, the logarithmic was chosen for the regressand. Using logarithmic functional form for strictly positive variables, often satisfy the OLS assumptions better than using level variables. The interpretation of the effect of a change in an explanatory variable is described as a constant percentage change in square meters demanded. This allows for a different square meter impact from a change in income depending on the level of household income. Heteroskedastic or skewed conditional distributions are often the case for strictly positive variables. Using the logarithm of these variables may reduce or eliminate both of these problems. The logarithmic transformation also makes the estimates less sensitive to outliers. Moreover, R^2 values were on average 3-5 percentage points higher when taking the natural log of the explained variable, indicating a better specified model (Wooldridge, 2012).

The effect of income on house size is given by the coefficient β_1. As both the regressor and regressand have a logarithmic functional form we interpret β_1 as the elasticity of house size with respect to real net household income, later shortened as the income elasticity. When this income changes with 1 percent, house size demanded changes with β_1 percent. INCdummy is a binary variable set equal to 1 if household income is 900 000 NOK and 0 otherwise. The coefficient α_1 gives the relative change in house size demanded for households with this income level compared to those with income marginally below.

Also the price variable has a logarithmic functional form in equation (2.1). The interpretation of the coefficient β_2 is the elasticity of house size with respect to the square meter price, later shortened as the price elasticity. β_2 gives the percentage demand response on house size from a 1% change in the square meter price.

The effect of one more child in the household on house size is given by β_3. Since the variable appears in its level form, can β_3 be interpreted as the semi-elasticity of children on house size, which is the relative demand change from an incremental change in the number of children. In other words, once a child is added to the household, the demand for house size changes with approximately $100 \cdot \beta_3$ percent. Same interpretation follows for β_4 and β_5, which give the relative demand change from a member increase in adults and elderly respectively.

In the present analysis, regional dummies are applied with Oslo and Akershus as the reference category. For the other regions, the binary variable equals 1 for households situated in that region and 0 otherwise. In the model (2.1), regional fixed effects is given by α_r. The interpretation of the parameter is the relative change in house size demanded when a household is situated in a specific region compared to Oslo and Akershus. Again, multiplying the coefficient of the regional dummy with 100 provides us with the percentage change in house size demanded of living in that region compared to Oslo and Akershus, all else equal.

Each category of populated area is also modeled as dummy variables with α_p as the fixed effects. The parameter values in front of the dummy variables are interpreted in the same way as the regional dummies, comparing the demand change to the reference group which here is living in a sparsely populated area.

The relative importance of education is given by $\alpha_e$in equation (2.1). Primary and lower secondary education is the base group, and the coefficients in front of the dummies for the other education groups give the demand change for an individual in that group compared to the reference group.

Fixed effects of housing type is given by α_h. Coefficients of the different house type dummies have the same interpretation as the other dummy variables, giving the relative demand change of the household living in that housing type compared to apartments which is the base group.

5 Empirical analysis

5.1.1 Income elasticity

Reported income elasticity in my analysis is defined as the percentage increase in demanded house size from a percentage increase in after tax household income. If income rises with 10% in 2001, demand will increase with 1.36% as shown in in Table 5.1. Three years later, demand increases with 1.79% from the same income increase. The estimated income elasticity reaches its highest value of 0.279 in 2007 noticeable higher than all other years. Note that estimated coefficients differ considerably for some variables in the sample from 2007. The last two sample years have identical income elasticity of 0.207. Overlooking 2007, the estimates show a steady increase in the income elasticity increase over time. Over the 13 year period, households seem to invest larger parts of their wage increases into housing. The truncated income dummy coefficient gives the difference in demand for households with income of 900 000 NOK to those marginally below (remember all higher incomes above were set equal to 900 000). The estimated coefficient reaches its maximum value in 2004, where demand is 18.2% higher for the top income group. However, the data indicates no changes over time as demand increases with 14 % both in 2001 and 2013. Even when more households enter the top income group, as seen in Table 3.1, their collective effect on demand for house size remains unaltered.

5.1.2 Price elasticity

The price elasticity of demand is on the other hand more stable. A 10% square meter price increase lower the demand for house size with about 2.0 % both in 2001 and 2013, showing no constant changes over time. Although the average square meter price sharply increases over the time period, the negative demand response from a square meter price increase does not deviate in the estimates of 2001 and 2013.

5.1.3 Household composition

The ceteris paribus effect from number of children on house size shows a clearly significant and constant effect when 2007 is excluded. Overall, the demand increases with approximately 5.5% with each additional child in the household. Looking at the estimated coefficients for the number of adults, inform us of a steady decline in magnitude ranging from 7.4% in 2001 to 4.0% in 2013, all significant at 5% significance level again excluding 2007. Continuing with the effect of the number of elderly demonstrates an identical pattern. Demand for house size in 2001 increases with 10.0% with each old member in the household, a number which falls to 7.2% in 2013. In summary, the effect on house size from an additional adult and old member in the household is reduced over the time period, but remains constant for an additional child. Again, the results from 2007 are excluded due to low parameter values and less significance.

Table 5.1 Demand for house size, OLS-regression results

	2001			2004			2007			2012			2013		
Variable	Coeff.	Std. Err.	P-value	Coeff.	Std. Err.	P-value	Coeff.	Std. Err.	P-value	Coeff.	Std. Err.	P-value	Coeff.	Std. Err.	P-value
After tax household income	0.136	0.0204	0	0.177	0.0240	0	0.279	0.0289	0	0.207	0.0192	0	0.207	0.0192	0
Price per square meter	-0.198	0.0193	0	-0.232	0.0216	0	-0.2351	0.0205	0	-0.207	0.0197	0	-0.202	0.0180	0
Truncated income dummy	0.144	0.0487	0.003	0.182	0.0274	0	0.181	0.0230	0	0.127	0.0167	0	0.142	0.0153	0
Household characteristics															
Number of children	0.057	0.0086	0	0.052	0.0070	0	0.032	0.0071	0	0.052	0.0053	0	0.058	0.0055	0
Number of adults	0.074	0.0313	0.018	0.055	0.0154	0	0.004	0.0166	0.801	0.044	0.0132	0.001	0.040	0.0125	0.001
Number of elderly	0.100	0.0244	0	0.085	0.0188	0	0.040	0.0200	0.046	0.073	0.0144	0	0.072	0.0147	0
Region															
Oslo_Akershus	(omitted)														
Hedemark_Oppland	-0.180	0.0340	0	-0.148	0.0308	0	-0.170	0.0315	0	-0.105	0.0282	0	-0.102	0.0260	0
Sør-Østlandet	-0.147	0.0272	0	-0.154	0.0246	0	-0.1207	0.0247	0	-0.063	0.0206	0.002	-0.064	0.0202	0.001
Agder_Rogaland	-0.115	0.0275	0	-0.126	0.0249	0	-0.0959	0.0249	0	0.024	0.0198	0.225	0.013	0.0186	0.480
Vestlandet	-0.085	0.0273	0.002	-0.044	0.0244	0.071	-0.0774	0.0246	0.002	0.001	0.0185	0.936	-0.019	0.0181	0.287
Trøndelag	-0.160	0.0296	0	-0.155	0.0275	0	-0.1641	0.0283	0	-0.040	0.0230	0.085	-0.046	0.0213	0.029
Nord-Norge	-0.208	0.0318	0	-0.235	0.0331	0	-0.220	0.0312	0	-0.103	0.0256	0	-0.106	0.0242	0
Size of populated area															
Sparsely populated area (< 200)	(omitted)														
Urban settlement (200 ≤ > 20 000)	0.051	0.0216	0.018	0.035	0.0224	0.121	0.069	0.0232	0.003	0.069	0.0172	0	0.059	0.0170	0.001
Densely populated area (20 000 ≤ > 100 000)	0.138	0.0260	0	0.139	0.0262	0	0.134	0.0271	0	0.108	0.0222	0	0.085	0.0212	0
Cities (> 100 000)	0.105	0.0311	0.001	0.134	0.0288	0	0.109	0.0311	0	0.132	0.0229	0	0.118	0.0224	0
Level of attained education															
Primary and lower secondary	(omitted)														
Upper secondary education	0.073	0.0220	0.001	0.051	0.0205	0.012	0.021	0.0202	0.289	0.055	0.0162	0.001	0.041	0.0161	0.011
Higher education lower level (≤ 4 years)	0.115	0.0279	0	0.098	0.0250	0	0.028	0.0225	0.217	0.091	0.0180	0	0.061	0.0173	0
Higher education higher level (> 4 years)	0.190	0.0454	0	0.129	0.0352	0	0.048	0.0324	0.141	0.089	0.0227	0	0.058	0.0230	0.012
Doctorate	0.133	0.0679	0.050	0.118	0.0679	0.084	-0.0016	0.0877	0.985	0.145	0.0493	0.003	0.105	0.0534	0.050
Missing education	0.171	0.0830	0.039	0.068	0.0390	0.080	-0.029	0.0385	0.451	-0.071	0.0854	0.404	-0.028	0.0840	0.742
Type of housing															
Apartment	(omitted)														
Detached house	0.581	0.0276	0	0.587	0.0239	0	0.587	0.0258	0	0.586	0.0199	0	0.577	0.0174	0
Linked house	0.276	0.0271	0	0.295	0.0272	0	0.327	0.0272	0	0.266	0.0193	0	0.263	0.0193	0
Semi-detached house	0.283	0.0358	0	0.241	0.0304	0	0.261	0.0302	0	0.242	0.0218	0	0.238	0.0207	0
Other housetypes	0.150	0.1216	0.217	0.458	0.0939	0	0.109	0.1320	0.408	0.052	0.2581	0.840	0.135	0.2974	0.651
Constant	4.343	0.8081	0	4.206	0.3636	0	3.013	0.4188	0	3.597	0.3603	0	3.565	0.3073	0
Number of observations	2282			2455			2348			4350			4592		
R-Squared	0.530			0.553			0.573			0.566			0.571		

Dependent variable is square meter house size. P-values < 0.001 are written as 0. The household weight is used in the estimation.

5.1.4 Type of housing

Keeping all other variables equal, a strong effect on demand is shown if the dwelling is detached. This should come as no surprise since detached houses on average of all years are about $90m^2$ larger than apartments. Table 5.1 shows that a household living detached demand 58.4% more square meters as an average effect compared to living in a multidwelling house. The separate regressions exhibit small deviations from each other, showing no sign of a time trend. Further, households with linked houses demand more size compared to households living in apartments, ranging from 26.3% in 2013 to 32.7% in 2007, following a hump-shaped curve over time. Semi-detached houses demand roughly the same percentage increase as linked houses, possibly slightly lower. All reported coefficients for house types come out clearly significant different from zero with p-values less than 0.001 for all sample years, emphasizing the importance of type of housing structure on the square meter demand. Last category, showing other house types, comes out insignificant for all years except 2004. Here a surprisingly high parameter value is found. However, only 15 individuals report living in other house types, casting doubt regarding variable importance.

5.1.5 Region

Interestingly, most of the geographical location dummies come out significant different from zero compared to Oslo and Akershus. This holds in particular for the three first sample years where demand is, with one exception, significantly lower than zero (at 5%) for all regions compared with Oslo and Akershus. The output tells us that regions including the largest metropolitan areas of Oslo, Bergen, Stavanger and Trondheim have all higher demand. While on the other side, sparsely populated regions like Nord-Norge and Hedmark and Oppland have lower demand than Oslo and Akershus significant at 1%. The effect is largest in 2004 where a household, all else equal, demand 23.5% less square meters situated in Nord-Norge instead of Oslo and Akershus.

Intuitively, we can think of regions with less people and more free space also have larger dwellings, something I also found for all regions comparing with Oslo and Akershus. Although, when controlling for dwelling type and square meter house price the coefficients turn negative. In 2013, 33% of the households in Oslo and Akershus live in a detached house whereas 43% live in an apartment. The figures deviate substantially comparing with the other regions combined, where 65% of the households live in a detached house compared to 15% in apartments. Another reason comes from the large discrepancies in price level between regions. For instance, Oslo and Akershus the mean price square meter price in 2013 is almost twice that of Nord-Norge. Hence, households demand for a given house type and square meter price more square meters in the capital region than elsewhere in the country. Further, measurement error in house size can result in too strong price elasticity in the square meter price compared to the true value. Too high price elasticity in absolute value will push down the demand more in areas with higher price level. I will come back to the nature of this estimation bias in Chapter 7.2.3.

Other reasons for higher demand include better job opportunities and other positive urban externalities. Higher price expectations in the capital region, making buying a dwelling for a given price a better investment is another possible explanation. Also regional divergence in prices of other consumption goods can potentially increase housing demand in the regions where the prices are lower. Few municipalities in Oslo and Akershus have enforced a voluntary property tax. Thus, the demand, all else equal, is lower in other regions where the tax is more common. I will come back to this issue later in the study. Last potential explanation come from how income for households above the capped value of 900 000 NOK are distributed. In a case where households in Oslo and Akershus have income levels far above the threshold, whereas other regions mainly lie just above would implicate a negative demand effect of other reasons versus Oslo and Akershus.

The regional effect drops substantially for the last two survey years, and in 2012 the demand is only 10.3% lower in Nord-Norge. In 2012 and 2013, Vestlandet, Agder and Rogaland and Trøndelag are all not different from zero at 5% significance level. The decline is largest for Agder and Rogaland, where the demand in 2001 is 11.5% lower than the capital region, compared to a positive, though highly insignificant, demand response for the two last years.

5.1.6 Size of the populated area

Particular households living in densely populated areas and cities have significantly higher demand compared with households in sparsely populated areas. From 2001 to 2007, the effect is slightly higher for densely populated areas with an average of 13.7% higher demand compared with 11.6% for cities. In the description of the data material, I mentioned that a change in the information collection from self-reported to register gave major changes in the distribution between the groups for the last two survey years. In particular there was a strong movement from densely populated areas to cities. The slope parameter for cities increases from 10.9% in 2007 to 13.2% in 2012 whereas the effect of densely populated areas falls for both of the last two years, ending at 8.5% demand increase in 2013. For the last two survey years ,the effect of the two groups have changed making city dwellers demanding larger homes than those living in densely populated areas, everything else equal.

Again, caution should be made due to the possible lack of comparability over time because of the change from self-reported to register information. Also migration within the time period could affect demand if a populated area changed category from 2001 to 2013.

5.1.7 Educational attainment

Looking at the ceteris paribus effect of education, reveals a greater square meter demand for respondents with higher education with the exception of 2007, where all estimates are insignificant even at 10% significance level. The effect of education is highest in 2001, where an individual with higher level higher education demand 19.0% more house size than respondents with below upper secondary education. Interestingly, the effect drops sharply over the time period, excluding 2007, ending with a 5.8% demand increase compared to the reference group in 2013. Lower level higher education follows the same trend, although in lower magnitudes, with a demand increase of 11.5% in 2001 falling to 6.1% in 2013. A potential reason to my findings is the increase in higher education we have seen in the descriptive statistics. As a consequence, the return to education falls as more people get educated lowering the marginal benefit of education on housing demand. A somewhat surprising result, however, is that the coefficients of long higher education are slightly lower than short higher education in 2012 and 2013. For households where the interview object has a PHD or doctorate, a fairly stable effect over time is found with an average of 12.5% higher demand than the base group. Caution should be advised, as the variable is borderline significant at 5% and contains only a few individuals.

5.2 Predicted demand over time

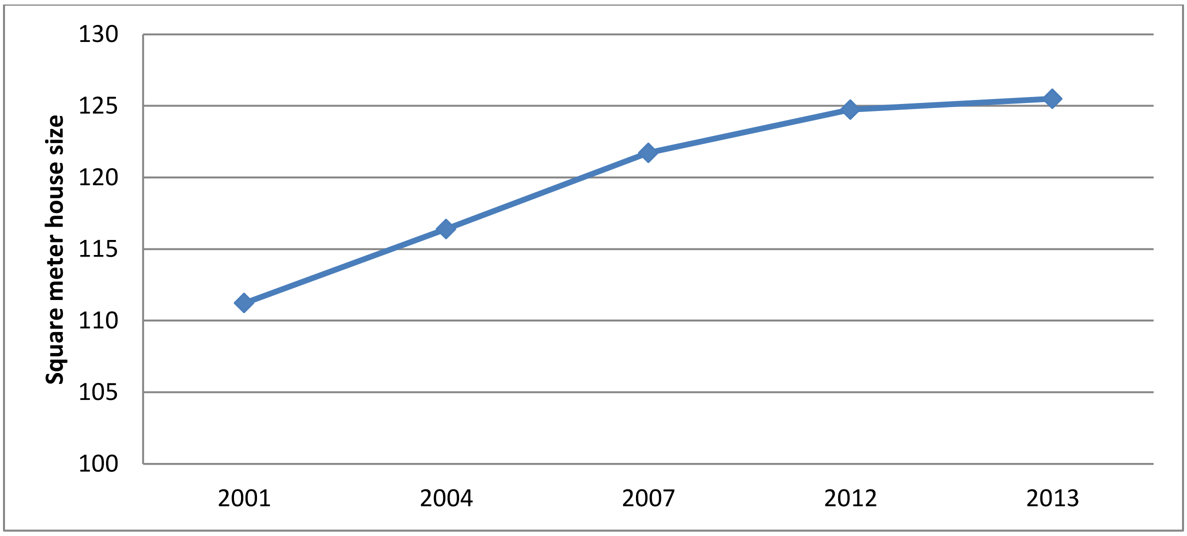
So far I have estimated the effect of a range of variables on the square meter demand for five samples. Next concern of this thesis is to analyze the change in house size demand over the time period given all the independent variables treated so far. Every household and house differs in their characteristics, which naturally leads to different demand for all households. To measure changes in the demand over time, the same values of the explanatory variables are inserted each survey year. The reason for looking at identical households, income, price, housing structure *et cetera*, is to separate out the effect of my explanatory variables from the other factors determining the demand for house size.

A normal procedure is to insert mean values of the independent variables. This could be averages for a given year, but in this analysis I use the overall average over the time period 2001-2013 given by Table 3.1. There is of course not possible to have 0.7 kids or to live 10% in Nord-Norge but averages are as a statistical identity valid. The average values are multiplied with the corresponding estimated coefficient for the specific year. Next, all these values and the estimated constant term are summarized into what the predicted demand for house size for that year. Specifically, equation (2.2) is computed for every survey year. The predicted demand for house size for the five surveys is plotted together and presented in level form in Figure 5.1. This curve will be called the predicted demand curve for the remainder of this paper.

$$\ln(Size_i) = \widehat{\beta_0} + \widehat{\beta_1} ln(\overline{Income_i}) + \widehat{\beta_2} ln\left(\overline{Sq.meterPrice_i}\right) + \widehat{\alpha_1}\overline{INCdummy_i} + \widehat{\beta_3}\overline{Children_i} + \widehat{\beta_4}\overline{Adults_i} + \widehat{\beta_5}\overline{Elderly_i} + \widehat{\alpha_r} + \widehat{\alpha_p} + \widehat{\alpha_e} + \widehat{\alpha_h} \quad (2.2)$$

We have already seen that average number square meters increase with 8.4m^2 between 2001 and 2013. In the case of a downward sloping predicted demand curve, would changes to the right side variables explain more than all of the increase in square meters, leaving other variables and preferences demanding less size combined. A flat curve indicates that the whole size increase can be explained by changes to the independent variables of my analysis. Finally, a positive slope tells us that given all the included variables, demand increases showing that other factors than those included in the regression are of importance to why households buy larger dwellings.

Figure 5.1 Predicted demand given all explanatory variables



At first look, we see that the predicted demand for house size rises every sample year, ranging from 111.9m^2 in 2001 to 126.3m^2 in 2013, leaving a total increase of 14.4m^2 or an average of 1.21m^2 per year. The magnitude is largest between 2001 and 2004 with a yearly increase of 1.82m^2. Between 2004 and 2007, demand increases with 1.77m^2 per year but falls down to a 0.62 m^2 increase between 2007 and 2012. Between the two last surveys with only a year apart we see a 0.73 m^2 increase. In brief, there is a quite large but diminishing increase over the period with a small upturn last survey year. Still, every year the actual increase in mean house size is larger than what can be explained by the explanatory variables of my analysis, indicating that there are other factors explaining the growth in the number of square meters house size. Potential other factors are discussed in the next section.

5.3 Macroeconomic variables and preferences

There are many elements affecting households in their decision in which home to buy. Some of these are macro variables which are equal for all households, thus not possible to include in a cross sectional analysis. Following, I present the in my opinion most important factors of demand with particular concern of any changes over the 13 year period that potentially can explain movement in the predicted demand curve in Figure 5.1.

5.3.1 The role of interest rates

Arguably a factor of major importance to the housing demand is the interest rate as it determines the actual cost of buying a house through the real mortgage costs. The real interest rate matters even if the dwelling is financed by savings as it represents the opportunity cost of investing in housing. In other words, a lower real interest rate, all else equal, raises the ability to pay for a house, thus increasing demand. The positive demand effect for households from a reduction in the real rate increases with the level of debt. Renters and other individuals not owning their dwelling are net losers as mortgage rates go down and housing prices increase (NOU, 2002).

First, I try to isolate the part of the interest rate which matter for households' housing demand. I have used the private average loan rate of private banks subtracted for debt deduction and inflation[3]. By studying the time period of current concern, we see that the average real interest loan rate of the private banks fell from 3.19% in 2001 to 2.88 % in 2004 (SSB, 2015d). In addition, inflation was at 3.0% in 2001 compared to only 0.4% in 2004, widening the interest rate differential only looking in nominal terms. The interest rate reduction between the two first survey years is arguably an important explanation for the sharp increase in predicted demand seen in Figure 5.1.

Even when the rate increased to 5.18 % in 2007, there was a strong demand increase. Then again, the rate falls to 3.18% in 2012 in a period with relatively low demand growth. For 2013, the nominal rate remains the same, but lower price growth leaves the real interest rate at 1.88 %.

[3] In order to isolate the real interest rate subtracted for debt deduction, the real interest rate is computed in this way: $(\text{Nominal rate})_i * (1-0.28) - (\text{inflation})_i$, where i = year.

5.3.2 Unemployment

Another macro variable of potential importance for the demand for house size is the unemployment rate. A person without a job is, most likely, not in a position to buy a larger home. Intuitively, high rates of unemployment reduce the aggregated housing demand.

The rate of unemployment is here given as the number of unemployed persons as percentage of the labor force collected from SSB's labor force survey (AKU). Looking at the rates, we see an increase in unemployment from 3.3% in 2001 to 4.2% in 2004[4]. Separately, higher unemployment should have reduced housing demand contrary to the effect observed for the predicted demand in Figure 5.1. However, in 2007, the rate was down to 2.6% which is a reduction of almost 50 percentage points in 3 years. Thus, the considerable fall in unemployment is a plausible explanatory factor for the predicted demand increase between 2004 and 2007. Last two survey years, the unemployment rate rose to 3.1% and 3.5% in 2012 and 2013 respectively. Higher rates seem sensible with the lower demand growth seen in the last part of the time period.

5.3.3 Expectations about the future

High unemployment rate can negatively affect demand by reducing expectations regarding the future. An explanation for the low demand increase from 2007-2012 is the global financial crisis that started late 2007. The Norwegian market was not as severe hit as other markets, and an expansive financial policy was practiced by the government to avoid a serious recession. The housing prices fell the first half of 2008, stabilized in 2009, but have increased since then (SSB, 2015b). As seen is the predicted demand growth at its lowest between 2007 and 2012. Fundamental factors and the fall in expectations are valid reasons for lower housing demand this period.

Regarding owner-occupied housing only as an investment, highlights the importance of future price expectations. Higher price expectations increase the value today making housing a better investment, boosting the willingness to pay for a given price. Periods with high price growth may positively affect households' price expectations, thinking that the prices will continue to go up. Hence, observed tendencies in the market often overcome fundamental factors like building cost in the price determination (Case, Quigley, & Shiller, 2003). Dusansky et al. (2010) found in some "hot" markets an upwards sloping demand curve in its own price indicating higher demand when prices go up.

Moreover, the interest rate and expectations often go in opposite directions. Economies with high aggregate output should have high interest rates to dampen an over-heated economy. In Norway,

[4] Figures are seasonally adjusted and given as first half year averages to correspond with the time of survey (SSB, 2015a).

however, interest rates are currently at a record low much because of external factors while the price expectations are high. The Norwegian central bank decided to maintain the interest rate at 1.25 % at its meeting 19.3.2015, surprising a market that expected a fall. The reason the rate was kept constant was mainly the fear of the galloping house prices causing financial instability (Singsaas, 2015).

5.3.4 Credit availability

From the late 1980s to the early 1990s, Norway suffered under a serious banking crisis. The crisis was mainly caused by losses in commercial and saving banks following a price collapse in the housing market. The crisis peaked in 1991, a year where the banks faced total losses of 20 billion NOK and the unemployment rate was almost as high at 6 %. As a consequence, some banks had to close down while others were forced to merge. Moreover, the government had to guarantee for private banks' liabilities and, in some cases, banks were put under administration (Krogh, 2010).

In order to limit market risk of default, the Norwegian authorities raised in 1998 the risk weight from 50% to 100% for mortgages with a loan-to-value ratio from 60% to 80%. As a result, the demand for credit was somewhat reduced. This policy was reversed in 2001 boosting the credit availability of households. Following the Basel II accords[5] implemented 2007, were risk weights reduced further making house loans cheaper and more available (Krogh, 2010). The reduction of risk weights has given an impact on the households' level of debt, seen in 2007, a year when 41 percent of existing homeowners increased their mortgage (Vatne, 2009).

In the mid-2000s, "flexible mortgages" (*rammelån)* got more popular providing flexible credit for homeowners to spend home equity more freely at a given mortgage constraint. This has contributed in making the housing asset more liquid (Krogh, 2010). The households' ability to pay has increased in the present time period through the increased credit availability. As the changes mostly happened over time there is not easy to pinpoint the exact increase in the predicted demand curve, however, more lenient lending policies have arguably been a substantial drive of the growing demand.

On the other hand, The Financial Supervisory Authority (*Finanstilsynet*) implemented a normal requirement of 10 percent equity on mortgages secured on the dwelling. This requirement rose to 15 percent in 2012 after concerns about the financial stability. There have been discussions about how strict these requirements should be interpreted and The Norwegian Minister of Finance, Siv Jensen, called for a more flexible interpretation of the 15 percent equity requirement with particular concern for the first time buyer's problem entering the market. Most recent, Siv Jensen wrote a letter dated

[5]A set of international banking regulations put forth by the Basel Committee on Bank Supervision.

6.March 2015 asking The Financial Supervisory Authority for measures to dampen the fast growing house prices and levels of household debt. Their response called for stricter regulation and follow-up action, tightening the banks' flexibility in the lending practice including the interpretation of the 15 percentage rule (Baltzersen, 2015).

5.3.5 Tax issues

Another component that may increase housing demand is the tax advantage by investing in housing compared to other capital assets. In Norway, interest payments on private debt are subject to tax deductions (*rentefradrag)*. For all the years I have data, taxes were deducted with 28% of the debt payments. However, the rate was reduced to 27% as of 2014 (Finansdepartementet, 2013).

Earlier, a tax on income from owner-occupied housing was implemented in Norway. Income from owner-occupied housing represents the amount the owner would have to pay if he rented his own house. The net housing income was interest rate payments and other housing related expenditure subtracted. For simplicity, net housing income was set as 2.5 % of the assessed house value (*ligningsverdi*) which is set maximum at 30% of market value. This value subtracted from a tax free allowance (*bunnfradrag*) was subject for taxation. The political foundation of the minority government lead by Kjell Magne Bondevik, also called the Sem-declaration, promised a removal of the income tax from owner-occupied housing. The law was implemented from the beginning of 2005. Also rental income of the primary residence became tax free under the condition that less than half of the structure was rented out. Justification was made on the grounds that the tax hit different types of houses arbitrarily, especially favoring older dwellings where price growth not were taken into consideration (Finansdepartementet, 2004).

In order for a better redistribution through the tax system, a wealth tax is implemented in Norway. The tax is levied on all personal assets including owner-occupied housing. Lower taxes are imposed on housing since the value of the dwelling is given by the assessed house value which on average counts for 25% of market price for the primary residence. Hence, owner-occupied housing is favored as a capital asset since other assets like stocks and bank deposits are valued at market price (Finansdepartementet, 2009). In 2001 and 2004, the state levied tax on wealth between 120 000 NOK and 540 000 NOK at a rate of 0.2%. Higher wealth was taxed at 0.7%. There is also a municipal wealth tax, which was set at 0.7% of wealth above 120 000 NOK for the two first survey years. Although the same rates were in force also for 2007, the tax free allowance slightly increased (Skatteeaten, 2015). In 2010, the tax was changed substantially raising the tax free allowance to 700 000 NOK for both the state and local tax, and changing to an area based system for assigning assessed house value. Total tax relief was estimated to be 760 million NOK and only 17% of the tax payers would pay wealth tax compared to 30% in 2005 (Finansdepartementet, 2009). Also for the two last survey years, the tax free allowance increased further (Skatteeaten, 2015).

A task to determine the total effect on demand from the modification in the wealth tax is complicated by the presence of both an income and a substitution effect. Households which either paid no tax before or always paid the tax are not affected by the changes as the tax rate remained stable at 0.7%. However as the tax free allowance increase, more households avoid the tax, thus lifting their housing demand because of higher disposable income. On the other hand, other financial assets will no longer be disfavored versus owner-occupied housing, shifting demand towards stocks and bank deposits, at the expense of housing, lowering the demand. A task to analyze which of the two effects that dominate is beyond the scope of this text.

There exists a local voluntary property tax decided on by each municipality. Yearly tax rate is fixed between 0.2% and 0.7% of market value decided on by each municipality. In 2001, 53% of all Norwegian municipalities enforced a property tax on owner-occupiers. The percentage increased to 63% as of 2007, whereas 77% or 329 municipalities had a property tax in 2013 (Refling, 2013). There seem to be a constant trend toward more property tax, which all else equal reduce the housing demand. An implication of the municipal property tax is that it might affect the regional dummies of my analysis. In a region where many municipalities have enforced property tax, might negatively affect the estimated coefficients of that region. Oslo and all of the municipalities in Akershus with the exception of three have no property tax, which potentially bias the estimates of all regional dummies downwards compared to a situation where property tax is controlled for (Refling, 2013).

5.3.6 Preferences and neighborhood effects

As already mentioned does “taste” account for an important component of the housing demand. According to Torstein Veblen, large homes are a way of showing off status, thus, all else equal, increasing preferences of housing higher than the actual need.

Housing preferences, although undoubtedly important, are hard to measure and highly correlated with other variables like income. A task to quantify the effect of preferences for house size and neighborhood effects in Norway are beyond the scope of this text. Still, a change in “taste” for more space and influence from neighbors may account for some of the increase in the predicted demand curve that is not explained by macroeconomic factors.

5.3.7 Macroeconomic variables and the housing demand

In order to get an overview over how the macroeconomic variables may have affected the demand throughout the time period, I have constructed a table displaying main findings. Table 5.2 shows how changes to the discussed factors may affect housing demand. Particularly, the interest rate might explain why predicted demand jumps from 2001 to 2004. Also higher expectations and more available credit are believed to boost demand in this period. Demand grew from 2004 to 2007 even with higher interest rates which could have been caused by an almost 50% cut in the unemployment rate as well as higher expectations. Low demand growth from 2007 to 2012 is thought mainly to be caused by lower expectations caused by the global financial crisis as the other variables indicate higher demand. Lower real interest rate from 2012 to 2013 can be one explanation why demand has increased in the last period.

Table 5.2 Effects of macroeconomic variables on the housing demand

Variable	2001-2004	2004-2007	2007-2012	2012-2013	Comments
Real interest rate	+ (2.9% in 2004)	÷ (5.2% in 2007)	+ (3.2% in 2012)	+ (1.9 % in 2013)	Positive sign when the rate falls.
Unemployment	÷ (4.2% in 2004)	+ (2.6% in 2007)	÷ (3.1% in 2012)	÷ (3.5 % in 2013)	Positive sign when rate falls.
Expectations	+	+	÷	+	Assuming positive expectation in periods with price increase.
Credit availability	+ (reversing higher risk weights from 1998)		+ (further reduction of risk weights following Basel II)	(÷) (Implementation of 15% equity on mortgage)	There has been a tendency towards more lenient lending including implementation of flexible mortgages.
Tax system		+ (elimination of tax on housing income)			An increasing tax base for property tax reduces isolated demand the whole period.

Positive sign stands for higher demand.

5.4 Shift analysis

As shown in the descriptive statistics part, several of the explanatory variables like income, prices and household characteristics change over time. Household incomes have increased substantially and house prices have gone up, while household size is reduced. This part analyses if changes to any of these or to other included explanatory variables can explain the observed increase in square meter house size. I start with determining the importance of changes to the right side variables combined by comparing the predicted demand curve with the plotted average square meters for each survey year. Next, the relative importance of changes to income, price and household composition on demand is studied separately.

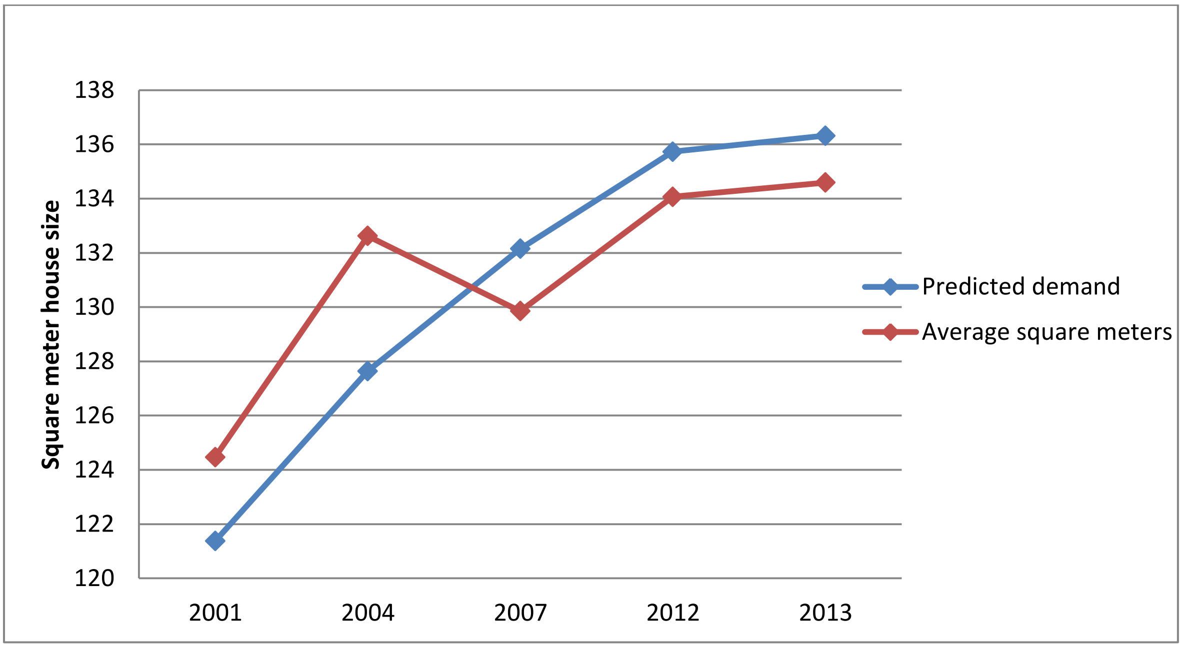
5.4.1 Changes to the explanatory variables combined

First, separate but identical regressions are preformed every sample year with all explanatory variables included. Specifically, (2.2) is again computed, identically as done in Chapter 5.2.

$$\ln(Size_i) = \widehat{\beta_0} + \widehat{\beta_1} ln(\overline{Income_i}) + \widehat{\beta_2} ln\left(\overline{Sq.meterPrice_i}\right) + \widehat{\alpha_1}\overline{INCdummy_i} + \widehat{\beta_3}\overline{Children_i} + \widehat{\beta_4}\overline{Adults_i} + \widehat{\beta_5}\overline{Elderly_i} + \widehat{\alpha_r} + \widehat{\alpha_p} + \widehat{\alpha_e} + \widehat{\alpha_h} \qquad (2.2)$$

Inserting for averages of all explanatory variables over the time period given by Table 3.1 yields an identical predicted demand curve to that in Figure 5.1. Again, we find the number of square meters demanded each survey year for a household with identical characteristics, income, price, house type and geographical location. Note that the household weight is dropped from this analysis, leading to somewhat higher square meter values compared to Figure 5.1. However, the relation between the predicted demand of each survey year remains unaltered. The predicted demand curve and plotted average square meter values are presented in Figure 5.2.

Figure 5.2 Predicted demand controlled for all explanatory variables graphed with average square meters



Looking at changes over the 13 year period, the predicted demand increases more than the actual mean in square meters. Specifically, given household characteristics, income, price, structural- and spatial components, the increase in square meter house size over the period is 14.9m^2 compared with the actual mean increase of 10.1m^2. Thus, the predicted growth in square meter house size is 4.8m^2 higher than the actual observed average growth in square meter house size. In other words, demand for house size given outside the model or unexplained demand increases over time.

A higher increase in predicted number of square meters than in the actual observed house size suggests that there have to be some change to the explanatory variables over the time period and how they affect house size. Looking at (2.2), this has to be apparent in either the estimated constant term or in the estimated coefficients, since identical average values are inserted for all sample years. For simplification, I say that the demand for square meters inherent in each level of the explanatory variables combined, differ depending on year. When controlled for the explanatory variables, the demand increase is higher than what is actually observed. This indicates that there have to be some changes to these variables over the time period that in combination reduce demand to what is actually observed. Figure 5.2 suggests that the demand inherent in each level of the independent variables combined is 4.8m^2 less in 2013 than in 2001. Rather than explaining the square meter increase, changes to the explanatory variables over time instead seem to dampen the increase in explained demand. So far we have only looked at the impact of changes to right side variables combined. In the next section, however, the relative impact on demand from each of the most important independent variables is studied separately.

5.4.2 Changes to income and prices

In order to separate the effect of increasing income on demand, a simple regression with income as the right side variable is estimated and the predicted demand given by equation (2.3) is computed for each survey year. Similar to previous analysis are the average values over the time period inserted for each year. Plotting the predicted house sizes together gives us a predicted demand curve for income.

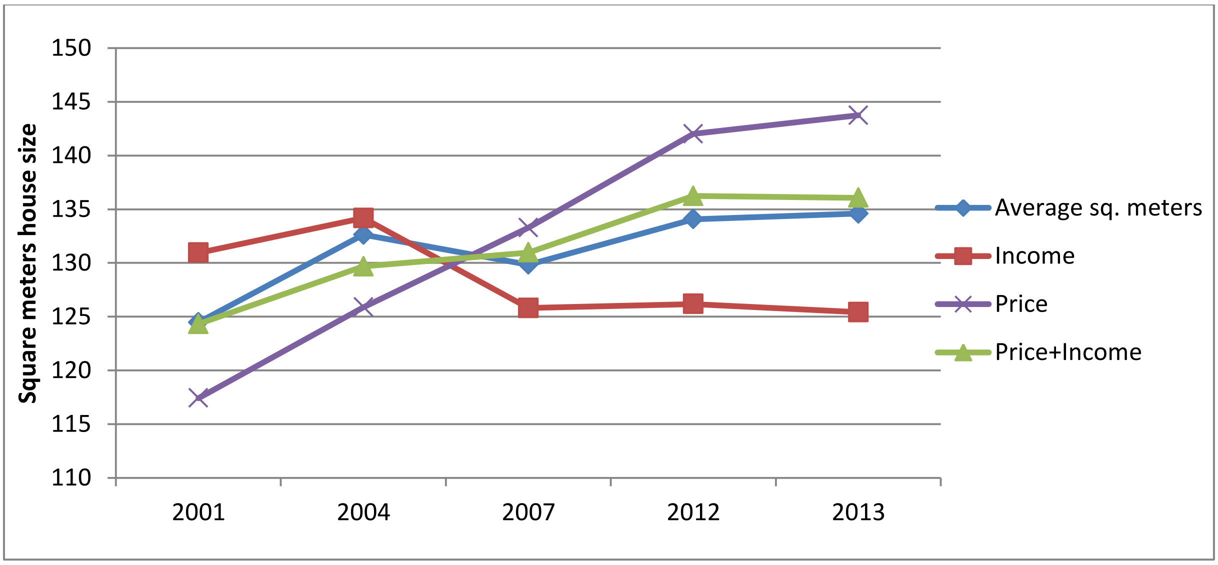
$$\ln(Size_i) = \widehat{\pi_0} + \widehat{\pi_1} ln(\overline{Income_\iota}) \qquad (2.3)$$

$$\ln(Size_i) = \widehat{\gamma_0} + \widehat{\gamma_1} ln(\overline{Sq.meterPrice_i}) \qquad (2.4)$$

$$\ln(Size_i) = \widehat{\delta_0} + \widehat{\delta_1} ln(\overline{Income_\iota}) + \widehat{\delta_2} ln(\overline{Sq.meterPrice_i}) \qquad (2.5)$$

Further, (2.4) is computed treating square meter price as the sole explanatory variable. Finally, a multivariate regression with income and price together is run and predicted values are computed by equation (2.5) following the same procedure. Also, the average number of square meter house size is included. All curves are graphed in Figure 5.3.

Figure 5.3 Predicted square meters controlled for income and price



The average values simply state the mean square meter for each year ignoring how price and income affect the demand in a given year. By controlling for income we get the predicted house size for each survey year if income was fixed. In this case the number of square meters is reduced with 5.5m^2 from 2001 to 2013. However, the actual increase in house size is 10.1m^2 which is 15.6m^2 higher than given by the predicted demand curve for income. A higher actual demand indicates that the demand inherent in households' level of income in 2013 is 15.6m^2 higher than in 2001. This finding should come as no surprise since there has been a steady growth in real income, as we have seen in the descriptive statistic part. In addition, as we saw in 5.1.1, the income elasticity increases over time. The fact that higher demand follows as a consequence of both higher income level and higher income elasticity seems natural.

Considering the square meter house price in a univariate analysis yields opposite results. When controlling for the square meter price, we find a constant increase in predicted demand. In total, predicted demand increases by 26.2m^2 from 2001 to 2013, 16.2m^2 higher than the actual demand. A lower actual number of square meters observed shows that households face prices in 2013 that lower their demand with 16.2m^2 compared to the price level of 2001. As seen in the empirical analysis, the price elasticity does not deviate much for the different years. Hence, the increasing square meter prices seen in the period negatively affect the demand for house size.

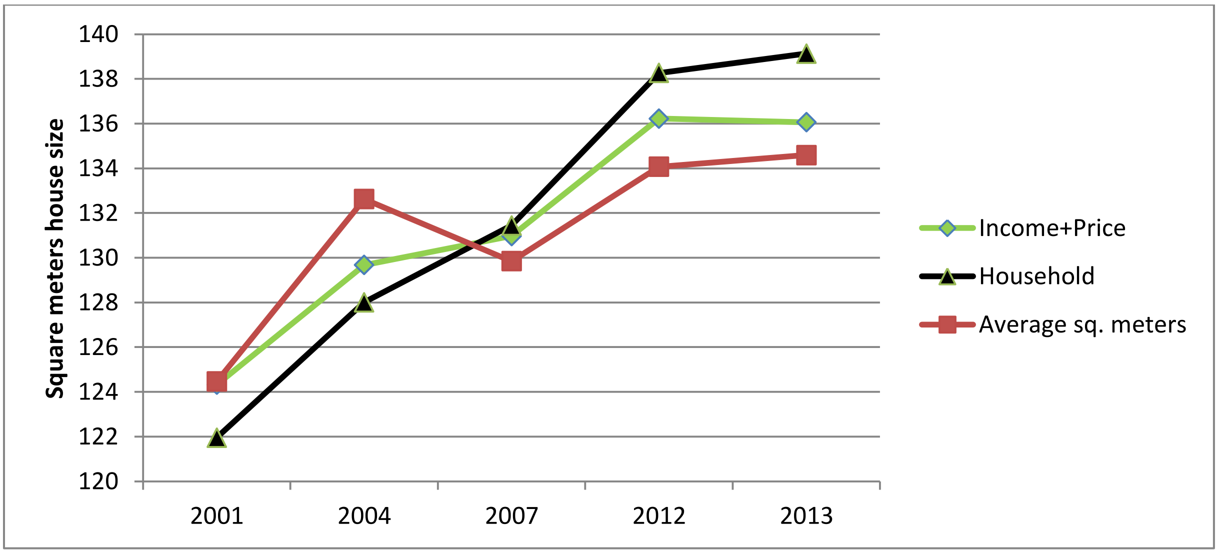
Finally, computing (2.5) creates a graph closer to the graphed average. The predicted demand curve lies below the average for the first two years and slightly above for the last two. In the latter years, both price and income seem to matter a lot, evident by the discrepancy between the predicted demand curves controlled for price and income, indicating the importance of increasing house prices and higher incomes on the number of square meters demanded. However, the effects are contrary to each other, resulting in only small deviations from average values when both are included in the regression. Said differently, income and price changes are not explaining the increased house sizes seen in the period. Specifically, once controlled for both price and income the total increase in households' demand is 1.5m^2 higher in 2013 than in 2001. Instead of explaining the increased demand, price and income changes over the period combined decrease the demand slightly, indicating a marginal domination of the price effect. Note that the income effect might be underestimated since the income variable is truncated and increases in income above the capped value will not be picked up by the model.

5.4.3 Changes to household composition

$$\ln(Size_i) = \widehat{\lambda_0} + \widehat{\lambda_1} ln(\overline{Income_i}) + \widehat{\lambda_2} ln(\overline{Sq.meterPrice_i}) + \widehat{\lambda_3}\overline{Children_i} + \widehat{\lambda_4}\overline{Adults_i} + \widehat{\lambda_5}\overline{Elderly_i} \quad (2.6)$$

To reveal the effect of changes to household composition on demand, the number of children, adults and elderly is included in the regression with price and income. Predicted demand is found by computation of equation (2.6). The predicted house sizes are graphed over time in Figure 5.4 together with mean values. Also, the graph for predicted demand from the regression with income and price given by (2.5) is included in the figure.

Figure 5.4 Predicted demand controlled for income, price and household composition



Summarizing over the time period shows a 10.1m^2 increase in mean square meter values from 2001 to 2013. Once controlling for price, income and household composition, the demand increases with 17.1m^2. The unexplained demand increases with 5.5m^2 as an isolated effect when I control for household composition. Interestingly, instead of explaining the observed demand increase, changes to price, income and household composition call for lower house size demand, increasing the unexplained demand in the model. More precisely, households have characteristics, income and face prices in 2013 that demand 7.0m^2 less compared to 2001. As already discussed, the household size is reduced over the time period, giving a plausible reason why the demand for housing inherent in the composition of the household is lower. However, lower marginal effects of adults and elderly over time go in the direction of lower estimates. Although this effect decreases the importance of household composition, the tendency of lower parameter values for adults and elderly is not as clear in the simple regression as in the main analysis.

5.4.4 Changes to other explanatory variables

The shift analysis was carried out for all explanatory variables by including each of them in a regression with price and income. The main result of this process reveals that changes to these variables and their separate effects on house size had low impact on demand compared to the ones already treated. Although they are not all elaborated on separately, figures are included in the appendix. One of these is the effect of education, which remains constant even as educational attainment increases considerably. A plausible explanation is the lower marginal effects of education already discussed.

A surprising finding, however, comes apparent once dummies of dwelling type are included in the regression with price and income. Summarizing over the time period shows a slightly lower predicted demand increase than the increase in actual values, that is, lowering the unexplained demand once I have controlled for house type. I find these findings counter-intuitive since there has been a trend from detached toward multidwelling houses in the sample period. One would think that households in 2013 live in housing types that demand less square meters than in 2001 controlling for income and price but findings indicate that they demand 1.6m^2 more as seen in Figure A.1.

Finally, a caution should be made regarding this method. By excluding important explanatory variables in the regressions, as done in this chapter, may bias the estimated coefficients. As already mentioned, the marginal effect of adults and elderly in the full model was reduced over time. However, in the simpler model with price, income and household characteristics as the only right side variables, this decline in parameter values was not as apparent. As a consequence, the demand reduction that most likely is caused by the trend towards smaller households may be weaker than the true value. For the other variables, estimated coefficients were similar and showing the same trends as the main empirical analysis. Later, the problem with omitted variable bias will be discussed more generally.

6 Correct-sized dwellings

6.1 House size demand for correct-sized dwellings

A mentioned problem with the housing market is that actors do not act according to their underlying demand because of high information cost and psychological costs of moving. In the survey on living conditions, one of the questions sounds: "How does the house size fit you?". The response alternatives are "too large", "correct-sized" and "too small". The answer distribution for the surveys in 2001, 2004, 2007 and 2012 are given by Table 6.1. In 2001, 11.1% or 242 households answered that they live in a too small sized dwelling which is about the same percentage as of households reporting living too large. This distribution is fairly stable over time although a higher percentage of households report they live too large for the other years. In 2012, 8.7% say they live too small while 12.4% report they live too large.

Table 6.1 Household opinion of house size

	2001		2004		2007		2012	
How does the house size fit you?	Frequency	Percentage	Frequency	Percentage	Frequency	Percentage	Frequency	Percentage
Correct-sized	1784	77.7 %	1906	77.0 %	1832	76.8 %	3448	78.8 %
Too large	257	11.2 %	247	12.8 %	255	12.1 %	398	12.4 %
Too small	242	11.1 %	301	10.3 %	263	11.1 %	505	8.7 %

Note, this question regarding house size is not asked as a part of the 2013 survey, thus this year is dropped. Numbers are reported as number of households and as the percentage of all households. The household weight is applied.

Households living too small or too large both consume another amount of square meters compared to what they ideally demand. Here ideal demand is defined as the house size they would demand if the household could move into a right sized dwelling. By looking only at right sized dwellings make us closer to identification of households' ideal demand since those that want to live in another sized house are dropped out of the sample. As a consequence higher estimated coefficients may be expected as seen in the literature with other approaches, for instance only looking at recently moved persons (R. K. Green & Malpezzi, 2003). Households not acting according to their ideal demand may take house size decisions on other grounds than what suggested by my theoretical model. By dropping these observations may remove noise in the model, thus produce stronger effects in absolute value of the estimated coefficients. I test if the effects of my explanatory variables are different when excluding households that do not feel they live in a correct sized dwelling. Considerable changes in estimates indicate that these "wrong-sized" dwellings bias my results casting doubt on my main findings.

Intuitively, of the families living cramped there is an overweight of households with children while older people stay in their large houses after kids move out making them overrepresented in the too large category. This tendency is found in the dataset for 2012, where 25% of the single parent households and 23% of couples with children aged 0-6 respond they live too small. Also about 20% of the 45 years and older answer they live too large (Sandlie & Grødem, 2013).

A regression with the main empirical specification given by equation (2.1) is run but dropping all observations from households living either too small or too large. Estimated coefficients, standard errors and p-values are given in Table 6.2. In addition, output from the main analysis of 2012 is included to ease the comparison. The ceteris paribus effect of number of children increases with 3-4 percentage points for all years once the wrong-sized dwellings are dropped from the samples. Estimating the model dropping too large and too small observations separately indicates that the increased demand effect of number of children mainly come from dropping households living in too small dwellings. Also, about a percentage point increase arise from dropping the too large group, similar effects all years. Moreover, effects of adult and old on demand are also higher than the main analysis although in smaller magnitude compared with children.

Table 6.2 Demand for house size, correct-sized dwellings, OLS-regression results

	2001			2004			2007			2012			Main analysis		2012
Correct-sized houses	**Coeff.**	**Std. Err.**	**P-value**	**Coeff.**	**Std. Err.**	**P-value**	**Coeff.**	**Std. Err.**	**P-value**	**Coeff.**	**Std. Err.**	**P-value**	**Coeff.**	**Std. Err.**	**P-value**
After tax household income	0.104	0.0209	0	0.164	0.0257	0	0.231	0.0303	0	0.201	0.0200	0	0.207	0.0192	0
Price per square meter	-0.179	0.0211	0	-0.224	0.0251	0	-0.225	0.0215	0	-0.195	0.0157	0	-0.207	0.0197	0
Truncated income dummy	0.161	0.0469	0.001	0.178	0.0289	0	0.202	0.0243	0	0.124	0.0168	0	0.127	0.0167	0
Household characteristics															
Number of children	0.098	0.0090	0	0.081	0.0077	0	0.067	0.0078	0	0.084	0.0058	0	0.052	0.0053	0
Number of adults	0.089	0.0304	0.003	0.060	0.0159	0	0.025	0.0170	0.138	0.059	0.0122	0	0.044	0.0132	0.001
Number of elderly	0.094	0.0248	0	0.084	0.0197	0	0.055	0.0201	0.006	0.087	0.0150	0	0.073	0.0144	0
Region															
Oslo_Akershus	(omitted)														
Hedemark_Oppland	-0.168	0.0373	0	-0.159	0.0318	0	-0.190	0.0311	0	-0.108	0.0266	0	-0.105	0.0282	0
Sør-Østlandet	-0.14	0.0280	0	-0.123	0.0267	0	-0.148	0.0257	0	-0.075	0.0198	0	-0.063	0.0206	0.002
Agder_Rogaland	-0.118	0.0276	0	-0.113	0.0277	0	-0.110	0.0259	0	0.032	0.0199	0.108	0.024	0.0198	0.225
Vestlandet	-0.078	0.0288	0.007	-0.023	0.0265	0.381	-0.085	0.0253	0.001	-0.005	0.0191	0.813	0.001	0.0185	0.936
Trøndelag	-0.149	0.0312	0	-0.156	0.0296	0	-0.152	0.0294	0	-0.074	0.0222	0.001	-0.040	0.0230	0.085
Nord-Norge	-0.198	0.0341	0	-0.239	0.0358	0	-0.231	0.0328	0	-0.103	0.0258	0	-0.103	0.0256	0
Size of populated area															
Sparsely populated area (< 200)	(omitted)														
Urban settlement (200 ≤ > 20 000)	0.064	0.0225	0.004	0.039	0.0237	0.100	0.052	0.0227	0.021	0.061	0.0181	0.001	0.069	0.0172	0
Densely populated area (20 000 ≤ > 100 000)	0.124	0.0265	0	0.141	0.0285	0	0.129	0.0260	0	0.095	0.0222	0	0.108	0.0222	0
Cities (> 100 000)	0.099	0.0322	0.002	0.144	0.0309	0	0.101	0.0310	0.001	0.112	0.0220	0	0.132	0.0229	0
Level of attained education															
Primary and lower secondary	(omitted)														
Upper secondary education	0.049	0.0233	0.037	0.046	0.0218	0.037	0.027	0.0202	0.182	0.054	0.0160	0.001	0.055	0.0162	0.001
Higher education lower level (≤ 4 years)	0.094	0.0283	0.001	0.106	0.0275	0	0.036	0.0231	0.115	0.097	0.0171	0	0.091	0.0180	0
Higher education higher level (> 4 years)	0.176	0.0472	0	0.163	0.0376	0	0.038	0.0347	0.272	0.085	0.0229	0	0.089	0.0227	0
Doctorate	0.215	0.0705	0.002	0.156	0.0626	0.013	0.121	0.0543	0.026	0.148	0.0524	0.005	0.145	0.0493	0.003
Missing education	0.175	0.0764	0.022	0.079	0.0435	0.069	-0.001	0.0369	0.977	-0.048	0.0978	0.620	-0.071	0.0854	0.404
Type of housing															
Apartment	(omitted)														
Detached house	0.513	0.0286	0	0.517	0.0260	0	0.527	0.0259	0	0.511	0.0195	0	0.586	0.0199	0
Linked house	0.230	0.0288	0	0.256	0.0297	0	0.261	0.0281	0	0.239	0.0201	0	0.266	0.0193	0
Semi-detached house	0.268	0.0368	0	0.216	0.0332	0	0.213	0.0330	0	0.208	0.0227	0	0.242	0.0218	0
Other housetypes	0.147	0.1023	0.151	0.410	0.1032	0	0.076	0.1274	0.553	0.150	0.1125	0.183	0.052	0.2581	0.840
Constant	4.583	0.7729	0	4.286	0.3942	0	3.548	0.4292	0	3.545	0.2821	0	3.597	0.3603	0
Number of observations	1784			1906			1830			3448			4350		
R-Squared	0.560			0.580			0.612			0.599			0.566		

Another difference by looking only at correct-sized dwellings is the reduced effect of other house types versus apartments for all years. The decrease is largest for detached houses where a household demand 7.5 percentage points less size only looking at correct-sized houses in 2012. Also the other house types experience lower effect compared to multidwelling houses, similar effects all years. Here the impact is largely driven by removing too large sized houses from the samples. Mean square meter house size deviate substantially depending on how respondents report their opinion on house size. Interview objects who answered they live too large have an average house size of 183.5m^2 in 2012. This number drops down to 136.2m^2 for correct-sized and to 102.2m^2 for those who answered too small. Moreover, the reported opinion on house size is largely reflected in the housing type of current dwelling. Only 38.2% of households who report living too small have a detached house compared to 83.7% for too large. Thus, removing large and detached houses from the sample reduces the demand effect versus apartments.

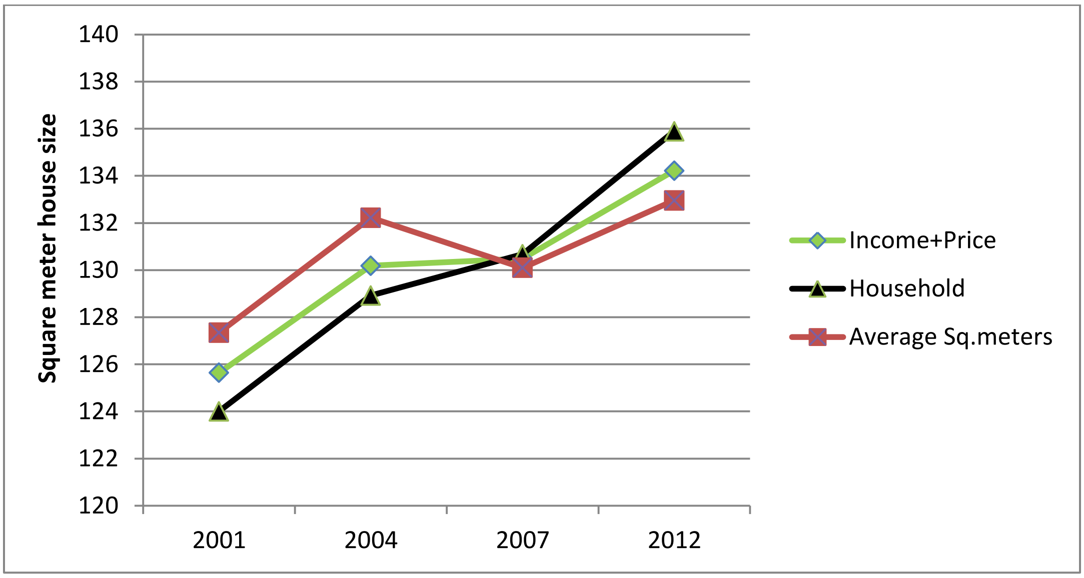
On the whole, comparing estimates for right-sized households with the main analysis reveals that size of households becomes a more important component of demand looking at all years. Moreover, all house types lower their demand response versus apartments. Still, the other parameter values only changes marginally showing robustness of my main result.

6.2 Shift analysis for correct-sized dwellings

A replication of the shift analysis shown in Figure 5.4 is carried out only looking at right-sized dwellings. If the shift analysis for households living correct-sized looks similar to the one including all households, this is an indication of strength of my main model. Hence, demand for house size is not determined by households living wrong-sized making their demand decision on other grounds than by my model.

$$\ln(Size_i) = \widehat{\lambda_0} + \widehat{\lambda_1} ln(\overline{Income_i}) + \widehat{\lambda_2} ln(\overline{Sq.meterPrice_i}) + \widehat{\lambda_3}\overline{Children_i} + \widehat{\lambda_4}\overline{Adults_i} + \widehat{\lambda_5}\overline{Elderly_i} \quad (2.6)$$

Figure 6.1 Predicted demand for correct-sized dwellings controlled for income, price and household composition



None of my conclusions change by dropping the wrong-sized dwellings in the shift analysis displayed in Figure 6.1. The predicted demand of income, price and household composition, given by equation (2.6), increases with more than the average square meter values, that is, an increase in unexplained demand, equivalent to what we saw in Figure 5.4. Changes to price, income and household composition increase the unexplained demand with 6.2m^2 from 2001 to 2012 for right-sized dwellings compared to an increase of 6.7m^2 looking at all households. As a consequence of only minor demand changes, we see that the importance of income, price and household composition changes are not driven by households in disequilibrium acting far from their ideal demand. Other shift-analysis was preformed looking at changes to the other explanatory variables for correct-sized dwellings. However, none of the conclusions changed, thus not elaborated on further.

7 Econometric challenges and limitations

7.1 Omitted variable bias

7.1.1 Quality and the age of the structure

In the current analysis, no variables describing the quality of the dwellings are included. Naturally, the quality of the house impacts the decision of how many square meters to acquire. Specifically, for a given house size the square meter price will increase with the quality of the dwelling, all else equal. A potential variable capturing quality is the age of the dwelling. In an old house, maintenance is costly, making the dwelling for a given price less attractive. Other reasons favoring a newer house include more favorable design, superior floor plan and more practical solutions. I here assume a negative correlation between the demand for house size and the age of the dwelling. Correlation between number of square meter demanded and age of structure is isolated not a problem. However, an omitted variable bias arises when there is correlation between the age of structure and other explanatory variables. In this case, the zero conditional mean assumption for OLS no longer holds. With k independent variables we get that:

$$E(u|x_1, x_2, \dots, x_k) \neq 0$$

Here, the age of structure will be in the error term and any correlation between included variables will break the assumption. Of particular concern is the correlation between square meter price and age of the dwelling. As a simplification we assume the true regression is given by:

$$\ln(KVM_i) = \beta_0 + \beta_1 ln(PriceKVM_i) + \beta_2 AgeDwelling_i + u_i$$

Instead we estimate a simple regression with price per square meter as a sole independent variable. The OLS estimator for the price coefficient can be shown to equal:

$$E\left(\widehat{\beta_1}\right) = \beta_1 + \beta_2\hat{\delta}$$

$\hat{\delta}$ is the estimated regression coefficient in a regression between square meter price and age of the dwelling. Also, this coefficient is assumed to be negative as the factors affecting the demand for house size also affect prices. We get that $E\left(\widehat{\beta_1}\right) > \beta_1$. Since both β_2 and $\hat{\delta}$ are negative, OLS will estimate the price elasticity to be less negative than the true value. In other words, the OLS estimator is, under the assumptions made, potentially biased towards zero, showing lower price elasticity than the actual value. A caution should be made regarding the direction of the bias in my main model as pairwise correlation between the explanatory variables can affect the direction of the bias (Wooldridge, 2012). Unfortunately, no variable giving the age of the structure was found in the data sets.

A way to incorporate the quality aspect into the price is by running a hedonic regression. As described by Zabel (2004) is the price run against variables describing the quality of the dwelling including age of the structure, number of rooms, bedrooms, bathrooms *et cetera*. A new adjusted square meter variable could have been constructed. First, by including the number of square meter house size alongside quality variables on the right side in a regression with price as left side variable will let us estimate the importance of size and quality on the price. Then each quality variable is weighted with its importance versus size on the price given by the hedonic regression. Adjusted number of square meters is given by house size and weighted quality variables. As a result, dwellings with favorable treats that increase their price have a higher number of adjusted square meters. I was unable to use an adjusted square meter as my dependent variable because of the limited choice of quality variables in the data sets.

7.1.2 Other variables

The danger of omitting important variables is a lingering issue in regression analyses. Other variables that may cause problems of endogeneity are preferences or unobserved heterogeneity in general. When preferences for house size are systematically correlated with included variables, an omitted variable bias of the OLS estimator follows. Unobserved heterogeneity is a general problem with cross sectional estimation that could potentially been solved with panel data (Wooldridge, 2012).

7.2 Measurement error

When there are discrepancies between the theoretical correct variable and the one we observe, then our model contains a measurement error.

7.2.1 In explanatory variables

Measurement error in an independent variable can be a serious problem if the error is uncorrelated with the true unobserved explanatory variable, also called the classical errors-in-variables (CEV). In the presence of CEV, the measurement error is correlated with an included explanatory variable, violating the zero conditional mean assumption. Estimation bias for the OLS estimator due to CEV is called the attenuation bias. A consequence of this type of bias is that OLS underestimates the absolute value of the coefficient which has a measurement error. The direction of the bias is towards zero, making the estimator less significant (Wooldridge, 2012).

An example of this type of measurement error already discussed is the use of current rather than permanent income. If permanent income enters in the true model, OLS underestimates the income elasticity of house size demand. Specifically, current income underestimates the effect of a wage increase on the demand for house size.

7.2.2 In the dependent variable

There is a possibility that some households report an incorrect number of square meters house size. The observed house size is given as:

$$Size = Size^{*} + \varepsilon$$

where $Size^{*}$ is the true unobserved number of square meter house size and ε is the measurement error. In order for OLS to be an unbiased estimator, there can be no correlation between the measurement error and the error term, assuming no measurement error in the explanatory variables. A normal assumption states that the measurement error in the depended variable cannot be correlated with any of the included variables. As a consequence, the conditional expectation of the measurement error on explanatory variables is zero.

$$E(\varepsilon | x_1, x_2, \dots, x_k) = 0$$

Under this assumption, measurement error in the dependent variable gives no bias for the OLS estimator. The only consequence relates to the variance of the error term, which increases by the inclusion of the measurement error (Wooldridge, 2012).

7.2.3 In the price elasticity

In my analysis, the assumption regarding no correlation between the measurement error in number of square meters (ε) and explanatory variables does not hold. The reason for this lies in the price variable as it is constructed from the dependent variable. A measurement error house size will be correlated with price per square meter.

$$Sq.meterPrice = \frac{Assumed\ market\ price}{Size} = \frac{Assumed\ market\ price}{Size^{*} + \ \varepsilon}$$

If a household reports larger house size than the real value (ε↑) then the square meter price will be lower than the true value. On the other hand, undervaluation of house size (ε↓) positively affects the square meter price as the house price is divided by a too small house size. In other words, the relationship between square meters demanded and the price per square meter becomes more negative in the presence of measurement error in house size. As a result, estimated price elasticity will be more negative compared to a situation without measurement error in house size.

However, measurement error in the assumed sales price will be positively correlated with the square meter price. Under the assumptions for CEV, measurement error in the sales price will bias the price elasticity towards zero. On the whole, measurement error in both the house price and number of square meters affects the estimated price elasticity in opposite directions, leading to an ambiguous direction of the bias.

A possible solution to the problem of endogenous square meter price is by using instrument variable (IV) estimation. Here we use another variable as instrument for the square meter price which is not subject to the measurement error mentioned. The instrument needs to fulfill two requirements:

$$Cov(z, u) = 0$$
$$Cov(z, Sq.meterPrice) \neq 0$$

The exclusion restriction states that instrument (z) needs to be uncorrelated with the error term. In addition, the instrument needs relevance, that is, correlation between the instrument and the variable being instrumented is required (Wooldridge, 2012). A potential instrument could be the average square meter price for each region, size of populated area and housing type. IV-estimation should fix the exact problem of measurement error in the square meter price caused by incorrectly reported number of square meters. However, there could also be measurement error in using average square meter prices.

7.3 Other remarks

7.3.1 Heteroskedasticity

Another OLS assumption states that the error term has a constant variance given all the explanatory variables:

$$Var(u \mid x_1, x_2, \ldots, x_k) = \sigma^2 < \infty$$

If this assumption holds, we can say that the error term is independent and identically distributed, or homoscedastic. Even with heteroskedasticity, OLS is an unbiased estimator for the true coefficient. However, OLS no longer is the best linear unbiased estimator (BLUE), since the variance tends to be higher. Another consequence is that the normal methods of interference no longer are valid because of the affected variance. A common way to correct for heteroskedasticity is by using robust standard errors where the squared residuals of the regression $(\hat{u}_i^2)$ are inserted for variance of the error term(σ_i^2) (Wooldridge, 2012).

To study whether heteroskedasticity is a problem in my analysis is complicated by the household weight. The weight assigns lager importance to observations with few adults. Using robust standard errors also adjust the weighting of observations by giving less weight when the disturbance variance is high. In fact, using weights allow us to solve the problem with heteroskedasticity also called Weighted Least Square (Williams, 2015). The classical ways to test whether heteroskedasticity is present does not allow for probability weights. There is a possibility to weight each variable in order to adjust for heteroskedasticity and keeping the household weight, but the process is tedious and beyond the scope of this text.

7.3.2 Net versus gross income

As already mentioned, the income variable for some years is truncated, affecting the mean values. There is also a potential endogeneity problem in using after tax income. Because of tax deduction will the size of the mortgage affect the after tax income. Specifically, by comparing two households with the same gross income, a highly indebted household will have larger after tax income than a household with lower debt, all else equal. Hence, net income is correlated with level of debt, which may in turn matter for house size.

On the other hand, gross income does not give a precise measurement of the disposable or permanent income, which is the variable of interest. The measurement error is given by the discrepancy between gross income and the true income variable affecting the income elasticity. In order to use an income variable closer to the true value, after tax income was chosen in spite of the endogeneity problem mentioned.

8 Conclusion

The purpose of this thesis is twofold. First goal is to identify which factors determine the owner-occupiers' demand for house size. Second task is to explain why the average number of square meters has increased from 2001 to 2013.

House size is only one of many components a household considers when buying a house. Still, house size is an important element of the housing package and can be understood as a housing service. A theoretical model is presented showing how a household maximizes its utility under a budget constraint. An equilibrium house size is found graphically, where the optimal size increases with income while the interest rate and the square meter price reduces the optimal number of square meters.

For the empirical analysis, I use the Survey of Living Conditions collected by Statistics Norway for five survey years between 2001 and 2013. An identical empirical model with the number of square meter house size as dependent variables is specified. Explanatory variables include net household income, square meter price, number of children, adults and elderly in the household, education, house type and geographical variables. Separate Ordinary Least Squares (OLS) regressions are run for the five sample years. Estimated income elasticity of house size is 0.207 in 2013, an increase from 0.136 in 2001. Although the estimates are similar in magnitude to what is found in the literature by Larsen (2014) and Zabel (2004), is comparison not possible due to differences in how the dependent variables are defined. Interestingly, the demand effect from number of adults and elderly fall over the time period while remaining stable for children, increasing house size with roughly 5.5 percent with each new child in the household. Note the regression output from 2007 deviate considerably compared to other years, showing low levels of significance in particular for household composition and education.

To measure how demand changes given all my explanatory variables, I insert for the average values for all independent variables and compute a predicted house size for each survey year. I find that predicted demand rises every year, indicating the potential presence of other factors explaining the 8.4 square meters increase in average house sizes found in my data. Potential candidates include macroeconomic variables and housing preferences. In particular, lower interest rates, higher expectations, more available credit and a favorable tax system can explain why households buy larger homes.

In a shift analysis, I look at how inherent changes to my right side variables in the data period affect the demand for house size. My findings suggest that changes to price and income to a large extent go against each other, leaving no explanation to why larger house sizes are observed. Changes to household composition, on the other hand, reduce demand caused by a tendency towards smaller households. The result of the shift analysis including all variables combined shows an increase in unexplained demand of 4.8 square meters, demonstrating that inherent changes to all variables over the 13 year period together drive in the direction of smaller dwellings. Thus, factors not included in the cross-sectional analysis like the macroeconomic variables mentioned are even more important in explaining why we live bigger.

An identical empirical analysis is pursued, only looking at households which report they live in a correct-sized dwelling. Generally, few discrepancies from the main analysis indicate that households act close to what they ideally demand showing strength of my empirical specification.

There are, however, some clear weaknesses of my analysis. First, households do not decide on house size alone but rather on a range of considerations, both with a consumption and investment motivation, making up the housing package. One aspect of the housing good not considered is the quality of the dwelling. The zero conditional mean assumption of the error term is violated in the presence of a quality variable, for instance the age of structure, which is correlated with included explanatory variables. Also unobserved heterogeneity within each household, neighborhood and structure may give omitted variable bias of the OLS estimator.

Another concern relates to measurement error which can be issue with self-reported variables. Normally measurement error in explanatory variables bias estimators towards zero making them less significant. In my analysis, however, price per square meter is defined from the dependent variable. A measurement error in the reported number of square meters will affect the estimated price elasticity making it more negative than the true value. Last, there may be an endogeneity problem by using after tax income because of correlation with the level of debt which arguably affects house size.

References

Angelini, V., & Laferrere, A. (2011). Residential Mobility of the European Elderly. *CESifo Economic Studies, 58*(3), 544-569.

Artle, R., & Varaiya, P. (1978). Life cycle consumption and homeownership. *Journal of Economic Theory, 18*(1), 38-58.

Baltzersen, M. (2015). Finanstilsynet proposes establishing requirements on residential mortgage lending practices in the form of regulations. [Press release]. Retrieved from http://www.finanstilsynet.no/en/Document-repository/Press-releases/2015/Q1/Finanstilsynet-proposes-establishing-requirements-on-residential-mortgage-lending-practices-in-the-form-of-regulations/

Case, K. E., Quigley, J. M., & Shiller, R. J. (2003). Home-buyers, Housing and the Macroeconomy. *Berkeley Program on Housing and Urban Policy*.

Dusansky, R., Koç, Ç., & Onur, I. (2010). Household Housing Demand: Empirical Analysis and Theoretical Reconciliation. *The Journal of Real Estate Finance and Economics, 44*(4), 429-445.

Finansdepartementet. (2004). *Skatte- og avgiftsopplegget 2005 - lovendringer*. Retrieved from https://www.regjeringen.no/contentassets/8023dc53f8574179a8ecc2cc3aeea4ce/no/pdfs/otp200420050001000dddpdfs.pdf.

Finansdepartementet. (2009). Mer rettferdig formuesverdsetting av boliger [Press release]. Retrieved from http://www.statsbudsjettet.no/Statsbudsjettet-2010/Satsinger/?pid=38789#hopp

Finansdepartementet. (2013). *Prop. 1 LS (2013–2014), Skatter, avgifter og toll 2014*.

Friedman, M. (1957). The permanent income hypothesis *A theory of the consumption function* (pp. 20-37): Princeton University Press.

Gillingham, R., & Hagemann, R. (1983). Cross-sectional estimation of a simultaneous model of tenure choice and housing services demand. *Journal of Urban Economics, 14*(1), 16-39.

Goodman, A. C., & Kawai, M. (1982). Permanent income, hedonic prices, and demand for housing: New evidence. *Journal of Urban Economics, 12*(2), 214-237.

Green, R., & Hendershott, P. H. (1996). Age, housing demand, and real house prices. *Regional Science and Urban Economics, 26*(5), 465-480.

Green, R. K., & Malpezzi, S. (2003). *A primer on US housing markets and housing policy*: The Urban Insitute.

Ioannides, Y. M., & Zabel, J. E. (2003). Neighbourhood effects and housing demand. *Journal of applied Econometrics, 18*(5), 563-584.

Krogh, T. S. (2010). Credit regulations in Norway, 1970-2008.

Larsen, E. R. (2014). The Engel curve of owner-occupied housing consumption. *Journal of Applied Economics, 17*(2), 325-352.

Mankiw, N. G., & Weil, D. N. (1989). The baby boom, the baby bust, and the housing market. *Regional Science and Urban Economics, 19*(2), 235-258.

Mayo, S. K. (1981). Theory and estimation in the economics of housing demand. *Journal of Urban Economics, 10*(1), 95-116.

Nesbakken, R. (2008). Selveiernes boligkonsum og vekt i KPI: estimeringer og diskusjon av metode.

NOU. (2002). *Boligmarkedene og boligpolitikken: utredning fra Boligutvalget oppnevnt ved kgl. res. 22. desember 1999 : avgitt til Kommunal- og regionaldepartementet 15. februar 2002* (Vol. NOU 2002: 2). Oslo: Statens forvaltningstjeneste. Informasjonsforvaltning.

Refling, D. (2013). Eiendomskatt i kommunene 2001-2013: Huseiernes landsforbund.

Rothenberg, J. (1991). *The Maze of urban housing markets: theory, evidence, and policy*. Chicago: University of Chicago Press.

Sandlie, H. C., & Grødem, A. S. (2013). *Bolig og levekår i Norge 2012* (Vol. 14/13). Oslo: Norsk institutt for forskning om oppvekst, velferd og aldring.
Singsaas, H. (2015). Styringsrenten uendret på 1,25% [Press release]. Retrieved from http://www.norges-bank.no/Publisert/Pressemeldinger/2015/Pressemelding-19-mars-2015/
Skatteeaten. (2015). Formuesskatt. Retrived from http://www.skatteetaten.no/no/Tabeller-og-satser/Formuesskatt/?ssy=2001#formulaDiv
Sribney, W. (1997). Probability weights, analytic weights, and summary statistics. Retrived from http://www.stata.com/support/faqs/statistics/weights-and-summary-statistics/
SSB. (2000). Norsk standard for utdanningsgruppering. Retrieved 24.3.2015 from http://www.ssb.no/a/publikasjoner/pdf/nos_c617/nos_c617.pdf
SSB. (2015a). Arbeidskraftundersøkinga, sesongjusterte tal. Retrieved 15.04.15, from Statistics Norway https://www.ssb.no/statistikkbanken/selecttable/hovedtabellHjem.asp?KortNavnWeb=akumnd&CMSSubjectArea=arbeid-og-lonn&checked=true
SSB. (2015b). Boligprisindeksen, alle boliger. Retrieved 14.04.15 from https://www.ssb.no/a/aarbok/tab/tab-272.html
SSB. (2015c). Folkemengde, etter fylke 1. januar. Retrieved 13.03.15 from http://www.ssb.no/befolkning/statistikker/folkemengde/aar/2015-02-19?fane=tabell&sort=nummer&tabell=218731

SSB. (2015d). Gjennomsnittlig utlåns- og innskuddsrente i bankene, prisstigning, beregnet realrente, effektiv rente for innenlandske obligasjoner (10 års) og Nibor (3 md.) per 31. desember. Prosent. Retrieved 29.03.15 from https://www.ssb.no/statistikkbanken/selecttable/hovedtabellHjem.asp?KortNavnWeb=akumnd&CMSSubjectArea=arbeid-og-lonn&checked=true
SSB. (2015e). Konsumprisindeks, historisk serie. Retrieved 10.02.2015, from Statistics Norway http://www.ssb.no/213809/konsumprisindeks-historisk-serie-1998-100
Thorsen, L. R., & Revold, M. K. (2014). Levekårsundersøkelsen EU-SILC 2013, Tema: Idretts- og kulturaktiviteter. (Vol. Notater 2014/7): Statistics Norway.
Trigg, A. B. (2001). Veblen, Bourdieu, and conspicuous consumption. *Journal of Economic Issues*, 99-115.
Veblen, T. ([1899]1994). The theory of the leisure class. In the collected works of Thorstein Veblen, vol. 1, Reprinted: London: Routledge.
Williams, R. (2015). Hetroskedasticity. Retrieved 20.06.15 from https://www3.nd.edu/~rwilliam/stats2/l25.pdf
Wilson, A., & Boehland, J. (2005). Small is beautiful US house size, resource use, and the environment. *Journal of Industrial Ecology, 9*(1-2), 277-287.
Wooldridge, J. (2012). *Introductory econometrics: A modern approach*: Cengage Learning.
Zabel, J. E. (2004). The demand for housing services. *Journal of Housing Economics, 13*(1), 16-35. doi: 10.1016/j.jhe.2003.12.002

Appendix

Figure A.1 Predicted demand controlled for income, price and housing type

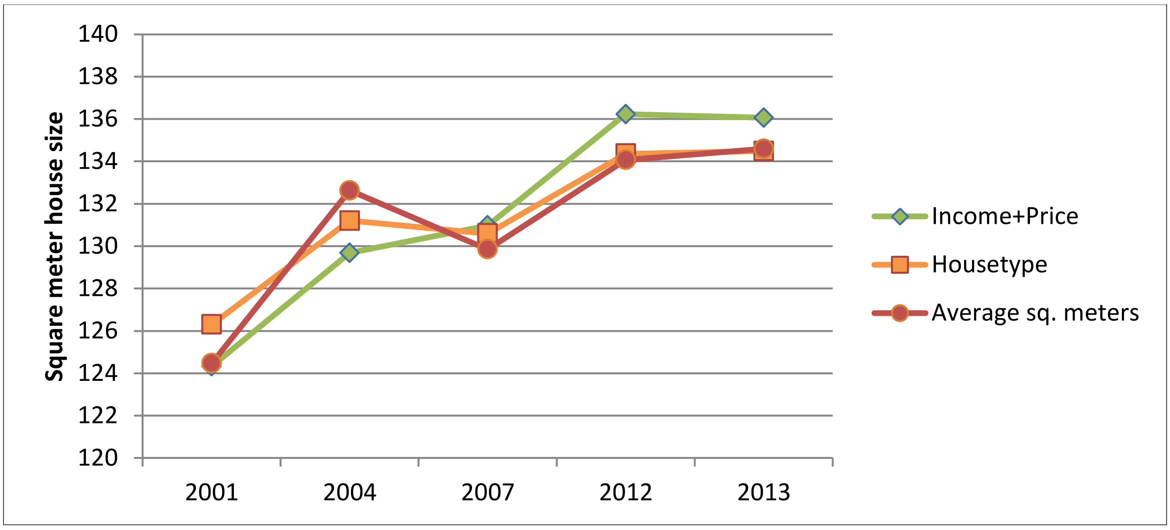


Figure A.2 Predicted demand controlled for income, price and region

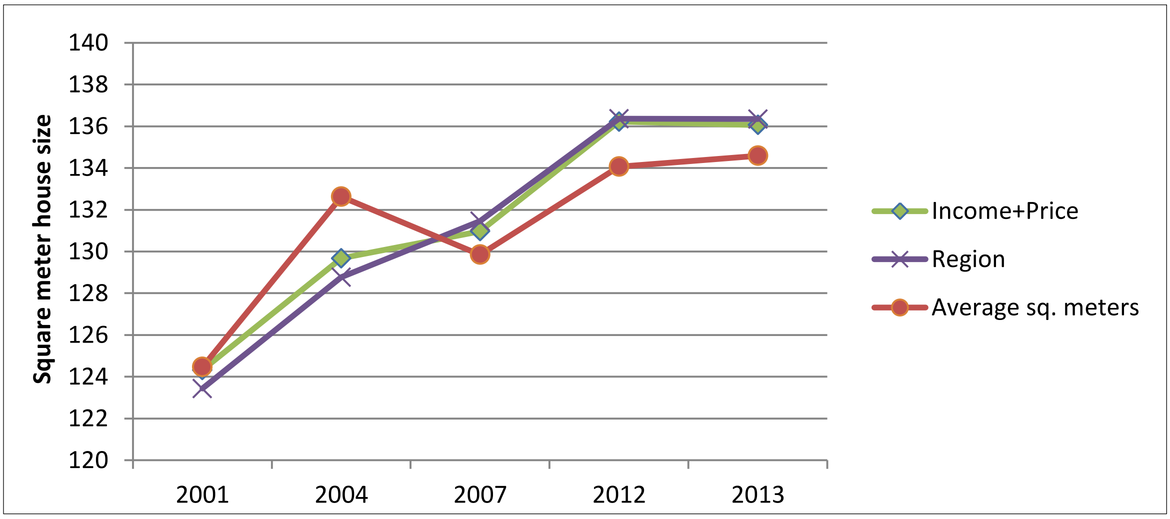


Figure A.3 Predicted demand controlled for income, price and size of populated area

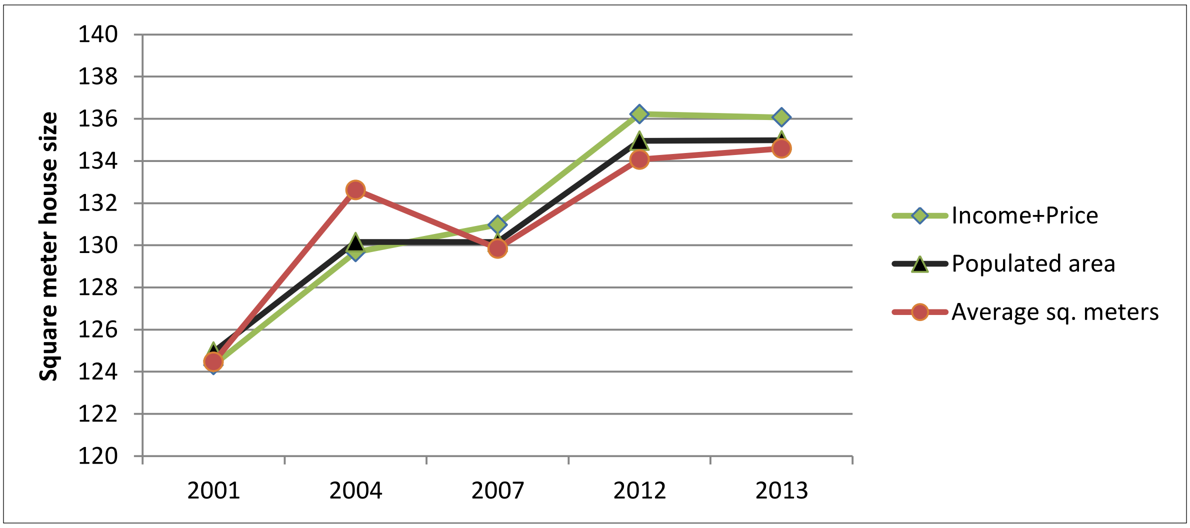


Figure A.4 Predicted demand controlled for income, price and educational attainment