Leverage Regulation and Housing Inequality

Jens Soerlie Kvaerner, Nicola Pavanini, Yushi Peng*

July 2023

Abstract

We estimate an equilibrium model of housing demand and supply. The model allows us to quantify the distributional effects of leverage regulation on mobility and access to high-quality housing. We match the population of households in Norway in 2010-2018, with demographic and financial characteristics, to the universe of housing transactions. Our model features households' dynamic renting and owning choices, speculators' housing portfolio rebalancing, and equilibrium pricing across housing products via a market clearing condition. We recover households' willingness to pay for housing quality and moving costs across the income distribution. Our counterfactuals quantify the regressive effects of tighter loan-to-income (LTI) limits, and document how these depend on household preferences and can be offset limiting speculators' real estate trading.

JEL: D15; D31; G21; G51; R21; R31

Keywords: Housing demand, housing supply, leverage regulation, housing quality, dynamic discrete choice

*Kvaerner: Tilburg University, jkverner@gmail.com. Pavanini: Tilburg University and CEPR, n.pavanini@tilburguniversity.edu. Peng: Tilburg University, Y.Peng@tilburguniversity.edu. We thank seminar participants at Copenhagen Business School, Tilburg TILEC, Tilburg Structural Econometrics Group, Bank of England. Nicola Pavanini gratefully acknowledges financial support from the European Research Council under the Grant Agreement No. 101042069.

1 Introduction

Housing and neighborhood choices impact households' welfare and wellbeing. Growing up in high-quality neighborhoods positively affects physical and mental health (Ludwig et al., 2012), future college attendance and earnings (Chetty et al., 2016; Chetty and Hendren, 2018b), as well as fertility and marriage patterns (Chetty and Hendren, 2018a). In addition, expected excess returns on real estate wealth are substantially higher than those on financial and pension wealth (Bach et al., 2020). This evidence highlights how access to housing wealth has the potential to affect intergenerational mobility and reduce wealth inequality.

Leverage is the key driver of access to housing, but excessive leverage can generate negative effects on the economy. Mian et al. (2017) provide evidence that in the last 50 years, higher household debt to GDP ratio predicts lower GDP growth and higher unemployment, and Mian and Sufi (2009) highlight how excessive lending to subprime borrowers was one of the leading causes of the 2008 financial crisis. To address these negative consequences of the leverage cycle, Geanakoplos (2009) provides theoretical grounds for macro-prudential interventions, suggesting that central banks should regulate leverage with tools such as Loan-To-Value (LTV) and Loan-To-Income (LTI) limits.

While there is evidence of the positive role of macro-prudential regulations on financial stability, as tighter LTI and LTV limits reduce originations of risky mortgages (DeFusco et al., 2020) and improve household debt solvency (van Bekkum et al., 2019), little is known about their distributional effects on household mobility and access to high-quality housing. However, in order to uncover these effects, it is necessary to separately identify the role of households' financial constraints from that of preferences for neighborhood quality in residential choices. This presents two challenges, as observing financial frictions requires detailed household-level data on income and wealth, and recovering preferences can only be achieved estimating a model of household residential choice.

Our paper addresses these challenges by developing a structural framework of housing demand and supply with two novel features. First, it explicitly incorporates households' affordability constraints, thanks to detailed household-level data on income and wealth. Second, it estimates households' heterogeneous preferences for housing and neighborhood quality, as well as moving costs, across the income distribution. We use our model to simulate counterfactual scenarios with tighter LTI limits, and document how this affects mobility choices and access to high-quality housing across the income distribution. Our model and data allow us to separate the effect of financial constraints from that of households' willingness to pay for quality. We quantify the regressive effects of these tighter limits, show how they depend on households' preferences, and document how they can be offset by limiting speculators' trading in real estate markets.

This framework extends the most recent literature on structural models of residential choices. We generalize the dynamic framework of households' residential location developed by Bayer et al. (2016), introducing five new features. First, we allow for two different agents to demand and supply housing products: financially constrained households, who mostly own and exchange a single housing product and face transaction costs, and financially unconstrained speculators, who own portfolios of properties and face no transaction costs. Second, we distinguish between owners and renters among households, allowing renters to become owners and vice versa. Third, all households can decide every period to stay in their current property, not just those who purchased a house within our sample period. Fourth, having data on households' balance sheets we can explicitly model their heterogeneous affordability constraints. Last, we model the equilibrium pricing in housing markets via a market clearing condition incorporating households' and speculators' demand and supply of properties.

Part of the modeling innovations that we introduce can be implemented due to administrative data from Norway. The dataset covers the period from 2010 to 2018 and is geographically restricted to the capital Oslo. It contains three key components. First, an individual-year-level dataset for the entire population that includes detailed information on income, debt, wealth, house value, liquid assets, cash, social security payments, demographics, education, and location of the residence. The data also links different individuals belonging to the same household, allowing us to calculate the income and wealth of a household unit, which will be key determinants of housing and mortgage affordability in the model. This data is provided by Statistics Norway. Second, a transaction-level dataset for the universe of housing transactions in Norway from Eiendomsverdi AS. For the period 2010-2018 in Oslo there are approximately 200,000 housing transactions. The data includes the unique tax identifier of both buyer and seller for each transaction, allowing us to merge it with the first dataset. The data also contains detailed information on the exact location of the transacted house, its size in square meters, its age, and the transaction price. Third, a set of district-level characteristics from the official statistics published by the municipality of Oslo. We use this data to construct a measure of neighborhood quality.

We define three levels of neighborhood quality and apply a K-means algorithm to allocate neighborhoods to low, medium, and high quality. We use five variables to perform this allocation, measuring education, poverty, crime, health, and happiness. We then show that our quality categorization strongly correlates with house prices and the income and financial wealth of neighborhood residents. Furthermore, we document how life outcomes of individuals between 26 and 35 years of age in 2015 are influenced by the quality of the neighborhood they grew up in as children. Our results show that growing up in a high-quality neighborhood, relative to a low-quality one, significantly increases the probability of holding a master's degree, holding a Ph.D., being a homeowner, and participating in the stock market. However, once we control for parents' education as in Heckman and Landersø (2022), we do not find a significant effect for homeownership and stock market participation anymore. This is consistent with parents passing on their wealth to children, and with neighborhood quality being an important determinant of children's educational attainments.

To provide descriptive evidence on the importance of leverage regulation, we show the impact of introducing an LTI limit in Norway on households' choice sets. Norway set in 2017 the LTI limit to 5, while before that, no explicit restrictions were imposed. We divide housing products across size (small vs large) and quality (low vs high). We show that households in the top 10% of the income distribution, who could afford any house before the LTI limit, can still afford all low- and high-quality small properties after 2017, but lose access to around 10% of low-quality large houses and to about 50% of high-quality large properties. On the other hand, households with median income, who could afford all small properties, lose access to 50% of small low-quality and 70% of small high-quality. In terms of large properties, median-income households could afford 90% of low quality and 50% of high-quality before 2017, but once the LTI limit is introduced can only afford 15% of the small low-quality and almost none of the large high-quality properties.

We use these descriptive results as motivating evidence for our structural framework, where we model the housing choices of speculators and households. Speculators are defined as owners of more than one property, and we model their optimal housing portfolio allocation every period. Households can be homeowners or renters and can choose whether to stay in their current residence or to move to another one chosen among a set of housing options, differentiated based on the location and size of the property. Each year, households' housing decisions will be the result of the maximization of their lifetime expected utility, whose preference parameters we will estimate, recovering willingness to pay for house and neighborhood quality across the income distribution. When households move to purchase a house, they incur financial moving costs of both buying the new property and selling the old one. Additionally, when households move as either owners or renters, they incur psychological moving costs. These two types of fixed moving costs will be expressed as a function of household income and demographic characteristics, and justify the dynamic nature of our model, where households form expectations over their future valuation and price of housing products. Alternatively, households can also choose to move out of the Oslo area.

Households are divided into types, defined by their disposable income, wealth, family size, and home ownership status. This will allow for heterogeneous choice sets across household types, which will be bound by affordability constraints based on LTI and LTV restrictions, paired respectively with household disposable income and net wealth. If a household moves, its new type reflects the reduction in wealth due to moving costs. Data on homeownership allows us to distinguish between five alternative decisions that households might be making, which the literature had not been able to fully capture before. First, homeowners selling their current property to buy a new one. Second, homeowners selling their current property to become renters. Third, renters buying a property and becoming homeowners. Four, renters moving to another rented property. Last, any type of household remaining inactive.

Our estimation delivers two sets of results. First, we recover financial and psychological moving costs across the income distribution and find that richer households face lower costs across both dimensions, while older households and larger families experience larger psychological moving costs. When recovering the determinants of households' flow utilities that drive their housing choice, we find that households' willingness to pay for housing quality is increasing with income, and that low-income households' willingness to pay is increasing in house size, while the opposite is true for high-income ones.

We use our model to simulate a counterfactual scenario with a tighter LTI limit than what was implemented in 2017, setting it to 3 instead of 5. We show that a more stringent LTI limit reduces the share of housing products and the share of high-quality houses differently across the income distribution. The choice sets of the lowest and highest income groups are largely unaffected, because the former could already only afford a small fraction of properties, and for the latter, a tight LTI limit is still mostly not binding. However, households in the middle of the income distribution experience at most an 18.6% reduction in the share of housing products they can afford, and at most a 25% reduction in the share of high-quality property they can consider buying.

We then focus only on households who actually moved during our sample period, which should be more directly affected by the policy in their mobility pattern. We find that a tighter LTI limit reduces the probability of going from renting to owning, equivalent to the probability of buying for first-time buyers, by 9.6% for the lowest income group, and has no effect on the highest income group. Overall, we document the regressive effect of tighter LTI limits quantifying the reduction in mobility across income and housing quality distributions, and the reallocation of households from ownership to rental.

We conduct two additional counterfactuals to propose policies that can offset these regressive effects of LTI limits. First, we simulate a scenario in which speculators' trading is limited, by making them more price inelastic. This results in an increase in prices for properties that are demanded by high-income households, due to their higher willingness to pay, while the opposite happens for houses preferred by low-income households. This leads to an increase in household mobility, as speculators partially withdraw from trading, and to a reallocation of households from low- to high-quality neighborhoods. Second, we simulate a scenario where households' preferences are shifted closer to those of the top income group, which could be achieved with policies such as schooling vouchers or public housing assistance programs. This also increases households' mobility, and leads to a reallocation from low- to high-quality neighborhoods, mostly for low-income households.

Related Literature. We contribute to three main strands of the literature. First, due to the recent introduction of macro-prudential regulations, empirical evidence on their effects is still scarce and has developed only recently. Acharya et al. (2022) investigate how changes to LTI and LTV limits in Ireland affect mortgage credit and house prices. They find that mortgage credit is reallocated from low- to high-income borrowers and from urban to rural counties, slowing down house price growth. Peydrò et al. (2020) use UK mortgage data to show that banks more constrained by a larger exposure to high-LTI mortgages cut credit supply more to low-income borrowers, lowering house price growth. DeFusco et al. (2020) show that the Dodd-Frank act in the U.S., introducing a rule akin to a tighter LTI limit, has managed to substantially curb originations of risky mortgages. van Bekkum et al. (2019) use Dutch data and a reduction in LTV limits to show that liquidity constrained households reduce leverage and are less likely to buy a property, but have better solvency on their debt. Similar results are found by Han et al. (2021) for Canada. More generally, Baker (2018) documents how heterogeneity in households' consumption elasticity is entirely driven by credit and liquidity, highlighting the key role of financial constraints to address household inequality.

Our contribution to this first strand is twofold. First, most of the papers mentioned above focus on how stricter leverage limits affect mortgage outcomes, but none quantifies the effects of these interventions on mobility and access to high-quality housing across the income distribution. Second, the state-of-the-art literature has so far focused on reduced form methods that mostly identify local average treatment effects, but have no ability to predict how alternative limits would impact all households across the income and wealth distribution. From a policymaker's perspective, it is instead crucial to have access to comprehensive predictions on the overall impact of regulatory changes, together with general equilibrium effects on housing demand and supply, and house prices. This can only be delivered by a structural model of housing demand and supply, which explicitly incorporates LTV and LTI constraints, and is able to predict how household leverage and residential choices would change across a range of counterfactual leverage limits. This is the framework that we develop in this paper.

The second branch of the literature that we contribute to is on structural equilibrium frameworks to model housing choices. Bajari et al. (2013), Bayer et al. (2016), and Epple et al. (2020) are examples of dynamic structural housing models.[1] Peng (2021) applies these models to the Chinese housing market, incorporating into the model financing conditions, but only with aggregate data. Almagro and Domínguez-Iino (2021) build a model of residential sorting that quantifies the importance of endogenous location amenities for inequality. There is also an important part of the literature developing quantitative general equilibrium macroeconomic models that study how changes in financing conditions affect housing choices and equilibrium house prices (Kiyotaki et al., 2011; Sommer et al., 2013; Favilukis et al., 2017). A limitation common across these papers is that households' financing decision is not considered, or for the macro models it is considered from an aggregate perspective, without household level data on mortgages and detailed loan conditions. The macro approach cannot, therefore, derive distributional implications.

Last, we contribute to the literature on the effects of social mobility on households' economic outcomes. Using a randomized housing mobility experiment, recent work documents that households moving to neighborhoods with less poverty experience long-term improvements in physical and mental health (Ludwig et al., 2012), as well as in children's future college attendance and earnings (Chetty et al., 2016). Chetty and Hendren (2018a) report similar findings, showing that longer exposure to better neighborhoods significantly improves children's outcomes. Chetty and Hendren (2018b) complement these results with evidence of adulthood increase in income due to exposure to better U.S. counties. Despite these benefits of moving to high-quality neighborhoods, Bayer et al. (2007) find that U.S. households prefer to self-segregate on the basis of race and education. Our framework complements this strand of literature showing how the combination of leverage regulation and households' preferences for neighborhood characteristics affect social mobility, as well as the distributional effects of relaxing those regulations.

The rest of the paper is organized as follows. Section 2 describes the data and presents some stylized facts that motivate our model. Section 3 introduces the model, while Section 4 outlines the estimation. Section 5 presents the results, Section 6 shows the counterfactuals, and Section 7 concludes.

2 Data and Stylized Facts

In this section, we describe our data sources, define the variables we use, and present a set of summary statistics and stylized facts that motivate our analysis.

[1]A comprehensive survey of the literature on structural estimation in urban economics is Holmes and Sieg (2015).

2.1 Data Sources and Variable Definitions

We rely on three data sources. First, the real estate property and transaction data from Eiendomsverdi AS (henceforth EV). EV estimates the market value for the Norwegian residential real estate market, both for individual properties and for portfolios of properties.[2] The data are available from 2010 to 2019. The dataset includes all housing transactions in Oslo and a rich set of housing attributes. We observe the identity of the buyer and the seller, the date of listing, the transaction price, the number of livable square meters, the number of rooms, and the district where the property is located.

Second, from the municipality of Oslo, we get district-level data on the stock of housing, rental prices, and district-level characteristics. Rental prices are reported by room number for five aggregate districts. District-level characteristics, which we use to assign a quality score to each district, are available at a more granular level than rental prices.[3] We explain all district-level variables and how we use them in Section 2.2.2.

Third, the Norwegian Tax Registry (NTR) and Statistics Norway (SSB) provide the household-level data. NTR is responsible for collecting income and wealth taxes in Norway. Employers, banks, and public agencies are obliged by law to submit personal information on income, total assets, and transfers to the NTR before the end of April each year, which is when individuals are required to submit their tax returns. Individuals are accountable for the information in their tax returns, and the submission of inaccurate information is punishable by law.

We observe the birth date of each individual, together with the number of children. We merge data on demographics with data on financial information. For each individual, we define income Y as the sum of gross salary and pension plus net capital income and total government transfers. We define net worth, A, as the sum of financial wealth and total assets minus the value assessment of principal residence and debt. We exclude the value assessment of the principal residence because we later add the price of the house to the net worth. We define home ownership as a variable that takes the value of one for all individuals with a positive value assessment of the principal residence. We use the variable $h \in \mathcal{H} = \{0, 1\}$ to separate renters ($h = 0$) from homeowners ($h = 1$).

We distinguish between individuals living alone and individuals with a partner. We obtain the national identity number of the spouse/registered partner from the SSB's population statistics and use this information to classify an individual into a one-adult household or more than one adult household. We refer to the two household types simply as singles and couples and let G denote the set of demographic variables, including family size, age of the

[2] More information is available here: https://eiendomsverdi.no/

[3] All data is freely available here: https://statistikkbanken.oslo.kommune.no/webview/

household head, and family type. For tax purposes, the household can allocate wealth in a way that gives the lowest wealth tax. Thus, there are no incentives for tax-motivated asset allocation within the household. All households are one-family households.

We calculate the same statistics for both household types. However, for couples, we aggregate total income (Y), net worth (A), and the number of children at the household level. For age and home ownership, we select the maximum in the household. We keep the anonymized identifier of the oldest individual in the household and refer to this individual as the household head. Finally, we require all households to have non-negative financial wealth, debt, and total income and have a household head of at least 18 years of age at the end of each year.

2.1.1 The Housing Choice Set

Oslo is divided into 18 districts. The 15 largest districts cover approximately 99.5% of the housing stock, so we focus on those. We define the collection of these 15 districts as the set $\mathcal{D} = \{1, 2, ..., 15\}$. Each district $d \in \mathcal{D}$ is populated with a set of housing units $u \in \mathcal{U}$, distinguished by the number of rooms in the housing unit, $\mathcal{U} = \{1-2, 3, 4^{+}\}$. $1-2$ includes small housing units with at most two rooms, 3 includes medium units with three rooms, and 4^{+} includes large units with four or more rooms. We choose this grid to ensure we have multiple transactions for each housing product at each point in time. The Cartesian product $\mathcal{J} = \mathcal{D} \times \mathcal{U} = \{(d, u) | d \in \mathcal{D}, u \in \mathcal{U}\}$ gives a total of $\mathcal{J} = \{1, 2, ..., 45\}$ housing products. The 45 housing products account for more than 95% of the housing stock in Oslo.[4]

We verify that our discretization scheme of the housing market accounts for a large share of the price variation in the data. Specifically, we decompose the natural logarithm of transaction prices each year into within-product variability and between-product variability. Table 1 presents the results of the variance decomposition. Overall, between 55% and 65% of the total variation in house prices are attributable to between-housing-product variation. We calculate the price $P_{j,t}$ for each product $j \in \mathcal{J}$ in each year t as the average transaction price across all transactions for product j.

2.2 Summary Statistics

2.2.1 The Sample

We construct our sample dynamically. We begin in 2010 and select all households who live in Oslo at the end of the year. For all other years, we consider all households that

[4] The excluded fraction of the housing stock (ca. 5%) comes from 3 excluded districts (0.5%) and missing information on the number of rooms (4.5%).

Table 1 VARIANCE DECOMPOSITION OF HOUSE PRICES

Year	Total Variance	Between Variance	Within Variance	Between % Total	Within % Total
2010	0.29	0.16	0.13	0.55	0.45
2011	0.28	0.15	0.13	0.54	0.46
2012	0.26	0.14	0.11	0.56	0.44
2013	0.23	0.13	0.10	0.58	0.42
2014	0.23	0.13	0.09	0.59	0.41
2015	0.20	0.12	0.08	0.59	0.41
2016	0.17	0.10	0.07	0.60	0.40
2017	0.17	0.11	0.07	0.63	0.37
2018	0.18	0.12	0.06	0.65	0.35

Notes: This table reports variance decomposition of the natural logarithm of house prices by year. The groups are the 45 housing products constructed as follows: we first define the collection of 15 districts as the set $\mathcal{D} = \{1, 2, ..., 15\}$. Each district $d \in \mathcal{D}$ is populated with a set of housing units distinguished by the number of rooms in the housing unit, $\mathcal{U} = \{1-2, 3, 4^+\}$. $1-2$ includes 1 and 2 rooms apartments and 3^+ includes housing units with four or more rooms. The Cartesian product $\mathcal{J} = \mathcal{D} \times \mathcal{U} = \{(d, u)|d \in \mathcal{D}, u \in \mathcal{U}\}$ gives a total of $\mathcal{J} = \{1, 2, ..., 45\}$ housing products.

either lived in Oslo last year or ended up living in Oslo this year. The mobility options are, therefore: moving to Oslo from outside Oslo, staying in Oslo, moving to the same district within Oslo, moving to another district within Oslo, or leaving Oslo. We refer to the latter as the outside option. Details about the sample construction are in Appendix B. In total, we have approximately 580,000 unique households from 2010 to 2018 and roughly 3.2 million observations. Table 2 presents descriptive statistics for 2015.

2.2.2 Identifying Neighborhood Quality

We use three levels of neighborhood quality and apply K-means to generate three quality clusters. Districts of the same quality have comparable scores on five indicators. These indicators are GPA of primary school, which is the sole criterion for admission to upper secondary school, the number of reports to the child welfare service per capita,[5] criminal offenses by individuals between 0 and 17 years of age per capita, and answers to two survey questions. The first survey question is a self-assessment of health.[6] The second survey question is the score on a happiness index.[7] We use the scores at the end of 2015 because it

[5]The total number of reports by district by year is divided by district population and then multiplied by 1,000. This normalization is applied to make all the variables in comparable units.

[6]The question is: "How satisfied or dissatisfied are you with your health?" The options are: "Very dissatisfied", "Slightly dissatisfied", "Neither satisfied nor dissatisfied", "Slightly satisfied", or "Very satisfied". We count those who answer "Slightly satisfied", "Very satisfied" as a proportion of all those who responded.

[7]The question is: "How satisfied are you with your local environment?". The options are: "Very dissatisfied", "Slightly dissatisfied", "Neither satisfied nor dissatisfied", "Slightly satisfied", or "Very satisfied".

Table 2 Descriptive Statistics

Panel A: Demographic and Financial Data

	Homeowners				
	Mean	Std Dev	10th	50th	90th
Fraction couples	0.73	0.44	0.00	1.00	1.00
Age	46	16	28	42	69
Number of children	1.2	1.2	0.0	1.0	3.0
Total income	699	1,553	262	540	1,075
Gross wealth	5,010	20,767	2,529	3,590	7,248
Debt	1,575	2,196	14	1,277	3,362

	Renters				
	Mean	Std Dev	10th	50th	90th
Fraction couples	0.43	0.50	0.00	0.00	1.00
Age	37	15	23	32	59
Number of children	0.73	1.24	0.00	0.00	2.00
Total income	329	693	36	279	594
Gross wealth	301	5,205	0	39	448
Debt	282	1,025	0	50	560

	Homeowners		**Renters**	
	N of Obs	Share	N of Obs	Share
Stayers	167,146	0.83	101,476	0.60
Leaving Oslo	7,031	0.03	8,669	0.05
Entering Oslo	7,189	0.04	26,368	0.16
Move within district in Oslo	8,037	0.04	10,519	0.06
Move between district in Oslo	13,181	0.07	21,453	0.13
Total	202,584	1.00	168,485	1.00

Notes: This table reports descriptive statistics of our sample in 2015. Panel A reports demographic and financial data. Panel B reports mobility statistics. We report the descriptive statistics for homeowners and renters separately. All financial variables are reported in NOK thousands.

is the only year we observe scores on all five quality indicators.

Table 3 presents the results from a cross-sectional regression of neighborhood characteristics on a constant and two dummy variables for neighborhood quality. Each regression includes 15 data points. The low (high) quality dummy variable takes the value of one if a district belongs to the low (high) quality neighborhood. The constant serves as a reference point. We refer to it as the baseline.

Table 3 HOUSING QUALITY REGRESSIONS

	Outcome variable						
	Price	Income	Financial Wealth	Age	N of Children	Covid Cases	Not Vaccine
Baseline	5,191.56***	420.06***	110.60***	39.43***	0.74***	0.05***	0.43***
	(275.76)	(14.18)	(14.58)	(1.56)	(0.15)	(0.005)	(0.04)
Low Quality	−1,670.54***	−42.20**	−39.66*	4.26*	0.53**	0.03***	−0.03
	(356.00)	(18.30)	(18.83)	(2.02)	(0.20)	(0.01)	(0.05)
High Quality	1,625.81***	67.62***	72.09***	5.97**	0.45**	−0.01	−0.10*
	(369.97)	(19.02)	(19.57)	(2.10)	(0.21)	(0.01)	(0.05)
Districts	15	15	15	15	15	15	15
Adjusted R^2	0.87	0.74	0.73	0.31	0.30	0.81	0.15

Notes: This table reports the results from regressing a set of household and neighborhood characteristics on a constant and two dummy variables for neighborhood quality. The "Low Quality" ("High Quality") dummy variable takes the value of one if the district belongs to the low (high) quality area. The constant represents the average score of a district between low and high quality. We include the following variables as dependent variables: Price for a 3-room apartment, median income and wealth, age and number of children, Covid-19 cases per 10,000 citizens as of June 2021, and the fraction of individuals above 18 years old not vaccinated against Covid-19.

The first column reveals large price differences for a 3-room apartment between neighborhoods. The high-quality neighborhood is about 30% more expensive than the baseline, while the low-quality neighborhood is about 30% cheaper. This shows that households' willingness to pay for a 3-room apartment is increasing in our measure of neighborhood quality. The second and third columns show that the average household income and financial wealth increase monotonically with neighborhood quality. In contrast, the age of the household and the number of children do not. The fifth column shows the number of Covid-19 cases per 10,000 citizens as of June 2021. The regression shows that the districts in low-quality neighborhoods had the most cases. The last column has as the dependent variable the fraction of

We count those who answer "Slightly satisfied", "Very satisfied" as a proportion of all those who responded.

individuals older than 18 years old not vaccinated against Covid-19.[8] We see that districts in the high-quality neighborhood had the highest vaccination rates.

2.3 Neighborhood Quality and Life Outcomes

With a measure of neighborhood quality, we can calculate the correlation between the quality of the neighborhood people grow up in and their outcomes later in life, in line with Chetty and Hendren (2018a). Following Heckman and Landersø (2022), we also control for parents' education. We restrict the analysis to 2015 and include 10 cohorts that are between 26 and 35 years old. In 1990 these people were 10 years old or younger. We run two sets of cross-sectional regressions, with dummies for whether an individual i in 1990 was resident in a low or high-quality neighborhood, and use the middle-quality neighborhoods as the reference group (i.e., $D_{i,1990} = 0$) in both specifications:

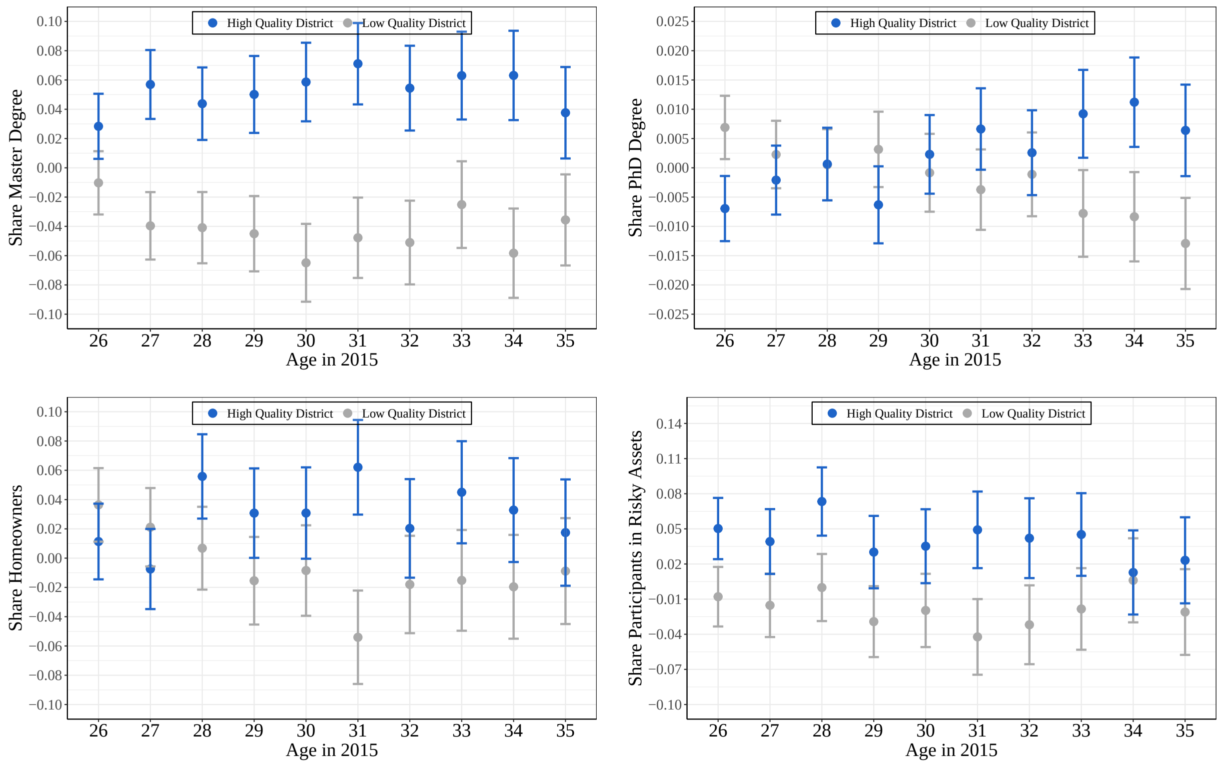
$$I_{i,2015} = \sum_{j=26}^{35} \gamma_j \mathbb{1}\{age_{i,2015} = j\} + \sum_{j=26}^{35} \eta_j \mathbb{1}\{age_{i,2015} = j\} \times D_{i,1990} + \delta E_{i,1990} + \epsilon_{i,2015}, \quad (1)$$

where $E_{i,1990}$ are dummies for each parent's years of education in 1990, and $I_{i,2015}$ is an indicator variable that measures life outcomes in 2015. We include four life outcomes: having a master's degree, a Ph.D. degree, being a homeowner, and participating in the stock market. In our sample 26% of individuals have at least a master's degree, 1% have a Ph.D., 55% are homeowners, and 58% are stock market participants. We use the middle-quality neighborhood to identify γ_j. It measures the proportion of individuals at a given age in a middle-quality neighborhood with a dependent variable of one. The coefficients of interest are η_j. They measure the difference in the proportion of individuals at a given age in 2015, with a dependent variable of one, who grew up in either a high or a low-quality neighborhood relative to a neutral one. Figure 6 presents the results from the estimation, showing that growing up in a high-quality neighborhood delivers significantly higher achievements across all four outcomes, while the opposite is true when growing up in a low-quality neighborhood.

In Appendix A we report the same figures based on regressions without controlling for parents' education. The comparison with Figure 6 highlights the importance of controlling for parents' education to identify neighborhood effects. Parents' education explains most of the difference in homeownership and stock market participation between high and low-quality districts, which is consistent with parents' passing on their wealth to children. On the

[8] We use not vaccinated rather than having 1 or 2 vaccines as our outcome measures. In mid-2021, all citizens in Oslo had received at least one offer to be vaccinated. Therefore, at that time, the lack of at least one vaccine reflects an active choice and is not due to different waiting lists within Oslo.

other hand, even controlling for parents' education, neighborhood quality still determines a significant difference between children's educational attainments.



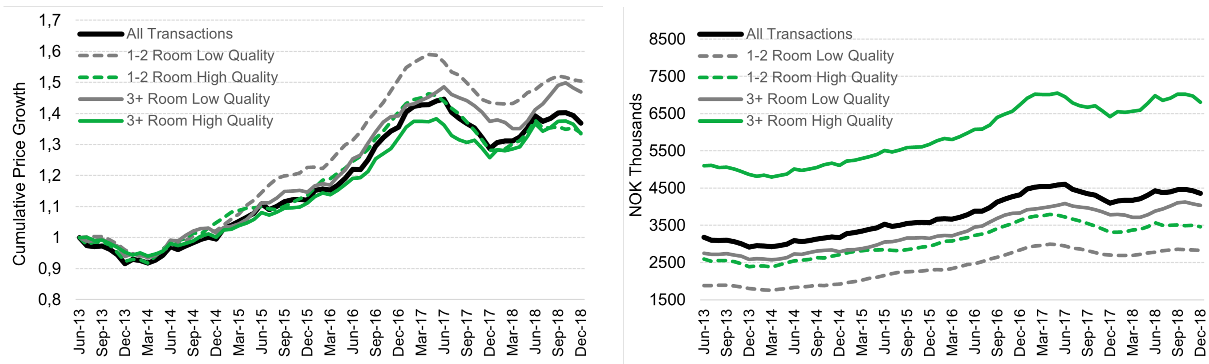
Notes: The figure plots the regression coefficients $\hat{\eta}_j$ from equation (1) with confidence intervals.

Figure 1 Life Outcomes when Growing up in Low vs High-Quality Neighborhood

2.4 The Effects of Mortgage Regulation on Prices and Choice Sets

Our last preliminary analysis shows the effect of the loan-to-income (LTI) cap on house prices and individuals' choice sets. The event we study is the introduction of an LTI cap in Norway in 2017. To provide context, starting January 1, 2017, mortgages were limited to five times the borrower's income. Prior to 2017, Norway did not have any income-based measures to regulate mortgage borrowing. In contrast, a loan-to-value (LTV) cap was already in introduced in 2010. In our sample period, the LTV cap is 0.85. We refer to Aastveit et al. (2020) for additional details on LTV regulation in Norway.

Figure 2 shows that the LTI regulation had a negative impact on the growth of house prices. The left plot shows the relative price growth. The right plot shows the 6-month rolling mean of average monthly transaction prices. Both plots include all transactions as well as four housing categories. These categories are based on two size groups (1 and 2-room apartments and 3 rooms or larger) and two district qualities (high and low). The period from January 2017 until June 2017 was when the LTI limit of 5 was gradually introduced.



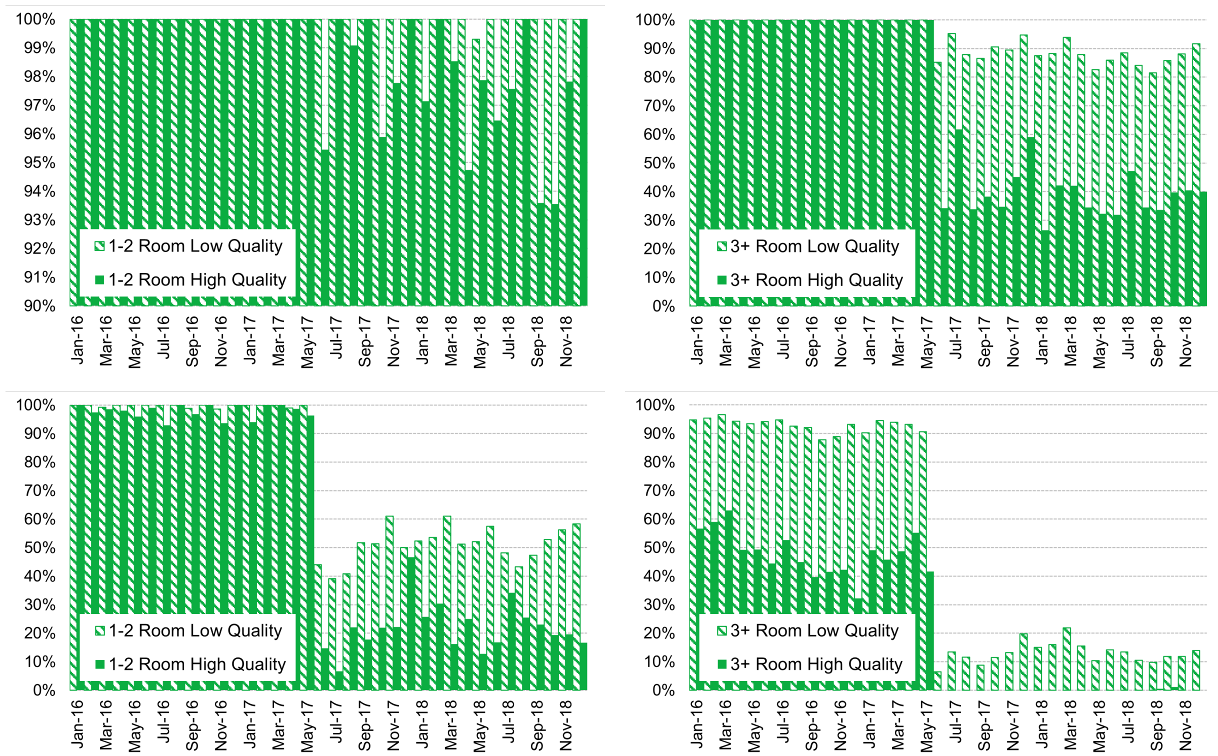
Notes: These figures plot the cumulative price growth (left panel) and the price level (right panel) across time of housing products of different size and quality.

Figure 2 Effect of LTI Regulation on House Prices over Time

We next analyze how the LTI regulation affected individuals' choice sets. We report the relative choice sets available to two sets of households. The first is the median household and the second is the top 10% households in terms of income distribution. We define the relative choice as the fraction of housing transactions within a month that is affordable given the resources of the household and the prevailing mortgage regulations. We focus on the period from January 2016 to December 2018. We use the same four housing categories as in Figure 2. Although the LTI cap was introduced in January 2017, it applied to everyone only from June 2017 onwards. The reason is that mortgage certificates (that is, the right to borrow a certain amount) are valid for six months at a time. As a result, we assume that LTI was binding from June 2017. Before then, the household only needs equity equal to 15% of the purchase price. We use the equity and income of the median and top 10% households at the end of the year to calculate these choice sets. Figure 3 presents the results.

3 Model

The previous evidence suggests that households vary in their willingness to pay for at least some features of neighborhoods. This raises several questions. First, is the issue of how willingness to pay relates to underlying household preferences. A related question is how variation in the willingness to pay for house and neighborhood characteristics affects both mobility patterns and house prices in the long-term. This is important both because the neighborhood during childhood is a strong predictor of future sociodemographic outcomes and housing wealth is the main source of wealth for the median household. Finally, it sheds light on a new dimension of mortgage regulation. In this section, we develop an equilibrium



Notes: These figures plot the share of housing products available in households' choice sets across time, before and after the implementation of the LTI limit in June 2017. The top two figures refer to households in the 10% of the income distribution, while the bottom two figures refer to the median households in terms of income. The two left figures refer to small properties (1-2 rooms), while the two right figures refer to large properties (3 or more rooms). Within each figure the solid vertical lines refer to high quality properties and the dashed vertical lines refer to low quality properties.

Figure 3 Effect of LTI Regulation on Housing Choice Sets

model of housing demand with mortgage affordability constraints. As we now explain, the model clarifies the relationship between the distribution of preferences and housing demand under different types of mortgage regulation.

We consider $i = 1, ..., \mathcal{N}$ potential investors in the real estate market in the capital of Norway, Oslo. These investors make a housing decision every period (year) $t = 1, ..., T$,[9] and can be either households or speculators. We define as households all the family units who own at most one property in Oslo and another property (such as a holiday home) outside Oslo. Households can be homeowners who live in their own residences without other property ownership in Oslo, or they can be renters who live in rented houses. We define instead as speculators all households and agents that own more than one property in Oslo at any point in time during the sample period and generate profits out of their real estate through rents and capital gains. We group households based on their type $\tau := \tau(Z_{i,t})$, where the variable

[9] We will use 2011-2018 for our estimation, and 2010 to determine the initial allocation of housing.

$Z_{i,t}$ includes households' total income ($Y_{i,t}$), net worth ($A_{i,t}$), and demographics ($G_{i,t}$) such as family size and age of the household head.[10]

3.1 Households

Every period t a household of type τ decides its ownership status $h \in \mathcal{H} = \{0, 1\}$ (i.e., to be a renter or an owner, respectively) and which type of housing product j across $J_{\tau,t}$ available options to consume. The housing product options are differentiated based on their location and the number of bedrooms in the property, as a proxy for size. In addition, the household can choose to move out of Oslo ($j = 0$). Thus, the set of housing product options is denoted by $\mathcal{J}_{\tau,t} = \{0, ..., J_{\tau,t}\}$. We denote a household's new type after the housing decision as $\bar{\tau} := \tau(\bar{Z}_{i,t})$, reflecting the potential change in wealth in the case of moving and property transactions. Household's i decision is defined as $d_{i,t} = \{j, h\} \in \mathcal{D}_{\tau,t} = \mathcal{J}_{\tau,t} \times \mathcal{H}$, and moving occurs when a household changes her current housing option, i.e., $d_{i,t} \neq d_{i,t-1}$.

Households in our model can make any of the following choices every period. Renters can choose to stay in their current house, move to another rented property, or buy a house of type j and become a homeowner. Similarly, homeowners can stay in their current property, sell their house to buy another one, or sell their property to become a renter in another house.

We use the information on income and wealth for each type τ in every period t, together with the average house prices $P_{j,t}$ across locations and property sizes, and the actual Loan To Income (LTI_t) and Loan To Value (LTV_t) constraints, to bound the housing choice set of households.[11] To be specific, if the household chooses to purchase a property, the affordable options must have a price satisfying both the LTI and LTV constraints:[12]

$$\mathcal{J}_{\tau,t} = \left\{ j \mid P_{j,t} \leq \min\left(\frac{A_{i,t}}{1 - LTV_t}, \frac{Y_{i,t} LTI_t}{LTV_t}\right) \right\}. \tag{2}$$

We assume that households can afford to rent any housing product.[13] Given a set of af-

---

[10] In order to have a finite number of household types, we discretize the variables that define types as follows. Income is divided in five groups (in NOK, where 1 USD in 2015 is about 8 NOK): <200k, 200k-400k, 400k-600k, 600k-800k, 800k-2,000k. Net worth is divided in five groups (in NOK): <100k, 100k-1,000k, 1,000k-2,500k, 2,500k-5,000k, 5,000k-15,000k. Family size is divided in two groups: singles and couples. Age of the household head is divided in three groups: 18-34, 35-49, and over 50.

[11] In Norway there was a suggested (not imposed) LTI limit of 3 until mid 2017, while afterwards a binding upper bound of 5 was introduced. LTV was .9 until 2012, and .85 onwards.

[12] While LTI constraints are straightforward to compute based on our detailed data on household income, LTV constraints are harder to measure. This is because, even though our data has precise information about household wealth, households might receive wealth from parents and/or relatives to purchase a property (Benetton et al., 2022).

[13] We motivate this assumption based on data evidence, as in our dataset we observe households of all types renting all property types.

fordable options, each household makes a sequence of housing decisions $\{d_{i,r}\}_{r=t}^{T}$ to maximize lifetime expected utility:

$$\max_{\{d_{i,r}\in\mathcal{D}_{\tau,r}\}} \mathbb{E}\left[\sum_{r=t}^{T}\beta^{r-t}u\left(P_{j,r},X_{j,r},\xi_{j,h,r},\chi_r,Z_{i,r},d_{i,r},d_{i,r-1},\varepsilon_{i,j,h,r}\right)|P_{j,t},X_{j,t},\xi_{j,h,t},\chi_t,Z_{i,t},d_{i,t},\varepsilon_{i,j,h,t}\right], \tag{3}$$

where $X_{j,t}$ includes observed house and neighborhood characteristics; χ_t are aggregate state variables, such as Loan To Income LTI_t and Loan to Value LTV_t limits; β is the discount factor; $\xi_{j,h,t}$ are unobserved house and neighborhood characteristics; $\varepsilon_{i,j,h,t}$ is the latent demand of household i for housing product j with the ownership status h, distributed as Type 1 Extreme Value. Assuming the household's problem follows a Markovian structure and has an infinite horizon, we can express the present value of lifetime expected utility as the sum of current utilities and the present discounted value of future utilities:

$$\begin{aligned} V(Z_{i,t},\chi_t,\varepsilon_{i,t}) = \max_{d_{i,t}\in\mathcal{D}_{\tau,t}} \{ & u\left(P_{j,t},X_{j,t},\xi_{j,h,t},\chi_t,Z_{i,t},d_{i,t}\right) - \mathbb{1}_{\{d_{i,t}\neq d_{i,t-1}\}}F(Z_{i,t},d_{i,t},d_{i,t-1}) + \varepsilon_{i,t} \\ & +\beta\mathbb{E}\left[V(Z_{i,t+1},\chi_{t+1},\varepsilon_{i,t+1})|\chi_t,Z_{i,t},\varepsilon_{i,t},d_{i,t}\right]\}, \end{aligned} \tag{4}$$

where $F(Z_{i,t},d_{i,t},d_{i,t-1})$ represents moving costs that the household incurs if it chooses to move. Following Rust (1987), we assume additive separability between per-period utility, moving costs, and the unobserved state variable, as well as conditional independence between the Markovian transition processes of the observed and unobserved state variables.

3.2 Speculators

We assume that all properties that are not owned by households are owned by speculators. Speculators hold portfolios of properties, and can buy or sell any property each period as they do not have affordability constraints. We also assume that speculators do not face transaction costs when rebalancing their portfolios. Variables and parameters referring to speculators will be indexed by s. There are two key differences between the modeling of households' moving decisions and the portfolio rebalancing decisions of speculators. First, while households' moving decision is just a discrete choice among mutually exclusive alternatives, speculators solve an optimal portfolio allocation problem across multiple assets simultaneously. Second, for households it is appropriate to explicitly incorporate a dynamic dimension in their model similarly to Bayer et al. (2016), as they trade-off the sunk cost of moving versus the evolution of house prices and of their valuations for housing products in the coming periods. However, for speculators a similar dynamic model with non-mutually exclusive alternatives across their portfolio would be challenging to solve. For this reason, we model speculators' optimal portfolio choice following the literature on characteristics-based asset demand (Koijen and Yogo, 2019).

We assume there are $i = 1, ..., \mathcal{N}^s$ speculators, and every period t they allocate wealth $A^s_{i,t}$ across properties in their investment universe $\mathcal{J}_t = \{0, ..., J_t\}$ and an outside asset. Let $w^s_{i,j,t}$ measure the portfolio weight for speculator i of product j at time t, expressed as the number of properties of that product type that it holds. We let the speculator choose the portfolio weights every period that maximizes the log utility over terminal wealth at date T, solving the following problem:

$$\max_{w^s_{i,j,t}} \mathbb{E}_{i,t}[\log(A^s_{i,t})] \quad (5)$$

subject to an inter-temporal budget constraint, where $\mathbb{E}_{i,t}$ is speculator's expectation at time t.

3.3 Econometric Model

We define $t_{0,i}$ as the first period in which we observe a household in our sample, and T_i as the total number of periods during which we observe household i. Let $v^\tau_{j,h,t} = \overline{V}(Z_{i,t}, \chi_t, d_{i,t} = \{j, h\})$ denote the expected choice-specific value function for a household with characteristics $Z_{i,t}$ and decision $d_{i,t} = \{j, h\}$. To simplify the notation, we use τ to denote the household type $\tau(Z_{i,t})$. Let $u^\tau_{j,h,t}$ instead denote the deterministic component of flow utility for households of type τ. If a household moves, its new type will be $\bar{\tau} := \tau(\bar{Z}_{i,t})$, reflecting the reduction in wealth due to moving costs. As standard in discrete choice models, we require some normalization to be able to identify the vector of lifetime utilities $v^\tau_{j,h,t}$. We, therefore, estimate a normalized lifetime utility $\tilde{v}^\tau_{j,h,t} = v^\tau_{j,h,t} - m^\tau_t$, where m^τ_t is a normalizing constant that reflects the average lifetime utility of household type τ in time t. We will discuss how to estimate m^τ_t in detail in the next section.

3.3.1 Households

We start by characterizing the households' decisions. Household i who is considering moving and becoming of type $\bar{\tau}$, chooses option j if $\tilde{v}^{\bar{\tau}}_{j,h,t} + \varepsilon_{i,j,h,t} > \tilde{v}^{\bar{\tau}}_{k,h,t} + \varepsilon_{i,k,h,t}$ $\forall k \neq j$. Conditional upon moving to an inside option (i.e., for $j \neq 0$), the probability of a household of type $\bar{\tau}$ choosing housing product j with ownership h in period t is:

$$\Pr^{\bar{\tau}}_{j,h,t} = \frac{\exp(\tilde{v}^{\bar{\tau}}_{j,h,t})}{\sum_{l=0}^{1} \sum_{k=1}^{J_{\bar{\tau}t}} \exp(\tilde{v}^{\bar{\tau}}_{k,l,t})}. \tag{6}$$

Let $t_{1,i}$ denote the time period in which household i decides where to move (conditional on moving to an inside option). A household's likelihood contribution for this decision is denoted by $L_i^{\text{prod}}(\tilde{v})$, where $\tilde{v}$ is the vector of all values of $\tilde{v}^{\bar{\tau}}_{j,h,t}$ and is given by:

$$L_i^{\text{prod}}(\tilde{v}) = \prod_{h=0}^{1} \prod_{j=1}^{J_{\bar{\tau}t}} \left(\Pr^{\bar{\tau}}_{j,h,t_{1,i}} \right)^{\mathbb{1}[d_{i,t_{1,i}} = \{j,h\}]}. \tag{7}$$

Similarly, the probability that a household chooses the outside option in time period t, conditional on moving, is given by:

$$\Pr^{\bar{\tau}}_{0,t} = \frac{\exp(\tilde{v}^{\bar{\tau}}_{0,t})}{\exp(\tilde{v}^{\bar{\tau}}_{0,t}) + \sum_{l=0}^{1} \sum_{k=0}^{J_{\bar{\tau}t}} \exp(\tilde{v}^{\bar{\tau}}_{k,l,t})}, \tag{8}$$

using which we can form the likelihood that a household chooses the outside option conditional on moving. Let $t_{2,i}$ denote the time period in which household i is considering the outside option (conditional on moving). The likelihood, which is denoted by L_i^{out}, is given by:

$$L_i^{\text{out}}(\tilde{v}) = {\Pr^{\bar{\tau}}_{0,t_{2,i}}}^{\mathbb{1}[j=0]} (1 - \Pr^{\bar{\tau}}_{0,t_{2,i}})^{\mathbb{1}[j \in \{1,\ldots,J_{\bar{\tau}t}\}]}. \tag{9}$$

In any given period, a household will not move from its current housing product if the indirect utility of staying exceeds the utility value of the best moving alternative. Recalling that if a household of type τ moves we denote their new type as $\bar{\tau}$, a household who is currently in housing type j with property ownership h will choose to stay if:

$$v^{\tau}_{j,h,t} + \varepsilon_{i,j,h,t} > \max_{\{l,k\} \in \mathcal{D}_{\tau,t}} [v^{\bar{\tau}}_{k,l,t} + \varepsilon_{i,k,l,t}] - \text{PMC}^{\bar{\tau}}_{i,t}, \tag{10}$$

where $\text{PMC}^{\bar{\tau}}_{i,t} = \bar{Z}'_{i,t}\gamma_{\text{pmc}}$ represents the psychological cost of moving for a household of type $\bar{\tau}$. These are any costs that households incur on top of the monetary cost of moving, which we instead introduce below. We allow these costs to depend on the household demographics that define their type. Employing the definition of the normalized choice-specific value functions, $\tilde{v}^{\tau}_{j,h,t}$, where $\tilde{v}^{\tau}_{j,h,t} = v^{\tau}_{j,h,t} - m^{\tau}_t$, and substituting in the above equation gives:

$$\tilde{v}^{\tau}_{j,h,t} + \varepsilon_{i,j,h,t} > \max_{\{l,k\} \in \mathcal{D}_{\tau,t}} [\tilde{v}^{\bar{\tau}}_{k,l,t} + \varepsilon_{i,k,l,t}] - (m^{\tau}_t - m^{\bar{\tau}}_t) - \text{PMC}^{\bar{\tau}}_{i,t}. \tag{11}$$

The term $(m^{\tau}_t - m^{\bar{\tau}}_t)$ captures the decrease in household lifetime utility caused by the reduction in wealth due to financial moving costs, which change the household type from τ to $\bar{\tau}$. Since $(m_t^{\tau} - m_t^{\bar{\tau}})$ is unobserved, we parametrize it as a function of financial moving costs, depending on household characteristics $\bar{Z}_{i,t}$, homeownership status in the previous and the current period. Formally, we define it as:

$$m_t^{\tau} - m_t^{\bar{\tau}} = \text{FMC}_{i,t}^{\bar{\tau}} \gamma_{i,\text{fmc}}^{\bar{\tau}}. \tag{12}$$

Financial moving costs can be of four types. If a renter is buying a house at price $P_{d_{i,t}}^{\text{buy}}$, where $d_{i,t} = \{j, h\}$ defines the housing product j that is purchased, then $\text{FMC}_{i,t}^{\bar{\tau}} = 0.03 \times P_{d_{i,t}}^{\text{buy}}$. If instead a homeowner is selling its house at price $P_{d_{i,t-1}}^{\text{sell}}$ to buy a new one at price $P_{d_{i,t}}^{\text{buy}}$, then $\text{FMC}_{i,t}^{\bar{\tau}} = 0.03 \times \left(P_{d_{i,t}}^{\text{buy}} + P_{d_{i,t-1}}^{\text{sell}}\right)$. If a homeowner is selling its house at price $P_{d_{i,t-1}}^{\text{sell}}$ to become a renter, then $\text{FMC}_{i,t}^{\bar{\tau}} = 0.03 \times P_{d_{i,t-1}}^{\text{sell}}$. This captures the idea that selling a property also implies facing financial costs proportionally similar to the purchasing process. Last, if instead, the household is a renter that moves to another rented property, then we set the financial moving cost to zero, $\text{FMC}_{i,t}^{\bar{\tau}} = 0$. Hence, $\text{FMC}_{i,t}^{\bar{\tau}}$ has the superscript $\bar{\tau}$ to capture how the financial cost varies depending on the household type. We let $\gamma_{i,\text{fmc}}^{\bar{\tau}} = \bar{Z}_{i,t}' \gamma_{\text{fmc}}$ to allow a marginal change in wealth to have a different impact on household utility depending on household characteristics, and $\text{PMC}_{i,t}^{\bar{\tau}} = \bar{Z}_{i,t}' \gamma_{\text{pmc}}$. The probability that a household stays in its current property of type j with ownership status h in a given period t becomes:

$$\Pr_{\text{stay},i,t}^{\tau,\bar{\tau}} = \frac{\exp(\tilde{v}_{j,h,t}^{\tau})}{\exp(\tilde{v}_{j,h,t}^{\tau}) + \sum_{l=0}^{1} \sum_{k=0}^{J_{\bar{\tau}t}} \exp(\tilde{v}_{k,l,t}^{\bar{\tau}} - \text{FMC}_{i,t}^{\bar{\tau}} \gamma_{i,\text{fmc}}^{\bar{\tau}} - \bar{Z}_{i,t}' \gamma_{\text{pmc}})}. \tag{13}$$

The likelihood contribution of each household's sequence of move/stay decisions is denoted $L_i^{\text{stay}}(\tilde{v}, \gamma_{\text{fmc}}, \gamma_{\text{pmc}})$ and is given by:

$$L_i^{\text{stay}}(\tilde{v}, \gamma_{\text{fmc}}, \gamma_{\text{pmc}}) = \prod_{t=t_{0,i}}^{t_{0,i}+T_i} (P_{\text{stay},i,t}^{\tau,\bar{\tau}})^{\mathbb{1}[d_{i,t}=d_{i,t-1}]} (1 - P_{\text{stay},i,t}^{\tau,\bar{\tau}})^{\mathbb{1}[d_{i,t} \neq d_{i,t-1}]}. \tag{14}$$

3.3.2 Speculators

We now characterize speculators' decisions. While regular households can stay or move to another property, either purchased or rented, speculators do not move, just buy and sell. Speculators also do not change their type over time, as they are always assumed to be unconstrained in their wealth and income levels, and are always owners of multiple properties. We assume speculators hold a portfolio of properties and are deep-pocketed investors for which transaction costs are negligible. Hence, differently from regular households, we will assume that they do not consider the financial or psychological costs of exchanging properties.

Moreover, we assume that the stock of newly built properties that have not yet been sold to households is also part of speculators' portfolios.

We let the aggregate stock of properties held by all speculators be defined as $\mathcal{S}_t^s$. Conditional on the choice of this aggregate stock, we model how speculators allocate their resources across property types as follows. Following Koijen and Yogo (2019), the optimal solution to equation (5) can be expressed as:

$$\frac{w_{i,j,t}^s}{w_{i,0,t}^s} = \exp\{\alpha^s P_{j,t} + \beta^s X_{j,t} + \xi_{j,t}^s\}\epsilon_{i,j,t}^s, \tag{15}$$

where $w_{i,0,t}^s$ represents the weight of the outside option, that is the number of properties that speculators decided not to purchase in that period. The determinants of these portfolio weights are product price $P_{j,t}$, other observed ($X_{j,t}$) and unobserved ($\xi_{j,t}^s$) product attributes, and a latent demand $\epsilon_{i,j,t}^s$ assumed to be Type 1 Extreme Value distributed. As we assume that all speculators have the same sensitivity to prices α^s and product characteristics β^s, we can estimate the parameters of equation (15) with a linear model aggregated to represent all speculators' portfolio weights $w_{j,t}^s$, as follows:

$$\log \frac{w_{j,t}^s}{w_{0,t}^s} = \alpha^s P_{j,t} + \beta^s X_{j,t} + \xi_{j,t}^s. \tag{16}$$

The stock of housing for product j held by speculators at time t will therefore be given by $\mathcal{S}_{j,t}^s = \mathcal{S}_t^s \times w_{j,t}^s$.

3.4 Market Clearing

We now define the market clearing condition that will be used in our counterfactuals to determine the new equilibrium prices of properties. Our model delivers predictions on households' housing demand and supply, and of the housing portfolio held by speculators.

For every product j at time t, the market clears when the housing supply of that product equals demand, that is:

$$\underbrace{\#\text{New Houses} + \#\text{Houses net supplied by speculators} + \#\text{Houses sold by owners}}_{\text{Supply}} =$$
$$\underbrace{\#\text{Houses bought by renters} + \#\text{Houses bought by switchers} + \#\text{Houses bought by entries},}_{\text{Demand}}$$

where in the demand part switchers are homeowners who change residency and entries are new households that enter the sample. Based on households' and speculators' decision probabilities we can construct a market clearing condition that determines the equilibrium price $P_{j,t}$ of each housing product j at time t as:

$$\underbrace{\mathcal{S}_{j,t}^{New} + (\mathcal{S}_{j,t-1}^{s} - \mathcal{S}_{j,t}^{s}) + \sum_{\tau}\sum_{i\in\tau}\left(1 - \mathrm{Pr}_{i,t}^{\tau,\bar{\tau}}\left(\{j,1\} \mid \{j,1\}\right)\right)}_{\text{Supply}} \tag{17}$$
$$= \underbrace{\sum_{\tau}\sum_{k}\sum_{i\in\tau}\mathrm{Pr}_{i,t}^{\tau,\bar{\tau}}\left(\{j,1\} \mid \{k,0\}\right) + \sum_{\tau}\sum_{k}\sum_{i\in\tau}\mathrm{Pr}_{i,t}^{\tau,\bar{\tau}}\left(\{j,1\} \mid \{k,1\}\right) + \sum_{\bar{\tau}}\mathcal{N}_{t}^{\bar{\tau}}\mathrm{Pr}_{j,h=1,t}^{\bar{\tau}}}_{\text{Demand}},$$

where $\mathcal{S}_{j,t}^{New}$ is the number of newly constructed houses of type j in year t, and $\mathcal{S}_{j,t-1}^{s} - \mathcal{S}_{j,t}^{s}$ is the net supply of type j properties by speculators at time t. The probability $\mathrm{Pr}_{i,t}^{\tau,\bar{\tau}}\left(d_{i,t} \mid d_{i,t-1}\right)$ indicates the probability that household i with previous decision $d_{i,t-1}$ makes decision $d_{i,t}$ at time t, changing its household type from τ to $\bar{\tau}$ with this decision. The supply is determined by the number of new houses, the net supply of speculators, and the number of properties sold by homeowners, which corresponds to the last term in the supply (left hand side) part of equation (17). More precisely, we sum the probabilities that each individual homeowner belonging to type τ and owning product j does not remain in her property $1 - \mathrm{Pr}_{i,t}^{\tau,\bar{\tau}}\left(\{j,1\} \mid \{j,1\}\right)$; this gives us the number of type τ homeowners that sell product j at time t. Next, we sum across all types τ to obtain the aggregate supply of owned housing product j at time t.

Demand for each product j at time t is determined by the number of housing products j purchased by renters plus that purchased by homeowners and new entries, which corresponds to the first, the second, and the third term in the demand (right hand side) part of equation (17), respectively. For the first two terms, summing the probabilities that each household belonging to type τ and renting (or owning) product k chooses to purchase product j gives us the number of products j bought by renters (or homeowners) of type τ at time t who used to live in product k. We sum over all types of properties that a household used to live in, and overall household types, to obtain the total housing demand. For the last term, $\mathrm{Pr}_{j,h=1,t}^{\bar{\tau}}$ is the probability that a household chooses to own type j conditional on the decision of moving, defined in equation (6); $\mathcal{N}_{t}^{\bar{\tau}}$ is the number of households of type $\bar{\tau}$ that enter into Oslo at time t. Summing $\mathcal{N}_{t}^{\bar{\tau}}\mathrm{Pr}_{j,h=1,t}^{\bar{\tau}}$ over all household types gives us the number of product j bought by new entries.

More specifically, the decision probabilities can be obtained as:

$$\Pr_{i,t}^{\tau,\bar{\tau}}(d_{i,t}=\{j,h\}\mid d_{i,t-1})=\frac{\exp(\tilde{v}_{j,h,t}^{\bar{\tau}}-\mathrm{FMC}_{i,t}^{\bar{\tau}}\widehat{\gamma}_{i,\mathrm{fmc}}^{\bar{\tau}}-\bar{Z}_{i,t}'\widehat{\gamma}_{\mathrm{pmc}})}{\exp(\tilde{v}_{0,t}^{\bar{\tau}})+\sum_{l=0}^{1}\sum_{k=0}^{J_{\bar{\tau}t}}\exp(\tilde{v}_{k,l,t}^{\bar{\tau}}-\mathrm{FMC}_{i,t}^{\bar{\tau}}\widehat{\gamma}_{i,\mathrm{fmc}}^{\bar{\tau}}-\bar{Z}_{i,t}'\widehat{\gamma}_{\mathrm{pmc}})}. \quad (18)$$

Note that the financial moving costs $\mathrm{FMC}_{i,t}^{\bar{\tau}}$ is determined by $d_{i,t}$ and $d_{i,t-1}$, as it is a function of the property price to be purchased or sold. The price that will clear each j,t combination will determine the value of the housing asset for sellers, therefore to which type $\bar{\tau}$ they will transition to if they sell, as well as the value of the purchased housing asset for buyers, hence to which type $\bar{\tau}$ they will transition to if they buy. We do not require the rental market to clear, as renters can always decrease housing consumption by living in shared housing units. Moreover, renters are assumed to be price-takers. All rented houses are owned by speculators who set a fixed rent-to-price ratio for each housing product.

4 Estimation

We estimate the model in four steps. First, we estimate $\tilde{v}$ according to households' housing product choices. Second, we estimate $\gamma_{\mathrm{fmc}}, \gamma_{\mathrm{pmc}}$ from the decisions to stay or move taking $\tilde{v}$ as given. Third, we recover the determinants of households' flow utility. Last, we estimate the parameters of the speculators' portfolio problem.

4.1 First Stage: Value Functions

The closed-form solution for maximizing the likelihood of choosing a housing product, conditional on moving, and of choosing the outside option of moving outside Oslo, which is $\sum_{i=1}^{N}\Big(\ln(L_i^{\mathrm{product}}(\tilde{v}))+\ln(L_i^{\mathrm{out}}(\tilde{v}))\Big)$, is given by:

$$\widehat{\tilde{v}}_{j,h,t}^{\bar{\tau}}=\ln(\widehat{\Pr}_{j,h,t}^{\bar{\tau}})-\frac{1}{2J_{\bar{\tau}t}+1}\left(\ln(\widehat{\Pr}_{0,t}^{\bar{\tau}})+\sum_{l=0}^{1}\sum_{k=1}^{J_{\bar{\tau}}}\ln(\widehat{\Pr}_{k,l,t}^{\bar{\tau}})\right), \quad (19)$$

where $\widehat{\Pr}_{j,h,t}^{\bar{\tau}}$ is the empirical probability of type $\bar{\tau}$ household choosing housing product j with ownership status h in time t conditional on moving. Instead of using the observed probabilities of a given household type directly, we use the kernel smoothing method similar to Bayer, McMillan, Murphy, and Timmins (2016) to account for the product choice decisions of similar household types. The kernel assigns weights across household types depending on the similarity in household characteristics, which allows us to overcome some small sample issues caused by the relatively large number of household types.[14]

4.2 Second Stage: Moving Costs

Once $\widehat{\tilde{v}}$ is recovered, we find $\gamma_{\text{fmc}}, \gamma_{\text{pmc}}$ to maximize the decision of stay and move, taking $\tilde{v}$ as given, that is:

$$\max_{\gamma_{\text{fmc}},\gamma_{\text{pmc}}} \sum_{i}^{N} \ln(L_i^{\text{stay}}(\widehat{\tilde{v}}, \gamma_{\text{fmc}}, \gamma_{\text{pmc}})). \tag{22}$$

The estimated $\widehat{\gamma}_{\text{fmc}}, \widehat{\gamma}_{\text{pmc}}$ can be used to recover the true choice-specific value functions $v^{\tau}_{j,h,t} = \tilde{v}^{\tau}_{j,h,t} + m^{\tau}_{t}$. Notice that $m^{\tau}_{t} - m^{\bar{\tau}}_{t}$ captures how the decrease in wealth due to financial moving costs affects the lifetime utility of household type τ. If we normalize the average utility of households with no wealth to zero, the impact of wealth on households' utility is recovered by multiplying household wealth with the marginal utility of wealth. Thus, we set $m^{\tau}_{t} = A^{\tau}_{i,t}\gamma^{\tau}_{i,fmc}$, where i is the median household of type τ, to recover the wealth effect on household utility $v^{\tau}_{j,h,t}$ for all household types.

4.3 Third Stage: Indirect Utility

To recover the per-period utilities we need to define the transition process of our state variables. We model the transition of the choice-specific value functions $v^{\tau}_{j,h,t}$ and house prices $P_{j,t}$ as:

$$v^{\tau}_{j,h,t} = \psi^{\tau}_{0,j,h} + \sum_{l=1}^{2} \psi_{1,l} v^{\tau}_{j,h,t-l} + \sum_{l=1}^{2} \psi_{2,l} P_{j,t-l} + \psi^{\tau}_{3,j,h} t + \omega^{\tau}_{j,h,t}, \tag{23}$$

[14] More precisely, we calculate the empirical probability of a type $\bar{\tau}$ household choosing housing product j with ownership h in time t conditional on moving, taking into account all household types that made the same housing decision:

$$\widehat{\text{Pr}}^{\bar{\tau}}_{j,h,t} = \frac{\sum_{i=1}^{N} \mathbb{1}_{[d_{i,t}=\{j,h\}]} \cdot W^{\bar{\tau}}\left(\bar{Z}_{i,t}\right)}{\sum_{i=1}^{N} W^{\bar{\tau}}\left(\bar{Z}_{i,t}\right)}, \tag{20}$$

where $W^{\bar{\tau}}\left(\bar{Z}_{i,t}\right)$ is the weight assigned to household i with characteristics $\bar{Z}_{i,t}$. We assign higher weights to household types that have higher similarity to $\bar{\tau}$ in the household characteristic space. The weight is the product of L normal kernels N:

$$W^{\bar{\tau}}\left(\bar{Z}_{i,t}\right) = \prod_{l=1}^{L} \frac{1}{b^{\bar{\tau}}_{l}} N\left(\frac{\bar{Z}_{i,t}(l) - \bar{Z}^{\bar{\tau}}(l)}{b^{\bar{\tau}}_{l}}\right), \tag{21}$$

where L is the dimension of Z, $Z(l)$ is lth attribute of household characteristics, and $b^{\bar{\tau}}_{l}$ is the bandwidth of the lth attribute determined by cross validation.

$$P_{j,t} = \phi_{0,j} + \sum_{l=1}^{2} \phi_{1,l} P_{j,t-l} + \phi_{2,j} t + \rho_{j,t}. \tag{24}$$

Knowing $v^{\tau}_{j,h,t}$, γ_{pmc}, γ_{fmc}, and the transition probabilities allows us to calculate mean flow utilities for each type and product, $u^{\tau}_{j,h,t}$, according to:

$$u^{\tau}_{j,h,t} = v^{\tau}_{j,h,t} - \beta \mathbb{E} \left[\ln \left(e^{v^{\tau}_{j,h,t+1}} + \sum_{l=0}^{1} \sum_{k=0}^{J_{\bar{\tau}}} e^{v^{\bar{\tau}}_{k,l,t+1} - \text{FMC}^{\bar{\tau}}_{i,t+1} \widehat{\gamma}^{\bar{\tau}}_{i,\text{fmc}} - \bar{Z}'_{i,t+1} \widehat{\gamma}_{\text{pmc}}} \right) \mid s_{i,t}, d_{i,t} = \{j, h\} \right], \tag{25}$$

where, in practice, $s_{i,t}$ includes all the variables on the right-hand side of equations (23) and (24), and β is set to 0.95. For each type, τ, product, j, ownership status h, and time, t, we now have the necessary information to simulate the expectation on the right-hand side of equation (25). To do this, we draw a large number of $v_{j,t+1}$ and $P_{j,t+1}$ from their empirical distributions. Specifically, using r to index random draws, each of these variables is generated by drawing from the empirical distribution of errors obtained when estimating each of these processes, respectively $\omega^{\tau}_{j,h,t}$ and $\rho_{j.t}$, and using the observed values of the current states. The draws of house prices are used to determine housing wealth, which determines households' type in the next period τ_{t+1}. For each draw r, we can then calculate a per-period flow utility $u^{\tau}_{j,t}$ using equation (25). The simulated $u^{\tau}_{j,t}$ is then calculated as the average across the draws.

We adopt a similar strategy as Bayer et al. (2016) to estimate the determinants of flow utility of household type τ choosing housing product j with ownership status h at time t as follows:

$$u^{\tau}_{j,h,t} = \alpha^{\tau}_0 + \alpha^{\tau}_h + \alpha^{\tau}_t + X_{j,t} \alpha^{\tau}_x + \alpha^{\tau}_r R_{j,t} + \xi^{\tau}_{j,h,t}. \tag{26}$$

The unobserved house and neighborhood attributes $\xi^{\tau}_{j,h,t}$ are assumed to have a different impact on flow utility for different household types, which is regarded as the error in the regression model. The parameters $\alpha^{\tau}_0, \alpha^{\tau}_h, \alpha^{\tau}_t$ are household-type-specific constant, utility of homeownership, and year fixed effects. The parameter α^{τ}_x captures the impact of observed housing attributes on household utility. Besides, α^{τ}_r captures the disutility of user costs, which are assumed to be equivalent to rental prices $R_{j,t}$. To address the endogeneity of prices, we exploit the estimated $\widehat{\gamma}^{\tau}_{\text{fmc}}$ that reflects the marginal utility of wealth. Assuming that households have the same marginal utility of wealth as that of income, equation (26) can be rewritten as:

$$u^{\tau}_{j,h,t} + \widehat{\gamma}^{\tau}_{\text{fmc}} R_{j,t} = \alpha^{\tau}_0 + \alpha^{\tau}_h + \alpha^{\tau}_t + X_{j,t}\alpha^{\tau}_x + \xi^{\tau}_{j,h,t}. \tag{27}$$

The parameters $\alpha^{\tau}_0, \alpha^{\tau}_h, \alpha^{\tau}_t, \alpha^{\tau}_x$ can be estimated by linear regressions.

4.4 Fourth Stage: Speculators' Model

We estimate equation (16) substituting product and time fixed effects to $X_{j,t}$, and calibrating the price sensitivity parameter α^s based on the following procedure. We search within the parameter space bounded by households' largest and smallest price sensitivity, corresponding respectively to the lowest and highest income quintiles, based on the estimates of $\widehat{\gamma}_{\text{fmc}}$.[15] For every guess of α^s, we use our model estimates and the market clearing condition in equation (17) to recover model-predicted equilibrium prices. We then choose the α^s that minimizes the difference between observed and model-predicted equilibrium prices. We implement this calibration as we only use 360 observations (45 products $\times$ 8 years) to estimate the speculator's model, the inclusion of product fixed effects absorbs most of the variation in prices, preventing us from precisely estimating the price sensitivity parameter.

5 Results

5.1 Model Estimates

We use the likelihood function (14), based on households' decisions to move or stay, to estimate the financial and psychological moving cost parameters γ_{fmc} and γ_{pmc}. Table 4 reports these estimates. For the financial moving cost, the constant reflects the average marginal disutility households derived from 1,000 NOK ($\sim$ 80 USD) of financial moving costs. We allow financial moving costs to depend on income, using as income the median values in each of the five groups described in footnote 10. We find that financial moving costs have less impact on higher income households. We allow the psychological moving cost to depend on households' income, age group, and family size. We find that psychological moving costs decrease with households' income, and are particularly high if the household head is above 55 years of age, while couples have higher psychological moving costs than singles.

[15]Note that $\widehat{\gamma}_{fmc}$ measures the impact of financial moving costs on marginal utility, which is a function of house prices depending on the cost of buying and selling. We consider an average transaction cost during the sample period, i.e., 5% of transaction prices. To be comparable with households' price sensitivity, we set the speculators' price sensitivity such that each households' income group and speculators' demand elasticities with respect to prices are the same.

Table 4 FINANCIAL AND PSYCHOLOGICAL MOVING COST

	Estimate
Financial Moving Cost	
Constant	0.01***
	(0.00)
Income	-0.01***
	(0.00)
Psychological Moving Cost	
Constant	5.87***
	(0.01)
Income	-0.14***
	(0.01)
Age: 35 - 54	0.96***
	(0.01)
Age: 55 +	2.07***
	(0.01)
Couple	0.40***
	(0.01)

Notes: This table presents the estimated parameters of financial moving costs γ_{fmc} and psychological moving costs γ_{pmc}. *Income* is the median income (in millions of NOK) of each income group. *Age: 35 -54* is a dummy variable taking the value of one if the head of a household is between 35 and 54 years old, and zero otherwise. *Age: 55+* is a dummy variable taking the value of one if the head of a household is above 55 years old, and zero otherwise. *Couple* is a dummy variable if a household is a couple, and zero otherwise.

5.2 Decomposition of Flow Utilities

Table 5 shows the determinants of households' flow utility. We regress the flow utility $u^{\tau}_{j,h,t}$ on property and neighborhood characteristics, excluding user costs in line with equation (27). In column (1) our estimates show that households' utility is increasing in neighborhood quality and property size, and homeownership delivers higher utility than renting. In column (2) we interact neighborhood quality, property size, and homeownership with household income. This is the specification that we use for our counterfactuals, as it allows us to simulate a scenario where all households have the same willingness to pay for housing attributes as the highest income quintile. We find that households' utility for high-quality neighborhoods is increasing in income, that low-income households prefer larger properties, but this is reversed for high income ones, and that homeownership delivers higher utility than renting at an increasing rate with income.

To be able to interpret the magnitudes of the determinant of flow utilities, Table 6 reports households' willingness to pay for different housing attributes based on their income. We calculate these numbers by dividing the estimated coefficients from equation (27) by the marginal utility of wealth $\hat{\gamma}^{\tau}_{\text{fmc}}$.

Table 5 Determinants of Flow Utility

	Flow Utility $\hat{u}$	
	(1)	(2)
High Quality	0.08***	-0.07***
	(0.03)	(0.04)
Low Quality	-0.13***	-0.00
	(0.04)	(0.02)
Three Rooms	0.08***	0.36***
	(0.04)	(0.04)
Four Rooms and Above	0.06	0.69***
	(0.05)	(0.05)
Home-Ownership	0.11***	-0.06***
	(0.02)	(0.02)
High Quality × Income		0.22***
		(0.02)
Low Quality × Income		-0.18***
		(0.05)
Three Rooms × Income		-0.39***
		(0.02)
Four Rooms and Above × Income		-0.87***
		(0.05)
Home-Ownership × Income		0.24***
		(0.03)
Year FE	Yes	Yes
Household Type FE	Yes	Yes
Observations	81,780	81,780
Adjusted R-Squared	0.88	0.90

Notes: This table reports the determinants of estimated flow utility. *High Quality* and *Low Quality* are dummy variables, with the omitted category being *Middle Quality*. *Three Rooms* is a dummy variable equal to one if a property type has three bedrooms, and zero otherwise. *Four Rooms and Above* is a dummy variable equal to one if a property type has four or more bedrooms, and zero otherwise. *Income* is the median income (in millions of NOK) of each income group.

To guide the interpretation of the numbers in Table 6, we find that an average household with annual income between 600,000 and 800,000 NOK (∼ 48,000-64,000 USD) is willing to pay 14,700 NOK (∼ 1,176 USD) every year for living in a high-quality neighborhood relative to a middle quality one. More generally, we find that all households other than the lowest income group are willing to pay more for living in high-quality neighborhoods relative to middle-quality ones, and this willingness to pay is increasing in income. Similarly, we find that all households would need to be compensated for living in low-quality neighborhoods relative to middle-quality ones, but high-income households need to be compensated more.

We find that households in the three lowest income groups are willing to pay increasingly more for larger properties, those in the fourth income group are willing to pay the highest for middle-sized properties, and the top-income group's willingness to pay is decreasing in property size. We interpret this as a reflection of family composition, as families with young children, who need more living space, are more likely to have a lower level of income than senior households, who instead need a smaller living space. Last, our results suggest that the willingness to pay for homeownership is increasing across income groups.

Table 6 Willingness to Pay for Housing Attributes

	High Quality	Low Quality	**3 Rooms**	**$\geq$ 4 Rooms**	**Ownership**
Income					
<400k	-1.45	-7.16	32.31	56.95	1.61
400 - 600k	5.10	-14.50	26.99	41.62	9.52
600 - 800k	14.70	-25.28	19.18	19.12	21.12
800 - 1000k	32.83	-45.64	4.43	-23.37	43.04
>1000k	157.73	-185.82	-97.15	-315.96	194.00

Notes: This table reports the average willingness to pay for housing attributes of households with different income levels. The first column suggests how much more an average household is willing to pay (in thousands of NOK) for being in a High-Quality district relative to a Middle-Quality district every year. The second column suggests how much more an average household is willing to pay (in thousands of NOK) for living in a Low-Quality district relative to a Middle-Quality district every year. The third (fourth) column suggests how much an average household in the income group is willing to pay (in thousands of NOK) for living in an apartment with three rooms (more than four rooms) compared with living in a one- or two-room apartment. The last column shows the average willingness to pay for living in their own house compared to a rented house.

6 Counterfactuals

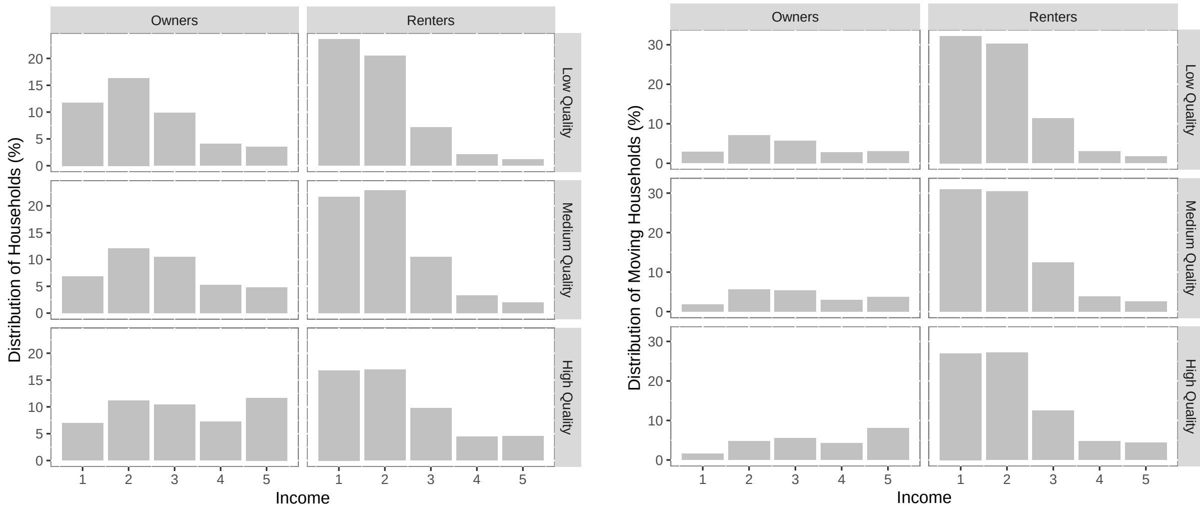
We use our structural model to simulate three counterfactual scenarios that help us quantify the distributional effects of leverage limits, and of other policies that would mitigate its regressive effects. More specifically, in 2017 in Norway the LTI limit was set to 5, while before there was no such restriction. We simulate a scenario for 2017 alone with a more stringent leverage limit, setting the LTI to 3. We investigate how our counterfactual affects prices in the Oslo housing market, and how it changes mobility patterns across the income distribution.

To determine the role of households' preferences, we compare the effect of the change in LTI to a scenario where all households have preferences closer to those of the highest income group. Simulating this change is not only a way to quantify the importance of preferences in residential choices, but also a way to proxy for housing policies, such as schooling vouchers or public housing assistance programs, that would encourage low-income households to become homeowners and/or move to high-quality neighborhoods. Last, we simulate a policy that curbs speculators' incentive to transact in housing markets, akin to higher taxation on second homes or rental income, by making speculators more price inelastic. As speculators mostly own properties in low-quality districts, we want to verify whether making them more inelastic would be an alternative way of incentivizing low-income households to move into high-quality neighborhoods.

Before showing the counterfactual results, we present in Figure 4 the baseline distribution of households in 2017 across housing quality levels and income groups, with breakdown between owners and renters. This is the benchmark to which we will eventually compare the counterfactuals. The left panel presents the "stocks", that is the distribution of all households, whereas the right panel shows the "flows", that is the distribution of moving households, computed as the model predicted probability of moving into a specific property type times the probability of moving. Within each of the two figures, the sum of the vertical bars in every row between owners and renters adds up to 100%. Before interpreting the figures, we should point out that the total number of households across income groups is not equal. The two lowest income groups represent roughly 50% of households, and the top income group has the smallest number of households. We specify this uneven distribution to mimic the skewed distribution of income, and to capture the increasingly different preferences that households in the top income groups have relative to the bottom ones.

There are three takeaways from the left figure. First, across all quality levels, households in the first two income groups are always more likely to be renters than owners, whereas the opposite is true for the two top income groups. This is consistent with low-income households being mostly young people who, for instance, have not yet accumulated enough wealth and income to afford owning a property. Second, the proportion of owners relative to renters increases with district quality. This is driven by speculators' real estate portfolios, which are mostly focussed on properties in low-quality districts placed in the rental market. Third, as expected, the largest proportion of high-income households is in high-quality districts, whereas the largest proportion of low-income households is in low-quality districts.

Similar patterns arise from the right figure, which focuses on movers, and presents two extra features. First, movers are more likely to end up as renters, partly due to renters' lower cost of moving. Second, low-income households are much more mobile than high-income ones, consistent with them being younger individuals who have not yet settled.



Notes: These figures plot the distribution of all households (left panel) and of moving households (right panel) across housing quality levels (low quality in the top, middle quality in the middle, high quality at the bottom), by households' income groups, and across owners and renters.

Figure 4 Stocks and Flows of Households across Quality and Income

6.1 More Stringent LTI Limit

Table 7 shows the results of our counterfactual LTI change relative to the baseline. For every row of the table, we present shares or probabilities across five income groups, both for the baseline level of LTI (labeled as *Base*) and for the percentage change between counterfactual and baseline (labeled as Δ). The top panel shows outcomes for all households in our data, while the bottom panel shows outcomes for households who moved in 2017. Starting from the top panel, our model shows that for the baseline level of LTI homeowners in the lowest (highest) income group have access to 52.9% (76.9%) of all 45 housing products. As the LTI becomes more stringent, low (high) income households' choice set shrinks, as they lose access to 14.3% (1.2%) of housing products. Both the lowest and highest income group are only marginally affected by a lower LTI in the share of properties they can access, but for different reasons. Rich households are unaffected because the LTI constraints are mostly not binding, while poor households are affected to a small extent because several houses were already unaffordable for them even with a higher LTI. This is supported by what happens to households in the middle of the income distribution, with those in the third group for example having 69.3% of products in their choice set in the baseline, but experiencing an 17.3% drop in the counterfactual.

Similarly, while in the baseline 14.2% (24.4%) of housing products are of high quality for households in the lowest (highest) income group, in the counterfactual their share of high-quality products drops by 12.5% (3.6%). Again, both the lowest and highest income groups are least affected by a lower LTI in the share of high-quality properties they can access, for the same argument described above. For households in the middle of the income distribution instead, we find that those in the fourth group have 19.5% of high-quality products in their choice set in the baseline, but experience a 25% drop in the counterfactual.

The other two rows of the top panel of Table 7 are meant to provide a quantification of mobility patterns of homeowners and renters across the income distribution. There are three takeaways that emerge. First, as expected, renters are more likely to move than homeowners. Second, low-income homeowners are more likely to move than high-income ones. Third, a stricter LTI limit has almost no effect on homeowners' probability of moving, marginally reduces by around 1% the probability of moving for renters, and has no effect on high-income renters' mobility. An important message that these results highlight is that within a year the fraction of households moving is very limited compared to those who do not.

The second panel of Table 7 focuses on households' (homeowners and renters) conditional probabilities, where we condition on the probability of moving. We find that the probability of renters becoming homeowners, which mostly represents first-time buyers, is 6.4% (55.9%) for the lowest (highest) income group, and a tighter LTI limit reduces it by 9.6% (0%). We also show that among movers the probability of moving from the lowest to the highest quality neighborhood is unaffected by the change in LTI for the lowest and highest income groups, for the reasons explained above on choice set heterogeneity, but it is reduced by at most 1.1% for the third and fourth income groups.

6.2 Restrictions on Speculators

In Table 8 we report changes in equilibrium prices between baseline and counterfactual LTI under two scenarios. In the first scenario we focus on the baseline speculators' price elasticity, while in the second we simulate the case of speculators becoming inelastic, that is reducing their price elasticity by 50% relative to the baseline. We think of this change in speculators' elasticity as equivalent to any interventions aimed at regulating their behavior in housing markets, such as higher transaction taxes for properties that are not the owner's main residence.

Under the baseline level of LTI, we find that making speculators inelastic significantly increases prices, but with substantial heterogeneity across the size and quality distribution of properties. While small and high-quality houses exhibit a price increase of respectively 4.5% and 5.6%, large properties experience a 1.4% drop in prices, and prices of low-quality houses only increase by 1.3%. We interpret these results as follows. Inelastic speculators are less responsive to price changes, so households who want to purchase properties held by

Table 7 Effect of Change in LTI Limit on Choice Sets and Moving Probabilities

	LTI	Income Groups 1	2	3	4	5
All households						
Owners' Share of Total	Base	52.9%	62.2%	69.3%	73.3%	76.9%
Products in Choice Set	Δ	-14.3%	-18.6%	-17.3%	-15.2%	-1.2%
Owners' Share of High Quality	Base	14.2%	16.9%	19.5%	21.3%	24.4%
Products in Choice Set	Δ	-12.5%	-18.4%	-25.0%	-20.8%	-3.6%
Owners' Moving Probability	Base	3.5%	4.2%	4.5%	4.5%	5.4%
	Δ	0.0%	0.0%	-0.3%	-0.4%	0.0%
Renters' Moving Probability	Base	12.4%	12.7%	11.9%	11.4%	11.7%
	Δ	-0.5%	-1.2%	-0.9%	-1.0%	0.0%
Movers						
From Renting to Owning	Base	6.4%	15.3%	27.8%	40.6%	55.9%
	Δ	-9.6%	-9.2%	-3.4%	-2.0%	0.0%
From Low to High Quality	Base	14.6%	15.1%	15.6%	18.4%	21.5%
	Δ	-0.4%	0.0%	-1.1%	-1.1%	0.0%

Notes: This table reports the mobility of households for each income quintile under the baseline scenario (Base) and the percentage changes under the counterfactual scenario of changing LTI limits (Δ). The first panel reports the share of available products for home purchasing (Owners' Share of Total Products in Choice Set), the share of available high-quality products for home purchasing (Owners' Share of High-Quality Products in Choice Sets), the probability of moving for homeowners (Owners' Moving Probability), and the probability of moving for renters (Renters' Moving Probability). The second panel considers only movers, where the probability of movers changing from renting to owning (From Renting to Owning) and the probability of movers changing from low-quality districts to high-quality districts (From Low to High Quality) are reported.

speculators will need to pay a greater price for markets to clear. High-income households will be the ones who can bear the largest price increase, as they are the most inelastic. As a result, we find that the largest price increase occurs for small and high-quality properties, as these are the houses that high-income households prefer the most.

Under the baseline level of speculators' elasticity, we find that simulating a more stringent LTI reduces prices by up to 2%. Our results show that the greatest price reduction occurs for large and high-quality properties. This happens because those are the most expensive properties, and as affordability constraints tighten those houses drop out from middle-income households' choice sets, as displayed in the top panel of Table 7. This also implies that middle-income households will shift demand towards less expensive houses, namely smaller and lower quality ones, reducing the impact of stricter LTI limits on their prices. With inelastic speculators these price effects of tighter LTI are preserved, with slightly larger magnitudes.

Table 8 Effect of Change in LTI Limit and Speculators' Elasticity on Prices

	Elastic Speculators (base)		**Inelastic Speculators**	
	LTI = 5	**Δ LTI**	**LTI = 5**	**Δ LTI**
House Size				
Small	3.30	-0.4%	3.45 (+4.5%)	-0.7%
Medium	4.34	-0.2%	4.44 (+2.3%)	-0.4%
Large	6.48	-2.0%	6.39 (-1.4%)	-2.2%
District Quality				
Low	3.08	-0.6%	3.12 (+1.3%)	-0.8%
Medium	3.47	-0.5%	3.56 (+2.6%)	-0.7%
High	4.32	-1.5%	4.56 (+5.6%)	-1.7%

Notes: This table reports the average house prices (in millions NOK) in 2017 under speculators' baseline price elasticity (*Elastic Speculators*), and under a 50% lower counterfactual price elasticity (*Inelastic Speculators*). The second and fourth columns report prices under the baseline LTI of 5, while the third and fifth columns report the percentage change in prices when LTI is reduced to 3. The fourth column also reports the percentage change in prices between scenarios with inelastic and elastic speculators. The first panel reports the average house prices for small (1-2 bedrooms), medium (3 bedrooms), and large (4 bedrooms and above), while the second panel for low, medium, and high-quality districts.

6.3 Role of Preferences

The last counterfactual we run aims at quantifying the importance of preferences in residential choices. To do so, we adjust all households' preferences to be closer to those of the top income group. We simulate a scenario where we reduce the distance between preferences of the bottom four income groups and the top one by 50%. Simulating this change allows us to quantify the importance of preferences in residential choices, and to proxy for housing policies that encourage low-income households to become homeowners and/or move to high-quality neighborhoods. In Figure 5 we compare the effects of this counterfactual, labelled as "Preferences", to the effects of the other two counterfactuals presented above, labelled as "Lower LTI" and "Inelastic Speculators". We show counterfactual results on the distributions of households' stocks and flows of properties, presenting percentage changes relative to the baseline in Figure 4.

In line with the results in Table 7, both figures show how tighter LTI limits reduce the extent of homeownership and increase rentals across all quality levels, both for movers and for all households. The effect of a lower LTI limit is stronger for low income households, highlighting its regressive effect.

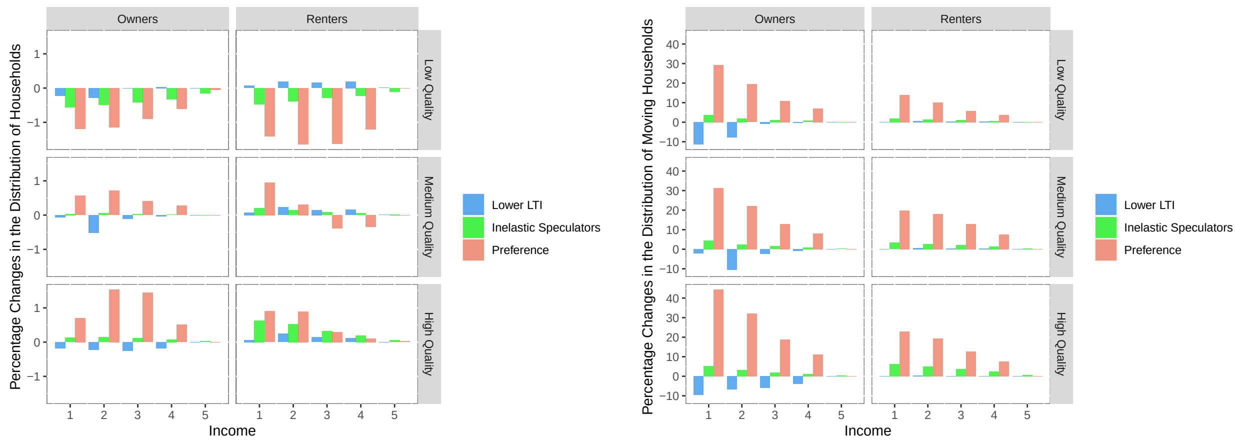
The counterfactual of inelastic speculators generates two overall effects. First, as the right panel of Figure 5 shows, there is an increase in mobility across quality levels for both owners and renters. This happens because inelastic speculators trade less in housing markets, resulting in households substituting for some of those missing real estate trades. Second, as the left panel of Figure 5 shows, there is a reduction in owners and renters at low quality districts, and an increase of mostly renters in medium- and high-quality neighborhoods. The reason for this is that speculators mostly hold properties in low-quality districts, and becoming inelastic makes them less likely to trade those products. As a consequence, households that were going to reside in low-quality districts are now more likely to rent in a high-quality one.

When we shift households' preferences closer to those of the top income group, we find two main effects. First, as the right panel of Figure 5 shows, the likelihood of moving increases across the income distribution, with stronger changes for low-income households, and across quality levels, with stronger changes for high-quality districts. Second, as the left panel of Figure 5 shows, there is a reallocation of households from low- to high-quality neighborhoods, both for owners and renters. These effects materialize because the change in preferences reduces moving costs and increases the benefit from being homeowners and living in high-quality neighborhoods.

The overall takeaway of this comparison is the following. On the one hand, more stringent LTI limits can have regressive effects and lead to an increase in segregation, as they reduce homeownership rates mostly for low-income households. On the other hand, influencing households' preferences or limiting speculators' trade in housing markets can have a countervailing effect, as both types on interventions encourage household mobility, with a greater effect on low income ones, and create an incentive to move to high-quality neighborhoods.

7 Conclusion

In this paper we develop and estimate a structural model of housing demand and supply to quantify the distributional effects of leverage regulation. We match demographic and financial characteristics of the population of households in the capital of Norway, Oslo, to the universe of housing transactions between 2010 and 2018. Our model features housing decisions of financially constrained households and financially unconstrained speculators, and derives equilibrium prices via a market clearing condition. Our detailed data on income and wealth allows us to measure precisely the affordability constraints of households, namely due



Notes: These figures plot the percentage changes in the distribution of all households (left panel) and of moving households (right panel) across housing quality levels (low quality in the top, middle quality in the middle, high quality at the bottom), by households' income groups, and across owners and renters. These changes are presented for three different counterfactuals. The blue bars represent the case of tighter LTI limits, the green bars represent the case of inelastic speculators, and the orange bars represent the case of households' preferences becoming closer to those of the top income group.

Figure 5 CHANGES IN STOCKS AND FLOWS OF HOUSEHOLDS ACROSS QUALITY AND INCOME

to Loan-To-Value and Loan-To-Income limits. We estimate households' moving costs and willingness to pay for neighborhood and property attributes across the income distribution.

We use the model estimates to conduct three counterfactual exercises. We start imposing a tighter leverage limit relative to baseline, which delivers the following results. First, we show that the housing choice sets of the lowest and highest income groups are only marginally affected by the change in LTI, as the former already have access to a limited number of properties in the baseline, and for the latter the constraint is still not binding. This means that households in the middle of the income distribution are the most affected, with a reduction of up to 18.6% in the size of the choice set, and at most a 25% drop in the share of high-quality houses available for purchase. Second, focussing on movers, we show that a tighter LTI limit reduces the probability of becoming homeowners by 9.6% for low-income households, and has no effect for the richest ones.

We then show that imposing restrictions on speculators' real estate trading, by making them more price inelastic, has two effects. First, it results in an increase in house prices mostly for properties that are demanded by high-income households, who have the highest willingness to pay. Second, it increases households' mobility, as they substitute for the reduction in speculators' trading, and it determines a reallocation of households from low- to high-quality districts, countervailing the regressive effect that a tighter LTI limit generates.

Last, we simulate a scenario where we shift households' preferences closer to those of the top income group. This also results in an increase in household mobility, in the rise of homeownership, and in a reallocation from low- to high-quality districts. Similarly to the previous counterfactual, this is another policy that regulators can implement to avoid the regressive effects of stricter LTI limits.

References

Aastveit, K. A., R. E. Juelsrud, and E. G. Wold (2020). Mortgage regulation and financial vulnerability at the household level. Norges Bank Working Paper.

Acharya, V. V., K. Bergant, M. Crosignani, T. Eisert, and F. McCann (2022). The anatomy of the transmission of macroprudential policies. *Journal of Finance 77*(5), 2533–2575.

Aikman, D., J. Bridges, A. Kashyap, and C. Siegert (2019). Would macroprudential regulation have prevented the last crisis? *Journal of Economic Perspectives 33*(1), 107–130.

Almagro, M. and T. Domínguez-Iino (2021). Location sorting and endogenous amenities: Evidence from amsterdam. Working Paper.

Bach, L., L. E. Calvet, and P. Sodini (2020). Rich pickings? risk, return, and skill in household wealth. *American Economic Review 110*(9), 2703–2747.

Bajari, P., P. Chan, D. Krueger, and D. Miller (2013). A dynamic model of housing demand: Estimation and policy implications. *International Economic Review 54*(2), 409–442.

Baker, S. R. (2018). Debt and the response to household income shocks: Validation and application of linked financial account data. *Journal of Political Economy 126*(4), 1504–1557.

Bayer, P., F. Ferreira, and R. McMillan (2007). A unified framework for measuring preferences for schools and neighborhoods. *Journal of Political Economy 115*(4), 588–638.

Bayer, P., R. McMillan, A. Murphy, and C. Timmins (2016). A dynamic model of demand for houses and neighborhoods. *Econometrica 84*(3), 893–942.

Benetton, M., M. Kudlyak, and J. Mondragon (2022). Dynastic home equity. Working Paper.

Calvet, L. E., J. Y. Campbell, and P. Sodini (2009). Fight or flight? portfolio rebalancing by individual investors. *Quarterly Journal of Economics 124*(1), 301–348.

Chetty, R. and N. Hendren (2018a). The impacts of neighborhoods on intergenerational mobility i: Childhood exposure effects. *Quarterly Journal of Economics 133*(3), 1107–1162.

Chetty, R. and N. Hendren (2018b). The impacts of neighborhoods on intergenerational mobility ii: County-level estimates. *Quarterly Journal of Economics 133*(3), 1163–1228.

Chetty, R., N. Hendren, and L. F. Katz (2016). The effects of exposure to better neighborhoods on children: New evidence from the moving to opportunity experiment. *American Economic Review 106*(4), 855–902.

DeFusco, A. A., S. Johnson, and J. Mondragon (2020). Regulating household leverage. *Review of Economic Studies 87*(2), 914–958.

Epple, D., L. Quintero, and H. Sieg (2020). A new approach to estimating equilibrium models for metropolitan housing markets. *Journal of Political Economy 128*(3), 948–983.

Fagereng, A., C. Gottlieb, and L. Guiso (2017). Asset market participation and portfolio choice over the life-cycle. *The Journal of Finance 72*(2), 705–750.

Favilukis, J., S. C. Ludvigson, and S. Van Nieuwerburgh (2017). The macroeconomic effects of housing wealth, housing finance, and limited risk-sharing in general equilibrium. *Journal of Political Economy 125*(1), 140–223.

Geanakoplos, J. (2009). The leverage cycle. *NBER Macroeconomics Annual 2009 24*, 1–65.

Han, L., C. Lutz, B. Sand, and D. Stacey (2021). The effects of a targeted financial constraint on the housing market. *Review of Financial Studies 34*(8), 3742–3788.

Heckman, J. J. and R. Landersø(2022). Lessons for americans from denmark about inequality and social mobility. *Labour Economics 77*, 1–14.

Holmes, T. J. and H. Sieg (2015). Structural estimation in urban economics. *Handbook of Regional and Urban Economics 5A*, 69–114.

Kiyotaki, N., A. Michaelides, and K. Nikolov (2011). Winners and losers in housing markets. *Journal of Money, Credit and Banking 43*(2-3), 255–296.

Koijen, R. S. J. and M. Yogo (2019). A demand system approach to asset pricing. *Journal of Political Economy 127*(4), 1475–1515.

Ludwig, J., G. J. Duncan, L. A. Gennetian, L. F. Katz, R. C. Kessler, J. R. Kling, and L. Sanbonmatsu (2012). Neighborhood effects on the long-term well-being of low-income adults. *Science 337*(6101), 1505–1510.

Mian, A. and A. Sufi (2009). The consequences of mortgage credit expansion: Evidence from the u.s. mortgage default crisis. *Quarterly Journal of Economics 124*(4), 1449–1496.

Mian, A., A. Sufi, and E. Verner (2017). Household debt and business cycle worldwide. *Quarterly Journal of Economics 132*(4), 1755–1817.

Peng, Y. (2021). Mortgage credit and housing markets. Working Paper.

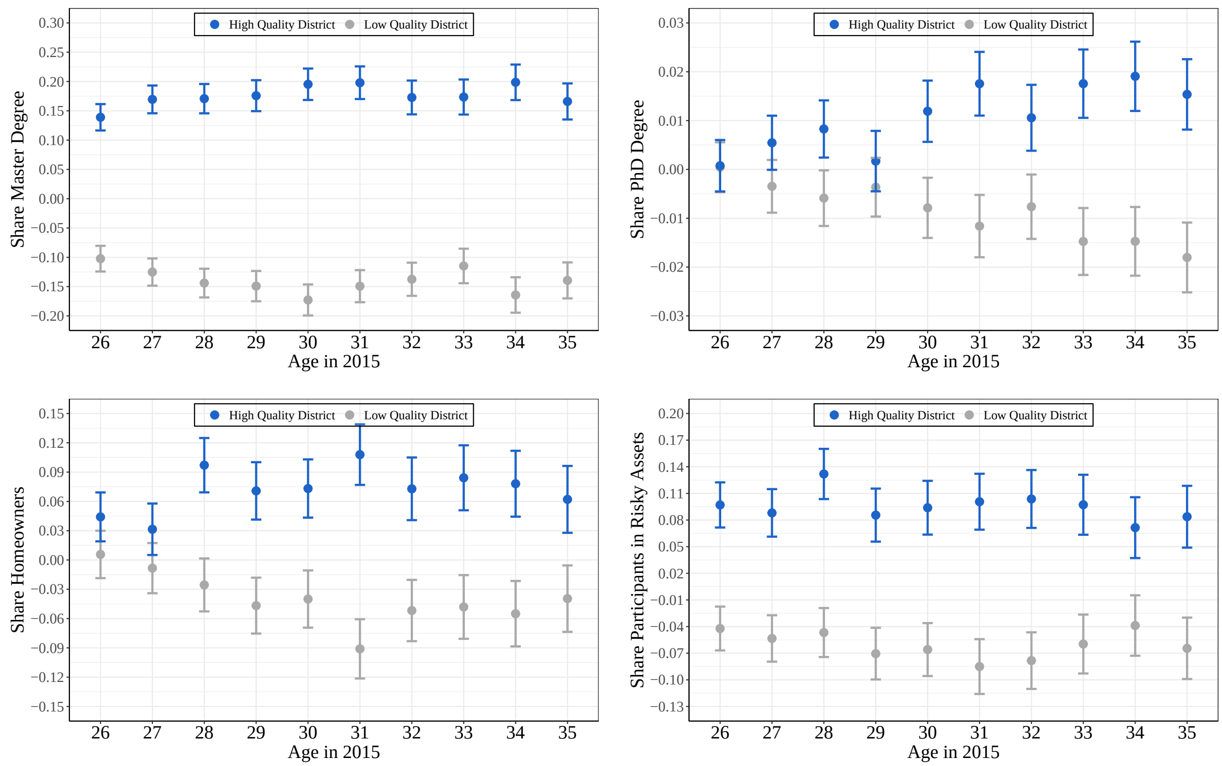
Peydrò, J.-L., F. Rodriguez-Tous, J. Tripathy, and A. Uluc (2020). Macroprudential policy, mortgage cycles and distributional effects: Evidence from the uk. Bank of England Staff Working Paper No. 866.

Rust, J. (1987). Optimal replacement of gmc bus engines: An empirical model of harold zurcher. *Econometrica 55*(5), 999–1033.

Sommer, K., P. Sullivan, and R. Verbrugge (2013). The equilibrium effect of fundamentals on house prices and rents. *Journal of Monetary Economics 60*(7), 854–870.

van Bekkum, S., M. Gabarro, R. M. Irani, and J.-L. Peydrò (2019). Take it to the limit? the effects of household leverage caps. Working Paper.

Appendix A Figures



Notes: The figure plots the regression coefficients $\hat{\eta}_j$ from equation (1) with confidence intervals.

Figure 6 LIFE OUTCOMES WHEN GROWING UP IN LOW VS HIGH-QUALITY NEIGHBORHOOD

Appendix B Data

This section provides additional details on how we construct our sample.

B.1 Variable Definitions

For each individual in our sample, we observe the birth date (variable name: "foedsels_aar_mnd") from the population database (In Norwegian: "Befolkning"). In the same database, we observe the number of child(ren) each individual has and the birth dates of the child(ren) (variable name: 'fodselsdato_barn_01-10"). In addition, we observe the ID (anonymized) of the spouse (variable name: 'ekt_fnr_aaaa"), or cohabitant (variable name: "sambo_snr_aaaa"). We use this information to classify an individual into a one-adult household (not registered ID for spouse or cohabitant) or more than one adult household. We refer to the two household types simply as singles and couples. The tax authority collects information on the complete wealth holdings of all households at the end of every year. For tax purposes, the household can allocate wealth in a way that gives the lowest wealth tax. Thus, there are no incentives for tax-motivated asset allocation within the household.

The financial information comes from the Norwegian Tax Registry (NTR) and reflects individuals' tax returns. We obtain this data from Statistics Norway, which merges it with the above demographic data. The NTR is responsible for collecting income and wealth taxes in Norway. By law, employers, banks, and public agencies must disclose personal information on income and wealth to the Tax Administration. The tax return includes all sources of income, as well as detailed information on wealth and debt. Individuals are accountable for the information provided in their tax returns, and the submission of inaccurate information is punishable by Norwegian law.

For each individual, we include total income (variable name: "wsaminnt"), which is the sum of gross salary income and pension plus net capital income and total government transfers, debt (variable name: "gjeld"), the value assessment of principal residence, which we label simple real estate (variable name: "prim_mark"), financial wealth (variable name: "bruttofin"), and total assets (variable name: "ber_brform"). We define homeownership as a variable that takes the value of one for all individuals with a positive value assessment of the principal residence. We also calculate adjusted total assets as total assets minus real estate. As we explain below, we net out the value assessment of the home because we include it in the estimation of the household's net worth, which is based on house product prices, $P_{j,t}$.

With this data, we need to aggregate individual data into household data. To be included in our sample, we follow standard practice in household finance and ensure that the households included in the analysis have a minimum cash balance (see, e.g., Calvet et al., 2009; Fagereng et al., 2017). In our case, we require the household to have at least 5,000 NOK in financial wealth. Because we study the mobility pattern of households, we also exclude from our sample the 20% with the lowest income at the end of the year and restrict the sample to households that are at least 18. Everyone that satisfies these criteria and does not have a registered spouse or cohabitant is classified as single.

The corresponding definition of couples is a bit more involved. We start by calculating the same financial data as for singles. We then aggregate total income, adjusted total assets, debt, and financial wealth to the household level. For age and homeownership we select the maximum in the household. We keep the ID (anonymized) of the oldest individual in the household and refer to this individual as the household head. District $d \in \mathcal{D}$ of residence and the number of children in the household are based on the household head. For the transactions, we include all transactions done by any of the two adults in the household. We impose the same financial and age requirements on couples as singles. The sum of single and couples satisfying our basic requirements comprise our sample of households. For each household, we define net worth, $A_{i,t}$, as:

$$A_{i,t} = P_{j,t} + \text{Financial Wealth}_{i,t} + \text{Other Real Estate}_{i,t} - \text{Debt}_{i,t}, \tag{28}$$

where $P_{j,t}$ is the price for house product $j \in \mathcal{J}$ in year t.

B.2 The Sample

We need the number of rooms and district for each unit and information about the homeowner or the renter to estimate the model. If these data are missing, we impute them. In our final sample, all homeowners and renters live in a particular housing product. In addition to households, we define a speculative sector that owns part of the housing stock. In what follows, we explain how we deal with missing values and define homeowners, renters, and speculators.

Regarding housing characteristics, the number of rooms $u \in \mathcal{U}$ is missing for 14,801 transactions (4.4%). For those observations, we use a multinomial logistic regression to predict the number of rooms based on the size of the apartment, the transaction price, and the district. The model predicts correctly in 74% of the cases. In comparison, randomly selecting the number of rooms $u \in \mathcal{U}$ would predict correctly in only 20% of the cases.

Regarding information about the homeowner or the renter, we observe it for households who transact in the market. For everyone else, we predict their housing product. Because we know the district where each household lives every year, it is sufficient to predict the number of rooms $u \in \mathcal{U}$ to identify their house product $j \in \mathcal{J}$. We begin by selecting all transactions in our sample period in which a household purchases a house with the number of rooms $u \in \mathcal{U}$. We use a rich set of characteristics for this sample to predict the number of rooms in their units. These characteristics include age, age^2, age^3, a dummy variable for being single, number of children, total income, and financial wealth. And the following dummies: D_{1i} takes the value of 1 if household i's total income is in the top 10 percent of the income distribution, D_{2i} takes the value of 1 if household i's financial wealth is in the top 20 percent of the financial wealth distribution, and D_{3i} takes the value of 1 if household i has more than four children. The idea with the indicator variables is to let income and wealth matter differently for very wealthy individuals relative to the rest of the sample. The model predicts correctly in approximately 54% of the cases. In comparison, randomly allocating the number of rooms $u \in \mathcal{U}$ would give a success rate of 33.3%.

Having identified a model that takes household characteristics as input and assigns the number of rooms $u \in \mathcal{U}$ as output, we create our sample. To do so, we start by selecting all households who live in any district $d \in \mathcal{D}$ at the end of 2010. For those households that also bought a housing product in the same year, we assign their actual product choice j to them. For the remaining households, regardless of homeownership status, we predict their housing product $j \in \mathcal{J}$ as we now explain.

Starting with 2010, the first year in our sample, we predict the number of rooms $u \in \mathcal{U}$ in the housing unit for all households for which we do not observe it. Given that we use a multinomial logistic regression model for this prediction, the output is a probability distribution for the number of rooms $u \in \mathcal{U}$. Since we have data on the number of housing units with $u \in \mathcal{U}$ number of rooms in each district $d \in \mathcal{D}$, we ensure that we never assign more housing units to a particular type $u \in \mathcal{U}$ than what is reported in official statistics. In addition, we ensure that the relative frequency distribution of housing units with $u \in \mathcal{U}$ number of rooms match official statistics. Given these restrictions, we assign the most likely choice, as predicted by our model, to each household.

An example illustrates what we do. Assume official statistics report that in district $d = 1$ there are 1,000 units with one room ($u = 1$) and 2,000 units with two rooms ($u = 2$). The total number of households in district $d = 1$ is 10,000. In our sample, assume that 8,000 households that live in district $d = 1$ satisfy the requirements to be included in the sample. Of those 8,000, we observe 500 households buying a housing unit with one room ($u = 1$) and 500 buying a housing unit with two rooms ($u = 2$). The number of households for which we need to predict the housing product is 8,000 - 1,000 = 7,000. We then assign the housing product $j = \{(1, 1)|d \in \mathcal{D}, u \in \mathcal{U}\}$ to: max{(Total j units/Total households in d) $\times$ Total households in d in our sample$-$Number of housing we observed buying product $j, 0\}$, which in this example is $\max\{1,000/10,000 \times 8,000 - 500, 0\}$.

In the next step, we first exclude the 800 households we just assigned a housing unit, then repeat the exercise for housing units with two rooms ($u = 2$). We continue until all the households in district $d = 1$ have a housing product. In all other years (i.e., the period from 2011 to 2018), we use the same method to assign the number of rooms $u \in \mathcal{U}$ in a housing unit for households that enter the sample without buying a housing unit or move to another district. Households entering the sample by purchasing a housing unit are given the housing product $j \in \mathcal{J}$ they choose.

To separate homeowners from renters, we calculate for each year the average book value of housing for those that transacted in the market during the year. Households that do not transact in the market but have a book value of housing above this estimate are classified as homeowners, while the rest of the households are renters. The threshold for homeownership is product specific.

In addition to the household sector, we include a speculative sector that transacts in the housing market to maximize risk-adjusted profits. We define speculators as all market participants that bought or sold a housing unit $j \in \mathcal{J}$ from 2011 to 2018 that do not satisfy our requirements to be classified as households. In addition, we classify households that buy or sell multiple units in a year as speculators.