Secondary Housing Supply*

This version: January 31, 2023; first version: January 10, 2020.

Andreas Mense

IAB Nuremberg & London School of Economics

Abstract

I estimate the impact of new housing supply on the local rent distribution, exploiting delays in housing completions caused by weather shocks during the construction phase. Increasing the flow of new supply by 1% lowers average rents by 0.2%, and increases disproportionately the number of second-hand units offered for rent. The supply shock affects the entire rent distribution. Employing a quantitative model, I explain this pattern by secondary supply: New supply triggers a cascade of moves that frees up units in all segments of the local market. The impact on rents is similarly strong in locations experiencing growing housing demand.

Keywords: housing supply, rental housing, rent distribution, secondary markets, market integration.

JEL classification: D15, D40, R21, R31.

*Andreas Mense: IAB Nuremberg, Regensburger Straße 100, 90478 Nuremberg, Germany, andreas.mense@iab.de, and Visiting Scholar at the Department of Geography and Environmant at LSE. I thank Gabriel Ahlfeldt, Rainald Borck, Thiess Büttner, Yongheng Deng, 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 UCLA/TAU/ESCP Affordable Housing Conference Tel Aviv 2022, the Potsdam Research Seminar in Economics, the Public Economics Research Seminar at CES/LMU Munich, and the brownbag seminar at University of Erlangen-Nuremberg for valuable comments and suggestions. This research was funded through grant ME 4963/1-1 by the German Research Foundation. I am grateful to LSE for its hospitality. All errors are my own.

1. Introduction

Housing markets are to a large extent secondary markets. In fact, by a huge margin, the majority of units traded in a given year are not new.[1] This paper addresses the question how a shock to new housing supply affects the distribution of rental prices in the primary and secondary local rental housing markets. Usually, second-hand units are of considerably lower quality — and may thus be poor substitutes for new housing. Although a lack of substitutability is a potential barrier to the propagation of such a shock, in secondary markets such as the housing market, substitutability is not a necessary condition for market integration across different market segments. The reason is that considerable adjustment costs prevent households from updating frequently their housing choices. As a consequence, many renters moving into new housing provide units of relatively low quality to the secondary market. Moreover, each move triggers a cascade of further moves that frees up additional second-hand housing units. Such cascades are central to market integration and to the propagation of shocks in the housing market.

The housing market is a particularly relevant example of a secondary market, but the core idea applies to other second-hand markets as well. 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.

In this paper, I consider the impact of new market-rate housing supply on the local distribution of private-market rents in Germany.[2] I exploit unusual weather conditions during the construction phase that cause considerable delays as an exogenous supply shifter, making use of a unique administrative data set comprising the universe of building completions in Germany between 2010 and 2017, in conjunction with data on rental

[1] According to the German data used in this study, 5.2% of the units offered for rent are newly built.

[2] The German homeownership rate is low by international standards — 45.7% according to the 2011 census. 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 also applies to the propagation of supply shocks inside the owner-occupier market.

housing units covering Germany as a whole from 2011 to 2018.

Long periods of rainfall during the summer, as well as unusually deep frost in February, reduce significantly the number of housing completions in November and December of the same year. The weather shocks affect all types of units, but the relationship is much stronger for single-family homes.[3] I document that the weather-induced delays have a long-lasting impact on the number of housing completions at the level of the local housing market, consistent with tight capacity constraints among housing developers during the most recent housing boom in Germany starting in 2010, and with evidence for the U.S. (Coulson and Richard, 1996; Fergus, 1999).

According to the baseline estimate, a 1% increase in yearly new housing supply causes the average local rent level to fall by 0.2%. This estimate does not vary much across housing unit types or local markets. First, there is no statistically significant difference between the impact on rents of high- versus low-quality units, as measured by the unit's position in the local rent/sqm distribution. Effects at the lower end are somewhat weaker, and they increase in magnitude towards the upper end, rainging from -0.14 to -0.29. Hence, new housing supply at market rates shifts the entire rent distribution to the left. Second, consistent with this result, the effect size varies only modestly with building age and housing unit size. The effects are slightly weaker for newly built and for units with two to three rooms. Overall, this pattern cannot be explained by substitution relationships between the new housing and units in the rental housing market. To the contrary, secondary supply triggered by the shock to new supply may explain well why the effect spreads speedily across the entire local market. Consistent with the secondary supply channel, the number of second-hand rental housing units that appear on the local market increases by 4.8 for every newly constructed housing unit.

From a policy perspective, local markets experiencing increasing housing demand are of particularly high relevance. The study period, 2011-2018, is well-suited to address

[3] About 25-30% of newly built single family homes are completed within 12 months after having obtained the building permit, and 58-65% within 18 months. The shares are substantially lower for multi-family homes (7 and 28%) (Schwarz, 2018). Weather shocks in a single year are arguably much less important for multi-year construction projects.

the question whether new supply is effective as a means to curbing rent growth in such high-demand markets. During this time, fueled by a robust economic development in Germany, with employment growing from 28.6 to 32.9 million persons, rental prices increased strongly in many locations. When restricting the sample to locations with above-median growth in employment, average gross labor income, and household income, respectively, the resulting estimates remain close to the baseline estimate of -0.2.

Arguably, the weather shocks affect rents only through the supply of new housing. One potential concern is that the instrumental variable picks up the long-lasting negative effects of local floods. I address this by showing that the baseline estimate is robust to excluding years with large flood events. Similarly, particular sectors such as tourism and agriculture could be directly affected by weather shocks. Yet, the baseline estimate is robust to controlling for housing demand factors that may correlate with the weather shocks. The weather – in particular, summer heat waves – might also affect behaviors on the housing market more directly.[4] However, the weather shocks are also uncorrelated with the pre-treatment outcome and with potential observable confounders prone to being affected directly by the weather, such as total work hours. Finally, I exploit the fact that February frost depth is almost orthogonal to the summer rainfall instrument, which makes it highly unlikely that the two variables share important unobserved confounders. The results are very similar when using either of the two instruments, and when using alternative defnitions of the rainfall instrument. Overall, these results lend strong support to the claim that the weather shocks are plausibly exogenous.

In the second part of the paper, I develop a structural model of a local housing market with 10×4 sub-segments representing combinations of housing quality and size. The purpose of the model is to investigate more deeply why rental prices for low-quality housing are affected swiftly by shocks to new housing supply, even if the new supply is catering mostly to owner-occupiers. The model characterizes both housing demand and

[4]Deng et al. (2021) show that temperatures above 32.2°C lead to a greater number of non-recourse mortgage defaults in the U.S., most likely because high temperatures affect the borrowers' home valuation. However, the argument does not apply to recourse loans, as common in Germany, and the number of hot days is much lower than in the US and exhibits much less regional variation, see https://www.umweltbundesamt.de/bild/anzahl-der-tage-einem-lufttemperatur-maximum-ueber.

secondary housing supply to the rental market. It is different from existing models in that movers in the market appear simultaneously on the demand side *and* on the supply side – the latter because they provide one vacant housing unit to the market. The secondary housing supply introduces strong cross-connections between different market segments that are absent in models where the supply side is either ignored or modeled from the perspective of a housing developer.

In the model, each renter moving into a newly built home triggers a series of adjustments across rental market segments 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, irrespective of the substitutability between particular segments.

The paper ties into the following strands of the literature: First, it adds to the growing empirical literatures on the impact of new housing supply on housing costs (Nathanson, 2019; Pennington, 2021; Mast, 2019) and filtering (Rosenthal, 2014, 2019). The most closely related papers are Pennington (2021)and Mast (2019) . Both papers focus on the effects of new housing supply on income-based sorting, gentrification, and housing costs at the level of the neighborhood. They do not, however, consider the aggregate effects of new housing supply at the level of the local or regional housing market, and they do not investigate the role of secondary housing supply.

Second, the results complement work studying housing choices of owner-occupiers and renters in the local housing market and the relationships between different market segments (Landvoigt et al., 2015; Piazzesi et al., 2020; Epple et al., 2020). Landvoigt et al. (2015) and Epple et al. (2020) develop a structural framework where households optimize housing and goods consumption without frictions in every period. Under standard assumptions, this implies perfect sorting of households by income into housing units ordered by quality. The dynamic framework proposed in this paper breaks up this perfect sorting and, in addition, models explicitly secondary supply. These two ingredients allow for a complex dependence structure across market segments that may help to explain high degrees of market integration, as observed, for example, across local markets in the U.S. prior to the Great Financial Crisis (Cotter et al., 2015).

Third, the paper is relevant for the large literature on the role of housing supply constraints for prices and rents, housing affordability, and the local housing market more generally (Büchler et al., 2019; Glaeser et al., 2005; Gyourko et al., 2013; Molloy et al., 2020; Hilber and Vermeulen, 2016; Hilber and Mense, 2021; Quigley and Raphael, 2004, 2005; Saks, 2008; Saiz, 2010; Van Nieuwerburgh and Weill, 2010). Most of this literature studies the impact of a given demand shock on housing prices in locations that differ in terms of their housing supply constraints. More recent work also considers the impact on housing rents, e.g., Büchler et al. (2019), Molloy et al. (2020) and Hilber and Mense (2021). Yet, the evidence from these papers regarding the effects of new housing supply on housing costs is only indirect. Moreover, it is less clear whether new supply at market rates is an effective means for achieving housing affordability for low-income households in locations experiencing strong demand pressure. Finally, while it is well understood that lack of housing supply has large effects on house prices, the impact of new housing supply to owner occupiers on the distribution of rental prices—in particular in the lower-quality segments—is less clear-cut, since credit constraints may represent a barrier between the two market segments (Ortalo-Magne and Rady, 2006).

The paper makes three main contributions: First, it provides estimates of rent price elasticities with respect to the flow of new housing supply. The preferred reduced-form estimate for the average effect is -0.2, suggesting that a 1% increase of new supply lowers rents by 0.2%.[5] This parameter in an important ingredient for quantitative regional economic models. It is also highly policy-relevant, since it helps local governments to understand how much average rental prices will decrease when issuing a higher number of building permits in a location. Moreover, changes in housing costs are an important component of consumer price inflation.

Second, to the best of my knowledge, this paper is the first to provide clean, quasi-ex-

[5] I corroborate the magnitude of this estimate using model-based simulations building on a structural dynamic housing choice model. The model-based elasticity is somewhat larger with -0.65 when the supply shock is to new owner-occupied housing, and -0.83 when the supply shock is to new rental housing.

perimental evidence on the connection between new housing supply and the distribution of rents in the local housing market as a whole. It documents that new housing supply effectively improves housing affordability of renters across the board, even in markets experiencing strong housing demand growth. This finding has significant implications for housing policy in general, suggesting that the focus should be on the supply side. This is especially important given that the rising housing costs in high-demand locations around the world have triggered various types of often distortionary and mostly ineffective demand-side policy responses (see Metcalf, 2018, for a recent survey).

Third, the paper proposes secondary housing supply as a key determinant of market integration between rental and owner-occupier markets. The degree of segmentation plays an important role in models of dual housing markets, e.g., Favilukis et al. (2017), Greenwald and Guren (2020), and Kaplan et al. (2020). In a nutshell, moving costs restrain households from making gradual adjustment of housing choices. This loosens the relationship between household income and housing quality, which in turn creates cross-connections between market segments and hence fosters market integration without requiring substitutability. In other words, it is not a necessary condition for market integration that there exist marginal buyers who are indifferent between buying or renting in different segments.

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 local rent distribution. Section 3 is devoted to the structural model, which is used to investigate the underlying mechanism. The final section draws conclusions and offers suggestions for policy and future research.

2. Reduced-Form Evidence: The Impact of New Housing Supply on Rents

2.1. Data

The administrative Building Completions Statistic reports information on all new housing units completed in Germany between 2010 and 2017, including municipality and month of completion.[6] Unfortunately, it is not possible to separate the supply of social housing from the supply of private-market housing in the empirical analysis. 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, a type of housing that rarely qualifies for subsidies in the German institutional setting.

According to the German Socio-Economic Panel (SOEP), 49.1% of new unsubsidized housing supply in Germany is absorbed by renters transitioning to owner-occupier status and 19.3% by former owner-occupiers. The remaining 31.6% are rental housing units.[8] 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 modal buyer of newly built owner-occupied housing in Germany is a renter.

The instrumental variables are 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.[9]

The rent data were collected from three large online real estate market places 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 first appearance, and a list of housing characteristics.

[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 does not provide information on subsidies.

[8] These numbers 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 in the SOEP. The Census 2011 reports very similar shares for housing built between 2009 and 2011, with 61% owner-occupied housing, and 39% rental housing (including subsidized housing).

[9] 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.

The outcome of interest is a log hedonic index based on the rent per square meter net of utilities and heating costs.[10]

Posted rents are advantageous in the present setting for several reasons. First, as long as there is no correlation between the measurement error when using posted instead of contractual rents and the instrument, the measurement error does not affect the estimate. Since the instrument is a lagged, weather-based instrument, this seems highly unlikely. Second, surveyed rents may be less precise than posted rents to the extent that households have difficulties to determine their *net rent*, as opposed to their total costs for shelter including heating and other services.[11]

In Germany, households typically pay the gross rent including heating services (consisting of net rent, property services, utilities, and heating). The different rent components and the floor size are posted separately in rental housing offers, whereby measurement is regulated by German bylaw.[12]. This increases the reliability as compared to information from surveys. Finally, posted rents are available on a small geographic scale and with detailed housing characteristics – which is not the case for surveyed rents.

The main analysis is conducted at the level of local housing markets, using German planning regions (PR) [*Raumordnungsregionen*]. Housing units and weather shocks are assigned to PRs based on their geocodes and the municipality identifier. For each PR, I employ ordinary and quantile hedonic regressions to compute quality adjusted local rent indices. The resulting panel is balanced and covers 94 PRs over eight years.[13] I merge these data with additional control variables capturing important determinants of local housing demand from the INKAR database of BBSR. Table 1 provides summary statistics for the PR panel data.

[10]The housing completions and rent data and the hedonic rent index are described in greater detail in Appendix O-A.

[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*].

[13]In total, there are 96 PRs, but for Bremen and Saar, the month of completion is not available. Since this is a key variable in the empirical strategy, I exclude these two PRs from the analysis throughout.

Table 1: Descriptive statistics for the PR panel (N=94, T=8)

	Min	Mean	Q25	Median	Q75	Max
A. Rents and hedonic rent indices, 2011–2018						
Real monthly rent/sqm (index-based, 2018 EUR)	4.33	7.64	6.15	7.23	8.66	26.17
Log mean real rent index (2011 = 0)	-0.045	0.059	0.003	0.043	0.095	0.322
Log real rent index 1st decile (2011 = 0)	-0.164	0.042	0.000	0.027	0.069	0.321
Log real rent index 3rd decile (2011 = 0)	-0.034	0.055	0.001	0.038	0.091	0.344
Log real rent index 5th decile (2011 = 0)	-0.041	0.063	0.002	0.046	0.105	0.379
Log real rent index 7th decile (2011 = 0)	-0.044	0.073	0.006	0.053	0.118	0.395
Log real rent index 9th decile (2011 = 0)	-0.082	0.086	0.007	0.063	0.140	0.516
B. New housing completions and weather shocks, 2010–2017						
New supply in Nov+Dec per yearly avg. # of newbuilds	0.040	0.394	0.270	0.350	0.482	1.051
Log new supply (whole year)	4.70	7.28	6.78	7.28	7.79	9.29
Avgerage summer rainfall spell (deviation)	-6.355	-0.001	-1.247	0.022	1.258	6.171
Feb. frost depth (deviation)	-13.900	0.001	-5.040	-2.853	0.596	41.304
Longest rainfall spell (deviation)	-3.279	0.000	-0.735	-0.086	0.536	4.320
Number of spells w/ 5+ days (deviation)	-1.113	0.003	-0.280	-0.057	0.280	1.825
C. Control variables in year of weather shock, 2010–2017						
Employment (1,000's)	62	311	155	214	356	1,426
Unemployment rate	0.021	0.065	0.040	0.060	0.083	0.148
U & college students per 1,000 residents	0.0	27.1	11.1	26.9	38.6	100.0
Share w/o school degree	0.028	0.062	0.046	0.056	0.074	0.159
Hours worked per worker in year	1,252	1,336	1,304	1,320	1,355	1,680
Gross average labor income	1,765	2,488	2,243	2,444	2,690	3,745
Dummy: Heavy flood in federal state	0.000	0.114	0.000	0.000	0.000	1.000

Notes: The real monthly rent per sqm is based on the average rent per sqm as observed in 2011 and the real average rent index, deflated by the CPI (2018=1). The rent indices are constant-quality hedonic indices, see Appendix O-A for details. Control variables are taken from the INKAR regional data base. Data on hours worked is not available for four PRs (1601, 1602, 1603, 1604) in the years 2010–2013. The share withouth school degree is the share of pupils leaving school without a school degree. The heavy flood dummy captures years with severe floods in the federal state the planning region belongs to (2013: Lower Saxony, Hesse, Rheinland-Palatinate, Baden-Wuerttemberg, Bavaria, Saxony, Saxony-Anhalt, Thuringia; 2017: Lower Saxony, Saxony-Anhalt, Thuringia).

2.2. Weather shocks as instrument for new housing supply

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. 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.[14] 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 evening without electric light. To the contrary, on a rainy day, "effective daytime" is much shorter and workers might be less motivated. Third, concrete, bonding agents, and certain other materials cannot be applied when there is heavy rainfall or rainfall continuing over multiple days.[15]

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. Therefore, most construction work pauses during wintertime.[16]

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 last much longer if capacity constraints in the construction sector are binding, preventing developers from catching up in the next year.

Definitions of the instrumental variables

I use four instruments in the regressions that build on these considerations. The main instrument is the average longest spell of consecutive rainfall days ($>$ 20 mm per sqm) in each summer month. I use two alternative definitions of the rainfall instrument, the longest overall spell, and the number of spells with at least five days of consecutive rainfall between July and September.

The fourth instrumental variable is 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 regions (e.g., in the Rhine-Main and coastal areas). However, frost depth in February is unrelated to summer rainfall, and

[14]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.

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

[16]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).

hence provides a source of variation that is orthogonal to the rainfall instrument.

The rainfall shocks are constructed from daily rainfall data on a 1×1 km^2 grid. For the main instrument, I compute the largest number of consecutive days with rainfall above 20mm per square meter by grid cell and month (July, August, September), which I refer to as a "rainfall spell". To remove time-constant differences in weather across locations, I subtract the grid cell mean of the particular calendar month. Hence, the identifying variation comes from weather conditions that deviate from the usual conditions at the location. The final step is to aggregate to the PR and year. Figure O-B1 shows that the instrument exhibits substantial spatio-temporal variation.

February frost depth is also provided for 1×1 km^2 grid cells by the German Weather Service. The three alternative instrumental variables are defined in an analogous fashion.

First-stage relationship

Table 2 summarizes the results from a set of regressions with three different summer rainfall variables and February frost depth as the explanatory variables. In this table, all instruments are scaled to have a standard deviation of one and a mean of zero. The unit of observation is the municipality by year. In columns (1) to (5), the dependent variable is the number of new housing units completed in November and December relative to the yearly average number of newly built housing units. When using the longest average summer rainfall spells during the summer months in column (1), the coefficient is highly significant and negative, with an F-statistic of 44.0. However, 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 standard deviation reduces new housing supply in the given year by about 1.93%. Nonetheless, it provides very useful instrumental variable variation, and, beyond the instrument's relevance, the quantitative magnitude of the first-stage relationship is not important.

The two other variants of the rainfall instrument yield comparable results, albeit with lower F-statistics. Deeper frost depth in February also reduces the number of units completed end-of-year, as shown in column (4). When adding the average summer rainfall spell and frost depth jointly in column (5), both coefficients are significant and stable, arguably due to the very low correlation between the two instruments at municipality level of 0.09.

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 have a much smaller influence in these cases. In column (6), the dependent variable is the number of units in multi-family buildings completed in November and December, again as a share of the average yearly supply. Although the signs of the instruments do not change, both instruments have a much smaller impact than in columns (1) and (5) and are less significant, lending support to the hypothesis that larger buildings are less strongly affected by the weather shocks. Hence, the weather-induced supply shock is mainly a shock to the supply of single-family housing.[17]

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

Dependent variable:	New housing units completed Nov+Dec as share of average yearly supply					
	in all types of buildings					in MFH's
	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS
Avg. summer rainfall spell (deviation from local average)	-0.0193*** (0.0029)				-0.0205*** (0.0029)	-0.0004** (0.0002)
Longest summer rainfall spell (deviation from local average)		-0.0095*** (0.0030)				
# of rainfall spells 5+ days (deviation from local average)			-0.0172*** (0.0030)			
Frost depth in February (deviation from local average)				-0.0194*** (0.0058)	-0.0236*** (0.0058)	-0.0009** (0.0004)
Year FE	yes	yes	yes	yes	yes	yes
Municipality FE	yes	yes	yes	yes	yes	yes
F statistic (proj. model)	44.0	9.9	33.3	11.3	30.4	4.5
Observations	83,632	83,632	83,632	83,632	83,632	83,632

Notes: Standard errors are clustered by municipality; *: $p < .1$, **: $p < .05$, ***: $p < .01$. In columns (1) to (5), the dependent variable is the number of housing units completed in November and December as share of average yearly supply in the municipality. In column (6), the dependent variable is the number of housing units in multi-family housing completed in December, as share of average yearly supply in the municipality. The explanatory variables are scaled to a mean of zero and a standard deviation of one.

During housing booms, when the construction sector operates near its maximum

[17] Figure O-B2 displays estimates for the imapact of the rainfall and frost instruments on housing completions in each month separately, analogous to column (5) of Table 2. It shows that there is virtually no effect between January and September, but both instruments marginally increase the number of completions in October. This is consistent with developers shifting attention away from projects affected adversely by a weather in July to September, to projects that are almost finished and may safely be completed before the end of the construction season despite the poor weather conditions. Almost-finished projects likely already have a closed building hull and require mostly inside work.

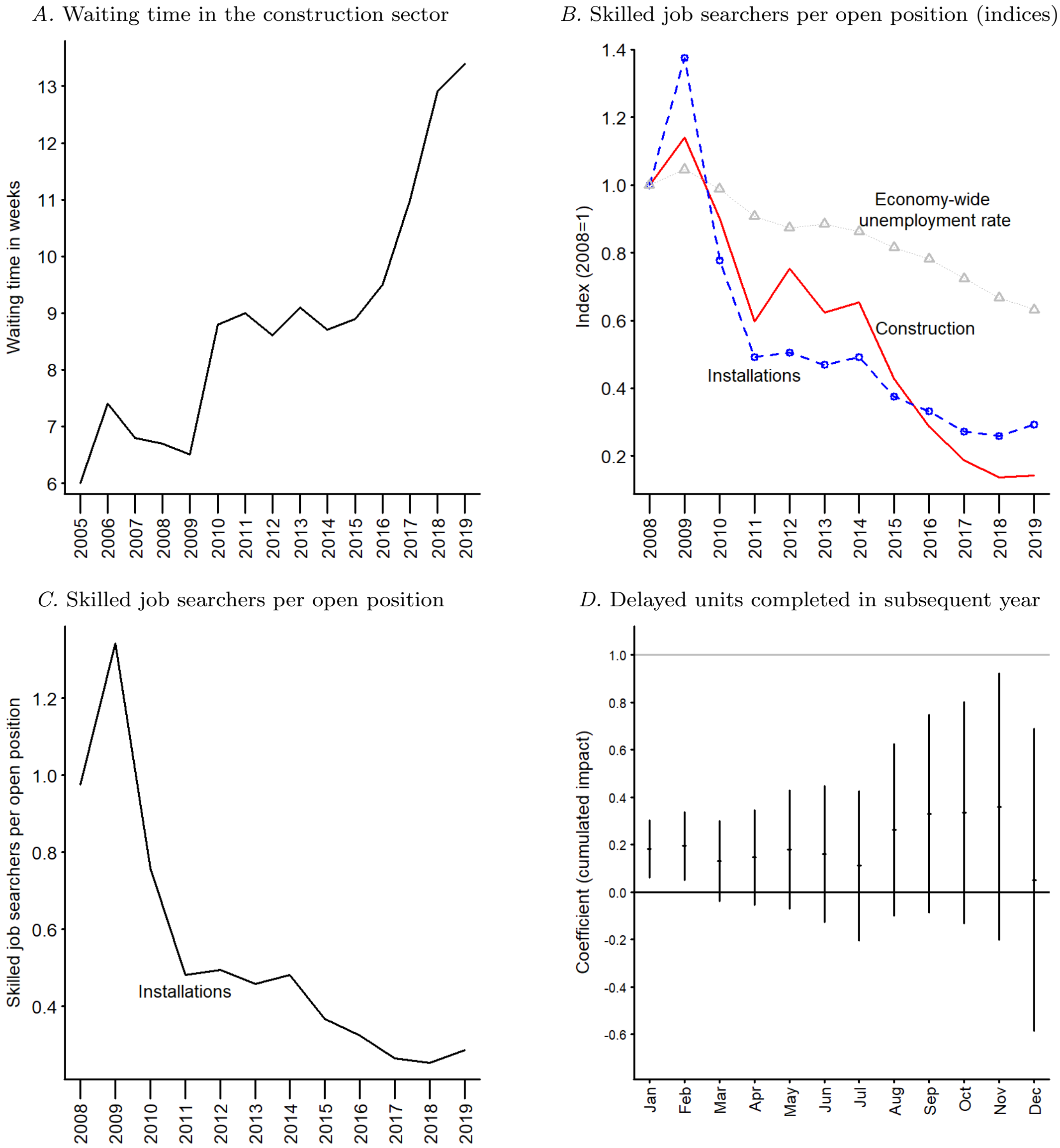
capacity, temporary reductions of construction volumes may lead to a quasi-permanent reduction of housing supply. This characterizes very well the situation in Germany since the start of the latest boom in 2010. Waiting times for construction firms (time between signing a contract and the start of its execution) more than doubled, from 6.5 weeks in 2009 to 13.4 weeks in 2019, and never decreased markedly 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 weather-induced delays, 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 not by much. The catching-up is never above 40%, and it falls close to zero when considering the whole year. This strongly suggests that further projects get delayed as the initially-delayed projects get completed, consistent with the construction industry working at the capacity limit. 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).

IV balance

Figure A1 summarizes a series of balancing tests that scrutinize the assumption that the local rental housing market is affected by summer rainfall only through its impact on new housing supply. The figure displays coefficient estimates of the rainfall shock

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



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) with 90% confidence intervals; standard errors clustered by municipality.

instrument along with 95% confidence intervals for a series of panel FE regressions using different standardized variables as outcomes, where the unit of observation is the PR by year. The first two coefficients represent the reduced-form and first-stage relationships. Longer summer rainfall spells decrease housing completions in November and December at the level of PRs (first stage) and increase the local hedonic rent index in the subsequent year (reduced form).

However, there is virtually no relationship between the summer rainfall shock and the hedonic rent index in the year of the rainfall shock. 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 correlate with broader trends in local housing demand or supply.

There is also no statistically significant relationship between the instrument and typical shifters of local housing demand, captured here by log employment and the log unemployment rate, the share of university and college students at the location, log GDP per capita, log gross labor income, and log household income, despite the fact that most of these estimates are relatively precise. Moreover, these coefficients are small relative to the reduced-form and first-stage relationships.

2.3. Estimation results

Baseline effects on average rents

I start by studying the impact of new housing supply on average local rents in panel IV-FE regressions at the level of PRs, with the hedonic rent index as the dependent variable. The housing completions in November and Decemberof the preceding year as a share of the average new housing supply is instrumented by the summer rainfall shock. PRs 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

$$\ln \text{Index}_{rt} = \gamma \left[\frac{S_{r,t-1}^{\text{Nov, Dec}}}{H_r} \right] + \psi_r + \phi_t + x'_{r,t-1}\beta + \varepsilon_{rt}, \tag{1}$$

where Index_{rt} is a hedonic rent index of planning region 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 average number of units supplied per year in r, and ψ_r and ϕ_t denote PR- and year-fixed effects. $x_{r,t-1}$ are control variables at the PR×year level that capture important determinants of housing demand. In the baseline regression, these are log employment, the log unemployment rate, and the number of university and college students per capita, all measured in the year of the rainfall shock. The latter group is likely to rent, and represents an important demand factor in many mid-sized cities. Employment opportunities attract demand for housing in the PR, and unemployed persons are restricted in their housing demand. Standard errors are clustered at the PR level.

Panel *A* of Table 3 displays the results. Column (1) includes as controls the log employment and the fixed effects only. The coefficient of main interest is both highly significant and negative. It suggests that a 1%-increase in yearly new supply lowers rents by about 0.2%, hence a rent price elasticity with respect to the flow of new housing supply of -0.2. Adding the log unemployment rate in column (2) and the share of university and college students in column (3) hardly affects this estimate, and the Kleibergen-Paap F statistics of all three regressions do not indicate weak instrument problems. The first-stage relationships are summarized in Panel *B*.

Robustness of baseline results

The identification strategy relies on variation in weather that is arguably exogenous to the state of the local housing market. Even though the local housing market clearly cannot affect the weather in the previous summer, weather may affect the local economy in ways that could, in theory, introduce a spurious correlation between the weather and the local rent level — despite the fact that the agricultural and tourism sectors in Germany are rather small, and even if most industries in Germany are unaffected by summer rainfall. I therefore test more rigourously how such and other potential confounders affect the baseline estimate.

First, severe weather conditions during the summer may lead to floods that have lasting effects on the local economy. In column (1) of Table A1, I control for federal

Table 3: Impact of new housing supply on average rents

A. Second Stage			
Dependent variable:	Log hedonic rent index		
	(1) IV	(2) IV	(3) IV
Units completed Nov + Dec in $t-1$ per avg. # units completed annually	-0.207*** (0.077)	-0.216*** (0.078)	-0.199*** (0.068)
Log employment, year $t-1$	1.042*** (0.146)	1.079*** (0.152)	0.997*** (0.144)
Log unemployment rate, year $t-1$		-0.053 (0.043)	-0.079* (0.045)
U & college students per 1,000 inh., year $t-1$			0.003** (0.001)
Year FE	yes	yes	yes
PR FE	yes	yes	yes
Kleibergen-Paap F	16.7	16.1	18.4
Number of PRs	94	94	94
Observations	752	752	752

B. First Stage			
Dependent variable:	Units completed Nov + Dec in $t-1$ per avg. # units completed annually		
	(1) OLS	(2) OLS	(3) OLS
Rainfall spell instrument (avg. length, Jul-Sep of year t-1)	-0.009*** (0.002)	-0.009*** (0.002)	-0.009*** (0.002)
Log employment, year $t-1$	0.880*** (0.307)	0.951*** (0.339)	0.730** (0.314)
Log unemployment rate, year $t-1$		-0.132 (0.124)	-0.228* (0.128)
U & college students per 1,000 inh., year $t-1$			0.009*** (0.003)
Year FE	yes	yes	yes
PR FE	yes	yes	yes
Adj. R squared	0.845	0.839	0.856
Number of PRs	94	94	94
Observations	752	752	752

Notes: Standard errors are clustered by PR; *: $p < .1$, **: $p < .05$, ***: $p < .01$. The instrument in columns (1) and (3) is the rainfall shock in year $t-1$. Columns (2) and (4) show the respective first stage regressions.

state-years with severe floods by using a dummy variable. In column (2), I exclude all observations for which this dummy is equal to one. The estimates are highly robust in both cases.

Second, the weather shocks could be spuriously correlated with determinants of local housing demand not already included in the baseline regression. I test this conjecture by adding further control variables to the baseline regression in Table A2. In particular, I add controls for the share of adult residents witout a school degree in column (1), and for the log number of hours worked and log gross labor income in column (2). Work hours and labor income could be affected by rainfall if rainfall-dependent sectors ask workers to reduce work hours in rainy years, or if workers choose to work more in rainy years as rainfall reduces the value of leisure time. Column (3) adds contemporaneous and lagged demand factors jointly. Due to the year and location fixed effects already included in the regression, these controls capture trends and changes in these trends in the local economy. Importantly, the main estimate remains very stable in all cases.

Third, instead of adding observable potential confounders as controls, I consider three alternative weather instruments. This addresses the concern that *unobserved* determinants of the rental price may be spuriously correlated with weather conditions in the previous summer, even if rainfall during the summer does not affect significantly the local economy as measured by important *observable* variables. The first and second alternative instruments are based on summer rainfall, but use a different definition for the rainfall shock. Columns (1) and (2) of Table A3 display the corresponding results. The Kleibergen-Paap F is lower in both cases, but the coefficient of main interest remains very stable, strongly suggesting that the functional form of the summer rainfall shock is not driving the results. The third alternative instrument is the frost depth in February, which is almost uncorrelated with the summer rainfall instrument (PR-level correlation of 0.117), working through a different mechanism: Rather than affecting construction work during the summer, frost depth delays starting dates at the beginning of the year. Hence, the two instruments likely do not share common unobserved confounders. In particular, concerns that summer rainfall may affect business and worker behavior do not apply to the frost depth instrument used in column (3). This instrument — albeit almost orthogonal to the main instrument — leads to a very similar point estimate of -0.257. When using the summer rainfall spell and the frost depth instruments jointly in column (4), the coefficient is again very close to the baseline estimate.

Fourth, the interpretation of the effect size is complicated by the fact that the weather shocks affect housing completions in particular months of the year only. Therefore, Table A4 displays results for a specification more akin to the standard approach for estimating elsticities, employing the log overall number of housing completions in year $t-1$ as the endogenous variable. Despite the lower Kleibergen-Paap F statistic of 10.0, the coefficient is virtually identical to that obtained with the baseline regression, lending support to the interpretation an (inverse) rental demand elasticitiy.

Fifth, one might be concerned that the fixed-effects specification is not adquately capturing the effect of new supply on the *change* in rent levels. Two alternative specifications are a regression of the change in the log rent index on new supply (i.e., the change in the housing stock), and on the change of new supply (i.e., the change in the flow). Table A5 shows the results. Both regressions include the changes in the baseline control variables in addition to year and location fixed effects, the latter capturing average location- and year-specific changes in rental prices. Despite these much more demanding controls, the main coeffficient is robust in both cases.

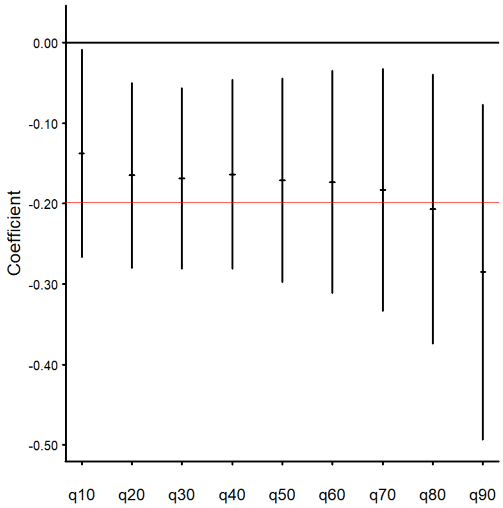
Finally, I consider two alterantive spatial delineations of the local housing market. In Column (1) of Table A6, local housing markets are defined as commuting zones, using the delination of BBSR based on commuter flows between German districts. All variables are at this level of aggregation in this specification. The coeffient is statistically significant but smaller, with -0.124. It is almost identical when instead using districts as the spatial unit in column (2). Overall, a smaller effect relative to the baseline estimate is consistent with the conjecture that the supply shock induced spillovers across smaller areas within the larger PR, hence a smaller measurable effect when using smaller geographies.

Overall, these results lend strong support to the assumption that the weather instrument is exogenous to local economic conditions and other determinants of local housing rents. In the next sections, I consider effect heterogeneity in three important dimenions: along the local rent distsribution, by building age, and across local housing markets.

Effects on the local rent distribution

This section addresses the question to what extent new housing supply affects the tails of the local rent distribution. To this end, I replace the hedonic index in equation (1) that captures average conditional rents, by conditional rent quantile indices. The quantile indices are estimated from hedonic quantile regressions, and are hence quality-adjusted. Details are given in Appendix O-A. The estimating equation is otherwise identical to the one defined in equation (1) and used in Table 3, column (3).

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



Note: The figure displays coefficient estimates for equation (1), using indices for the conditional quantile of the local rent/sqm distribution (constant-quality) as outcome. The housing completions in November/December are instrumented by the rainfall shock. Vertical bars represent cluster-robust 95% confidence intervals.

Figure 2 displays the impact of the housing supply shock on the first to ninth decile of the PR-level rent distribution. The red horizontal line shows the impact on average rents reported in column (3) of Table 3. All coefficients are negative and significant at least at the 5% level, with a slightly stronger impact at the top of the distribution. However, this variation is not large, ranging from -0.138 at the first decile to -0.285 at the ninth decile. Overall, these results suggest that integration between the market for new (single-family) homes and all quality segments of the rental market is relatively tight.

Heterogeneity across housing units

The housing completions data do not provide information about whether units are going to be rented out owner-occupied. Although the instrument mainly picks up variation in single-family housing completions, there could be a direct effect on rental prices for new units. Moreover, larger units could be affected more strongly if large rental housing units are close substitutes to newly built single-family housing.

Table 4 displays estimates of the impact of new supply on rents by age class, with building age defined as year of construction minus year of observation.[18] The baseline

[18]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.

sample in column (1) is the full sample of rental units used to construct the hedonic indices. The regression controls for the set of housing characteristics employed in the PR-level hedonic regressions used to construct the local rent indices. Moreover, it includes the same set of PR-level controls as in the baseline regression. Conceptually, the main differences to that regression are, first, the implicit weighting of each local housing market by the number of observations, and, second, the fact that coefficients of the housing characteristics controls are not location-specific in Table 4, while they vary by PR when constructing the hedonic indices. Apparently, both of these differences have only little impact on the coefficient estimate, despite the lower Kleibergen-Paap F statistic.

In column (2), the regression is weighted by the inverse of the local housing market's size, as captured by the number of units in the housing stock in 2011, in order to make the regressio more comparable to the baseline regression. Column (3) excludes units lacking information on the year of construction. The coefficient is somewhat smaller in magnitude, but it retains significance. Columns (4) to (7) consider different building age brackets. The impact on rents for newly built units shown in column (4) is negative, but not significant and much smaller in magnitude than the baseline estimate, showing that the bulk of the effect does not come from a direct supply effect of rental housing development. The effects for older buildings are somwhat stronger and (marginally) significant in all cases.

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	(7) IV
Age class (years)	any/NA	any/NA	any	0	1–10	11–50	51+
Units completed Nov + Dec in $t-1$	-0.185***	-0.159**	-0.123**	-0.137	-0.171*	-0.146**	-0.184*
per yearly avg. # of new units	(0.066)	(0.067)	(0.059)	(0.121)	(0.102)	(0.066)	(0.100)
Year FE	yes	yes	yes	yes	yes	yes	yes
PR FE	yes	yes	yes	yes	yes	yes	yes
Other controls	yes	yes	yes	yes	yes	yes	yes
Kleibergen-Paap F	8.1	12.7	9.7	14.5	10.9	13.3	4.1
Number of PRs	94	94	94	94	94	94	94
Observations	6,926,371	6,926,371	4,693,150	360,387	394,489	2,142,367	1,795,901

Note: Standard errors are clustered by PR; *: $p < .1$, **: $p < .05$, ***: $p < .01$. All regressions control for housing characteristics, location and year fixed effects, and the controls used in Table 3. The endogenous dependent variable is the number of housing units completed in the preceding November/December as a share of the average yearly supply of new housing, instrumented by the summer rainfall shock. Column (1) reports results for the entire unweighted sample. Regressions in all other columns are weighted using the inverse size of the housing stock, in order to achieve comparability with the panel IV regressions summarized in Table 3. Columns (3)–(7) exclude units with missing information on the year of construction. In columns (4)–(7), the sample is partitioned by building age (year of observation minus year of construction).

Table 5 explores effect heterogeneity by unit size – arguably, the single most important component of overall housing quality. Using the same approach as before, I partition the sample into four groups by total number of rooms (bedrooms plus other rooms). As columns (1) to (4) show, the effects are significantly negative for all four groups, but strongest for very small (one room) and very large rental housing units (four or more rooms). Although a tight substitution relationship between owner-occupied single-family housing and large rental units and may explain the greater impact on the latter, the overall pattern cannot be explained by substitution relationships between unit types alone.

Table 5: Effect heterogeneity by unit size

Dependent variable:	Log rent per sqm			
	(1) IV	(2) IV	(3) IV	(4) IV
Number of rooms	1	2	3	4+
Units completed Nov + Dec in $t-1$ per avg. # units completed annually	-0.228** (0.114)	-0.108* (0.064)	-0.134** (0.064)	-0.211** (0.095)
Year FE	yes	yes	yes	yes
PR FE	yes	yes	yes	yes
Other controls	yes	yes	yes	yes
Kleibergen-Paap F	13.7	11.1	12.0	14.7
Number of PRs	94	94	94	94
Observations	872,904	2,591,727	2,502,667	959,073

Note: Standard errors are clustered by PR; *: $p < .1$, **: $p < .05$, ***: $p < .01$. All regressions control for housing characteristics, location and year fixed effects, and the variables displayed in Table 3. The endogenous dependent variable is the number of housing units completed in the preceding November/December as a share of the average yearly supply of new housing, instrumented by the summer rainfall shock. All regressions are weighted by the inverse size of the local housing stock to achieve comparability with the panel IV regressions.

Impact in markets with increasing housing demand

A particularly policy-relevant question is whether new housing supply can effectively curb rent increases in markets experiencing sustained demand growth. Therefore, this section considers PRs with above-median demand growth during the sample period, as captured by the long-difference (2011 to 2018) in log employment at workplace, log average labor income, and log household income. The German economy was in a sustained boom phase during this time, with a median PR-level change in log employment of 0.14 from 2011 to 2018. Table 6 reports the results for the high-demand PRs using the baseline specification.

As shown in column (1), in PRs with a strong positive trend in log employment, the impact of the supply expansion on rents is still significantly negative and of a similar magnitude as in the baseline regression. This also holds for locations with strong growth in average gross labor income in column (2), and average household income in column

Table 6: Effect of new supply in markets with increasing housing demand

Dependent variable:	Log rent per sqm		
	(1) IV	(2) IV	(3) IV
Sample restricted to locations w/ above-median growth of	employment	avg. gross labor income	avg. household income
Units completed Nov + Dec in $t-1$ per avg. # units completed annually	-0.162**	-0.185*	-0.299**
	(0.062)	(0.108)	(0.118)
Year FE	yes	yes	yes
PR FE	yes	yes	yes
Other controls	yes	yes	yes
Kleibergen-Paap F	22.7	9.7	13.4
Number of PRs	47	47	47
Observations	376	376	376

Note: Standard errors are clustered by PR; *: $p < .1$, **: $p < .05$, ***: $p < .01$. All regressions control for location and year fixed effects, and the controls used in Table 3. The endogenous dependent variable is the number of housing units completed in the preceding November/December as a share of the average yearly supply of new housing, instrumented by the summer rainfall shock. The functional form for all three regressions is identical to that of the baseline regression, column (3) of Table 3.

(3), where the latter regression produces a larger effect with -0.299. The overall picture suggests that the estimated rent price elasticity does not shrink in markets experiencing strong demand growth. Overall, expanding housing supply is a very effective means for achieving housing affordability in markets with surging housing demand.

2.4. Impact on the quantity of rental housing traded in the market

The evidence provided so far is consistent with the idea that secondary supply works as a transmission channel that spreads the shock to new supply throughout the rental market. If this is the case, the newly built units should trigger a series of moves in the rental market, implying an effect on the quantity of housing offered for rent.

To test this conjecture, I compute the number of units offered for rent by PR and year from the rent data, which cover the private German rental housing market almost completely, and run the baseline regression with the number of rental units as a share of the average yearly supply of new housing as the outcome. Results are reported in Table 7. In column (1), the coefficient is positive, but it is not significant at conventional levels of confidence. When considering new rental units in column (2), the coefficient is close to zero, lending further support to the conjecture that the instrument affects mainly newly built owner-occupied rather than rental housing. Columns (3) and (4) restrict the outcome to second-hand rental units, which yields a slightly larger effect. It is 3.95 when using the baseline controls. When adding the log gross labor income as further control to capture the impact of income changes on the propensity to move house, the coefficient increases slightly and becomes marginally significant. In terms of magnitude, this regression suggests that one newly supplied housing unit triggers about 4.75 moves in the rental housing market in the subsequent 12 months.

Table 7: Effect of new supply in markets with increasing housing demand

Dependent variable:	# units offered for rent in t per avg. # of units completed annually			
	(1) IV	(2) IV	(3) IV	(4) IV
Rental unit types	all	new only	existing only	existing only
Units completed Nov + Dec in $t-1$ per avg. # of units completed annually	3.646	-0.308	3.954	4.752*
	(2.818)	(0.457)	(2.678)	(2.787)
Log employment, year $t-1$	-31.058***	-1.916	-29.142***	-27.903***
	(6.412)	(2.018)	(5.784)	(5.709)
U & college students per 1,000 inh., year $t-1$	-0.014	0.012***	-0.026	-0.021
	(0.040)	(0.004)	(0.039)	(0.040)
Log unemployment rate, year $t-1$	-2.724	-0.264	-2.460	-0.317
	(2.512)	(0.367)	(2.449)	(2.618)
Log avg. gross labor income, year $t-1$				20.446**
				(8.751)
Year FE	yes	yes	yes	yes
PR FE	yes	yes	yes	yes
Kleibergen-Paap F	18.4	18.4	18.4	17.1
Number of PRs	94	94	94	94
Observations	752	752	752	752

Note: Standard errors are clustered by PR; *: $p < .1$, **: $p < .05$, ***: $p < .01$. All regressions control for location and year fixed effects. The endogenous dependent variable is the number of housing units completed in the preceding November/December as a share of the average yearly supply of new housing, instrumented by the summer rainfall shock. The outcome variable is the number of housing units offered for rent in the subsequent year as a share of the average yearly supply of new housing.

3. Quantitative Secondary Supply Model of a Rental Housing Market

This section develops and estimates a secondary supply model of a local rental housing market to investigate further the channels through which new supply affects the rent distribution. The main feature distinguishing this model from existing ones is the explicit treatment of secondary supply. In the model, renters determine the demand for rental housing, but they also contribute to the supply of rental housing when moving house.

3.1. Secondary supply model: setting and definitions

Dynamic discrete choice model of housing quality and tenure choice

Setting. The main building block of the secondary supply model is a dynamic discrete choice model in discrete time that features moving costs. The choice model determines aggregate demand for and secondary supply of rental housing.

Choice set. In each period, the household faces a set of $J = 44$ mutually exclusive alternatives $j = 0, ..., 43$. The baseline choice $j = 0$ is to stay in the current accommodation. Rental housing units differ by quality $q \in \{1, ..., 10\}$ and number of rooms $s \in \{1, 2, 3, 4+\}$. Quality q is measured as the normalized rank in the local distribution of rent per square meter, binned into deciles. This definition is akin to that in Landvoigt et al. (2015) and Epple et al. (2020) and does not involve value judgements regarding the attributes of the unit including neighborhood characteristics, some of which are unobserved. In contrast to Landvoigt et al. (2015) and Epple et al. (2020), it allows for separate valuation of quality and size. Moreover, it is consistent with the reduced-form analysis. Choices $j \in \{1, ..., 40\}$ correspond to moving into a rental housing unit with quality and size (q, s).

Households may buy and self-occupy an existing ($j = 41$) or a new housing unit ($j = 42$), or leave the local housing market ($j = 43$). The subsequent choice path following one of these three choices is not modeled explicitly, i.e., these choices are terminal. This simplifies considerably the estimation, but it does not interfere with the purpose of the model, namely to determine preferences that shape demand and secondary supply in the rental market. Moreover, the lifetime utilities associated with these choices capture the possibility that the household becomes a renter again in the future.

State space. Households are characterized by a set of observables $x_t = (r_t, q_t, s_t, \tau_t, y_t, w_t, a_t, m_t, k_t, a_t^k, (r_t^q)_{q=1,..,10})$ and an unobserved type $z \in \{1, .., 8\}$. The observables are the net rent r_t, quality $q_t \in \{1, .., 10\}$ and size $s_t \in \{1, .., 4\}$ of the apartment currently occupied, as well as the length of tenure τ_t. y_t is household income net of taxes and social security contributions, w_t is financial wealth[19], and a_t is the age of the household head. $m_t \in \{0, 1\}$ is an indicator for a couple household, $k_t \in \{0, 1, 2\}$ is the number of dependent children, and $a_t^k \in \{0, ..., 16\}$ is the age of the oldest child.[20] r_t^q is the current market rent per sqm for a unit of quality q. The type z captures unobserved preferences for residential

[19] I use financial assets reported in the 2002, 2007, and 2012 SOEP 'wealth modules' and the savings of the household reported in each survey year to calculate forward and backward the financial wealth. For simplicity, I ignore potential returns through interest, as well as withdrawals.

[20] $s_t = 4$ for units with at least four rooms, and $k_t = 2$ if at least two dependent children are present.

mobility and the two owner-occupier choices (strong/weak, $2^3 = 8$ combinations).

State transitions. Household income, financial wealth, the couple indicator, and the number of children follow a stochastic transition path. The income transition depends on current income, the number of adults and children, and on the age of the household head and of the first child, incorporating life-cycle effects, earnings persistence, and labor supply effects from having (young) children. The wealth transition is a function of disposable income net of housing costs, the lead income change, and a move indicator, since moving costs may reduce the amount saved. The transitions of the couple indicator and the number of children depend flexibly on household composition and age. Appendix O-C.1 provides technical details.

The other state variables evolve in straightforward ways: $a_{t+1} = a_t + 1$, $a^k_{t+1} = a^k_t + 1$ if $k_t > 0$, and $\tau_{t+1} = \tau_t + 1$ if $j = 0$, and $\tau_{t+1} = 1$ otherwise. Moreover, I assume that the household expects real rents to remain constant in the next period.

Flow utility of rental housing. Living in rental housing of quality and size (q, s) provides deterministic flow utility of

$$\begin{aligned} u_{jt}(x_t) =& \theta_0^{k_t=0} dispinc_{jt} + \theta_1^{k_t=0} dispinc_{jt}^2 + \theta_0^{k_t>0} dispinc_{jt} + \theta_1^{k_t>0} dispinc_{jt}^2 \\ &+ \sum_{s=1}^{4} \left[\theta_2^{s,\text{single}} \mathbb{1}(s_{jt} = s, m_t = 0) + \theta_2^{s,\text{couple}} \mathbb{1}(s_{jt} = s, m_t = 0) + \theta_2^{s,\text{kids}} k_t \mathbb{1}(s_{jt} = s) \right] \\ &+ \theta_3 q_{jt} e^{-\delta \tau_{jt}} + \theta_4 \left[q_{jt} e^{-\delta \tau_{jt}} \right]^2 + \theta_5 \tau_{jt} + \theta_6 \tau_{jt}^2 + \varepsilon_{jt}, \qquad j \in \{0, ..., 40\}. \end{aligned} \tag{2}$$

$dispinc_{jt} = (y_t - r_{jt} - 2.5 \cdot S(s_t)) \cdot (1 + m_t + k_t/2)^{-0.5}$ is the equivalized disposable household income net of costs for shelter, where $2.5 \cdot S(s_t)$ is the total cost for utilities and $S(s)$ is the average floor size of units with s rooms. For $j > 0$, (q_{jt}, s_{jt}) equals the pair (q, s) corresponding to $j > 0$, and $r_{jt} = r_t^{q_{jt}} S(s_{jt})$. I follow Calder-Wang (2019) in assuming that households take the rent distribution as given because they face a competitive housing market with an atomistic demand side.

Households gain utility from a quadratic in $dispinc_{jt}$, where parameters vary by presence of children. The household's valuation of size s_{jt} in line 2 depends on the number of adults and children. Households also value housing quality $q_{jt}e^{-\delta\tau_{jt}}$, $\delta > 0$, which depreciates over time, as captured by the negative exponential decay terms in line 3. Attachment to the unit is captured by the quadratic in the length of tenure τ_{jt}.

Moving costs. Following Kennan and Walker (2011), Buchinsky et al. (2014), and others, I allow moving costs to depend on the household's characteristics.

$$MC_{jt}(x_t, z) = \mathbb{1}(j > 0)\left(\mu_0^z + \mu_1 a_t + \mu_2 a_t^2 + \mu_3 m_t + \mu_4 k_t + \mu_5 \mathbb{1}(a_t^k > 5)\right). \tag{3}$$

In contrast to Kennan and Walker (2011) and Buchinsky et al. (2014), these moving costs reflect renters' costs of moving *within* a local housing market. They depend on the age of the household head, and the presence of a partner, children, and school children.

Lifetime utility of terminal choices. The valuation of the terminal choices is modeled in reduced form. The total deterministic payoff of choosing $j = 41, 42, 43$ is

$$v_{jt}(x_t, z) = \gamma_{j0}^z + \gamma_{j1}\ln(y_t) + \gamma_{j2}w_t + \gamma_{j3}w_t^2 + \gamma_{j4}a_t + \gamma_{j5}a_t^2 + \gamma_{j6}m_t + \gamma_{j7}k_t + \gamma_{j8}\mathbb{1}(a_t^k > 5). \tag{4}$$

For $j = 41, 42$, $\gamma_{j,0}^z$ depends on the unobserved household type and may take on two values.

Idiosyncratic component of utility. I assume that the payoffs for each choice have an idiosyncratic component ε_{jt} that represents household- and period-specific preferences for alternative j. The preference shocks are drawn independently over time and alternative from a Type-I Extreme Value distribution. The unobserved heterogeneity across household types implies that the model does not suffer from 'independence of irrelevant alternatives'. From the perspective of the econometrician, the errors exhibit dependence over choices.

Choice problem. The household maximizes lifetime utility by selecting an optimal choice sequence $d^*(t) := (d_{t'}^*)_{t' \geq t}$, where $d_t = (d_{0t}, ..., d_{43t})$ and d_{jt} is an indicator for choosing alternative j in period t. Letting $\chi_{t,t'} = \prod_{\tilde{t}=t}^{t'-1}\left(1 - \sum_{j=41}^{43} d_{j\tilde{t}}\right)$ be an indicator for not having made a terminal choice between periods t and $t'-1$, and defining $\tilde{u}_{jt} = u_{jt} - MC_{jt}$ and $\tilde{v}_{jt} = v^r_{jt} - MC_{jt}$, the expected discounted sum of payoffs for choice j is

$$\max_{d(t)} \sum_{t'=t}^{T} \chi_{t,t'} \beta^{t'-t} \left[\sum_{j=0}^{40} d_{jt'} \mathbb{E}_t[\tilde{u}_{jt'}(x_{t'}) + \varepsilon_{jt'}] + \sum_{j=41}^{43} d_{jt'} \mathbb{E}_t[\tilde{v}_{jt'}(x_{t'}, z) + \varepsilon_{jt'}] \right]. \tag{5}$$

β is the discount factor, and $\mathbb{E}_t$ represents the expectation at time t.

Market equilibrium

The dynamic discrete choice model determines choice probabilities for the different housing choices. I use these choice probabilities to construct aggregate demand and supply for the 40 housing types by aggregating over a sample of households indexed by n.

Aggregate supply of rental housing. Rental supply of units with quality q and size s is

$$S_{q,s}(r) = S^{\text{primary}}_{q,s} + S^{\text{secondary}}_{q,s}(r). \tag{6}$$

$S^{\text{primary}}_{q,s}$ is exogenously fixed to match the share of new (q, s) units in the rental housing data, and scaled such that $\sum_{q,s} S^{\text{primary}}_{q,s}$ makes up 31.6% of total new supply, i.e., the share of new rental housing supply in the SOEP data. Secondary supply is given by

$$S^{\text{secondary}}_{q,s}(r) = \sum_n \mathbb{1}(s_n = s)(1 - p_0(r, q_n, s_n, r_n, \tau_n; x^-_n))\ell(q|q_n, \tau_n) w_n. \tag{7}$$

$1 - p_0(r, q_n, s_n, \tau_n; x^-_n)$ is the probability that a household with characteristics x^-_n facing a rent vector r and currently occupying a unit with quality q_n and size s_n at rent r_n for τ_n years chooses to move out of the current housing unit.

I assume that landlords ugrade a unit of quality q_n occupied for τ_n years to quality q with probability $\ell(q|q_n, \tau_n)$, and that $\ell(q|q_n, \tau_n) = 0$ if $q_n e^{-\delta\tau_n} < .8$, representing a 'minimum quality requirement'.[21] Finally, w_n is a sampling weight.

[21] I rule out that landlords self-occupy rental units. The function accounts for depreciation over the τ_n periods and scrappage of units at the bottom of the quality distribution. Appendix O-C provides further details.

Aggregate Demand. Aggregate demand consists of local and external demand,

$$D_{q,s}(r) = D_{q,s}^{\text{local}}(r) + D_{q,s}^{\text{external}}(r), \tag{8}$$

$$D_{q,s}^{\text{local}}(r) = \sum_{n} p_{(q,s)}(r, q_n, s_n, \tau_n; x_n^-)w_n, \tag{9}$$

$$D_{q,s}^{\text{external}}(r) = \sum_{n} p_{(q,s)}(r, q_n, s_n, \tau_n; x_n^-)w_n'. \tag{10}$$

The weight w_n' reflects household n's propensity to move to the local market.

Market equilibrium. The equilibrium rent vector $r^* \in \mathbb{R}_+^{40}$ satisfies

$$D_{q,s}(r^*) = S_{q,s}(r^*) \quad \forall (q, s) \in \{1, ..., 10\} \times \{1, ..., 4+\}. \tag{11}$$

I also require that the demand for new and existing owner-occupied housing matches the exogenously given supply of these housing types.

Discussion of model mechanism

In this framework, a reduction of new supply of a particular housing type shifts unmet demand from that housing type to other housing types. Close substitutes experience the largest increases in demand.

At the same time, part of the unmet demand is re-directed to the 'stay' choice, i.e. some households are going to decide to stay in their current home instead. This leads to a reduction of secondary supply, as suggested by the reduced-form evidence from Table 7. The distribution of this secondary supply effect depends on the distribution of housing types occupied by the now-stayer households. This distribution does not necessarily depend on the substitutatbility between units, as households adjust housing choices only infrequently.

3.2. Estimation of structural parameters

Household panel data

The main data source for the model is the SOEP, 2001-2017. The sample starts in 2001 because a novel move indicator is available from that year onwards. Housing quality is captured by the unit's position in the local distribution of rent/sqm. To measure the rent distribution going back to 2001, I employ rich data on rents from the Mikrozensus, a large repeated cross-section of about 400,000 households.[22]

There are 2,957 households in the sample with full information on all variables. Table 8 reports summary statistics and Table O-C1 reports the number of households by number of consecutive years a household was observed. The sample consists of renter households that moved house at least once between 2001 and 2017 and are hence relatively mobile as compared to the German population as a whole. Moreover, the terminal choices remove a household from the sample. 642 households appear in the data for ten or more consecutive years. 281 renters move into existing units as owner-occupiers and 117 move into new units. There are 2429 moves within the local rental market, and 401 moves out of the local market.

Table 8: Descriptive statistics for the SOEP household sample

			Quantile				
	Mean	SD	.25	.5	.75	Min	Max
housing quality	5.84	2.92	3.00	6.00	8.00	1.00	10.00
housing unit size (1/2/3/4+)	2.88	0.86	2.00	3.00	4.00	1.00	4.00
length of tenancy	2.41	2.76	0.00	2.00	4.00	0.00	15.00
rent	557.4	253.3	395.1	503.6	667.0	15.5	4157.2
rent per sqm	7.26	2.31	5.80	6.94	8.33	0.28	35.71
rent per sqm (size-adjusted)	7.20	2.26	5.76	6.86	8.24	0.27	34.74
monthly net real hh income (1k EUR)	2.42	1.21	1.51	2.20	3.09	0.41	8.27
yearly real savings (1k EUR)	2.59	5.19	0.00	0.72	3.08	0.00	150.00
real acc. savings (imputed, 1k EUR)	48.3	211.5	1.6	12.2	40.7	0.0	10453.9
age of household head	44.11	15.37	32.00	41.00	54.00	18.00	94.00
couple household	0.55	0.50	0.00	1.00	1.00	0.00	1.00
number of children (0/1/2+)	0.61	0.82	0.00	0.00	1.00	0.00	2.00
age of oldest child	2.62	4.73	0.00	0.00	3.00	0.00	16.00
year	2010	4	2007	2010	2013	2002	2017

Notes: Sample of SOEP households used in the estimation, exlcuding the period when the household was first observed. Housing quality is determined by the position in the local rent/sqm distribution at the time of moving. The size-adjusted rent/sqm is corrected for the correlation between size and rent/sqm, using a regression estimated from the rent data employed in Section 2. Accumulated savings were imputed from SOEP waves 2002, 2007, and 2012 ('wealth module'), using the savings variable (reported in all waves). Euro values refer to the price level in 2017.

Figure O-C1 shows that the quality of rental housing units occupied by subsequent first-time buyers of new homes is relatively dispersed, partly due to the influence of quality deprecation, suggesting that the initial impact of a shock to new supply on rental prices may be equally dispersed across quality levels.

[22]Source: RDC of the Federal Statistical Office and Statistical Offices of the Federal States, *Mikrozensus*, survey years 2006, 2010, 2014, 2018. Details are provided in Appendix O-C.3

Discount factor and housing quality decay

I follow the literature in assuming $\beta = .95$. In the present context, β is relevant only one period ahead. Discount factors in periods beyond $t + 1$ are subsumed into the non-parametric control factor, see Appendix O-C.5. Hence, the estimation allows for the possibility that discount rates are downward-sloping over long horizons (Giglio et al., 2014, 2021).

Depreciation of housing quality captures the change of the unit's position in the local rent/sqm distribution and is estimated from the rent data. The estimated depreciation factor is 4% p.a., capturing pure depreciation excluding effects of maintenance. Appendix O-C.4 provides details.

Dynamic discrete choice problem

The discrete choice model is estimated using the maximum-likelihood-based EM algorithm of Arcidiacono and Miller (2011). Technical details are given in Appendix O-C.5.

Flow utility of rental housing and moving costs. Table 9 displays parameter estimates for the flow utility of rental housing in Panel A and for the moving cost component in Panel B, for two versions of the model. Model 1 does not allow for unobserved heterogeneity, while Model 2 is the unrestricted model. Panel C reports the log likelihood and an LR ratio test, which supports Model 2. The distribution of unobserved types is reported in Table O-C3.

Regarding the parameter estimates, flow utility increases in disposable income, rental housing quality, and the time since the last move capturing attachment to the unit, but at decreasing rates. Although the two age coefficients are not significant separately, the overall effect of age on moving costs is significant, with older renters being less mobile.

The flow utility of housing size depends on household composition in a flexible way. Figure 3 displays the flow utility from having two, three, or four or more rooms instead of just one room, respectively. The two panels correspond to Models 1 and 2. The patterns are generally consistent with the conjecture that larger households prefer larger apartments. However, singles prefer two- and three-room apartments over single- and

Table 9: Estimated flow utility parameters and model summary statistics

	Model 1 no unobserved household types		Model 2 8 unobserved household types	
A. Rental housing utility parameters, eq. (2)	Coefficient	SE	Coefficient	SE
disp. income in 1k EUR, hh w/ children	1.752***	0.270	1.666***	0.149
disp. income in 1k EUR squared, hh w/ children	-0.514***	0.073	-0.546***	0.046
disp. income in 1k EUR, hh w/o children	0.950***	0.205	0.856***	0.110
disp. income in 1k EUR squared, hh w/o children	-0.204***	0.040	-0.212***	0.026
housing quality	-0.042	0.026	-0.069***	0.018
housing quality squared	0.007***	0.002	0.009***	0.001
tenancy duration	0.223***	0.021	0.108***	0.015
tenancy duration squared	-0.010***	0.002	-0.005***	0.002
housing unit size	see Figure 3A		see Figure 3B	
B. Moving cost parameters, eq. (3)	Coefficient	SE	Coefficient	SE
intercept (high MC type)	4.687***	0.190	5.078***	0.141
intercept (low MC type)	—	—	3.277***	0.139
age / 100	0.384	0.872	1.967***	0.660
age / 100 squared	0.839	0.910	0.065	0.693
couple household	-0.166***	0.046	-0.173***	0.033
number of children in household	-0.281***	0.033	-0.256***	0.024
school child in household	-0.034	0.059	-0.034	0.046
C. Model summary statistics				
Log Likelihood	-18,049		-17,136	
LR statistic (critical value $\chi^2_{11} = 21.92$)	—		1,826	

Notes: Standard errors (in parentheses) were obtained by block bootstrapping over individuals, with 500 repetitions. *: $p < .1$, **: $p < .05$, ***: $p < .01$. Panel *A* displays the parameter estimates for the flow utility of rental housing eq. (2). Panel *B* displays estimates for the moving cost component eq. (3). For Model 1, the uncertainty related to estimating the conditional choice probability of $j = 42$ is taken into account in the calculation of the standard errors. For Model 2, the bootstrap procedure takes the distribution over unobserved types and the model for the conditional choice probability of $j = 42$ as given, see the technical details on the EM algorithm in Appendix O-C.5.

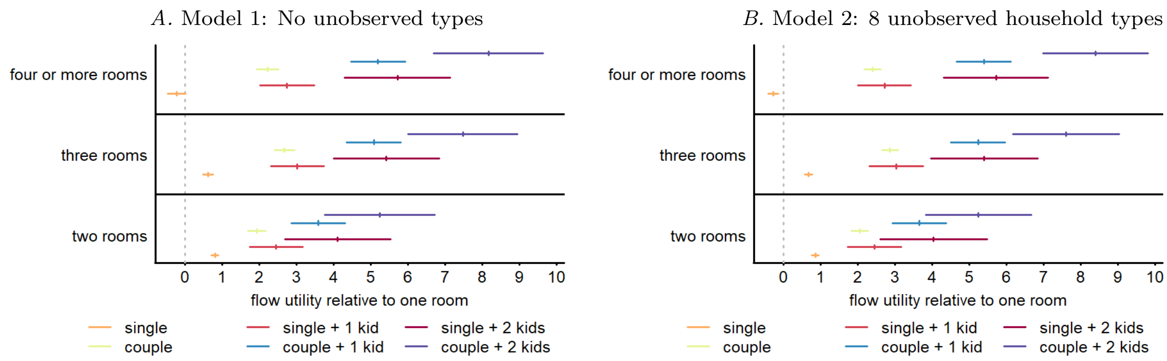
four-room apartments, and couples without children prefer three over four rooms.

Terminal utility. The parameter estimates for the terminal choices, are displayed in Table O-C4. The coefficients in eq. (4) are generally difficult to interpret because they capture the valuation of the three terminal choices *relative to alternative lifetime utilities*, whereby the latter depend on the household's characteristics. It is more straightforward to compare how the variables in eq. (4) affect the relative valuations of the three terminal choices: Higher income increases the propensity to buy a new rather than an existing home, whereas the effect of wealth is not significant. Moreover, older persons are less likely to buy a new home, but the household composition does not matter in this dimension.

Relative to buying an existing home, the propensity to move long-distance decreases in household wealth and with the age of the household head, the latter at an increasing rate. Moreover, the presence of a partner, children, and, in particular, school children decrease the propensity to leave the local market relative to buying an existing house.

Transition functions. The parameter estimates are summarized in Appendix O-C.5.

Figure 3: Flow utility of housing unit size



Notes: The graphs display the valuation of housing unit size by household compositon, relative to a one-room apartment, for Models 1 and 2, respectively. The horizontal bars denote 95% confidence intervals based on bootstrapping with 500 replications.

3.3. Model-based simulations

I combine the discrete choice model together with the system defined by eq.'s (6)–(11) and a population of model households drawn from the SOEP in 2014, the middle of the sample period of the reduced-form analysis. Since the rent distribution is PR-specific, I focus on a single PR, Berlin.

New supply of owner-occupied and rental housing is set such that (i) the shares match the shares of these housing types in the rents data used in Section 2, (ii) the share of total new supply to owner-occupiers matches the respective share in the SOEP data, and (iii) the number of new units supplied to owner-occupiers matches demand.

For each of the 93 sample household, I draw a total of 20 housing choices from the empirical distribution to create a more dense distribution of housing choices. To ensure that the distributions of income, wealth, age, and household composition and the rent distribution in the baseline equilibrium match the observed distributions in the data, I reverse-engineer sample weights w_n^d for each household n and draw d such that the baseline rental price vector solves the equilibrium equation (11). In doing so, I require that each household n gets the same overall weight, $\sum_d w_n^d = \bar{w}\ \forall n$.[23]

The model counterpart of the reduced-form analysis is an exogenous change in new housing supply in the model. The rental price vector adjusts to bring back the model

[23] As population of households immigrating from other locations, I use the simulation sample. The weight of each immigrant household equals the propensity to make a long-distance move. I fix these weights throughout. Appendix O-C.5.1 provides further details on the construction of the simluation sample.

economy into equilibrium. This allows to determine a simulation-based elasticity of the rental price with respect to new supply, and the impact on quantities traded by housing quality segment.

Scenario 1: Reduction of new supply to owner-occupiers

Scenario 1 is a shock to new owner-occupied housing supply, while existing housing supply is completely inelastic.[24] Panel *A* of Figure 4 shows the impact on rental prices and quantities traded, aggregated by housing quality bin. The rent price elasticity in Panel *A1* is somewhat larger than the reduced-form baseline estimate. All quality segments are affected. The model-based elasticities are larger in magnitude for lower qualities. One potential reason is the decreasing marginal utility in disposable income in model, which is why model agents react relatively strongly to rent changes in more expensive market segments.

Panel A2 displays the corresponding effect on the quantities traded in each segment. Strikingly, the first to the seventh quality segment of the rental market exhibit similarly strong increases in quantities traded when new supply to the owner-occupied market increases. Hence, owing to the secondary supply effect, the cascades of moves triggered by the new units reach all quality levels of the rental market. The highest-quality bins are relatively less affected. The model-based estimate of the overall number of units traded for each newly built owner-occupied unit is 4.6, very close to the size of the reduced-form estimate of 4.8 from Table 7.

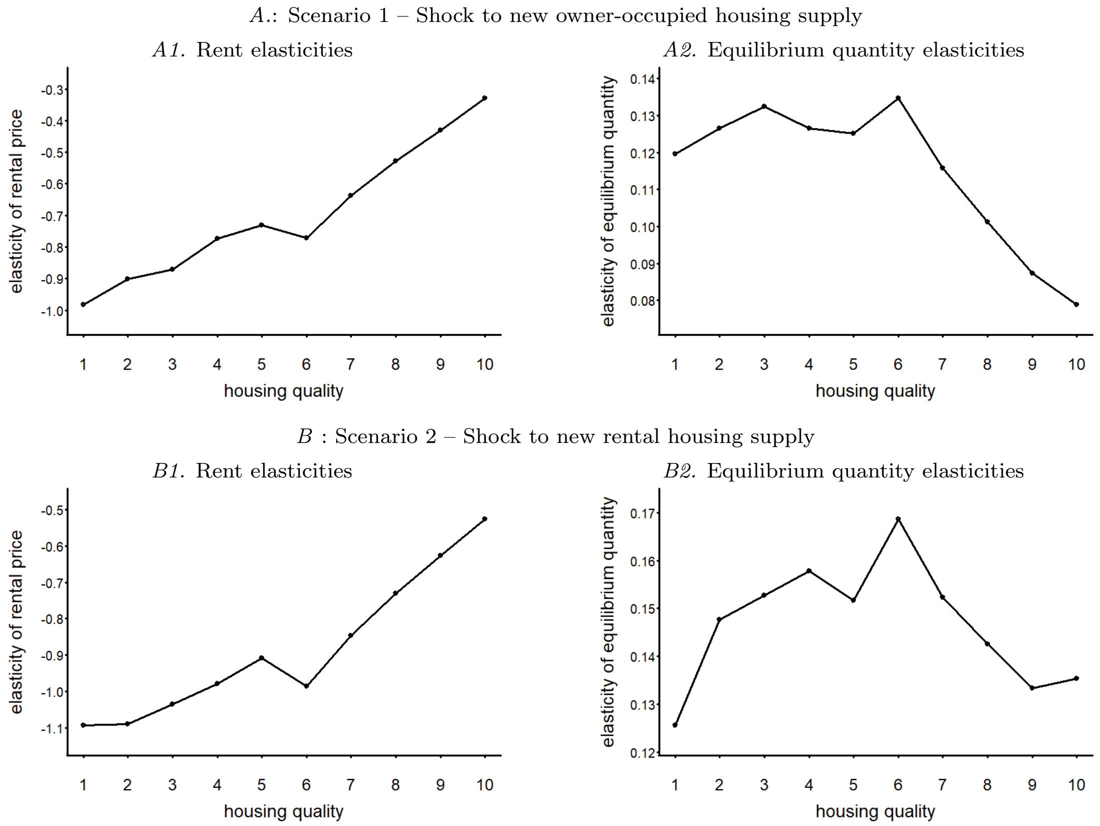
Scenario 2: Reduction of new rental housing supply

In Scenario 2, the supply shock is to new rental housing. New supply of rental housing differs across the 40 segments, and the shock reduces supply in each segment by a common factor. The supply of existing and new housing to owner-occupiers remains fixed.

Panel *B* of Figure O-C3 shows the resulting elasticities. The rent elasticities are slightly larger than in Scenario 1, but exhibit a similar pattern. Moreover, the impact on the quantity traded is now relatively larger for the middle segments, and smaller for very

[24]Figure O-C3 shows results when the supply of existing housing is elastic.

Figure 4: New housing supply: Price and quantity elasticities by housing quality bin



Notes: Panel *A* display the impact of a shock to new owner-occupied housing supply on rental prices and equilibrium quantities traded, aggregated by housing quality bins, represented as an elasticity. In this case, the supply of existing owner-occupied housing is constant. In Panel *B*, the supply shock is to new rental housing. In this case, the supply of new and existing owner-occupied housing is constant. Figure O-C3 shows results for the two scenarios assuming elastic supply of owner-occupied housing.

high and very low-quality units. In this scenario, there are about 6.2 rental units traded in the rental market for each newly built rental unit.

4. Conclusions

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, by identifying the impact of new housing supply at market rates on rental prices in different segments of the local housing market.

The channel through which these effects operate is secondary housing supply: Units freed up by renters moving into the newly built housing trigger a cascade of moves. Through this cascade, the supply effects quickly reach all parts of the local rent distribution, contributing crucially to market integration.

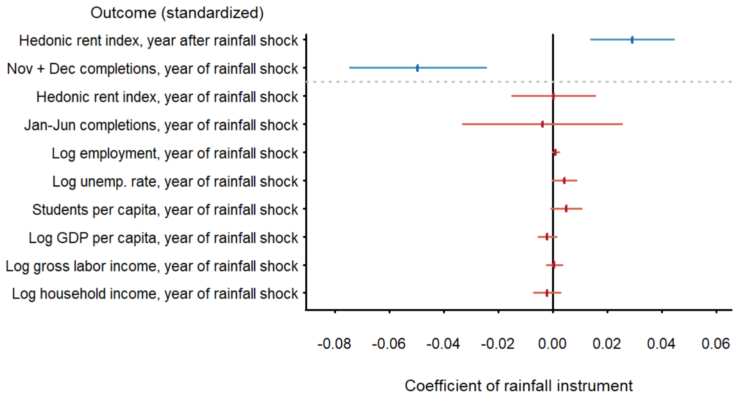
The results imply that restrictions to market-rate housing supply are harmful to low-income renters, as even the supply of single-family homes can lower this group's housing cost burden. The model-based simulations suggests that the supply of new multi-family housing at market rates has even greater potential to curb surging housing costs of low-income households in expensive locations. 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.

The effectiveness of other housing policies likely depends both on the forward-looking nature of housing choices and on the peculiarities of the housing market as a secondary market. Taking into account these factors and their distributional consequences when evaluating housing policies seems to be a fruitful avenue for future reserach.

Appendix

A. Robustness of Baseline Results

Figure A1: Reduced form, first stage, and placebo outcomes



Notes: Vertical bars denote 95% confidence intervals clustered by PR. 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.

Table A1: IV rent regressions controlling for extreme weather events

Dependent variable	Log hedonic rent index	
	(1) IV	(2) IV
Units completed Nov + Dec in $t-1$ per avg. # units completed annually	-0.216*** (0.071)	-0.245*** (0.084)
Dummy: severe flood in federal state in year $t-1$	-0.012* (0.006)	
Year FE	yes	yes
PR FE	yes	yes
Other controls	yes	yes
Kleibergen-Paap F	19.5	15.5
Number of PRs	94	94
Observations	752	666

Notes: Standard errors are clustered by location; *: $p < .1$, **: $p < .05$, ***: $p < .01$. The dependent variable is the log hedonic rent index. The instrument for the supply variable is the summer rainfall shock. In column (1), the dummy variable captures years with severe floods in the federal state the planning region belongs to (2013: Lower Saxony, Hesse, Rheinland-Palatinate, Baden-Wuerttemberg, Bavaria, Saxony, Saxony-Anhalt, Thuringia; 2017: Lower Saxony, Saxony-Anhalt, Thuringia). In column (2), observations from state-years with severe floods were excluded.

Table A2: IV rent regressions controlling for additional local demand factors

Dependent variable	Log hedonic rent index		
	(1) IV	(2) IV	(3) IV
Units completed Nov + Dec in $t-1$ per avg. # units completed annually	-0.187*** (0.067)	-0.215*** (0.075)	-0.212*** (0.071)
Log employment, year $t-1$	0.956*** (0.145)	0.928*** (0.167)	0.524* (0.310)
U & college students per 1,000 inh., year $t-1$	0.003** (0.001)	0.003** (0.001)	0.004** (0.001)
Log unemployment rate, year $t-1$	-0.089** (0.045)	-0.111* (0.058)	-0.065 (0.048)
Share w/o school degree (*Hauptschulabschluss*)	0.003 (0.002)		
Log hours worked, year $t-1$		0.373 (0.535)	
Log gross labor income, year $t-1$		-0.152 (0.253)	
Log employment, year t			0.558* (0.320)
U & college students per 1,000 inh., year t			-0.001 (0.001)
Log unemployment rate, year t			-0.038 (0.054)
Year FE	yes	yes	yes
PR FE	yes	yes	yes
Kleibergen-Paap F	17.0	15.9	16.7
Number of PRs	94	94	94
Observations	752	736	752

Notes: Standard errors are clustered by location; *: $p < .1$, **: $p < .05$, ***: $p < .01$. The dependent variable is the log hedonic rent index. The instrument for the supply variable is the summer rainfall shock. The control variables are taken from the INKAR regional data base. Year $t-1$ refers to the year of the rainfall shock, and year t is the year when the rent index is measured. Data on hours worked is not available in 16 cases (years 2010–2013 in PRs 1601, 1602, 1603, and 1604).

Table A3: IV rent regressions employing alternative instrumental variables

Dependent variable	Log hedonic rent index			
	(1) IV	(2) IV	(3) IV	(4) IV
Units completed Nov + Dec in $t-1$ per avg. # units completed annually	-0.222** (0.104)	-0.220** (0.110)	-0.257* (0.136)	-0.214*** (0.064)
Year FE	yes	yes	yes	yes
PR FE	yes	yes	yes	yes
Other controls	yes	yes	yes	yes
Kleibergen-Paap F	9.6	6.9	6.0	14.2
Number of PRs	94	94	94	94
Observations	752	752	752	752
Instruments	longest rainfall spell	# rainfall spells > 4 days	February frost depth	Feb. frost depth, avg. rainfall spell

Notes: Standard errors are clustered by location; *: $p < .1$, **: $p < .05$, ***: $p < .01$. The dependent variable is the log hedonic rent index. The supply variable is instrumented as indicated in each column.

Table A4: IV rent regressions using an alternative supply variable

Dependent variable	Log hedonic rent index
	(1) IV
Log # of units completed in $t-1$	-0.214**
	(0.089)
Year FE	yes
PR FE	yes
Other controls	yes
Kleibergen-Paap F	10.0
Number of PRs	94
Observations	752

Notes: Standard errors are clustered by location; *: $p < .1$, **: $p < .05$, ***: $p < .01$. The dependent variable is the log hedonic rent index. The supply variable is instrumented by the summer rainfall instrument.

Table A5: IV rent regression specification in changes

Dependent variable	Δ Log hedonic rent index	
	(1) IV	(2) IV
Units completed Nov + Dec in $t-1$ per avg. # of units completed annually	-0.221**	
	(0.100)	
Δ Units completed Nov + Dec in $t-1$ per avg. # of units completed annually		-0.318*
		(0.191)
Year FE	yes	yes
PR FE	yes	yes
Other controls	yes	yes
Kleibergen-Paap F	9.4	3.8
Number of PRs	94	94
Observations	658	658

Notes: Standard errors are clustered by location; *: $p < .1$, **: $p < .05$, ***: $p < .01$. The dependent variable is the change in the log hedonic rent index. The supply variable is instrumented by the summer rainfall instrument.

Table A6: IV rent regressions using alternative spatial delineations

Dependent variable	Log hedonic rent index	
	(1) IV	(2) IV
Units completed Nov + Dec in $t-1$ per avg. # units completed annually	-0.124**	-0.123**
	(0.060)	(0.049)
Year FE	yes	yes
Location FE	yes	yes
Other controls	yes	yes
Spatial unit	commuting zone	district
Number of spatial units	252	392
Kleibergen-Paap F	13.0	11.5
Observations	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. The supply variable is instrumented in each regression, as indicated in each column. The control variables are taken from the INKAR regional data base.

References

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

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

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

Calder-Wang, S. (2019), 'The distributional impact of the sharing economy on the housing market', Working paper, Available at SSRN 3908062.

Cotter, J., S. Gabriel, and R. Roll (2015), 'Can Housing Risk Be Diversified? A Cautionary Tale from the Housing Boom and Bust', *The Review of Financial Studies, 28*(3), pp. 913–936.

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

Deng, Y., C. Han, T. Li, and T. J. Riddiough (2021), 'Whither Weather?: High Temperature, Climate Change and Mortgage Default', Proceedings of paris december 2021 finance meeting eurofidai - essec.

Epple, D., L. Quintero, and H. Sieg (2020), 'A New Approach to Estimating Equilibrium Models for Metropolitan Housing Markets', *Journal of Political Economy, 128*(3), pp. 948–983.

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

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), pp. 63–87.

Giglio, S., M. Maggiori, K. Rao, J. Stroebel, and A. Weber (2021), 'Climate Change and Long-Run Discount Rates: Evidence from Real Estate', *The Review of Financial Studies, 34*(8), pp. 3527–71.

Giglio, S., M. Maggiori, and J. Stroebel (2014), 'Very Long-Run Discount Rates', *The Quarterly Journal of Economics, 130*(1), pp. 1–53.

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

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

Greenwald, D. L. and A. Guren (2020), 'Do credit conditions move house prices?', MIT Sloan School of Management, mimeo.

Gyourko, J., C. Mayer, and T. Sinai (2013), 'Superstar Cities', *American Economic Journal: Economic Policy, 5*(4), pp. 167–99.

Hilber, C. A. and W. Vermeulen (2016), 'The Impact of Supply Constraints on House Prices in England', *The Economic Journal, 126*(591), pp. 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), pp. 497–529.

Kaplan, G., K. Mitman, and G. L. Violante (2020), 'The housing boom and bust: Model meets evidence', *Journal of Political Economy, 128*(9), pp. 3285–3345.

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

Landvoigt, T., M. Piazzesi, and M. Schneider (2015), 'The Housing Market (s) of San Diego', *American Economic Review, 105*(4), pp. 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), pp. 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.

Ortalo-Magne, F. and S. Rady (2006), 'Housing market dynamics: On the contribution of income shocks and credit constraints', *The Review of Economic Studies 73*(2), pp. 459–485.

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), pp. 720–59.

Quigley, J. M. and S. Raphael (2004), 'Is Housing Unaffordable? Why Isn't It More Affordable?', *Journal of Economic Perspectives, 18*(1), pp. 191–214.

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

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), pp. 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), pp. 1253–1296.

Saks, R. E. (2008), 'Job creation and housing construction: Constraints on metropolitan area employment growth', *Journal of Urban Economics 64*, pp. 178–195.

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', DIW Berlin/SOEP.

Van Nieuwerburgh, S. and P.-O. Weill (2010), 'Why Has House Price Dispersion Gone Up?', *Review of Economic Studies 77*, pp. 1567–1606.

Online Appendix — NOT FOR PUBLICATION

O-A. Background Information on Data and Hedonic Rent Indices

Building completions data. The main explanatory variable in the rent regressions is the number of housing units completed in November and December, by PR and year. This variable is aggregated from individual observations in the administrative Building Completions Statistic, which covers 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.

Rental housing data. 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]. 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), and 84% of all rental units were offered on at least two different real estate market places.

Table O-A1 diplays summary statistics for the rents sample. The monthly rent refers to the rent posted on the day the offer appears online for the first time.

Local rent indices. To calculate the local rent indices, I run separate hedonic regressions for each location (PR, district, commuting zone), with the log rent per square meter as the dependent variable, and 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.

Table O-A1: Descriptive statistics for the rents sample

A. Non-categorial and binary variables	Min	Mean	Q25	Median	Q75	Max
Monthly rent per sqm	1.6	8.0	5.5	7.0	9.3	85.2
Living area in sqm	15.0	71.7	52.9	67.0	85.0	300.0
Year of construction	1800	1969	1954	1973	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	0.18	0.00	0.00	0.00	1.00
Fitted kitchen	0.00	0.33	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	0.18	0.00	0.00	0.00	1.00
Balcony or terrace	0.00	0.61	0.00	1.00	1.00	1.00

B. Categorial variables (shares)	0	1	2	3	4	5	6	7	8
Dwelling type	0.597	0.109	0.128	0.008	0.032	0.002	0.006	0.010	0.108
Quality	0.017	0.147	0.831	0.005					

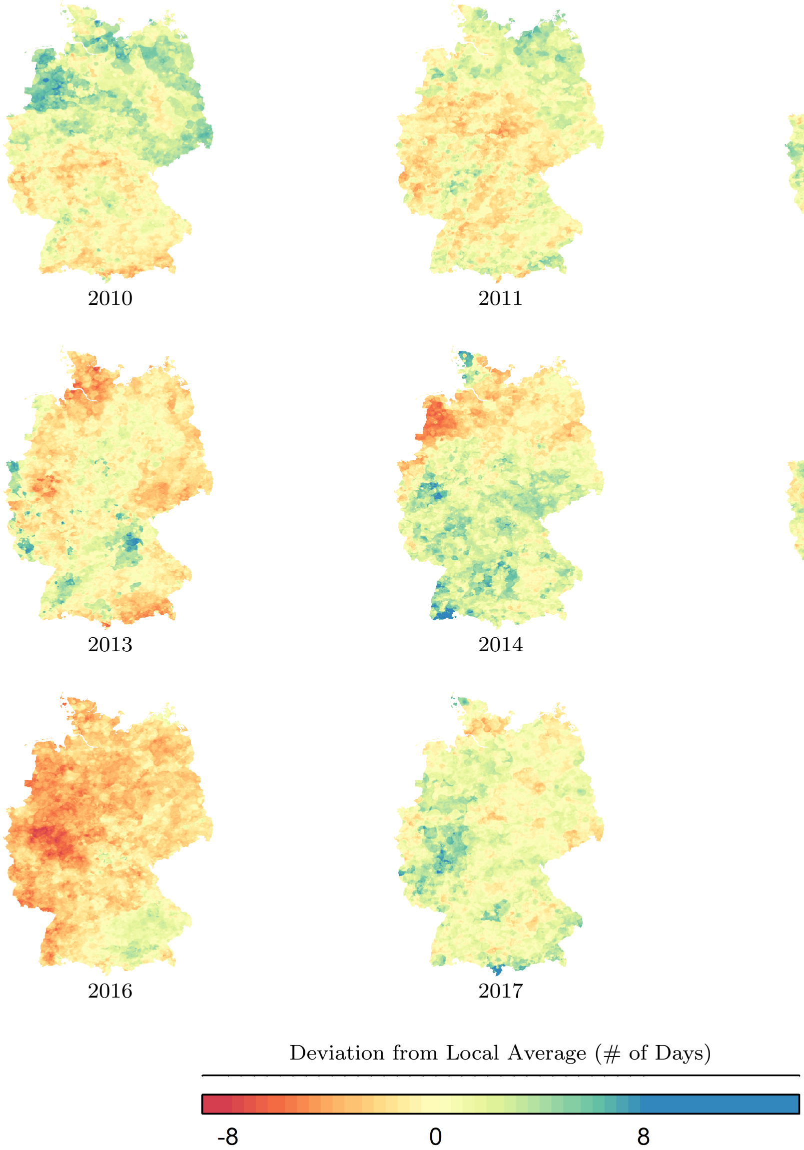
C. Number of observations	
Observations	6,926,371

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.

The controls are the log floor area, a second-order polynomial in the year of construction, an indicator variable for missing year of construction, 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.

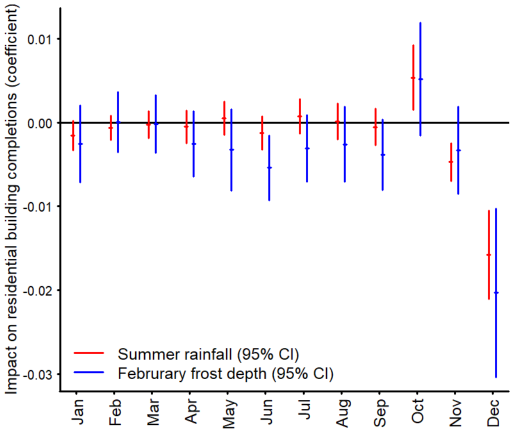
O-B. Weather shocks as temporary shifters of new housing supply

Figure O-B1: Spatial and temporal variation in the summer rainfall shock instrument



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-B2: Impact of the weather shocks on new housing supply throughout the year



Note: The graph displays coefficient estimates of regressions with the number of new units completed in month m relative to the yearly average number of new units as the dependent variable, on the summer rainfall shock and February frost depth, as measured in the same year. The spatial units of observation are municipalities. Vertical bars indicate 95% confidence intervals.

O-C. Further Details on the Quantitative Model

This section provides details on the quantitative model developed in the main text.

O-C.1. Transition functions

Income is modeled as a continous variable. The income process implicitly captures labor-related changes to household income, as well as the accumulation of skill over the life-cycle:

$$
\begin{aligned}
\ln y_{t+1} =& \phi_0^y a_t + \phi_1^y a_t^2 + \phi_2^y \ln y_t \\
&+ \phi_3^y m_t + \phi_4^y m_{t+1} + \phi_5^y \mathbb{1}(a_t^k < 2) + \phi_6^y \mathbb{1}(a_t^k \in \{2,3\}) + \phi_7^y \mathbb{1}(k_t = 2) + \varepsilon_t^y.
\end{aligned} \tag{O-C1}
$$

The transition depends on age and current income, incorporating life-cycle effects and earnings persistence, e.g., due to skill accumulation. Future household income also depends on the number of potential earners, the age of the first child, and whether there are two or more children in the household. The latter capture potential negative labor supply effects from having (young) children. The household forms an expectation over the distribution of one-period-ahead income changes based on eq. (O-C1), drawing from the stored regression residuals ε_t^y.

Accumulated savings are treated as continuous and modeled in an analogous way:

$$
w_{t+1} - w_t = \phi_0^w (y_t - r_t) + \phi_1^w (y_{t+1} - y_t) + \phi_2^w \mathbb{1}(j > 0) + \varepsilon_t^w. \tag{O-C2}
$$

ϕ_0^w and ϕ_1^w represent the savings rates with respect to disposable income net of housing costs, and the contemporaneous income change. ϕ_2^w is the savings reduction due to moving house, allowing for the possibility that financial moving costs are financed byreducing savings.

The remaining two transition functions are modeled via multinomial logits. The latent variable for the couple indicator is

$$
\begin{aligned}
g^m(m_{t+1}=1) =& \phi_0^m + \phi_1^m a_t + \phi_2^m a_t^2 + \phi_3^m \mathbb{1}(k_t>0) + \phi_4^m \mathbb{1}(k_t=2) \\
&+ \mathbb{1}(a_t^k>3)\big[\phi_5^m + \phi_6^m \mathbb{1}(k_t=2)\big] + \varepsilon_t^m, \quad \varepsilon_t^m \sim \text{i.i.d. Type-I EV.}
\end{aligned}
\tag{O-C3}
$$

This equation is estimated separately for $m_t = 0$ and $m_t = 1$.

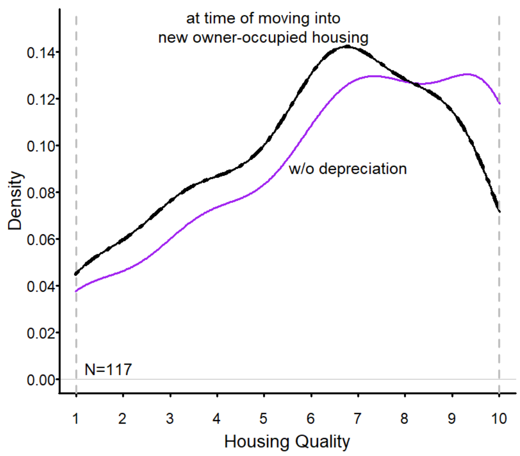
The latent variable for having $k_{t+1} = h$ children in the household is

$$
\begin{aligned}
g^k(k_{t+1}=h) =& \phi_{0,h}^k + \phi_{1,h}^k a_t + \phi_{2,h}^k \mathbb{1}(a_t>40)\mathbb{1}(k_t=0) + \phi_{3,h}^k m_t + \mathbb{1}(k_t>0)\big[\phi_{4,h}^k \\
&+ \phi_{5,h}^k \mathbb{1}(a_t^k=0) + \phi_{6,h}^k \mathbb{1}(a_t^k=1)\big] + \phi_{7,h}^k \mathbb{1}(k_t=2) + \varepsilon_t^{k,h}, \quad \varepsilon_t^{k,h} \sim \text{i.i.d. Type-I EV.}
\end{aligned}
\tag{O-C4}
$$

These transition functions allow for a rich dependence structure of household composition and age on future household composition, which in turn affects household income and wealth accumulation.

O-C.2. Descriptive evidence

Figure O-C1: Rental housing quality distribution of subsequent first-time buyers of new housing



Note: The graph shows the density of housing quality among renters moving into newly built owner-occupied housing (including and net of depreciation).

O-C.3. Additional details on the SOEP household data and the measure of housing quality

Sample households by number of periods observed. Table O-C1 lists the number of households observed by total number of periods the household was observed in the data. Households making a terminal choice are dropped from the sample. For instance, a household observed over 13 periods stayed in the local rental housing market for at least 12 periods and was surveyed for all 13 periods. If that household leaves the local rental market in the 13th period, it may still be interviewed in subsequent SOEP waves, but it does not re-appear in the analysis sample.

Table O-C1: Number of complete cases in the SOEP household sample

years observed	2	3	4	5	6	7	8	9
# of households	333	353	327	307	372	301	178	144
years observed	10	11	12	13	14	15	16	17
# of households	118	114	86	72	67	64	54	67
Total # of hh	2,957							

Housing quality measure in the SOEP. The Mikrozensus included housing modules in 2006, 2010, 2014, and 2018, allowing the estimation of the PR-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. 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 and assigning quality based on the size-adjusted rent/sqm. This strategy renders the quality and size dimensions orthogonal to each other, consistent with the reduced-form approach in Section 2, where hedonic indices are conditional on the size of the unit. This approach puts a restriction on the sample, because housing quality can only be measured if a household moved into a rental housing unit between 2001 and 2017. I exclude households that were not observed to have moved in that period.

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 a yearly panel.[25] 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.

O-C.4. 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 PR-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.

Depreciation of housing quality is estimated from the rental housing offer data. 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 time difference between two offers is 29 months, and the mean is 33.5 months. Rents per square meter increased by 0.075 log points on average.

The goal is to estimate pure depreciation net of maintenance because, arguably, units are not renovated during tenancies. I therefore restrict the sample further to pairs of units where observable characteristics of the unit (condition, fitted kitchen, flooring) remain unchanged. This applies to 94,706 pairs. The mean and median time differences are one month smaller in this sub-sample. A potential reason is that landlords removed units with altered characteristics from the market temporarily for renovation works. The average rent change shrinks to 0.062 log points in the group of units with unchanged characteristics, the difference of about 0.013 log points possibly representing the average value of the alterations. The measure of housing quality $q_i \in \{1, ..., 10\}$ is defined by

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

using the 10%-, ..., 90%-quantiles of the local rent distribution (per square meter) as breaks, which are measured in the full rent sample (by PR and year).

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

$$\Delta \ln q_i = \delta \Delta \text{years}_i + \text{postcode}_i + \eta_i. \tag{O-C5}$$

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 PR. δ is the quality decay factor. I restrict the sample to units that start at a quality level of 3 to 10.[26] Table O-C2 displays the estimation results.

Table O-C2: Estimated housing quality decay factor

	Dependent variable: $\Delta \ln q$	
	OLS (1)	OLS (2)
Δ years	-0.039*** (0.005)	-0.047*** (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: Standard errors (in parentheses) clustered by ROR; *: $p < .05$, **: $p < .01$, ***: $p < .001$. q is the discretized, normalized rank of the unit in the PR-level rent/sqm distribution ($q \in \{1, \dots, 10\}$). The sample is restricted to units observed at least twice, without observable changes to unit characteristics. Units offered as being renovated or refurbished when observed the second time were excluded. The initial position in the rent distribution is above 2 and the time difference between two observations is at least 12 months.

Table O-C2 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-

[26]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 captured well by a common exponential discounting factor.

order polynomial in the time difference does not lead to a better fit.

O-C.5. Estimation of the dynamic discrete choice model

Technical details

Choice Problem. According to Arcidiacono and Miller (2011) and Hotz and Miller (1993), the difference in the expected payoffs of choices $j = j'$ and $j = 0$ in period t, net of the idiosyncratic components in period t, are logistically distributed and can be expressed as

$$
\begin{aligned}
v_{j't}(x_t, z) - v_{0,t}(x_t, z) &= u_{j't}(x_t, z) - u_{0,t}(x_t, z) \\
&+ \beta \sum_{x_{t+1}} \left[f_{j't}(x_{t+1}|x_t) - f_{0,t}(x_{t+1}|x_t) \right] \left[v_{41,t+1}(x_{t+1}, z) - \ln p_{41,t+1}(x_{t+1}, z) \right]. \quad \text{(O-C6)}
\end{aligned}
$$

The term on the right-hand side of the first line represents the current-period flow utility difference. The sum in the second line is over all attainable states x_{t+1} in period $t + 1$. The first term in brackets represents a probability weight for state x_{t+1}, which can be positive or negative. Depending on the initial choice, the probability to reach a particular x_{t+1} is given by $f_{jt}(x_{t+1}|x_t)$, which is implicitly defined by the transition rules described above.

The second term in brackets is the terminal utility of choosing to live in existing owner-occupied housing, $j = 41$, net of a correction factor $\ln p_{41,t+1}(x_{t+1})$. This factor corrects for the fact that $j = 41$ may not be the optimal choice when coming into period $t + 1$ being in state x_{t+1}. Intuitively, the correction is large if the probability to choose $j = 41$ is small, since the latter implies that $j = 41$ is not a very common choice, suggesting that the true utility of being in state (x_{t+1}, z) is much higher than $v_{41,t+1}(x_{t+1}, z)$. If, conversely, $p_{41,t+1}(x_{t+1}, z) \approx 1$, the correction is close to zero. I use a regression forest (Athey et al., 2019) for predicting the empirical conditional probability to choose $j = 41$, $\hat{p}_{41,t+1}(x_{t+1}, z)$, which replaces $p_{41,t+1}(x_{t+1}, z)$ in the estimation.

Transition functions. The log likelihood is separable in the parameters of the transition and utility functions. I therefore estimate the transition functions in a separate step.

Expectation-Maximization algorithm. I plug into (O-C6) the estimated transition processes, the housing quality decay factor, and $\beta = .95$. I then form the log likelihood over the full choice sequence and for all households, using (O-C6). For known conditional probabilities of the unobserved household types, this is the standard maximum likelihood estimator for multinomial logit models. These probabilities are found using the expectation-maximization algorithm of Arcidiacono and Miller (2011), which iterates back and forth between the maximum likelihood step and an 'expectation step' used to update the conditional probabilities of the unobserved household types. I repeat this procedure using different starting values for the conditional probabilities of the unobserved types.

Coefficient estimates for the terminal choices and transition functions

Table O-C3: Unobserved types (Model 2)

type z	1	2	3	4	5	6	7	8
unconditional probability	0.057	0.015	0.296	0.065	0.128	0.015	0.300	0.123
intercept moving costs	-5.08***	-3.28***	-5.08***	-5.08***	-3.28***	-3.28***	-5.08***	-3.28***
	(0.14)	(0.14)	(0.14)	(0.14)	(0.14)	(0.14)	(0.14)	(0.14)
intercept buy existing home	2.38***	2.38***	4.97***	2.38***	4.97***	2.38***	4.97***	4.97***
	(0.86)	(0.86)	(0.86)	(0.86)	(0.86)	(0.86)	(0.86)	(0.86)
intercept buy new home	-9.82***	-9.82***	-9.82***	-7.74***	-9.82***	-7.74***	-7.74***	-7.74***
	(1.38)	(1.37)	(1.38)	(1.37)	(1.38)	(1.37)	(1.37)	(1.37)

Notes: Standard errors (in parentheses) were obtained by block bootstrapping over individuals, with 500 repetitions, takeing the distribution over unobserved types and the model for the conditional choice probability of $j = 41$ as given, see the technical details on the EM algorithm in O-C.5. *: $p < .1$, **: $p < .05$, ***: $p < .01$. The table displays the estimated unconditional probability to be of type $z = 1, \ldots, 8$ and the corresponding coefficient estimates of the three parameters that differ by unobserved household type.

Table O-C4: Estimated terminal utility parameters

	Model 1 no unobserved household types				Model 2 8 unobserved household types			
	buy new vs. buy existing		long-distance move vs. buy existing		buy new vs. buy existing		long-distance move vs. buy existing	
	Coef.	SE	Coef.	SE	Coef.	SE	Coef.	SE
intercept (avg. over types)	-9.828***	1.903	5.913***	1.058	-11.686***	2.179	4.475***	0.737
log household income	0.795***	0.224	-0.143	0.144	1.092***	0.211	0.120	0.100
wealth in 100k EUR	0.024	0.213	-0.230***	0.073	0.052	0.256	-0.248***	0.055
wealth in 100k EUR sq.	0.000	0.037	0.002	0.004	-0.000	0.041	0.002	0.003
age of hh head	0.112*	0.064	-0.147***	0.030	0.132**	0.065	-0.136***	0.027
age of hh head squared	-0.001*	0.001	0.001***	0.000	-0.002**	0.001	0.001***	0.000
couple household	0.331	0.230	-0.812***	0.149	0.312	0.209	-0.825***	0.136
number of children	0.106	0.106	-0.375***	0.080	0.085	0.096	-0.427***	0.075
school child in hh	-0.181	0.204	-0.377**	0.150	-0.044	0.188	-0.286**	0.138

Notes: Standard errors (in parentheses) were obtained by block bootstrapping over individuals, with 500 repetitions. *: $p < .1$, **: $p < .05$, ***: $p < .01$. The table displays the coefficient differences between the terminal choices denoted in the column header, as defined in eq. (4). For Model 1, the uncertainty related to estimating the conditional choice probability of $j = 41$ is taken into account in the calculation of the standard errors. For Model 2, the bootstrap procedure takes the distribution over unobserved types and the model for the conditional choice probability of $j = 41$ as given, see the technical details on the EM algorithm in O-C.5.

Table O-C5: Parameter estimates for the income and wealth transition functions

A. Income transition, eq. (O-C1)		
Dep. variable: lead log household income	Coef	SE
age of household head	0.009***	0.001
age squared $\times 1e-3$	-0.102***	0.008
log income in t	0.975***	0.002
couple household in t	-0.425***	0.014
couple household in $t+1$	0.452***	0.013
first child < 2 years old	-0.015**	0.006
first child 2–3 years old	0.038***	0.009
2+ children in household	-0.015***	0.004
Adj. R squared	0.765	
Observations	17,025	
B. Savings transition, eq. (O-C2)		
Dep. variable: savings in current year in EUR	Coef	SE
disposable income (net of rent)	0.107***	0.006
income change	0.132***	0.009
mover household	-911.647***	101.070
Adj. R squared	0.111	
Observations	17,025	

Notes: Cluster-robust standard errors in parentheses; *: $p < .1$, **: $p < .05$, ***: $p < .01$.

Table O-C6: Parameter estimates for the couple and children transition functions

A. Couple transition, eq. (O-C3)

Outcome: 2+ adults in hh (lead)	one adult in household (1)	two adults in household (2)
intercept	-2.394***	0.291
	(0.459)	(0.437)
age of household head	0.052**	0.121***
	(0.024)	(0.021)
age squared $\times 1e-3$	-1.247***	-1.107***
	(0.299)	(0.212)
1+ children in household	-0.336*	0.437**
	(0.203)	(0.207)
2+ children in household	0.530	-0.908***
	(0.420)	(0.247)
First child > 3 yrs	-0.018	-0.309
	(0.250)	(0.273)
First child > 3 yrs × 2+ children	-0.282	1.161***
	(0.473)	(0.341)
Log Likelihood	-1,950	-1,490
Observations	7,890	9,135

B. Children transition, eq. (O-C4)

	Coef	SE
Outcome: 1 child in household		
intercept	-2.791***	0.190
age of household head ($k = 1$)	-0.011*	0.006
age > 40× no kids	-1.764***	0.234
couple household	0.691***	0.096
1+ children	7.217***	0.202
2+ children	-3.367***	0.198
first child born this year	-2.117***	0.213
first child born last year	-0.350	0.539
Outcome: 2+ children in household		
intercept	-4.840***	0.269
age of household head	-0.027***	0.006
age > 40× no kids	0.175	0.329
couple household	1.212***	0.122
1+ children	7.629***	0.266
2+ children	2.596***	0.202
first child born this year	-3.195***	0.219
first child born last year	-1.173**	0.570
Log Likelihood	-3,722	
Observations	17,025	

Notes: Standard errors; *: $p < .1$, **: $p < .05$, ***: $p < .01$ (both panels). Column (1) of Panel *A* refers to the transition probability to becoming a couple household when not being a couple household in the current period, whereas column (1) of Panel *A* refers to the respective transition probability for couple household in the current period.

O-C.5.1. Construction of the simulation sample

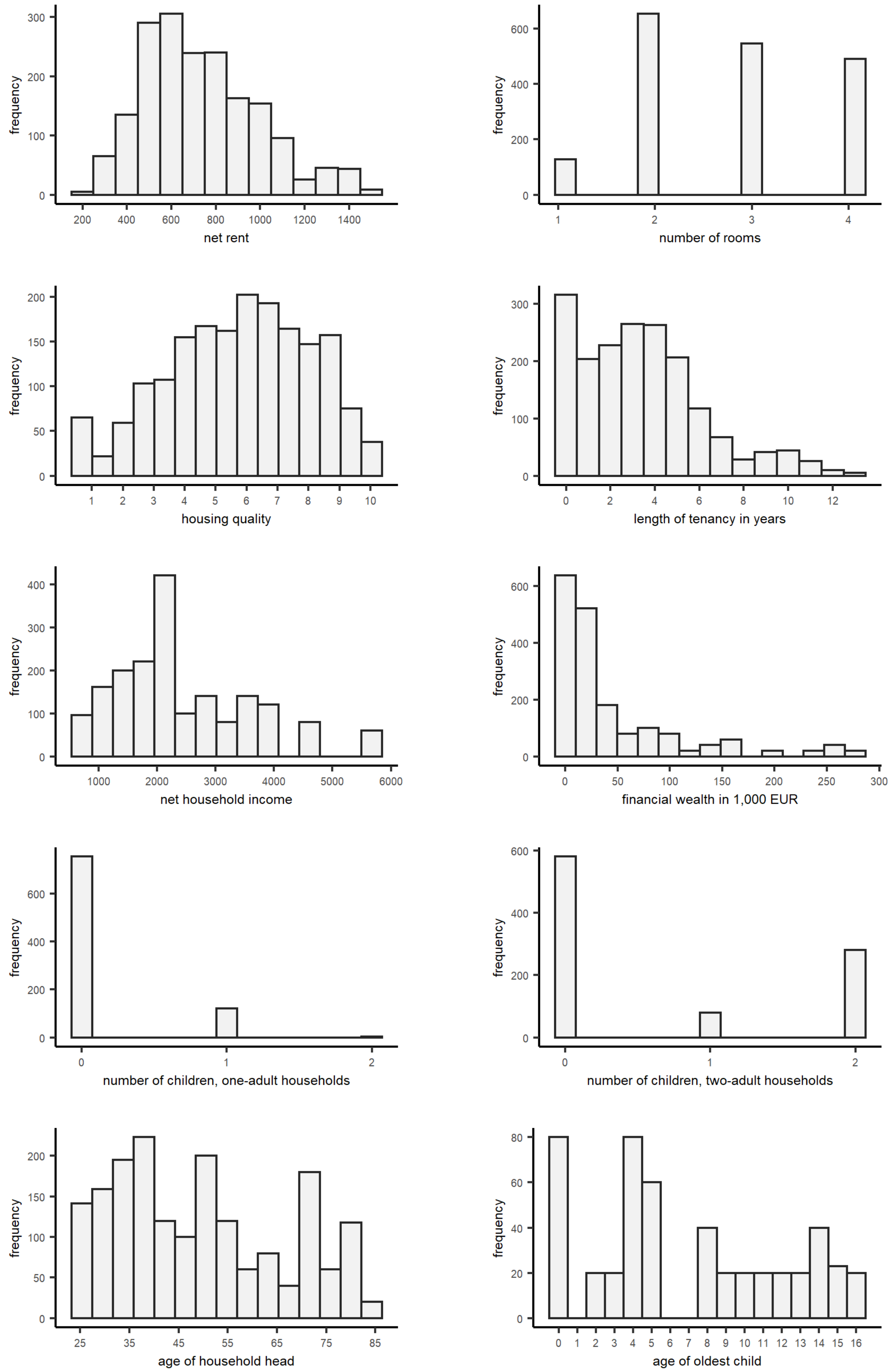
I construct a simulation sample that reflects the income, age and household structure, and financial wealth of households in Berlin in 2014, in the middle of the sample period of the reduced-form analysis.

In detail, the simulation sample is formed as follows: (i) Select all 93 renter households observed in 2014 in Berlin. (ii) Estimate OLS models for size, housing quality, and stay length of each sample household, with household composition, age, household income, and financial wealth as explanatory variables. (iii) For each sample household, draw 20 times from the distributions of regression errors and compute the size, housing quality, and stay length as the predicted value plus the error draw, rounded to the nearest category. (iv) Compute the rent for the housing unit currently occupied by using the local rent distribution and the unit's size and quality as inputs. I assume that the yearly real rent increase was 2% and that rents are fixed nominally during a tenancy due to tenancy rent control. Draws with rent expenditure shares below 5% or above 80% are discarded. (v) The sample weights are then determined by fixing the initial equlibrium rent vector to the observed rent distribution in Berlin in 2014, and by making use of the derivatives of demand and suppy with respect to w_n^d and requiring that the sum of weights for each household is the same, $\sum_d w_n^d = \bar{w}\ \forall n$. The resulting sample has an income, financial wealth, and age and household structure distribution as observed in the data. Figure O-C2 displays histograms for the simulation sample.

O-C.6. Additional scenarios

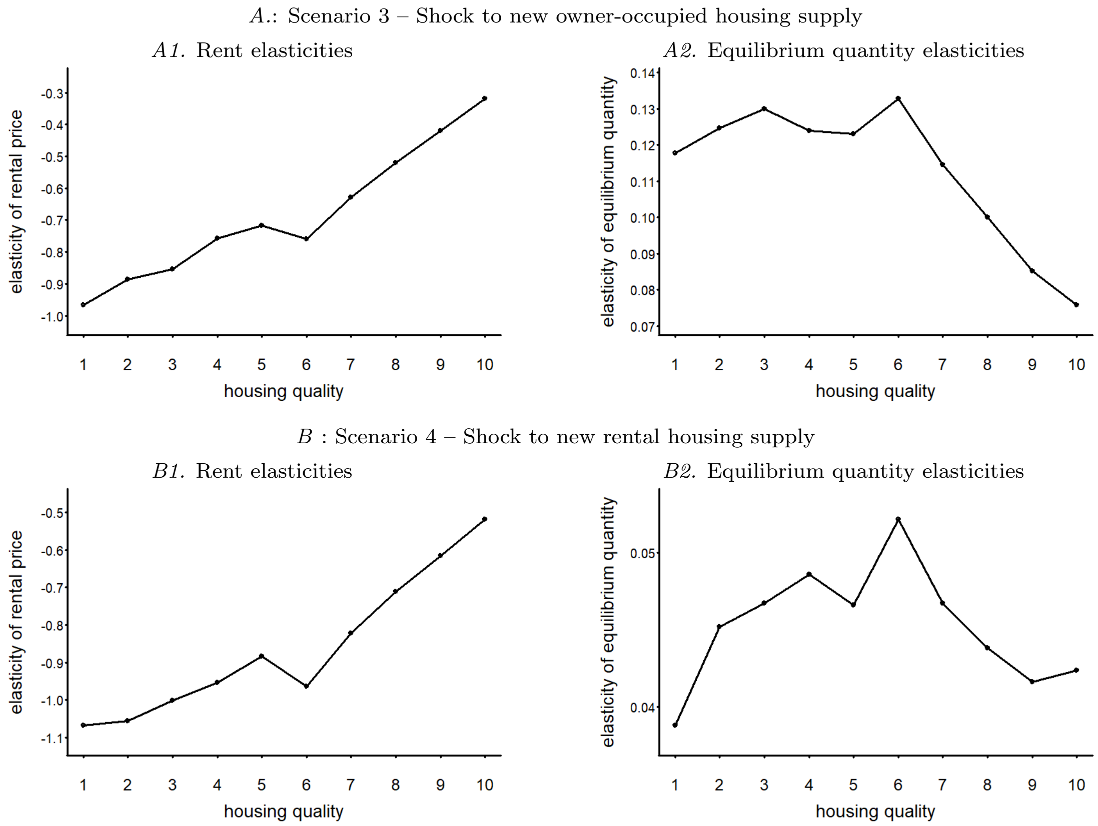
Figure O-C3 displays results analogous to Figure 4 in the main text for the case when there is a shock to new supply of owner-occupied housing and the supply of existing owner-occupied housing is fully elastic (Scenario 3), and when there is a shock to new rental housing supply, and the supply of new and existing owner-occupied housing is fully elastic (Scenario 4).

Figure O-C2: Simulation data: histograms



Notes: The graphs display histograms for the simulation sample.

Figure O-C3: New housing supply: price and quantity elasticities by housing quality bin



Notes: Panel *A* display the impact of a shock to new owner-occupied housing supply on rental prices and equilibrium quantities traded, aggregated by housing quality bins, represented as an elasticity. In this case, the supply of existing owner-occupied housing is fully elastic In Panel *B*, the supply shock is to new rental housing. In this case, the supply of new and existing owner-occupied housing are fully elastic.