2015-21

Simon Juul Hviid

PhD Thesis

Dynamic Models of the Housing Market

DEPARTMENT OF ECONOMICS AND BUSINESS ECONOMICS
AARHUS UNIVERSITY • DENMARK

Dynamic Models of the Housing Market

Simon Juul Hviid

A dissertation submitted to
Business and Social Sciences, Aarhus University,
in partial fulfilment of the requirements of
the PhD degree in
Economics and Business Economics

December 2015

CREATES
Center for Research in Econometric Analysis of Time Series

PREFACE

This dissertation is the result of my PhD studies at the Department of Economics and Business Economics at Aarhus University and Department of Economics at Duke University in the time period September 2012 to August 2015. I thank Realdania and the Knowledge Center for Housing Economics for financing my PhD position. I am grateful to the departments for providing excellent research facilities and Department of Economics and Business Economics at Aarhus University for generous financial support, which has allowed me to attend numerous courses, summer schools, workshops, and conferences in Denmark and abroad. I also would like to acknowledge the financial support from CREATES - Center for Research in Econometric Analysis of Time Series, Norges Bank, Oticon Fonden, Knud Højgaards Fond, and Augustinus Fonden.

There are a number of people I would like to thank. First and foremost, I would like to thank my supervisors Tom Engsted, Rune M. Vejlin and Torben M. Andersen for encouraging me to enroll as a PhD student and for being huge inspirations throughout my PhD studies. I would thank Tom Engsted for his valuable guidance, for always sustaining a high academic standard, and for his insightful contributions to our joint research project. A special thanks to Rune M. Vejlin for always taking the time for me for constructive discussions and for an inspiring and contagious curiosity that has often made me work long nights. I feel truly privileged to have worked with such capacities throughout the years. I am thankful to Solveig N. Sørensen for always being available for help and for proof-reading the chapters in this dissertation.

During the spring 2014, I had the privilege of visiting Christopher Timmins at Department of Economics, Duke University, North Carolina. I am grateful to Chris for his hospitality and attention, and working closely with him has inspired me enormously. There are also a number of people that contributed to a truly pleasant stay, especially James; thanks for sharing an office with me and introducing me to the life around Duke.

At the Department of Economics and Business Economics, Aarhus University, I wish to thank the faculty and seminar participants for providing feedback on my work. The PhD students deserve thanks for a friendly work environment and many insightful discussions. Special thanks goes to Anders and Palle for many hours spend on a broad range of economic endeavors and even more hours spend on non-academic diversions. Jonas M, deserves a thank you for co-founding the "Rust" reading group and the subsequent endless discussions on estimation and identification. Additionally, I would like to thank Morten, Sanni, Anne, Jonas E, Rune L, and many others for countless coffee breaks and many shared laughs.

Finally, and most importantly, I would like to deeply thank my family and friends for unconditional love and support. It means everything to me. A heartfelt thank to my parents and my siblings for unrelenting believe, and for bearing over with me being absentminded at times.

Simon Juul Hviid
Aarhus, August 2015

UPDATED PREFACE

The predefense took place on October 27, 2015. The assessment committee consists of Stig Vinther Møller, Aarhus University and CREATES, Mette Gørtz, University of Copenhagen, and Esteban Aucejo, London School of Economics. I am grateful to the members of the assessment committee for their careful reading of the thesis, and their constructive comments and insightful discussion of my work. Many suggestions have been incorporated in the current version of the thesis, while others remain for the future research.

Simon Juul Hviid
Copenhagen, December 2015

CONTENTS

Summary **vii**

Danish summary **ix**

1 Explosive Bubbles in House Prices? Evidence from the OECD Countries **1**

1.1 Introduction 2
1.2 The Bubble Model 4
1.3 The Econometric Methodology 6
1.4 Empirical Results 14
1.5 Conclusions 23
Acknowledgments 24
References 28

2 Dynamic Residential Sorting - Investigating the Distribution of Capital Gains **33**

2.1 Introduction 34
2.2 Data 36
2.3 The Theoretical Model 44
2.4 The Empirical Model 48
2.5 Empirical Results 54
2.6 Conclusions 65
Acknowledgments 66
References 67
Appendices 71

3 Valuation of Non-Traded Amenities in a Dynamic Demand Model **73**

3.1 Introduction 74
3.2 The Dynamic Demand Model 76
3.3 Data 78
3.4 Estimation 82

3.5 Results 91
3.6 Conclusions 101
Acknowledgments 102
References 103
Appendices 106

SUMMARY

This dissertation comprises three self-contained chapters that all relate to the implications of expectations in the housing market. In general, house price changes are caused by changes to 'fundamentals' and/or speculative behavior. The fundamental price is notoriously difficult to asses as it is influenced by a large set of variables including, but not limited to, income, mortgage rates, property taxes, local factors such as neighborhood attributes, and expectations to future values of these factors. Additionally, households move infrequently as housing is an illiquid asset and transaction costs in the housing market are substantial. Furthermore, households might face borrowing constraints such that they are restricted from some parts of the housing market. Therefore, expectations to the future state of the housing market are an important driver of household choices and, by extension, housing prices. In this context, the three chapters of this dissertation investigate the role of expectations and their implications in inherently dynamic housing markets.

The first chapter "Explosive Bubbles in House Prices? Evidence from the OECD Countries" is co-authored with Tom Engsted (Aarhus University, CREATES) and Thomas Q. Pedersen (Aarhus University, CREATES). In this chapter, we conduct an econometric analysis of speculative bubbles in housing markets. With econometric methods that explicitly allow for explosiveness, i.e. a rational bubble, we investigate the explosive nature of bubbles. First, we apply a univariate right-tailed unit root test procedure on the price-rent ratio in order to identify periods of exuberance. With this sample we then apply a co-explosive VAR framework to test for explosive bubbles. Using quarterly OECD data for 18 countries from 1970 to 2013, we find evidence of explosiveness in many housing markets, thus supporting the bubble hypothesis.

A slightly shorter version of the first chapter has been accepted for publication in *Journal of International Financial Markets, Institutions, and Money.*

The second chapter "Dynamic Residential Sorting - Investigating the Distribution of Capital Gains" estimates a dynamic residential sorting model of housing owners. The model explicitly takes account of transactions costs, borrowing constraints of

households, and allows for forward looking behavior. The focus in this chapter is on how capital gains are distributed geographically and across the wealth distribution. The model is estimated using unique Danish register data from 1992 to 2011 of housing owners. The chapter finds substantial differences in capital gains as the highest wealth decile, i.e. the 10 percent wealthiest households, over the sample receives almost a 2 percentage points larger annual capital gain than the wealth type with the lowest housing investment. Furthermore, I find substantial differences in capital gains geographically. Lastly, it is found that the freeze of property taxes in 2002 enhanced capital gains dispersion and counterfactual simulations show that the progressive Danish taxation scheme from before 2002 could have mitigated parts of the dispersion.

The third chapter "Valuation of Non-Traded Amenities in a Dynamic Demand Model" is co-authored with Christopher Timmins (Duke University) and Rune M. Vejlin (Aarhus University) and was partly written during my stay at Duke University. Using the population-wide Danish register data with precise measures of households' wealth, income, and socio-economic status, we specify and estimate a dynamic structural model of residential neighborhood demand. Our model includes moving costs, forward looking behavior of households, and uncertainty about the evolution of neighborhood attributes, wealth, income, house prices, and family composition. We estimate marginal willingness to pay for non-traded neighborhood amenities with a focus on air pollution. We allow household willingness to pay to vary in household characteristics and argue that low wealth and low income households face borrowing constraints. The willingness to pay of households who are likely borrowing constrained is found to be much more sensitive to changes in wealth than for other households. Our application finds that the dynamic approach adjusts for various biases relative to a comparable static approach.

DANISH SUMMARY

Denne afhandling består af tre selvstændige kapitler, der alle relaterer sig til konsekvenserne af forventninger på boligmarkedet. Generelt er boligprisændringer forårsaget af ændringer til fundamentale forhold og/eller spekulativ adfærd. Den fundamentale pris er notorisk vanskelig at vurdere, da den påvirkes af en lang række variabler, herunder, men ikke begrænset til, indkomst, renter på realkredit, ejendomsskatter, lokale faktorer såsom nabolagskarakteristika og forventninger til fremtidige værdier af disse faktorer. Derudover flytter boligejere sjældent, eftersom boliger er illikvide aktiver og transaktionsomkostningerne på boligmarkedet er betydelige. Desuden kan husholdninger stå overfor lånebegrænsninger, således at de er afholdt fra visse dele af boligmarkedet. Derfor er forventningerne til den fremtidige tilstand af boligmarkedet en vigtig faktor for husholdningernes valg og derigennem boligpriserne. I dette henseeende vil de tre kapitler i denne afhandling undersøge forventningernes rolle og deres konsekvenser i boligmarkeder, som grundlæggende er dynamiske.

Det første kapitel "Explosive Bubbles in House Prices? Evidence from the OECD Countries" er skrevet sammen med Tom Engsted (Aarhus Universitet, CREATES) og Thomas Q. Pedersen (Aarhus Universitet, CREATES). I dette kapitel foretager vi en økonometrisk analyse af spekulative bobler på boligmarkedet. Med økonometriske metoder, som eksplicit tillader eksplosivitet, det vil sige en rationel boble, undersøger vi den eksplosive karakter af boligprisbobler. Først anvender vi en univariat højresidet testprocedure for en enhedsrod i forholdet mellem boligpris og huslejeniveau for at identificere perioder med boligprisbobler. På denne periode anvender vi derefter et co-eksplosivt VAR framework til at teste for eksplosive bobler. Med kvartalsdata for 18 OECD-lande fra 1970 til 2013 finder vi eksplosivitet på flere boligmarkeder, hvilket dermed understøtter boblehypotesen.

En lidt kortere version af dette kapitel er blevet accepteret til publicering i *Journal of International Financial Markets, Institutions and Money*.

Det andet kapitel "Dynamic Residential Sorting - Investigating the Distribution of

Capital Gains" estimerer en dynamisk boligsorteringsmodel for boligejere. Modellen tager eksplicit højde for transaktionsomkostninger, lånebegrænsninger og fremadskuende adfærd. Fokus i dette kapitel er på, hvordan kapitalgevinster er fordelt geografisk og på tværs af formuefordelingen. Modellen estimeres på de unikke danske registerdata fra 1992 til 2011. Kapitlet finder væsentlige forskelle i kapitalgevinsterne, hvor de 10 procent mest formuende over perioden har næsten 2 procentpoint større årlige kapitalgevinster end formuetypen med de laveste boliginvesteringer. Desuden findes der væsentlige geografiske forskelle i kapitalgevinsterne. Slutteligt findes det at fastfrysningen af ejendomsværdiskatterne i 2002 har øget kapitalgevinsternes spredning og kontrafaktiske simulationer viser, at den progressive ejendomsværdibeskatning fra før fastfrysningen i 2002 kunne have afbødet en del af spredningen.

Det tredje kapitel "Valuation of Non-Traded Amenities in a Dynamic Demand Model" er skrevet sammen med Christopher Timmins (Duke University) og Rune M. Vejlin (Aarhus Universitet) og blev delvist skrevet under mit ophold på Duke University. Ved anvendelse af de omfattende danske registerdata, som inkluderer præcise mål for husholdningernes formue, indtægter og socioøkonomiske status, specificerer og estimerer vi en dynamisk strukturel boligefterspørgselsmodel. Vores model inkluderer flytteomkostninger, fremadskuende adfærd og usikkerhed om udviklingen i lokale nabolagskarakteristika, formue, indtægter, boligpriser og familiesammensætning. Vi estimerer marginale betalingsvilligheder for ikke-handlede nabolagskarakteristika med fokus på luftforurening. Vi tillader husstandes betalingsvillighed at variere med husholdningskarakteristika og hævder, at husholdsninger med lav formue og lav indkomst er lånebegrænsede. Betalingvilligheden for husholdninger, som sandsynligvis er lånebegrænsede, viser sig at være mere følsomme over for ændringer i formuen end for andre husholdninger. Vores estimation fastslår, at den dynamiske tilgang justerer for forskellige skævheder i forhold til en sammenlignelig statisk tilgang.

CHAPTER 1

Explosive Bubbles in House Prices? Evidence from the OECD Countries

A shorter version of this chapter has been accepted for publication in Journal of International Financial Markets, Institutions, and Money.

Tom Engsted
Aarhus University and CREATES
Simon J. Hviid
Aarhus University and CREATES
Thomas Q. Pedersen
Aarhus University and CREATES

Abstract

We conduct an econometric analysis of bubbles in housing markets in the OECD area, using quarterly OECD data for 18 countries from 1970 to 2013. We pay special attention to the explosive nature of bubbles and use econometric methods that explicitly allow for explosiveness. First, we apply the univariate right-tailed unit root test procedure of Phillips et al. (2015) on the individual countries price-rent ratio. Next, we use Engsted and Nielsen's (2012) co-explosive VAR framework to test for bubbles. We find evidence of explosiveness in many housing markets, thus supporting the bubble hypothesis. However, we also find interesting differences in the conclusions across the two test procedures. We attribute these differences to how the two test procedures control for cointegration between house prices and rent.

JEL Classification: C22, C32, G12
Keywords: Co-explosive VAR model, right-tailed unit root tests, date-stamping bubble periods, price-to-rent ratio

1.1 Introduction

Many countries have experienced dramatic movements in house prices over the past 15-20 years, with large increases during the 1990s and first half of the 2000s followed by price drops since 2006-2007. This pattern is also visible in the ratio of prices to rents and has been especially pronounced in countries such as Spain, Ireland, Denmark, Sweden, the Netherlands, the UK, and the US. Understanding these developments is important, not least because the recent international financial crisis to a large extent originated from the housing markets, e.g. the subprime mortgages in the US and the overinvestment in housing in many European countries.

In general, changes in house prices must be due to one of two causes (or a combination of them), either changing 'fundamentals' or speculative bubbles. In the literature, *rents* are usually considered an important part of fundamentals for house prices, see e.g. Hamilton and Schwab (1985), Meese and Wallace (1994), Himmelberg et al. (2005), Gallin (2008), Brunnermeier and Julliard (2008), Campbell et al. (2009), Plazzi et al. (2010), Cochrane (2011), Ghysels et al. (2013), Engsted and Pedersen (2014a, 2015), and Gelain and Lansing (2014). For the owner of a house who also lives in the house, rents can be seen as a proxy for the unobservable housing service flow and thus are the equivalent to the dividends that an owner of a stock obtains in the equity market.

However, the recent boom-bust developments in housing markets have generated a heated discussion of whether speculative bubbles could be a major factor in house price movements in addition to changing fundamentals. During the boom period several observers, most notably Shiller (2005), raised the possibility that a bubble was driving US house prices, while others, e.g. Himmelberg et al. (2005), McCarthy and Peach (2004), and Krainer and Wei (2005), argued that the US housing market was not inflated by a bubble.

After the end of the boom period a few studies have investigated the bubble hypothesis for the US housing market using formal econometric tests. Phillips and Yu (2011) basically use the econometric methods from Phillips et al. (2011), which rely on forward recursive regressions coupled with right-tailed unit root tests, to document explosive behavior in US house prices. Kivedal (2013) uses the co-explosive vector-autoregressive (VAR) methodology from Engsted and Nielsen (2012), and he also finds US house prices to be explosive. Thus, both these recent studies find evidence in support of the bubble hypothesis for the US.

Systematic econometric analyses of explosiveness in house prices outside the

US have been sparse. Exceptions are Yiu et al. (2013), Caspi (2014), and Pavlidis et al. (2014). However, neither of these studies investigate the cointegrating relationship between prices and fundamentals in addition to explosiveness. In this paper we fill this gap in the literature. We conduct a thorough econometric analysis of bubbles in housing markets in the OECD area, using quarterly OECD data for 18 countries from 1970 to 2013. We pay special attention to the explosive nature of bubbles and use econometric methods that explicitly allow for explosiveness. First, we apply the univariate right-tailed unit root test procedure of Phillips et al. (2015), which is a generalization of the test procedure of Phillips et al. (2011), on the individual countries price-rent ratio. Next, we use Engsted and Nielsen's (2012) co-explosive VAR framework to test for bubbles while at the same time allowing prices to be cointegrated with fundamentals and estimating the cointegrating relationship.

The appealing feature of the co-explosive VAR framework of Engsted and Nielsen (2012) is that it allows prices to contain both an explosive component - coming from the bubble - and an $I(1)$ component coming from the non-stationary part of fundamentals, i.e. prices and fundamentals may 'cointegrate' despite the explosive root in prices. This is an important feature of traditional bubble models that is often neglected in empirical bubble studies although emphasized by Diba and Grossman (1988a) and Engsted (2006). The drawback of the co-explosive VAR methodology is that it assumes that the bubble period can be identified á priori; in principle the method does not allow for bursting or partially bursting bubbles during the sample period. Thus, the sample period needs to end before or at the peak of the bubble.

By contrast, the univariate right-tailed unit root test procedure of Phillips et al. (2015) is explicitly designed to capture bursting bubbles and to date-stamp the beginning and end of the bubble. Thus, this procedure can handle a sample period that contains both bubble and non-bubble sub-periods. The main drawback of the procedure is that it does not allow for both an explosive root and a unit root. The null hypothesis underlying the test is that the relevant time series is an $I(1)$ process, while it is explosive under the alternative. Applying the test on the price-rent ratio and rejecting the null hypothesis it is thus implicitly assumed that prices and rent cointegrate, which need not be the case. The co-explosive VAR framework does allow for both an explosive root and a unit root, and hence the two test procedures applied in this paper can potentially result in different conclusions regarding the presence of bubbles.

The main results of our analysis are as follows. First, using the univariate right-tailed unit root test procedure on the price-rent ratio we find evidence of bubbles in 16 of the 18 OECD countries' housing markets. Only in Germany and Italy do we not detect a bubble during the sample period from 1970 to 2013. Second, there appears to be a large degree of housing bubble synchronicity across the OECD countries. Most countries experience bubbles in their housing markets in the early 2000s. The exceptions are Japan, Switzerland, and Finland, who in contrast experienced bubbles around 1990. Third, using the co-explosive VAR framework we obtain results that for four countries (Ireland, Norway, Sweden, and the US) are consistent with the univariate analysis in that we find evidence of both explosiveness and cointegration between prices and rents. For another four countries (Australia, Denmark, Finland, and France) we still find evidence of bubbles but also of no common $I(1)$ trend in the relation between prices and rents. For the remaining countries the co-explosive VAR analysis does not indicate the presence of an explosive housing bubble.

The rest of the paper is organized as follows. In the next section the bubble model is described. Section 1.3 describes the econometric methodologies. That section also contains a comparison of our approach with earlier bubble tests. In section 1.4 we present the empirical results using data from the OECD countries. Section 1.5 concludes.

1.2 The Bubble Model

We start by considering the standard model for asset price determination that is often used in house price studies. Let P_t denote the house price and X_t the service flow, which we proxy by housing rents, both at time t. Given a constant and positive expected one period housing return, $R > 0$,[1] P_t is given as

$$P_t = \frac{1}{1+R} E_t \left(P_{t+1} + X_{t+1} \right), \tag{1.1}$$

where the expectation operator, E_t, is conditioned on information at time t. From (1.1) the general solution for P_t is

$$P_t = \sum_{i=1}^{\infty} \left(\frac{1}{1+R} \right)^i E_t X_{t+i} + B_t, \tag{1.2}$$

where

$$B_t = \frac{1}{1+R} E_t B_{t+1}. \tag{1.3}$$

B_t is the bubble component that reflects self-fulfilling expectations: the bubble is only present at time t if it is expected to be present at time $t+1$. From (1.3) it follows that $B_{t+1} = (1+R)B_t + \xi_{t+1}$, where ξ_{t+1} is a rational forecast error such that $E_t \xi_{t+1} = 0$. Note that the restriction $R > 0$ implies that any rational bubble must have an explosive root in its autoregressive representation.

The bubble component can be eliminated by imposing a transversality condition $\lim_{T \to \infty} (1+R)^{-T} E_t P_{t+T} = 0$ when solving (1.1) recursively forward for P_t. In that

[1] In most of the bubble literature expected returns are assumed to be constant. In section 3.3 we briefly discuss this assumption.

case house prices are determined only by the present value of expected future rents, i.e. the first term on the right-hand side of (1.2). In this paper, however, we do not impose the no-bubble transversality condition.[2]

We can, following Campbell and Shiller (1987), define the 'spread' as $S_t \equiv P_t - \frac{1}{R}X_t$ and rewrite (1.2) into the following equation

$$S_t = \frac{1+R}{R}\sum_{i=1}^{\infty}\left(\frac{1}{1+R}\right)^i E_t\Delta_1 X_{t+i} + B_t, \tag{1.4}$$

where $\Delta_1 X_{t+i} \equiv X_{t+i} - X_{t+i-1}$ is the first-difference of X_{t+i}. Equation (1.4) shows that if there is no bubble ($B_t = 0$) and rents have a unit root, i.e. $X_t \sim I(1)$, then prices, P_t, will share the unit root with X_t such that the spread S_t is stationary, $I(0)$, i.e. P_t and X_t cointegrate with cointegrating vector $(1, -\frac{1}{R})$.

If there is a bubble, $B_t > 0$, then P_t and X_t still share a stochastic $I(1)$ trend such that S_t does not contain a unit root. However, in that case P_t will also contain the explosive root coming from the bubble. Thus, P_t and X_t are still 'cointegrated' in the sense that the linear combination given by S_t has no $I(1)$ component but it is not stationary since it contains the explosive bubble component.

Our VAR analysis in Section 1.4 will be based on the parameterization in (1.4). Following Craine (1993), another useful reparameterization of (1.2) is to divide both sides by X_t to give an expression for the price-rent ratio in terms of expected future compounded growth rates in rents and the bubble-rent ratio:

$$\frac{P_t}{X_t} = \sum_{i=1}^{\infty}\left(\frac{1}{1+R}\right)^i \frac{E_t X_{t+i}}{X_t} + \frac{B_t}{X_t} \tag{1.5}$$

If $B_t = 0$ and P_t and X_t have a common stochastic $I(1)$ trend, the price-rent ratio P_t/X_t will be stationary (cf. Craine, 1993). However, if there is a bubble, $B_t > 0$, the price-rent ratio will be explosive. If there is a bubble and prices and rents do not have a common $I(1)$ trend, i.e. they do not 'cointegrate', P_t/X_t may contain both a unit root and an explosive root. A unit root in P_t that is unrelated to X_t may come from other parts of fundamentals besides rents, e.g. household disposable income. Part of our analysis in Section 1.4 will be based on the expression in (1.5).

In theory the above model setup does not imply riskless arbitrage opportunities despite the bubble. Since the bubble evolves according to (1.3), it is consistent with the no-arbitrage relation (1.1), cf. Diba and Grossman (1988b). Thus, despite the presence of a bubble, the model describes an informationally efficient market with positive and constant expected returns.

From (1.1) we can define,

$$M_t \equiv P_t + X_t - (1+R)P_{t-1} \tag{1.6}$$

[2] In principle, the bubble model implies that there is no predictability of returns.

where $E_t M_t = 0$, i.e. M_t is a martingale difference sequence, that is one-period returns in excess of a constant are unpredictable given past observations, i.e. the classical efficient markets hypothesis as described by Leroy (1989). We can reformulate M_t in the spread as

$$M_t = \Delta_{1+R} S_t + (1 + R^{-1})\Delta_1 X_t \tag{1.7}$$

where $\Delta_{1+R} \equiv 1 - (1 + R)L$, thus for M_t to be a martingale difference it is allowed for the spread to be explosive and for the service flow to be non-stationary.

However, it is well-documented that housing markets are not informationally efficient, see e.g. Case and Shiller (1989), so in our empirical analysis we will not test the strict efficient markets implications of the model. Instead, the focus will be on the cointegration and explosive properties that follow from (1.4) and (1.5). For tests of the full set of implications of the bubble model, see Engsted and Nielsen (2012)[3].

1.3 The Econometric Methodology

Right-tailed Unit Root Tests and Date-Stamping

As noted in Section 1.2 a key issue in identifying rational bubbles is whether we can detect explosive behavior in the spread $P_t - \frac{1}{R} X_t$ and the price-rent ratio P_t / X_t. With the aim of identifying potential rational bubbles in the stock market, Diba and Grossman (1988a) suggest the use of right-tailed unit root tests to detect explosiveness, i.e. instead of testing the null of a unit root against stationarity one should test against the explosive alternative. However, using a simulation study Evans (1991) shows that this test procedure has very low power in detecting periodically collapsing bubbles.

Motivated by the idea of Diba and Grossman (1988a) and the findings of Evans (1991), Phillips et al. (2011) propose to conduct a series of right-tailed unit root tests based on an expanding window (with a fixed start date) of data with the largest of these test statistics being the relevant test statistic for explosiveness. They name this test the supremum augmented Dickey-Fuller (SADF) test and show that it has much better power properties in the presence of periodically collapsing bubbles than a unit root test based on the entire sample as proposed by Diba and Grossman (1988a). Phillips et al. (2015) generalize the test procedure by allowing both the start and end date to vary and find that this approach has even higher power than the SADF test in detecting periodically collapsing bubbles. An important feature of both the SADF and generalized SADF (GSADF) test procedure is that it enables us to date-stamp periods with explosive behavior. We will make use of GSADF and the date-stamping

[3] A prerequisite for testing the efficient market hypothesis in this model is that the restrictions listed below are not rejected. In our application we do not identify any bubble period in which we do not reject one of these restrictions and, hence, tests for the full set of implications of the efficient market hypothesis are omitted from the paper.

procedure proposed by Phillips et al. (2015) to test for explosiveness in the price-rent ratio and to pinpoint periods with explosive behavior.

The GSADF Test

The null hypothesis in the GSADF test procedure is that the relevant time series, y_t, follows a random walk process with an asymptotically negligible drift

$$y_t = dT^{-\eta} + \theta y_{t-1} + \varepsilon_t, \quad \varepsilon_t \overset{iid}{\sim} N\left(0, \sigma^2\right), \quad \theta = 1, \tag{1.8}$$

where d is a constant and $\eta > 1/2$ is a coefficient that controls the magnitude of the drift as the sample size, T, goes to infinity.[4] Before explaining the GSADF test procedure it is useful to introduce some notation. Let r_1 and r_2 denote fractions of the total sample with $r_2 = r_1 + r_w$, where $r_w > 0$ is the fractional window size used in the auxiliary regressions underlying the test procedure. This implies that in the regressions we will use $T_w = \lfloor Tr_w \rfloor$ observations, where $\lfloor \cdot \rfloor$ denotes the integer part of the argument. Also, let r_0 denote the smallest fractional window size used.

Based on the data-generating process under the null hypothesis (1.8), Phillips et al. (2015) use the following auxiliary regression when performing the unit root tests

$$y_t = \mu_{r_1,r_2} + \delta_{r_1,r_2} y_{t-1} + \sum_{j=1}^{k} \phi_{j,r_1,r_2} \Delta y_{t-j} + \varepsilon_t, \quad \varepsilon_t \overset{iid}{\sim} N\left(0, \sigma^2_{r_1,r_2}\right), \tag{1.9}$$

where the subscript r_1, r_2 is used to illustrate that (1.9) is estimated using a sample that begins with observation $\lfloor Tr_1 \rfloor$ and ends with $\lfloor Tr_2 \rfloor$. Based on the auxiliary regression (1.9) we can test the null hypothesis of a unit root ($\delta_{r_1,r_2} = 1$) against the explosive alternative ($\delta_{r_1,r_2} > 1$). We denote the corresponding ADF test statistic by $ADF_{r_1}^{r_2}$.

The GSADF test statistic, which is a function of r_0, is defined as

$$GSADF\left(r_0\right) = \sup_{\substack{r_2 \in [r_0, 1] \\ r_1 \in [0, r_2 - r_0]}} \left\{ ADF_{r_1}^{r_2} \right\}.$$

The ending point of the sample sequence varies from $\lfloor Tr_0 \rfloor$ to T, while the starting point varies from the first observation to $\left\lfloor T\left(r_2 - r_0\right) \right\rfloor$. The GSADF test is the supremum of the corresponding sequence of ADF test statistics. The SADF test statistic is a special case of the GSADF test where $r_1 = 0$. GSADF and SADF both have nonstandard limiting distributions, so critical values are obtained by means of simulation.

[4] Phillips et al. (2015) set d and η to 1, while Phillips et al. (2011) effectively set $\eta \to \infty$, corresponding to a null hypothesis of a random walk without drift.

Date-Stamping

Phillips et al. (2015) propose a backwards supremum ADF (BSADF) test procedure to pinpoint periods with explosive behavior. The BSADF test statistic, which is also a function of r_0, is defined as

$$BSADF_{r_2}\left(r_0\right) = \sup_{r_1 \in [0, r_2 - r_0]} \left\{BADF_{r_1}^{r_2}\right\},$$

where $BADF_{r_1}^{r_2}$ denotes the ADF test statistic using a sample with starting point $\lfloor Tr_1 \rfloor$ and ending point $\lfloor Tr_2 \rfloor$. For each value of r_2 we thus get a sequence of BADF statistics, and the BSADF test statistic is the supremum of these. Based on the series of BSADF test statistics, Phillips et al. (2015) define the (fractional) origination and termination points of explosive behavior ($\widehat{r}_e$ and $\widehat{r}_f$) as

$$\widehat{r}_e = \inf_{r_2 \in [r_0, 1]} \left\{r_2 : BSADF_{r_2}\left(r_0\right) > cv_{r_2}^{\alpha_T}\right\},$$

$$\widehat{r}_f = \inf_{r_2 \in [\widehat{r}_e, 1]} \left\{r_2 : BSADF_{r_2}\left(r_0\right) < cv_{r_2}^{\alpha_T}\right\},$$

where $cv_{r_2}^{\alpha_T}$ denotes the $100\left(1-\alpha_T\right)\%$ critical value of the SADF test statistic based on $\lfloor Tr_2 \rfloor$ observations. This identification scheme implies that the origination date of the bubble $\lfloor T\widehat{r}_e \rfloor$ is the first chronological observation where the BSADF test statistic exceeds the critical value, and correspondingly the termination date $\left\lfloor T\widehat{r}_f \right\rfloor$ is the first chronological observation after $\lfloor T\widehat{r}_e \rfloor$ where the BSADF test statistic falls below the critical value.

As mentioned in Section 1.2 a rational bubble, B_t, reflects self-fulfilling expectations meaning that the bubble is only present at time t if it is expected to be present at time $t+1$. In addition, given rationality it must hold that $E_t \xi_{t+1} = 0$ in $B_{t+1} = (1+R)B_t + \xi_{t+1}$. As a consequence, a rational bubble cannot burst and reappear but must have been present in the asset from the first time it was traded, cf. Diba and Grossman (1988b). This property appears to conflict with the idea of date-stamping bubble periods. However, rational bubbles can be periodically collapsing (see e.g. Evans, 1991) which implies that bubbles can burst *partially* and then begin to build up again. If the bubble component is small compared to fundamentals it will be hard to detect using econometric techniques meaning that the bubble has to be of a certain size before one can identify it. Date-stamping the presence of rational bubbles can thus be understood as identifying periods, where the bubble component has a significant size compared to fundamentals.

The Co-Explosive VAR Methodology

The bubble testing procedure proposed by Engsted and Nielsen (2012) builds on the co-explosive framework developed in Nielsen (2010). Consider the vector $V_t =$

$(P_t, X_t)'$ and assume that it follows the k'th order vector autoregression (VAR)

$$M: \qquad V_t = \mu_v + \sum_{j=1}^{k} A_j V_{t-j} + \epsilon_t, \tag{1.10}$$

where $A_j \in \mathbb{R}^{2\times 2}$ are the VAR parameter matrices, $\mu_v \in \mathbb{R}^2$ is the vector of deterministic terms, and ϵ_t is the vector of error terms. In accordance with Engsted and Nielsen (2012) we call this model M. If the characteristic polynomial for the autoregressive structure of (1.10) has an explosive root, $\rho > 1$, the VAR model can be reparameterized in vector error correction (VECM) form as

$$\Delta_1 \Delta_\rho V_t = \mu_v + \Pi_1 \Delta_\rho V_{t-1} + \Pi_\rho \Delta_1 V_{t-1} + \sum_{j=1}^{k-2} \Phi_j \Delta_1 \Delta_\rho V_{t-j} + \epsilon_t,$$

where $\Delta_\rho = (1 - \rho L)$, L is the lag-operator, and $\Pi_1, \Pi_\rho, \Phi_j \in \mathbb{R}^{2\times 2}$. Note that here V_t is assumed to have a unit root in addition to the explosive root.

The VECM representation is linked to the VAR representation through the standard error correction form that follows from Granger's Representation Theorem,

$$\Delta_1 V_t = \mu_v + \Pi V_{t-1} + \sum_{j=1}^{k-1} \Gamma_j \Delta_1 V_{t-j} + \epsilon_t,$$

and the following set of identities

$$\Pi_1 = \frac{\Pi}{1-\rho}, \quad \Pi_\rho = -\rho \left(I_p + \Pi_1 + \sum_{j=1}^{k-1} \rho^{-j} \sum_{l=j+1}^{k} A_l \right), \quad \Phi_j = -\sum_{l=j+1}^{k-1} \rho^{j-l} \sum_{i=l+1}^{k} A_i,$$

where $\Pi, \Gamma_j \in \mathbb{R}^{2\times 2}$.

In addition to the explosive root, V_t has either one or two unit roots. If there is one unit root it implies that P_t and X_t have a common stochastic $I(1)$ trend, i.e. they are 'cointegrated'. If P_t and X_t are not cointegrated, there are two independent unit roots in the system. In the former case we have the following reduced rank restrictions $\Pi_1 = \alpha_1 \beta_1'$ and $\Pi_\rho = \alpha_\rho \beta_\rho'$, where $\alpha_1, \beta_1, \alpha_\rho, \beta_\rho \in \mathbb{R}^2$. Here β_1 has the common interpretation of a cointegrating relation of V_t, equivalent to the one identified by the spread, $S_t = P_t - \frac{1}{R} X_t$, in Section 1.2. Therefore we have the model implied restriction $\beta_1 = (1, -1/R)'$. Furthermore, β_ρ is interpreted as a co-explosive vector that captures the common explosive component in P_t and X_t. The case that we will be particularly interested in is when rents are non-explosive but prices are explosive due to a bubble. This case implies the following restriction on the co-explosive vector: $\beta_\rho = (0, 1)'$.

We now explain in more detail these restrictions and how to test them.

The Cointegration Restriction

The starting point in the estimation procedure is to estimate the unrestricted model M, equation (1.10). Next we compute the characteristic roots from the characteristic polynomial and if the largest root is larger than unity, $\hat{\rho} > 1$, this serves as preliminary evidence of explosiveness. If the largest root is equal to or less than one, the system is just a standard $I(1)$ or stationary system. Hereafter, we apply the sequential cointegrating rank test of Johansen (1995). This corresponds to the null hypothesis

$$H_0^1: \qquad (\Pi_1, \mu_v) = \alpha_1(\beta_1', \zeta_1), \quad \Pi_\rho = \alpha_\rho \beta_\rho', \tag{1.11}$$

where $\zeta_1 \in \mathbb{R}$ is a constant that is restricted to the cointegrating space. The null hypothesis implies a cointegrating relation between P_t and X_t. The Johansen rank test is valid even in the presence of an explosive root, cf. Nielsen (2010). For $\hat{\rho} > 1$, if we use M_1 to denote M when restricted by H_0^1, then

$$M_1: \qquad \Delta_1\Delta_\rho V_t = \mu_v + \alpha_1\beta_1'\Delta_\rho V_{t-1} + \alpha_\rho\beta_\rho'\Delta_1 V_{t-1} + \sum_{j=1}^{k-2}\Phi_j\Delta_1\Delta_\rho V_{t-j} + \epsilon_t, \tag{1.12}$$

and we calculate the updated largest characteristic root, $\hat{\rho}_1$. If this root is found to be strictly lager than unity we proceed through the subsequent steps of the estimation procedure. If the largest characteristic root is less than one, the system again reduces to a standard cointegrated $I(1)$ system. In any case, $\widehat{\beta}_1'$ gives the estimated cointegrating vector between P_t and X_t.

Non-Explosiveness of Rents

Under the assumption of H_0^1 and $\rho_1 > 1$ we test the null hypothesis of non-explosive rents, that is imposing the restriction

$$H_0^X: \qquad \beta_\rho = (0,1)'. \tag{1.13}$$

This additional hypothesis restricts the model to

$$M_{1X}: \qquad \Delta_1\Delta_\rho V_t = \mu_v + \alpha_1\beta_1'\Delta_\rho V_{t-1} + \alpha_\rho\Delta_1 X_{t-1} + \sum_{j=1}^{k-2}\Phi_j\Delta_1\Delta_\rho V_{t-j} + \epsilon_t. \tag{1.14}$$

The likelihood of this model is determined by a numerical profile argument. Specifically, we apply a grid search over a range of values for ρ. The likelihoods are then determined by reduced rank regressions of $\Delta_1\Delta_\rho V_t$ on $\Delta_\rho V_{t-1}$, where we correct for lagged rent growth, $\Delta_1 X_{t-1}$, and differences, $\Delta_1\Delta_\rho V_{t-j}$. Maximizing the likelihood gives $\hat{\rho}_{1X}$. The likelihood-ratio test of H_0^X under the model M_1 is asymptotically $\chi^2(k-1)$, cf. Engsted and Nielsen (2012). If H_0^X is rejected, X_t cannot be taken to be non-explosive. The interpretation of an explosive root in the system as due to a rational bubble hinges on the non-rejection of H_0^X.

The Spread Restriction

Suppose that H_0^1 in (1.11) and H_0^X in (1.13) are both not rejected. In this case we proceed by testing the restriction on the spread

$$H_0^S: \qquad \beta_1 = (1, -1/R)', \quad \text{where } \rho = 1 + R. \tag{1.15}$$

The restriction $\rho = 1 + R$ comes from equation (1.3) in Section 1.2. The restriction (1.15) implies that the model takes the form

$$M_{1XS}: \qquad \Delta_1 \Delta_\rho V_t = \mu_v + \alpha_1 \Delta_\rho S_{t-1} + \alpha_\rho \Delta_1 X_{t-1} + \sum_{j=1}^{k-2} \Phi_j \Delta_1 \Delta_\rho V_{t-j} + \epsilon_t, \tag{1.16}$$

where the spread is inserted for $S_t = P_t - \frac{1}{\rho-1} X_t$. Maximization of the likelihood is done along the same numerical lines as in the previous step whereby we obtain $\hat{\rho}_{1XS}$. The likelihood-ratio test statistic of H_0^S in M_{1X} is asymptotically $\chi^2(1)$ and the simultaneous test of H_0^X and H_0^S in M_1 is asymptotically $\chi^2(k)$, cf. Engsted and Nielsen (2012). If H_0^S is rejected, the procedure does not provide a unique estimate of the expected return, R, since the estimates provided by the explosive root $\widehat{\rho}_{1X}$, and the cointegrating vector $\widehat{\beta}_1$, respectively, are statistically different. However, rejection of H_0^S does not affect the conclusion that P_t contains an explosive component.

Visual Inspection

An advantage of the Engsted and Nielsen (2012) procedure is that it explicitly allows for an explosive component along with cointegration between prices and fundamentals. This approach enables us to plot the two components of the series in which both a stochastic and an explosive trend are present. The explosive component can be calculated from the formula

$$\hat{P}^e = (\rho - 1)^{-1}(1, 0) C_\rho \rho^t \left(A_\rho + \sum_{s=1}^{t} \rho^{-s} \varepsilon_s \right),$$

and the stochastic component is calculated as

$$\hat{P}^s = (1 - \rho)^{-1}(1, 0) C_1 \left(A_1 + \sum_{s=1}^{t} \varepsilon_s \right).$$

Here C_x and A_x are defined in Theorem 2 of Engsted and Nielsen (2012) for $(x, y) = (1, \rho)$ and $(x, y) = (\rho, 1)$ as

$$\Psi_x = I_2 + \frac{\alpha_y \beta_y'}{y-1} - \sum_{j-1}^{k-2} x^{-j} \Phi_j,$$

$$C_x = b\beta_{x\perp} \left(\alpha_{x\perp}' \Psi_x \beta_{x\perp}\right)^{-1} \alpha_{x\perp}',$$

$$A_x = \Psi_x \Delta_y V_0 - \frac{\alpha_y \beta_y'}{y-x} \Delta_x X_V + \sum_{j=1}^{k-2} \Psi_j \sum_{h=0}^{j-1} \Delta_x \Delta_y V_{-h}.$$

Comparison with earlier Bubble Tests

West (1987) developed an often cited specification test for rational stochastic bubbles. The test compares two sets of estimates of the underlying asset pricing model. The first set of estimates is consistent both with and without a bubble, while the second set is only consistent in the absence of a bubble. Equality of the two sets of estimates is then tested using a Hausman (1978) type specification test. The null hypothesis is no bubble, while the presence of a bubble should lead to rejection of the hypothesis. A problem with this procedure (noted by West himself in West, 1985) is that the test is not consistent. Under the alternative hypothesis that a bubble is present, the probability that the test will reject the null does not go to unity asymptotically. This is a direct consequence of the explosiveness of prices under the alternative. The Engsted and Nielsen (2012) procedure that we apply in this paper does not face this problem because in this procedure the null hypothesis explicitly involves a bubble.

Diba and Grossman (1988a) proposed to test for rational bubbles by using Bhargava's (1986) von Neumann-like statistic to test the null hypothesis of a unit root in prices against the explosive alternative. They also tested for cointegration between prices and fundamentals arguing that with a constant discount factor cointegration precludes bubbles while no cointegration would be consistent with the presence of a bubble. By using Bhargava's (1986) test for explosiveness the Diba and Grossman methodology assumes that the variables are at most a first-order autoregressive process, and the discount factor cannot be estimated but must be specified a priori. The Engsted and Nielsen (2012) procedure extends Diba and Grossman's procedure by specifying a general VAR for the variables that allows for an explosive root in addition to a possible common stochastic $I(1)$ trend (i.e. cointegration) between prices and fundamentals. In addition, the procedure allows estimation of the discount factor instead of prefixing it a priori as in the Diba-Grossman procedure.

Using a linear VAR for prices and fundamentals requires that the discount factor is constant. Most previous bubble studies in fact assume that the discount factor is constant. In the empirical finance literature this assumption is controversial since returns are often found to be predictable (see e.g. Cochrane, 2008). However, Engsted et al. (2012) show that a rational bubble may make returns appear predictable even when expected returns (and thereby the discount factor) is constant. In addition, even if expected returns are time-varying, Craine (1993) and Timmermann (1995) show that unless expected returns are highly persistent, the cointegrating relationship between prices and fundamentals implied by the constant discount factor present value model will hold approximately when the discount factor is time-varying.[5]

Evans (1991) showed in a simulation study that in a finite sample unit root and cointegration based tests will often not identify the explosive component of periodically collapsing rational bubbles (see also Hall et al., 1999). Thus, the Engsted and Nielsen (2012) framework may not work well in that situation. By contrast, the Phillips et al. (2015) recursive procedure is explicitly designed to account for the periodically collapsing nature of the Evans type bubbles. However, the Phillips et al. (2015) procedure does not allow for estimation of the cointegrating relationship between prices and fundamentals because it does not allow both a unit root and an explosive root. If prices and fundamentals are cointegrated, i.e. have a common $I(1)$ trend, then, as we noted in Section 1.2, a solution would be to use instead of prices the ratio of prices to fundamentals. That eliminates the unit root. However, in contrast to the Engsted and Nielsen (2012) approach, the Phillips et al. (2015) methodology does not allow testing for a common $I(1)$ trend between prices and fundamentals. Furthermore, if prices have a unit root in addition to the explosive root, and this unit root is not shared with the unit root in fundamentals, the price-fundamental ratio will still contain both a unit root and an explosive root. It is not clear what the properties of the Phillips et al. procedure are in this case.

In fact, based on Diba and Grossman (1988a) many earlier empirical bubble studies have claimed that cointegration between prices and fundamentals rules out bubbles. For example, Phillips et al. (2011, p.206) state: "In the presence of bubbles, p_t [price] is always explosive and hence cannot co-move or be cointegrated with d_t [fundamental] if d_t is itself not explosive. Therefore, an empirical finding of cointegration between p_t and d_t may be taken as evidence against the presence of bubbles." This statement is at best incomplete. As shown above the stochastic $I(1)$ trend in d_t will be part of p_t also in the presence of an explosive bubble, and the multivariate cointegrated VAR analysis based on reduced rank regressions will capture this feature. However, it is true that univariate regression based cointegration analysis will not be able to estimate the common $I(1)$ trend in p_t and d_t if p_t also involves an explosive trend because the regression residuals will always be non-stationary in that case.

[5] Craine (1993) shows that with a time-varying (but stationary) discount factor the *ratio* between prices and fundamentals will be stationary under no bubbles. Thus, testing for explosiveness of this ratio is robust to the assumption about the discount factor. However, no other testable restrictions follow from Craine's approach.

1.4 Empirical Results

The Data

We use official OECD data for 18 countries:[6] Australia, Belgium, Canada, Switzerland, Germany, Denmark, Spain, Finland, France, the UK, Ireland, Italy, Japan, the Netherlands, Norway, New Zealand, Sweden, and the US. The dataset provided by the OECD contains quarterly observations from 1970Q1 to 2013Q4 for all countries, except Australia (begins 1972Q3), Belgium (begins 1976Q2), Spain (begins 1971Q1), Norway (begins 1979Q1), and Sweden (begins 1980Q1). The dataset contains seasonally adjusted real house prices and the price-rent ratio, which allows us to back out a real rent series for each country.[7] The price indices and the price-rent ratios from the OECD are indexed with 2010 as base year. This indexation does not present a problem for the unit root tests for explosiveness of the price-rent ratio and most of the testable restrictions in the co-explosive VAR framework. The exception is the spread restriction (1.15), which requires that the level of rents is linked to prices such that we can calculate returns based on these two series. To facilitate this link we fix the price-rent ratios in 2013Q4 to the actual price-rent levels in 2013 reported by Numbeo.[8] From the actual price-rent ratios and rent growth obtained from the backed out rent series, we obtain new rent series, which are directly linked to prices.

Our right-tailed unit root tests and date-stamping procedure will be based on the price-rent ratio. Figure 1.2 shows the (indexed) price-rent ratios. For most countries we observe a relatively stable ratio until the mid 1990s after which large increases in the ratio suggest that house prices started to decouple from rents. A group of countries (Denmark, Spain, Ireland, the Netherlands, and the US) has since the mid 2000s experienced noticeable decreases in the price-rent ratio caused by large drops in house prices, while for others (Australia, Belgium, Canada, Finland, France, the UK, Norway, New Zealand, and Sweden) the ratio has either stabilized or continued to increase. A few countries have experienced movements in the price-rent ratio different from this overall pattern. In Germany the ratio has generally been downward trending, while it has been relatively stable in Italy. Switzerland and Japan experienced large increases in house prices up to around 1990 after which prices fell dramatically in both countries. In recent years house prices have recovered somewhat in Switzerland, while prices continue to fall in Japan. Finland experienced price movements around 1990 similar to Switzerland and Japan, but since the mid 1990s prices have here generally been upward trending.

---

[6] These data have been used in earlier analyses by OECD, e.g. Girouard et al. (2006).

[7] Ideally, we would like to use prices and rents from matched properties; however, such data are not readily available. If this cause a systematic difference, then the effect is limited to affect the test of the spread restriction in the Engsted and Nielsen (2012) approach.

[8] http://www.numbeo.com/property-investment/rankings_by_country.jsp.

Right-tailed Unit Root Tests and Date-Stamping

Table 1.1 shows the empirical results when performing the GSADF test on the price-rent ratio. In implementing the test we set the minimum window size $\lfloor Tr_0 \rfloor$ to 40 for each country and the lag length in the ADF regressions to 1. To obtain critical values we perform 5,000 simulations for each sample size. According to the GSADF test the price-rent ratio has displayed explosive behavior in our sample period even at a 1% significance level in 15 of the 18 countries. In Norway the evidence of explosiveness in the price-rent ratio is less strong, but using a 10% significance level we still reject the null of a random walk against the explosive alternative. Only for two countries, Germany and Italy, we cannot reject the null hypothesis at any conventional significance level. The results from the GSADF test thus reveal evidence of bubbles in all of the 18 OECD countries' housing markets, except Germany and Italy.

In condensed form Figure 1.1 shows the results of the date-stamp procedure. The grey areas denote periods with explosiveness in the price-rent ratio as determined by the BSADF test. To be consistent with the GSADF test, we set the minimum window size to 40, the lag length in the ADF regressions to 1, and perform 5,000 simulations at each point in time to obtain critical values. We exclude Germany and Italy from the date-stamping analysis since the GSADF test does not indicate explosiveness in these two countries' price-rent ratio. Abstracting from very short-lived explosive periods there appears to be a large degree of bubble synchronicity across the OECD countries. In the early 2000s we find evidence of explosiveness in 13 countries. For most of these countries the explosive behavior terminates around 2006 and 2007. For Belgium and Canada do we in 2013 still reject the null hypothesis of a random walk against the explosive alternative. The three countries for which there is no evidence of explosiveness in the price-rent ratio in the early 2000s are Japan, Finland, and Switzerland. Interestingly these countries all experience an explosive price-rent ratio around 1990. For Switzerland this is again the case in the last three years of our sample period.

When we apply the SADF and GSADF procedure we need to choose parameter values of r_0 and k. We set $\lfloor r_0 \rfloor = 40$ in line with simulations in Phillips et al. (2015) with sample sizes close to ours. In this way we strive a balance between allowing for multiple bubbles but at the same time minimizing the impact of short term blips in the prices-rent series. We set the autoregressive lag length to be fixed at $k = 1$. The simulation study in Phillips et al. (2015) finds that the size of the SADF and GSADF tests is better for fixed lag lengths than when compared to rules implied by information criteria or by significance tests on the lags. When comparing the size of the test for different lag lengths it is found that shorter fixed lag lengths imply better size, closer to the nominal significance level - especially for the SADF test which has the same properties as the BSADF test statistic used in the data-stamping procedure. Therefore, we set $k = 1$ as we consider this a conservative estimate when we use the BSADF test statistic to pin down periods that exert explosive behaviour. In Pavlidis et al. (2014) the methodology of Phillips et al. is applied to test for house price exuberance in a range of countries which is partially overlapping with our data set[9] and fix the autoregressive lag length $k = 4$. Pavlidis et al. (2014) primarily investigate the price-income ratio, whereas our application focus is on the price-rent ratio, and we find interesting differences. As in this paper, Pavlidis et al. (2014) identify the same synchronicity in explosiveness up to 2007. However, periods of explosiveness are identified over longer periods. This difference can be a result of the less conservative choice of $k = 4$, however, the main difference probably stems from rents co-moving more with house prices than income levels do.[10]

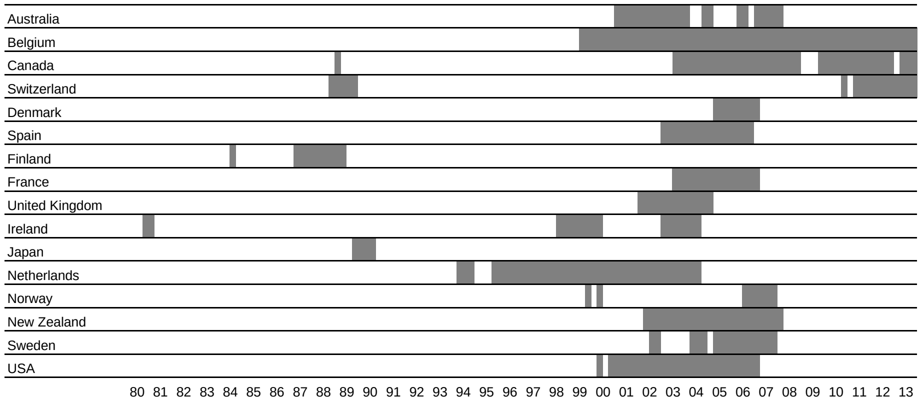
Table 1.1. Right-tailed unit root tests on the price-rent ratio.

Country	T	**GSADF**	**CV** 90%	95%	99%
Australia	166	5.545	1.515	1.7860	2.3364
Belgium	151	7.279	1.438	1.7208	2.2651
Canada	176	4.944	1.545	1.8199	2.3261
Switzerland	176	3.572	1.545	1.8199	2.3261
Germany	176	1.160	1.545	1.8199	2.3261
Denmark	176	2.558	1.545	1.8199	2.3261
Spain	172	2.915	1.537	1.8225	2.3089
Finland	176	3.782	1.545	1.8199	2.3261
France	176	4.103	1.545	1.8199	2.3261
United Kingdom	176	2.670	1.545	1.8199	2.3261
Ireland	176	4.658	1.545	1.8199	2.3261
Italy	176	0.260	1.545	1.8199	2.3261
Japan	176	2.412	1.545	1.8199	2.3261
Netherlands	176	6.666	1.545	1.8199	2.3261
Norway	140	1.646	1.404	1.6910	2.2191
New Zealand	176	4.388	1.545	1.8199	2.3261
Sweden	136	2.818	1.396	1.6650	2.2097
United States	176	5.580	1.545	1.8199	2.3261

Note: In calulating the GSADF test we set the minimum window size to 40 observations and the lag length to 1 in the auxiliary regressions. CV denotes the finite-sample critical values based on 5000 simulations.

[9] Pavlidis et al. (2014) investigate Croatia, Luxembourg, South Africa, and South Korea in addition to the countries covered in this paper.

[10] Our results are reasonably robust to the choice of fixed lag lenght, k, but setting $k = 1$ is slightly more conservative than longer lag lengths.



Note: Explosiveness is identified using the BSADF test with a minimum window size to 40 observations, a lag length to 1 in the auxiliary regressions, and finite-sample critical values based on 5000 simulations.

Figure 1.1. Date-stamping explosiveness in the price-rent ratio.

The Co-Explosive VAR Analysis

The main drawback of using univariate right-tailed unit root tests to detect bubbles, as in the previous section, is that the tests do not allow for both an explosive root and a unit root. Under the null hypothesis the time series process has a unit root, while under the alternative it has an explosive root. As mentioned in Section 1.2 if there is a bubble and prices and rents do not have a common $I(1)$ trend, i.e. they do not 'cointegrate', the price-rent ratio may contain both a unit root and an explosive root. This scenario is not a part of the hypotheses in the univariate test procedure. The conclusion that housing markets in most OECD countries have been subject to bubbles is thus based on the implicit assumption that prices and rents cointegrate. In this section we will use a co-explosive VAR framework to test for bubbles. This test procedure allows for both an explosive root and a unit root and thus accommodates the main drawback of the univariate right-tailed unit root tests. It does however have its own drawback in not allowing for bursting or partially bursting bubbles during the sample period. We address this issue by fixing the end of our sample period in the VAR analysis to the last occurence of explosive behavior in the price-rent ratio as documented in Figure 1.1. We highlight these 'bubble collapses' in Figure 1.2.

Table 1.2 reports the results of the co-explosive VAR analysis for those countries where the univariate GSADF test in Table 1.1 indicates explosiveness in the price-rent ratio, i.e. all countries except Germany and Italy. The lag-length, k, in the VAR analyses is chosen to make the errors serially uncorrelated.

As seen, the unrestricted largest characteristic root, $\widehat{\rho}$, is larger than one for most of the countries, in accordance with the univariate analysis. However, for Belgium, the Netherlands, and New Zealand, $\hat{\rho}$ is slightly less than one, which indicates no explosiveness in these three countries. Due to the downward finite-sample bias of $\hat{\rho}$ in VAR models (cf. Engsted and Pedersen, 2014b) the true ρ may in fact be larger than one also for these countries, in accordance with the GSADF procedure. For Belgium and Netherlands the cointegration test for a rank of $r = 0$ rejects while $r = 1$ cannot be rejected at a 5% level, and the implied values for the expected quarterly real housing return from the estimated cointegrating vectors are $\widehat{R} = 1/188.69 = 0.0053$ (0.53%) and $\widehat{R} = 1/32.59 = 0.0307$ (3.07%), respectively, for Belgium and the Netherlands. For New Zealand the rank test does not indicate cointegration between prices and rents.

In five of the countries (Australia, Denmark, Finland, France, and New Zealand) the estimated VAR system has an explosive root and there is no evidence of cointegration between prices and rents. This suggest the presence of an explosive housing bubble in these countries and that prices have no low-frequency relation to fundamentals measured by rents.

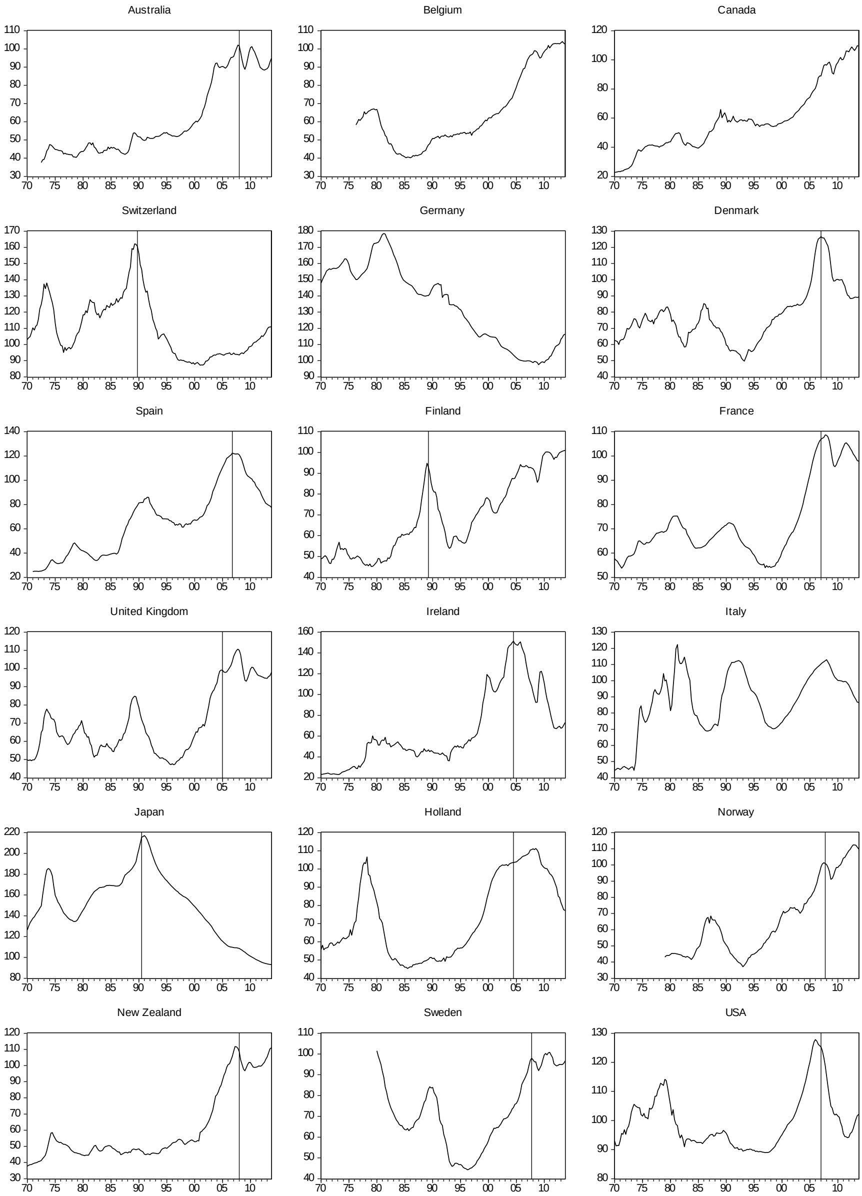
For another group of countries (Canada, Spain, the UK, and also Switzerland and Japan for the early part of the sample) there is evidence of a common $I(1)$ trend, i.e. cointegration, between prices and rents with estimated expected quarterly real returns ranging from 0.19% (Canada) to 3.88% (the UK). However, in these cases the explosive root disappears when imposing a cointegrating rank of $r = 1$. Thus, either the $\widehat{\rho}$ estimates in Table 1.2 are indistinguishable from one (such that the VAR is a standard cointegrated $I(1)$ system) or there is a common explosive root between rents and prices in these five countries. A possible explanation for this result is a constant long-run growth rate in rents which will show up as an exponential increase in the level of rents and thereby imply an explosive root in the VAR system.

For the remaining four countries (Ireland, Norway, Sweden, and the US) there is evidence of both a common $I(1)$ trend between prices and rents and an explosive root. As seen from Table 1.2, the largest characteristic root remains larger than one when imposing a cointegrating rank of $r = 1$ ($\widehat{\rho}_1 > 1$). For Ireland, Norway and Sweden, the H_0^X hypothesis, $\beta_\rho = (0,1)'$, is not rejected at a 5% level implying no explosive rents. Thus, for these three countries there is evidence of an explosive bubble in house prices but also that there is a low-frequency relation between prices and rents. For the US, however, the hypothesis of non-explosive rents is statistically rejected. This result contrasts with the results by Kivedal (2013) who on US data - and also using the Engsted and Nielsen (2012) approach - does not reject the hypothesis of non-explosive rents. In all four countries the largest characteristic root remains larger than one when imposing the spread restriction, H_0^S: $\beta_1 = (1, -1/R)$ with $\rho = 1 + R$, ($\widehat{\rho}_{1X} > 1$; $\widehat{\rho}_{1XS} > 1$). However, the restriction is statistically rejected in all four cases. As seen from the estimated β_1' vector, the expected quarterly real housing return, R, is negative which contrasts with the estimate of R from the explosive root which gives $\widehat{R}$ values ranging from 0.1% (the US) to 4.5% (Norway). These large differences in the estimates of R from the cointegrating vector β_1' on the one hand and the explosive root, ρ, on the other hand probably explain why we reject the spread restriction.

The analysis in this section has documented some interesting cross-country differences in the dynamics of house prices and rents in the OECD area, and also some important differences to the analysis using the right-tailed test procedure. For four of the countries the VAR results are consistent with the univariate analysis on the price-rent ratio in that we find both a common $I(1)$ trend between prices and rents and an explosive root. However, for the remaining 12 countries the VAR results are qualitatively different from the univariate results. In particular, in several cases the explosive root is eliminated when imposing the cointegration restriction between prices and rents. Also, in a number of cases there is an explosive root but no common $I(1)$ trend, which in theory invalidates the GSADF test procedure applied to the price-rent ratio, since this ratio contains a unit root in addition to an explosive root.

In general, for 8 of the 18 countries in our sample, the results in this section lend support to the common claim that a housing bubble has been a main driver of house prices throughout the OECD area, while for the remaining 10 countries the co-explosive VAR analysis does not indicate the presence of an explosive housing bubble. The country specific differences in country differences in part be explained by institutional differences. [11]

[11] Campbell (2013) provides an overview of such differences in mortgage markets for a broad selection of countries.



Note: The vertical lines indicate the end of a bubble period identified using the BSADF test with minimum window size of 40 observations, lag length of 1 in the auxiliary regressions, and finite-sample critical values obtained using 5000 simulations.

Figure 1.2. Price-rent ratio for the 18 OECD countries.

Table 1.2. Co-explosive VAR analysis for prices and rents.

Sample	k	$\hat{\rho}$	**Cointegration** H_0: r=0	H_0: r=1	β_1'	$\hat{\rho}_1$	$\hat{\rho}_{1X}$	$\hat{\rho}_{1XS}$	**LR**$(M_{1X}\|M_1)$ χ^2(k-1)	**LR**$(M_{1XS}\|M_{1X})$ χ^2(1)	**LR**$(M_{1XS}\|M_1)$ χ^2(k)
Australia											
1973Q1 - 2008Q1	6	1.018	[0.386]	[-]	(-,-)	-	-	-	-	-	-
Belgium											
1976Q1 - 2013Q4	2	0.982	[0.018]	[0.064]	(1,-188.69)	-	-	-	-	-	-
Canada											
1970Q1 - 2013Q4	2	1.011	[0.000]	[0.174]	(1.-528.54)	-	-	-	-	-	-
Switzerland											
1970Q1 - 1989Q4	4	1.017	[0.099]	[0.946]	(1,-59.66)	-	-	-	-	-	-
Denmark											
1970Q1 - 2007Q1	5	1.018	[0.848]	[-]	(-,-)	-	-	-	-	-	-
Spain											
1971Q1 - 2006Q4	2	1.023	[0.000]	[0.203]	(1,-34.33)	-	-	-	-	-	-
Finland											
1970Q1 - 1989Q1	2	1.070	[0.254]	[-]	(-,-)	-	-	-	-	-	-
France											
1970Q1 - 2007Q1	3	1.008	[0.337]	[-]	(-,-)	-	-	-	-	-	-
United Kingdom											
1970Q1 - 2005Q1	2	1.001	[0.024]	[0.848]	(1,-25.75)	-	-	-	-	-	-
Ireland											
1970Q1 - 2004Q3	2	1.031	[0.001]	[0.174]	(1,2.094)	1.031	1.031	1.016	[0.667]	[0.000]	[0.000]
Japan											
1970Q1 - 1990Q3	6	1.006	[0.087]	[0.473]	(1,-151.58)	-	-	-	-	-	-
Netherlands											
1970Q1 - 2004Q3	3	0.998	[0.002]	[0.621]	(1,-32.59)	-	-	-	-	-	-
Norway											
1979Q1 - 2007Q4	4	1.025	[0.005]	[0.083]	(1,29.05)	1.025	1.002	1.045	[0.065]	[0.005]	[0.004]

Continued on next page

Table 1.2 – *Continued from previous page*

Sample	k	$\hat{\rho}$	**Cointegration** H_0: r=0	H_0: r=1	β_1'	$\hat{\rho}_1$	$\hat{\rho}_{1X}$	$\hat{\rho}_{1XS}$	**LR**$(M_{1X}\|M_1)$ χ^2(k-1)	**LR**$(M_{1XS}\|M_{1X})$ χ^2(1)	**LR**$(M_{1XS}\|M_1)$ χ^2(k)
New Zealand											
1970Q1 - 2008Q1	6	0.988	[0.460]	[-]	(-,-)	-	-	-	-	-	-
Sweden											
1980Q1 - 2007Q4	2	1.019	[0.074]	[0.491]	(1,16.40)	1.016	1.017	1.025	[0.092]	[0.015]	[0.013]
United States											
1970Q1 - 2007Q1	2	1.013	[0.000]	[0.703]	(1,0.73)	1.008	1.012	1.001	[0.000]	-	[0.000]

Note: Numbers in [•] denote p-values of the relevant test statistic. The three likelihood-ratio tests LR$(M_{1X}|M_1)$, LR$(M_{1XS}|M_{1X})$, and LR$(M_{1XS}|M_1)$ correspond, respectively, to testing (1.13) in (1.12), testing (1.15) in (1.14), and testing (1.13) and (1.15) in (1.12) simultaneously. k refers to the lag length of the model specification, which has been chosen to ensure no autocorrelation of the model residuals. $\hat{\rho}$, $\hat{\rho}_1$, $\hat{\rho}_{1X}$, $\hat{\rho}_{1XS}$ are the largest characteristic roots in the systems (1.10), (1.12), (1.14), and (1.16) respectively.

Visual Inspection

Figure 1.3 to 1.5 plots the explosive and the stochastic components of the price series for Ireland, Sweden and the USA[12]. In principle the components can be plotted for any restricted or unrestricted model, but the plots in Figure 1.3 through 1.5 are based on the model M_{1X}. For all three cases we note that the explosive trend explains a substantial part of the movement in housing prices. Secondly, the stochastic trend captures a large degree of the remaining variance. However, some persistent deviations appear when large changes to the individual markets occur, such as e.g. the dramatic increase in rent levels in Sweden in the beginning of the 1990s. The wedge that appears in the Irish stochastic component is a result of a linear increase in price during the boom rather than an decidedly explosive one, hence, the explosive trend is overshooting the actual price developments, which is unclear from the graphs due to the scales. The USA differs from Ireland and Sweden as the non-explosiveness of rents was clearly rejected. However, the visual inspection of the series clearly supports that prices have behaved explosively, and we conclude that prices and rents must have shared an explosive root implying that fundamentals to some degree have contributed to the US house prices up to the financial crisis.

The fit of the model is assessed by the plot of the actual and the fitted series of $\Delta\Delta_\rho X_t$ in Figures 1.3 to 1.5 and in general the model seems to fit the data well.

1.5 Conclusions

In this paper we investigate the empirical relation between house prices and rents with the purpose of detecting and assessing explosive behavior.

First, we use the univariate right-tailed unit root test procedure of Phillips et al. (2015) on the price-rent ratio in order to identify and date-stamp periods with explosive behaviour. For all but two of the 18 OECD countries (Germany and Italy) we find evidence of explosiveness. In date-stamping the bubble periods we detect a large degree of bubble synchronicity in the early 2000s and to a certain degree also around 1990.

The explosive alternative of the univariate right-tailed test procedure implicitly assumes that the price-rent ratio does not contain a unit root, i.e. it is assumed that if rents contain a unit root, it is shared with the unit root in prices such that the two variables 'cointegrate'. To explicitly investigate the cointegrating relation between prices and rents - in addition to explosiveness -, we also test for bubbles using the co-explosive VAR framework of Engsted and Nielsen (2012). From this analysis we find that there are large differences in the dynamics of housing markets across countries.

[12] We do not present these graphs for Norway, as the $\hat{\rho}_{1X}$ is close to unity and hence, the two components are almost indistinguishable.

In eight countries we find explosiveness with either cointegration or no cointegration between prices and rents. For the remaining countries the co-explosive VAR analysis does not indicate the presence of an explosive housing bubble.

Acknowledgments

A slightly shorter version of this paper has been accepted for publication in Journal of International Financial Markets, Institutions, and Money. The authors acknowledge support from The Danish Council of Independent Research (DFF – 4003-00022) and CREATES - Center for Research in Econometric Analysis of Time Series (DNRF78), funded by the Danish National Research Foundation. We thank Jérôme Brezillon for providing us with the OECD data along with Bent Nielsen and participants at the 14th OxMetrics User Conference, the Danish Academic Economists in North America Meeting, the 1st Conference on Recent Developments in Financial Econometrics and Applications, Danmarks Nationalbanks seminarrække, and anonymous referees, for useful comments.

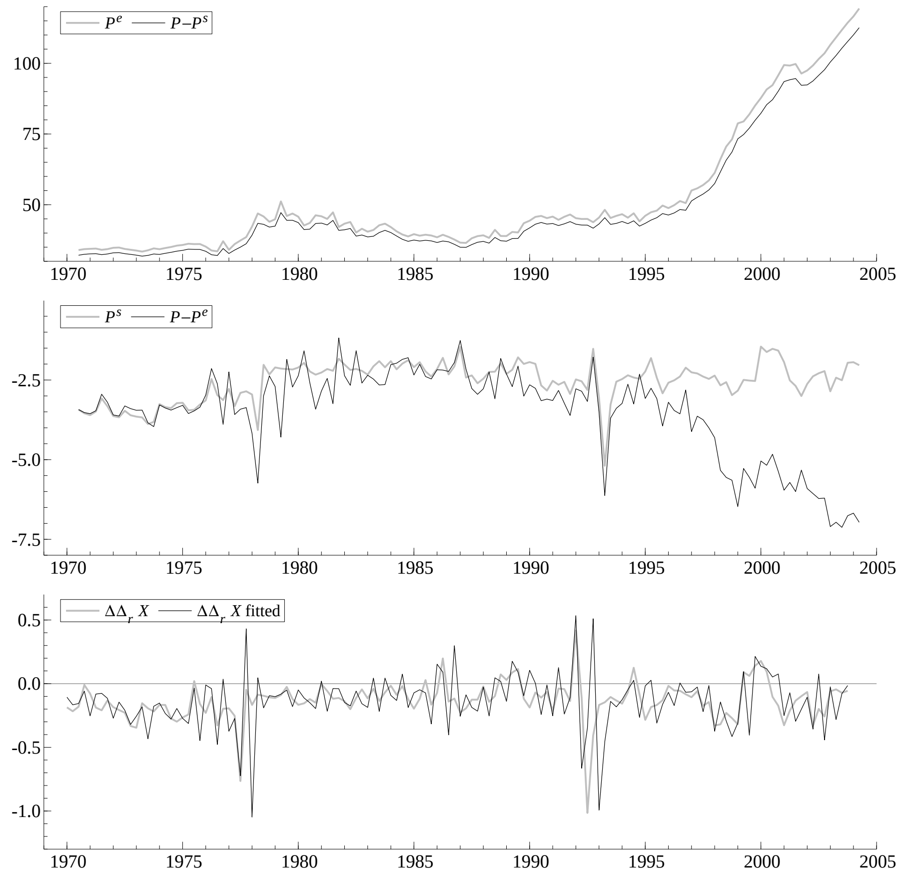


Figure 1.3. Ireland: Visual representation of the explosive and stochastic trends along with the fit of the stochastic and explosive trend stationary rent processes.

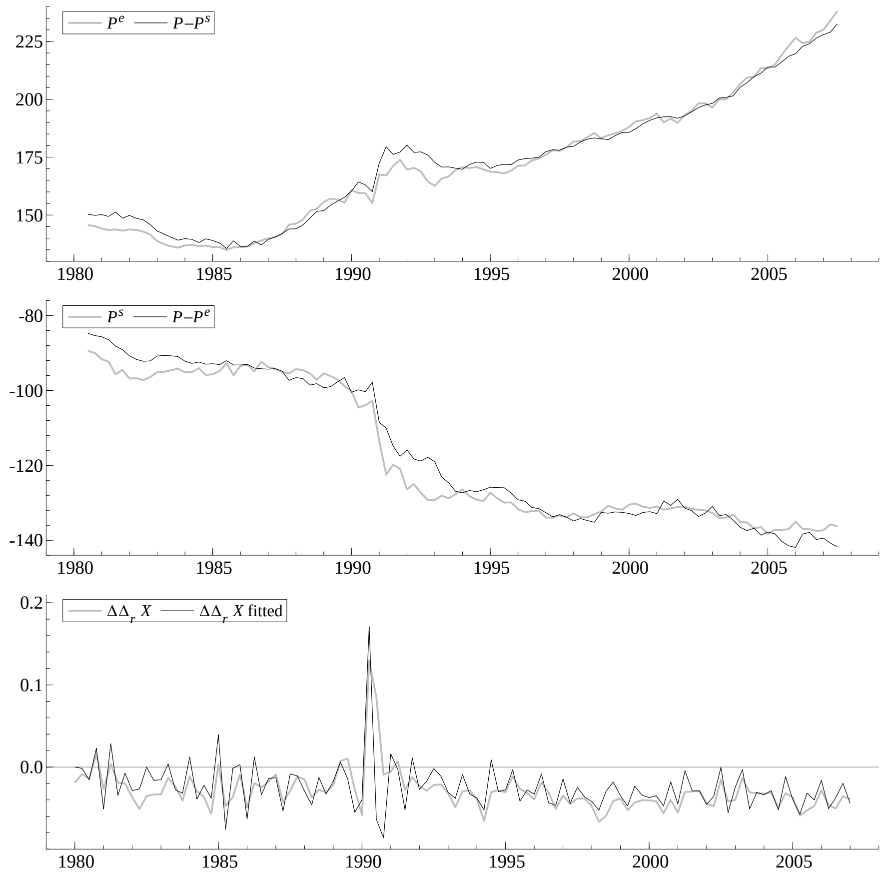


Figure 1.4. Sweden: Visual representation of the explosive and stochastic trends along with the fit of the stochastic and explosive trend stationary rent processes.

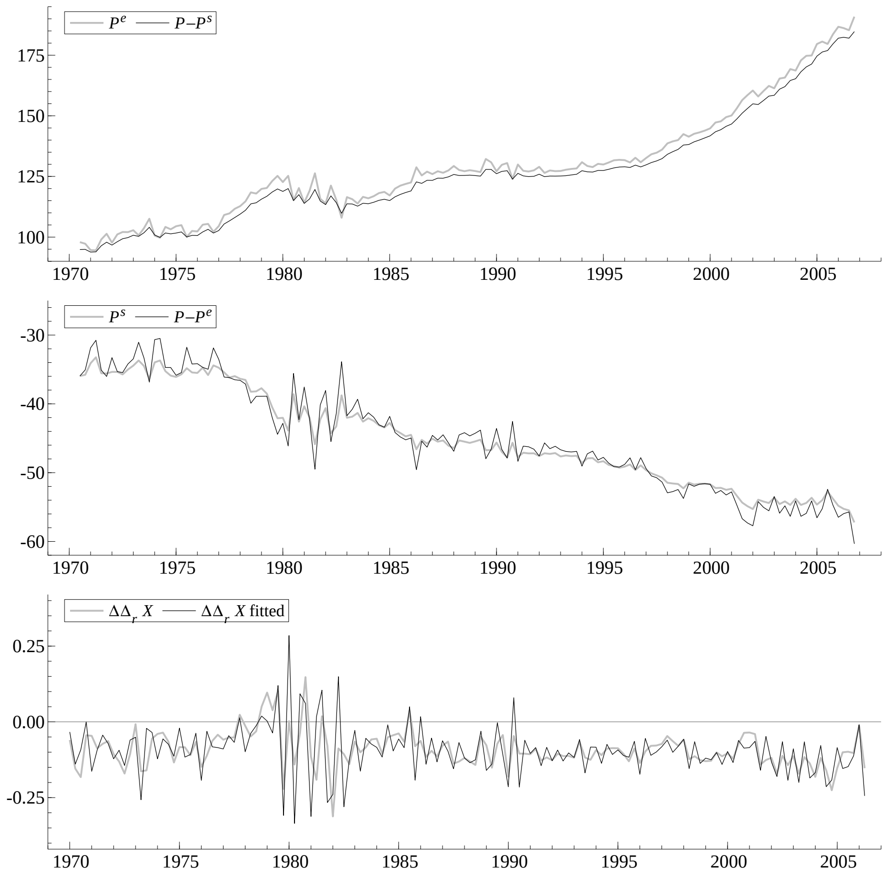


Figure 1.5. USA: Visual representation of the explosive and stochastic trends along with the fit of the stochastic and explosive trend stationary rent processes.

References

BHARGAVA, A. (1986): "On the Theory of Testing for Unit Roots in Observed Time Series," *The Review of Economic Studies*, 53, pp. 369–384.

BRUNNERMEIER, M. K. AND C. JULLIARD (2008): "Money Illusion and Housing Frenzies," *Review of Financial Studies*, 21, 135–180.

CAMPBELL, J. Y. (2013): "Mortgage Market Design," *Review of Finance*, 17, 1–33.

CAMPBELL, J. Y. AND R. J. SHILLER (1987): "Cointegration and Tests of Present Value Models," *Journal of Political Economy*, 95, 1062–88.

CAMPBELL, S. D., M. A. DAVIS, J. GALLIN, AND R. F. MARTIN (2009): "What moves housing markets: A variance decomposition of the rent-price ratio," *Journal of Urban Economics*, 66, 90–102.

CASE, K. E. AND R. J. SHILLER (1989): "The Efficiency of the Market for Single-Family Homes," *American Economic Review*, 79, 125–37.

CASPI, I. (2014): "Testing for a Bubble in Israeli Home Prices," Working Paper, Bar-Ilan University.

COCHRANE, J. H. (2008): "The Dog That Did Not Bark: A Defense of Return Predictability," *Review of Financial Studies*, 21, 1533–1575.

——— (2011): "Presidential Address: Discount Rates," *The Journal of Finance*, 66, 1047–1108.

CRAINE, R. (1993): "Rational Bubble: A Test." *Journal of Economic Dynamics and Control*, 17, 829–846.

DIBA, B. T. AND H. I. GROSSMAN (1988a): "Explosive Rational Bubbles in Stock Prices?" *American Economic Review*, 78, 520–30.

——— (1988b): "The Theory of Rational Bubbles in Stock Prices," *Economic Journal*, 98, 746–54.

ENGSTED, T. (2006): "Explosive bubbles in the cointegrated VAR model," *Finance Research Letters*, 3, 154–162.

ENGSTED, T. AND B. NIELSEN (2012): "Testing for rational bubbles in a coexplosive vector autoregression," *The Econometrics Journal*, 15, 226–254.

ENGSTED, T. AND T. Q. PEDERSEN (2014a): "Bias-Correction in Vector Autoregressive Models: A Simulation Study," *Econometrics*, 2, 45–71.

——— (2014b): "Housing market volatility in the OECD area: Evidence from VAR based return decompositions," *Journal of Macroeconomics*, 42, 91–103.

——— (2015): "Predicting returns and rent growth in the housing market using the rent-price ratio: Evidence from the OECD countries," *Journal of International Money and Finance*, 53, 257–275.

ENGSTED, T., T. Q. PEDERSEN, AND C. TANGGAARD (2012): "The Log-Linear Return Approximation, Bubbles, and Predictability," *Journal of Financial and Quantitative Analysis*, 47, 643–665.

EVANS, G. W. (1991): "Pitfalls in Testing for Explosive Bubbles in Asset Prices," *American Economic Review*, 81, 922–30.

GALLIN, J. (2008): "The long-run relationship between house prices and rents," *Real Estate Economics*, 36, 635–958.

GELAIN, P. AND K. J. LANSING (2014): "House prices, expectations, and time-varying fundamentals," *Journal of Empirical Finance*, 29, 3–25.

GHYSELS, E., A. PLAZZI, W. N. TOROUS, AND R. I. VALKANOV (2013): *Handbook of Economic Forecasting*, chap. Forecasting real estate prices, 509–580.

GIROUARD, N., M. KENNEDY, P. VAN DEN NOORD, AND C. ANDRE (2006): "Recent House Price Developments: The Role of Fundamentals," OECD Economics Department Working Papers 475, OECD Publishing.

HALL, S. G., Z. PSARADAKIS, AND M. SOLA (1999): "Detecting periodically collapsing bubbles: a Markov-switching unit root test," *Journal of Applied Econometrics*, 14, 143–154.

HAMILTON, B. W. AND R. M. SCHWAB (1985): "Expected appreciation in urban housing markets," *Journal of Urban Economics*, 18, 103 – 118.

HAUSMAN, J. (1978): "Specification Tests in Econometrics," *Econometrica*, 46, 1251–71.

HIMMELBERG, C., C. MAYER, AND T. SINAI (2005): "Assessing High House Prices: Bubbles, Fundamentals, and Misperceptions," NBER Working Papers 11643, National Bureau of Economic Research, Inc.

JOHANSEN, S. (1995): *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*, Oxford University Press.

KIVEDAL, B. R. K. (2013): "Testing for rational bubbles in the US housing market," *Journal of Macroeconomics*, 38, 369–381.

KRAINER, J. AND C. WEI (2004): "House prices and fundamental value," *Economic Letter*, November-October, 1–3.

LEROY, S. F. (1989): "Efficient Capital Markets and Martingales," *Journal of Economic Literature*, 27, pp. 1583–1621.

MCCARTHY, J. AND R. PEACH (2004): "Are home prices the next "bubble"?" *Economic Policy Review*, December, 1–17.

NIELSEN, B. (2010): "Analysis of Coexplosive Processes," *Econometric Theory*, 26, 882–915.

PAVLIDIS, E., A. YUSUPOVA, I. PAYA, D. PEEL, E. MARTINEZ-GARCIA, A. MACK, AND V. CROSSMAN (2014): "Episodes of exuberance in housing markets: in search of the smoking gun," Working paper, Lancaster University, Department of Economics.

PHILLIPS, P., S.-P. SHI, AND J. YU (2015): "Testing for multiple bubbles: Historical Episode of Exuberance and the Collapse in the S&P 500." *International Economic Review*, Forthcoming.

PHILLIPS, P. C., S.-P. SHI, AND J. YU (2013): "Testing for Multiple Bubbles: Historical Episodes of Exuberance and Collapse in the S&P 500," Cowles Foundation Discussion Papers 1914, Cowles Foundation for Research in Economics, Yale University.

PHILLIPS, P. C. B., Y. WU, AND J. YU (2007): "Explosive Behavior in the 1990s Nasdaq: When Did Exuberance Escalate Asset Values?" Working Papers 222007, Hong Kong Institute for Monetary Research.

PHILLIPS, P. C. B. AND J. YU (2010): "Dating the Timeline of Financial Bubbles during the Subprime Crisis," Cowles Foundation Discussion Papers 1770, Cowles Foundation for Research in Economics, Yale University.

PLAZZI, A., W. TOROUS, AND R. VALKANOV (2010): "Expected Returns and Expected Growth in Rents of Commercial Real Estate," *Review of Financial Studies*, 23, 3469–3519.

RICHARD, M. AND W. NANCY (1994): "Testing the Present Value Relation for Housing Prices: Should I Leave My House in San Francisco?" *Journal of Urban Economics*, 35, 245–266.

SHILLER, R. (2009): *Irrational Exuberance: (Second Edition)*, Princeton University Press.

TIMMERMANN, A. (1995): "Cointegration tests of present value models with a time-varying discount factor," *Journal of Applied Econometrics*, 10, 17–31.

WEST, K. D. (1986): "A Specification Test for Speculative Bubbles," NBER Working Papers 2067, National Bureau of Economic Research, Inc.

——— (1987): "A Specification Test for Speculative Bubbles," *The Quarterly Journal of Economics*, 102, 553–80.

YIU, M. S., J. YU, AND L. JIN (2013): "Detecting bubbles in Hong Kong residential property market," *Journal of Asian Economics*, 28, 115–124.

CHAPTER 2

DYNAMIC RESIDENTIAL SORTING - INVESTIGATING THE DISTRIBUTION OF CAPITAL GAINS

Simon Juul Hviid
Aarhus University and CREATES

Abstract

This paper estimates a dynamic residential sorting model that explicitly takes account of transaction costs and borrowing constraints and focuses on how capital gains differ across the wealth distribution. The model is estimated using unique Danish register data from 1992 to 2011 for housing owners. In a simulation exercise the paper finds that there are substantial differences in capital gains across the wealth distribution, with the highest decile receiving larger capital gains than lower wealth types. The dispersion is only influenced by transaction costs to a limited extent. Counterfactual simulations show that a progressive taxation scheme could have an inequality mitigating effect.

JEL Classification: D31, D63, D91, H20, R00, R21, R51
Keywords: income distribution, dynamic discrete choice models, residential sorting

2.1 Introduction

Recently, the long standing literature on income inequality has received considerable interest after the publication of Piketty (2014), and papers such as Piketty and Zucman (2013) find that the capital share of income has risen remarkably over the past decades.[1] Subsequently, Rognlie (2014) and Rognlie (2015) have investigated the same data further and found that the shift in income shares follow only from housing related income and writes that *"In the long run, there is a moderate increase in the aggregate net capital share [of income], but this owes entirely to the housing sector - where, since ownership is less concentrated, the consequences for inequality may be less severe." (Rognlie, 2015, p. 3).*

This paper investigates how the income from housing capital might differ across the wealth distribution and to what extent these differences enhance wealth inequality. My model draws from the literature of sorting in dynamic discrete choice models, which have been increasingly popular over the past decades as computational power and feasible estimation techniques have evolved.

The sorting literature was hatched with the notion of *voting with your feet* of Tiebout (1956).[2] With the technical contributions of McFadden (1974), the residential sorting literature evolved with the first notable contribution in McFadden (1978) and many papers have followed especially in recent years of which Kuminoff, Smith, and Timmins (2013) provide an excellent review. Most of the work on residential sorting takes a static rather than dynamic approach. However, the housing market is inherently dynamic in several respects as large financial transaction costs limit households from frequent optimization and therefore location choices will affect households for some time as local housing markets evolve over time leading to having a direct effect on wealth accumulation and future investment possibilities. Therefore, the model of this paper follows the literature of dynamic residential sorting that emerged after the seminal contribution of Rust (1987) that explicitly takes account of transaction costs in a dynamic structural model. Many of the advances are related to the progress in the industrial demand literature of dynamic demand of which Aguirregabiria and Nevo (2010) make an insightful review.

The modeling framework explicitly accounts for the wealth implications of the household location decision and follows Bayer, McMillan, Murphy, and Timmins (2015) whom develop a novel estimation approach that treats households as owners

[1] A review of this literature can be found in e.g. Piketty and Zucman (2013).

[2] Singleton (2013) provides a thorough analysis of the thoughts of Charles Tiebout and their implications.

rather than renters. The latter has been the dominant approach in the residential sorting literature. This allows for endogenous household wealth accumulation from capital gains and decreases in wealth from financial transaction costs. This approach extends the framework of Kennan and Walker (2011), which abstracts from the wealth dimension in an otherwise comprehensive dynamic model of residential sorting.

The advantage of estimating a structural model of residential choice that explicitly treats households as owners is that that the model identifies the structural parameters that drive household choices. With these parameters in hand, the model can predict household behavior under scenarios in which the structural settings are changed. As households are owners, this allows for measurement of the distributional impact that such changes might have on capital gains.

To the best of my knowledge there exists no prior investigation of the relation between capital gains on housing and wealth inequality implications. However, there exists a strong literature on the marginal propensity to consume out of housing wealth, which naturally touches upon many of the same features that are addressed in this paper.[3] This paper will not investigate the relation between housing capital gains and the consumption decision, but will only identify housing wealth evolution. I abstract from the consumption decision of the individual households.

The model is estimated using the rich Danish register data, which measures wealth, income, and socio-economic information very precisely. The data set is unique in an international comparison and the data allow estimation of the dynamic model, in which it is important that individuals and households can be followed over time. The sample covers the years 1992 to 2011 and includes the population of house-owning households. The estimation allows for capital gains differences over the price distribution within municipalities and allows for preference heterogeneity across the wealth distribution and across different educational levels of household members. The estimation is concluded by a simulation of a representative household to investigate how differences in initial conditions affect relative capital gains and moving costs across the wealth distribution, where the simulation strategy takes explicit account of borrowing constraints when households choose to relocate.

The simulation exercise finds that there are large differences in capital gains. In particular, there is a difference of almost 2 percentage points in average annual capital gains from households with no initial wealth to the highest wealth decile. Given the large sample size this is a very precise result. When compared to the average annual capital gain of 5.8%, this finding supports the hypothesis that income inequality derives from housing assets.

When there is a rising capital share of income, Piketty (2014) argues that there is

[3] Prominent papers in this literature include Campbell and Cocco (2007), Case, Quigley, and Shiller (2005), Carroll, Otsuka, and Slacalek (2011), Mian and Sufi (2011), Iacoviello, Edelstein, and Kim (2004), Browning, Gørtz, and Leth-Petersen (2013), and Leth-Petersen (2010).

a need for progressive capital taxation in order to mitigate rising inequality. This is an argument that appears in Piketty (2011), Piketty and Saez (2013), and Piketty and Zucman (2013). As Rognlie (2015) finds that inequality rises solely due to housing income, the proper measure for inequality mitigation would be taxation of housing capital. An appealing feature of housing capital is that it is internationally immobile compared to other capital, therefore, the problems of tax evasion as identified in Zucman (2013) should be limited.

In Denmark, housing wealth taxation is unevenly distributed both geographically and over the population. After a freeze of housing taxation in 2002 the dispersion was aggravated. From the model, a single household is simulated in a counterfactual world with and without facing a progressive effective property tax rate in all years and municipalities for the years after 2002. The simulated household does not affect prices and sorting behavior, so even though distributional gains are small, there is an additional general equilibrium effect if the tax rate is faced by all households. This follows as an increase in the effective tax rate for all individuals would imply relatively higher costs of housing in the metropolitan areas, which have experienced the largest house price increases. This increase could put a lid on the appreciation in these market segments and only lower the regional and wealth based dispersion of the housing implied wealth accumulation. Hence, the simulation strategy identifies an upper bound for the dispersion of capital gains in the wealth distribution from dropping the freeze of housing taxation.

The remainder of the paper is organized as follows. The next section describes the Danish register data and sample selection along with an overview of the institutional settings of the housing market and a brief descriptive overview of the distribution of housing prices and how moving patterns vary. Section 2.3 formally sets up the economic model of the paper. Section 2.4 explains the estimation strategy of the empirical model. In Section 2.5 results from the estimation and simulations are presented before Section 2.6 concludes.

2.2 Data

The basis for the estimation of the dynamic model is the rich Danish register based data sets and related to the one used in Hviid, Timmins, and Vejlin (2015), but with a longer time horizon and extended with data on educational attainments of individuals. The registers contain individual level data and are unique in an international comparison. It allows us to follow individuals and households over time, which is important for the model used in this paper. Ultimately, the data set covers Danish housing owners and buyers from 1992 to 2011. It includes data from several sources that are merged together and covers a variety of data on households and housing units. This section will describe the individual sources, the institutional settings of the housing market, and will include a descriptive investigation of the housing market.

Register Data

The unit of observation in the model will be a household and the starting point is the entire Danish population of individuals. This data set contains a few individual characteristics such as age and marital status, but the key variable is a unique family identifier that links families not only if they are married, but also if they consist of cohabiting adults and their children. There are strict requirements on the definition of a family; in particular these conditions include in particular that the family is living at the same address, which is an essential feature for the model of this paper.[4] The data set also contains unique address information, which links households to municipalities. Children are linked to their parent by social security numbers, and they are part of the family up to the age of 25 as long as they live at home. However, children's wealth and income are deleted from family characteristics.[5]

Individuals are merged with data from the Integrated Database for the Labour Market (IDA). IDA is a registry of the Danish population aged 15-74 and includes a range of socio-economic variables that are directly or indirectly related to the labor market. Several variables from IDA are relevant for this project. First of all it includes very detailed information of the highest educational attainment of the individual for every year, which allows assignment of individuals to different educational groups.[6] The IDA also contains very detailed income measures and in particular it separates earned income from other types of income. This will be utilized later on. There is a system in place of third party income reporting that makes the income variables very accurate, which is supported by the findings of Kleven, Knudsen, Kreiner, Pedersen, and Saez (2011). The essential variables from IDA describe detailed wealth information, which is a key component for the estimation of the model. Denmark had a wealth tax in the years 1987-1996 and during these years there was established another third party reporting system in which financial institutions, such as banks and mortgage institutions, reported all assets and liabilities of individuals including stocks, bonds, any kind of debt, and deposits. The only drawback is that pension savings are not observed, and given the aggregate size of these savings, the impact could be substantial. However, households can not borrow against pension savings for housing investments, so the impact on the estimation should be limited as long

[4] Further details on the family formation can be found at `http://www.dst.dk/da/Statistik/dokumentation/Times/moduldata-for-befolkning-og-valg/familie-id`.

[5] Children are assigned to families based on legal rather than biological relations.

[6] Specifications of educational attainments can be found at `http://www.dst.dk/da/Statistik/dokumentation/Times/uddannelsesdata/befolkningens-uddannelse/hfaudd`.

as households above age 60 are discarded, i.e. close to retirement.[7] After the wealth tax was eliminated, the system of reporting from financial intermediaries was kept in place and the wealth data is considered to be of very high quality.

A separate register contains information on sales of properties which includes the actual sales price, the unit sold and the exact date of registration and take over date. The register contains sales of commercial as well as private properties. However, the focus here will be on private properties.

In order to appropriately tax land and properties, the tax authorities have developed a system for valuation of these entities and the sum of the property and the land value should be in the region of (0%, -15%) in deviation of actual sales prices and are thus systematically undervalued.[8] Valuations are made bi-annually, where commercial properties are re-valued in even years and private properties are valued in uneven years.[9] These data are merged onto the data set which will be used to determine the subjective choice set under household borrowing constraints and the general movements in the house price distribution. As the public valuations by construction deviate from actual sales values at the mean, valuations are adjusted by the average percentage deviation from actual sales prices within municipalities year by year.[10]

The final data set used is one of homeowner information. It links individuals to their housing assets with specific address information that makes it possible to ensure that the household actually occupies the housing unit by comparing to address information from the first data set.

All data sets are measured ultimo of the year, except for the public valuations that are constructed as of October 1^{st}, and the sales data which is a collection of observations from the prior year.

The data sets are merged by linking them via either the social security number or the address information and relatively few observations are lost during the merging procedure. The information of adults in the families are aggregated, but individual educational attainments are kept. The selection is conditioned on housing ownership or purchase within the year and the top and bottom 1% of the wealth and house price distribution are deleted.[11]

[7] In principle individuals can get hold of private pension savings, but as this is very expensive it is rarely executed.

[8] The lower valuation is put in place to ensure that no household pays more than it ought to in property taxes and households can complain if they believe the valuation is too high.

[9] The public valuations have recently received some criticism due to some larger deviations from sales prices than expected.

[10] Ideally, I would adjust valuations by actual sales prices within municipality, year, and price decile; however, There are very few sales in some deciles. In particular, this is an issue at the top and the bottom of the price distributions.

[11] The data contains observations with a price of zero or very high prices, which are e.g. farms, and the top and bottom 1% are deleted as the choices of these households are likely governed by something else

Table 2.1 shows some descriptive statistics for the selected sample.

Table 2.1. Descriptive Statistics for Households

Co-variate	Mean	Std. Dev.	Minimum	Maximum
Fraction of buyers[a]	5.20 %	22.20%	0	1
Fraction of first-time buyer[a]	1.47%	12.05 %	0	1
Wealth[b]	951,277	1,599,002	-2,800,000	24,700,000
House price[b]	1,740,401	1,327,397	230,000	14,600,000
Observations	22,460,778			

[a] Buyers are measured in percentage of the population of housing owners.
[b] Wealth and prices are measured in 2011 Danish kroner.

Institutional Setting

This section will briefly outline the structure of the Danish housing and mortgage market.

Campbell (2013) provides a description of international mortgage markets including the Danish housing market. Roughly 50% of Danish households own housing assets and home ownership is evenly distributed across the wealth distribution. 50% home ownership is not large in an international comparison. When compared to other financial assets, home ownership is evenly distributed in the wealth dimension, whereas only about 25% of households own stocks and bonds and these assets are highly concentrated at the top of the wealth distribution. The degree of home ownership is larger for older households as seen in Figure 2.3.

When households invest in housing assets, the investment can be financed by low-interest mortgage loans and bank loans. Lower mortgage rates can be obtained financing up to 80% of the house's value, and up to rougly 95% can be financed by the high interest bank loans. Therefore, the remaining 5% down-payment should be private equity[12].

The Danish mortgage market is similar to the system in the U.S., but some differences should be noted. The mortgage market is the largest in the world relative to GDP, being above 100%.[13] The mortgage lending consists of covered bonds, such that

than those of the general population.

[12]This is not a legislative requirement and in principle households with a good financial standing can borrow above this limit, but in practice this requirement is often met.

[13]The size of the mortgage market is large, but Denmark also has one of the world's largest pension savings compared to GDP, i.e. a lot of the saving for retirement is done through pension funds rather than in housing. See Campbell (2013).

lending to house buyers is backed by an equivalent bond with maturity, amortizations, and mortgage rates that satisfy the underlying loan.[14] The maximum time to maturity is 30 years, which is also the most popular time to maturity with an average of 20 years. See Andersen, Campell, Nielsen, and Ramadorai (2014) for a thorough description of the Danish mortgage market.

Saving and borrowing in housing wealth is considered easy and housing wealth is fairly liquid due to the institutional setting of the mortgage market. First of all, households can at any time pre-pay their mortgage at face value without incurring any penalties. Furthermore, households can refinance their mortgage and in that way lower interest payments when rates are low. The possibility of refinancing minimizes the interest rate risk that households face when entering the housing market and makes timing of entrance less relevant. Refinancing is also possible for households with negative equity as long as it does not increase the outstanding mortgage debt. In 1992 the entire mortgage system was reformed from which point it was made possible for households to withdraw equity and increase the principle balance of the loan.[15] This was made possible also for consumption loans, which gained popularity in the years before the financial crisis.

In 1997, variable rate mortgage loans were introduced, but they did not gain much interest before another reform of the mortgage market was implemented in 2004. This reform allowed mortgage loans with deferred amortizations with deferral up to 10 years. Both deferred amortization loans and loans with variable interest rate became increasingly popular from this point. The possibility to defer amortizations allowed households to smooth consumption and has in particular been popular among younger households.

Prior to 'Pinsepakken', a tax and fiscal policy reform from 1998 that was implemented over the years 1998 to 2001, interest paid on mortgage and bank loans were 50% tax deductible but that was lowered to 32.4% over the period in order to stabilize the economy. Before 'Pinsepakken', housing was taxed based on imputed rents as a source of income, a system that was in place from 1903. After the implementation of 'Pinsepakken', Danish housing assets are taxed along two dimensions as there are both a property tax and a land tax. The property tax is received the central government and is set to 1% of the property value and 3% of the value that exceed 3m in 2012 Danish kroner. In order to appropriately tax properties, which are not sold annually, the tax authorities make a bi-annual valuation of all property and land values. In 2002, a nominal freeze of the property tax was introduced, which meant that all future property taxes would be based on the 2002 valuation. This freeze has had the implication that the effective tax rate for properties located at the outskirts of

[14] The mortgage market has seven mortgage institutions.

[15] Equity withdrawal is conditional on a certain credit standing. However, there is only a simple credit scoring system in place that stamps households with outstanding unpaid debt.

Denmark, which have experienced low if any price increases since 2002, is around or just below 1%, whereas properties located in east Jutland and in particular in the center of Copenhagen, which have appreciated about 100% in value since 2002, only face an effective tax rate below 0.5%. The system has hence implied a redistribution of income from the outskirts to the metropolitan areas. The land tax rate is set and collected by municipalities within the bounds of 0.16%-0.34%.[16] The ability of municipalities to manage the land tax has resulted in counter cyclical taxation in some municipalities, which could have a destabilizing effect on the general economy. In the years from 1987 to 1996 there was also a wealth tax in place, which meant that equity in housing was double taxed. There is no clear pattern in the geographical distribution of land taxes, but municipalities with more high income households tend to have lower land taxes.

As pointed out in Rangvid et al. (2013) many of the reforms to the mortgage market contributed to the dramatic house price increases up to the burst of the bubble in 2007, but also pro-cyclical fiscal policy and the state of the international economy have been contributing factors.

The Housing Market

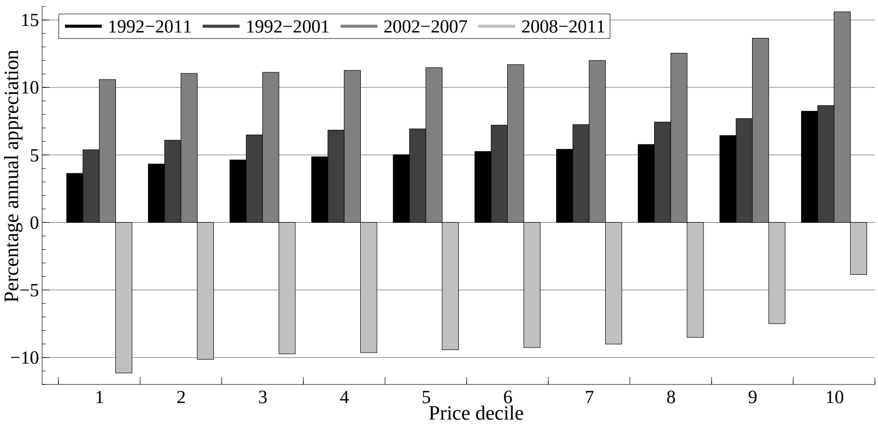
Over the past decades, house prices in Denmark have increased dramatically as in many other countries. Many factors have contributed to this development leading to an explosive bubble in house prices up to the burst in 2007.[17] The country mean real increase in the house prices of owner occupied housing over the period from 1992 to 2007 was more than 154.3% leading to large capital gains for housing owners. Subsequently, real housing prices fell by 24.7% in 2008-2011. However, the capital gains have not been evenly distributed geographically and across the population.

In the past decades, house prices have evolved heterogeneously across groups of houses. First of all, prices of more expensive houses tend to appreciate at a higher rate than for less expensive houses. Figure 2.1 shows how annual house price appreciation is distributed across the house price distribution from 1992 to 2011. The appreciation rate is monotonically increasing in house price with a difference of almost 5 percentage points from the lowest to the highest decile.[18] Figure 2.1 also shows appreciations for different sub samples starting with 1992 to 2002, before the

[16] The municipal reform in 2007 merged 270 municipalities into 98, which meant that differences in land taxes were evened out within the new municipalities. Prior to the reform, 0.1% of land taxes was collected to the counties, which was closed in the reform. There are certain restrictions on how large the annual changes to land taxes can be.

[17] See Engsted, Hviid, and Pedersen (2015) for a thorough econometric analysis of the explosiveness of, among others, Danish house prices.

[18] The pattern is the same within municipalities as across the entire price distribution. Figure 2.1 does not distinguish between municipalities.



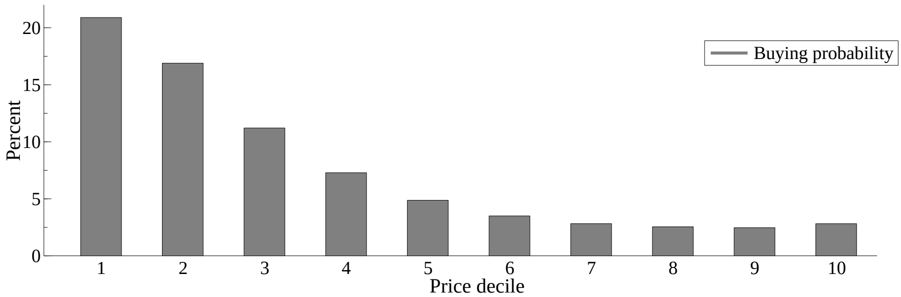
Note: Price decile '1' contains the 10% houses with lowest prices, decile '2' contains the subsequent 10%, etc.

Figure 2.1. Annual appreciation rates of owner occupied housing units for deciles of the house price distribution from 1992-2011 and sub periods.

tax freeze was introduced and the deferred amortization mortgages gained popularity. This sub sample has the lowest dispersion between the lowest and highest house price decile of 3.56 percentage points. The second sub sample covers the period from 2002 to 2007, where the dramatic house price movements occurred, leading to above 10% annual appreciation for all price groups and more than 15.9% annual appreciation for the most expensive houses. Lastly, in the post bubble years 2008-2011 real house prices fell, but much more for the least expensive housing units than for the most expensive ones and this is the period where the largest dispersion from the lowest to the highest price decile occurs, i.e. 7.1 percentage points. When the insight from Figure 2.1 is combined with the fact that households face borrowing constraints, which means that some households can not enter the housing market at any point in the house price distribution, then it suggests that capital gains are not evenly distributed across the population, which would aggravate the dispersion of the wealth distribution. As a rule of thumb, banks and mortgage institutions are willing to lend 3.5 times pre-tax income plus net wealth for housing investments.

Moving is costly with average costs being 4% of house value[19] and hence, every time a household chooses to move they accept this wealth reduction. For the mean wealth level and the mean house price the cost is 7.32% of total wealth. Figure 2.2 shows the probability of housing owners choosing to move for the deciles of the price

[19] Correspondence with several real estate agencies confirm that this would be the average total cost of housing transactions.



Note: Price decile '1' contains the 10% houses with lowest prices, decile '2' contains the subsequent 10%, etc.

Figure 2.2. Transaction probabilities for deciles of the price distribution.

distribution. Moving probabilities are relatively high for the least expensive housing units and are decreasing to the median of the price distribution from where moving probabilities are more or less constant over the top half of the price distribution. For the least expensive units, the fraction of cumulative moving costs over wealth will be higher than for the higher price deciles, which is another channel that potentially aids the dispersion of the capital gains distribution.

Obviously there are many underlying factors explaining the differences in moving probabilities and one of the major factors is that younger less wealthy households buy more frequently as they enter the housing market and maybe increase housing size when children are born. Figure 2.3 shows the distribution of sales and ownership at different ages of the oldest member of the household. Buying probabilities increase steadily throughout the twenties and the highest intensity of the house buying behavior is in the beginning of the thirties, from which point it is slowly decaying. The largest share of house-owning households occurs at age of fifty.

Housing choices and implications for wealth accumulation differ substantially with geography. Figure 2.4 shows two maps of Denmark divided into municipalities as of the municipal reform in 2007 where 270 municipalities were merged into 98. The left hand side map shows how the price per square meter differs across municipalities from zero (black)[20] to the largest price appreciation (white), which is Frederiksberg in Copenhagen. That municipality appreciated from being 27% to 202% above the country mean between 1992 and 2011. The right hand side map shows the net migration for municipalities ranged from lowest (white) which is the island Læsø at the top of the map, which experienced a decrease in population of 13.7%, to the largest increase which was in Copenhagen at 15.8%, but also East Jutland has experienced

[20] The lowest appreciation in nominal price per square meter from 1992-2011 occurred in the municipality of Lolland with only 25% nominal price increase over the sample.

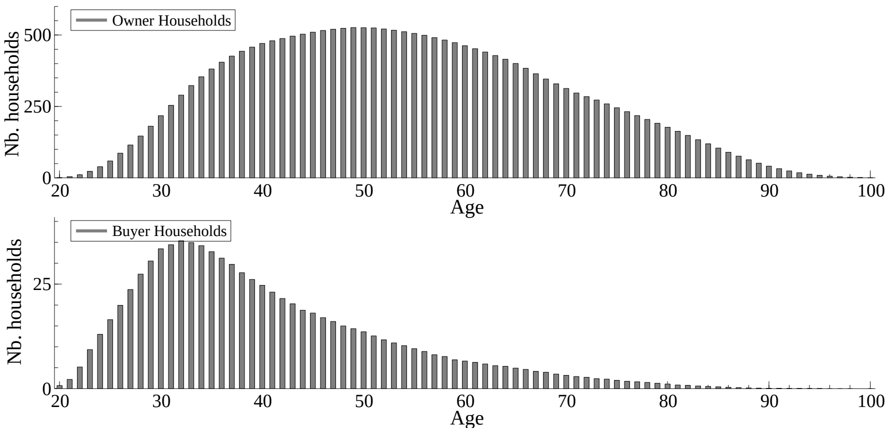


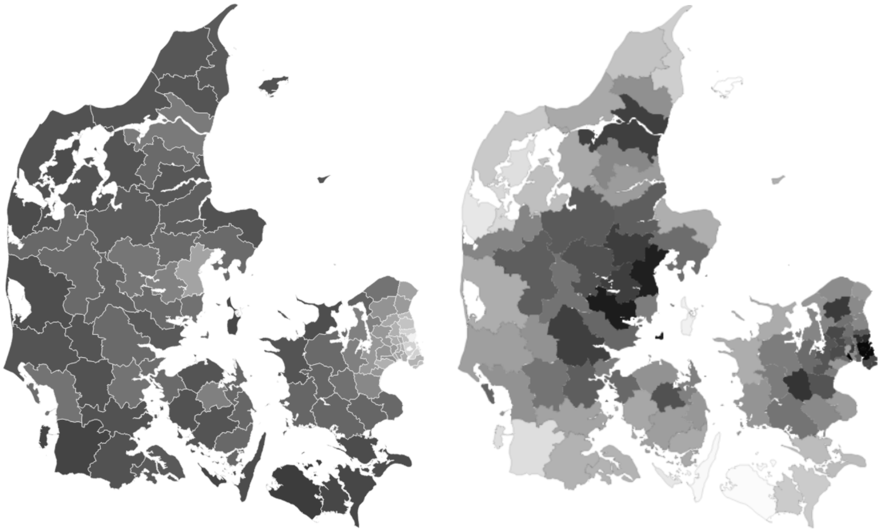
Figure 2.3. Ownership and purchases of housing for ages of the oldest member of households. The frequency is measured in 1000's of households.

significant population growth. The migration pattern and appreciation distribution across municipalities are interrelated; in the context of sorting, an equilibrium is obtained when prices are such that there is no excess demand; i.e., under a borrowing constraint all households have weighted prices and municipality amenities etc. and chosen the optimal location.

2.3 The Theoretical Model

Based on the dynamic structural model of Rust (1987) and the advances made in the discrete choice literature up to Bayer, McMillan, Murphy, and Timmins (2015), I develop a structural model to explain how ownership and location choices influence wealth accumulation of households. The model accommodates key dynamic features of the housing market such as moving costs, borrowing constraints, and forward looking behavior. With a structural model and identification of the structural parameters, it is possible to change the institutional setting and predict household behavior and measure the impact of household wealth accumulation through housing ownership.

In this model of dynamic residential sorting, households, denoted i, will choose at the end of every period, denoted t, whether to stay in the current housing unit or move to a different one. Define the binary choice variable $y_{i,t}$ to be one if a move occurs and zero otherwise. If a household moves, it chooses between J different municipalities to reside in. In addition to the J municipalities, housing owners are allowed the choice of leaving the housing market to rent rather than own and the outside option is assigned



House price appreciations based on m^2 prices of sales from 1992 to 2011 where black is zero and white is the largest appreciation.
Source: Association of Danish Mortgage Banks, Housing market statistics.

Net migration from 2006 to 2013 where white is the lowest and black is the largest change.
Source: Statistics Denmark, Population Statistics.

Figure 2.4. Municipality differences in house price appreciation (left) and net migration in Denmark (right).

the identifier 0.[21] Let $d_{i,t} = j \in \{0, 1, \ldots, J\}$ define the location specific choice variable, where $d_{i,t} = d_{i,t-1}$ if the household does not move, but stays in the current housing unit.

The choices of households are determined by two observed state vectors and an unobserved state vector containing idiosyncratic shocks. The first state vector, $z_{i,t}$, includes household specific information, which covers e.g. wealth information, income, and educational level of household members. Included in the first state vector is information on the housing unit where house value, $h_{i,t}$, is of particular interest. The second state vector, $x_{j,t}$, contains municipality specific information. The municipality specific state vector includes the house price distribution, local wage levels, and differences in relative taxes. The model includes an idiosyncratic shock that is unobserved and varies across municipalities, households and time. Denote the idiosyncratic shock by $\varepsilon_{i,j,t}$ and note that it covers both taste shocks and shocks to

[21] The model is one of house ownership and therefore there is no distinction between renting locations. This option also covers households that choose to leave Denmark.

moving costs. Including the unobserved state vector is necessary, as it is unlikely that the observed states are sufficient to explain the complex nature of residential choices and can explain differences between observed and expected behavior. Define $\Omega_{i,t} \equiv \{z_{i,t}, x_t, x_{t-1}, d_{i,t-1}\Omega_{i,t-1}\}$, as the set that contains all observable information sets about the household, its housing unit, municipalities, and past resided municipalities which will be the basis for the choices of household i at time t.

When a household resides in a given municipality it yields a flow utility that depends on individual characteristics and municipality amenities, which will be denoted by $u_{i,j,t} = u\left(z_{i,t}, x_{j,t}, \varepsilon_{i,j,t}\right)$. However, when the household chooses to move it incurs moving costs to the flow utility, $c_{i,t} = c(z_{i,t})$, and continuation values are affected as wealth is reduced. The flow moving costs are measured in utility and include a financial component, which reflects the disutility from the wealth decrease, and a psychological component. Only households that move are subject to the moving costs. The union of flow utility and flow moving costs is the measure that households face when making location decisions, which is denoted $u^c_{i,j,t} = u_{i,j,t} - c_{i,t} I_{[y_{i,t} \neq 0]}$, where $I_{[\cdot]}$ is an indicator function.

The households choose to move to increase flow utility today and in the future. However, households are constrained by a borrowing constraint; i.e., households can only move to the fraction of houses in municipality that falls within the subjective borrowing constraint, $b(\Omega_{i,t})$. In particular, $h_{i,t} \leq b(\Omega_{i,t})$.

The state variables have Markovian transition probabilities,

$$q_{i,t} = q(\Omega_{i,t+1}, \varepsilon_{i,t+1} | \Omega_{i,t}, d_{i,t}, \varepsilon_{i,t}). \tag{2.1}$$

Thus, $q_{i,t}$ denotes the probability of household characteristics, prices, and municipality amenities to transition from one state to another. Assume that all households face a constant discount factor, β, which is crucial for the identification of the model.[22] Then, the primitives of the model are constituted by (u, c, b, q, β).

Given the primitives of the model, the household's infinite horizon objective function is given by the expected sum of all future discounted flow utilities,

$$U_{i,t} = u^c_{i,j,t} + E_t \left[\sum_{\tau=t+1}^{\infty} \beta^{\tau-t} u^c(z_{i,\tau}, x_{j,\tau}, \varepsilon_{i,\tau}) \,\middle|\, \Omega_{i,t}, \varepsilon_{i,t}, d_{i,t}, y_{i,t} \right]. \tag{2.2}$$

Assuming an optimal control sequence for the two choice variables, the lifetime value function can be defined,

$$V_t(\Omega_{i,t}, \varepsilon_{i,t}) = \max_{\{y_{i,t}, d_{i,\tau}\}_{\tau=t}^{\infty}} E_t \left[\sum_{\tau=t}^{\infty} \beta^{\tau-t} u^c(z_{i,\tau}, x_{j,\tau}, \varepsilon_{i,\tau}) \,\middle|\, \Omega_{i,t}, \varepsilon_{i,t}, d_{i,t}, y_{i,t} \right]. \tag{2.3}$$
$$\text{s.t.} \quad h_{i,\tau} \leq b(\Omega_{i,\tau}) \quad \text{if } y_{i,\tau} = 1.$$

[22] See e.g. Rust (1994) or Magnac and Thesmar (2002) for thorough discussion of identification in dynamic discrete choice models.

The time independent structure of the infinite horizon specification allows us to ignore the t subscript on the lifetime value function. From here the Bellman equation can be defined

$$V(\Omega_{i,t},\varepsilon_{i,t}) = \max_{y_{i,t},d_{i,t}} \left[u^c_{i,j,t} + \beta E_t \left[V\left(\Omega_{i,t+1},\varepsilon_{i,t+1} \middle| \Omega_{i,t},\varepsilon_{i,t},d_{i,t}\right)\right]\right], \tag{2.4}$$
$$\text{s.t.} \quad h_{i,\tau} \leq b(\Omega_{i,\tau}) \quad \text{if } y_{i,\tau} = 1,$$

which is a contraction mapping in $V(\cdot)$ under certain regularity condition described in Rust (1994).

Identification of this type of dynamic discrete choice model is difficult in general, which leads us to impose some restrictions on the unobserved state vector that were first formalized by Rust (1987).[23] The first assumption is that of Additive Separability (AS) between the flow utility and the idiosyncratic error term, which gives the flow utility function the following form

$$u^c_{i,j,t} = u_{i,j,t} - c_{i,t} I_{[y_{i,t}]} + \varepsilon_{i,j,t}. \tag{2.5}$$

Secondly, impose the Conditional Independence assumption (CI), which puts some limiting structure on the dependence of the observed and unobserved state variables. Specifically, it imposes that, conditional on the current states, $\Omega_{i,t}$, the current idiosyncratic shocks, $\varepsilon_{i,t}$, have no predictive power on the future states, $\Omega_{i,t+1}$. Hence, the transition density of the Markov process is the product of the marginal transition densities of the state vectors, q_Ω, and the idiosyncratic shocks, q_ε,

$$q(\Omega_{i,t+1},\varepsilon_{i,t}|\Omega_{i,t},d_{i,t},\varepsilon_{i,t}) = q_\Omega(\Omega_{i,t+1}|\Omega_{i,t},d_{i,t})q_\varepsilon(\varepsilon_{i,t+1}|\varepsilon_{i,t}). \tag{2.6}$$

In models of inter-temporal optimization the subjective beliefs of households need to be aligned with the Markov process governing the transition probabilities of the state vectors, which was first noted by Magnac and Thesmar (2002) implying the Rational Expectations assumption (RE). Denote the transition density of the Markov process of subjective beliefs by $\mu(\Omega_{i,t+1},\varepsilon_{i,t+1}|\Omega_{i,t},d_{i,t},\varepsilon_{i,t})$. Then the assumption implies

$$\mu(\Omega_{i,t+1},\varepsilon_{i,t+1}|\Omega_{i,t},d_{i,t},\varepsilon_{i,t}) = q(\Omega_{i,t+1},\varepsilon_{i,t+1}|\Omega_{i,t},d_{i,t},\varepsilon_{i,t}). \tag{2.7}$$

Furthermore, assume that the idiosyncratic shocks are i.i.d. type 1 extreme value distributed. Along with AS, CI, and RE, this makes it possible to define the choice specific value function

$$v^c_j(\Omega_{i,t}) = u^c_{i,j,t} + \beta E\left[log\left(e^{v_j(\Omega_{i,t+1})} + \sum_{k=0}^{J} e^{v^c_k(\Omega_{i,t+1})}\right) \middle| \Omega_{i,t},d_{i,t} = j\right]. \tag{2.8}$$

[23] Several papers have investigated the identification of dynamic discrete choice models in detail i.e. Rust (1994), Magnac and Thesmar (2002) and Abbring and Heckman (2007). Additionally, Berry and Haile (2009) have analyzed identification of static equilibrium discrete choice models.

Based on the differences in lifetime value as determined in this equation, households will choose to move and choose a municipality in each period. As for the flow utility, it is possible to define $v_j(\Omega_{i,t}) = v_j^c(\Omega_{i,t}) + c_{i,t} I_{[y_{i,t}]}$, which will be the basis for parts of the estimation below.

2.4 The Empirical Model

The estimation of the dynamic demand model follows three steps that each estimate one of the primitives of the model (b, v, c).[24] The estimation commences with estimating the lifetime values from conditional choice probabilities for every year, municipality, and type combination, where types are defined by the household wealth and the educational attainments of household members. In the second stage move/stay probabilities are matched to the data and both financial and psychological moving costs are identified. The estimation concludes with a wage regression, which identifies the municipality differences in wages and hence the effective borrowing constraint of the household at different locations. Having estimated the primitives, the final step simulates households through the sample and measures the implications for wealth accumulation across the wealth distribution based on the model and various counterfactual models for a treated household. The estimation of the model requires information of households and states over time, for which the Danish population-wide registers' data are unique and appropriate sources.

Conditional Choices

Based on the conditional choice probabilities of Hotz and Miller (1993), the first stage of the estimation emerges from Proposition 1 of the same article and is based on the inversion of the conditional choice probabilities in Berry (1994).

In every period, households choose if they wish to stay in the current housing unit or move to a different one. The household will choose to move to the municipality that provides the highest lifetime value based on (2.8). Hence, the household will move to a different location if the value of living there is larger than the value of living in the current location added the moving costs and larger than any other place it could move to.

The value of living in the different locations can be obtained in different ways given the assumption that the idiosyncratic error term, $\varepsilon_{i,j,t}$, is i.i.d type 1 extreme value distributed. Some papers have used a parametric specification, but here a non-parametric version is applied to uncover the lifetime values. As in Berry (1994)

[24] Based on the identifying assumptions, the remaining primitive q_ε is assumed known and the estimation is independent of the subjective discount factor β.

households are divided into different household types based on observable characteristics. These types are defined discretely and denoted by $\tau = \tau(z_{i,t})$. Lifetime values are then assigned based on the type of the household, such that all households of type τ get the value $v^{\tau}_{j,t}$ from living in municipality j at time t. The types are defined discretely from 10 wealth bins and 9 education bins resulting in a total of 90 types. The education bins are created such that there are three educational levels and that the household can consist of either one or two adults.[25] These particular dimensions are chosen to capture differences in borrowing constraints as good as possible.[26]

Based on (2.8) the value function of a household of type τ is determined as

$$v^{\tau_t}_{j,t} = u^{\tau_t}_{j,t} - c^{\tau_t} + \beta E_t \left[log \left(e^{v^{\tau_{t+1}}_{j,t+1}} + \sum_{k=0}^{J} e^{v^{\tau_{t+1}}_{j,t} - c^{\tau_{t+1}}} \right) \middle| \Omega_{i,t}, d_{i,t} = j \right]. \tag{2.9}$$

As is clear from the structure, the moving cost component drops out of the maximization as an additive constant when conditioning on a household having chosen to move. Hence, with the assumption that the idiosyncratic error term is type 1 extreme value distributed and conditional upon moving, the conditional choice probability, i.e. the probability of a household of type τ choosing municipality j, is given by

$$P^{\tau}_{j,t}(v^{\tau}_t) = \frac{1}{\sum_{k=0}^{J} e^{v^{\tau}_{k,t} - v^{\tau}_{j,t}}}. \tag{2.10}$$

The conditional choice probabilities are determined based on the value differences within every type and therefore the value functions are unique to an additive constant. This constant will later be identified and utilized to determine the value difference within types. This stage therefore identifies the normalized value function $\tilde{v}^{\tau}_{j,t} = v^{\tau}_{j,t} - m^{\tau}_t$, where m^{τ}_t is the normalizing constant.

The estimation commences by obtaining non-parametric estimates of the conditional choice probabilities by calculating the share of each type that chooses to move to each municipality

$$\tilde{P}^{\tau}_{j,t} = \frac{\sum_{i=1}^{N_1} I_{[d_{i,t}=j]} \cdot I_{[Z_{i,t} \in Z^{\tau}]}}{\sum_{i=1}^{N_1} I_{[Z_{i,t} \in Z^{\tau}]}}, \tag{2.11}$$

[25]The first spouse can be of any of the three types and the second spouse can additionally be of type 0 if the household consists of only a single adult. The three educational levels applied are no education, vocational training, or tertiary education. Thus, there are $3! + 3 = 9$ education bins. If all housing owners are divided based on these three educational levels then close to 1/3 are in each bin.

[26]The main concern is that some young households might be able to borrow more than their observables would normally imply.

where N_1 is the total number of buyers. The model allows explicitly for an outside option, where N_2 owners choose to rent and the probability is estimated as

$$\hat{P}^{\tau}_{0,t} = \frac{\sum_{i=1}^{N_2} I_{[d_{i,t}=0]} \cdot I_{[Z_{i,t} \in Z^{\tau}]}}{\sum_{i=1}^{N_2} I_{[Z_{i,t} \in Z^{\tau}]}}. \tag{2.12}$$

The inside conditional choice probabilities are adjusted down to correct for the outside option from $\hat{P}^{\tau}_{j,t} = (1 - P^{\hat{\tau}}_{0,t})\tilde{P}^{\tau}_{j,t}$. Lastly, after having obtained the non-parametric estimate of the conditional choice probabilities then (2.10) can be inverted to estimates of the normalized value functions, as proved in Hotz and Miller (1993), [27]

$$\hat{v}^{\tau}_{j,t} = \log(\hat{P}^{\tau}_{j,t}) - \frac{1}{J+1} \sum_{k=0}^{J} \log(\hat{P}^{\tau}_{k,t}). \tag{2.13}$$

Moving Costs

When households choose to move they pay a moving cost that reduces wealth instantly. Moving costs are decomposed into two terms: financial moving costs, $f(z_{i,t})$, and psychological moving costs, $p(z_{i,t})$. The costs are dependent on $z_{i,t}$, which is the vector of household characteristics including the value of the currently occupied housing unit

$$c(z_{i,t}) = f(z_{i,t}) + p(z_{i,t}). \tag{2.14}$$

The financial part is calculated as 4% of the actual price of the house sold for non-first-time buyers and is treated as observed.[28] As defined above $y^{\tau}_{i,t}$ is the binary decision variable of the household, where $y^{\tau}_{i,t} = 0$ is equal to choosing not to move and $y^{\tau}_{i,t} = 1$ is a choice of moving to the neighborhood that yields the highest lifetime value but incurring the moving cost. Suppose a household chooses to move, then the endogenous wealth reduction from paying the financial moving cost will change the household type τ down to $\bar{\tau}$ with the associated lifetime value of location choice k to $v^{\bar{\tau}}_{k,t}$, which yields the choice function,

$$y^{\tau}_{i,t} = I\left[v^{\tau}_{j,t} + \zeta_{i,j,t} < \max_{\{k\}_0^J} [v^{\bar{\tau}}_{k,t} + \varepsilon_{i,k,t}] - p(z_{i,t}) \right]. \tag{2.15}$$

[27] As it follows from the normalized value function, $\hat{P}^{\tau}_{j,t} = 0$ will constitute a problem, hence, it is assumed that there exists a strictly positive lower bound, $\mu \sim 0$, on the probability of any type moving to any neighborhood, i.e. $\hat{P}^{\tau}_{j,t} = \max[\mu, \hat{P}^{\tau}_{j,t}]$. This occurs for 7% of the type choices. Investigations of the sensitivity of the structural parameters to the choice of the lower bound show limited impact of the choice.

[28] This is based in information from several real estate institutions. Note that a wrong percentage would pass through to the estimated moving cost parameters, but not impact the final results.

The estimation in the preceding section provides estimates of the normalized value function, so the lifetime values can be expressed in terms of the normalized values $v^{\tau}_{j,t} = \hat{v}^{\tau}_{j,t} + m^{\tau}_{t}$. Inserting the normalized value and normalizing constants, the binary move/stay choice can be expressed as

$$y^{\tau}_{i,t} = I\left[\hat{v}^{\tau}_{j,t} + \zeta_{i,j,t} < \max_{\{k\}^{J}_{0}}[\hat{v}^{\bar{\tau}}_{k,t} + \varepsilon_{i,k,t}] - (m^{\tau}_{t} - m^{\bar{\tau}}_{t}) - p(z_{i,t})\right]. \tag{2.16}$$

Here $(m^{\tau}_{t} - m^{\bar{\tau}}_{t})$ determines the value difference of being of the wealth type τ and the lower wealth type $\bar{\tau}$. The difference is unobserved but measures the (dis)value of paying the moving costs and thus, $f(z_{i,t}) = (m^{\tau}_{t} - m^{\bar{\tau}}_{t})$.

The moving costs are parameterized to household characteristics, assuming linearity in the arguments.[29] Hence,

$$f(z_{i,t}) = (m^{\tau}_{t} - m^{\bar{\tau}}_{t}) = h_{i,t} \cdot I_{[j\neq 0]} \cdot 0.04 \cdot z'_{i,t}\gamma_{f}, \tag{2.17}$$

$$p(z_{i,t}) = z'_{i,t}\gamma_{p}, \tag{2.18}$$

where $h_{i,t}$ is the value of the household's housing unit. The moving cost parameters γ_f and γ_p are estimated by maximum likelihood according to

$$P(y^{\tau}_{i,t} = 1, j) = \frac{e^{\hat{v}^{\tau}_{j,t}}}{e^{\hat{v}^{\tau}_{j,t}} + \sum_{k=0}^{J} e^{\hat{v}^{\bar{\tau}}_{k,t} - 0.04 \cdot h_{i,t} \cdot z'_{i,t}\gamma_f - z'_{i,t}\gamma_p}}. \tag{2.19}$$

McFadden (1974) shows that the log-likelihood of the choice probability is globally concave in the parameters corresponding to γ_f and γ_p in this model.

Having parameterized and estimated $(m^{\tau}_{t} - m^{\bar{\tau}}_{t})$ it is possible to back out the inter-type consistent estimates of the lifetime values, as mentioned in the above stage. To do so a choice has to be made on some reference type and level of lifetime value. Here that choice is to set the mean choice specific lifetime value to zero for the lowest wealth type for each combination of wealth and education type and year. Then the $\hat{m}^{\tau}_{t}$'s can be recursively calculated and, hence, the unnormalized choice specific lifetime value from $v^{\tau}_{j,t} = \hat{v}^{\tau}_{j,t} + \hat{m}^{\tau}_{t}$ is identified.

Borrowing Constraints

When a household buys a housing unit, it faces a certain borrowing constraint based on what the bank and mortgage institution are willing to lend it. The two key components that determine the borrowing constraint are wealth and income, which

[29] A constant and income level are included in γ_f, and γ_p includes a constant, wealth level, and a dummy for household composition, but could include a lot of other variables such as age, children, etc.

are measured very precisely in the Danish register data. Across Denmark, there are significant geographical differences in wages, such that two identical individuals living in two different municipalities might not require the same wage for their labor, which would also lead to different borrowing constraints. Therefore the estimation of the borrowing constraint starts with identifying the wage differentials across municipalities for different household types. The same problem obviously does not apply for household wealth.

Many papers have investigated wage differentials, where Roback (1982) is one of the first to estimate a hedonic model which is explicitly identified from estimation of geographical wage differentials. More recently, Kennan and Walker (2011) and Gemici (2007) estimate dynamic models of migration, emphasizing the effects of income, which is also utilized in Bishop (2012) for dynamic hedonic valuation. The estimation of the model in the present paper differs from these papers as it utilizes the advances from the labor literature made after the seminal paper Abowd, Kramarz, and Margolis (1999) that identify firm and worker fixed effects. Here the estimates of the 'firm' fixed effects are substituted by municipality fixed effects for different educational levels. For the model and later simulations it is important to correct for the individual fixed effects as high wage individuals will sort into certain municipalities and have better borrowing options and neglecting this feature would distort the results.

The wage equation is set such that it captures the key features of the model as the type space is discretized and the wage equation is adjusted accordingly. For individual s in municipality j at time t the wage is modeled as

$$w_{s,j,t} = \mu_{j,\tau} + \eta_s + z'_{i,t}\alpha + \omega_{s,t}, \tag{2.20}$$

where $\mu_{j,\tau}$ is the location specific fixed effect on income for individuals of type τ and η_s is the individual specific fixed effect. The wage equation includes $z'_{i,t}\alpha$, which is the wage effect from observable individual characteristics. Lastly, $\omega_{s,t}$ is an i.i.d. transient component that captures random shocks as well as unobserved ability of the individual. Kennan and Walker (2011) show that a generalized version of this type of wage equation is identified.

When the wage equation is estimated, the borrowing constraint can be identified for every household at every location. Isolating the individual permanent component will be important for the simulations in the final stage in order to accommodate changes to the initial location of the household without changing the unobserved ability of the household. This would be the outcome from simply using municipality mean wages. As described in Section 2.2, the borrowing constraint is based on net wealth and pre-tax income of the household, so in municipality j at time t the borrowing constraint of household i is

$$b_{i,j,t} = \sum_{s\in i} 3.5 \cdot w_{s,j,t} + wealth_{s,t}, \tag{2.21}$$

where wealth includes both housing and non-housing wealth.

Capital Gains Simulations

There are two key dynamic components to the residential sorting model. These are the financial and psychological moving costs and realized housing asset appreciations. Households have to determine whether they wish to move today and reduce their wealth or postpone a move and save for a better future housing investment. At the same time location matters for wealth accumulation as expected appreciation rates differ across locations. This section describes how agents are simulated through the model, which takes account of these effects. I then use the model to analyze how initial conditions affect the outcomes of the household. In particular the focus will be on how differences in initial wealth affect wealth accumulation with respect to housing ownership. The simulation concludes by making various counterfactual simulations for a treated household to see how different institutional settings would affect capital gains over the wealth distribution.

The simulations first take a representative household of type τ in year 0 with mean characteristics located in municipality j_0. Then the household determines whether to move in this period to the municipality that yields the largest lifetime value or to stay in the current housing unit. The decision is determined by drawing a $J+1$ vector of the type 1 extreme value distributed error terms, ε, and decide to move based on (2.16).[30] It is important to note that the correct distribution to draw the preference shocks from is the standard type 1 extreme value distribution with dispersion parameter of one as this is the distribution implied by the estimation of the normalized value functions. This is shown in Appendix A. When the household chooses to move, then the same draw of ε determines the subsequent location d_1 according to the rule

$$d_1 = \underset{j \in J}{\operatorname{argmax}} \left[v^{\tau}_{j,1} + \varepsilon_{j,1} \right]. \tag{2.22}$$

When the household has chosen to move to a given location, the household chooses how much to invest in the housing market. The investment will be based on the average housing investment of the type in the new location, but conditioning on the amount being below the borrowing constraint. If the borrowing constraint is binding the household will invest what the constraint allows. Then at the end of period 1 the household updates its type based on the potential moving cost implied wealth reduction and the appreciation from housing ownership where the appreciation

[30] To limit the model to one of housing ownership, the simulations abstract from the outside option and the entry and exit decisions.

rate is allowed to vary across the price distribution within municipality.[31] Then the period 2 simulations begin and iterate the first period simulations until the end of the sample at time T. Appendix B includes a time line of the simulation strategy.

For every point in the simulation housing investments, moving costs and property taxes are determined. As for prices the tax ratio of property to land valuations is also allowed to differ in the same respect across the price distribution of housing within each municipality.

Capital gains are expected to differ across initial location, initial wealth type, and across years. All of these aspects can be measured by the baseline simulations. If there are differences in the capital gains across the wealth distribution then there could be support for the claim that the dispersion might spur to wealth inequality. Therefore, some counterfactual simulations are conducted in which the institutional setting is altered. It is important to note that the estimation of the model is based on equilibrium prices and sorting behavior, so any non-marginal change to the housing market, which would lead to a new price equilibrium and a different sorting outcome, can not be encompassed in the counterfactual simulations. However, the counterfactual simulations provide some bounds for the minimal effects on the capital gains distribution. The simulations are based on a single household that is referred to as the treated household, which is not big enough to alter the equilibrium.

The exercise is concluded by decomposing capital gains calculated from the simulations to obtain the wealth effect, where I control for household education type, municipality fixed effect, and the general movements on the housing market by year fixed effects. There are large differences in the number of households of each type that inhabit the different municipalities. Therefore the entries in the decomposition are weighted by the respective frequency in the initial period.

2.5 Empirical Results

The model is estimated using 10 ordered wealth bins, in between which the household can transition endogenously when housing wealth increases or decreases. Additionally, there are 9 education types for which there is assumed to be no inter type transition across time, which means that the model could be estimated separately for each education type. However, the model is estimated jointly and moving cost parameters are restricted to be constant across education groups. It should be noted that the marginal value of wealth as measured by the financial moving costs are allowed some flexibility by including wages that differ across education types.

[31] The simulations do not allow for exogenous type changes such as e.g. Hviid, Timmins, and Vejlin (2015) as the focus is to isolate the wealth implications of housing only.

Moving Costs

The estimation of the moving cost parameters is conducted by matching expected and observed moves in the data by maximum likelihood. The estimated moving cost parameters are presented in Table 2.2 where the financial moving cost parameters are multiplied by 4% of house value as this is the average total financial moving cost, which is treated as observed.

Table 2.2. Financial and Psychological Moving Cost Parameteres

Financial, $\hat{\gamma}_f$		Psychological, $\hat{\gamma}_p$	
Constant	88.1055	Constant	4.9830
	(0.0940)		(0.0020)
Income	2.5275	Wealth	0.4385
	(0.1145)		(0.0013)
		Couple	-0.2823
			(0.0023)
Observations	22,460,778		

Prices, wealth, and income measured in millions DKK.
Standard errors in parenthesis are obtained numerically.

The financial moving costs, which are measured in units of lifetime utility, are positive and slightly increasing in income. The parameters are such that there is no type where the moving costs are negative. Psychological moving costs are positive, increasing in wealth, and lower if the household consists of a couple.[32]

One important observation is that the financial moving costs provide a link between wealth and lifetime value, such that wealth can be determined in units of value and vice versa. Taking a representative household with mean characteristics living in an average house valued 1,740,401 DKK this means that the financial moving costs would constitute 69,616 DKK and the psychological moving costs can be converted to a monetary cost of 57,060 DKK, which corresponds to a total cost of 7.28% of the house value. For comparison, reallocating housing wealth from the municipality with the lowest mean appreciation rate to the municipality with the highest mean appreciation rate would increase appreciation with 6.31 percentage points, enabling the household to be compensated for the flow cost and wealth reduction within two periods.

Additionally, it is important to note that the interpretation of the random shock to current location, ζ_t, can be interpreted as either a taste shock or a shock to moving

[32] The negative impact of being a couple on the psychological moving costs can be explained by these housholds often include kids, motivating up-sizing and down-sizing. Additionally, there is a potential selection effect of single-housholds choosing to enter the housing market.

costs without any possibility to distinguish between the two.[33] Interpreting the shock as a shock to moving costs, households that actually move will on average face significantly lower moving costs than those faced by the average household, as a negative shock to moving costs improves the likelihood of a move decision. Suppose that a household living in municipality j chooses to move and moves to municipality k, then the expected moving cost will be $\bar{c}_{i,t} = c_{i,t} - E_t\left[\varepsilon_{i,j,t} - \zeta_{i,t} | d_{i,t} = k\right]$ where the cost term has been identified. From Kennan (2008) it follows that

$$E_t\left[v^c_{k,t} + \varepsilon_{i,t} - v_{j,t} - \zeta_{i,t} | d_{i,t} = k\right] = -\frac{\log(P(y_{i,t} = 1, j))}{1 - P(y_{i,t} = 1, j)}, \tag{2.23}$$

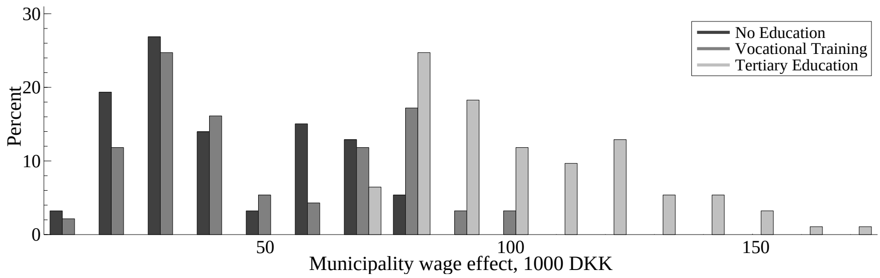
which is a result derived from the Independence of Irrelevant Alternatives property of the model. Assuming that the household chooses to move to a municipality in which the value gain equals the flow costs, then using the average moving probability the average moving cost is 4,973 DKK less than the estimated moving cost. This lowers the average flow costs to 6.99% worth of the housing value compared to 7.28% before.

Wage Differentials

In the third stage of the estimation, the wage equation in (2.20) is estimated to determine the wage differentials across municipalities for the three different educational categories. This produces 279 municipality permanent wage effects and 2.63 million individual permanent effects of which the standard deviations are 34,734.5 DKK and 155,271.9 DKK in 2011 Danish kroner, respectively. Figure 2.5 shows a histogram of the distribution of the wage differentials of the three educational groups for increments of 10,000 DKK, where the effects are measured in differences from the lowest permanent effect. The estimation controls for annual differences along with age and age squared, which are common controls in the labor market literature for experience.

The figure reveals that there are large differences between the three groups suggesting that no education individuals get the lowest wages, individuals with vocational training get slightly more, and individuals with tertiary education get remarkable higher wages. Whereas the histogram is fairly smooth for tertiary education with a heavy right tail, the labor market seems to be clearly partitioned in two for the vocational training and the no education type as they have a bi-modal shape. Interestingly, the highest wages for these two groups are found north of Copenhagen and it is the areas to the east of the Great Belt that constitute the right hand side top, whereas individuals residing in Jutland and on Funen in general get lower wages. The pattern for tertiary education wages is smoother suggesting that the local labor

[33] Kennan and Walker (2011) provide a comprehensive analysis of the interpretation of moving costs in dynamic discrete choice models.



Note: Wage differentials are measured in differences from the lowest municipality and type fixed effect.

Figure 2.5. Histogram of municipality wage differentials interacted with education type. Prices are in 1,000s Danish kroner.

markets are better integrated. However, it is still the eastern municipalities that have the highest wages.

With respect to wealth accumulation there is a positive correlation between municipalities with high wage levels and high house price appreciation, which could only spur the capital gains dispersion. When households have more income they can get a larger mortgage loan and invest relatively more in the housing market. On top of this comes that the house prices appreciate more in the same areas, such that there is a positive second order effect.

Capital Gains

The implied capital gains from the simulation exercise in the final stage of the estimation are decomposed in a linear regression in order to investigate how average capital gains are distributed over the initial wealth distribution. Capital gains are highly influenced by geographical differences and differences change over time, especially during the years where the housing bubble was emerging and subsequently collapsing. Therefore the decomposition includes controls for initial municipality and year dummies. As location choices and borrowing constraints might differ for different household and education types the decomposition also allows for differences between household types.

Table 2.3 reports the results from the decomposition for the initial wealth deciles in 1992. The first two columns report the mean initial wealth and housing assets for the wealth deciles. Interestingly, the housing investment at the initial state is not monotonically increasing in wealth and in particular the lowest wealth decile has notably larger investments in the housing market than nearby wealth types. Many factors might explain this non monotonicity, which is not pursued here but it is most likely related to risk seeking behavior and the fact that these households have probably previously been located in one of the higher wealth deciles. The non monotonicity is important to have in mind when the decomposition is interpreted.

Table 2.3. Capital gains and moving costs from baseline simulations for wealth deciles.

Wealth	Initial State		Outcomes		
Decile	Wealth	Housing Assets	Capital Gains	Moving Costs	Total
1^{st}	-581.1	1,072.5	0	0	0
			-	-	-
2^{nd}	-149.4	840.8	-0.665***	0.0357***	-0.701***
			(0.00706)	(0.000166)	(0.00706)
3^{rd}	-13.0	793.7	-0.662***	0.0431***	-0.705***
			(0.00712)	(0.000167)	(0.00713)
4^{th}	103.6	836.6	-0.677***	0.0403***	-0.717***
			(0.00708)	(0.000166)	(0.00708)
5^{th}	257.6	856.5	-0.654***	0.0424***	-0.697***
			(0.00708)	(0.000166)	(0.00708)
6^{th}	429.0	856.6	-0.623***	0.0541***	-0.677***
			(0.00711)	(0.000167)	(0.00712)
7^{th}	607.4	858.8	-0.601***	0.0399***	-0.641***
			(0.00714)	(0.000168)	(0.00714)
8^{th}	827.9	946.6	-0.376***	0.0144***	-0.390***
			(0.00714)	(0.000168)	(0.00715)
9^{th}	1,177.6	1,076.7	0.0228***	-0.0164***	0.0392***
			(0.00715)	(0.000168)	(0.00715)
10^{th}	2,389.6	1,389.6	1.182***	-0.0454***	1.227***
			(0.00717)	(0.000168)	(0.00718)

Based on 2,000 simulations from 8,370 initial types and locations from 1992 to 2011.
Wealth and housing assets are in 1000's 2011 Danish kroner.
Standard errors in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The first of the last three columns in Table 2.3 reports the capital gains across the wealth distribution in percentage points differences from the lowest wealth decile. The same non monotonicity as is seen in the average housing investment arises in the capital gains distribution as the lowest wealth decile receives larger capital gains than many of the higher deciles. However, from the third wealth decile, which has approximately no net wealth, and the lowest initial housing investment, the capital gains are monotonically increasing in wealth type and the difference in capital gains from this type to the highest wealth type is 1.80 percentage points. The increase in capital gains is larger between the higher wealth groups where the average initial wealth and housing asset differences are also largest. The differences in moving costs are reported in the fourth column and the relative moving costs are increasing to the 6^{th} decile and then decreasing. This pattern follows from the structural moving cost parameters, as psychological moving costs are increasing in household wealth and financial moving costs are increasing in house price. Overall the impact from moving costs is limited compared to the capital gains in the third column.[34] In the last column the total gains from housing ownership when costs are subtracted from the capital gains are reported and as high wealth households move less frequently, the difference in total gains is slightly increased from the raw capital gains to 1.94 percentage points. Differences in moving costs are relatively small compared to the differences in the distribution of capital gains and the overall picture is not altered much.

Overall the results show that there are significant differences in capital gains across the wealth distribution which increases the wealth inequality within housing owners. In particular, the highest wealth groups receive substantially larger gains on their housing investment than the lowest wealth types.

The decomposition also identifies the permanent effect on capital gains for different household composition and education types. Table 2.4 shows the capital gains including the impact from moving costs for the decomposition in differences from a household consisting of a single adult with no education. The first row contains households consisting of a single adult and the three rows below report the results for various compositions of two adults.

Table 2.4. Capital gains including moving costs for different household education types.

	No Education	Vocational Training	Tertiary Education
Single Adult	0	0.0775***	0.427***
	-	(0.00784)	(0.00976)
No Education	0.0795***	-	-
	(0.00537)		
Vocational Training	0.175***	0.306***	-
	(0.00557)	(0.00593)	
Tertiary Education	0.628***	0.601***	0.769***
	(0.00784)	(0.00697)	(0.00725)

Based on 2,000 simulations from 8,370 initial types and locations from 1992 to 2011.
Standard errors in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The results show that there are some differences in capital gains between house-

[34]The negative impact of moving costs to the highest wealth deciles is a result of less frequent moves compared to the lowest wealth decile.

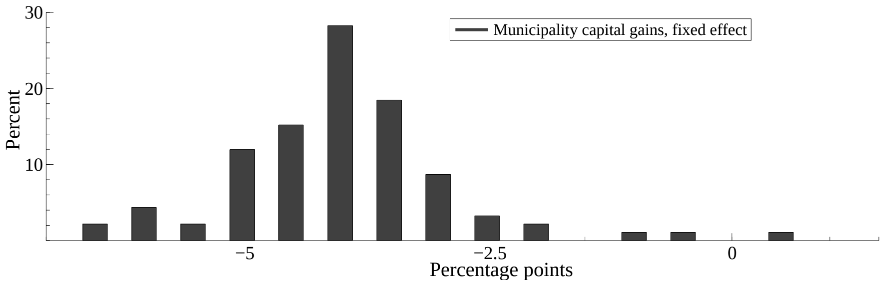


Figure 2.6. The municipality permanent effect in percentage capital gains in deviations from Copenhagen.

hold types. In general, households consisting of a single adult receive lower capital gains on their housing investment. This can to a large extent be explained by differences in borrowing constraints as these single households have only a single source of income to pay for the mortgage and bank loans and therefore can not enter the expensive end of the housing market which appreciate at a higher rate. There are also large differences across education levels, where households with vocational training get higher gains than households with no education, but it is the households with tertiary education that receive the largest gains. This difference can be explained by many factors and of those included in the model are differences in borrowing constraints and location choices. Households with higher education are in general better at making choices that increase their capital gains, which is related to the labor market as tertiary education households are more likely to move to the larger metropolitan areas as this is where labor market conditions are better in terms of the amount of potential jobs. These areas are also those where house prices appreciate the most.

Lastly, to shed some light on the impact of initial location, Figure 2.6 shows the differences in capital gains including moving costs for the municipalities in differences from Copenhagen Municipality. Clearly, most municipalities have much lower capital gains than in Copenhagen and there is a 7.1 percentage points difference from the highest to lowest gains.

Mobility Improvements

The housing market is in many ways an imperfect market and the relatively large financial and psychological moving costs are contributing factors as they limit households to move only when the lifetime value gain is of an order that would outweigh these costs. For the same reason households can not increase or decrease the investment in the housing market without incurring large transaction costs. The first order derivative of the moving probability is globally negative in the moving cost. Hence, reducing e.g. the financial moving costs would lead to more frequent reallocation decisions of households and let households utilize capital gains of households in relatively good locations to step into a higher price decile in a less well performing municipality in terms of house price appreciations. The households in locations with low capital gains prospects would be able to leave for a better location at a higher rate with lower moving costs.

Table 2.5 reports the capital gains distribution for the treated household when the financial moving costs are set to 2% rather than the actual 4% of house value. This corresponds to eliminating the property registration fees from a housing transaction.[35] The capital gains in the first column share the profile from the baseline simulation in Table 2.3 but there is a slight decrease in the difference from the lowest to the highest capital gains in the wealth distribution, i.e. 1.79 percentage points. The largest relative changes occur for the lowest wealth decile as the borrowing constraint is more often binding and the increased moving frequency forces these households to lower their housing investment, which is already relatively large. The second column in Table 2.5 reports the differences in moving costs, where the dispersion across the wealth distribution is increased compared to the baseline model, as more weight is put on psychological moving costs that decrease in wealth. When the two effects are combined in the third column there is a slight increase in dispersion to 1.93 percentage points, but there is no overall difference compared to the baseline model.

The simulation is based on treating a single household with lower financial moving costs thereby abstracting from general price and sorting effects. In an equilibrium setting, improving mobility would increase the demand for housing in e.g. the capital region and increase the supply in rural areas, increasing the geographical dispersion of capital gains and isolating more low wealth households from the expensive locations.

Tax Policy Evaluation

When there are large differences in capital gains across the price and wealth distribution, one policy that might mitigate these differences could be a taxation scheme of housing wealth. This section investigates the current taxation scheme and the freeze of housing taxes that was introduced in 2002, by conducting a baseline simulation from 2002 to 2010 to see how housing wealth taxation can affect the distribution of capital gains. Subsequently, a treated household is simulated to identify an upper

[35] The registrations fees, collected to the state, and brokerage fees are set as a percentage of the investment.

Table 2.5. Capital gains distribution across with lowered financial moving costs.

Wealth		Outcomes	
Decile	Capital Gains	Moving Costs	Total
1^{st}	0	0	0
	-	-	-
2^{nd}	-0.326***	0.0313***	-0.357***
	(0.00485)	(0.000148)	(0.00486)
3^{rd}	-0.257***	0.0331***	-0.290***
	(0.00490)	(0.000149)	(0.00490)
4^{th}	-0.296***	0.0297***	-0.325***
	(0.00487)	(0.000148)	(0.00487)
5^{th}	-0.298***	0.0303***	-0.328***
	(0.00487)	(0.000148)	(0.00487)
6^{th}	-0.257***	0.0330***	-0.290***
	(0.00489)	(0.000149)	(0.00490)
7^{th}	-0.190***	0.0156***	-0.206***
	(0.00491)	(0.000150)	(0.00491)
8^{th}	0.0154***	-0.00542***	0.0209***
	(0.00491)	(0.000150)	(0.00492)
9^{th}	0.432***	-0.0451***	0.477***
	(0.00492)	(0.000150)	(0.00492)
10^{th}	1.472***	-0.0983***	1.570***
	(0.00493)	(0.000150)	(0.00494)

Based on 2,000 simulations from 8,370 initial types and locations from 1992 to 2011.
Standard errors in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

bound on the dispersion of capital gains in the wealth distribution. This exercise investigates how the capital gains distribution would have been affeceted by not passing the property tax freeze.

As described above housing taxes are divided into taxation of land and taxation of the property. This analysis will focus on the national level property tax, which in general constitutes the majority of the tax on housing wealth. The top graph in Figure 2.7 shows a plot of the effective tax rate out of total housing value for the wealth deciles with and without the 2002 property tax freeze. Both distributions start off just above 0.8% with about 0.05 percentage points difference from the lower to the highest wealth types as more expensive housing units often have a larger fraction of the total housing value embedded in the land valuation. As time passes the effective tax rate is almost constant in the case without a tax freeze, whereas it is dramatically decreasing under the scheme with a property tax freeze and the dispersion widens and reaches 0.15 percentage points in 2010. Thereby, the taxation scheme is adding to the inequality of net housing income in the wealth distribution. The taxation scheme is progressive in housing value as value that exceeds 3 million in 2012 DKK is taxed by 3% which could help reducing the inequality in capital gains. The middle part in Figure 2.7 shows the effective tax rate when this progressive component is added to the tax scheme. Only households of the highest wealth deciles have housing assets in 2002 that exceed the limit and, hence, the effective tax rate for the highest wealth decile is increased slightly to be placed in the middle of the distribution, but overall this progressive component has had a limited effect under the frozen taxation scheme. On the contrary, in the counterfactual world, where the tax freeze is not introduced, the progressive component affects all types in 2007 at the peak of the bubble period and the highest wealth type is affected in all periods leading to a difference in the effective tax of 0.5% in 2007 and 0.3% in 2010 where the households with larger housing investments pay more in taxes.

The bottom part in Figure 2.7 shows the mean effective tax rates across municipalities in 2010 under the freeze of housing taxes. The figure shows that there are large differences in the effective tax rate and municipalities that pay the highest tax are by construction also those that have gained the least on the house price appreciations over the period 2002 to 2010.[36]

In order to analyze the housing taxes and implications for capital gains, the taxes, which by nature are user costs, are treated as wealth components.

Table 2.6 reports the initial states from early 2002 and capital gains over the wealth distribution for the four regimes in Figure 2.7. The differences in wealth are substantially larger in 2002 than in 1992 and also the values of the housing asset are larger for all groups. The lowest wealth decile has a larger investment in the housing market than nearby deciles, but the difference is smaller than in 1992. The first column of results reports capital gains in a simulation that abstracts from housing taxes and include moving costs, comparable to the last column in Table 2.3. The difference from the household with the least housing assets to the top decile of the wealth distribution is 2.33 percentage points and larger than over the period from 1992 to 2010. However, the capital gains increase almost linearly, whereas the slope is steeper for the top deciles over the entire sample. The fourth and fifth column include the flat tax from the tax scheme with the tax freeze and without, respectively. Interestingly, the dispersion of capital gains has increased under the tax freeze scheme, which has then contributed to the inequality in capital gains by 0.05 percentage points, whereas a tax regime without a frozen property tax would not have increased the dispersion significantly. The sixth and seventh column report the

[36]Within some municipalities there exist housing units for which the effective tax rate has increased over the sample. However, most of these are deleted in the sample selection.

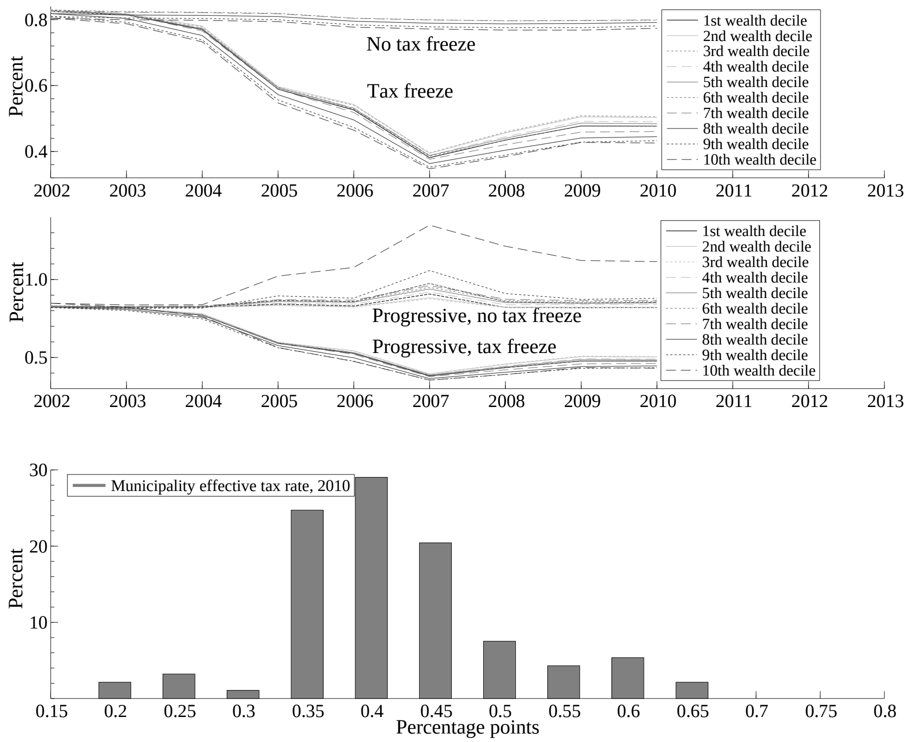


Figure 2.7. Effective property tax rates property after the tax freeze and counterfactual taxes for a treated household excluding progressive profile (top), effective property tax rate including the progressive profile (middle), and municipality differences in effective tax rate under the tax freeze as of 2010 (bottom).

decomposition when the progressive component is added to the taxation schemes in column 4 and 5. Under the tax freeze, the relative gains of the highest wealth types are only lowered by an insignificant amount. On the contrary, if the housing tax was not frozen from 2002 the dispersion of capital gains would be lowered relatively much from 2.39 percentage points to 2.12 percentage points.

It is important to note that the simulations of the treated household do not account for the stabilizing effect that increased housing taxes would have on the house price distribution. But the difference of 2.44 percentage points in the simulation of the flat tax system can be viewed as an upper bound on the dispersion of capital gains of the wealth deciles and to the extent that flat housing taxation can not sufficiently decrease the inequality in capital gains, a progressive taxation scheme seems effective

Table 2.6. Simulations of capital gains that control for housing taxes under different schemes.

Initial states			Flat Tax		Progressive Tax	
Wealth	Housing	Baseline	Freeze	No Freeze	Freeze	No Freeze
-563.6	1,463.8	0	0	0	0	0
		-	-	-	-	-
-141.8	1,311.5	-0.560***	-0.572***	-0.561***	-0.572***	-0.563***
		(0.0166)	(0.0167)	(0.0166)	(0.0167)	(0.0166)
49.5	1,346.3	-0.409***	-0.417***	-0.409***	-0.417***	-0.415***
		(0.0168)	(0.0168)	(0.0168)	(0.0168)	(0.0167)
252.9	1,449.7	-0.117***	-0.119***	-0.116***	-0.119***	-0.135***
		(0.0167)	(0.0167)	(0.0167)	(0.0167)	(0.0166)
479.1	1,524.1	0.0116	0.0137	0.0137	0.0129	-0.0129
		(0.0167)	(0.0167)	(0.0167)	(0.0167)	(0.0166)
729.4	1,600.5	0.292***	0.301***	0.296***	0.299***	0.259***
		(0.0167)	(0.0168)	(0.0167)	(0.0168)	(0.0167)
1,025.9	1,723.3	0.728***	0.747***	0.732***	0.744***	0.689***
		(0.0168)	(0.0168)	(0.0168)	(0.0168)	(0.0168)
1,411.5	1,926.6	1.330***	1.361***	1.336***	1.358***	1.294***
		(0.0168)	(0.0169)	(0.0168)	(0.0169)	(0.0168)
2,018.5	2,276.4	1.758***	1.801***	1.767***	1.796***	1.699***
		(0.0168)	(0.0169)	(0.0168)	(0.0169)	(0.0168)
3,963.8	2,993.0	1.774***	1.816***	1.783***	1.803***	1.563***
		(0.0169)	(0.0169)	(0.0169)	(0.0169)	(0.0168)

Based on 2,000 simulations from 8,370 initial types and locations from 2002 to 2011.
Wealth and housing assets are in 1000's 2011 Danish kroner.
Standard errors in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

to decrease the dispersion further.

2.6 Conclusions

This paper investigates the distribution of capital gains from housing ownership across the wealth distribution in Denmark over the period 1992 to 2011. The data applied is the population-wide Danish register data which includes detailed information on housing units and individual wealth, income, and socio-economic details. The range and quality of the data are exceptional compared to what has previously been applied in related studies. The basis for the evaluation of the capital gains distribution is a dynamic residential sorting model that takes account of geographical variations in anything that affects municipality quality including expected capital gains and price. The estimation of the model allows for preference heterogeneity between household types, where types are determined from wealth levels and educational attainments of household members. To accommodate the dynamic nature of the housing market, household decisions are explicitly allowed to be influenced by financial and psychological moving costs and households that change location are subject to a borrowing constraint that limits the access to certain parts of the housing market.

An advantage of estimating a structural model is that it identifies the structural parameters that affect household decisions. When the decision rules are known, the model can be used to project how a household would behave in a counterfactual world and ultimately determine how this would affect the capital gains of the household.

In a simulation exercise it is found that there are substantial differences in capital gains across the wealth distribution that enhance wealth inequality. Financial transaction costs are found to have a limited effect on the dispersion of capital gains. Additionally, the simulation identifies large geographical variations in the distribution of capital gains, which has only been magnified by the current Danish taxation scheme. A counterfactual analysis supports that such a progressive taxation scheme could have an attenuating effect on the dispersion of capital gains and an appropriately formed progressive taxation scheme could prove inequality mitigating.

Acknowledgments

The author gratefully acknowledges support from Aarhus University, Department of Economics and Business Economics as well as from Realdania and The Knowledge Center for Housing Economics and from CREATES - Center for Research in Econometric Analysis of Time Series (DNRF78), funded by the Danish National Research Foundation. Comments from Rune Vejlin, Christopher Timmins, Tom Engsted and participants at the Danish Graduate Programme in Economics Workshop and at the Labor and Public Policy Seminars are also gratefully acknowledged.

References

ABBRING, J. H. AND J. J. HECKMAN (2007): "Econometric Evaluation of Social Programs, Part III: Distributional Treatment Effects, Dynamic Treatment Effects, Dynamic Discrete Choice, and General Equilibrium Policy Evaluation," in *Handbook of Econometrics*, ed. by J. Heckman and E. Leamer, Elsevier, vol. 6 of *Handbook of Econometrics*, chap. 72.

ABOWD, J. M., F. KRAMARZ, AND D. N. MARGOLIS (1999): "High Wage Workers and High Wage Firms," *Econometrica*, 67, 251–333.

AGUIRREGABIRIA, V. AND A. NEVO (2010): "Recent Developments in Empirical IO: Dynamic Demand and Dynamic Games," MPRA Paper 27814, University Library of Munich, Germany.

ANDERSEN, S., J. CAMPELL, K. M. NIELSEN, AND T. RAMADORAI (2014): "Inattention and inertia in household Finance: Evidence from the danish mortgage market," Working paper, Harvard University.

BAYER, P., R. MCMILLAN, A. MURPHY, AND C. TIMMINS (2015): "A Dynamic Model of Demand for Houses and Neighborhoods," *Econometrica*, Forthcomming.

BERRY, S. T. (1994): "Estimating Discrete-Choice Models of Product Differentiation," *RAND Journal of Economics*, 25, 242–262.

BERRY, S. T. AND P. A. HAILE (2009): "Nonparametric Identification of Multinomial Choice Demand Models with Heterogeneous Consumers," Cowles Foundation Discussion Papers 1718, Cowles Foundation for Research in Economics, Yale University.

BISHOP, K. (2012): "A Dynamic Model of Location Choice and Hedonic Valuation," Working Paper, Olin Business School.

BROWNING, M., M. GØRTZ, AND S. LETH-PETERSEN (2013): "Housing Wealth and Consumption: A Micro Panel Study," *The Economic Journal*, 123, 401–428.

CAMPBELL, J. Y. (2013): "Mortgage Market Design," *Review of Finance*, 17, 1–33.

CAMPBELL, J. Y. AND J. F. COCCO (2007): "How do house prices affect consumption? Evidence from micro data," *Journal of Monetary Economics*, 54, 591–621.

CARROLL, C. D., M. OTSUKA, AND J. SLACALEK (2011): "How Large Are Housing and Financial Wealth Effects? A New Approach," *Journal of Money, Credit and Banking*, 43, 55–79.

CASE, K. E., J. M. QUIGLEY, AND R. J. SHILLER (2005): "Comparing Wealth Effects: The Stock Market versus the Housing Market," *The B.E. Journal of Macroeconomics*, 5, 1–34.

ENGSTED, T., S. J. HVIID, AND T. Q. PEDERSEN (2015): "Explosive bubbles in house prices? Evidence from the OECD countries." *Journal of International Financial Markets, Institutions, and Money*, Forthcoming.

GEMICI, A. (2007): "Family Migration and Labour Market Outcomes," Working Paper, University of Pennsylvania.

HOTZ, V. J. AND R. A. MILLER (1993): "Conditional Choice Probabilities and the Estimation of Dynamic Models," *The Review of Economic Studies*, 60, pp. 497–529.

HVIID, S. J., C. TIMMINS, AND R. VEJLIN (2015): "Valuation of Non-Traded Amenities in a Dynamic Demand Model," Working paper, Aarhus University.

IACOVIELLO, M., R. H. EDELSTEIN, AND K.-H. KIM (2004): "House Prices and Collateral Constraints: A Structural Econometric Analysis," *Journal of Housing Economics*, 1, 304–320.

KENNAN, J. (2008): "Average Swithcing Cost in Dynamic Logit Models," Unpublished Manuscript, University of Wisconsin.

KENNAN, J. AND J. R. WALKER (2011): "The Effect of Expected Income on Individual Migration Decisions," *Econometrica*, 79, 211–251.

KLEVEN, H. J., M. B. KNUDSEN, C. T. KREINER, S. PEDERSEN, AND E. SAEZ (2011): "Unwilling or Unable to Cheat? Evidence From a Tax Audit Experiment in Denmark," *Econometrica*, 79, 651–692.

KUMINOFF, N. V., V. K. SMITH, AND C. TIMMINS (2013): "The New Economics of Equilibrium Sorting and Policy Evaluation Using Housing Markets," *Journal of Economic Literature*, 51, 1007–62.

LETH-PETERSEN, S. (2010): "Intertemporal Consumption and Credit Constraints: Does Total Expenditure Respond to an Exogenous Shock to Credit?" *American Economic Review*, 100, 1080–1103.

MAGNAC, T. AND D. THESMAR (2002): "Identifying Dynamic Discrete Decision Processes," *Econometrica*, 70, 801–816.

MCFADDEN, D. (1974): "Conditional logit analysis of qualitative choice behavior," Tech. Rep. 105-142, Frontiers in Econometrics, New York.

——— (1978): "Modelling the Choice of Residential Location," Cowles Foundation Discussion Papers 477, Cowles Foundation for Research in Economics, Yale University.

MIAN, A. AND A. SUFI (2011): "House Prices, Home Equity-Based Borrowing, and the US Household Leverage Crisis," *American Economic Review*, 101, 2132–56.

PIKETTY, T. (2011): "On the Long-Run Evolution of Inheritance: France 1820–2050," *The Quarterly Journal of Economics*, 126, 1071–1131.

——— (2014): *Capital in the Twenty-First Century*, Harvard University Press.

PIKETTY, T. AND E. SAEZ (2013): "A Theory of Optimal Inheritance Taxation," *Econometrica*, 81, 1851–1886.

PIKETTY, T. AND G. ZUCMAN (2013): "Capital is Back: Wealth-Income Ratios in Rich Countries, 1700-2010," CEPR Discussion Papers 9588, C.E.P.R. Discussion Papers.

RANGVID, J., A. G. FINNSTRUP, P. MØGELVANG-HANSEN, H. F. JENSEN, J. THOMSEN, P. SCHUUTZE, J. GALBO, C. ØLGAARD, N. K. FREDERIKSEN, AND ET AL. (2013): "Den finansielle krise i Danmark: Årsager, konsekvenser og læring." Tech. rep., Erhvervs of Vækstministeriet.

ROBACK, J. (1982): "Wages, Rents, and the Quality of Life," *Journal of Political Economy*, 90, 1257–78.

ROGNLIE, M. (2014): "A note on Piketty and diminishing returns to capital," Working paper, Massachusetts Institute of Technology.

——— (2015): "Deciphering the fall and rise in the net capital share," Conference Draft, Brookings papers on Economic Activity.

RUST, J. (1987): "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica*, 55, pp. 999–1033.

——— (1994): "Structural estimation of markov decision processes," in *Handbook of Econometrics*, ed. by R. F. Engle and D. McFadden, Elsevier, vol. 4 of *Handbook of Econometrics*, chap. 51, 3081–3143.

SINGLETON, J. (2013): "Sorting Charles Tiebout: The Construction and Stabilization of Postwar Public Good Theory," Tech. rep., Center for the History of Political Economy Working Paper Series 2013-20.

TIEBOUT, C. M. (1956): "A Pure Theory of Local Expenditures," *Journal of Political Economy*, 64, 416.

ZUCMAN, G. (2013): "The Missing Wealth of Nations: Are Europe and the U.S. net Debtors or net Creditors?" *The Quarterly Journal of Economics*, 128, 1321–1364.

Appendix A: A Note on Conditional Choice Probabilities

The conditional choice probability of an individual to choose municipality j is

$$P_j = \text{Prob}(\varepsilon_k < \varepsilon_j + \text{v}_j - \text{v}_k \ \forall k \neq j),$$

where ε is identically and independent type 1 extreme value distributed. The distribution has two parameters where α is an additive location parameter and $\lambda > 0$ is the multiplicative dispersion parameter. The location parameter can take any value without affecting the conditional choice probability as only differences matter for choices. However, for fixed probabilities, the dispersion parameter affects values in a multiplicative way.[37] When it comes to simulations, the dispersion parameter matters for the draws of ε. Normalize values and idiosyncratic shocks to get a unit dispersion parameter,

$$P_j = \text{Prob}\left(\frac{\varepsilon_k}{\lambda} < \frac{\varepsilon_j + \text{v}_j - \text{v}_k}{\lambda} \ \forall k \neq j\right).$$

Defining $\text{v}_j/\lambda = v_j$ and conditioning on a realization of ε_j, the conditional probability can be expressed as the product of the cumulative distribution of $\varepsilon_k \ \forall k \neq j$, i.e. $P_j|\varepsilon_j = \Pi_{k\neq j} e^{-e^{-(\varepsilon_j/\lambda + v_j - v_k)}}$. Then integrating the density weighted values of ε_j yields the conditional choice probability

$$\begin{aligned} P_j &= \int_{-\infty}^{\infty} \left(\Pi_{k\neq j} e^{-e^{-(\varepsilon_j/\lambda + v_j - v_k)}}\right) \frac{1}{\lambda} e^{-e^{\varepsilon_j/\lambda}} e^{-\varepsilon_j/\lambda} d\varepsilon_j \\ &= \int_{-\infty}^{\infty} \left(e^{-e^{-\varepsilon_j/\lambda} \sum_j e^{(v_j - v_k)}}\right) \frac{1}{\lambda} e^{-\varepsilon_j/\lambda} d\varepsilon_j. \end{aligned}$$

If $s = e^{-\varepsilon_j/\lambda}$ implying $ds = -\frac{1}{\lambda} e^{-\varepsilon_j/\lambda} d\varepsilon_j$ and noting $\lim_{\varepsilon_j \to -\infty} s = \infty$ and $\lim_{\varepsilon_j \to \infty} s = 0$ then

$$\begin{aligned} P_j &= \int_0^{\infty} e^{-s\sum_k e^{(v_j - v_k)}} ds \\ &= \left[\frac{e^{-s\sum_k e^{(v_j - v_k)}}}{-\sum_k e^{(v_j - v_k)}}\right]_0^{\infty} = \frac{1}{\sum_k e^{-(v_j - v_k)}}. \end{aligned}$$

Hence, the values, v rather than v, obtained from inverting the conditional choice probabilities in the estimation are those that correspond to idiosyncratic shocks drawn from a standard type 1 extreme value distribution, i.e. with location zero and dispersion one. This implies that the correct distribution to draw shocks from in the simulations will be $T1EV(0,1)$.

[37] This holds for all cases except the case when $v_i = v_j \ \forall i, j$, which is uninteresting as any such value would lead to $P_k = P_j = \frac{1}{J}$.

Appendix B: Outline of Simulation Strategy

The timing of my simulation strategy for simulation i is as follows:

t_0 - Choose initial type τ_0 and initial municipality $d_{i,t_0} \in \{1, \ldots J\}$.
- Draw vector ε_{i,t_0}.
- Choose to stay or move from (2.15).
  - if 'stay' then $d_{i,t_1} = d_{i,t_0}$.
  - if 'move' then:
    - choose $d_{i,t_1} \in \{1, \ldots, J\}$ based on (2.22).
    - choose housing investment: $\min\left(b_{i,t_0}, \text{Avg. housing assets of } \tau_0\right)$.
- Update type to τ_1
  - Endogenous: Wealth loss from moving costs: $-0.04 \cdot h_{i,t_0}$.

t_1 - Observe type τ_1 and municipality d_{i,t_1}.
- Update type after capital gains accumulation in period t_1 based on price decile.
- Draw vector ε_{i,t_1}.
- Choose to stay or move from (2.15).
  - if 'stay' then $d_{i,t_2} = d_{i,t_1}$.
  - if 'move' then:
    - choose $d_{i,t_2} \in \{1, \ldots, J\}$ based on (2.22).
    - choose housing investment: $\min\left(b_{i,t_1}, \text{Avg. housing assets of } \tau_1\right)$.
- Update type to τ_2
  - Endogenous: Wealth loss from moving costs: $-0.04 \cdot h_{i,t_1}$.

t_2 - Repeat as for t_1

Continue the procedure until the end of the sample at time T.

CHAPTER 3

VALUATION OF NON-TRADED AMENITIES IN A DYNAMIC DEMAND MODEL

Simon J. Hviid
Aarhus University and CREATES

Christopher Timmins
Duke University

Rune M. Vejlin
Aarhus University

Abstract

Using unique population-wide Danish register data with precise measures of households' wealth, income, and socio-economic status, we specify and estimate a dynamic structural model of residential neighborhood demand. Our model includes moving costs, forward looking behavior of households, and uncertainty about the evolution of neighborhood attributes, wealth, income, house prices, and family composition. We estimate marginal willingness to pay for non-traded neighborhood amenities with a focus on air pollution. We allow household willingness to pay to vary in household characteristics and argue that low wealth and low income households are borrowing constrained. Our application finds that the dynamic approach adjusts for various biases relative to a static approach and that the willingness to pay of households who are likely borrowing constrained are much more sensitive to changes in wealth.

JEL Classification: D31, D63, D91, H20, R00, R21, R51
Keywords: income distribution, dynamic discrete choice models, residential sorting

3.1 Introduction

An extensive literature has developed around the residential location decision. As one of the largest and most important purchases most households will ever make, the home purchase decision will be carefully considered. As such, economists have used these decisions to learn about the attributes of local public goods and amenities that are capitalized into house values, and to build models of how location decisions might be altered by policy changes.[1] Indeed, models of residential sorting have become ubiquitous in non-market valuation and local public finance.[2]

Most who have ever bought or sold a house would agree that the process is characterized by large transactions costs (both financial and psychological). Some fraction of the house value is paid over to a realtor, and one is forced to box all of one's possessions to be fit on to the back of a truck. The costs of searching for a new home may be substantial (particularly for long-distance moves). Faced with all of these costs, we would expect that households would want to minimize the number of moves that they make. Combine this with a world in which amenities and local public goods change over time (sometimes very rapidly), and we would expect homebuyers to be forward looking with respect to the way attributes evolve. Add to this the fact that, for many households, their home is their primary store of wealth. Evolving amenities imply evolving wealth if those amenities are capitalized into house prices. We would therefore expect homebuyers to consider the evolution of price.

With all this in mind, it is surprising that the literatures dealing with residential sorting have, with few exceptions, ignored forward looking behavior, typically only considering current amenity and local public good values. Home buyers are usually treated as renters, where annualized price is treated as a flow cost of occupancy; there is subsequently no value to rising house prices. The obvious reasons for this omission are twofold. On the one hand, computational costs associated with traditional dynamic programming techniques become prohibitive in the residential sorting context. Hundreds of neighborhoods each have multiple attributes, each of which can take multiple values. Together, this leads to a prohibitively large state space for traditional dynamic programming approaches. The consumer durables literature

[1] Some examples that use crime data include Pope (2008) and Linden and Rockoff (2008). Chay and Greenstone (2005) and Bayer, Keohane, and Timmins (2009) take very different methodological approaches to the valuation of air quality. Black (1999) uses district boundary discontinuities to value school quality, and Davis (2004) values health perceptions arising from a cancer cluster. Greenstone and Gallagher (2008) and Gamper-Rabindran and Timmins (2013) come to different conclusions about the value of remediating Superfund sites. Palmquist (2006) provides a review of the property value hedonics literature.

[2] See Kuminoff, Smith, and Timmins (2013) for a review.

in the industrial organization field has made some progress in this dimension, but has typically done so by imposing strong restrictions on the way evolving attributes enter into utility.[3]

The second factor that has stalled progress in the modeling of residential location decisions in a dynamic context is data. Identifying preferences in a model of forward looking homebuyers inherently requires 'dynamic' information - at a minimum, we need to know when a homeowner sells a property after he or she buys it. Ideally, we would like to know the house that he or she buys after making a sale. Many residential decisions are driven by important life events (e.g. a marriage or divorce, the birth of a child or becoming an 'empty nester'). Decisions about house purchases, especially to the extent that the house purchase is viewed as an investment, are driven by evolution of household wealth. Finding a data set that tracks family units over multiple residential decisions, keeps track of major family events, and follows family wealth (in addition to income) is not easy. In this paper, we overcome these constraints by using a unique data source from Denmark.

The data set is based on the population-wide individual level Danish registers. We merge individual level data from several data sources. Almost all of the data sources originate from public registers and thus the data quality is considered very high. These include detailed socio-economic information, housing unit characteristics, information on sales prices, public valuations of houses, and neighborhood characteristics from geographical data, air pollution data, and data on crime convictions. Through unique identifiers of families and housing units linked with address information, we are able to follow household's characteristics and locations over time. In an international comparison, such a comprehensive data set is unique.

With these data, we are well-positioned to study the residential location decision accounting for forward looking agents and a suite of important evolving state variables, including wealth and children. In particular, we adapt the method described in Bayer, McMillan, Murphy, and Timmins (2015) to account for the special features of these data. Our application focuses on house size and a number of neighborhood attributes including air quality (measured by particulate matter, carbon monoxide and ozone concentrations) and public safety (measured by the presence of convicted criminals residing in the neighborhood). Results suggest a bias in the static model similar to what is found in previous studies. We go beyond that simple interpretation of the results, however, and focus on the special role in the model played by wealth.

[3] See, for example, Melnikov (2013), Gowrisankaran and Rysman (2012), Hendel and Nevo (2006) for applications to computer printers, digital camcorders, and household stocks of laundry detergent, respectively. Schiraldi (2011) extends Gowrisankaran and Rysman (2012) to account for re-sale in the used car market, and is therefore similar to our analysis in many ways. Our approach differs from this literature primarily in that we develop a multi-stage estimator that exploits unique features of the housing market. This allows us to avoid ever having to solve the homebuyer's dynamic programming problem explicitly (which requires other, more restrictive assumptions in the papers cited above).

In particular, we identify individuals who, because of their wealth and income combination, are likely constrained in their residential location decision. These individuals may not have the resources to make a down payment on a house with nice air quality; similarly, they may not have the income stream to borrow to enable that choice. We find that, for individuals under these constraints, the marginal effect of wealth on marginal willingness to pay to avoid pollution is much greater than for individuals who are likely constrained, suggesting that such constraints may be important.

Focusing on the unique role of family structure in our data, we similarly demonstrate the bias in the estimated value of house size that arises from failing to account for forward looking expectations with respect to children (in particular, expectations regarding the exit of grown children from the household). Failing to account for those expectations leads to an overstatement of the value on house size of over 25%.

The remainder of this paper proceeds as follows. Section 3.2 sets up the dynamic model of residential location. Section 3.3 describes the comprehensive data set that we use for the estimation described in Section 3.4. Section 3.5 presents our empirical results and Section 3.6 concludes.

3.2 The Dynamic Demand Model

This section will formally set up the dynamic model of demand. The basic unit in the model is a household. The aim is to identify the flow utility of households living in specific neighborhoods from which we can obtain estimates of marginal willingness to pay for a selection of locally non-traded amenities.

At time t every household makes a choice between staying in their current housing unit or selling and moving to another unit, potentially in a new neighborhood. Denote the binary move-stay decision variable of a household by $y_{i,t}$, where the subscript i denotes the household in question. The decision variable identifies whether the household will move or not. The choice variable $d_{i,t}$ specifies in which neighborhood the household chooses to reside. Let J be the number of neighborhoods and let 0 denote the outside option, i.e. if a household chooses to leave the Danish housing market and then rent rather than own.[4] Hence, the choice set is $d_{i,t} = j \in \{0, 1, \ldots, J\}$ and the number of choice options are $J + 1$. The decision rule is simply such that if a household observes a lifetime value gain from moving, then the household will move to the neighborhood with the highest expected lifetime value.

For each point in time there are two observable state vectors $\{x_{j,t}, z_{i,t}\}$. Here, j identifies neighborhood and $x_{j,t}$ is a vector of neighborhood specific attributes that are allowed to evolve over time. Such attributes will cover air quality, and neighbor

[4]This choice could embed a large set of options such as moving to a different country or if the household ceases to exist. We do not explicitly model any of these decisions and hence limit the model to one of home owner behavior with a uniform renting based alternative.

characteristics. The vector $z_{i,t}$ contains household specific characteristics (i.e., wealth, income, number of children, and housing unit value).

In addition to the three observed state variables, there are two unobservable variables in the model. The first is the unobserved quality of the neighborhood, which is a scalar summary of a collection of neighborhood characteristics that are not observed by the econometrician, denoted by $\xi_{j,t}$. The other is an idiosyncratic stochastic shock to the flow utility of a household associated with each neighborhood, $\varepsilon_{i,j,t}$.

We define $\Omega_{i,t} \equiv \{z_{i,t}, x_t, \xi_t, d_{i,t-1}, z_{i,t-1}, x_{t-1}, \xi_{t-1}, d_{i,t-2}, \dots\}$ to be a set that contains all information about the household and neighborhoods that possibly helps to predict future characteristics of the household and neighborhoods.

Neglecting the cost of moving, let $u_{i,j,t} = u\left(z_{i,t}, x_{j,t}, \xi_{j,t}, \varepsilon_{i,j,t}\right)$ be the flow utility function of the individual household from living in neighborhood j. Then, let $c_{i,t} = c(z_{i,t}, x_{h_{i,t},t})$ denote the flow costs associated with a move decision. The costs of moving are allowed to cover both financial and psychological factors that affect the flow utility of the household. The moving costs are zero if the household chooses not to move and it consists only of psychological moving costs if the household is a first-time buyer, as the financial transaction costs are paid by sellers. As in Bayer, McMillan, Murphy, and Timmins (2015) this specification only allows the moving costs to be a function of the characteristics of the neighborhood that is left and household characteristics. This assumption implies that we can define the corrected flow utility function by $u^c_{i,j,t} = u_{i,j,t} - c_{i,t} I_{[y_{i,t} \neq 0]}$. Hereby it is implicitly assumed that the utility is additively separable in the cost of moving.

The transition probabilities of the state variables are Markovian and denoted by

$$q_{i,t} = q(\Omega_{i,t+1}, \varepsilon_{i,t+1} | \Omega_{i,t}, \varepsilon_{i,t}, d_{i,t}). \tag{3.1}$$

Along with the discount factor, β, these components make the primitives of the model, $\{u, c, q, \beta\}$. With a constant discount factor the household's infinite horizon objective function is the expected discounted sum of all future flow utilities,

$$U_{i,t} = u^c_{i,j,t} + E_t \left[\sum_{\tau=t+1}^{\infty} \beta^{\tau-t} u^c(x_{j,\tau}, \xi_{j,\tau}, \varepsilon_{i,\tau}, d_{i,\tau-1}) \,\middle|\, \Omega_{i,t}, \varepsilon_{i,t}, d_{i,t} \right]. \tag{3.2}$$

Note that this is an autonomous problem, hence, assuming that $\{d^\star_{i,t}\}_{t=\tau}^{\infty}$ is the optimal control sequence we can define the lifetime value function,

$$V_t(\Omega_{i,t}, \varepsilon_{i,t}) = \max_{\{d_{i,\tau}\}_{\tau=t}^{\infty}} E_t \left[\sum_{\tau=t}^{\infty} \beta^{\tau-t} u^c(x_{j,\tau}, \xi_{j,\tau}, \varepsilon_{i,\tau}) \,\middle|\, \Omega_{i,t}, \varepsilon_{i,t}, d_{i,t} \right]. \tag{3.3}$$

With an infinite time horizon we can drop the time subscript on the value function. At time t the optimal control will yield a total reward over all subsequent periods equal to $V(\cdot)$ and the immediate reward will be the flow utility, such that the Bellman equation for the household becomes

$$V(\Omega_{i,t},\varepsilon_{i,t}) = \max_{d_{i,t}}\left[u^c_{i,j,t} + \beta E_t\left[V\left(\Omega_{i,t+1},\varepsilon_{i,t+1}|\Omega_{i,t},\varepsilon_{i,t},d_{i,t}\right)\right]\right]. \tag{3.4}$$

Given that that the lifetime value function is appropriately bounded, this Bellman equation will be a contraction mapping in $V(\cdot)$ (See Rust (1994) for sufficient conditions).

3.3 Data

The data consists of the full Danish population of households from 1992-2005. We merge micro data from several different data sources. These include detailed socio-economic household information (including detailed accurate wealth information), housing unit characteristics, information on sales values, the tax authorities valuation of houses, neighborhood characteristics from geographical data, air pollution data, and finally data on crime convictions. Being able to link all these data sets across individuals and time is unique in an international comparison and allows us to extend the current modeling setup along several dimensions. As the data consists of multiple sources, we first describe the data creation process and then move to our sample selection.

Data Creation

The point of departure is the Integrated Database for Labour Market Research (IDA), which is a panel of the Danish population aged 15-74 from the years 1992 to 2005. Each individual has a social security number that allows us to follow him or her across years. For each individual, we have socio-economic information such as education, marital status, family ties (mother and father), exact address information, several income measures, and detailed wealth information. We have both biological and legal family ties. Below, we describe the family definition used in this paper, which is based on legal ties, not biological. The above information in IDA is drawn from various sources by Statistics Denmark. The educational information is compiled using registers going back to the mid 1970's. Income information is drawn from tax records and is considered highly reliable.[5]

We have detailed wealth information since Denmark from 1987 to 1996 had a wealth tax and the tax reports from individuals were randomly audited. House value,

[5] See Kleven, Knudsen, Kreiner, Pedersen, and Saez (2011) who 'find that the tax evasion rate is close to zero for income subject to third-party reporting'. Third-party reporting includes both reporting from firms who tax wages and reporting from banks, mortgage institutions, brokers, etc.

bank depostits, value of listed stocks, mortgage debt and bank debt were reported automatically by banks and other financial intermediaries to the tax authorities for all Danish taxpayers and are therefore considered to be very reliable. In the period of the wealth tax, individuals self-reported information about non-deposited bonds and cash holdings, as well as the value of investment objects and high value objects such as cars and boats. After the abolishment of the wealth tax the automated systems continued as before, but individual self-reporting stopped. However, in the aggregate series no impact can be seen from this. The detailed wealth information allows us to model variation in wealth due to non-housing decisions. This is further elaborated on in the modeling section.

As the unit of observation in the model is a family, we use the definition by Statistics Denmark.[6] We sum income and wealth across family members. One of the reasons for using this definition of a family is that it requires all family members to live together. Using the address information for the family, we merge data from several administrative sources describing the housing unit that the family currently lives in, including actual sales prices and public valuations. Data on public valuations are available for the entire period. These form the basis for taxation and individuals can complain if they do not think that the valuation is fair.

Using the address information we assign each family to a neighborhood. The neighborhoods are constructed by Damm and Schultz-Nielsen (2008) mapping 100 times 100 meter cells into larger neighborhoods.[7] The largest neighborhood constructed by Damm and Schultz-Nielsen (2008) consists of around 600 households. These are too small for our study, so we aggregate them further.[8] This results in a division of Denmark into 1263 neighborhoods.

One drawback of the neighborhood identifiers is that they are only calculated using 2003 data. I.e. our data divides all inhabited houses in 2003 into the neighborhoods. This raises some dynamic problems, since houses might be built later, torn down before, or simply be uninhabited in 2003. In these cases we can not assign a neighborhood to the address of the family. Specifically, this means that our sample covers 5.16m Danes in 2004 (ultimo 2003) but only 4.66m in 1992 and 4.98m in 2005. For further details see appendix A. Given that we lose around 10 percent of the sample in the most critical year, we do not regard this as a major problem.

[6] A family could be a single household, couples (married or cohabiting), or a couple with children. Children are included if they live on the same address as at least one of the parents and are under the age of 25. The definition of children is based on a legal perspective and not on biology. For a detailed description of the definition of a family, see `http://www.dst.dk/da/Statistik/dokumentation/Times/moduldata-for-befolkning-og-valg/familie-id.aspx`

[7] See appendix B for a detailed description for what they have done.

[8] See appendix B for a detailed description of our aggregation algorithm.

Neighborhood Characteristics

Each neighborhood is going to be associated with a set of characteristics. These come from various sources.

Properties With housing unit specific information, including sales prices, we are able to construct a representative household for each neighborhood. To do this we merge very detailed information on the housing unit from the Buildings and Housing Registry (BBR), which is administered by the Danish government for tax purposes and for the public valuations. The register contains a large amount of descriptive variables for houses, around 100 variables, however, due to changes in definitions across time, many missing values, and non-quantifiable measures many of the variables are of low interest. The register contains only property specific information.

The value of a house is measured as the sum of the public valuation of the plot and of the house.[9] This information is obtained from a separate register, which is managed by the tax authorities. The reason why the plot and house values are measured individually is that Denmark has both a property tax and a land tax, where the property tax is recovered to the state and the land tax is set by and recovered to municipalities.[10]

Air Pollution Furthermore, we have acquired data from the Danish Center for Environment and Energy on air pollution. They have combined two models into one model. To the best of our knowledge this is the best model for air pollution in Denmark. The model has a regional background level and a local urban level. The pollution contribution from the regional level is calculated for all of Europe and measures pollution across 17 by 17 square meter cells. This is done using the DEHM model, see Christensen (1997) for further details. The main variation that we are going to use is that of the local urban level (UBM), see Berkowicz (2000) for further details. The model divides Denmark into 1 by 1 square kilometer cells using the 100 times 100 meter grid that was also used to construct the neighborhoods.

For each center of a 1 by 1 square kilometer cell we calculate the average pollution over that area across time. The model simulates pollution patterns for each hour. The

[9]The public valuation is used for tax purposes and is in general considered to be of reasonable quality. For houses sold we can calculate the correlation between the selling price and the public valuation. This correlation is found to be 0.79 and the R^2 of a simple regression pooled over the entire sample is 0.63. In the sample, actual sales prices are found to be 7.45% higher than public valuations with yearly variations from -2% to 18%. This makes sense since individuals are allowed to complain if they believe that their valuation is to high.

[10]The valuations are updated only every second year, and the algorithm for measuring house values is not publicly available.

concentration of pollutants is the average levels, hence not accounting for local short lived blips.[11]

We have measures for the concentration of O_3, $PM_{2,5}$, and CO measured by μg per m^3, which we standardize to have zero mean and variance one.

Criminals Exposure to local crime might also be an important determinant of residential choices. We have access to data on convictions, which includes the type of crime and the social security number of the convicted. We link these data to the population data described above and then aggregate for each neighborhood the number of convicted individuals living in that neighborhood. The data set on convictions starts in 1980 and our sample starts in 1992, so to ensure consistency in our measure we use a rolling window of 12 years when we aggregate convicts in the neighborhoods. This implies an assumption that society will forget about the past criminal activities of a convict after 12 years of law-abiding behavior. We measure the crime intensity by the number of convicted offenders per 1000 inhabitants in the neighborhoods, where we divide convictions into those convicted of property, violent, and sexual crime, respectively.[12]

Physical Description For each neighborhood we have calculated the center as the centroid where weights are given only to inhabited hectare cells.[13] From here we find distances to various physical amenities like nearest freeway and airport using GIS data provided by the Danish Geodata Agency. The data also contains information about land usage. From this we determine the amount of land within the neighborhood that is devoted to forest and urban areas. All of these physical measures are calculated in 2012 and assumed to have been constant over our sample period, even though some minor changes have occurred.

Neighbor Observables From the IDA register (with full population) mentioned above and the neighborhood identifiers, we can observe characteristics of inhabitants within each neighborhood. We determine the mean unemployment rate of inhabitants along with the fraction of inhabitants that are ethnic Danes. Lastly, we calculate a proxy of population density. Neighborhood sizes are defined from the approximated amount of land that is forest, urban area, or another vegetation. We divide neighborhood populations by this size measure to get our estimate of population density.

[11] Further details are available upon request.

[12] Post 2002 we also have access to data that identifies victims. From this data we find that the correlation between the number of convicts and victims within each neighborhood is .44, .68, and .19 for property, violent, and sexual crime, respectively. This gives a strong indication, especially for property and violent crime, that our measure of convicts is a good proxy for the probability of being a victim.

[13] This is the simple average over the north and east coordinates of hectare cells respectively.

Descriptive Statistics

Here we present some descriptive statistics for the described sample. Table 3.1 shows descriptive statistics for the characteristics of the households used. The table includes the three variables that will later be used to discretize the type space, which is wealth, income, and whether there are children in the household. Furthermore, the table contains the fraction of the sample that chooses to move within one year, the sales price, which is the determinant of moving costs and one of the key elements of household wealth accumulation.

Table 3.2 shows descriptive statistics for the neighborhood amenities. The first part contains the dynamic amenities for which we will either determine the marginal willingness to pay or include as controls for unobserved neighborhood characteristics in a utility decomposition. The second part contains static amenities.

Further description of the sample can be found in appendix A, where some additional descriptive statistics of the sample are included including the persistence and geographical variation of amenities.

Table 3.1. Descriptive Statistics for Households

Co-variate (Z)	Mean	Std. Dev.	Minimum	Maximum
Wealth	616,267	1,052,103	-2,600,000	20,400,000
Income	479,792	280,820	47,000	4,400,000
Children	34.74%	47.61%	0	1
Buyer	4.48%	21.43%	0	1
Price	1,724,585	3,786,239	150,000	79,000,000

3.4 Estimation

The estimation of the dynamic demand model is conducted in four steps that depart from the conditional choice probabilities of Hotz and Miller (1993) and account for the dynamics of expectations inspired by Bayer, McMillan, Murphy, and Timmins (2015). First, some assumption needs to be imposed on the model to ensure identification. We use the conditional choice probabilities to identify the lifetime values of owning in the Danish neighborhoods. Secondly, we estimate the structural parameters of the moving costs and at the same time identify the marginal value of wealth. Third, we simulate the future in order to form appropriate expectations hereof, which is then used to isolate the flow utility. In the last step we decompose the flow utilities based on various neighborhood characteristics and form estimates of the marginal

Table 3.2. Descriptive Statistics for Amenities

Amenity (X)	Unit	Obs.	Mean	Std. Dev.	Min.	Maxi.
Unit size	m^2	17,682	128.91	19.74	52.76	206.57
$PM_{2.5}$ [a]	$\mu g/m^3$	17,682	0	1	-2.52	3.60
CO [a]	$\mu g/m^3$	17,682	0	1	-3.55	2.64
O_3 [a]	$\mu g/m^3$	17,682	0	1	-1.25	7.57
Violent neighbors	$/1,000$	17,682	1.83	1.34	0	13.47
Ethnic Danes	%	17,682	93.84	6.08	18.68	100
Unemployment	%	17,682	5.07	2.52	0.42	36.50
Population density	$/km^2$	17,682	28.83	7.97	4.64	67.40
Distance freeway	km	1,263	6.42	11.30	0.01	116.20
Urban land area	ha	1,263	1,931.82	1,833.51	6.60	6,716.10
Forest area	ha	1,263	3,022.64	2,224.32	10.30	9,962.40

[a] Pollutants are measured in standard deviations of annual mean.

willingness to pay (MWTP) for these amenities by utilizing the estimates of marginal value of wealth.

Identifying Assumptions

In order to identify and extract the flow utility from the Bellman equation, it is necessary to place some structural assumptions on the model like the assumptions that can be found in Rust (1987), which draw from the evolution in the static discrete choice literature (see e.g. McFadden (1981) for a review). The identification of dynamic discrete choice processes is addressed in Rust (1994) and Magnac and Thesmar (2002) and has evolved to a set of standard assumptions in the literature.

Additive Separability It will be assumed that the flow utility function, $u_{i,j,t}$, and the idiosyncratic error term, $\varepsilon_{i,j,t}$, are additive separable. Hence, it follows from the definitions above that the moving cost-corrected flow utility will be given as

$$u^c_{i,j,t} = u(x_{j,t}, \xi_{j,t}, z_{i,t}) - c(z_{i,t}, x_{d_{i,t-1},t}) I_{[y_{i,t}=1]} + \varepsilon_{i,j,t}. \tag{3.5}$$

Conditional Independence To obtain conditional independence it is assumed that the idiosyncratic stochastic shock at time t, $\varepsilon_{i,j,t}$, does not have any predictive power for the realization of the state variables in the subsequent period, $\Omega_{i,t+1}$. Thereby it is possible to break the conditional probability of the realizations of future states in equation (3.1) into multiplicative transition densities,

$$q\left(\Omega_{i,t+1}, \varepsilon_{i,j,t+1} | \Omega_{i,t}, \varepsilon_{i,j,t}, d_{i,t}\right) = q_\omega(\Omega_{i,t+1} | \Omega_{i,t}, d_{i,t}) q_\varepsilon(\varepsilon_{i,t+1}). \tag{3.6}$$

In particular, note that the previous periods' location choice, $d_{i,t-1}$, will have no predictive power for any future state, as this information is fully captured by the current periods' location decision, $d_{i,t}$.

Rational Expectations As noted by Magnac and Thesmar (2002) we need to impose some structure on agents subjective beliefs about the Markovian law of motion of the state variables, denoted $\mu\left(\Omega_{i,t+1}, \varepsilon_{i,j,t+1} | \Omega_{i,t}, \varepsilon_{i,j,t}, d_{i,t}\right)$. Here we will assume that expectations are perfect such that they are in line with actual transition probabilities

$$\mu\left(\Omega_{i,t+1}, \varepsilon_{i,j,t+1} | \Omega_{i,t}, \varepsilon_{i,j,t}, d_{i,t}\right) = q\left(\Omega_{i,t+1}, \varepsilon_{i,j,t+1} | \Omega_{i,t}, \varepsilon_{i,j,t}, d_{i,t}\right). \tag{3.7}$$

The Idiosyncratic Stochastic Shock Following the standard assumption (See Berry (1994) for an application in a comparable model framework) we assume that the idiosyncratic stochastic shock, $\varepsilon_{i,j,t}$, is independently and identically distributed Type 1 Extreme Value (T1EV). It is important that this error term is independently distributed both across families and across neighborhoods. This assumption makes it possible to identify the value function from the observed 'market' share of households that chose a given neighborhoods at any point in time.

Given these assumptions we can obtain the choice-specific value function,

$$v_j^c(\Omega_{i,t}) = u_{i,j,t} - c_{i,t} I_{[y_{i,t}=1]} + \beta E_t \left[\log\left(e^{v_j(\Omega_{i,t+1})} + \sum_{k=0}^{J} e^{v_k^c(\Omega_{i,t+1})} \right) \middle| \Omega_{i,t}, d_{i,t} = j \right]. \tag{3.8}$$

From here it also follows that the choice-specific value function can be divided into two additive components, the first being a choice specific value function, neglecting the cost associated with a move in the current period, and the other being the potential moving cost incurred by the optimal location decision, $v_j^c(\Omega_{i,t}, h_{i,t}) = v_j(\Omega_{i,t}) - c(z_{i,t}, x_{h_{i,t}}) I_{[y_{i,t}=1]}$. Here the value function is not a function of the current location, $h_{i,t}$, as it follows from the right hand side of equation (3.8) since only the potential moving cost is a function hereof. Hence,

$$v_j(\Omega_{i,t}) = u_{i,j,t} + \beta E_t \left[\log\left(e^{v_j(\Omega_{i,t+1})} + \sum_{k=0}^{J} e^{v_k(\Omega_{i,t+1}) - c_{i,t}} \right) \middle| \Omega_{i,t}, d_{i,t} = j \right]. \tag{3.9}$$

This is the fundamental equation that will be estimated. The equation is used to subtract flow utilities that households receive from living in neighborhoods with particular compositions of the various locally non-traded amenities.

Conditional Value Function

At the beginning of every period a household faces a choice between staying at the current housing unit or moving to another unit, potentially in a different neighborhood, but then facing substantial moving costs. At first we consider only households that have chosen to move. As described above, the cost of moving is assumed only to be dependent on the housing unit and neighborhood that is vacated. Hence, conditional on moving the moving cost is the same across location choices and does not affect this decision. Therefore we can omit moving costs from this first stage of the estimation procedure.

Households will choose $d_{i,t} = j$ to maximize lifetime value $v_j(\Omega_{i,t}) + \varepsilon_{i,j,t}$, where $v_j(\Omega_{i,t})$ is defined in (3.9) and $\varepsilon_{i,j,t}$ was assumed to be i.i.d. T1EV distributed.

Ideally we would like to recover estimates of the lifetime value for each individual household, however, in practice this is unfeasible due to the binary nature of location choices. The curse of dimensionality often haunts dynamic discrete choice models with a large number of choices and states so we partition the sample in order to utilize the results from Hotz and Miller (1993) who identify the lifetime values from the conditional choice probabilities. We discretize the type space and let types be defined discretely and denote the types by τ. Based on the observable characteristics of household i, $Z_{i,t}$, we determine the type of every household and define $v^{\tau}_{j,t}$ to be lifetime value of households of type τ living in neighborhood j at time t where we ignore the current period moving costs. If we let $u^{\tau}_{j,t}$ and c^{τ}_{t} denote the corresponding flow utility and moving costs respectively then the discrete type version of (3.9) takes the form

$$v^{\tau}_{j,t} = u^{\tau}_{j,t} + \beta E_t \left[\log\left(e^{v^{\tau}_{j,t+1}} + \sum_{k=0}^{J} e^{v^{\tau}_{k,t+1} - c^{\tau}_{t+1}} \right) \middle| \, \Omega^{\tau}_{t}, d^{\tau}_{i,t} = j \right]. \tag{3.10}$$

In order to maximize lifetime value conditional on moving, household i chooses at time t over neighborhoods according to the rule $d_{i,t} = \text{argmax}_{j\in\{0,\ldots,J\}}[v^{\tau}_{j,t} + \varepsilon_{i,j,t}]$. Now, we can utilize the distributional assumption on the idiosyncratic shock to be i.i.d. T1EV as this implies that the probability of a household of type τ to move to neighborhood j at time t is

$$P^{\tau}_{j,t}(v^{\tau}_{t}) = \frac{e^{v^{\tau}_{j,t}}}{\sum_{k=0}^{J} e^{v^{\tau}_{k,t}}}.$$

Given that we can obtain non-parametric estimates of these conditional choice probabilities, $\hat{P}^{\tau}_{j,t}$, we can invert this relation and get estimates of the normalized lifetime values according to

$$\hat{v}^{\tau}_{j,t} = \log(\hat{P}^{\tau}_{j,t}) - \frac{1}{J+1}\sum_{k=0}^{J} \log(\hat{P}^{\tau}_{k,t}). \tag{3.11}$$

This set of normalized estimates of the lifetime value is calculated within type and year, hence, we need to add a normalizing constant, m_t^τ, to ensure inter type and year consistency of the lifetime values. The relation between the true and the estimated lifetime value is $\hat{v}_{j,t}^\tau = v_{j,t}^\tau - m_t^\tau$, where m_t will be estimated as part of the subsequent stage of the estimation procedure.

The type space is defined discretely from 10 wealth bins, 2 income bins and 2 children-/no children living at home bins.[14] The rich data allows us to follow household types for every year and know the exact type of the household even when they do not move. The types are defined from the entire population of housing owners rather than only by the households that actually move.

From the data we will obtain non-parametric estimates of the conditional choice probabilities by the observed shares. Type τ choose to move to neighborhood j at time t with probability

$$\tilde{P}_{j,t}^\tau = \frac{\sum_{i=1}^{N_1} I_{[d_{i,t}=j]} \cdot I_{[Z_{i,t} \in Z^\tau]}}{\sum_{i=1}^{N_1} I_{[Z_{i,t} \in Z^\tau]}} \tag{3.12}$$

where N_1 is the number of movers that choose to buy within the Danish housing market. The data set contains not only owners but also non-owners with their exact address. This allows us to determine the number of previous owners that choose to rent a housing unit on the Danish renting market rather, N_2, rather than own. These will choose the outside option with probability

$$\hat{P}_{0,t}^\tau = \frac{\sum_{i=1}^{N_2} I_{[d_{i,t}=0]} \cdot I_{[Z_{i,t} \in Z^\tau]}}{\sum_{i=1}^{N_2} I_{[Z_{i,t} \in Z^\tau]}}. \tag{3.13}$$

Inside probability estimates are corrected down for the outside choice probabilities by multiplication, $\hat{P}_{j,t}^\tau = (1 - \hat{P}_{0,t}^\tau)\tilde{P}_{j,t}^\tau$.[15]

Moving Costs

In the first stage of the estimation we have analyzed the conditional choices, $d_{i,t}$. In this stage we will address the decision to move, $y_{i,t}$. A household starts each period by making a move/stay decision, where the household can either choose to stay in the current housing unit with some observed and unobserved attributes and neighborhood amenities, or it can choose to make a costly re-optimization and move

[14] The wealth bins are determined from the deciles of the wealth distribution and the income bins from the mean of the income distribution.

[15] As it follows from the normalized value function in (3.11), zero-shares will constitute a problem, hence, we assume that there exists a strictly positive lower bound, $\mu \sim 0$, on the probability of any type moving to any neighborhood, i.e. $\hat{P}_{j,t}^\tau = \max\{\mu, \hat{P}_{j,t}^\tau\}$. In practice we set μ to 1/10 of an observed move.

to a different housing unit with some other characteristics, which in turn yields a change in expected lifetime value. This household decision is inherently dynamic for two reasons. First, there are large monetary costs associated with selling a housing unit, which will affect household wealth both today and in the subsequent periods, which again influences future decisions. On the other hand, the household might choose to move into a better housing unit that has a higher quality as a financial asset, implying larger capital gains in the future. Secondly, since re-optimization is costly, a household will not choose to move in every period and therefore the current period choice will not only affect the current flow utility, but also the flow utilities for some periods to come. This is obvious from (3.10).

The cost of moving is composed of two additive terms. First, there are psychological moving costs, $p(z_{i,t})$. The psychological costs comprise the dis-utility of any non-monetary costs related to moving, that could be spending time searching for a new housing unit, loss of sentimental value, and uncertainty about future neighborhood and neighbors. The second part is financial moving costs, $f(x_{h_{i,t},t}, z_{i,t})$. The financial moving costs are the loss of utility that the household faces due to the wealth reduction that follows from selling a housing unit.

$$c(z_{i,t}, x_{h_{i,t}}) = f(x_{h_{i,t},t}, z_{i,t}) + p(z_{i,t}). \tag{3.14}$$

We specify $y^{\tau}_{i,t}$ as the binary decision variable of the household, where $y^{\tau}_{i,t} = 0$ is equal to choosing not to move and $y^{\tau}_{i,t} = 1$ is a choice of moving to the neighborhood that yields the highest lifetime utility. Our type-space is among others defined in the wealth-dimension and therefore the household changes type endogenously whenever the household chooses to move. Let $\bar{\tau}$ denote the type of a household if the monetary costs of moving subtracted from the initial wealth of the period. In this case the move/stay decision of a household that live in neighborhood j is made according to the following rule

$$y^{\tau}_{i,t} = I\left[v^{\tau}_{j,t} + \epsilon_{i,j,t} < \max_{\{k\}^J_0}[v^{\bar{\tau}}_{k,t} + \varepsilon_{i,k,t}] - p(z_{i,t})\right]. \tag{3.15}$$

We distinguish between $\epsilon_{i,j,t}$ and $\varepsilon_{i,j,t}$ in order to insure that households can make moves within neighborhoods. $(\epsilon_{i,j,t} - \varepsilon_{i,j,t})$ can be thought of as a shock to moving costs. Inserting $\hat{v}^{\tau}_{j,t} = v^{\tau}_{j,t} - m^{\tau}_t$ from the first stage the decision rule has the form

$$y^{\tau}_{i,t} = I\left[\hat{v}^{\tau}_{j,t} + \epsilon_{i,j,t} < \max_{\{k\}^J_0}[\hat{v}^{\bar{\tau}}_{k,t} + \varepsilon_{i,k,t}] - (m^{\tau}_t - m^{\bar{\tau}}_t) - p(z_{i,t})\right]. \tag{3.16}$$

The unobserved term $(m^{\tau}_t - m^{\bar{\tau}}_t)$ should be interpreted as the baseline differences in expected lifetime value of being of the type τ versus having endogenously reduced wealth and being of type $\bar{\tau}$. In other words, the term $(m_t^{\tau} - m_t^{\bar{\tau}})$ is equal to the value equivalent of the financial moving costs. The financial moving cost is treated as an observable and set to 4% of the actual sales price of the housing unit sold[16], where we denote housing unit price of the unit owned by individual i at time $t-1$ by $\zeta_{i,t}$. Note that we have assumed that the financial moving costs are carried by sellers, so obviously first-time buyers only face psychological moving costs and anything else is considered a reallocation of wealth into the housing market.

The two moving cost components are parameterized to household characteristics, assuming linearity in the arguments, hence[17],

$$f(x_{h_{i,t},t}, z_{i,t}) = (m_t^{\tau} - m_t^{\bar{\tau}}) = I_{[j\neq 0]} \cdot 0.04 \cdot \zeta_{i,t} z'_{i,t}\gamma_f, \tag{3.17}$$

$$p(z_{i,t}) = z'_{i,t}\gamma_p. \tag{3.18}$$

We estimate γ_f and γ_p by maximum likelihood based on the observed choices and the model-based probabilities of observing those choices.

$$P(y_{i,t}^{\tau} = 0) = \frac{e^{\hat{v}_{j,t}^{\tau}}}{e^{\hat{v}_{j,t}^{\tau}} + \sum_{k=0}^{J} e^{\hat{v}_{k,t}^{\bar{\tau}} - I_{[j\neq 0]}\cdot 0.04\cdot\zeta_{i,t}\cdot z'_{i,t}\gamma_f - z'_{i,t}\gamma_p}}. \tag{3.19}$$

Having parameterized and estimated $(m_t^{\tau} - m_t^{\bar{\tau}})$ we can back out the inter-type consistent estimates of the lifetime values as mentioned in the first stage. All we need is to decide on some reference type and level of lifetime value. We choose to set the mean choice specific lifetime value to zero for the lowest wealth type for each combination of income-, child-type, and year. Then we can recursively calculate the m_t^{τ}'s and hence identify the unnormalized choice specific lifetime value from $v_{j,t}^{\tau} = \hat{v}_{j,t}^{\tau} + m_t^{\tau}$.

From this step there is an important gain from having identified γ_f. As we know the structural financial moving cost parameters, we essentially know how household observables load wealth changes into lifetime value changes; i.e., we have established a link between our measure of lifetime value and a monetary unit. This will be utilized below.

Flow Utility

At this stage we have identified all the relevant current period state variables and the aim of this section is to outline how expectations are formed by specifying the transition probabilities. When we know the transition probabilities we can simulate

[16] Anecdotal evidence from Danish real estate organizations support that the average cost of selling a housing unit is 4%. However, there are some fixed costs e.g. to the Land Registry.

[17] We include a constant, child dummy, and income level in γ_f and γ_p.

the expected value of the log-sum in (3.10) and from there it is straightforward to back out the flow utilities.

First of all, every household faces a probability of changing type, which we will consider as an exogenous event. Such changes occur when households are expanded by the arrival of their first child or if the last child leaves the nest. Income-type changes can be a result of a job change or loss and the exogenous change in wealth can be related to inheritance or the materialization of investment risks. Non-parametrically we calculate the exogenous probability that a family of type τ is of type $\tilde{\tau}$ in the subsequent period,

$$P(\tilde{\tau}|\tau) = \frac{\sum_{i=1}^{N} I_{[\tilde{\tau}|\tau]}}{\sum_{i=1}^{N} I_{[\tau]}}, \tag{3.20}$$

where N is the number of families that own, but do not move, across time.[18]

The wealth accumulation in the household's housing unit and the moving costs are assumed to be sufficient statistics for determining the endogenous wealth type transitions. Therefore, predicting future neighborhood prices suffices to determine the transition probabilities of household housing wealth. House prices for all neighborhoods are found to be non-stationary, hence, we predict house price changes, $\Delta\zeta_{j,t}$, based on an AR(L) model, where we pool information across neighborhoods,

$$\Delta\zeta_{j,t} = \varrho_0 + \sum_{l=1}^{L_1} \varrho_{1,l} \Delta\zeta_{j,t-l} + \nu_{j,t}. \tag{3.21}$$

We construct neighborhood prices from the data on public valuations and set $L_1 = 1$.

Current lifetime values depend critically on the evolution of the lifetime values of the various type and neighborhood combinations. Therefore households need to form expectations on the transition probabilities of the lifetime values of living in every neighborhood for every potential next period type. In practice we predict the lifetime values according to an AR(L) model, where we pool the regressions across types, implying that every type holds similar beliefs about the dynamic nature of a given neighborhood, i.e. is lifetime neighborhood value persistent or volatile across time. In order to cope with inter type level differences, we model first differences.

$$\Delta v_{j,t}^{\tau} = \rho_0^{\tau} + \sum_{l=1}^{L_1} \rho_{1,l}^{\tau} \Delta v_{j,t-1}^{\tau} + \varsigma_{j,t}^{\tau}. \tag{3.22}$$

Here we treat the logit inclusive value as a sufficient statistic to predict future states. This way we allow for both observed and unobserved neighborhood characteristics to influence the future states equally.

[18] We do not include movers, as these are expected to change wealth-type endogenously and potentially moving is correlated with having the first child.

With predictions of current and next period state variables we can determine the flow utility from (3.10) for every type, year, and neighborhood combination

$$u^{\tau}_{j,t} = v^{\tau}_{j,t} - \beta \sum_{\tilde{\tau}} P(\tilde{\tau}|\tau) E_t \left[\log\left(e^{v^{\tilde{\tau}}_{j,t+1}} + \sum_{k=0}^{J} e^{v^{\tilde{\tau}}_{k,t+1} - c^{\tilde{\tau}}_{t+1}} \right) \Big| \, \Omega^{\tau}_t, d^{\tau}_t = j \right]. \tag{3.23}$$

Due to the non-linearity in the expectation we simulate the future $n = 200$ times[19] and average over the simulated flow utilities. We denote the flow utility from the simulation s by $u^{\tau}_{j,t}(s)$, where we draw prices and lifetime values from their respective empirical distribution of the in-sample prediction errors based on (3.21) and (3.22) and exogenous type transitions from the transition matrix in (3.20). By averaging over simulations we calculate the flow utility of living in neighborhood j at time t for every type, τ.

$$u^{\tau}_{j,t} = \frac{1}{n} \sum_{s=1}^{n} u^{\tau}_{j,t}(s). \tag{3.24}$$

Marginal Willingness to Pay

Household move/stay decisions reflect the preferences over neighborhoods and associated amenities. The neighborhoods provide households with utility today as well as in the future, however, neighborhood amenities evolve dynamically over time and future states are unobserved. Hence, a decomposition of lifetime value from living in a given neighborhood would lead to misleading conclusions regarding household preferences as only today's amenity level is observed. Therefore, we will decompose the flow utility that is estimated in the foregoing steps, $u^{\tau}_{j,t}$, by the observed neighborhood amenities, $x_{j,t}$. The unobserved characteristics of the neighborhood, $\xi^{\tau}_{j,t}$, is included as the error term of the model.

One issue that must be addressed before proceeding to the decomposition is the cost of home ownership. Clearly, the user cost, denoted $uc_{j,t}$, varies across neighborhoods and in particular it varies with prices, as higher prices leads to larger mortgages. However, this implies that user costs will be endogenous as prices are higher in better neighborhoods. In the moving cost section we obtained estimates of the marginal value of wealth and by assuming that the marginal value of wealth equals the marginal utility of income, we can predetermine the dis-utility of paying user costs and adjust the flow utility accordingly. If we denote the neighborhood observables excluding user costs by $\tilde{x}_{j,t}$ then we estimate the following regression model

$$u^{\tau}_{j,t} + uc_{j,t} z'_{\tau,t} \gamma_f = \alpha^{\tau}_0 + \tilde{x}'_{j,t} \alpha^{\tau}_1 + FE^{\tau} + FE^{t} + FE^{r} + \xi^{\tau}_{j,t}, \tag{3.25}$$

[19] 200 simulations seem to be sufficient for robustness of the simulated continuation values.

where FE^{τ}, FE^{t}, and FE^{r} cover type, year, and region fixed effects respectively.

From this decomposition we obtain estimates of the marginal utility of the various neighborhood amenities included in $\tilde{x}_{j,t}$. In order to obtain our final estimates of marginal willingness to pay for the included amenities, we reapply the findings from financial moving costs as the link from utility to the monetary unit.

$$MWTP^{\tau} = \frac{\alpha_1^{\tau}}{z_{\tau}'\gamma_f}. \tag{3.26}$$

This concludes the estimation.

The marginal willingness to pay coefficients are allowed to vary across types and to explore such patterns we propose a linear varying coefficients decomposition in two steps. First, we decompose the flow utilities for each type and year combination. Then, after converting the marginal utilities to marginal willingness to pay, we decompose these based on type characteristics

$$MWTP_t^{\tau} = g(z_{\tau,t}') + \xi_t^{\tau}. \tag{3.27}$$

Here $g(\cdot)$ is a general function of type characteristics to be chosen. In our application we let the function contain wealth, income and interactions along with an indicator for potentially constrained households types.

3.5 Results

The estimation of the structural model provides us with a large amount of estimates of lifetime values, flow utilities, structural parameters on moving costs, and expectation formation. This section focuses on the moving costs and the marginal willingness to pay estimates. The structural parameters describing financial moving costs are of interest as they represent the marginal value of wealth, which we use in the final decomposition of flow utilities into marginal willingness to pay estimates of neighborhood amenities. The estimation approach allows us to evaluate the willingness to pay at the average as well as for different household types, which we exploit in the latter part of the section.

Moving Costs

Households refrain from frequent re-optimizing of their location decision primarily due to the psychological and financial moving costs that such choices entail. The moving cost parameters are identified from the choice between moving and staying and the parameters are presented in Table 3.3.[20] The estimates are measured in

[20] Standard errors are calculated by a numerical method.

lifetime value and hence they do not have any clear cut interpretation, though the sign and relative sizes are of a certain interest. We note that there are large positive financial and psychological moving costs. The financial moving cost is decreasing in income, suggesting that the value function has a concave profile in wealth[21] and the financial component is also decreasing for households with children, which is interpreted as the concern for one's children is larger than that for wealth accumulation. Considering the psychological moving costs, the impact of having children point in the same direction, but the cost is increasing in income. This seems obvious as most income is increasing in workload and hence limits the time to search for an alternative housing unit.

Table 3.3. Financial and Psychological Moving Cost Parameteres

	Estimate	Std. error
Financial		
Constant	5.1654^{***}	0.0024
Income	-0.1127^{***}	0.0001
Children	-1.4398^{***}	0.0031
Psychological		
Constant	10.4603^{***}	0.0024
Income	0.0454^{***}	0.0002
Children	-0.0487^{***}	0.0034

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Kennan and Walker (2011) provides a thorough discussion of the interpretation of moving costs in dynamic discrete choice models like this. One thing that should be mentioned is that the unexplained component of the flow utility, $\xi^{\tau}_{j,t}$, comprises both variations in unobserved neighborhood characteristics and shocks to psychological moving costs, hence, the mean psychological moving costs in Table 2.2 can be substantially larger than those faced by households that actually move.

The financial moving cost parameters are identified by the decrease in value for movers implied by paying 4% of house value. This link between the monetary unit and the value unit is crucial for the final stage of our estimation as it allows us to convert marginal utilities to marginal willingness to pay.

MWTP for Locally Non-traded Amenities

Table 3.4 reports the results from the last stage of the estimation. The results are based on the flow utilities after correcting for the impact of endogenous user costs.

[21] We assume that wealth and income are perfect substitutes.

Furthermore, the results are converted to marginal willingness to pay estimates by the marginal value of wealth measure obtained in the second estimation stage and should be interpreted as the willingness to pay for an additional unit of each of the components over a period of one year. This decomposition of the marginal utilities is conducted by pooling information across all type bins, where we have weighted the bins according to their relative sample sizes, such that willingness to pay of types with a low population, e.g. high wealth and low income, do not get overemphasized. The estimates are reported in 2005 Danish kroner.[22]

The primary variables of interest are the measures of the representative housing unit in each neighborhood and the air quality.[23] We additionally include a range of controls that fall into two categories; those that characterize your neighbors and those that are descriptive of the location. Additionally, the decomposition includes fixed effects for type, year and the region in which the neighborhood is located.

Across all types we find that the households are willing to pay 992.5 DKK for an additional square meter, which is in line with average prices from the Danish renting market. Households dislike bad air quality and are willing to pay significant amounts for a one standard deviation reduction of the ozone, particulate matter, and carbon monoxide levels.

Ethnic Danes and unemployment are both measured in percentage of neighborhood population and have the expected signs, where we note that the unemployment rate does not only control for the presence of unemployed individuals but also for poor local labor market conditions. The share of neighborhood inhabitants that have been convicted for a violent or sexual offense within the past twelve years are marginally significantly disliked. The remaining four controls are included to proxy for the general movement from the countryside into urban areas.

The results are resonably robust to changes in the specification of the type bins. Additionally, it is a concern that some housholds sort into newly build neighborhoods after 2003 when our neighborhoods are defined, hence, excluding some of the effect of forward looking behaviour. We have estimated the model ending the sample in 2003, which only affects the results in table 3.4 marginally.[24]

Elasticity of Substitution: Wealth and Income

Ever since Roback (1982), studies have investigated many relationships between income and willingness to pay for various climatic of pollution variables. More recent studies have more explicitly investigated the elasticity of substitution between in-

[22] The average exchange rate between USD and DKK in 2005 was 6.1 DKK/USD.

[23] The estimates of marginal willingeness to pay to avoid air pollutants is slightly larger than findings in previous studies.

[24] A simpler version of the model has also been estimated using smaller neighborhoods, to which the results seem robust.

Table 3.4. Pooled Marginal Willingness to Pay Estimates

	MWTP
m^2	992.5***
	(233.4)
$PM_{2.5}$	-4,170.9**
	(1,600.7)
CO	-158.5***
	(29.24)
O_3	-5,545.5***
	(1,055.7)
Violent neighbors	-1,702.7*
	(741.4)
Ethnic Danes	4,664.3***
	(682.4)
Unemployment	-7,546.1**
	(2,350.0)
Population density	1,399.8
	(903.8)
Forest area	-188.2
	(496.4)
Urban area	-90.08
	(521.1)
Distance to freeway	-237.6
	(368.5)
Fixed effects	Year, type, region
Observations	656,760

Standard errors in parentheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

come and pollutants. However, these studies all ignore the issue of some households being borrowing constrained.[25] Suppose that some households do not have a lot of savings, which we are able to identify in our comprehensive data; these households would be less likely to move as the inability to make a sufficiently large down-payment would restrict their choices to neighborhoods with low prices, but also with a higher level of dis-amenities. Furthermore, the choice is influenced by the price at which the household can sell its current housing unit. This issue implies that the low wealth, i.e. borrowing constrained, households are less likely to sell their houses and those that actually do move are more likely to accept a higher level of dis-amenities, e.g.

[25] As a rule of thumb households can borrow 3.5 times pre-tax income plus net wealth, but should always be able to make a minimum down-payment of 5% of the house price.

bad air quality. In models based on the Hotz and Miller (1993) conditional choice probabilities, these borrowing constraints would, falsely, lead to the conclusion that borrowing constrained households with low wealth and income positively value the dis-amenities. In a dynamic context the issue of borrowing constraints is only magnified as households in neighborhoods that become relatively worse can fall victim to lock-in effects as households will have lower housing wealth to transfer to the new housing unit. Oppositely, households in improving neighborhoods will ceteris paribus have higher moving probabilities and can move to even better neighborhoods as these households can exploit the capital gains received from the fortunate evolution on local conditions.

Tables 3.5 through 3.8 report the varying coefficients decomposition of willingness to pay estimates for square meters and a one-standard deviation in $PM_{2.5}$, CO, and O_3. We explicitly allow for differences in the slope for borrowing constrained households as wealth should have a larger marginal impact on willingness to pay for this group of households compared to households that can choose from all locations. We indicate housholds as being borrowing constrained if they are low income and belong one of the lowest four wealth deciles, meaning that they have wealth below 200,000 DKK and pre-tax income below 450,000 DKK. The tables report coefficients over the entire sample from 1993 to 2005 and for two sub periods, 1993 to 1998 and 2000 to 2005. The sample is partitioned at the period over which 'Pinsepakken' was implemented. As the mortgage deductability was decreased from 50% to 32% and the bottom tax was decreased by 3 percentage points, we would expect that the impact of wealth on the willingness to pay would increase and the impact of income would be less important. However, prices on the housing market increased substantially over the sample period, and the income could therefore have a larger effect of household options.

First, the willingness to pay for additional space increases in wealth and income and the interaction of wealth and income suggests that households can substitute between the two. For the wealth constrained households there is a substantially larger effect from having additional wealth than for the rest of the population, but the slope is smaller if the households can substitute with income. This observation is in line with expectations. Interestingly, income has a lower effect in the marginal willingness to pay for extra space than the rest for the rest of the population, which could be explained by some degree of hand to mouth consumption, as increasing housing size is not the primary concern of households that are credit constrained, but rather to increase consumption.

In order to make inference on the differences in willingness to pay for the different pollutants in the decomposition it is necessary to understand what drives the geographical distribution of the pollutants. All three pollutants are byproducts of traffic, however, certain other differences occur. Particulate matter is primarily

Table 3.5. Varying coefficient decomposition of the WTP for m^2.

	Sample period		
	1993-2005	1993-1998	2000-2005
W	253.8***	193.7***	303.2***
	(0.221)	(0.173)	(0.323)
I	804.9***	721.1***	926.4***
	(0.509)	(0.444)	(0.856)
$W \cdot I$	-24.34***	-20.88***	-28.72***
	(0.0358)	(0.0282)	(0.0473)
$D \cdot W$	1,059.2***	1,227.5***	1,181.5***
	(8.659)	(8.092)	(11.98)
$D \cdot I$	-532.1***	-524.7***	-556.8***
	(0.833)	(0.836)	(1.437)
$D \cdot W \cdot I$	-378.5***	-434.3***	-418.0***
	(2.419)	(2.250)	(3.442)
Children	-1,007.1***	-776.1***	-1,246.6***
	(1.665)	(1.342)	(3.121)
N	520	240	240

W: wealth, I: income, D: dummy for borrowing constrained.
Wealth and income are measured in 100,000's DKK.
Standard errors in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

a byproduct of chemical reactions in the air that are based on the polluting gasses from engines and power-plants and especially consumption of fossil fuels in e.g. ships motoring though Danish waters. Hence, particulate matter is the best indicator of industrial areas as only about 1/3 of particulate matter in metropolitan areas is related to traffic and local chimneys. Carbon monoxide is a poisonous gas that is a result of incomplete combustion in petrol and diesel engines, which is mainly a concern in the metropolitan areas and the best measure of traffic related pollution. Lastly, ozone emerges from nitrogen oxides, which again stem from diesel engines and power plants, but the majority of the ozone that is measured in Denmark comes from sources in the Central Europe where ozone is created in chemical reactions with sunlight in still weather. Some of the ozone in metropolitan areas reacts with other gases, such that ozone levels are in general lower in the cities than in the countryside.

Table 3.6 reports the varying coefficients' estimates for small particulate matter, $PM_{2.5}$. Wealth and, in particular, income matter a lot for location decisions and willingness to pay. As was the case for square meters, income and wealth are again substitutes as the interaction term is positive. Willingness again increases more for wealth constrained households in wealth, and as it did for square meters, income

Table 3.6. Varying coefficient decomposition of the WTP for $PM_{2.5}$.

	Sample period		
	1993-2005	1993-1998	2000-2005
W	-614.4***	-537.1***	-637.8***
	(1.225)	(1.654)	(1.632)
I	-4,772.1***	-4,587.8***	-4,361.2***
	(3.484)	(3.492)	(6.023)
$W \cdot I$	112.9***	126.8***	108.2***
	(0.176)	(0.231)	(0.228)
$D \cdot W$	-6,555.2***	-5,308.4***	-9,749.6***
	(60.47)	(35.25)	(91.39)
$D \cdot I$	2,221.3***	760.7***	3,795.1***
	(4.698)	(4.051)	(8.410)
$D \cdot W \cdot I$	1,695.3***	1,351.3***	2,470.0***
	(17.13)	(9.603)	(26.42)
Children	-430.8***	-1,629.4***	875.0***
	(11.52)	(8.293)	(23.02)
N	520	240	240

W: wealth, *I*: income, *D*: dummy for borrowing constrained.
Wealth and income are measured in 100,000's DKK.
Standard errors in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

has a smaller effect on demand than in the general population, though still negative. We also see that demand for clean air for this group is even less influenced by wealth when income is larger.

In table 3.7 the varying coefficients for carbon monoxide, CO, are reported. While paramter estimates are small they are statistically significant and the signs of the estimates are as expected and point in the same direction as for $PM_{2.5}$. Lastly, coefficients from the decomposition of willingness to pay for ozone, O_3, are reported in table 3.8. Especially, income matters a lot for the willingness to pay to avoid ozone, but as most ozone in Denmark stems from sources abroad and ships, it is hard to argue that moving patterns truly are driven by a consideration of ozone levels, as ozone levels are far lower than what causes serious health issues, but rather that neighborhoods in the southern part of Denmark is generally low income locations and these are closer to Central Europa from which ozone may drift.[26]

[26] See Ellermann, Brandt, Hertel, Loft, Andersen, Raaschou-Nielsen, Bønløkke, and Sigsgaard (2014) for an assessment of recent pollutant levels and their sources.

Table 3.7. Varying coefficient decomposition of the WTP for CO.

	Sample period		
	1993-2005	1993-1998	2000-2005
W	-1.237***	2.526***	-2.333***
	(0.0136)	(0.0213)	(0.0187)
I	-0.525***	2.994***	0.179**
	(0.0413)	(0.0499)	(0.0707)
$W \cdot I$	0.103***	-0.276***	0.154***
	(0.00209)	(0.00317)	(0.00281)
$D \cdot W$	-60.25***	-25.66***	-75.18***
	(0.433)	(0.689)	(0.526)
$D \cdot I$	13.99***	23.21***	14.71***
	(0.0499)	(0.0577)	(0.0873)
$D \cdot W \cdot I$	17.57***	10.41***	19.24***
	(0.119)	(0.184)	(0.147)
Children	40.03***	12.14***	80.62***
	(0.116)	(0.127)	(0.202)
N	520	240	240

W: wealth, I: income, D: dummy for borrowing constrained.
Wealth and income are measured in 100,000's DKK.
Standard errors in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Implications of Expectations

Recent literature has put a great effort into identifying the exact impact of neglecting the dynamic features of the housing market in various valuation exercises. Valuable contributions on this include Bayer, Keohane, and Timmins (2009), Bayer, McMillan, Murphy, and Timmins (2015) and Bishop (2012). However, little effort has been put on determining the impact of the various components of these dynamic features.

In this section we look into the effect of specific components in the dynamic model of this paper, but leave the discussion between the static versus the dynamic model to the existing literature.

One of the advantages of our model is that it explicitly takes expectations into account. The main focus of Bayer, McMillan, Murphy, and Timmins (2015) was on the impact of expectations and endogenous type changes implied by household move-stay decisions. However, household choices are also influenced by expectations about the future that are related to non-housing decisions and are hence exogenous to this model. These expectations include the chance of having a child or the last child moving out of the home. It also includes household decisions on saving for

Table 3.8. Varying coefficient decomposition of the WTP for *O*3.

	Sample period		
	1993-2005	1993-1998	2000-2005
W	-71.83***	21.79***	-109.5***
	(0.489)	(0.911)	(0.614)
I	-1,303.2***	-1,093.4***	-1,385.9***
	(1.404)	(1.699)	(2.476)
$W \cdot I$	11.89***	5.780***	15.29***
	(0.0653)	(0.126)	(0.0822)
$D \cdot W$	415.1***	3,368.8***	-1,552.6***
	(20.40)	(24.47)	(25.46)
$D \cdot I$	428.4***	381.0***	662.3***
	(1.801)	(1.954)	(3.003)
$D \cdot W \cdot I$	-165.0***	-953.4***	308.8***
	(5.715)	(6.659)	(7.330)
Children	-1,451.7***	-1,110.0***	-1,301.3***
	(4.346)	(4.381)	(8.232)
N	520	240	240

W: wealth, I: income, D: dummy for borrowing constrained.
Wealth and income are measured in 100,000's DKK.
Standard errors in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

wealth accumulation versus consuming, and the impact of expected job loss or promotion. These exogenous changes are embedded in (3.20) and impact household continuation values. With detailed annual wealth and income data and specific knowledge of household composition, we are able determine the impact from these exogenous processes via the type transitions matrix (3.20). From equation (3.23) it follows that the bias from neglecting the exogenous type changes is given by

$$\beta \sum_{\tilde{\tau}} \left(P(\tilde{\tau}|\tau) - I_{[\tilde{\tau}=\tau]} \right) E_t \left[\log \left(e^{v^{\tilde{\tau}}_{j,t+1}} + \sum_{k=0}^{J} e^{v^{\tilde{\tau}}_{k,t+1} - c^{\tilde{\tau}}_{t+1}} \right) \Bigg| \, \Omega^{\tau}_t, d^{\tau}_t = j \right]. \tag{3.28}$$

For comparison with earlier studies we include a static version of the model in table 3.9, where $u^{\tau}_{j,t}$ is replaced by the static equivalent $(1-\beta) \cdot v^{\tau}_{j,t}$, and confirm previous finding (i.e., that the static model underestimates the willingness to pay for pollutants and living in close proximity to convicts). The results for the decomposition of flow utility under the two versions of expectations are presented in the second and third columns of table 3.9. Overall, results from the dynamic model without the expectation of exogenous type changes are only slightly biased upward for the primary amenities of interest. The largest difference implied by not taking account of exogenous type variation is in house size and unemployment. Not accounting for the forward looking expectations of family characteristics leads to an overestimate of marginal willingness to pay for extra space over 25%. Even though the two estimates are not statistically different for the two groups at any conventional significance level, the economic impact is substantial. Take an average house of 130 square meters and assume that the marginal willingness to pay is constant across square meters, then the impact from this difference in the estimates would be equivalent to additional flow utility worth of 33,150 DKK per year and with a subjective discount rate of 95% this difference would increase the house value by 633,000 DKK.

Table 3.9. MWTP in dynamic models.

	Static	Dynamic	
		Same type	Type change
m^2	998.8***	1,247.5***	992.5***
	(6.914)	(232.2)	(233.4)
$PM_{2.5}$	-2,413.1***	-4460.7**	-4,170.9**
	(41.90)	(1,580.3)	(1,600.7)
CO	-23.78***	-159.8***	-158.5***
	(0.625)	(29.13)	(29.24)
O_3	-980.8***	-5,680.1***	-5,545.5***
	(19.26)	(1,050.8)	(1,055.7)
Violent neighbors	37.17*	-1,809.1*	-1,702.7*
	(15.90)	(738.6)	(741.4)
Ethnic Danes	-42.93**	4,387.3***	4,664.3***
	(15.54)	(679.7)	(682.4)
Unemployment	-5,079.9***	-8,354.4***	-7,546.1**
	(74.41)	(2,295.9)	(2,350.0)
Population density	-325.9***	1,369.0	1,399.8
	(20.46)	(900.4)	(903.8)
Forest area	-16.56	-162.9	-188.2
	(11.30)	(492.6)	(496.4)
Urban area	108.8***	-44.24	-90.08
	(11.91)	(517.1)	(521.1)
Distance to freeway	-8.388	-219.2	-237.6
	(8.370)	(367.0)	(368.5)
Observations	656,760	656,760	656,760

Standard errors in parentheses.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

This bias emerges to a large extent as the result of household life cycle dynamics. In particular we can point to two oppositely directed effects for the population of housing owners. In general, more households in the population transition from having children to not having children living at home than the opposite and many households do not enter the housing market before having had the first child. This transition implies that the households will have lower willingness to pay for extra square meters as they take account of the fact that in the future they might not have a need for this space as children depart. Oppositely, households tend to increase wealth and income over the life cycle. If households internalize this in their location choices in early life, they might wish to invest more today in order to limit moving costs as income actually increases. However, overall the life cycle income effect is dominated by the effect of children.

3.6 Conclusions

Models of residential sorting have become ubiquitous in the non-market valuation literature, but the literature on residential sorting has, with few exceptions, ignored forward looking behavior due to computational costs and limited access to sufficiently detailed data. However, large transaction costs and evolving neighborhoods suggest that the housing market is inherently dynamic.

Our approach differs from this literature as we set up a dynamic model of neighborhood demand and we develop a multi-stage estimator that exploits unique features of the housing market. We overcome the data constraint by using unique population-wide Danish registers. These registers include detailed socio-economic information, housing unit characteristics, information on sales prices, and public valuations of houses, which we append with neighborhood characteristics from geographical data, air pollution data, and data on crime convictions. Through unique identifiers of families and housing units, we are able to follow household's characteristics and locations over time. With these data, we are in a unique position to study the residential location decision accounting for forward looking behavior and evolving state variables.

Our application focuses on house size and a number of neighborhood attributes including air quality measured by particulate matter, carbon monoxide, and ozone concentrations, and public safety measured by the presence of convicted criminals residing in the neighborhood. We estimate marginal willingness to pay for these neighborhood attributes. Our empirical application allows us to tackle the computational challenge of estimating a dynamic model of residential sorting with relative ease and the results suggest a bias in the static model like that found in previous works. We go beyond the simple interpretation of the results, however; the unique Danish data allow us to explore the special role of wealth in measuring willingness to pay. In particular, some individuals may be borrowing constrained and not have the resources to make a down payment on a house in a good neighborhood, or, they may not have the income stream to borrow to enable that choice. We find that, for individuals under these constraints, the marginal contribution of wealth to marginal willingness to pay for neighborhood amenities is much greater than for individuals who are likely unconstrained, suggesting that such constraints are important. Our data and approach furthermore allow for forward looking expectations with respect to future family structure, in particular children leaving the nest, and we similarly demonstrate the bias in the estimated value of house size, which leads to an overstatement of the value on house size of over 25%.

Acknowledgments

We gratefully acknowledge support from CAP - Cycles, Adjustments, and Policy, CREATES - Center for Research in Econometric Analysis of Time Series (DNRF78), funded by the Danish National Research Foundation, Aarhus University, Duke University as well as from Realdania and The Knowledge Center for Housing Economics. We also would like to thank the Rockwool Foundation Research Unit for providing us with neighborhood identifiers and Anna Piil for helping us understand how these were constructed. We thank participants at the 40th Anniverary Conference of Dansk Økonometrisk Selskab, the Research Group Meeting in connection with the Conference in Memory of Dale T. Mortensen, the Danish Graduate Programme in Economics Workshop, and participants at lunch seminars at Duke University and Aarhus University for comments and suggestions.

References

BAYER, P., N. KEOHANE, AND C. TIMMINS (2009): "Migration and hedonic valuation: The case of air quality," *Journal of Environmental Economics and Management*, 58, 1–14.

BAYER, P., R. MCMILLAN, A. MURPHY, AND C. TIMMINS (2015): "A Dynamic Model of Demand for Houses and Neighborhoods," *Econometrica*, Forthcomming.

BERKOWICZ, R. (2000): "A Simple Model for Urban Background Pollution," *Environmental Monitoring and Assessment*, 65, 259–267.

BERRY, S. T. (1994): "Estimating Discrete-Choice Models of Product Differentiation," *The RAND Journal of Economics*, 25, pp. 242–262.

BISHOP, K. (2012): "A Dynamic Model of Location Choice and Hedonic Valuation," Working Paper,Olin Business School,Washington University in St. Louis.

BLACK, S. E. (1999): "Do Better Schools Matter? Parental Valuation Of Elementary Education," *The Quarterly Journal of Economics*, 114, 577–599.

CHAY, K. Y. AND M. GREENSTONE (2005): "Does Air Quality Matter? Evidence from the Housing Market," *Journal of Political Economy*, 113, 376–424.

CHRISTENSEN, J. (1997): "The Danish Eulerian Hemispheric Mode - a three-dimensional air pollution model used for the Arctic," *Atmospheric Environment*, 31, 4169–4191.

DAMM, A. P. AND M.-L. SCHULTZ-NIELSEN (2008): "Danish Neighbourhoods: Construction and Relevance for Measurement of Residential Segregation," *Danish Journal of Economics*, 146, 241–262.

DAVIS, L. W. (2004): "The Effect of Health Risk on Housing Values: Evidence from a Cancer Cluster," *American Economic Review*, 94, 1693–1704.

ELLERMANN, T., J. BRANDT, O. HERTEL, S. LOFT, Z. J. ANDERSEN, O. RAASCHOU-NIELSEN, J. BØNLØKKE, AND T. SIGSGAARD (2014): "Luftforureningens indvirkning på sundheden i Danmark," Technical Report from DCE - Danish Centre for Environment and Energy 94, Aarhus University.

GAMPER-RABINDRAN, S. AND C. TIMMINS (2013): "Does cleanup of hazardous waste sites raise housing values? Evidence of spatially localized benefits," *Journal of Environmental Economics and Management*, 65, 345–360.

GOWRISANKARAN, G. AND M. RYSMAN (2012): "Dynamics of Consumer Demand for New Durable Goods," *Journal of Political Economy*, 120, 1173 – 1219.

GREENSTONE, M. AND J. GALLAGHER (2008): "Does Hazardous Waste Matter? Evidence from the Housing Market and the Superfund Program," *The Quarterly Journal of Economics*, 123, 951–1003.

HENDEL, I. AND A. NEVO (2006): "Measuring the Implications of Sales and Consumer Inventory Behavior," *Econometrica*, 74, pp. 1637–1673.

HOTZ, V. J. AND R. A. MILLER (1993): "Conditional Choice Probabilities and the Estimation of Dynamic Models," *The Review of Economic Studies*, 60, pp. 497–529.

KENNAN, J. AND J. R. WALKER (2011): "The Effect of Expected Income on Individual Migration Decisions," *Econometrica*, 79, 211–251.

KLEVEN, H. J., M. B. KNUDSEN, C. T. KREINER, S. PEDERSEN, AND E. SAEZ (2011): "Unwilling or Unable to Cheat? Evidence From a Tax Audit Experiment in Denmark," *Econometrica*, 79, 651–692.

KUMINOFF, N. V., V. K. SMITH, AND C. TIMMINS (2013): "The New Economics of Equilibrium Sorting and Policy Evaluation Using Housing Markets," *Journal of Economic Literature*, 51, 1007–62.

LINDEN, L. AND J. E. ROCKOFF (2008): "Estimates of the Impact of Crime Risk on Property Values from Megan's Laws," *American Economic Review*, 98, 1103–27.

MAGNAC, T. AND D. THESMAR (2002): "Identifying Dynamic Discrete Decision Processes," *Econometrica*, 70, 801–816.

MCFADDEN, D. (1981): "Econometric Models for Probabilistic Choice among Products," *The Journal of Business*, 53, S13–29.

MELNIKOV, O. (2013): "Demand For Differentiated Durable Products: The Case Of The U.S. Computer Printer Market," *Economic Inquiry*, 51, 1277–1298.

PALMQUIST, R. B. (2006): "Property Value Models," in *Handbook of Environmental Economics*, ed. by K. G. Mäler and J. R. Vincent, Elsevier, vol. 2 of *Handbook of Environmental Economics*, chap. 16, 763–819.

POPE, J. C. (2008): "Fear of crime and housing prices: Household reactions to sex offender registries," *Journal of Urban Economics*, 64, 601–614.

ROBACK, J. (1982): "Wages, Rents, and the Quality of Life," *Journal of Political Economy*, 90, 1257–78.

RUST, J. (1987): "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica*, 55, pp. 999–1033.

——— (1994): "Structural estimation of markov decision processes," in *Handbook of Econometrics*, ed. by R. F. Engle and D. McFadden, Elsevier, vol. 4 of *Handbook of Econometrics*, chap. 51, 3081–3143.

SCHIRALDI, P. (2011): "Automobile replacement: a dynamic structural approach," *RAND Journal of Economics*, 42, 266–291.

Appendix A: Data

This section gives an overview of the broader data set. We also describe features of the data that are not directly relevant for this paper, but could shed some light on the potential for future research.

As mentioned in the Data section, we combine data from various sources, which implies that we lose some observations due to differences in the coverage of these data sets. Table 3.10 shows the consequence of each of the merging steps.

Table 3.10. Overview of the merge of datasets.

Data set	Merged by	Obs. in	Obs. out	Nb. families
IDA		79,744,817		40,427,282
+ Building info[a]	*year, unit ID*	35,545,479	76,320,921	37,960,766
+ Ownership data[b]	*ID, year, unit ID*	26,457,634	76,320,921	37,960,766
+ Plot data[c]	*year, unit ID*	20,105,520	76,320,921	37,960,766
+ Public valuation[d]	*year, unit ID*	20,137,582	76,320,921	37,960,766
+ Sales data[e]	*year, unit ID*	1,024,245	76,320,921	37,960,766
+ Neighborhood ID[f]	*year, ID*	75,840,656	73,467,584	36,227,978

[a] Individuals without a Danish address are deleted, that is 3,39%. 1.99% of BOL is uninhabited and deleted.

[b] 4,649,092 are only in owner data set and are deleted. 30.34% of raw sample individuals owns a housing unit.

[c] 1,658,282 are only in plot data set and are deleted. 99.65% of sample is in plot data set.

[d] 744,274 are only in public valuation and are deleted. 96.15% of sample has a public valuation.

[e] Some properties are sold multiple times in one year - not allowed. We keep the last sale and delete 50.451 observations. Furthermore 131,045 are only in sales data and deleted. 3.22% of the sample is sold per year.

[f] In 2003 the entire sample has a neighborhood identifier. We delete 2.853.337 observations that can not be assigned a neighborhood.

The most critical step is the merge with neighborhood identifiers. As mentioned the neighborhoods are only constructed for the inhabited housing stock in 2003. This has some implications for the sample size in the years around 2003. In Figure 3.1 we plot the sample size of individuals and households after having merged with the neighborhood identifiers. We also include Figure 3.2 that plots the fraction of individuals across years that could be assigned a neighborhood. Furthermore, Figure 3.3 plots the number of house owning households in the final sample after trimming of data etc.

In table 3.11 we include other measures that provide a broader understanding of the data. In this paper we only use the information of homeowners' behavior, but only 45% of our raw sample of households actually own housing assets. It is also

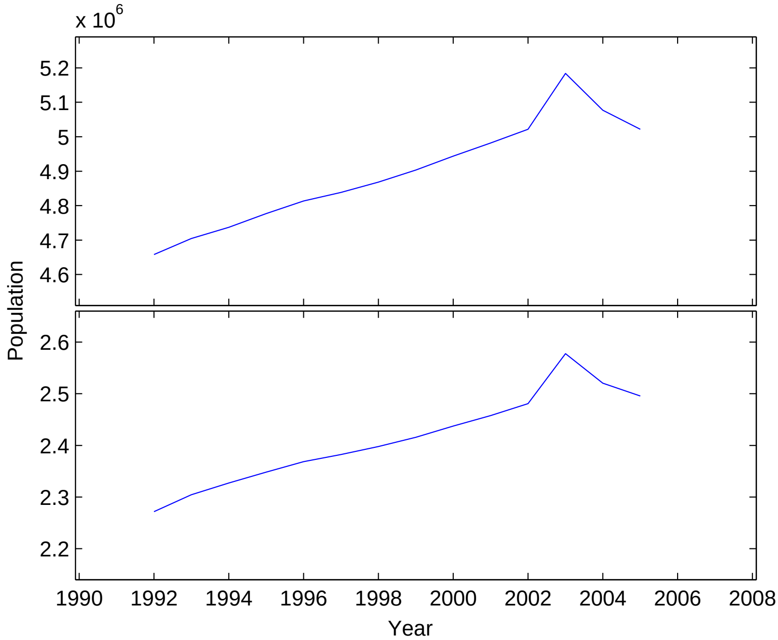


Figure 3.1. These graphs show the population size of individuals (top) and households (bottom) across years that are successfully assigned a neighborhood. The graphs cover both owners and renters.

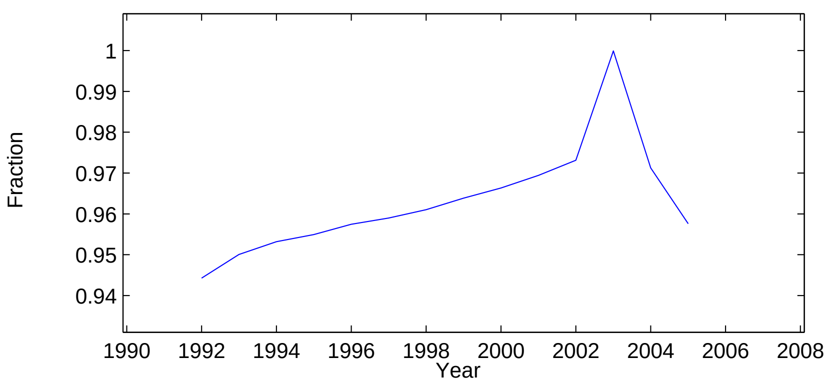


Figure 3.2. This graph shows the fraction of the raw sample that can be assigned a neighborhood identifier for each year.

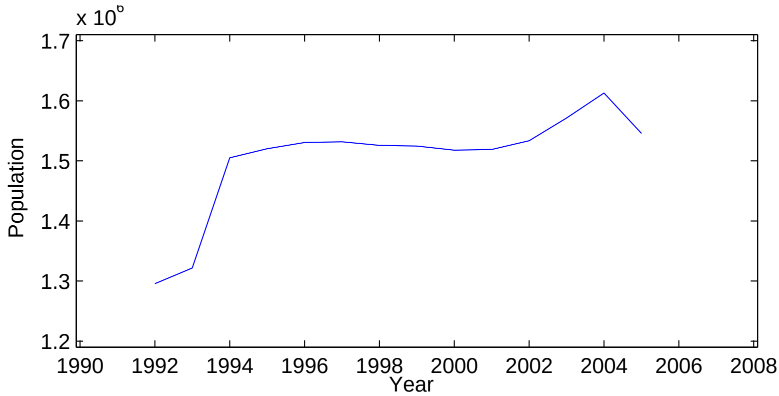


Figure 3.3. The graph shows the population size of households across years in the final sample, covering only owners.

important to note that the mean age of housing owners is in the fifties, the mean age of buyers is a lot lower as is evident from Figure 3.4 that shows the distribution of age for housing buyers. It is the choices of these individuals that we utilize in the paper, however, it also addresses the question of life-cycle importance, which we do strive to investigate.

The bottom of table 3.11 describes house specific attributes, which do not enter our model of neighborhood demand.

Table 3.11. Further Descriptive Statistics on Data.

Variable	Mean	Std. Dev.
Owner	44.58%	49.71%
Nb. Children	0.42	0.84
Nb. Adults	1.48	0.53
Age	52.27	15.50
Baths	1.17	0.47
Bathrooms	1.46	0.58
Floors	1.15	0.65
Rooms	4.63	1.50
Commercial m^2	0.83	7.88
Basement m^2	15.03	119.45
Plot m^2	16168.87	1740.62
House age	52.47	39.26

Table 3.12 shows the persistence of the amenities in the sence of a coefficient

from a somple AR(1) regression including a constant. Furthermore, the table includes a measure of geographical variation in the amenities, by looking af the standard deviation in observed amenities relative to the national mean.

Table 3.12. Variational Properties of Ameneties.

Variable	AR(1) coeff.	% Std. Dev.[a]
$PM_{2.5}$	0.801	9.69%
	(0.005)	
CO	0.950	4.45%
	(0.002)	
O_3	0.944	9.86%
	(0.003)	
Violent neighbors	1.011	6.18%
	(0.002)	
Ethnic Danes	1.033	7.06%
	(0.001)	
Unemployment	0.904	2.93%
	(0.002)	

[a] National standard deviation over national mean.

Appendix B: Construction of larger neighborhoods

Statistics Denmark has divided Denmark into 100 meter times 100 meter cells. Damm and Schultz-Nielsen (2008) in collaboration with an external firm with expertise in GIS (Geographical Information Software) produced the aggregation from the cells to two sets of neighborhoods. One is a small neighborhood where the neighborhood is at least inhabited by 150 households. This gives 9,400 neighborhoods in Denmark with an average of 572 individuals living in each one. The larger neighborhood contains at least 600 households. This gives 2,300 neighborhoods with an average of 2,300 individuals living in each one.

We have decided to aggregate the neighborhoods in order to get a reasonable number of sales per neighborhood. Therefore, we have aggregated the neighborhoods in a way that put weight on physical location and neighborhood 'alikeness' in terms of house types and ownership.

By regions in Denmark[27] the algorithm used for the aggregation is as follows

1 Pick the neighborhood with the largest east-coordinate.

[27] By regions we mean Bornholm, Zealand, Funen, and Jutland.

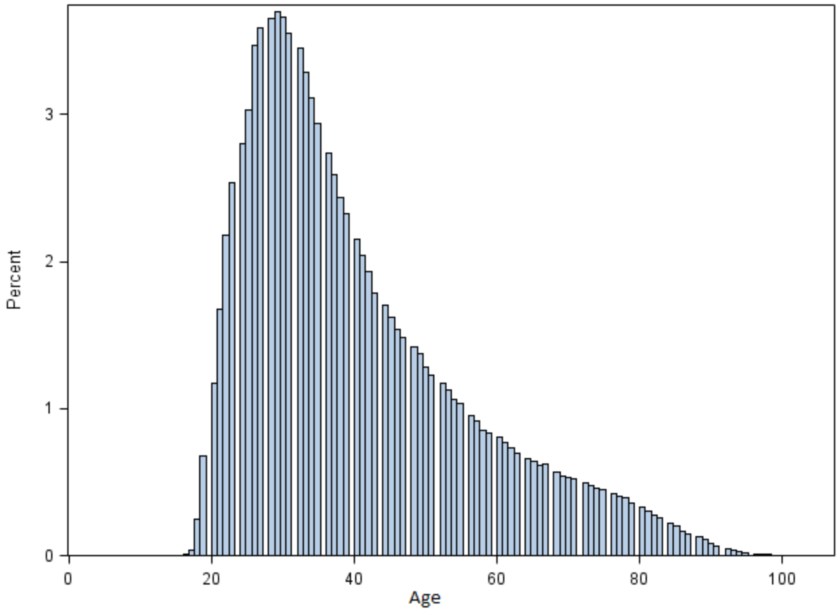


Figure 3.4. Histogram of the age distribution for housing buyers.

2 Based on centroid, calculate a metric for the nearest six neighborhoods. $\texttt{metric} = .5 \cdot (.7 \cdot \texttt{house type} + .3 \cdot \texttt{ownership}) + .5 \cdot \texttt{physical distance}$.

3 Merge the initial neighborhood with the one most alike based on the metric.

4 Run step 2-3 until there is a minimum of 3000 households in the new neighborhood.

5 Run 1-4 until 10 neighborhoods are left - these are aggregated/allocated manually.

DEPARTMENT OF ECONOMICS AND BUSINESS ECONOMICS

AARHUS UNIVERSITY
SCHOOL OF BUSINESS AND SOCIAL SCIENCES
www.econ.au.dk

PhD Theses since 1 July 2011

2011-4	Anders Bredahl Kock: Forecasting and Oracle Efficient Econometrics
2011-5	Christian Bach: The Game of Risk
2011-6	Stefan Holst Bache: Quantile Regression: Three Econometric Studies
2011:12	Bisheng Du: Essays on Advance Demand Information, Prioritization and Real Options in Inventory Management
2011:13	Christian Gormsen Schmidt: Exploring the Barriers to Globalization
2011:16	Dewi Fitriasari: Analyses of Social and Environmental Reporting as a Practice of Accountability to Stakeholders
2011:22	Sanne Hiller: Essays on International Trade and Migration: Firm Behavior, Networks and Barriers to Trade
2012-1	Johannes Tang Kristensen: From Determinants of Low Birthweight to Factor-Based Macroeconomic Forecasting
2012-2	Karina Hjortshøj Kjeldsen: Routing and Scheduling in Liner Shipping
2012-3	Soheil Abginehchi: Essays on Inventory Control in Presence of Multiple Sourcing
2012-4	Zhenjiang Qin: Essays on Heterogeneous Beliefs, Public Information, and Asset Pricing
2012-5	Lasse Frisgaard Gunnersen: Income Redistribution Policies
2012-6	Miriam Wüst: Essays on early investments in child health
2012-7	Yukai Yang: Modelling Nonlinear Vector Economic Time Series
2012-8	Lene Kjærsgaard: Empirical Essays of Active Labor Market Policy on Employment
2012-9	Henrik Nørholm: Structured Retail Products and Return Predictability
2012-10	Signe Frederiksen: Empirical Essays on Placements in Outside Home Care
2012-11	Mateusz P. Dziubinski: Essays on Financial Econometrics and Derivatives Pricing

2012-12	Jens Riis Andersen: Option Games under Incomplete Information
2012-13	Margit Malmmose: The Role of Management Accounting in New Public Management Reforms: Implications in a Socio-Political Health Care Context
2012-14	Laurent Callot: Large Panels and High-dimensional VAR
2012-15	Christian Rix-Nielsen: Strategic Investment
2013-1	Kenneth Lykke Sørensen: Essays on Wage Determination
2013-2	Tue Rauff Lind Christensen: Network Design Problems with Piecewise Linear Cost Functions
2013-3	Dominyka Sakalauskaite: A Challenge for Experts: Auditors, Forensic Specialists and the Detection of Fraud
2013-4	Rune Bysted: Essays on Innovative Work Behavior
2013-5	Mikkel Nørlem Hermansen: Longer Human Lifespan and the Retirement Decision
2013-6	Jannie H.G. Kristoffersen: Empirical Essays on Economics of Education
2013-7	Mark Strøm Kristoffersen: Essays on Economic Policies over the Business Cycle
2013-8	Philipp Meinen: Essays on Firms in International Trade
2013-9	Cédric Gorinas: Essays on Marginalization and Integration of Immigrants and Young Criminals – A Labour Economics Perspective
2013-10	Ina Charlotte Jäkel: Product Quality, Trade Policy, and Voter Preferences: Essays on International Trade
2013-11	Anna Gerstrøm: World Disruption - How Bankers Reconstruct the Financial Crisis: Essays on Interpretation
2013-12	Paola Andrea Barrientos Quiroga: Essays on Development Economics
2013-13	Peter Bodnar: Essays on Warehouse Operations
2013-14	Rune Vammen Lesner: Essays on Determinants of Inequality
2013-15	Peter Arendorf Bache: Firms and International Trade
2013-16	Anders Laugesen: On Complementarities, Heterogeneous Firms, and International Trade

2013-17 Anders Bruun Jonassen: Regression Discontinuity Analyses of the Disincentive Effects of Increasing Social Assistance

2014-1 David Sloth Pedersen: A Journey into the Dark Arts of Quantitative Finance

2014-2 Martin Schultz-Nielsen: Optimal Corporate Investments and Capital Structure

2014-3 Lukas Bach: Routing and Scheduling Problems - Optimization using Exact and Heuristic Methods

2014-4 Tanja Groth: Regulatory impacts in relation to a renewable fuel CHP technology: A financial and socioeconomic analysis

2014-5 Niels Strange Hansen: Forecasting Based on Unobserved Variables

2014-6 Ritwik Banerjee: Economics of Misbehavior

2014-7 Christina Annette Gravert: Giving and Taking – Essays in Experimental Economics

2014-8 Astrid Hanghøj: Papers in purchasing and supply management: A capability-based perspective

2014-9 Nima Nonejad: Essays in Applied Bayesian Particle and Markov Chain Monte Carlo Techniques in Time Series Econometrics

2014-10 Tine L. Mundbjerg Eriksen: Essays on Bullying: an Economist's Perspective

2014-11 Sashka Dimova: Essays on Job Search Assistance

2014-12 Rasmus Tangsgaard Varneskov: Econometric Analysis of Volatility in Financial Additive Noise Models

2015-1 Anne Floor Brix: Estimation of Continuous Time Models Driven by Lévy Processes

2015-2 Kasper Vinther Olesen: Realizing Conditional Distributions and Coherence Across Financial Asset Classes

2015-3 Manuel Sebastian Lukas: Estimation and Model Specification for Econometric Forecasting

2015-4 Sofie Theilade Nyland Brodersen: Essays on Job Search Assistance and Labor Market Outcomes

2015-5 Jesper Nydam Wulff: Empirical Research in Foreign Market Entry Mode

2015-6 Sanni Nørgaard Breining: The Sibling Relationship Dynamics and Spillovers

2015-7 Marie Herly: Empirical Studies of Earnings Quality

2015-8 Stine Ludvig Bech: The Relationship between Caseworkers and Unemployed Workers

2015-9 Kaleb Girma Abreha: Empirical Essays on Heterogeneous Firms and International Trade

2015-10 Jeanne Andersen: Modelling and Optimisation of Renewable Energy Systems

2015-11 Rasmus Landersø: Essays in the Economics of Crime

2015-12 Juan Carlos Parra-Alvarez: Solution Methods and Inference in Continuous-Time Dynamic Equilibrium Economies (with Applications in Asset Pricing and Income Fluctuation Models)

2015-13 Sakshi Girdhar: The Internationalization of Big Accounting Firms and the Implications on their Practices and Structures: An Institutional Analysis

2015-14 Wenjing Wang: Corporate Innovation, R&D Personnel and External Knowledge Utilization

2015-15 Lene Gilje Justesen: Empirical Banking

2015-16 Jonas Maibom: Structural and Empirical Analysis of the Labour Market

2015-17 Sylvanus Kwaku Afesorgbor: Essays on International Economics and Development

2015-18 Orimar Sauri, Lévy Semistationary Models with Applications in Energy Markets

2015-19 Kristine Vasiljeva: Essays on Immigration in a Generous Welfare State

2015-20 Jonas Nygaard Eriksen: Business Cycles and Expected Returns

2015-21 Simon Juul Hviid: Dynamic Models of the Housing Market

ISBN: 9788793195295