Discussion Papers in Economics

No. 05/2021

Secondary Housing Supply

Andreas Mense
FAU Erlangen-Nürnberg

ISSN 1867-6707

Friedrich-Alexander-Universität Erlangen-Nürnberg
Institute for Economics
https://www.iwf.rw.fau.de/discussion-papers/

Secondary Housing Supply*

This version: August 30, 2021; first version: January 10, 2020.

Andreas Mense

Friedrich-Alexander-University Erlangen-Nuremberg & London School of Economics

---

Abstract

I estimate the impact of new housing supply on the local rent distribution, exploiting weather shocks during the construction phase as an instrument. New supply decreases rents at all quality levels. Building on a quantitative dynamic model of housing quality and tenure choice, I explain this pattern by secondary housing supply: New supply to the owner-occupier market triggers a cascade of moves in the rental market, freeing up units across the housing quality spectrum. This mechanism has implications for housing policy, the integration of the housing market, and the integration of new- and used-product markets in general. [97 words]

Keywords: secondary markets, market integration, housing markets, rent distribution, dynamic housing choice.

JEL classification: D15, D40, R21, R31.

---

*Friedrich-Alexander-University Erlangen-Nuremberg, School of Business, Economics, and Society, Findelgasse 7, 90402 Nuremberg, Germany, Tel: +49 911 5302-956, andreas.mense@fau.de. I thank Gabriel Ahlfeldt, Rainald Borck, Thiess Büttner, Christian Hilber, Markus Nagler, Kate Pennington, Johannes Rincke, Sebastian Siegloch, Niko Szumilo, Matthias Wrede, and participants at the 2020 North-American Virtual Meeting of the Urban Economics Association, the Virtual Annual Conference of the VfS 2020, the 2021 European Virtual Meeting of the Urban Economics Association, the Virtual Annual Congress of the European Economic Association 2021, the Potsdam Research Seminar in Economics, and seminar participants at the University of Erlangen-Nuremberg for valuable comments and suggestions. I am solely responsible for all remaining errors. This research is funded through grant ME 4963/1-1 by the German Research Foundation. An earlier version of the paper circulated under the name "The Impact of New Housing Supply on the Distribution of Rents".

1. Introduction

Prices on second-hand markets depend on the price of the new product. Two important second-hand markets are the housing market and the market for used cars. The markets for used cell phones and second-hand fashion are continuing to grow rapidly, partly due to increasing environmental awareness.[1] As market places like *Ebay* show, there are numerous other examples of durable goods for which second-hand markets exist.

To date, there is very little research on the question how shocks to new supply affect the distribution of second-hand prices. One potential barrier to the propagation of a shock to new supply is the fact that used products may be of considerably lower quality — and may thus be poor substitutes for new products. This paper argues that substitutability is not a necessary condition for market integration of the first- and second-hand markets. The reason is that adjustment costs prevent agents in the market from updating their product choices. For instance, a person might be driving her new car until a mileage of 100,000. When purchasing a new car at that point, the purchase creates a direct link between the new-car segment and the 100,000-mileage segment. This is despite the fact that the two types of cars may be very poor substitutes, in the sense that they are likely bought by very different types of consumers.

This paper provides evidence on this mechanism, using the housing market as an example. New housing supply only amounts to a very small share of the overall stock of housing in an economy. However, new supply triggers a cascade of moves in the market that frees up housing units. Such cascades are central to market integration and to the propagation of shocks to (local) housing markets. The mechanism applies equally to other second-hand markets.

[1] See Scarsella, A. and Stofega, W. (2020), *Worldwide Used Smartphone Forecast, 2020–2024*, IDC, and Khusainova, G. (2021), *The Secondhand Market Is Growing Rapidly, Can Challengers Like Vinokilo Thrive And Scale?*, `www.forbes.com`, Jan. 21, 2021.

I consider the impact of new, market-rate, housing supply on the local distribution of private-market rents in Germany.[2] Arguably, lower-quality rental housing units are poor substitutes for new housing. This implies that a direct effect — driven by a shift in demand from low-quality rental housing to newly built housing — seems at best unlikely. Units freed up by movers in the market may be much better substitutes, facilitating the diffusion of the shock.

I exploit weather shocks during the construction phase as an exogenous supply shifter. Unusual rainfall spells during the summer, as well as unusually deep frost in February, reduce significantly the number of end-of-year completions of single-family homes (November/December).[3] I document that the weather-induced delays are long-lasting, consistent with tight capacity constraints among housing developers during the most recent boom in Germany (starting in 2010), and with evidence for the U.S. (Coulson and Richard, 1996; Fergus, 1999). Arguably, the weather shocks affect rents only through the supply of new housing — most sectors of the German economy do not depend on summer rainfall, and the results are robust to excluding years with larger floods. Moreover, the shocks are not correlated with placebo outcomes or typical determinants of local housing demand.

Instrumental variable quantile regressions (IVQR) show that new housing supply at market rates shifts the location of the rent distribution to the left. There is no statistically significant difference between the impact on rents of high- versus low-quality units. To explain this pattern, I develop and estimate a dynamic discrete choice model of housing choices that characterizes *secondary housing supply* to the

[2] The German homeownership rate is low by international standards — 45.7% according to the 2011 census. However, the mechanism applies in an analogous way to housing markets with higher shares of owner-occupied housing, as long as some buyers of new housing are former renters. Moreover, the mechanism applies also to the propagation of supply shocks inside the owner-occupier market.

[3] Multi-family homes are also marginally affected. Since it usually take more than one year to build a multi-family home, weather shocks in a single year are arguably much less important for larger construction projects such as multi-family homes.

rental market, building on the filtering framework (Sweeney, 1974b,a). The term 'secondary housing supply' refers to units freed up by mover households. The model is able to reproduce the location shift of the rent distribution in response to a shock to new housing supply. The reason is that each renter moving into a newly built home triggers about 3.5 additional moves in the rental market until a new equilibrium is reached. Moreover, renters typically 'jump up the housing ladder' — rather than taking small steps — because they face moving costs. These channels lead to tight integration of all quality segments in the rental market, and of the owner-occupier and rental markets.

The paper makes three main contributions: First, the paper proposes a mechanism that determines market integration of second-hand markets, using secondary housing supply to the rental housing market as an example. This mechanism has deep implications for the integration of different market segments inside the rental housing market, as well as between the rental and owner-occupier markets. In this sense, the results complement recent findings regarding the housing choices of owner-occupiers and the relationships between different segments of the market for home sales (Landvoigt et al., 2015; Piazzesi et al., 2020). Crucially, moving costs restrain households from making gradual adjustment of housing quality, which loosens the relationship between household income and housing quality. Higher moving costs speed up the adjustment process because they reduce the number of moves in the cascade to the new equilibrium, while leading to more direct cross-connections between different housing quality segments. More generally, the mechanism has implications for the market structure of durable goods markets, and for firm behavior in these markets (Rampini, 2019; Chen et al., 2013; Dana and Fong, 2011; Johnson, 2011; Waldman, 2003).

Second, to the best of my knowledge, the paper is the first to provide clean, quasi-experimental evidence on the connection between new housing supply and the tails of the rent distribution, documenting that new housing supply effectively improves housing affordability of renters across the board. This finding has significant implications for housing policy in general, adding to a small, but growing empirical literature on filtering (Rosenthal, 2014, 2019; Mast, 2019) and on the impact of new housing supply on housing costs (Büchler et al., 2019; Molloy et al., 2020; Hilber and Mense, 2021; Nathanson, 2019; Pennington, 2021).[4]

Third, the paper provides estimates of renters' *within-market* moving costs. Mobility frictions at the micro-level have significant consequences for the capitalization of amenities into house prices (Bayer et al., 2016), the effectiveness and efficiency of housing subsidies, rent controls (Diamond et al., 2019), social housing, neighborhood revival programs, or local labor market policies that aim to improve the situation of low-income households and to reduce spatial inequality.[5] They can also explain reservations of households against "downgrading" their housing unit when becoming older—a potential source of housing supply in aging societies. The structural estimates suggest that a move within the local market may reduce utility as much as giving up about half of the household's net yearly income. Compared to this, existing estimates from dynamic structural models relate to long-distance moves and are substantially larger (Kennan and Walker, 2011; Buchinsky et al., 2014; Oswald, 2019). Mover households staying within the local housing market do not have to change jobs, nor do they lose access to their local social networks, and smaller distances between

[4]In recent years, housing costs have increased substantially in many places around the world — most dramatically in productive and attractive places such as San Francisco, New York, London, Tokyo, or Paris, where housing supply elasticities are typically low. The rising housing costs have triggered various policy responses. However, most–if not all–of these policies lead to considerable distortions or have minuscule quantitative effects (see Metcalf, 2018, for a recent survey). While it is well understood that lack of housing supply has large effects on house prices (Hilber and Vermeulen, 2016; Gyourko et al., 2013; Saiz, 2010; Quigley and Raphael, 2005; Glaeser et al., 2005), the connection between new housing supply at market rates and the rental housing market—in particular the lower segments—is less clear-cut.

[5]For instance, high moving costs can explain relatively low take-up rates and large differences between intention-to-treat and treatment-on-treated effects in the Moving to Opportunity experiment, see Chetty et al. (2016).

locations make it easier to collect information about the new neighborhood. In this sense, the estimates are consistent with recent evidence on the role of information frictions for long-distance moving costs (Schmutz and Sidibé, 2019).

The remainder of this paper is structured as follows: In Section, 2, I first describe the housing supply, weather, and rent data, and motivate the instrumental variable strategy. Then, I analyze the effects of new housing supply on the distribution of rents. The structural dynamic model developed and estimated in Section 3 is used to investigate the underlying mechanism — secondary housing supply, in conjunction with moving costs. In the final section, I draw conclusions and offer suggestions for policy and future research.

2. The Effect of New Housing Supply on the Distribution of Rents

2.1. Data

Housing completions are provided by the administrative Building Completions Statistic (see Appendix O-A for more details on the data source).[6] It provides information on all new housing units completed in Germany between 2010 and 2017, including the location (municipality) and the month of completion. Unfortunately, it is not possible to separate the supply of social housing from the supply of private-market housing in the empirical analysis. However, in recent years, only a small share of new housing supply in Germany was subsidized social housing.[7] In all other cases, developers are free to sell their units at any price. Moreover, as I show below, the instrument mainly captures shocks to the supply of single-family housing.[8]

[6] Source: RDC of the Federal Statistical Office and Statistical Offices of the Federal States, *Statistik der Baufertigstellungen*, survey years 2010-2017.

[7] Since 2007, the German Länder (federal states) are responsible for social housing, and a unified statistic does not exists. According to a parliamentary interpellation from March 2017, about 6% of new housing supply was subsidized in 2013 and 2014 (Deutscher Bundestag, 18/11403). Unfortunately, the Building Completions Statistic also does not provide information on subsidies.

[8] Single-family housing rarely qualifies for subsidies in the German institutional setting.

According to the German Socio-Economic Panel (SOEP), 49.1% of new housing supply in Germany is absorbed by renters transitioning to owner-occupier status, while only 19.3% are former owner-occupiers. The remaining 31.6% are built-to-let developments.[9] Moreover, 90.4% of all movers were renters, and 9.6% were owner-occupiers. Roughly half of the owner-occupiers moved into owner-occupied housing (5% of all moves). The overall share of renters transitioning into owner-occupied housing was about three times larger (14.8% of all moves). These numbers underscore the importance of renters' decisions for understanding spillovers between rents and prices more generally, and they suggest that the marginal buyer of (newly built) owner-occupied housing in Germany is a renter.

The instrumental variables were derived from data on rainfall and frost depth, provided by the German Weather Service as grid cell data (1×1 km^2) for the years 2010–2017.[10]

The rent data are posted rents, collected from three large online real estate market places (Immonet, Immowelt, Immobilienscout24) on a monthly basis between July 2011 and December 2018, covering around 80–90% of the rental housing market in Germany. The data contain information on the net rent, the unit size in square meters, the postcode of the unit, the month of its first appearance, and a list of housing characteristics. The outcome of interest is the log rent per square meter, net of utilities and heating costs. Appendix O-B provides further background information.

Posted rents are advantageous in the present setting for several reasons. First, the instrumentation strategy takes care of the problem that posted rents may differ from concluded rents, as long as there is no correlation between the measurement error

[9] The shares refer to mover households for which the year of construction equals the year of observation, between 2010 and 2017 (excluding subsidized housing). 56 such moves were observed. The Census 2011 reports very similar shares for housing built between 2009 and 2011, with 61% owner-occupied, and 39% built-to-let developments (including subsidized housing).

[10] Source: DWD Climate Data Center (2010-2017): REGNIE grids of daily precipitation; DWD Climate Data Center (2010-2017): Monthly grids of the maximum frost depth under uncovered soil at midday.

and the instrument. Since the instrument is a lagged, weather-based instrument, this seems highly unlikely. Second, surveyed rents may be less precise than posted rents because households may have difficulties to determine their *net rent*, as opposed to their total costs for shelter.[11] In Germany, households typically pay the gross rent including heating services (consisting of net rent, property services, utilities, and heating). The different components of the gross rent are posted separately in the rent offers, making the net rent information arguably much more reliable than comparable information from surveys. Similar arguments apply to the exact floor size, for which measurement is regulated by German bylaw.[12] Finally, posted rents are available on a fine geographical scale and with very precise information on housing characteristics and the state of the unit – unlike surveyed rents. In the German setting, the latter are not available at yearly frequency, let alone on a small geographic scale. Moreover, existing surveys include new contracts and older contracts subject to tenancy rent control, and sample sizes shrink dramatically when focusing on recent movers alone.

Table 1 contains summary statistics. The average monthly rent per square meter is 7.8 Euro (median: 6.8 Euro). The monthly rent refers to the rent posted on the day the offer appears online for the first time.

2.2. Weather Shocks as Instrument for New Housing Supply

2.2.1. Technical Mechanism

In order to identify shifts in new housing supply, I exploit fluctuations in housing completions at the end of the year, caused by unfavorable weather conditions during spring and summer. Previous studies have found that local weather conditions influence the number of housing completions, creating persistent supply shocks (see, e.g.

[11]For instance, the SOEP has changed several times the way respondents are asked about their housing costs, see SOEP Group (2019), admitting that some households may have misunderstood the question or may simply not know how much they pay. In particular, the SOEP does not ask respondents to report their net contract rent.

[12]Real estate agents have to apply DIN 283/1951 and the Floor Area Act [*Wohnflächenverordnung*].

Table 1: Descriptive statistics for the rents sample

A. Non-categorial and binary variables	Min	Mean	Q25	Median	Q75	Max
Monthly rent per sqm	1.6	7.8	5.4	6.8	9.0	85.2
Living area in sqm	15.0	71.4	52.8	67.0	85.0	300.0
Year of construction	1800	1970	1955	1974	1996	2018
Floor heating	0.00	0.08	0.00	0.00	0.00	1.00
Parquet flooring	0.00	0.03	0.00	0.00	0.00	1.00
Elevator	0.00	0.16	0.00	0.00	0.00	1.00
Fitted kitchen	0.00	0.32	0.00	0.00	1.00	1.00
Second bathroom	0.00	0.15	0.00	0.00	0.00	1.00
Garden	0.00	0.19	0.00	0.00	0.00	1.00
Balcony or terrace	0.00	0.60	0.00	1.00	1.00	1.00
New units in December, relative to						
1000 units in the stock (2011)	0.00	1.37	0.29	0.72	1.68	1324.32
the avg. # of rental units on the market	0.00	0.14	0.02	0.05	0.15	14.24
Summer rainfall spell (deviation)	-8.20	-0.15	-1.70	-0.30	1.30	12.60
Feb frost depth (deviation)	-17.00	-0.83	-5.60	-3.10	-0.50	46.90

B. Categorial variables (shares)	0	1	2	3	4	5	6	7	8
Dwelling type	0.586	0.112	0.131	0.009	0.033	0.002	0.006	0.010	0.110
Quality	0.017	0.147	0.831	0.005					

Notes: Dwelling type categories are 0: regular, 1: roof storey, 2: ground floor, 3: souterrain, 4: maisonette, 5: loft, 6: penthouse, 7: other, 8: NA. Quality categories are 0: luxurious, 1: above average, 2: average, 3: below average.

Fergus, 1999, for the U.S.). Poor weather conditions as a reason for an extension of building time are recognized by German building law (see §6 Abs. 2 Nr. 1 VOB/B).

As soon as the soil has thawed up, developers begin groundwork, usually erecting the building walls until mid-summer. In the summer, rainfall may lead to delays, for a number of reasons. First, many building materials, such as concrete and mortar, need to dry before roof and windows can be closed. Otherwise, moisture can lead to damages, and it encourages mold to form inside the building. If the summer is too wet, this process takes longer, so that construction work cannot be completed before the winter.[13] Second, on sunny summer days, the "effective daytime" is longer, so that construction work can take place from the early morning hours until the late

[13]There is no official statistic on building starts in Germany, and I am not aware of a data set that documents the timing of the construction process. However, various newspaper and magazine articles suggest that most housing starts occur in late winter or early spring, and that walls are erected within approximately four to five months, e.g. https://www.immonet.de/service/zeitplanung-hausbau, https://www.hausausstellung.de, or https://www.n-tv.de/ratgeber. About 25-30% (58-65%) of newly built single family homes are completed within 12 (18) months after having obtained the building permit. The shares are substantially lower for multi-family homes (7 and 28%) (Schwarz, 2018).

evening without electric light. To the contrary, on a rainy day, "effective daytime" is much shorter, making it more costly to build at the same intensity. Third, concrete, bonding agents, and certain other materials cannot be applied when there is heavy rainfall.[14]

Winters in Germany are usually too cold and too windy to allow outside construction work on buildings, and most types of plaster and concrete cannot be handled below certain temperatures, and construction work can only resume once the winter is over.[15]

According to this reasoning, a later start in the spring, or less favorable conditions in the summer may lead to delays that prolong building times at least over the winter. Delays may also last longer if capacity constraints in the construction sector are binding, preventing developers from catching up in the next year.

2.2.2. Definitions of the Instrumental Variables

I use two instruments in the regressions that build on these considerations. The main instrument is the number of consecutive rainfall days ($>$ 20 mm per sqm). As an alternative instrument, I use frost depth in February. Rainfall has the advantage that it is a relevant factor in all parts of Germany — in contrast to snow and frost, which occur only rarely in the north- and north-western parts (e.g., in the Rhine-Main and coastal areas).

The rainfall shock is constructed from daily rainfall data on a 1×1 km^2 grid. I calculate, for each grid cell and month, the largest number of consecutive days with rainfall above 20mm per square meter which I refer to as a "rainfall spell". To control for time-constant differences in weather between different locations, I subtract from

[14] See https://www.nwzonline.de/bauen-wohnen/hausbau

[15] Many materials require outside temperatures above five to ten degree Celsius. Although it is technologically feasible to build also in a cold winter, this increases tremendously the construction costs (see, e.g., Wilke, F. (2016) "Fünf Grad, die magische Grenze" [*Five degree Celsius, the magic threshold*], *Sueddeutsche Zeitung* January 1 2016, https://www.sz.de/201601/bauen).

each grid cell the grid cell mean of the particular calendar month. I then aggregate the resulting variable by location and month. The instrument used in the regressions is the sum of this rainfall spell variable in the three months July, August, and September. February frost depth (in cm) is also provided for 1×1 km^2 grid cells by the German Weather Service, and I use an analogous procedure to construct the 'frost depth' instrumental variable.

To summarize, the identifying variation comes from weather conditions that deviate from the usual conditions at the location. Figure O-C3 displays the spatio-temporal variation in the instrument.[16]

2.2.3. First-Stage Relationship

Table 2 summarizes the results from a set of regressions with the summer rainfall shock and February frost depth as the explanatory variables. The units of observation are municipalities by year. In regressions (1) to (3), the dependent variable is the number of new housing units completed in November and December, per 1,000 units in the housing stock. When regressing the November and December completions on the summer rainfall shock in column (1), the coefficient is highly significant and negative. The *quantitative* impact of the rainfall shock on housing completions is very small. This is consistent with the fact that summer rainfall, a very common phenomenon, is not a key driver of new housing supply. An increase of the rainfall shock by one within-SD (2.3) reduces new housing supply in December by $0.043 \times 2.3 = 0.1$ units per 1,000 units in the stock (within-SD = 5.3). Nonetheless, it provides very useful instrumental variable variation – the quantitative magnitude of the first-stage relationship is not important, beyond the instrument's relevance.[17] Deeper frost in

[16]Taking the actual rainfall instead of locally-demeaned rainfall leads to qualitatively and quantitatively similar results.

[17]One might be concerned that a selection of the instrument based on the strength of the first-stage relationship (among a number of potential candidates) leads to invalid results. However, the exclusion restriction only requires that the instrument be uncorrelated with the regression error. If all candidate instruments meet this restriction individually, this trivially holds also for the selected

February also reduces the number of units completed end-of-year (column (2)). When adding both variables at the same time in column (3), both coefficients are significant and stable, arguably due to the low correlation between the two instruments of .13.

One question not addressed so far is whether the impact of the instruments differs by type of building. Larger buildings have longer construction times, typically exceeding one year. Weather conditions in a single year may thus have a much smaller influence in these cases. In column (4), I investigate this possibility. The dependent variable of the regression is the number of housing units in multifamily buildings completed in November and December, per 1,000 units in the overall residential housing stock. Although the signs of the instruments do not change, February frost depth is no longer significant and the summer rainfall shock has a much smaller impact than in columns (1) and (3), lending support to the hypothesis that larger construction projects are less strongly affected by the weather shocks. Hence, the identified supply shock is mainly a shock to single-family housing supply.

Table 2: Weather shocks and end-of-year completions

Dependent variable:	New housing units completed Nov+Dec per 1,000 units in the stock (2011)			
	in all units			in MFH's
	(1) OLS	(2) OLS	(3) OLS	(4) OLS
Summer rainfall spell	-0.051***		-0.054***	-0.009***
(deviation from local average)	(0.009)		(0.009)	(0.003)
Frost depth in February		-0.010***	-0.012***	-0.003
(deviation from local average)		(0.003)	(0.003)	(0.002)
Year-FE	yes	yes	yes	yes
Municipality-FE	yes	yes	yes	yes
adj. R^2	0.0815	0.0812	0.0816	0.0979
Observations	86,048	86,048	86,048	86,048

Notes: Standard errors are clustered by municipality; *: $p < .1$, **: $p < .05$, ***: $p < .01$. In columns (1) to (3), the dependent variable is the number of housing units completed in December, per 1,000 units in the stock. In column (4), the dependent variable is the number of housing units in multifamily housing (MFH) completed in December, per 1,000 units in the stock.

During housing booms, when the construction sector operates near its maximum capacity, temporary reductions of construction volumes may lead to a quasi-

instrument. It is a widely used strategy to select among valid instruments, based on the correlation with the endogenous variable, because this minimizes the bias.

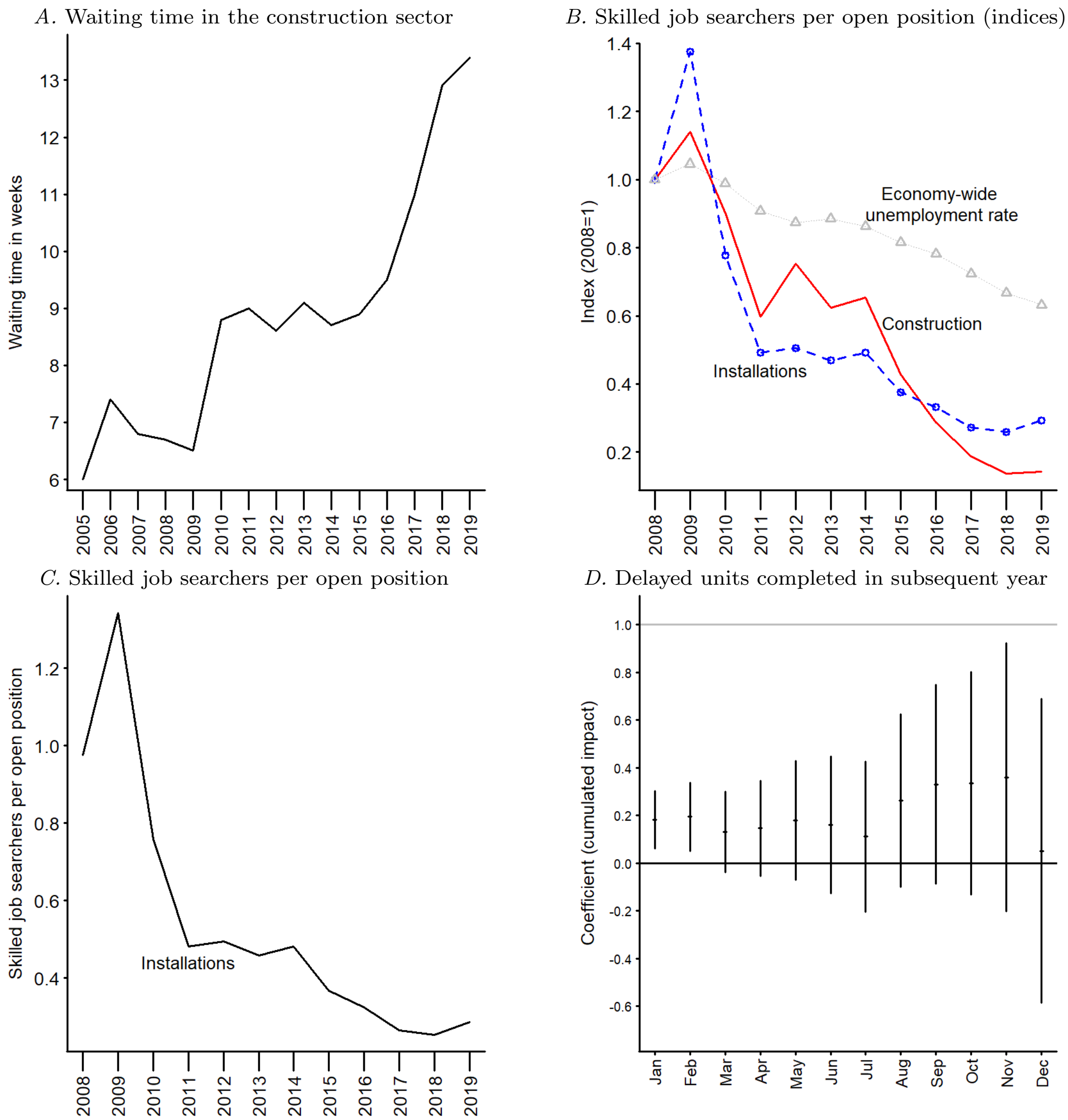
permanent reduction of housing supply. This characterizes very well the situation in Germany since the start of the latest boom in 2010. Waiting times for construction firms (time between signing a contract and the start of its execution) doubled, from 6.5 weeks in 2009 to 13.4 weeks in 2019, and never decreased after 2010 (Panel A of Figure 1). The ratio of skilled job searchers to open positions decreased by a factor of three (installations sub-sector) to five (building construction) (Panel B). In particular, skilled workers in the installations sub-sector were extremely scarce, with only about three skilled job searchers per ten open positions in 2018 (Panel C). This picture is consistent with reports about severe construction capacity constraints during the most recent boom (Gornig et al., 2019).

To investigate the average length of the delay in the sample, Panel D of Figure 1 displays the impact of one building not being completed due to poor weather conditions in the preceding November/December, on the number of buildings completed between January and the given month. The estimates are based on IV regressions of the number of residential building completions between January and month m of the year following the rainfall shock, on the number of November and December completions in the year of the shock (conditional on year and municipality fixed effects). According to the graph, fewer building completions due to unusually poor weather conditions increase the number of building completions in the subsequent year, but only about three out of ten delayed buildings are completed by mid-year. Overall, Figure 1 suggests that the effects of the weather-induced supply shocks lasted longer than one year, consistent with earlier evidence for the U.S. (Fergus, 1999).

2.3. Estimation Results

2.3.1. Effects on Average Rents

Before turning to the effects of a local housing supply shock on the rent distribution, this section investigates the impact of new housing supply on average local rents. I run panel fixed-effects IV regressions at the level of planning regions (ROR



Note: Panel *A* displays average waiting times in the construction industry, from signing of the contract to start of execution (source: ZDH Konjunkturbericht). Panel *B* plots indices for the number of skilled job searchers per open position in the building construction and installations sub-sectors, and for the overall unemployment rate in Germany (base year 2008; source: Federal Employment Agency). Panel *C* shows the number of skilled job searchers per open position in the installations sub-sector (source: Federal Employment Agency). Panel *D* displays the estimated share of delayed units completed by month m of the subsequent year (cumulative). Standard errors are clustered by municipality.

Figure 1: Delayed housing completions and capacity constraints in the building sector

[*Raumordnungsregionen*]), with a hedonic rent index[18] as dependent variable and the housing completions at the end of the preceding year (November and December) as endogenous regressor, instrumented by the rainfall shocks. RORs are a rather broad definition of a local housing market, so that – arguably – local spillovers triggered by the supply shock are contained within the location. The estimating equation is

$$\ln \text{Index}_{rt} = \gamma \left[\frac{S_{r,t-1}^{\text{Nov, Dec}}}{H_r} \times 1,000 \right] + \psi_r + \phi_t + \varepsilon_{rt}, \tag{1}$$

where Index_{rt} is a hedonic rent index of ROR r in year t, $S_{r,t-1}^{\text{Nov, Dec}}$ is the number of units completed in November and December of year $t-1$, H_r is the number of units in the housing stock in 2011, and ψ_r and ϕ_t denote ROR- and year-fixed effects. Standard errors are clustered at the ROR level.

Column (1) of Table 3 displays the results. The main coefficient, γ, is both highly significant and negative. To make sense of the effect size, consider a ROR with the median number of housing completions in November and December (1.65 new homes per 1,000 units in the stock) and assume that housing supply expands to the third quartile (2.47). The estimate suggests that this reduces mean rents in the subsequent year by about $0.82 \times 0.051 \approx 0.042$ log points. When moving from the median down to the first quartile (1.02), rents increase by $0.63 \times 0.051 \approx 0.032$ log points.[19]

According to the Kleibergen-Paap F statistic, there are no signs of weak instrument problems. The first stage relationship in column (1) confirms the municipality-level estimates from Table 2.

Table A1, displays robustness checks. In column (1), I exclude years with extreme rainfall events that led to floods in some federal states. The results do not seem to be driven by these events, with very little impact on the coefficient of interest. Columns

[18]The hedonic index is described in greater detail in Appendix O-B.

[19]Compared to this, the impact of a one standard deviation change in the rainfall shock is much smaller, leading to one delayed unit per 10,000 units in the stock, and hence an impact on rents of -0.0051 log point.

Table 3: Impact of new housing supply on average rents

	Log hedonic rent index	First stage
	(1) IV	(2) OLS
Units completed Nov + Dec of year $t-1$ per 1,000 units in the stock 2011	-0.051*** (0.019)	
Rainfall spell instrument (Jul-Sep of year t-1)		-0.048*** (0.011)
Year FE	yes	yes
ROR FE	yes	yes
Kleibergen-Paap F	19.2	-
Adj. R squared	-	0.801
Number of RORs	94	94
Observations	752	752

Notes: Standard errors are clustered by ROR; *: $p < .1$, **: $p < .05$, ***: $p < .01$. The dependent variable in column (1) is the log hedonic rent index. The instrument is the rainfall shock in year $t-1$.

(2) and (3) replicate the baseline regression, using travel-to-work areas (TTWA) and districts as spatial units. The coefficient is highly significant, but somewhat smaller in magnitude. This could be due to local spillovers: A given supply shock in one district may attract households from surrounding districts that belong to the same ROR.

Columns (4) to (6) display results for regressions using February frost depth as an alternative instrument. This instrument is largely unrelated to the summer rainfall shock (with a ROR-level correlation between the instruments of 0.12), working through a different mechanism: Rather than extending drying times during the summer, it delays starting dates at the beginning of the year. The coefficients are remarkably stable and remain (marginally) significant, although the first-stage relationship is considerably weaker.

Clearly, local housing demand shifters do not influence the spatial or temporal distribution of rainfall shocks. Moreover, most industries in Germany are largely unaffected by summer rainfall or February frost depth, suggesting that the exclusion restriction holds. Nonetheless, it could be that people spend more time working in wetter summers. This could affect local housing demand and supply in the year of the weather shock, and potentially also in the subsequent year. I therefore test the correlation between the instrument and placebo outcomes capturing local housing demand and supply *before or around the time of the weather shock*. These outcomes are the number of housing completions before the rainfall shock (January to June), around the time of the rainfall shock (July to October), log median income of renter households and log GDP in the year of the rainfall shock, and the log difference in these two variables, from the year of the shock to the year after. All outcomes are standardized to achieve comparability across regressions.

Figure A1 displays the resulting coefficient estimates of the rainfall shock instrument, along with 90% confidence intervals. The first two coefficients represent the reduced-form and first-stage impacts relating to Table 3. Longer summer rainfall spells decrease housing completions in November and December (first stage) and increase local rents in the subsequent year (reduced form). However, there is virtually no relationship between the summer rainfall spells and the contemporaneous hedonic rent index. The same holds true for the number of housing units completed between January and June (i.e., in the six months before the rainfall shock), suggesting that summer rainfall did not affect contemporaneous housing demand or supply.[20] There is also no significant relationship between the rainfall shock and housing completions in July to October (i.e., when the rainfall shock occurs). Typical demand shifters (local income, local GDP) are also not related to the rainfall shocks in a statistical sense, neither in levels nor in changes. Overall, Figure A1 strongly suggests that the rainfall shocks are not correlated with local trends in housing demand or supply.

2.3.2. Effects on the Local Rent Distribution

The preceding section provides evidence that a shock to (single family) housing supply shifts average local rents in the subsequent year. However, it is still an open question to what extent new housing supply affects the tails of the local rent distri-

[20] In a simple supply-demand framework, as supply or demand shifts, it must be that the quantity supplied, or the price, or both variables change in response.

bution.

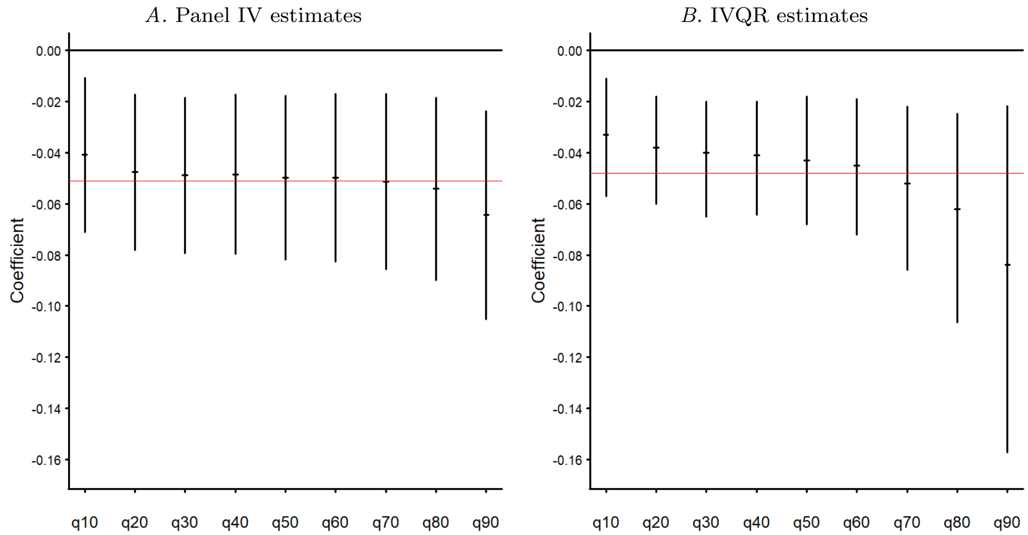
I use two complementary estimation strategies. First, I run a set of panel IV regression with conditional rent quantile indices as outcomes (as in equation 1). The quantile indices are estimated from hedonic quantile regressions, and are hence quality-adjusted, see Appendix O-B for details. Second, I estimate IVQRs (Chernozhukov and Hansen, 2005, 2006, 2008), based on individual-level rent data.[21] IVQR allows studying the effect of a non-randomly assigned treatment (new housing supply) on the distribution of an outcome variable (the rent distribution), if an instrumental variable is available. To simplify notation, let D be the endogenous variable of interest, and let X include the ROR and year fixed effects. The estimating equation is

$$\ln R = \beta(U)X + \gamma(U)D. \tag{2}$$

Here, R is the rent per square meter and $U \sim \mathbb{U}([0,1])$, whereby U may statistically depend on D, but it is independent of X and the instrument Z. The controls X include housing characteristics, and year and ROR fixed effects. Moreover, it is required that the right-hand side of equation (2) increases strictly in U almost everywhere (Chernozhukov and Hansen, 2006). I follow the estimation strategy and inference procedure developed in Chernozhukov and Hansen (2008).

Panel A of Figure 2 displays the impact of the housing supply shock on the first to ninth decile of the ROR-level rent distribution, based on the panel IV regression. The red horizontal line shows the impact on average rents, see column (1) of Table 3. All coefficients are negative and significant at the 5% level or higher, showing a slightly

[21] When using individual units, the first-stage relationship becomes is much weaker in the full sample. The reason is that the weather instruments do not affect much multi-family housing developments, but locations with high shares of multi-family housing also have larger rental housing markets, with many observations in the sample. To address this issue, I drop Berlin and Hamburg, where about 7 of ten new units are in multi-family buildings. The share of observations from these two RORs is 11% in the full sample. The results are qualitatively robust to including Berlin and Hamburg, and Berlin and Hamburg are also included in the ROR-level regressions.



Note: Panel *A* displays coefficient estimates for equation (1), using indices for the conditional quantile of the local rent/sqm distribution (constant-quality) as outcome. The IVQ regressions in Panel *B* control for housing characteristics, and year and ROR fixed effects. The outcome variable is the log rent/sqm, and housing completions in November/December are instrumented by the summer rainfall shock. The vertical bars denote cluster-robust 90% confidence intervals. For Panel *B*, the standard errors were block-bootstrapped based on 200 replications.

Figure 2: Impact of new housing supply on the distribution of rents per sqm

stronger impact at the top of the distribution. However, the variation in impacts is not large, ranging from -0.041 at the first decile to -0.062 at the ninth decile.

The same pattern emerges from the IVQR, depicted in Panel *B* of Figure 2, although the impact at the bottom (top) is somewhat weaker (stronger). The red horizontal line displays the estimate from the analogous linear IV regression (shown below in column (1) of Table 4). As before, all coefficients are significant at least at the 5% level.

Overall, these results indicate tight integration between the market for new single-family homes and all quality segments of the rental market. The next section explores further, complementary heterogeneities in the rent response to the supply shock.

2.3.3. Impact on Rents of Newbuilds

The housing completions data do not provide information about whether the unit will be rented or sold to owner-occupiers. Although the instrument mainly picks up variation in single-family housing completions, there could be a direct effect on rents in new units. Table 4 displays IV estimates of the impact of new supply on rents, by age class (year of construction minus year of observation).[22]

The baseline sample in column (1) is the same as in the IVQR. The coefficient estimate is very close to the estimate from Table 3. In column (2), observations with missing year of construction are excluded. In column (3), the sample consists of new units only (year of construction equal to year of observation). Although still negative, the coefficient is much smaller and insignificant in this regression, showing that the bulk of the impact on average rents does not come from a direct supply effect of built-to-let development. By implication, the effect of new supply on rents runs through secondary housing supply — units freed up by first-time buyers moving into the newly built homes.

Columns(4) to (6) display results for other age classes. Rents decrease only modestly when the building is between one and 20 years old, see column (4). The impact is larger and close to the average impact if building age is between 21 and 60 year in column (5), or more than 60 years in column (5).

Table 4: Effect heterogeneity by building age class

Dependent variable:	Log rent per sqm					
	(1) IV	(2) IV	(3) IV	(4) IV	(5) IV	(6) IV
Age class (years)	any	any	0	1–20	21–60	61+
Units compl. Nov + Dec $t-1$	-0.049***	-0.037***	-0.027	-0.022*	-0.034***	-0.041***
per 1k units in the stock 2011	(0.016)	(0.012)	(0.026)	(0.012)	(0.013)	(0.014)
Year FE	yes	yes	yes	yes	yes	yes
ROR FE	yes	yes	yes	yes	yes	yes
Kleibergen-Paap F	22.9	20.2	28.0	24.7	17.9	12.7
Number of RORs	94	94	94	94	94	94
Observations	6,202,788	4,137,700	314,555	835,029	1,985,742	1,002,360

Note: Standard errors are clustered by municipality; *: $p < .1$, **: $p < .05$, ***: $p < .01$. All regressions control for housing characteristics, location, and year fixed effects. The endogenous dependent variable is the number of housing units completed in the preceding November/December, per 1,000 units in the stock, instrumented by the summer rainfall shock. Column (1) oreports results for the entire sample (excluding Berlin and Hamburg). In Panel *A*, the sample is partitioned by number of rooms. Columns (2)–(6) excludes units with missing information on the year of construction. In columns (3)–(6), the sample is partitioned by building age (year of observation minus year of construction). When Berlin and Hamburg are included, the coefficients are qualitatively similar, but identification is weaker, as indicated by the Kleibergen-Paap F (results available on request).

[22]The year of construction is reported in the description of the unit and may refer to the original year of construction. Buildings may have been refurbished or redeveloped at a later point.

3. A Quantitative Dynamic Model of Housing Quality and Tenure Choice

The preceding section establishes that new housing supply shifts downwards the rent distribution as a whole, but the direct effect on rents for newbuilds was not significant. This section investigates whether *secondary housing supply* — units freed up by renters moving in the local housing market — can explain the reduced-form evidence. Secondary rental housing supply depends on the housing quality and tenure choices of renter households.I develop a dynamic discrete choice model of housing quality and tenure choice that delivers estimates of secondary housing supply in response to a primary supply shock.

The discrete choice model features forward-looking households and uncertainty about the evolution of household income, household savings, and the subjective quality of the match between household and housing unit. It is estimated from histories of renter households observed yearly in the SOEP (2001–2017).[23] Time is discrete and all decisions are made at the household level. As a key ingredient to the model, moving within the local market is costly. Households can move into other rental housing units, buy a new or an existing unit, or leave the local housing market (as an outside option). When making rental housing choices, households care about the match between income and housing quality, as in standard filtering models such as Sweeney (1974a); see also Landvoigt et al. (2015).

I then use the choice model to determine housing demand by renters and secondary rental supply in a local housing market. Secondary supply is given by the sum of inhabitable units vacated in the current period.[24] The standard equilibrium condition (demand equals supply) then determines a rental price for each quality level. A shock to new supply triggers a reordering of choices, with some mover households switching

[23]The sample is restricted to households living in private rental housing (excluding social tenants, tenants of former social housing, and tenants paying a “reduced rent” for other reasons). Households acquiring a home through inheritance or gift are also excluded.

[24]In the model, units may become uninhabitable because of depreciation.

to other quality levels and some stayer households becoming movers (or vice versa), leading to an adjustment of quantities traded in equilibrium, and of the associated rental price vector.

3.1. Data and Descriptive Evidence

The main data source for the model is the SOEP, 2001-2017. From 2001 onward, more precise move indicators are available for the SOEP, which I use to determine whether a household moved.[25] Households enter the estimation sample if they moved into their current rental housing unit between 2001 and 2017.

Throughout the paper, I measure housing quality as the quantile in the local distribution of rent/sqm. I employ rich data on rents from the Mikrozensus, a large repeated cross-section of about 400,000 households.[26] The Mikrozensus included housing modules in 2006, 2010, 2014, and 2018, allowing the estimation of the ROR-level rent/sqm distribution over time, and going back in time longer than the posted rents used in Section 2. Crucially, the distribution of rents/sqm allows assigning units to quality levels (the position in the distribution) without having to make assumptions about the valuation of observable and unobservable housing characteristics. This measure of quality captures all characteristics of the unit (including neighborhood characteristics), except the unit's size.[27]

In Germany, all long-term rental contracts are subject to tenancy rent control. Therefore, rent changes in the years after moving into a housing unit are strictly limited to inflation adjustment, and uncommon overall. I assign respondents in the Mikrozensus to the year they moved into the housing unit, allowing me to construct

[25] As a second source of moving information, respondents indicate the year when they moved into the unit.

[26] Source: RDC of the Federal Statistical Office and Statistical Offices of the Federal States, *Mikrozensus*, survey years 2006, 2010, 2014, 2018.

[27] In order to prevent confounding the quality and size dimensions, I control for size variation in rents/sqm, via a regression using a second-order polynomial in size. This strategy is consistent with the estimation strategy in Section 2, where all regressions control for size of the unit.

a yearly panel.[28] As plausibility checks, I compare the distributions for 2006, 2010, and 2014, as constructed from respondents interviewed in the respective Mikrozensus wave, and from respondents interviewed four years later. The correlations of the 10%, ..., 90% deciles are very high, exceeding .9 in almost all cases.

I define the household's income class as the position in the local income distribution among renters (net household income), again using the Mikrozensus. Here, I interpolate between the waves based on the median rent evolution in the SOEP. As long as all households value housing quality, there should be a positive relationship between housing quality and income, resulting in assortative matching. In the model, I discretize both variables into ten bins, based on the deciles of the ROR$\times$year-level distribution.

In total, there are 2,350 households in the sample with full information on all variables (16,092 household-year observations). There are 250 transitions from renter to owner-occupier status into existing units, 73 transitions into new units, 955 moves within the rental market segment, and 292 moves out of the local market.[29] Tables 5 and 6 report additional details on the sample and summary statistics of important variables. 531 households appear in the data for ten or more years.

Table 5: Number of complete cases in the SOEP household data

years observed	2	3	4	5	6	7	8	9
# of households	248	264	262	259	289	224	153	120
years observed	10	11	12	13	14	15	16	17
# of households	106	95	62	58	56	54	42	58
Total # of hh	2,350							

Panel A of Figure O-D4 plots the distribution of the housing quality-to-income difference for renters, at the time when the renter moved in (black line), and when s/he moved out, into another rental housing unit (red), or into an existing or new owner-

[28]The year of the last move is recorded in binned form only. I use interpolation techniques to construct values for each year.

[29]In interpreting these numbers, please note that households leave the sample when choosing to own, or when moving to a different local housing market, since these choices are modeled as terminal choices.

Table 6: Descriptive statistics for the SOEP household data

	Mean	SD	Quantile .25	.5	.75	Min	Max
Housing quality	5.8	3.0	3.0	6.0	8.0	1.0	10.0
Match quality	-0.1	0.5	0.0	0.0	0.0	-1.0	1.0
Length of tenancy	2.5	3.0	0.0	2.0	4.0	0.0	15.0
Rent per sqm	7.2	2.5	5.7	6.9	8.3	0.3	35.7
Rent per sqm (size-adjusted)	7.2	2.5	5.7	6.8	8.2	0.4	34.4
Local real rent growth since move (log points)	0.00	0.12	-0.05	0.00	0.03	-0.39	0.69
Implied tenancy RC discount (as share of income)	0.01	0.02	-0.00	-0.00	0.01	0.00	0.28
Age of household head	45.3	15.3	33.0	42.0	55.0	18.0	94.0
Monthly net real household income (EUR)	2457	1218	1549	2257	3131	429	8268
Local income class	6.0	2.7	4.0	6.0	8.0	1.0	10.0
Yearly real savings (1k EUR)	2.3	4.9	0.0	0.6	2.4	0.0	96.0
Real financial assets (imputed, 1k EUR)	55	512	1	12	41	0	33351
Year	2009	4	2006	2009	2012	2001	2016

Notes: Sample of SOEP households used in the estimation (periods with observable choices). Housing quality is determined by the position in the local rent/sqm distribution at the time of moving, accounting for depreciation. Match quality is a subjective assessment of the unit's size (-1/too small, 0/just about right, 1/too large). The rent control discount is defined as the difference in the household's rent expenditure share, between the counterfactual situation without tenancy rent control, and the actual rent expenditure share. The income class is the household's position (decile) in the local income distribution (at ROR level) among renters. Financial assets were imputed from SOEP waves 2002, 2007, and 2012 ('wealth module'), using the savings variable (reported in all waves).

occupied unit (blue). The black line shows that the median renter perfectly matches income class to housing quality level. Renters moving out typically have higher income than quality, but this is much more pronounced when moving into owner-occupied housing — mostly due to strong income increases in this group. Arguably contributing to this pattern, small changes in (rental) housing quality do not justify a move if moving costs are high, and life-cycle effects may contribute to strong income increases among future first-time homebuyers. As a consequence, rental housing quality of first-time homebuyers moving into newly built housing is relatively dispersed (Panel B). Although most units have above-average quality, there is considerable weight also at the lower end of the distribution.

3.2. Dynamic Discrete Choice Model

3.2.1. Model Setup

Choice Set. In each period, the household faces a set of $J = 14$ mutually exclusive alternatives $j = 0, ..., 13$. The baseline choice $j = 0$ is to stay in the current accommodation. The household may move to another rental accommodation. Rental housing units differ by quality $\in \{1, ..., 10\}$, mapped to $j = 1, ..., 10$. The size of the unit is not part of the choice set, and I assume that households are able to choose the 'correct size' that matches their needs (no size mismatch). However, the 'size match' may change stochastically from period to period.

The household may also become an owner-occupier and move to an existing ($j = 11$) or a new housing unit ($j = 12$). These two latter choices are terminal.[30] This formulation of the model simplifies the estimation considerably, but it does not interfere with the main goal behind the estimation, namely to determine secondary housing supply to the rental market. The third terminal choice is to move to another local housing market (ROR), $j = 13$.

State Space. Households are characterized by a set of observable and unobservable variables $(x_t, s) =: z_t$. $s \in \{1, 2, 3\}$ is the unobserved household type. The observable variables are the unit's housing quality, $\text{quality}_t \in \{1, ..., 10\}$, its subjective size match, $\text{size}_t \in \{-1, 0, 1\} \equiv \{\text{too small}, \text{just about right}, \text{too large}\}$, the age of the household head, age_t, the household's local income class, $\text{inc}_t \in \{1, ..., 10\}$, the lagged income class, inc_{t-1}, the time since the last move, length of tenancy_t, the implicit subsidy due to tenancy rent control, implicit RC $\text{subsidy}_t \geq 0$, and the accumulated savings of the household, $\text{savings}_t \geq 0$.

Housing quality is defined by the housing unit's position in the ROR-level rent distribution. Specifically, I assign a unit to quality level k if its rent per square meter falls between deciles $k - 1$ and k of the distribution, at the time when the household moves into the unit. Moreover, housing quality depreciates over time, at a negative exponential rate.

The household's income class captures a combination of the household's labor market skill(s), labor supply, and experience. It also includes non-labor income sources. Households move up or down the income distribution over time via a stochastic tran-

[30]This simply means that the utility of owner-occupied housing is modeled in reduced form. The lifetime utilities associated with the owner-occupier choices still capture the possibility that the household moves in the future.

sition function introduced below.

The implicit rent subsidy captures the rent increase on the market since the time the household moved in. Landlords cannot increase rents during a tenancy due to tenancy rent control (apart from inflation adjustments), shielding renters from local rent increases as long as they do not move. In the model, the household derives utility from this differential, in terms of the rent expenditure share. That is, the implicit RC subsidy captures the difference in the household's rent expenditure share, between the counterfactual situation without tenancy rent control, and the actual rent expenditure share. By construction, this variable is zero when the household moves, or when rents have decreased. Since all households in the sample moved in the period they enter the sample, all variation in this variable comes from income and local rent changes in later periods. This addresses the initial conditions problem and related endogeneity concerns.

The accumulated savings measure financial wealth of the household. They evolve through a stochastic transition process that depends on the choice and the income class (defined below). I construct the measure from the financial assets reported in the 2002, 2007, and 2012 SOEP 'wealth modules' and use the savings of the household (reported each year) to calculate forward and backward the financial wealth. In doing so, I ignore potential returns through interest, as well as withdrawals.

Transitions. The household's income class, financial wealth, and match quality transition stochastically. While I do not model explicitly the labor supply or job choice of the household, the transition function incorporates such labor-related changes to the household. Income class changes depend flexibly on the current and lagged income class, on age, and on the difference between today's and tomorrow's income class. Yearly savings are discretized to five levels (0, 1,000, 2,000, 4,000, 8,000 Euro per year), and the transition probability depends on income and the choice, capturing the idea that costly moves may infringe upon the ability of the household to save.

The match quality transition captures changes in household size, and income effects in the preferences for housing unit size. This transition depends on the choice, on the income change from this to the next period, on age, and on the current size mismatch. All transition probabilities are modeled via multinomial logits. The definitions are given in Appendix O-D.

Utility. For non-terminal choices $j < 11$, a household of unobserved type $s \in \{1, 2, 3\}$ receives per-period utility, according to a linear-in-parameters utility function

$$\begin{aligned}
u_{jt}(x_t, s) = {} & \theta_1^s \times \mathbb{1}\left(\Delta\text{qual}_{jt} < 0\right) \times |\Delta\text{qual}_{jt}| + \theta_2^s \times \mathbb{1}\left(\Delta\text{qual}_{jt} > 0\right) \times |\Delta\text{qual}_{jt}| \\
& + \theta_3^s \times \mathbb{1}(\text{size}_{jt} < 0) + \theta_4^s \times \mathbb{1}(\text{size}_{jt} > 0) \\
& + \theta_5^s \times \text{implicit RC subsidy}_{jt} + \theta_6^s \times \text{length of tenancy}_{jt} \\
& + \mathbb{1}(j > 0)\left(\theta_7^s + \theta_8^s \times \text{age}_t + \theta_9^s \times (\text{age}_t)^2\right) + \varepsilon_{jt}. \qquad (3)
\end{aligned}$$

There are two dimensions of rental housing services that are important to the household, housing quality and size. I define the difference between the housing quality and income quantile as $\Delta\text{quality}_{jt} := \text{income class}_t - \text{quality}_{jt} \times \exp(-\delta \times \text{length of tenancy}_t)$. Here, length of tenancy$_t$ is the length of tenancy at the address, δ is a decay factor that captures depreciation of housing quality over time, and $\text{quality}_{jt} = j$, length of tenancy$_t$ equals zero when the household makes a rent-rent move, i.e. for $j \in \{1, ...10\}$. Since both the income class and the housing quality are defined in terms of their underlying distributions, they can be readily compared. The first two components of the utility function reflect marginally decreasing utility of (housing and non-housing) consumption: For negative coefficients θ_1^s and θ_2^s, a larger mismatch decreases utility.

As described above, size$_{jt}$ captures the subjective quality of the household-to-unit size match (θ_3^s, θ_4^s). When the household moves, size$_{jt}$ is reset to zero in the current period.

θ_5^s captures benefit due to tenancy rent control, which shields tenants from rent increases. This may create lock-in effects. Finally, length of tenancy$_{jt}$ is a control for attachment effects and misspecification of housing quality depreciation (θ_6^s).

Moving is costly, and may depend on the age of the household through a quadratic in the household head's age (θ_7^s, θ_8^s, θ_9^s). Older persons might find it harder to move for physical reasons. Note that these moving costs reflect renters' costs of moving *within* a local housing (and labor) market, which may be substantially smaller than the costs of moving between locations or regions, as estimated in Kennan and Walker (2011) or Buchinsky et al. (2014).

The choices to purchase and move into an existing ($j = 11$) or new ($j = 12$) housing unit, and the choice to move to another local housing market ($j = 13$) are terminal. The household does not face any further choices in the future and gets lifetime utility

$$\begin{aligned} v_{jt}(x_t, s) = & [\bar{\theta}_{0j} + \theta_7^s] + \bar{\theta}_{1j} \times \ln(1 + \text{savings}_t) + \bar{\theta}_{2j} \times \mathbb{1}(\text{savings}_t = 0) \\ & + [\bar{\theta}_{3j} + \theta_8^s] \times \text{age}_t + [\bar{\theta}_{4j} + \theta_9^s] \times (\text{age}_t)^2 + \bar{\theta}_{5j} \times \text{income class}_t + \varepsilon_{jt}. \end{aligned} \tag{4}$$

Utility of owning a new or existing home may depend on the log accumulated savings (and a term capturing absence of savings), on the age and age squared, and on the income class. Terminal utility associated with long-distance moves has the same functional form. The coefficients $\bar{\theta}_{0j}, ..., \bar{\theta}_{5j}$ are specific to the alternative ($j = 11, 12, 13$).

Household Problem. Let d_{jt} be an indicator for choosing alternative j in period t, and $d_t = (d_{0,t}, ..., d_{13,t})$. The household's objective in period t is to maximize expected lifetime utility by selecting an optimal choice sequence $d^*(t) := (d^*_{t'})_{t' \geq t}$. Letting $\chi_{t,t'} = \prod_{\tilde{t}=t}^{t'-1} \left(1 - \sum_{j=11}^{13} d_{j\tilde{t}}\right)$ be an indicator for not having made a terminal choice between periods t and $t' - 1$, the household selects

$$d^*(t) = \arg\max_{d(t)} \sum_{t'=t}^{T} \chi_{t,t'} \beta^{t'-t} \left[\sum_{j=0}^{10} d_{jt'} \mathbb{E}_t[u_{jt'}(x_{t'}, s)] + \sum_{j=11}^{13} d_{jt'} \mathbb{E}_t[v_{jt'}(x_{t'}, s)] \right]. \quad (5)$$

3.2.2. Estimation Approach

This section provides a brief summary of the estimation strategy. The supplementary material (Appendix O-D) contains further details.

Depreciation of Housing Quality. I estimate one parameter outside of the model, the depreciation of housing quality. To do so, I resort to the rich rent data employed in Section 2. Depreciation captures the change in the unit's position in the local housing quality distribution. This links closely to filtering theory, where units move down the housing quality distribution over time due to depreciation. Landlords typically do not renovate or refurbish their units during a tenancy. I therefore estimate the depreciation rate from units observed repeatedly in the rent data that were not renovated or altered in any way. The estimated depreciation factor is 4% p.a., capturing pure depreciation net of maintenance. Appendix O-D.3 provides further details.

Discount Factor. I follow the literature in assuming $\beta = 0.95$.

Transition Functions. Since the transition functions do not depend on the unobserved state, they can be estimated in a separate step. The income transition is estimated based on the full sample. The savings transition is estimated from households observed to make a rental housing choice ($j \leq 10$). For the match transition function, the sample consists of stayer households ($j = 0$). These sample restrictions are consistent with the definitions of the transition functions and the model structure.

Dynamic Discrete Choice Problem. Estimation of all other parameters is based on the conditional choice probability (CCP) inversion and the "finite dependence" property (Arcidiacono and Miller, 2011; Hotz and Miller, 1993). I exploit the terminal choice

$j = 11$ of the model to guarantee one-period dependence, leading to a computationally cheap estimation procedure.

3.2.3. Estimation Results

Transition Function Parameters. Table B2 displays the estimated parameters of the transition functions. Larger income jumps are less common than smaller ones, and there is a tendency that past income class changes are undone in the current period. Moreover, the income class increases at a decreasing rate with the age of the household (Panel A). Households with higher incomes are more likely to save higher amounts, and moving reduces the amount saved (Panel B). Finally, households are more inclined towards finding their unit too small when they are younger, and when they experience an increase in income class. This latter result represents an incentive for households to adjust housing consumption to changes in incomes. Older households are relatively more likely to find their unit too large, potentially because children have moved out. Moreover, there is strong persistence in the subjective assessment of the unit's size (Panel C).

Utility Function Parameters. Table 7 displays the parameter estimates for the utility functions.

Panel A of Table 7 refers to the flow utility as a renter, see equation (3). The columns represent the three different unobserved types, with the unconditional shares given by π at the bottom of the panel.

Utility of all three types decreases in the difference between the income class and the housing quality level, both when income is higher or lower than housing quality (with four of the six coefficients significant at the 5% level), lending support to the hypothesis that households choose housing quality to match their income class. For all three types, utility decreases significantly if the unit is deemed too small or too large.

Types $s = 2, 3$ are less likely to move when benefiting from the implicit rent control subsidy – evidence for lock-in effects of tenancy rent control. To put the coefficients in perspective, if the implicit subsidy amounts to a rent expenditure share of 3.2% (the third quartile during the boom, 2011–2016), the lock-in effect for type $s = 2$ ($s = 3$) makes up for about 160% (22%) of the effect of living in a unit that is too small. In 2016, six years into the boom, the third quartile increased to 6.1%, thereby increasing the two fractions to 309% and 41%, respectively.

There is substantial heterogeneity in moving costs, both across unobserved types and along the age dimension. Moving costs are significantly negative for all types. Figure 3 displays the age-moving cost profiles resulting from the estimated coefficients. Putting these numbers into perspective, a moving cost of -5.95 for type $s = 3$ at age 35 is about as bad as staying in a unit that is too small for the next 18 years. This shows that adjustments of housing consumption are rather unlikely to occur, unless there are strong reasons for a move. Put differently, most moves occur because the idiosyncratic shock is large, in addition to experiencing a mismatch of household to housing unit. In many instances, this may not be the case in the years prior to transitioning into owner-occupied housing, leading to more dispersed housing quality among movers into owner-occupied housing.

For type $s = 2$, the coefficient on the implicit rent control subsidy allows for an interpretation of the moving cost coefficient in terms of income shares. At age 35, the moving costs amount to -5.95 in parameter space, the equivalent of 56% of net household income. By all standards, this is much lower than the USD $300,000$ estimated by Kennan and Walker (2011) for inter-regional moves. Nonetheless, it is still a very substantial amount that may well restrain a household from adjusting housing consumption to moderate changes in economic and personal circumstances, exceeding by far the direct financial costs of moving.

Panel B of Table 7 displays the coefficients of the terminal utility functions defined

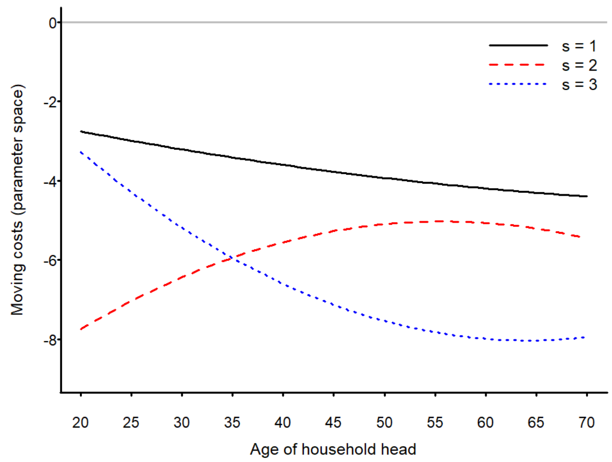
Table 7: Estimated parameters of the utility function

A. Renters			
	Unobserved types		
	$s = 1$	$s = 2$	$s = 3$
Income-quality distance × income lower than quality	-0.1906* [-0.2522, -0.1517]	-0.0374 [-0.0620, 0.0156]	-0.0846* [-0.1241, -0.0117]
Income-quality distance × income higher than quality	-0.1848* [-0.2367, -0.1334]	-0.0532* [-0.1102, -0.0123]	-0.0329 [-0.0523, 0.0355]
Unit too small	-0.3206* [-0.4692, -0.0025]	-0.2093* [-0.4019, -0.0449]	-0.5014* [-0.6978, -0.3368]
Unit too large	-0.5603* [-0.5469, -0.0886]	-1.5642* [-1.6371, -1.2606]	-0.2972* [-0.5640, -0.2263]
Lock-in due to rent control (implicit subsidy as share of income)	-3.3626* [-8.8151, -0.8884]	10.6062* [7.7696, 17.4285]	3.3787 [-2.1636, 10.2668]
Years since last move	-0.0385 [-0.0801, 0.0196]	-0.0706* [-0.1053, -0.0365]	0.0335 [-0.0018, 0.0735]
Moving costs, intercept	-1.6608* [-2.3889, -0.3590]	-11.6328* [-15.2531, -10.7881]	1.9901* [1.2192, 4.1980]
Moving costs, age	-0.0613* [-0.1161, -0.0333]	0.2368* [0.2030, 0.3616]	-0.3122* [-0.4209, -0.2893]
Moving costs, age squared	0.0003* [0.0001, 0.0008]	-0.0021* [-0.0032, -0.0018]	0.0024* [0.0023, 0.0034]
π (shares of unobserved types)	0.097	0.351	0.552

B. Terminal choices			
	Moves to ownership		Long distance
	Existing	new	
Intercept	-21.9364* [-39.7315, -7.4925]	-26.9009* [-46.0450, -14.3191]	-15.5108* [-32.2370, -2.3230]
Accumulated savings	0.2416* [0.0382, 1.3147]	0.4718* [0.2888, 1.5140]	0.1734 [-0.0856, 1.0504]
$\mathbb{1}$(Accumulated savings = 0)	2.0951* [0.5212, 10.4676]	3.0447* [2.3726, 11.6684]	1.6078 [-0.6083, 7.7138]
Age	0.9701* [0.1654, 1.8852]	1.0054* [0.3053, 1.8934]	0.8200* [0.1253, 1.7295]
Age squared	-0.0201* [-0.0363, -0.0068]	-0.0208* [-0.0371, -0.0082]	-0.0185* [-0.0347, -0.0064]
Local income class	0.1916* [0.0485, 0.3165]	0.3566* [0.2311, 0.5120]	-0.0892 [-0.2068, 0.0468]

Notes: 95% confidence intervals (in square brackets) calculated from 500 cluster-bootstrap replications (holding fixed the unobserved type); the estimated coefficients refer to equations (3) (flow utility of renters) and (4) (lifetime utility of owners and long-distance moves). *: $p < .05$, **: $p < .01$, ***: $p < .001$.

by equation (4). Columns (1) and (2) refer to the owner-occupier choices (existing and new housing). The parameters corresponding to long-distance moves are in column (3). Generally, higher accumulated savings and higher incomes increase transition rates into owner-occupied housing, and the effects are relatively stronger for new homes. Secondly, there is a negative, nonlinear effect of age. This is consistent with



Note: Each line represents the estimated age-moving costs profile of a household of the specific unobserved type, employing the coefficients reported in Table 7.

Figure 3: Estimated age-moving cost profiles

higher transaction costs requiring longer anticipated duration at the new address. The latter may be relatively shorter for older residents.

3.3. Secondary Housing Supply

The discrete choice model estimated in the preceding section lends itself to characterizing housing demand and secondary supply from renter households.[31] The discrete choice model did not take into account rental prices as state variables. However, if utility is linear in prices, relative prices merely shift the level of utility associated with a particular choice. In this section, I introduce a scarcity premium in utility-space that is subtracted from the lifetime utility of the respective option, representing the relative rental prices at each quality level (see Landvoigt et al., 2015, who use a similar concept).

[31] While I assume that some units are removed from the market (because of sub-par quality or conversion), I do not model explicitly the landlord decision whether to offer a vacant unit on the rental market. In the German institutional setting, eviction of tenants is regulated strictly by law, making the landlord's decision secondary to the tenant's move decision. With data on landlord decisions, one could incorporate this second-order layer.

3.3.1. Market Structure

I assume the following structure: All renters in the local market evaluate their options based on the discrete choice model. I then fix choices of households choosing to buy. The units vacated by these households add to the supply of rental housing on the market. All remaining local renters, and a number of movers coming from other local markets, then make their rental housing choices, subject to the scarcity premia for each quality level. They are determined such that demand equals supply at each quality level.

Supply. Supply of new and existing owner-occupied housing is taken as given. It equals the number of local households choosing to buy.[32]

Rental supply at quality k consists of

$$S_k(p) = S_k^{\text{new}} + S_k^{o,e} + S_k^{o,n} + S_k^r(p) + S_k^m(p), \tag{6}$$

where S_k^{new} is new supply of rental housing (built-to-let). Secondary supply consists of units vacated by renters buying existing ($S_k^{o,e}$) or new ($S_k^{o,n}$) homes, renters moving within the local market ($S_k^r(p)$), and renters leaving the local market ($S_k^m(p)$). The latter two quantities depend on the scarcity premium vector, $p \in \mathbb{R}^{10}$, whereas S_k^{new}, $S_k^{o,e}$ and $S_k^{o,n}$ are determined by the initial choices.[33] In the model, new rental supply amounts to 30% of overall new supply and is of quality level 10 (4 of 5 new units) and 9 (1 of 5 new units) .

Secondary supply is determined by the units freed up in the local market due to moves. Each vacated unit is assigned to one of ten quality levels, taking into account depreciation.[34] Units below the lowest threshold are assumed to be uninhabitable.

[32]This can be rationalized easily via a utility premium relative to the other choices.

[33]Supply at quality k depends on the full price vector, not just on a particular component.

[34]The assignment is based on thresholds $k \times \exp(-\delta \times 2)$ for $k = 1, ..., 10$, where δ is the quality decay factor. The thresholds are chosen such that units move down the first (discrete) quality level three years after a household moved in.

Moreover, at each quality level, a fraction $\exp(-\kappa k)$ of units is removed from the market, which decreases in quality k for $\kappa > 0$.[35]

Demand. Rental housing demand at quality level k is

$$D_k(p) = D_k^l(p) + D_k^f(p). \tag{7}$$

$D_k^l(p)$ and $D_k^f(p)$ are the demands from local and foreign renters willing to move into a unit of quality k, given the scarcity premia p. Foreign renters are a fixed set of households who choose among the rental housing options only, but they do not contribute to secondary supply. Local demand for quality k may change because movers opt for a different quality level, or because they switch to staying or to a long-distance move.

Equilibrium. The equilibrium scarcity premium vector p^* satisfies

$$D_k(p^*) = S_k(p^*) \quad \forall k \in \{1, ..., 10\}. \tag{8}$$

3.3.2. Simulated data

To build a model economy, I simulate data using the estimated discrete choice model. I start with the 58 households in the estimation sample that were at most 21 years old when first observed, employing the estimated conditional probabilities of being type s. From this set, I randomly draw with replacement a set of 1,000 households and their unobserved type. For this baseline sample of 1,000 households, I simulate outcomes for 2 to 45 periods into the future (44 cohorts). This leads to a sample of renters with a housing choice, wealth, income, and age structure as implied by the model. In the supply-demand exercise, I consider jointly all cohorts in their

[35] I set $\kappa = .0215$, because then the total number of units on the market matches the total number of mover households for $p = 0$.

final simulation period.

With this procedure, I construct two simulated samples. In Scenario 1, there are no lock-in effects due to rent control (assuming stable rents). In Scenario 2, rents increase steadily, and housing expenditure shares increase depending on the change in income, resulting in lock-in effects. Figures O-D5 and O-D6 displays graphs for the simulated data corresponding to Figure O-D4. In the simulated data, the same qualitative patterns emerge.

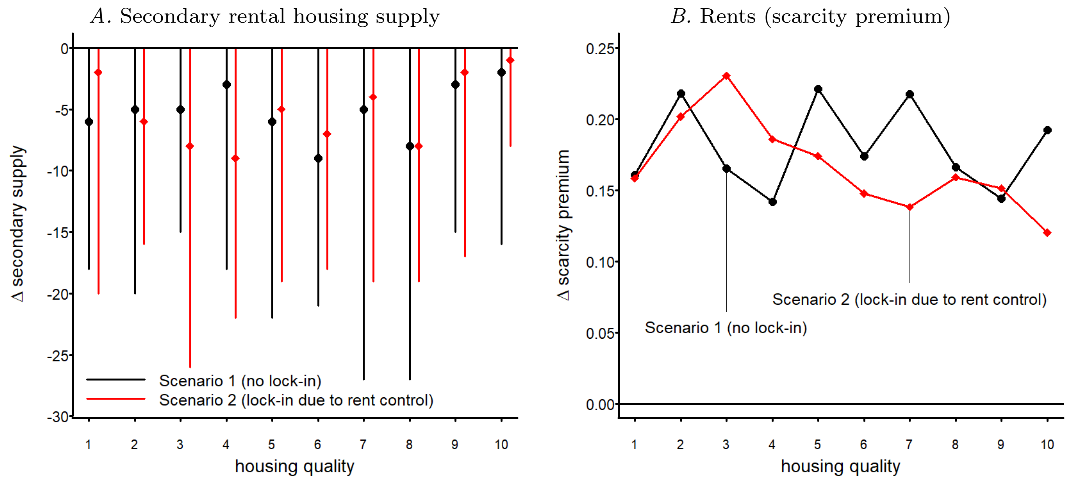
3.3.3. Simulated impact of new supply to the owner-occupier market

For the simulation exercise, the initial equilibrium is determined first, for a fixed level of new supply (owner-occupied and built-to-let). The supply shock caused by summer rainfall spells could be thought of as the random removal of new buildings from the market. In the model, I consider a shock that takes away randomly 50% of new supply to the owner-occupier market, without affecting the supply of built-to-let units.[36]

Equilibrium quantities. Panel *A* of Figure 4 displays the impact of the negative supply shock to the owner-occupier market, on secondary supply to the rental market. The dots represent the change in secondary supply from renters buying a newly built home (units left vacant by these renters). The distribution is relatively dispersed, with higher weight at quality levels 5–8 (Scenario 1) and 3–8 (Scenario 2). Quite strikingly, the initial moves by renters buying a new housing unit lead to many other secondary moves, as depicted by the vertical bars. In Scenario 1, one buyer of new housing triggers 3.8 additional moves on average. Lock-in effects due to rent control reduce this number to 3.5 in Scenario 2. Overall, these moves contribute to a more even distribution of the secondary supply shock in both scenarios, again with considerably

[36] In further simulations, I also consider shocks to new supply before choices are made (which affects the choices of marginal buyers, rather than a random subset of buyers). The results are very similar.

stronger impact at the lower end in Scenario 2.[37]



Notes: The graphs display the impact of a negative shock to new housing supply to the owner-occupier market, for Scenarios 1 and 2. Panel *A* shows the impact on secondary supply to the rental market. The dots represent the absolute change in secondary supply from renters buying a newly built home. The vertical bars represent the absolute change in total secondary supply of all movers in the rental market. Panel *B* shows the impact on the rent distribution (scarcity premium). A larger scarcity premium represents higher rents.

Figure 4: Simulated impact of a negative shock to new housing supply

Impact on rents. In a model where rents enter the utility function linearly, the scarcity premium is directly related to rents via a scaling factor (i.e., a coefficient). Panel *B* of Figure 4 plots the scarcity premia under the two scenarios (with and without lock-in effects due to rent control). Consistent with the results documented in Section 2, new supply to the owner-occupier market triggers secondary supply and thereby shifts the location of the rent distribution. The impact on lower-quality rents is somewhat larger with tenancy rent control: Lock-in effects increase the overall impact of depreciation, tilting the distribution of secondary housing supply towards lower quality levels.

4. Conclusions

Second-hand markets are expected to grow as more firms and consumers refurbish and reuse products in a global effort to combat climate change.[38] It is thus important

[37]Figure O-D7 replicates the graph with the change in log secondary supply, confirming that the differences in absolute values are not due to different magnitudes in the baseline equilibrium.
[38]Scarsella, A. and Stofega, W. (2020), *Worldwide Used Smartphone Forecast, 2020–2024*, IDC

to understand how demand and supply in the markets for second-hand and new products interact. Market integration in second-hand markets with heterogeneous products — such as the housing, car, and smartphone markets — depends crucially on direct links created by buyers of new and used products, who simultaneously act as sellers on the second-hand market. This paper provides a detailed account of such interactions in the housing market, by identifying the impact of new housing supply at market rates on the distribution of rents.

The channel through which these effects operate is secondary housing supply — units freed up by renters moving into the newly built units, triggering a cascade of moves. Through this cascade, new supply of single family housing may have strong effects on the rent distribution. To begin, many first-time homeowners live in rental housing units of moderate quality, because moving costs restrain renter households from adjusting housing quality in reaction to income changes. Hence, households move only if they can improve significantly their housing situation, but they typically do not respond to smaller income changes. Lower housing quality levels are affected by the shock because other renters adjust their choices when further units are left vacant in the process.

With higher moving costs, the impact of a shock to a given point in the housing quality distribution travels downstream more quickly, because the point is connected to a much lower quality level via a single mover. The quantitative model shows that the overall number of rental units traded on the market increases by about three and a half units for each newly built owner-occupied unit, and this amplification mechanism works throughout the housing quality distribution, contributing crucially to market integration. Moreover, it does not rely on substitutability.

In addition, the results suggest that restrictions to overall housing supply may be even more harmful to low-income households than previously thought. Even the supply of single-family homes can lower housing costs of renters. It is likely that the supply of new multi-family housing has similar–if not stronger–effects on the rent distribution. Policy makers should thus focus on removing barriers to the supply of new housing, and on creating a tax system that provides incentives encouraging optimal land use.

Future research on the effects of housing policies should take into account the forward-looking nature of housing choices. Moving locally is costly, especially in markets with tenancy rent control. Ignoring these moving costs may lead researchers to underestimate welfare effects of housing and labor market policies, such as rent controls, rental housing regulation, social housing, housing vouchers, and unemployment benefits.

References

Arcidiacono, P. and R. A. Miller (2011). Conditional Choice Probability Estimation of Dynamic Discrete Choice Models With Unobserved Heterogeneity. *Econometrica 79*(6), 1823–1867.

Athey, S., J. Tibshirani, S. Wager, et al. (2019). Generalized Random Forests. *Annals of Statistics 47*(2), 1148–1178.

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

Buchinsky, M., C. Gotlibovski, and O. Lifshitz (2014). Residential Location, Work Location, and Labor Market Outcomes of Immigrants in Israel. *Econometrica 82*(3), 995–1054.

Büchler, S. C., M. von Ehrlich, and O. Schöni (2019). The Amplifying Effect of Capitalization Rates on Housing Supply. CRED Research Paper No. 24.

Chen, J., S. Esteban, and M. Shum (2013). When Do Secondary Markets Harm Firms? *American Economic Review 103*(7), 2911–2934.

Chernozhukov, V. and C. Hansen (2005). An IV Model of Quantile Treatment Effects. *Econometrica 73*(1), 245–261.

Chernozhukov, V. and C. Hansen (2006). Instrumental Quantile Regression Inference for Structural and Treatment Effect Models. *Journal of Econometrics 132*(2), 491–525.

Chernozhukov, V. and C. Hansen (2008). Instrumental variable quantile regression: A robust inference approach. *Journal of Econometrics 142*(1), 379–398.

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.

Coulson, N. E. and C. Richard (1996). The Dynamic Impact of Unseasonable Weather on Construction Activity. *Real Estate Economics 24*(2), 179–194.

Dana, J. D. and Y.-F. Fong (2011). Product Quality, Reputation, and Market Structure. *International Economic Review 52*(4), 1059–1076.

Diamond, R., T. McQuade, and F. Qian (2019). The Effects of Rent Control Expansion on Tenants, Landlords, and Inequality: Evidence from San Francisco. *American Economic Review 109*(9), 3365–94.

Fergus, J. T. (1999). Where, when, and by how much does abnormal weather affect housing construction? *The Journal of Real Estate Finance and Economics 18*(1), 63–87.

Glaeser, E. L., J. Gyourko, and R. E. Saks (2005). Why have housing prices gone up? *American Economic Review 95*(2), 329–333.

Gornig, M., C. Michelsen, and M. Bruns (2019, January). Bauwirtschaft weiter im Vorwärtsgang – staatliche Impulse treiben die Preise. DIW Wochenbericht.

Gyourko, J., C. Mayer, and T. Sinai (2013). Superstar Cities. *American Economic Journal: Economic Policy 5*(4), 167–99.

Hilber, C. A. and W. Vermeulen (2016). The Impact of Supply Constraints on House Prices in England. *The Economic Journal 126*(591), 358–405.

Hilber, C. A. L. and A. Mense (2021). Why Have House Prices Risen so Much More than Rents in Superstar Cities? Centre for Economic Performance Discussion Paper No.1743.

Hotz, V. J. and R. A. Miller (1993). Conditional Choice Probabilities and the Estimation of Dynamic Models. *The Review of Economic Studies 60*(3), 497–529.

Johnson, J. P. (2011). Secondary markets with changing preferences. *The RAND Journal of Economics 42*(3), 555–574.

Kennan, J. and J. R. Walker (2011). The effect of expected income on individual migration decisions. *Econometrica 79*(1), 211–251.

Landvoigt, T., M. Piazzesi, and M. Schneider (2015). The Housing Market (s) of San Diego. *American Economic Review 105*(4), 1371–1407.

Mast, E. (2019). The Effect of New Market-Rate Housing Construction on the Low-Income Housing Market Income Housing Market. Working paper 19-307, Upjohn Institute.

Metcalf, G. (2018). Sand Castles Before the Tide? Affordable Housing in Expensive Cities. *Journal of Economic Perspectives 32*(1), 59–80.

Molloy, R., C. Nathanson, and A. Paciorek (2020). Housing Supply and Affordability: Evidence from Rents, Housing Consumption, and Household Location. Mimeo.

Nathanson, C. G. (2019). Trickle-Down Housing Economics. Working paper, Kellogg School of Management.

Oswald, F. (2019). The effect of homeownership on the option value of regional migration. *Quantitative Economics 10*(4), 1453–1493.

Pennington, K. (2021). Does Building New Housing Cause Displacement? The Supply and Demand Eects of Construction in San Francisco. mimeo.

Piazzesi, M., M. Schneider, and J. Stroebel (2020). Segmented Housing Search. *American Economic Review 110*(3), 720–59.

Quigley, J. M. and S. Raphael (2005). Regulation and the high cost of housing in California. *American Economic Review 95*(2), 323–328.

Rampini, A. A. (2019). Financing Durable Assets. *American Economic Review 109*(2), 664–701.

Rosenthal, S. S. (2014). Are Private Markets and Filtering a Viable Source of Low-Income Housing? Estimates from a “Repeat Income” Model. *American Economic Review 104*(2), 687–706.

Rosenthal, S. S. (2019). Owned Now Rented Later? The Effect of Housing Stock Transitions on Affordable Housing and Market Dynamics. Working paper, Syracuse University.

Saiz, A. (2010). The Geographic Determinants of Housing Supply. *The Quarterly Journal of Economics 125*(3), 1253–1296.

Schmutz, B. and M. Sidibé (2019). Frictional labour mobility. *The Review of Economic Studies 86*(4), 1779–1826.

Schwarz, T. (2018, August). Baufertigstellungen im Wohnbau 2017. Statistisches monathsheft, Landesamt für Statistik Baden-Württemberg.

SOEP Group (2019). SOEP-Core v34 HGEN: Household-Related Status and GeneratedVariables. SOEP Survey Papers 759: Series D – Variable Descriptions and Coding. Technical report, DIW Berlin, Berlin: DIW Berlin/SOEP.

Sweeney, J. L. (1974a). A Commodity Hierarchy Model of the Rental Housing Market. *Journal of Urban Economics 1*(3), 288–323.

Sweeney, J. L. (1974b). Quality, Commodity Hierarchies, and Housing Markets. *Econometrica 42*(1), 147–167.

Waldman, M. (2003). Durable Goods Theory for Real World Markets. *Journal of Economic Perspectives 17*(1), 131–154.

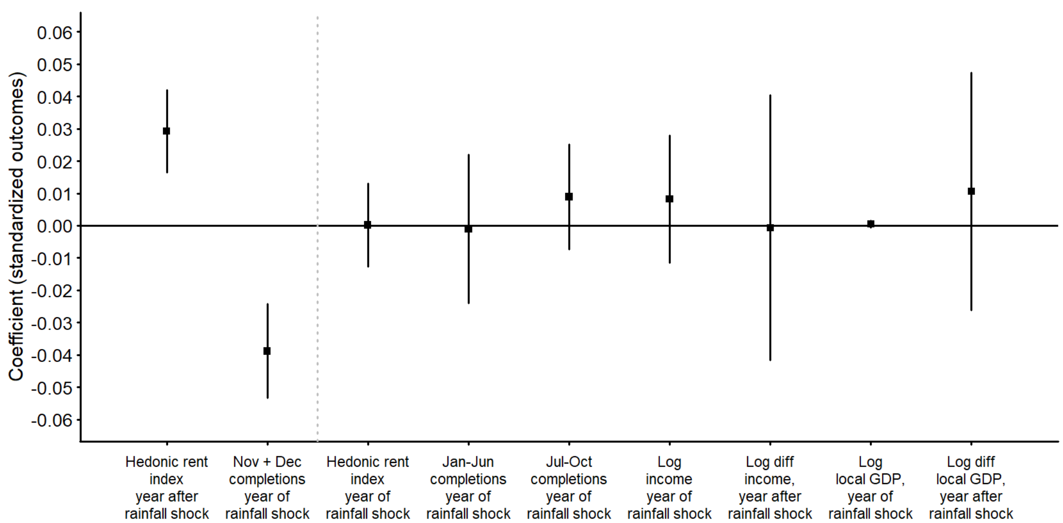
Appendix

A. Robustness and IV Balance

Table A1: Robustness of IV rent regressions to extreme weather events, location definitions, and choice of instrument

Dependent variable	Log hedonic rent index					
	(1) IV	(2) IV	(3) IV	(4) IV	(5) IV	(6) IV
Units completed Nov + Dec of year $t-1$ per 1,000 units in the stock 2011	-0.056*** (0.021)	-0.024** (0.010)	-0.028*** (0.009)	-0.042* (0.025)	-0.054** (0.024)	-0.044*** (0.017)
Year FE	yes	yes	yes	yes	yes	yes
Location FE	yes	yes	yes	yes	yes	yes
Instrument	rainfall	rainfall	rainfall	frost	frost	frost
Kleibergen-Paap F	18.1	18.0	19.8	6.3	7.0	9.3
Spatial unit	ROR	TTWA	district	ROR	TTWA	district
Number of spatial units	94	252	392	94	252	392
Observations	666	2,016	3,136	752	2,016	3,136

Notes: Standard errors are clustered by location; *: $p < .1$, **: $p < .05$, ***: $p < .01$. The dependent variable is the log hedonic rent index. In column (1), locations with severe floods during the summer were excluded (2013: Lower Saxony, Hesse, Rheinland-Palatinate, Baden-Wuerttemberg, Bavaria, Saxony, Saxony-Anhalt, Thuringia; 2017: Lower Saxony, Saxony-Anhalt, Thuringia). Columns (1) to (3) use the summer rainfall shock as instrument; in columns (4) to (6), the instrument is based on February frost depth. Location delineations are RORs in columns (1) and (4), TTWAs in columns (2) and (5), and districts in columns (3) and (6). TTWAs are based on the classification of the Federal Office for Building and Regional Planning.



Notes: Vertical bars denote 90% confidence intervals (clustered by ROR). Each variable denoted at the horizontal axis indicates an outcome variable in a regression of the outcome on the rainfall shock instrument, conditioning on location and year fixed effects. Median renter income is calculated from the SOEP. Local GDP is taken from the regional input-output tables [*Volkswirtschaftliche Gesamtrechnung*].

Figure A1: IV balance. Reduced form, first stage, and placebo outcomes

B. Dynamic Discrete Choice Model: State Transition Functions

Table B2: Parameter estimates for the transition functions

A. Income transition, see equation (O-D1)		
	Coef	SE
$\mathbb{1}(\Delta_{t+1}\text{inc} < 0) \times \Delta_{t+1}\text{inc}$	-0.820***	0.017
$\mathbb{1}(\Delta_{t+1}\text{inc} < 0) \times \Delta_{t}\text{inc}$	0.204***	0.014
$\mathbb{1}(\Delta_{t+1}\text{inc} < 0) \times \text{age}$	-0.015***	0.002
$\mathbb{1}(\Delta_{t+1}\text{inc} < 0) \times \text{age}^2 \times 1e^{-3}$	0.116***	0.031
$\mathbb{1}(\Delta_{t+1}\text{inc} > 0) \times \Delta_{t+1}\text{inc}$	-0.707***	0.015
$\mathbb{1}(\Delta_{t+1}\text{inc} > 0) \times \Delta_{t}\text{inc}$	-0.178***	0.014
$\mathbb{1}(\Delta_{t+1}\text{inc} > 0) \times \text{age}$	0.017***	0.002
$\mathbb{1}(\Delta_{t+1}\text{inc} > 0) \times \text{age}^2 \times 1e^{-3}$	-0.399***	0.030
Log Likelihood	-20,980	
Observations	13,742	
B. Savings transition, see equation (O-D2)		
	Coef	SE
1k: intercept	-1.540***	0.060
1k: hh moves in current period (rent-rent)	-0.207*	0.096
1k: income class	0.101***	0.010
2k: intercept	-2.796***	0.077
2k: hh moves in current period (rent-rent)	-0.468***	0.116
2k: income class	0.278***	0.011
4k: intercept	-3.880***	0.096
4k: hh moves in current period (rent-rent)	-0.453***	0.123
4k: income class	0.419***	0.013
8k: intercept	-7.445***	0.175
8k: hh moves in current period (rent-rent)	-0.456***	0.136
8k: income class	0.852***	0.021
Log Likelihood	-16,976	
Observations	13,200	
C. Match quality transition, see equation (O-D3)		
	Coef	SE
$h = 0$: intercept ($\tilde{\omega}_{0,0}$)	-2.732***	0.108
$h = 0$: Δ_{t+1}inc ($\tilde{\omega}_{1,0}$)	-0.066**	0.020
$h = 0$: age ($\tilde{\omega}_{2,0}$)	0.038***	0.002
$h = 0$: $\text{match}_{t+1} = 0$ ($\tilde{\omega}^0_{3,0}$)	3.273***	0.063
$h = 0$: $\text{match}_{t+1} = 1$ ($\tilde{\omega}^1_{3,0}$)	4.088***	0.280
$h = 1$: intercept ($\tilde{\omega}_{0,1}$)	-6.487***	0.319
$h = 1$: Δ_{t+1}inc ($\tilde{\omega}_{1,1}$)	-0.197***	0.036
$h = 1$: age ($\tilde{\omega}_{2,1}$)	0.039***	0.003
$h = 1$: $\text{match}_{t+1} = 0$ ($\tilde{\omega}^0_{3,1}$)	3.872***	0.286
$h = 1$: $\text{match}_{t+1} = 1$ ($\tilde{\omega}^1_{3,1}$)	8.163***	0.391
Log Likelihood	-6,116	
Observations	12,245	

Notes: Standard errors in parentheses; ; *: $p < .05$, **: $p < .01$, ***: $p < .001$. In Panel *A*, Δ_{t+1}inc is the lead change in income class. $\Delta_t$inc is the observed change in income class between periods $t-1$ and t. In Panel *B*, the outcome is a discretized savings variable (savings of 0, 1k, 2k, 4k, 8k EUR in the current year) and the baseline outcome is 'no savings this year'. The sample is restricted to renter households that make a rental housing choice (stay or move to another rental accommodation). In Panel *C*, the $\tilde{\omega}$'s refer to the parameters in the estimating equation, h denotes the match quality in the next period, and the sample is restricted to renter households that stay at their current address.

Online Appendix — NOT FOR PUBLICATION

O-A. Building Completions Data

The main explanatory variable in the rents regressions is the number of housing units completed in a municipality in December. This variable is aggregated from individual observations in the Building Completions Statistic. The Building Completions Statistic is an administrative statistic that contains all building completions in Germany. There are severe penalties for developers who do not acquire permission to build. Fines range from 500 to 50,000 Euro, and the authorities can oblige the owner to demolish the building at the owner's expense. Information on the month of completion is not provided in individual years by some federal states. I exclude the respective state-years from the analysis.

Figure O-A1 shows the variation in building completions by calendar month. Most buildings are reported to be completed in December (Panel *A*: shares; Panel *B*: completions by month per 1000 units in the stock (2011)).

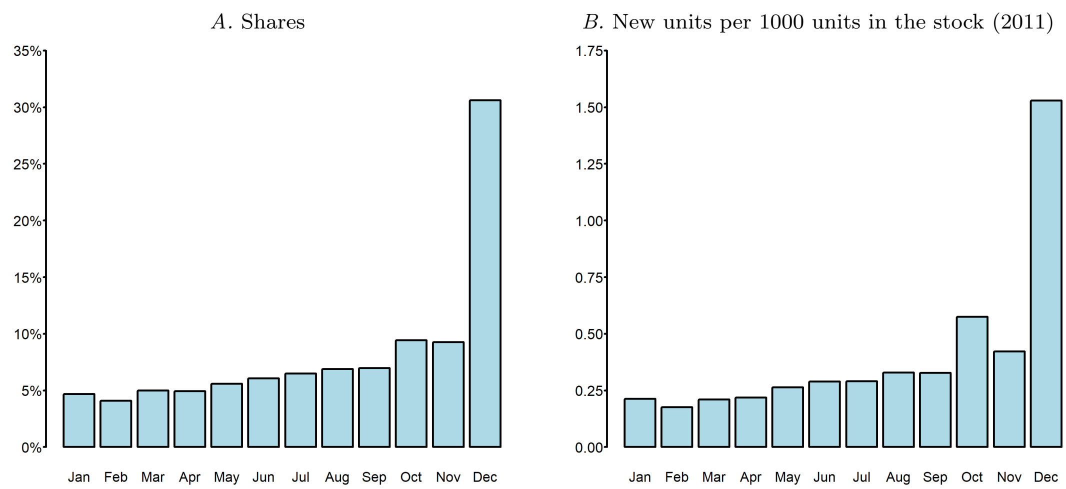


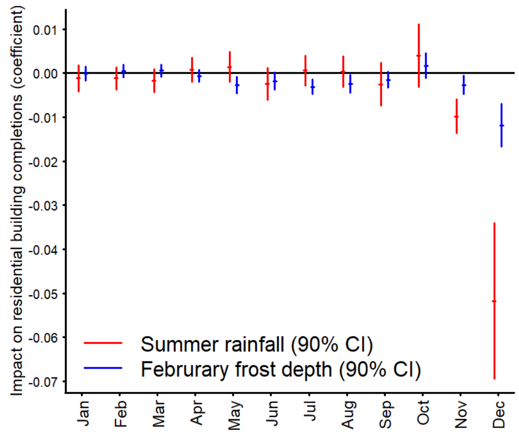
Figure O-A1: New housing units completed in Germany 2010–2017, by calendar month

O-B. Rental Housing Data

Data Source. The rents data were collected between July 2011 and December 2018 via web scraping from three large online real estate market places, Immoscout24, Immonet, and Immowelt. Immonet and Immowelt merged in 2015, but continue to coexist as websites. Duplicates were removed based on a comparison of key variables. The three websites have a combined market share of 80–90%, according to Immoscout24 and the Federal Cartel Office of Germany. All other market places are considerably smaller, see the report "Freigabe des Zusammenschlusses von Online-Immobilienplattformen", Bundeskartellamt B6-39/15 [Federal Cartel Office]. Immonet and Immowelt merged in 2015. In February 2018, Immobilienverband Deutschland conducted a survey "Usage of Real Estate Online Market Places" [*Nutzung von Immobilienportalen*] among 1,287 real estate agents, 99.3% of the respondents use third-party real estate market places for marketing purposes. 76% use Immonet/Immoscout, and 74.4% use Immobilienscout24 (multiple answers possible). Respondents also indicated that 84% of all rental units were offered on at least two different real estate market places.

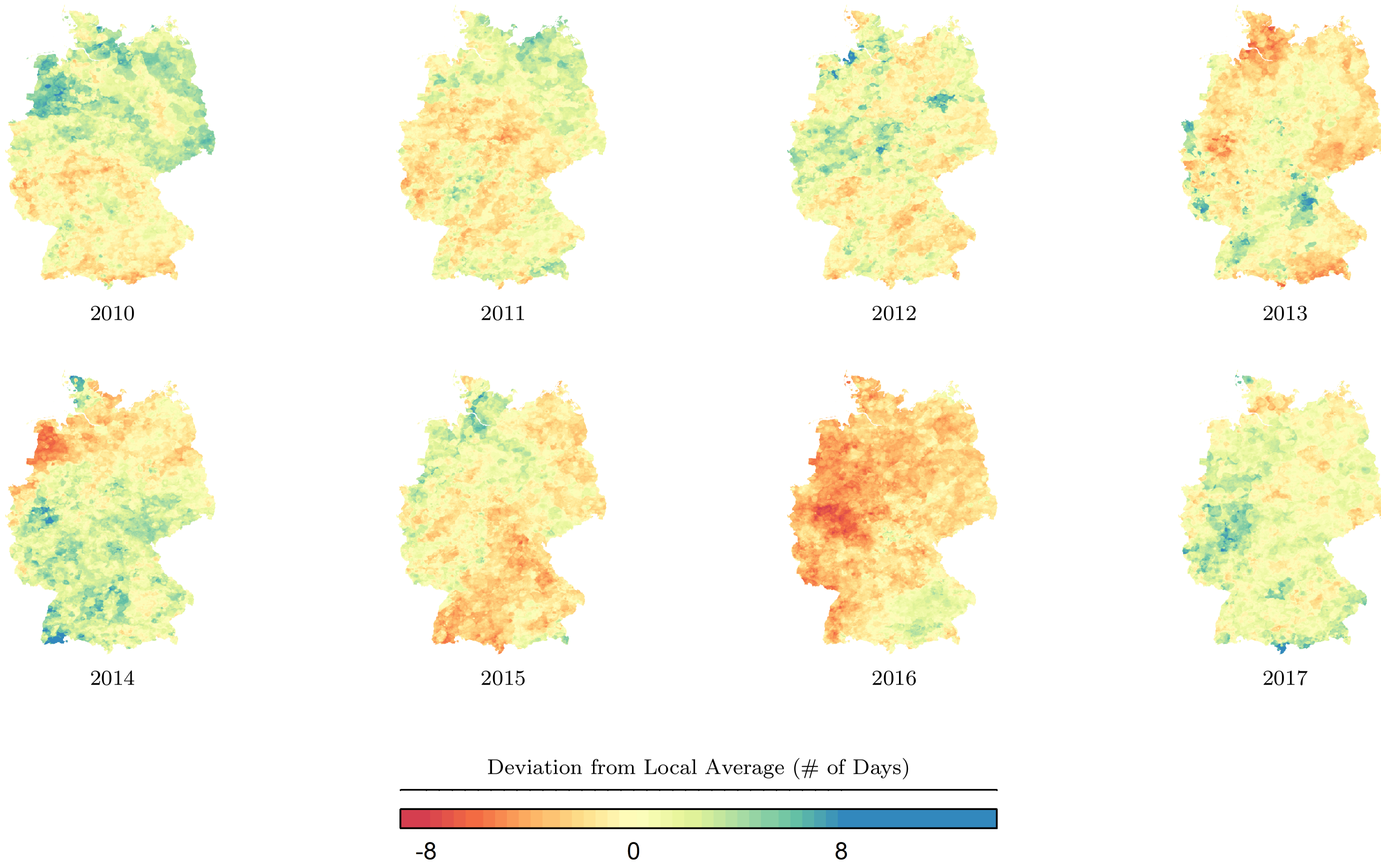
District-Level Rent Indices. In order to calculate the district-level rent indices, I run separate hedonic regressions for each district, with the log rent per square meter as the dependent variable, and a set of housing characteristics and year fixed effects as controls. The resulting index value for year t is given by $\exp(\mathrm{FE}_t)$, the exponential of year t's fixed effect. The controls are the log floor area, a second-order polynomial in the year of construction, an indicator variable for observations where the year of construction was not reported, dummies for the presence of floor heating, parquet flooring, an elevator, a fitted kitchen, a second bathroom, a balcony or a terrace, a garden, and categorial quality and condition indicators. The quantile indices are calculated from analogous quantile regressions.

O-C. Weather shocks as temporary shifters of new housing supply



Note: The graph displays coefficient estimates of regressions with the number of new units completed in month m per 1,000 units in the stock as the dependent variable, on the summer rainfall shock and February frost depth instruments (measured in the same year), at the level of municipalities. Vertical bars indicate 90% confidence intervals.

Figure O-C2: Impact of the weather shocks on new housing supply throughout the year

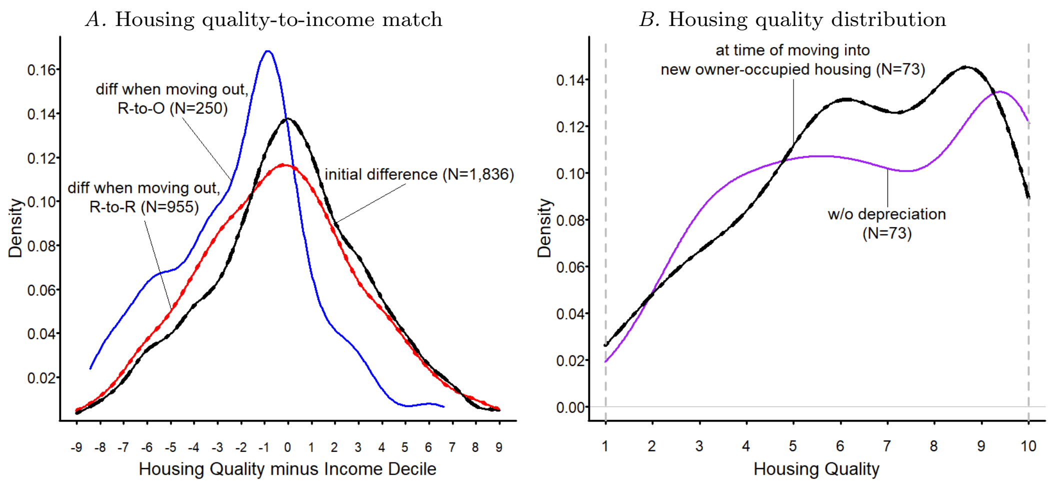


Notes: Each graph displays the variation in the rainfall shock, by municipality. The rainfall shock is measured as the number of consecutive days with rainfall above 20mm during the summer months (Jul-Aug-Sep), relative to the average number of consecutive rainfall days at the location during the summer months. A larger number indicates more rainfall in the particular year than in an average year. This variable is used as instrumental variable in the IV rents regressions below.

Figure O-C3: Spatial and temporal variation in the summer rainfall shock instrument

O-D. Estimation of the Dynamic Discrete Choice Model

O-D.1. Descriptive Evidence



Note: Panel *A* displays densities of the difference between housing quality level and income class in the estimation sample. The black line is the difference when moving in. The red (blue) line denotes the difference at the time of moving into another rental housing unit (owner-occupied housing). Panel *B* shows the density of housing quality among renters moving into newly built owner-occupied housing (including and net of depreciation).

Figure O-D4: Housing quality and income of renters and first-time homebuyers

O-D.2. Transition Functions

Definitions of the State Transition Functions. For income classes $l, k \in \{1, ..., 10\}$ in periods $t+1$ and t, the income class transition is defined by the equations

$$\begin{aligned} y_l^{\text{inc}} =& \mathbb{1}(\Delta_{t+1}\text{inc} < 0) \times \left(\tilde{\mu}_1^-|l-k| + \tilde{\mu}_2^-|\Delta_t\text{inc}| + \tilde{\mu}_3^-\text{age}_t + \tilde{\mu}_4^-\text{age}_t^2\right) + \\ & \mathbb{1}(\Delta_{t+1}\text{inc} > 0) \times \left(\tilde{\mu}_1^+|l-k| + \tilde{\mu}_2^+|\Delta_t\text{inc}| + \tilde{\mu}_3^+ + \text{age}_t + \tilde{\mu}_4^+\text{age}_t^2\right) + \varepsilon_l^{\text{inc}}. \end{aligned} \tag{O-D1}$$

$\varepsilon_l^{\text{inc}}$ is distributed iid Type-I extreme value, y_l^{inc} is a latent variable. The transition function captures in a flexible way the relationship between age and income, and path dependency of the income transition. It depends on the absolute difference between the current and new income classes, $|l-k|$, the sign of this difference, the age and age squared of the household head, and the change in income between period $t-1$ and t, denoted by $\Delta_t$inc.

The savings transition is defined via

$$y_w^{\text{savings}} = \tilde{\gamma}_{0,w} + \tilde{\gamma}_{1,w}\mathbb{1}(j > 0) + \tilde{\gamma}_{2,w}\text{inc}_t + \varepsilon_w^{\text{savings}}. \tag{O-D2}$$

y_w^{savings} is a latent variable for yearly savings $w \in \{0, 1k, 2k, 4k, 8k\}$ in EUR and $\varepsilon_w^{\text{savings}}$ is iid Type-I extreme value. The amount saved in a given year may depend on a move indicator and on the income class. If moves are costly also in monetary terms, households may be able to save less. Higher-income households should be able to save more. Accumulated savings are discretized, and next year's value is given by this year's accumulated savings plus the (stochastic) amount saved. Hence, although agents cannot decide how much to save in this setting, their choices nonetheless affect wealth accumulation.

While in the model, housing quality is a deterministic variable, I allow the match quality with regard to housing unit size to change in a probabilistic way. This captures the effects of all types of changes to the household's subjective space requirement, such as changes in household size or household income, some of which are not perfectly predictable for the household. The transition function is defined by

$$y_h^{\text{size}} = \tilde{\omega}_{0,h} + \tilde{\omega}_{1,h}\Delta_{t+1}\text{inc} + \tilde{\omega}_{2,h}\text{age}_t + \sum_{k\in\{0,1\}} \tilde{\omega}_{3,h}^{k}\mathbb{1}(\text{size}_t = k) + \varepsilon_h^{\text{size}}. \qquad \text{(O-D3)}$$

$\varepsilon_h^{\text{size}}$ is distributed iid Type-I extreme value and y_h^{size} is a latent variable. This transition function introduces a dependency between the income and match quality transitions, through Δ_{t+1}inc. Moreover, the size match is affected by age and the current match quality of the household. If households care about the match quality, expected income changes provide a dynamic incentive in the model for the household to adjust housing consumption. For choices $j \in \{1, ..., 10\}$, the current-period match quality is assumed to be zero.

Note that the transitions do not depend on the unobserved states. This simplifies the estimation insofar as the transition functions can be estimated separately from the other parameters of the model.

O-D.3. Depreciation of Housing Quality

This paper assigns housing quality to units based on its position in the local distribution of rents per square meter. That is, a unit's quality equals q if the unit's rent per square meter is the q-quantile of the ROR-level distribution. I use this rule to assign a quality level to each housing unit. To be consistent with the model, I assign each observation to one of ten quality bins.

Since many observations also include information on the address, the data allow identifying 'repeated rentals', by matching units based on the address, the floor, number of rooms, floor size, and presence of a balcony or terrace. I restrict the sample to matches with at least 12 months difference between the two offers. There are 175,962 such matched pairs in the data. The median (mean) time difference between two offers is 29 (33.5) months, and rents per square meter increased by 0.075 log points on average. The goal is to estimate pure depreciation, net of maintenance. I therefore restrict the sample further to pairs of units where observable characteristics of the unit (condition, fitted kitchen, flooring) remain unchanged, and to units with an initial quality above the lowest quality level. There are 94,706 pairs left in the sample, and the mean and median time differences are one month smaller, suggesting that landlords removed some units from the market temporarily for renovation works. Moreover, the rent change is only 0.062 log points, the difference of about 0.013 log points arguably representing the value of the alterations. The measure of housing quality $q_i \in \{1, ..., 10\}$ is defined by using the 10%-, ..., 90%-quantiles of the local rent distribution (per square meter) as breaks, which are measured in the full rent sample (by ROR and year).

In the model, the posited relationship between quality and time is log-linear. I therefore estimate the following equation:

$$\Delta \ln q_i = \delta \Delta \text{years}_i + \text{postcode}_i + \eta_i. \tag{O-D4}$$

For a unit i, Δyears_i is the difference in months between the two offers, divided by 12, and postcode_i is a postcode fixed effect that controls for gentrification effects (the up- or downward movement of a neighborhood's relative quality). η_i is an error term. Standard errors are clustered by ROR. δ is the quality decay factor. I restrict the sample to units that start at a quality level of 3 to 10.[39] Table O-D1 displays the estimation results.

Table O-D1 contains the results for the quality decay factor. Column (1) shows that rental housing quality decreases by 0.039 log-points per year, and the precision of the estimate is very high. This means that a unit in the highest quality bin ($q = 10$) has a quality of $q = 9$ after 2.5 years, and it reaches $q = 5$ after about 17.5 years. Column (2) tests whether the exponential discounting model is appropriate, finding that a second-order polynomial in the time difference does not yield a better fit.

O-D.4. Utility Parameters

Estimation Strategy. Estimation relies on the expectation-maximization algorithm developed in Arcidiacono and Miller (2011), exploiting one-period dependence via

[39]Obviously, units starting at $q = 1$ cannot depreciate further in this setting. At $q = 2$, the depreciation factor appears to be much lower (results available upon request). To keep the structure of the model simple, I focus on the depreciation factor that applies to the middle and top of the housing quality distribution (where it appears to be rather constant).

Table O-D1: Estimated housing quality decay factor

	Dependent variable: $\Delta \ln q$	
	OLS (1)	OLS (2)
Δ years	-0.039***	-0.047***
	(0.005)	(0.006)
Δ years squared $\times 10^{-3}$		1.153*
		(0.553)
Postcode FE	yes	yes
Adj. R^2	0.146	0.146
Observations	67,385	67,385

Notes: q is the discretized, normalized rank of the unit in the local (ROR) distribution of rents per square meters (values 1, ..., 10). The sample is restricted to units observed at least twice, without observable changes to unit characteristics, and all units are offered without renovation when observed the second time. The initial position in the rent distribution is above 2 and the time difference between two observations is at least 12 months. Standard errors (in parentheses) clustered by ROR; *: $p < .05$, **: $p < .01$, ***: $p < .001$.

the terminal choice $j = 11$. Let $v_{jt}(z_t)$ be the non-idiosyncratic component of the conditional value when choosing j in period t , being in state $z_t = (x_t, s)$. Using the mapping between the CCP's, $p_{jt}(z_t)$, and the conditional value functions, $v_{jt}(z_t)$, the difference between the baseline value and the conditional value of choosing $j' \in \{1, ..., 10\}$ is

$$v_{j't}(z_t;\theta) - v_{0,t}(z_t;\theta) = u_{j't}(z_t;\theta) - u_{0,t}(z_t;\theta) + \\ \beta \sum_{z_{t+1}=1}^{Z} \left[f_{j't}(z_{t+1}|z_t;\theta) - f_{0,t}(z_{t+1}|z_t;\theta)\right]\left[v_{11,t}(z_{t+1};\theta) - \ln\left(p_{11,t}(z_{t+1})\right)\right]. \quad \text{(O-D5)}$$

The first term on the right-hand side represents the difference in period-t flow utility. The discounted future value term has two components. The first bracket is the difference in the probability to reach z_{t+1} when being in initial state z_t and choosing $j = j'$, and the corresponding probability when choosing $j = 0$ in period t. For instance, moving to a different housing unit may change the size, match quality, housing type, and rent. The second term is the conditional value of choosing $j = 11$, a terminal choice, when being in state z_{t+1}, corrected for the fact that this choice is not necessarily optimal for the household. The correction factor is quantitatively large if the probability to choose $j = 11$ is small. Intuitively, the choice $j = 11$ does not have a very high value in this case, so that the true future value of being in state z_{t+1} is much higher than $v_{11,t}(z_{t+1})$. The conditional value function differences look very similar for the terminal choices $j' = 11, 12$, which are not stated explicitly in the interest of space.

Empirical CCP's. For a known $p_{11,t}(z_{t+1})$, the estimation problem is very similar to a standard multinomial logit model with unobserved heterogeneity, because all other functions are known up to the parameter vector θ (for a fixed discount factor β). I follow Arcidiacono and Miller (2011) and estimate $p_{11,t}(z_{t+1})$ as the empirical share of households choosing $j = 11$ when being in state z_{t+1}. In order to deal with the large state space, I employ proximity weighting when computing the empirical CCPs. I employ a generalized random forest (Athey et al., 2019) to calculate proximity weights. The random forest predicts the terminal choice $j = 11$ based on the observable covariates. The resulting weights are consistent estimators of the indicator function suggested in Arcidiacono and Miller (2011) because a forest trained on an infinite number of observations would grow each tree until all leaves only contain just one point of the (finite) state space. Hence, the forest-based weights preserve the consistency of the estimator.

Algorithm. The estimator proceeds iteratively. In iteration k, it starts from an initial set of parameters $\theta^{(k)}$, and values for the conditional probabilities of being in state s, denoted by $q_s^{(k)}$. In the expectation step of each iteration, $\theta^{(k)}$ and $q_s^{(k)}$ are used to calculate conditional and unconditional likelihoods of observed choices and state transitions, which determine $q_s^{(k+1)}$, the resulting unconditional probability ob being in state s, and the CCP's. The maximization step updates the parameter vector by maximizing the likelihood, taking the $q_s^{(k+1)}$'s and the CCP's as given. This procedure converges to a local maximum. Since the transition functions do not depend on unobserved states, they are not affected by the updating step and can be calculated by maximum likelihood prior to the execution of the algorithm.

I run the algorithm with different starting values, proceeding in the same order each time.

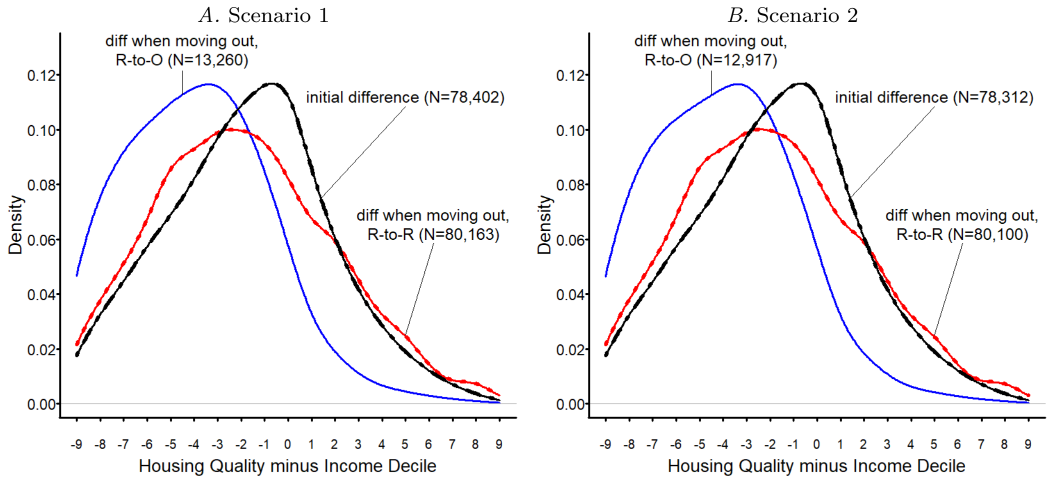
(1) Draw $q_s^{(0)}$ from independent distributions (normalized to $\sum_s q_s^{(0)} = 1$), for each household in the sample.

(2) Calculate $\pi_s^{(0)}$.

(3) Calculate $\psi^{(0)}(x_t, s_t)$, based on $q_s^{(0)}$ and the observed choices and state variables.

(4) Update the value function to get $\theta^{(0)}$, conditional on $\psi^{(0)}(x_t, s_t)$, $q_s^{(0)}$, and $\pi_s^{(0)}$.

These values serve as the starting point of the EM algorithm. I run the routine for ten different draws, with varying variances and means of the $q_s^{(0)}$'s.

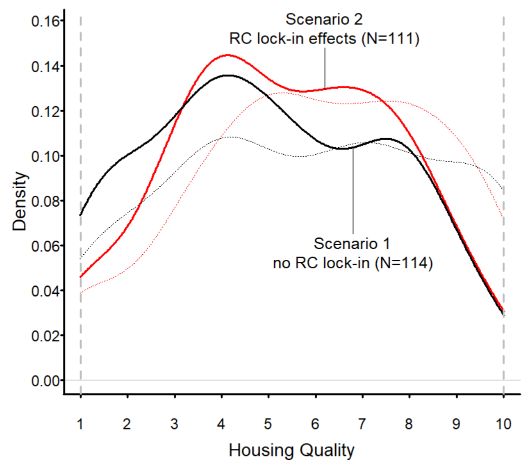


Notes: Both panels display densities of the difference between housing quality level and income class in the simulated data (Scenario 1: No lock-in; Scenario 2: lock-in due to rent control). The black line is the difference when moving in. The red (blue) line denotes the difference at the time of moving into another rental housing unit (owner-occupied housing).

Figure O-D5: Housing quality-to-income match in simulated data

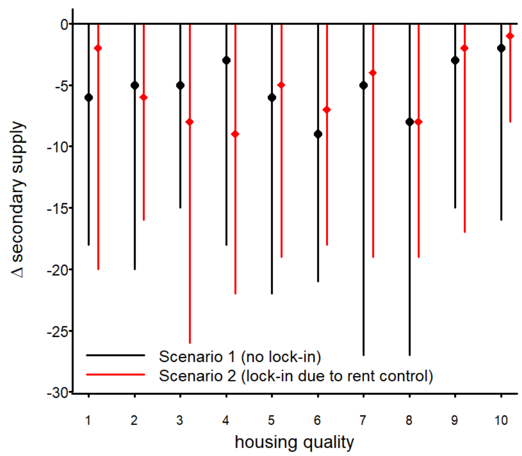
O-D.5. Simulation outcomes

Figure O-D5 displays differences between housing quality level and income class in the simulated data, see Section 3.3.2 in the main text. The corresponding graphs for the estimation sample are in Figure O-D4.



Notes: The figure displays the density of housing quality among renters moving into newly built owner-occupied housing in the simulated data (solid lines: including depreciation; dashed thin lines: net of depreciation). The black lines corresponds to Scenario 1 (no lock-in); the red lines corresponds to Scenario 2 (lock-in due to rent control).

Figure O-D6: Housing quality of renters moving into newly built housing in the simulated data



Notes: The graph displays the log change in secondary supply of rental housing due to the negative shock to new housing supply on the owner-occupier market, for Scenarios 1 (no lock-in) and 2 (lock-in due to rent control).

Figure O-D7: Simulated impact of a negative shock to new housing supply on log secondary supply of rental housing