Crowding (at) the Margins: Investigating the Unequal Distribution of Housing Space in Germany *

SEBASTIAN KOHL†
MAX F. STEINHARDT‡
LUCA STELLA§
SIMON VOSS¶

November 21, 2024

Abstract

This study explores the factors influencing household overcrowding using longitudinal survey data from Germany spanning the years 1985 to 2022. As average square meters per capita have declined for urban tenants, we find that overcrowding rates have substantially increased since 2012: By 2022, 11% of the population lived in overcrowded housing (Eurostat definition), while up to 19% of individuals subjectively felt overcrowded. At the same time, under-occupation also rose, with 39% of dwellings objectively classified as under-occupied, and 16% of residents subjectively perceiving their homes as under-occupied. We demonstrate that the likelihood of entering, experiencing, and remaining in overcrowded housing increases in early adulthood and decreases over the life cycle. Moreover, we find that, after controlling for socio-demographic characteristics such as the number of children or a migration background, economic factors contribute relatively little to explaining the likelihood of living in an overcrowded household. In policy terms, our paper highlights a misallocation of housing space and the need for housing policies to target particular vulnerable groups at high risk of overcrowding.

JEL Codes: D31, I32, J13, R21, R31

Keywords: Housing affordability, Living space, Inequality, Germany, Overcrowding, Well-being

*Acknowledgments: We would like to thank Dirk Löhr, Stijn Van Nieuwerburgh, Alkis Otto, Thomas Theobald, and Matthias Waltersbacher; and participants at the International SOEP User Conference, the European Network for Housing Research (ENHR) Conference, and the North American Meeting of the Urban Economics Association for valuable comments and feedback. We gratefully acknowledge funding from the Hans Böckler Foundation.

†Freie Universität Berlin, Email: sebastian.kohl@fu-berlin.de

‡Freie Universität Berlin, IZA, LdA, Email: max.steinhardt@fu-berlin.de

§University of Milan, Freie Universität Berlin, CESifo, IZA, Email: luca.stella@fu-berlin.de

¶Freie Universität Berlin, Email: simon.voss@fu-berlin.de

1 Introduction

In recent years, housing shortages and "shrinking homes" have become a major problem in many Western countries, driven by the affordability crisis (Gabor & Kohl 2022, Hubbard 2024) and further exacerbated during the COVID-19 pandemic by the shift of work and schooling to households. Square meters per capita have declined for tenants in many cities. According to Eurostat data, 16.8% of the EU population lived in overcrowded households in 2022, a trend that is growing in most Western countries, while nearly 34% of the population lived in under-occupied housing arrangements (Eurostat 2023). In some Eastern European countries, more than every second person lives in an overcrowded household. This issue not only challenges the United Nations' definition of the right to housing, which includes the requirement of 'adequate space', but also has far-reaching consequences on health, life satisfaction, family stability, and educational outcomes (Bujard et al. 2021, Wang & Liu 2023). It particularly concerns migrant populations (Huang & Yi 2015, Robinson et al. 2024). As a result, housing overcrowding represents a broader societal problem that extends well beyond housing.

It is important, therefore, to investigate the relatively under-researched factors contributing to overcrowding and the unequal endowment of households with housing space. To this end, we focus on the German case, as it represents a typical example with EU-average levels of under-occupation and slightly below-average rates of overcrowding (see Figure A.1 in the Appendix). Germany also provides an interesting context in which to study the determinants of overcrowding because it has the largest proportion of renters among EU countries (52.4% in 2023 relative to 30.8% in the EU)[1], and because its housing market regulations, although more stringent compared to most European countries, are relatively moderate on a global scale (see, for example, Kholodilin & Kohl 2021). Following decades of steady growth in per-capita housing space in Western Germany, and after 1990 also in the Eastern parts, urban households, particularly renters, have

[1]For details, see Eurostat (2014), available at this link.

seen a decline since 2010 (Deschermeier & Henger 2015). Since housing expenditures as a share of income and real estate property are unequally distributed in the German population, the issue of housing space is closely related to broader societal inequality of income and wealth (Pfeffer & Waitkus 2021). In particular, the proportion of income spent on housing has increased for those at the lower end of the income distribution, whereas it has declined for those at the upper end (Dustmann et al. 2022). Although prior empirical studies have shown that income, migration, and regional factors play a role in affecting housing overcrowding (Baldenius et al. 2019, Myers & Lee 1996, Painter & Yu 2010, Waltersbacher 2022), to the best of our knowledge, a comprehensive analysis that systematically examines the different driving forces using longitudinal micro data is still missing.

To address this research gap, we use data from the German Socio-Economic Panel (SOEP) which tracks more than 20,000 German households over the period 1985-2022. The results of our analysis indicate that overcrowding rates have substantially increased since 2011, up to 11% in 2022, with subjectively reported overcrowding affecting almost 19% of respondents in 2022. Considering under-occupation, which refers to a situation in which individuals have more housing space than is considered adequate, we observe an increase from about 33% to 39% over the period 1998[2] to 2022, whereas the respondents who considered their apartment too large increased from approximately 10% to 16% over the same period.

This reveals three interesting patterns: The number of people with excess housing space far exceeds those with too little; only a minority of households in under-occupied homes are aware of this situation; and the rising trend in overcrowding is paradoxically accompanied by an increase in under-occupancy.

We also find that the likelihood of overcrowding varies significantly by household characteristics. Our longitudinal analysis shows that the probability to enter, live and

[2]1998 is the first year in our observation period for which the subjective evaluation of dwelling size is available.

remain in overcrowded housing rises in early adulthood and decreases over the life cycle, with basic socio-demographics (i.e., being married, single-parent, immigrant, having children, with low education) explaining large parts of the total variation. Economic factors, such as low household income and tenancy, further increase the likelihood of overcrowding. These results are robust to an extensive set of tests, including different samples, alternative metrics for the outcome of interest, as well as the inclusion of individual fixed effects. Overall, our findings highlight the need for policymakers to implement nuanced housing policies that address both ends of the spectrum, i.e., implementing policies that provide affordable housing for young families from vulnerable groups, while also promoting better use of under-occupied homes. In this latter case, a policy example may include measures aimed at reducing the misallocation of housing space, such as providing incentives for individuals living in under-occupied housing to exchange their large apartments with families residing in overcrowded homes.

The paper is structured as follows: After reviewing the literature in Section 2, we present the data and stylized facts in Section 3. In Section 4, we describe the model specification, discuss our main results, and study the dynamics of overcrowding. Section 5 concludes and highlights the potential policy implications of our findings.

2 Previous Research

As a cursory Web of Science bibliographic search reveals, the study of overcrowding phenomena is most pronounced for hospitals or emergency units, transportation/mobility, prisons, or animals. Studies on housing overcrowding are comparatively less numerous, despite overcrowding having figured among the key concerns of historical reforms since the 19th-century "housing question", while also being a present housing concern.

Among the housing overcrowding studies, comparisons between different countries suggest that the distribution and availability of housing space strongly depend on the specific housing market characteristics as well as a country's demographic conditions, such as the age composition of a society (Schwartz & Seabrooke 2008). Therefore, country-specific evidence appears essential. However, only few studies focus on the German context which, despite its high tenancy rates, broadly falls in the middle of the European distributions of overcrowding and under-occupation (see Figure A.1 in Appendix A).

In Germany, the average size of disposable housing space has increased significantly since reunification in 1990. There has been an especially strong increase of housing space in the regions of former East Germany, where average apartment size has made up much ground to the numbers in the West, while still falling behind substantially (Frick & Grimm 2009, Deschermeier & Henger 2015). However, as more recent studies show, this trend has strongly decelerated in the years since 2010, with per-capita housing space even decreasing slightly between 2014 and 2018 (Waltersbacher 2022). This is also reflected in the trend of the overcrowding rate, which started rising roughly around this time after a decade-long decrease (see Figure 1).

Prior works have tried to disentangle age effects from those related to specific birth cohorts. According to an analysis by the German Economic Institute (Deschermeier & Henger 2015), the per-capita housing space in Germany increased annually by 1.15% between 1991 and 2013, of which 0.93% can be attributed to a cohort and 0.22% to an age effect. The cohort effect seems to be especially relevant for the described catching-up process of former East Germany, where, according to the study, its size of 1.46% was much larger than in the West. Furthermore, studying a sample of European countries, Sunega & Lux (2016) find that the overcrowding rates according to objective measures differ substantially from subjective perceptions of housing space adequacy by the household members. This suggests that an analysis using objective measures for overcrowding

should be accompanied by one that takes the households' own evaluations into account.

Previous studies highlight substantial differences in housing space across dimensions such as household age, income, city size, and tenure status (Waltersbacher 2022). Panel studies trace the rise in individual housing space to higher demand driven by changing household compositions. As people age, higher incomes often enable larger housing, and the 'empty-nest effect'—where household members remain in the same dwelling after children move out—leads to increased per-capita space (Henger et al. 2014, Siedentop et al. 2006).

In recent research, Waltersbacher (2022) examines the factors affecting per-capita housing space in Germany in 2018, finding that new rental price hikes, especially in big cities, contribute significantly to housing shortages. This trend reflects the growing demand for urban living, while rural areas often see a housing surplus. With rising prices, households with multiple children, migrants, and single earners often must choose between undersized units or relocating to suburbs (Gränitz 2022, Auspurg et al. 2017).

On average, homeowners have more space than renters—who make up about 55% of German households and more than 70% in cities with over 500,000 residents (Kohl et al. 2019, Dustmann et al. 2022). From 2006 to 2018, homeowners' housing space grew by 7.1%, compared to only 1.7% for renters (Waltersbacher 2022). Misallocated housing space, as seen in the empty-nest effect, compounds this issue: According to a report by the German Economic Institute, the number of overcrowded households nearly matches that of under-occupied ones (Sagner & Voigtländer 2023).

Going beyond Germany, previous research has identified several social groups particularly affected by overcrowding. For example, immigrants in the US typically have less housing space, and US counties with higher shares of immigrants, hispanic renters, or poor families with children experience higher overcrowding rates (Painter & Yu 2010, Myers et al. 1996, Huang & Yi 2015, Robinson et al. 2024). Further, housing space

per capita typically increases with longer stays in the host country (Myers & Lee 1996, Painter & Yu 2010). Additionally, overcrowding is more prevalent among renters, especially in urban areas (Gabor & Kohl 2022).

Previous studies have examined not only the determinants but also the potentially detrimental consequences of overcrowding, particularly with respect to childhood development. Research shows that children in overcrowded households face disadvantages in physical development and overall health (Booth & Johnson 1975, Richter 1989). Overcrowding is also strongly linked to psychological stress, family instability, poor schooling outcomes, and reduced well-being of children (Gómez-Jacinto & Hombrados-Mendieta 2002, Solari & Mare 2012, Lopoo & London 2016). These findings have been confirmed in other countries, showing negative impacts on physical health (Pevalin et al. 2008, Nkosi et al. 2019), schooling (Goux & Maurin 2005, Contreras et al. 2019) or social life (Hadziabdic & Kohl 2024). Historical studies, such as one comparing childhood mortality in Glasgow and Edinburgh from 1911 to 1961, highlight overcrowding as a significant factor (Cage & Foster 2002). Furthermore, overcrowded conditions in US immigrant housing during the 1920s and 1930s contributed to the spread of contagious diseases (Ager et al. 2020). Many studies also link overcrowded housing to negative mental health outcomes (Giel & Ormel 1977, Schwab et al. 1979, Duckitt 1983, Wells & Harris 2007, Evans et al. 2016, Foye 2016, Pierse et al. 2016), with recent research on Covid-19 lockdowns identifying evidence for increased domestic violence (Perez-Vincent et al. 2020, Arenas-Arroyo et al. 2021).

To the best of our knowledge, there are no studies so far that address the determinants of overcrowding using multiple regression techniques and making use of both cross-sectional and longitudinal analysis. We aim to fill this research gap by conducting a thorough analysis of the SOEP dataset, making use of the richness of information it provides and explicitly making use of its longitudinal dimension.

3 Data and Stylized Facts

The data for our analysis comes from the German Socio-Economic Panel (SOEP). Specifically, our study uses version 39 of the SOEP, covering the years 1984-2022 (SOEP 2024). The SOEP is a representative longitudinal dataset that has surveyed households and individuals in Germany since 1984. For a detailed description of the survey, refer to Goebel et al. (2019) and Schröder et al. (2020). One major advantage of the SOEP is that it contains a wide range of individual- and household-level information, making it an ideal source for investigating the potential determinants of housing overcrowding. A second reason for using the SOEP is that, in addition to a range of individual and household characteristics, the survey contains detailed self-reported information on the quantity and quality metrics of housing overcrowding. Since 1998, the respondents have been asked the following questions: "What do you think about the total size of your dwelling?". They respond using a scale ranging from "much too small", to "a bit too small", to "just right", to "a bit too large", and to "much too large". We create a binary variable for subjective overcrowding equal to one if the size of the dwelling is considered "much too small" or "a bit too small".

For the objective measure of overcrowding, we follow the definition used by Eurostat (Eurostat 2019). Specifically, a dwelling is considered overcrowded when the persons living there do not have enough rooms for the corresponding size of the household (any space with an area greater than 6 m^2 is classified as a room). This is true if the number of rooms is lower than the sum of one room for the household, one room for each couple, one for each single adult person, each pair of children, each pair of teenagers of the same gender, and each single teenager. This implies that even single-person households can be considered overcrowded if they have only one room larger than 6m^2 available to them. Similarly, we follow Eurostat's definition for under-occupation (Eurostat 2021). A dwelling is considered under-occupied when the persons living there have *more* rooms at their disposal than defined above. To not rely on the room and crowding defini-

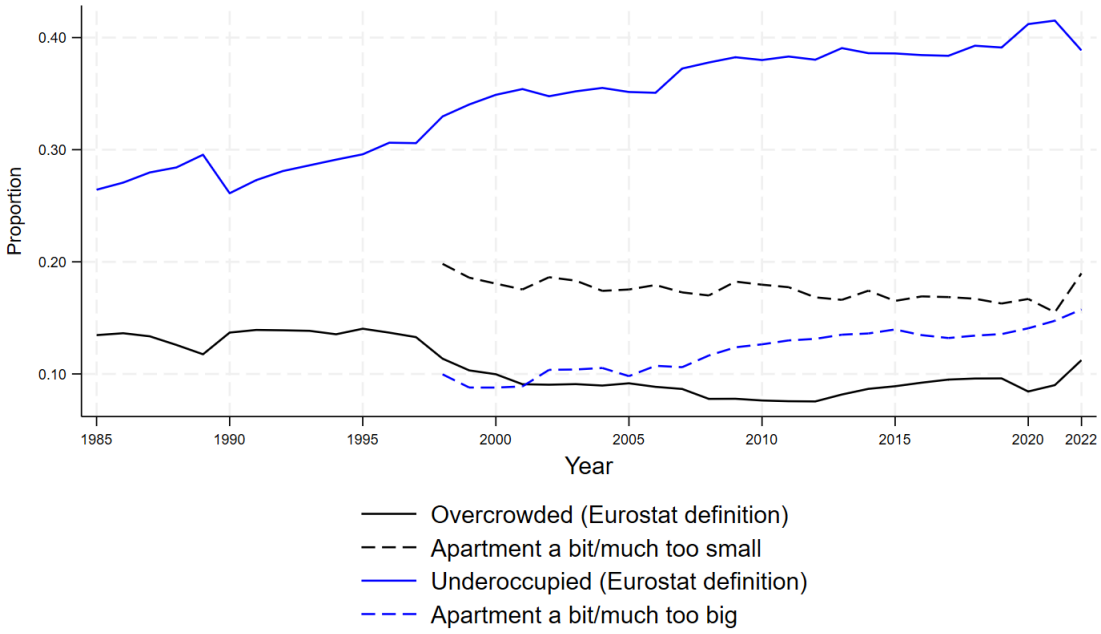
tion alone, we also work with the square-meter-per-capita provision as an alternative. Following the definition used by Eurostat (Eurostat 2024), we define housing cost burden as the proportion of total housing costs (excluding housing allowances) relative to the annual disposable household income (also excluding housing allowances). Because some components of total housing costs are missing in certain SOEP survey years, our definition includes only rent and heating expenses for tenants, and maintenance, mortgage interest, and mortgage principal payments for homeowners, thereby differing from Eurostat's measure, which includes additional components such as expenses on electricity and utilities. Finally, we deviate from Eurostat's definition by including the sum of mortgage interest and mortgage principal payment–instead of mortgage interest alone–since the SOEP does not contain separate information on mortgage interest payment only (Lozano Alcántara & Romeu Gordo 2020).

Finally, the SOEP allows us to exploit the longitudinal dimension of the data by controlling for time-invariant individual characteristics. In our main sample, we include data from 1985 to 2022, focusing on individuals of working age, specifically those between 18 and 64 years old at the time of the interview, to avoid changes in behavior due to retirement.[3] After these restrictions, we obtain a final longitudinal sample that contains 550,848 person-year observations resulting from 86,264 individuals. Descriptive statistics for our final analysis dataset are reported in Table A.1 in Appendix A. Figure 1 looks at how both objective and subjective metrics of overcrowding have evolved over the period 1985-2022. The share of respondents in objectively overcrowded dwellings has strongly increased since 2011, after a decade-long decline. While fluctuating considerably over the past few years, the rate of overcrowding has increased from approximately 8% in 2012 to approximately 11% in 2022, underscoring the relevance of the issue. By 2022, 11% of individuals lived in overcrowded households, while 39% lived in house-

[3]In a robustness analysis, we use the survey years 2000-2022 and consider alternative age groups (see Tables B.1 and B.2 in Appendix B).

holds considered under-occupied. [4] This pattern is remarkable, since the number of persons with too much housing space far exceeds that of persons with too little space. This suggests that the misallocation of the housing space is a substantial part of the problem.

Figure 1: Share of Different Housing Conditions over Years



Notes - Data are drawn from the SOEP version 39. Survey years: 1985-2022.

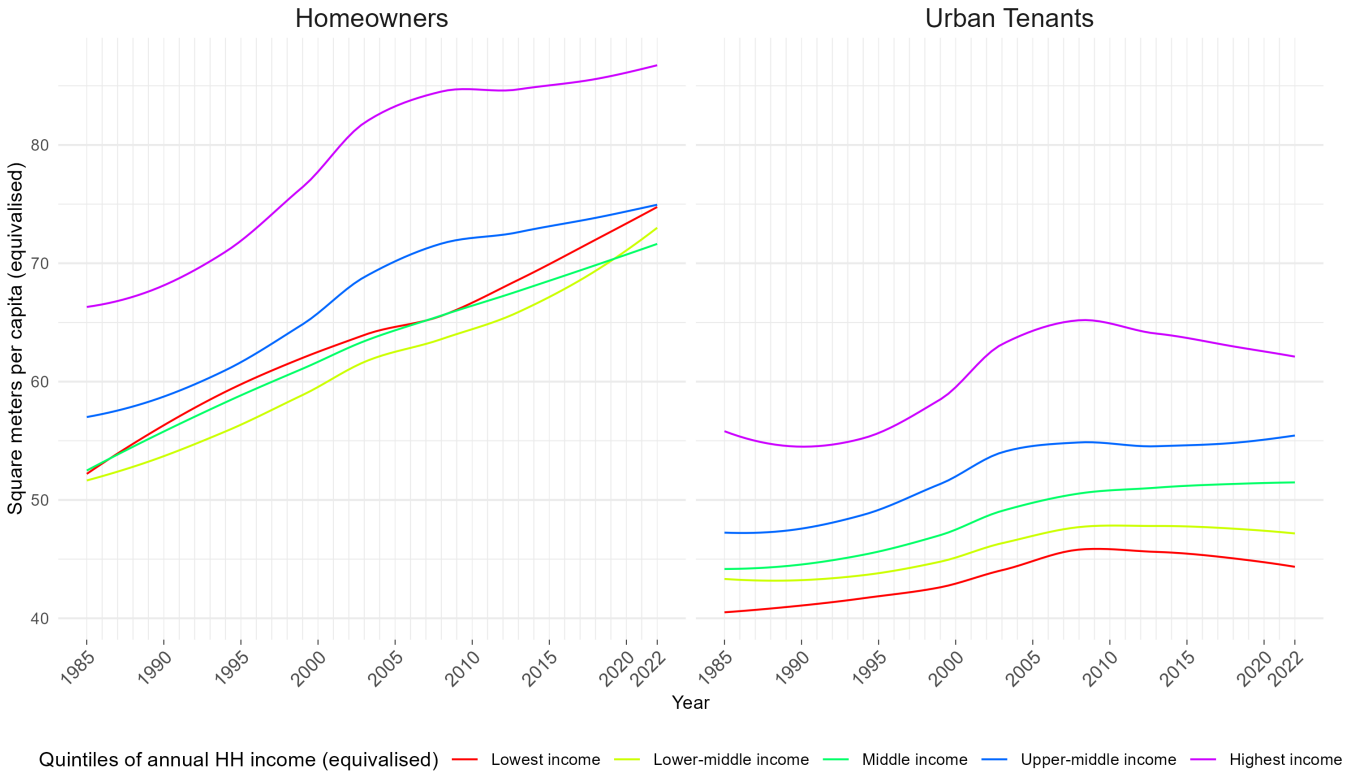
In addition, Figure 1 shows the development of subjective housing conditions over time. This reveals two further intriguing stylized facts: First, the share of subjective overcrowding generally exceeds its objective counterpart, particularly following a divergence after 2005 and parallel growth after 2011. By 2022, the gap lies at approximately 8 percentage points (19% vs. 11%). More respondents thus believe to live in an overcrowded dwelling than is actually the case according to the objective Eurostat measure. Second, for under-occupation the pattern is reversed, more pronounced than for overcrowding,

[4] The spike in objective overcrowding in 2022 is largely driven by a significantly higher proportion of cohabiting pairs of individuals who are not a couple. Additionally, part of the increase may be due to the fact that, for 2022, the SOEP currently provides only preliminary survey weights.

and showing an increasing gap over time. While 39% live in an under-occupied dwelling, only 16% of respondents consider their apartment too large by 2022. This implies that only a minority of households who live in an under-occupied dwelling are perceiving their situation in these terms. Further, Figures A.2 and A.3 indicate that overcrowding is more common in urban than in rural areas (13% vs. 8%), whereas under-occupation follows the opposite trend, being higher in rural (46%) than in urban areas (35%). This mirrors the political, public, and academic debates on housing, which focus on housing scarcity and small apartments, while under-occupation receives much less attention (Gränitz 2022). A media analysis of the main German newspapers confirms that the salience of housing topics has seen a pronounced increase over the last two decades, but with a focus on shortages, not under-occupation or vacancies (see Figure A.4 in Appendix A).

Overcrowding has not only increased when looking at the provision with rooms, but is also reflected in the per capita square meters. Figure 2 shows this metric for homeowners and–as a particular risk group in rental Germany–urban tenants on the left- and right-panel respectively, each time broken down by household-income quintiles since 1985. While homeowners have seen higher levels, but particularly steady increases of per-capita home sizes across income groups, urban tenants–following a catch-up in the post-unification years–were exposed to stagnating or even declining rates of square-meter provision, particularly among the bottom 40% and top 20% households. Different household classes are thus quite heterogeneous with regard to both levels and recent changes in the provision of floor area per capita.

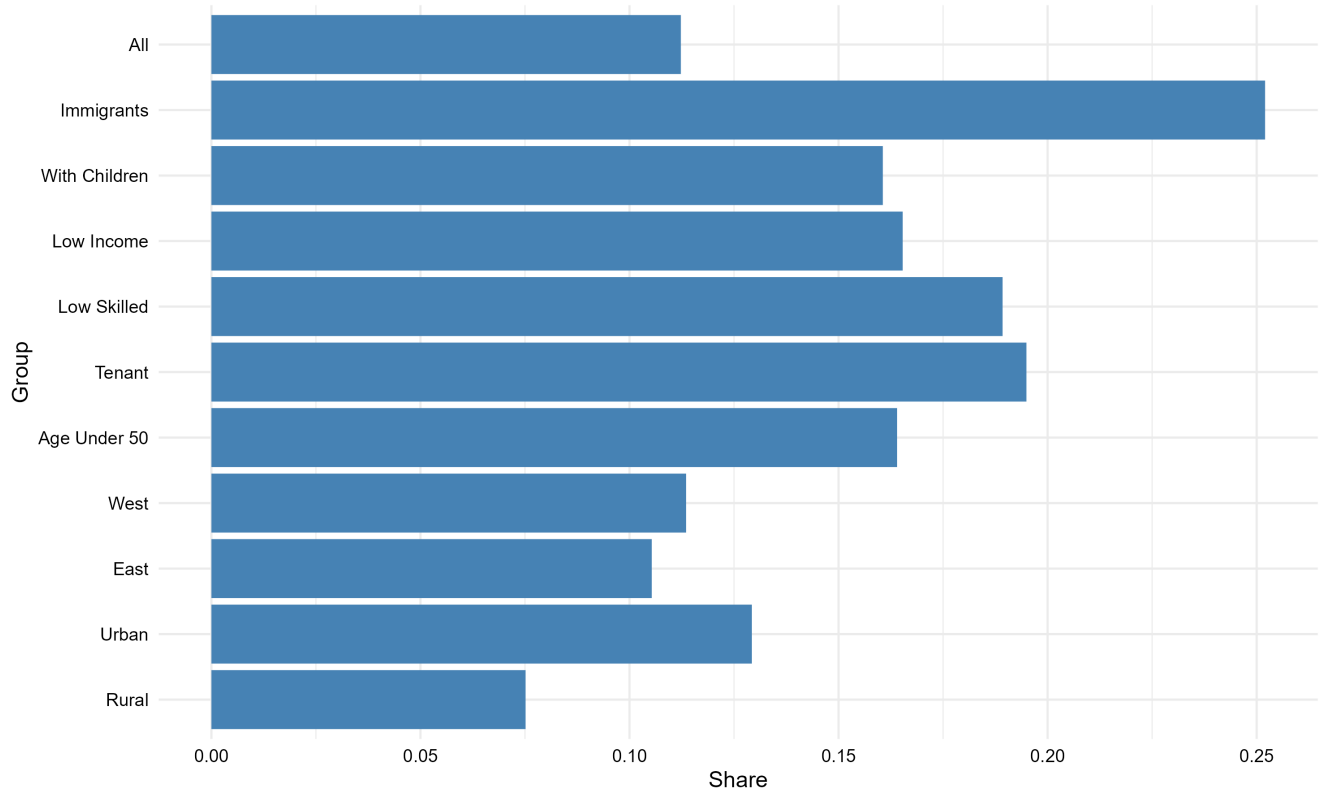
Figure 2: Square Meters per Capita by Income and Over Time



Notes - Data are drawn from the SOEP version 39. Survey years: 1985-2022. Square meters per capita and household income have been adjusted using OECD equivalized income weights.

To explore this heterogeneity further, Figure 3 shows the proportion of individuals living in overcrowded households by different groups, along socioeconomic and regional characteristics. This provides first suggestive evidence regarding which groups are most vulnerable to overcrowding. The likelihood of living in an overcrowded household is notably highest among immigrants. Higher overcrowding rates are also observed in households with children, renters, individuals with low skill levels (e.g., those without a high school diploma), low-income individuals, and people under 50 years old. However, the differences in overcrowding rates for these groups, compared to the overall population, are less pronounced.

Figure 3: Overcrowding – by Selected Characteristics in 2022



Notes - Data are drawn from the SOEP version 39. Survey year: 2022. Low income is defined as the lowest 40% in income distribution.

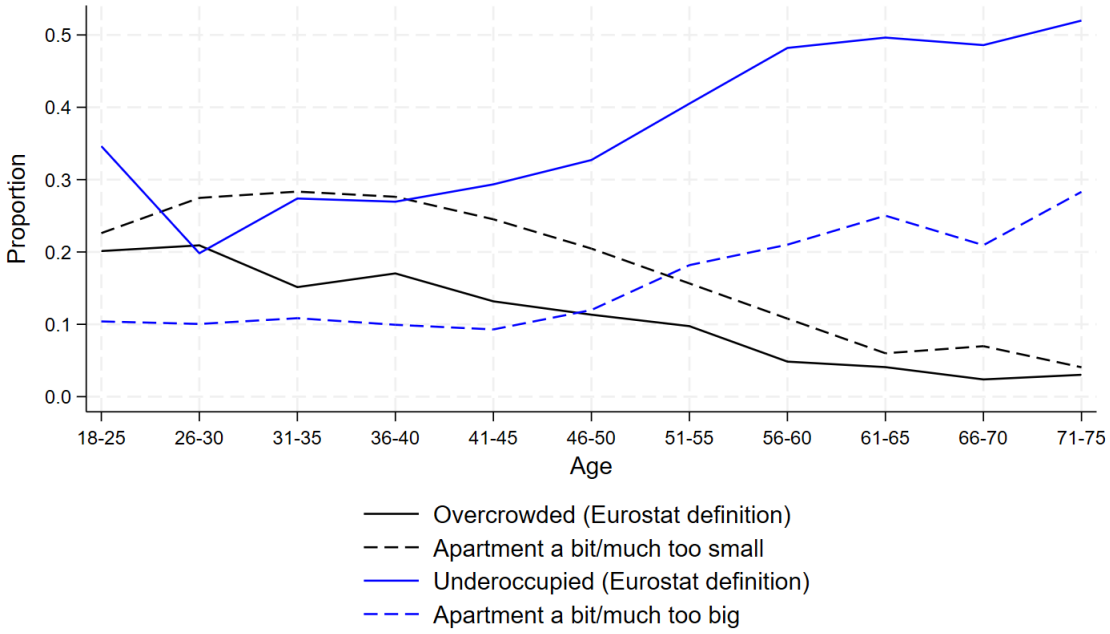
Figure 3 also breaks down the overcrowding rate by regional characteristics, revealing a clear disparity. On average, individuals in Western Germany are more likely to live in overcrowded households compared to those in the Eastern regions, which contradicts the generally higher overcrowding patterns in Eastern vis-à-vis Western Europe. Additionally, the proportion of people living in overcrowded conditions is significantly higher in urban areas than in rural regions. This is in line with findings from previous research (e.g. Waltersbacher 2022, Gabor & Kohl 2022) and highlights the necessity to consider regional factors in addition to individual characteristics when studying the determinants of overcrowding.

Finally, in Figure 4, we zoom in on the age-variable and entire life cycle by plotting

the rate of individuals in overcrowded as well as underoccupied dwellings against different age groups, both for the objective and subjective overcrowding measures.[5] We observe that the proportion of people living in overcrowded conditions decreases over the life cycle, starting at over 20% for the 18-25 age group. Thus, the lowest share of respondents living in overcrowded conditions is found in the oldest age groups, 66-70 and 71-75, where only between approximately 2% and 3% experience inadequate housing space. For the subjective measure of overcrowding, the pattern diverges somewhat: The proportion of individuals aged 18 to 25 who felt their dwelling is too small starts at a level comparable to the objective measure, but rises significantly for older age groups. After the 36-40 age group, the rate of subjective overcrowding follows a similar pattern as the objective measure, decreasing steadily with age. On the other hand, the share of individuals living in under-occupied housing (after an initial decline) increases significantly over the life cycle, highlighting the previously mentioned empty-nest syndrome. Additionally, there is a considerable gap between the objective and subjective measures, suggesting that many individuals with more than adequate housing space do not perceive their situation as such. The figure provides initial suggestive evidence that typical life-cycle events, such as having children in midlife or the death of a partner in later life, influence the prevalence of overcrowding. We will address this aspect explicitly in the following section.

[5]However, it is important to note that the figure presents only the average trends in overcrowding and does not differentiate between age effects and potential cohort effects.

Figure 4: Overcrowding-Age Profiles in 2022



Notes - Data are drawn from the SOEP version 39. Individuals aged 18-75. Survey year: 2022.

4 Empirical Analysis

4.1 Model Specification

To examine the relationship between housing overcrowding and socio-economic factors, we employ two complementary empirical approaches. First, we estimate a linear probability model. Therefore, we consider the following equation:

$$Y_{iht} = \alpha + \beta X_{iht} + \kappa_s + \tau_t + \epsilon_{iht} \tag{1}$$

where the index iht denotes an individual i, who belongs to household h, and was interviewed in year t. Y_{iht} denotes the outcome variable of interest: a binary variable coded as one if individual i lives in an overcrowded household h at time t. The variable

X_{iht} is a vector containing a wide range of individual-level socio-demographic, economic and regional characteristics, which are added in a stepwise fashion in our regressions. Specifically, regarding socio-demographic characteristics, we include gender, age and age squared, an indicator for migration background, marital status, the number of children, a dummy for living in a single-parent household and dummy variables for the individual's education (i.e., indicators for having a high-school education or a college-level education). Concerning economic characteristics, we use a dummy for excessive housing burden (that is, paying more 40% of household income on housing costs), the logarithm of net household monthly income, and a dummy for being a homeowner. To capture regional differences, we include a binary indicator for residing in an urban area, as well as a separate dummy variable representing whether the individual lives in a house rather than an apartment. Moreover, we include federal state fixed effects (κ_s) that capture unobservable, time-invariant differences across states (Länder) that may influence housing overcrowding. Such factors might include, for example, state-level differences in preferences and attitudes regarding housing due to discrepancies in cultural and institutional backgrounds as well as cross-state differences in housing market regulations. The model also contains survey year fixed effects (τ_t) to account for possible trends in housing behavior. Finally, ϵ_{iht} represents a disturbance term. To account for within-family correlation, we cluster standard errors at the household level for all estimates.

We then complement the analysis conducted using the linear probability model with a hazard model to better understand the dynamics of overcrowding over the life cycle. In particular, we aim to estimate the instantaneous risk or likelihood of an individual experiencing their first episode of entering or leaving overcrowding. Formally, we estimate the following model:

$$h_{iht}^{IN/OUT} = \lambda(t)\phi(\delta X_{iht} + v_{iht}) \tag{2}$$

where the dependent variable $h_{iht}^{IN/OUT}$ is the hazard that individual i enters (or leaves) an overcrowded household h in year t. The term $\lambda(t)$ represents the baseline hazard function. This function captures the time dependence of the transition into the state, and it is modeled using a flexible piecewise constant function. As for the hazard function ϕ, our preferred specification uses a logistic regression. X_{iht} and the error term v_{iht} are defined in the same way as in Equation (1). All estimates are weighted using survey weights.

4.2 Main Results

The estimated coefficients of the model described in Equation (1) are reported in Table 1. We add our control variables in a stepwise fashion, starting with variables reflecting socio-demographic characteristics (see columns 1 to 7), followed by economic characteristics (see columns 8 to 10), and finally regional characteristics (see columns 11 to 13). The purpose of this stepwise procedure is to first estimate the overall influence of socio-demographic characteristics on the likelihood of housing overcrowding, and then, in subsequent steps, to assess how these effects change as additional factors are included. When considering all socio-demographic determinants (see column 7), we find that the likelihood to live in an overcrowded household strongly differs by immigrant status and household composition. Namely, being an immigrant is on average associated with a 13.8 percentage points higher likelihood to live in an overcrowded household. While married persons are 8.5 percentage points less likely to live in overcrowded conditions, having one more child in the household is associated with a 5 percentage points higher likelihood of overcrowding. Moreover, individuals in single-parent households are especially at risk, with a 9.1 percentage points higher likelihood of living in overcrowded conditions than other households (when the other socio-demographic factors are held constant). Our results also document that education is strongly related to overcrowding: the higher the level of education the lower is the risk of living in an overcrowded house-

hold. For example, college graduates experience a lower probability of overcrowding by 7.6 percentage points compared to individuals with less than a high-school diploma.

Regarding gender, we initially find small differences between women and men which, however, increase in size after adding controls such as marital status and education. Overall, our results suggest that women are slightly less likely to live in overcrowded households than men.

Table 1: Potential Determinants of Overcrowding – 1985-2022

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Dep. var.: Overcrowding			**Socio-demographic characteristics**					**Economic characteristics**			**Regional characteristics**		
Female	-0.006**	-0.004	-0.005**	-0.005**	-0.008***	-0.013***	-0.017***	-0.017***	-0.019***	-0.020***	-0.020***	-0.020***	-0.020***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Age		-0.003***	-0.004***	0.001	-0.004***	-0.005***	-0.001	-0.001	-0.002***	-0.004***	-0.004***	-0.004***	-0.005***
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Age^2		-0.000	0.000	-0.000***	0.000***	0.000***	-0.000	-0.000	0.000	0.000***	0.000***	0.000***	0.000***
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Immigrant			0.152***	0.161***	0.150***	0.149***	0.138***	0.139***	0.127***	0.111***	0.111***	0.110***	0.103***
			(0.007)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Married				-0.070***	-0.105***	-0.086***	-0.085***	-0.086***	-0.054***	-0.044***	-0.043***	-0.042***	-0.039***
				(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Children					0.055***	0.051***	0.050***	0.049***	0.053***	0.055***	0.055***	0.055***	0.057***
					(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Single parent						0.096***	0.091***	0.091***	0.086***	0.086***	0.086***	0.087***	0.088***
						(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.008)
Medium educated							-0.055***	-0.055***	-0.048***	-0.045***	-0.046***	-0.047***	-0.047***
							(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
High educated							-0.076***	-0.077***	-0.048***	-0.050***	-0.053***	-0.056***	-0.057***
							(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Excessive housing burden								-0.001	-0.031***	-0.025***	-0.026***	-0.027***	-0.026***
								(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Monthly HH income									-0.085***	-0.063***	-0.063***	-0.065***	-0.061***
									(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Homeownership										-0.077***	-0.073***	-0.070***	-0.041***
										(0.004)	(0.004)	(0.004)	(0.004)
Urban area												0.031***	0.023***
												(0.005)	(0.005)
Dwelling: House													-0.058***
													(0.005)
Observations	550,848	550,848	550,848	550,848	550,848	550,848	550,848	550,848	550,848	550,848	550,848	550,848	550,848
R-squared	0.005	0.027	0.053	0.063	0.083	0.089	0.093	0.094	0.112	0.125	0.127	0.129	0.134
Mean of Dep. Var.	0.106	0.106	0.106	0.106	0.106	0.106	0.106	0.106	0.106	0.106	0.106	0.106	0.106
Std.Err. of Dep. Var.	0.308	0.308	0.308	0.308	0.308	0.308	0.308	0.308	0.308	0.308	0.308	0.308	0.308

Notes - Data are drawn from the SOEP version 39. Standard errors are reported in parentheses and are clustered at the household level. Individuals aged 18-64. All specifications include year FE. In columns 11 to 13 state FE are included. All estimates are weighted using survey weights.
*Significant at 10 per cent; ** Significant at 5 per cent; ***Significant at 1 per cent.

When we also take into account economic characteristics (see columns 8 to 10), we find that paying an excessive share of household income on housing costs is not significantly related to housing overcrowding. This might partly be driven by the fact that not all components of housing costs are included in our measure (see Section 3). We also find that even after controlling for basic demographic characteristics and migrant status,

as well as excessive housing burden, a higher monthly household income is associated with a lower likelihood of living in overcrowded conditions. Finally, homeownership, too, is significantly related to housing overcrowding. Our results imply that owning one's apartment or house is associated with a 7.7 percentage points lower likelihood of living in overcrowded conditions.

Finally, we include a set of regional variables in columns 11 to 13. We start with adding state fixed effects in column 11, which leaves most coefficients largely unchanged. This implies that controlling for state differences in overcrowding patterns does not alter our results in any meaningful way. However, as reported in columns 12 and 13, overcrowding appears to be related to both the type of geographic area and the type of dwelling a household is living in. Specifically, while living in an urban area is associated with a 3.1 percentage points higher likelihood of overcrowding, households living in a house are 5.8 percentage points less likely to be overcrowded compared to those living in an apartment, when all the other factors are held constant.

Overall, we find that including economic as well as regional characteristics to our estimations does not add a lot of explanatory power to our model. This highlights that socio-demographic factors, which are of course connected to economic and regional characteristics, are one important determinant of overcrowding. Of particular note is our discovery that immigrants, single parents, and families with children are disproportionately impacted by overcrowding, even when accounting for variables such as education, income, regional location, and homeownership.[6] The remaining significant and large gap between natives and migrants is indicative of ethnic discrimination in the housing market, which has been documented in various studies within the German context (see, for example, Diehl et al. 2013, Auspurg et al. 2017).

Our results also help to quantify the risk of overcrowding and identify high-risk groups. For example, our results indicate that a low-skilled migrant mother with two

[6]In Germany, immigrants are less likely to own a home, a pattern also observed in the US context (Luik et al. 2023).

children, renting an apartment in an urban area, faces a high risk of experiencing overcrowding, with a predicted probability of approximately 40.7%. In contrast, high-skilled native-born renting women without children living in apartments face a much lower risk of overcrowded conditions, even in urban areas, with a predicted probability of just 17.2%.[7] While these assessments should not be interpreted as causal, the estimates help identify vulnerable groups at high risk of experiencing overcrowding.

In Tables B.1 to B.5 in Appendix B, we demonstrate that the results are robust to an extensive set of sensitivity checks, such as changes to the sample, the inclusion of individual fixed effects, as well as alternative metrics for the outcome of interest including square meter per capita.

4.3 The Dynamics of Overcrowding

We begin our analysis by analyzing the potential determinants of transitioning into and out of overcrowded housing. For this purpose, we estimate Equation (1) using as the dependent variable a dummy coded as one if an individual has changed the overcrowding status–either entering or exiting overcrowding–at least once during the sample period. The corresponding results are reported in Table 2. As in Table 1, we proceed in a stepwise fashion, adding to our specification first the socio-demographic covariates and next the economic and regional determinants of overcrowding. The results remain substantially unchanged with respect to the direction, magnitude, and significance. Consistent with the results displayed in Table 1, we find that the likelihood of being a switcher is primarily driven by socio-demographic factors such as immigrant status, the number of children in the household, living in a single-parent household, and the education level. In comparison, regional and economic characteristics add only minimal explanatory power to our model.

[7] The calculation is based on average aged, married individuals not living in a single-parent household that do not face an excessive housing cost burden and have an average household income, living in North Rhine-Westphalia in 2022.

Table 2: Potential Determinants of Being a Switcher - 1985-2022

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Dep. var.: Being a switcher			**Socio-demographic characteristics**					**Economic characteristics**			**Regional characteristics**		
Female	0.017***	0.021***	0.019***	0.020***	0.017***	0.012***	0.006	0.006	0.003	0.001	0.001	0.001	0.001
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.004)
Age		0.001	-0.001	0.006***	0.001	0.000	0.006***	0.006***	0.004***	0.001	0.001	0.001	0.000
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Age^2		-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Immigrant			0.166***	0.179***	0.167***	0.166***	0.150***	0.149***	0.133***	0.107***	0.110***	0.108***	0.101***
			(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
Married				-0.100***	-0.139***	-0.122***	-0.121***	-0.119***	-0.076***	-0.059***	-0.060***	-0.059***	-0.056***
				(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Children					0.061***	0.058***	0.055***	0.055***	0.059***	0.063***	0.062***	0.062***	0.064***
					(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Single parent						0.086***	0.078***	0.078***	0.070***	0.071***	0.071***	0.072***	0.073***
						(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
Medium educated							-0.080***	-0.080***	-0.070***	-0.066***	-0.068***	-0.069***	-0.069***
							(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.008)
High educated							-0.120***	-0.119***	-0.080***	-0.083***	-0.089***	-0.092***	-0.092***
							(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
Excessive housing burden								0.034***	-0.008	0.002	0.004	0.002	0.003
								(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Monthly HH income									-0.116***	-0.078***	-0.076***	-0.077***	-0.073***
									(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Homeownership										-0.129***	-0.126***	-0.123***	-0.091***
										(0.008)	(0.008)	(0.008)	(0.009)
Urban area												0.032***	0.024**
												(0.010)	(0.010)
Dwelling: House													-0.063***
													(0.009)
Observations	550,848	550,848	550,848	550,848	550,848	550,848	550,848	550,848	550,848	550,848	550,848	550,848	550,848
R-squared	0.009	0.049	0.063	0.073	0.084	0.086	0.091	0.092	0.108	0.124	0.126	0.127	0.130
Mean of Dep. Var.	0.292	0.292	0.292	0.292	0.292	0.292	0.292	0.292	0.292	0.292	0.292	0.292	0.292
Std.Err. of Dep. Var.	0.455	0.455	0.455	0.455	0.455	0.455	0.455	0.455	0.455	0.455	0.455	0.455	0.455

Notes - Data are drawn from the SOEP version 39. Standard errors are reported in parentheses and are clustered at the household level. Individuals aged 18-64. All specifications include year FE. In columns 11 to 13 state FE are included. All estimates are weighted using survey weights.
*Significant at 10 per cent; ** Significant at 5 per cent; ***Significant at 1 per cent.

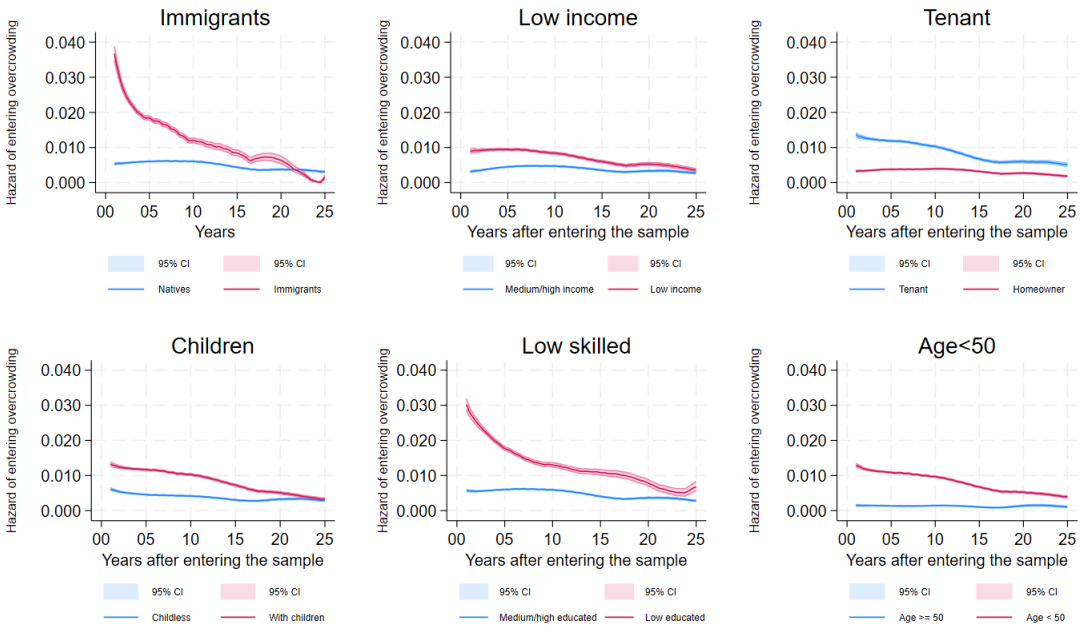
Next, we explicitly analyze the dynamics of overcrowding by showing the evolution of the hazard functions for both entering and exiting overcrowded conditions. Figure 5 illustrates the unconditional pattern of entering overcrowding for selected socio-economic groups. In general, this figure reveals several notable differences: a) in the first years after entering the sample, the risk of entering overcrowding among immigrants is considerably higher compared to that among natives;[8] b) families with children have significantly higher rates of entering overcrowding compared to those without children; c) low-income, low-educated individuals, tenants as well as younger people are at higher risk of entering overcrowding even as time passes, thereby providing further evidence of the long-term consequences of income, education, tenant status and age structure.

In the Appendix, we further illustrate the pattern of entering overcrowded conditions based on regional characteristics. Figure C.1 documents that in the first 15 years after entering the sample, individuals residing in East Germany face a higher risk of entering overcrowded conditions compared to those living in West Germany. In addition, in the first 20 years after entering the sample, individuals in urban areas face a higher risk of entering overcrowding compared to those residing in rural settings. Finally, those living in apartments have a higher risk of entering overcrowding compared to those living in houses.

We then estimate a discrete-time duration model for the hazards of individuals entering overcrowding as well as exiting overcrowding. Thus, each equation in Model (2) is estimated using a separate logistic hazard function. The results of this analysis are reported in Tables C.1 and C.2 in the Appendix. First, we look at the determinants for the hazard of entering overcrowding, shown in Table C.1. As in our main analysis, we follow a stepwise approach, adding to our specification first the socio-demographic, next the economic, and finally the regional characteristics. When looking at socio-demographic

[8]One may be concerned that selection explains the convergence in housing overcrowding between natives and immigrants. However, in the Appendix we replicate the main results including individual fixed effects in our analysis (see column 3 of Table B.4), at least partially mitigating the concern of a selection based on unobservable factors and a possible selective return migration.

Figure 5: Hazard of Entering Overcrowding - by Socioeconomic Group



Notes - Data are drawn from the SOEP version 39. Each figure shows the unconditional pattern of entering overcrowding. Individuals aged 18-64. Survey years: 1985-2022. Figures are displayed without survey weights in order to include confidence intervals. Results are very similar when survey weights are used.

characteristics only (see column 7), we find that immigrants as well as individuals with a higher number of children in the household and single parents are at a higher risk of entering overcrowding. At the same time, married individuals have a lower risk of entering overcrowding. These findings are robust to the inclusion of economic characteristics. A higher monthly income and homeownership lead to a reduction in the hazard of entering overcrowding (see column 10). When we consider the regional characteristics, we find that also living in apartments is associated with a higher risk of entering overcrowding, while residing in urban areas shows no significant association (see column 13). To see if the determinants for entering and leaving overcrowding are symmetrical, in Table C.2 in the Appendix we analyze the hazard rate of leaving overcrowding. Our estimation results regarding entering and leaving are fairly symmetrical, except for homeownership, which is associated with a lower likelihood of leaving overcrowding. Overall, the findings from the hazard model are in line with our main results (see Table 1).

5 Conclusions

Our study shows that housing overcrowding is a multifaceted issue that does not affect all individuals and households equally. The risk of living in an overcrowded household varies markedly depending on socio-demographic and economic factors, including place of birth, marital status, the number of children, education level, and income. Families with children, low income, and a migration background are among the groups at the highest risk of living in overcrowded conditions, even after accounting for regional factors. In contrast, high-skilled individuals without a migration background or children are much less likely to experience overcrowded housing. Furthermore, the hazard analysis has shown that socio-demographic changes, such as having children or changes in marital status, are key factors influencing the transition into overcrowded housing

conditions. A similar pattern applies to the transition out of overcrowding.

Our analysis also reveals that significantly more people in Germany live in under-occupied than in overcrowded households. This suggests a considerable potential for improving housing efficiency through the reallocation of existing housing and living resources, without the need for new, carbon-emitting construction. Policy-makers could address this in different ways. One possible solution is to offer incentives to specific groups, such as elderly individuals living in under-occupied housing, to exchange their large apartments with families currently living in overcrowded conditions. Moreover, policy-makers should take into account the unequal distribution of overcrowding along socio-demographic and economic factors and explicitly target high-risk groups when designing new housing policies aimed at reducing the extent of overcrowding.

References

Ager, P., Feigenbaum, J. J., Hansen, C. W. & Tan, H. R. (2020), 'How the other half died: Immigration and mortality in US cities', *Working Paper* **27480**.

Arenas-Arroyo, E., Fernandez-Kranz, D. & Nollenberger, N. (2021), 'Intimate partner violence under forced cohabitation and economic stress: Evidence from the COVID-19 pandemic', *Journal of Public Economics* **194**.

Auspurg, K., Hinz, T. & Schmid, L. (2017), 'Contexts and conditions of ethnic discrimination: Evidence from a field experiment in a German housing market', *Journal of Housing Economics* **35**, 26–36.

Baldenius, T., Kohl, S. & Schularick, M. (2019), 'Die neue Wohnungsfrage: Gewinner und Verlierer des deutschen Immobilienbooms', *Leviathan* **48**(2), 195–236.

Booth, A. & Johnson, D. R. (1975), 'The effect of crowding on child health and development', *American Behavioral Scientist* **18**(6), 736–749.

Bujard, M., von den Driesch, E., Ruckdeschel, K., Laß, I., Thönnissen, C., Schumann, A. & Schneider, N. F. (2021), 'Belastungen von Kindern, Jugendlichen und Eltern in der Corona-Pandemie', *BIB Bevölkerungs-Studien* **2**.

Cage, R. A. & Foster, J. (2002), 'Overcrowding and infant mortality: A tale of two cities', *Scottish Journal of Political Economy* **49**(2).

Contreras, D., Delgadillo, J. & Riveros, G. (2019), 'Is home overcrowding a significant factor in children's academic performance? Evidence from Latin America', *International Journal of Educational Development* **67**, 1–17.

Deschermeier, P. & Henger, R. (2015), 'Die Bedeutung des zukünftigen Kohorteneffekts auf den Wohnflächenkonsum', *IW-Trends* (3).

Diehl, C., Andorfer, V. A., Khoudja, Y. & Krause, K. (2013), 'Not in my kitchen? Ethnic discrimination and discrimination intentions in shared housing among university students in Germany', *Journal of Ethnic and Migration Studies* **39**(10), 1679–1697.

Duckitt, J. (1983), 'Household crowding and psychological well-being in a South African coloured community', *The Journal of Social Psychology* **121**, 231–238.

Dustmann, C., Fitzenberger, B. & Zimmermann, M. (2022), 'Housing expenditure and income inequality', *The Economic Journal* **132**(645), 1709–1736.

Eurostat (2014), 'Distribution of population by tenure status, type of household and income group', https://ec.europa.eu/eurostat/databrowser/view/ilc_lvho02/default/table?lang=en. Accessed: 2024-11-10.

Eurostat (2019), 'Statistics explained: Overcrowding rate', https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Overcrowdingrate. Accessed: 2024-06-06.

Eurostat (2021), 'Statistics explained: Under-occupied dwelling', https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Under-occupied_dwelling. Accessed: 2024-06-06.

Eurostat (2023), 'Living conditions in europe - housing', https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Living_conditions_in_Europe_-_housing&oldid=569706. Accessed: 2024-06-06.

Eurostat (2024), 'Statistics explained: Overcrowding rate', https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Housing_cost_overburden_rate&lang=en. Accessed: 2024-06-06.

Evans, G. W., Saegert, S. & Harris, R. (2016), 'Residential density and psychological health among children in low-income families', *Environment and Behavior* **33**(2), 165–180.

Foye, C. (2016), 'The relationship between size of living space and subjective well-being', *Journal of Happiness Studies* **18**(2), 427–461.

Frick, J. R. & Grimm, S. (2009), 'Wohnen in Deutschland nach dem Mauerfall: Eine Analyse für die Jahre 1990 bis 2008 auf Basis der Daten des Sozio-oekonomischen Panels (soep)', *SOEPpapers on Multidisciplinary Panel Data Research* **236**.

Gabor, D. & Kohl, S. (2022), 'My home is an asset class: The financialization of housing in Europe', *Working Paper for European Greens* .

Giel, R. & Ormel, J. (1977), 'Crowding and subjective health in the Netherlands', *Social Psychiatry* **12**(1), 37–42.

Goebel, J., Grabka, M. M., Liebig, S., Kroh, M., Richter, D., Schröder, C. & Schupp, J. (2019), 'The German socio-economic panel (SOEP)', *Jahrbücher für Nationalökonomie und Statistik* **239**(2), 345–360.

Goux, D. & Maurin, E. (2005), 'The effect of overcrowded housing on children's performance at school', *Journal of Public Economics* **89**(5-6), 797–819.

Gränitz, S. (2022), 'Zwischen Wohnen und Nicht-Wohnen: Dimensionen der neuen Wohnungsnot', *WSI Mitteilungen* **75**(3), 205–212.

Gómez-Jacinto, L. & Hombrados-Mendieta, I. (2002), 'Multiple effects of community and household crowding', *Journal of Environmental Psychology* **22**(3), 233–246.

Hadziabdic, S. & Kohl, S. (2024), 'A room of one's own? The consequences of living density on individual well-being and social anomie', *Social Forces* p. 1–23. Published online ahead of print.
URL: *https://doi.org/10.1093/sf/soae163*

Henger, R., Schier, M. & Voigtländer, M. (2014), 'Wohnungsleerstand: Eine wirtschaftspolitische Herausforderung', *Beiträge zur Ordnungspolitik aus dem Institut der deutschen Wirtschaft Köln* (62).

Huang, Y. & Yi, C. (2015), 'Invisible migrant enclaves in Chinese cities: Underground living in Beijing, China', *Urban Studies* **52**(15), 2948–2973.
URL: *https://journals.sagepub.com/doi/abs/10.1177/0042098014564535*

Hubbard, P. (2024), 'Small is beautiful? Making sense of 'shrinking' homes', *Urban Studies* pp. 1–7. Published online ahead of print.

Kholodilin, K. & Kohl, S. (2021), 'Social policy or crowding-out? Tenant protection in comparative long-run perspective', *Housing Studies* **38**(4), 1–24.
URL: *https://doi.org/10.1080/02673037.2021.1900796*

Kohl, S., Sagner, P. & Voigtländer, M. (2019), 'Mangelware Wohnraum: Ökonomische Folgen des Mietpreisbooms in deutschen Großstädten', *Working Paper* **Forschungsinstitut für gesellschaftliche Weiterentwicklung**.

Lopoo, L. M. & London, A. S. (2016), 'Household crowding during childhood and long-term education outcomes', *Demography* **53**(3), 699–721.

Lozano Alcántara, A. & Romeu Gordo, L. (2020), 'Measuring housing costs and housing affordability using SOEP: An example applied to older households', *SOEPpapers on Multidisciplinary Panel Data Research* (1111).

Luik, M.-A., Steinhardt, M. F. & Voss, S. (2023), 'Language proficiency and homeownership: Evidence from U.S. immigrants'. Discussion Papers 2023/3, Free University Berlin, School of Business & Economics.

Myers, D., Baer, W. C. & Choi, S.-Y. (1996), 'The changing problem of overcrowded housing', *Journal of the American Planning Association* **62**(1), 66–84.

Myers, D. & Lee, S. W. (1996), 'Immigration cohorts and residential overcrowding in Southern California', *Demography* **33**(1), 51–65.

Nkosi, V., Haman, T., Naicker, N. & Mathee, A. (2019), 'Overcrowding and health in two impoverished suburbs of Johannesburg, South Africa', *BMC Public Health* **19**(1).

Painter, G. & Yu, Z. (2010), 'Immigrants and housing markets in mid-size metropolitan areas', *International Migration Review* **44**(2), 442–476.

Perez-Vincent, S. M., Carreras, E., Gibbons, M. A., Murphy, T. E. & Rossi, M. A. (2020), 'Covid 19 lockdowns and domestic violence: Evidence from two studies in Argentina', *Inter-American Development Bank Technical Note* **IDB-TN- 1 956**.

Pevalin, D. J., Taylor, M. P. & Todd, J. (2008), 'The dynamics of unhealthy housing in the UK: A panel data analysis', *Housing Studies* **23**(5), 679–695.

Pfeffer, F. T. & Waitkus, N. (2021), 'The wealth inequality of nations', *American Sociological Review* **86**(4), 567–602.

Pierse, N., Carter, K., Bierre, S., Law, D. & Howden-Chapman, P. (2016), 'Examining the role of tenure, household crowding and housing affordability on psychological distress, using longitudinal data', *Journal of Epidemiology and Community Health* **70**(10), 961–966.

Richter, L. M. (1989), 'Household density, family size and the growth and development of black children a cross-sectional study from infancy to middle childhood', *South African Journal of Psychology* **19**(4), 191–198.

Robinson, H., Molenaar, J. & Van Praag, L. (2024), 'Navigating spatial inequalities: The micro-politics of migrant dwelling practices during COVID-19 in Antwerp', *Urban Studies* **61**(9), 1756–1772.
URL: *https://journals.sagepub.com/doi/abs/10.1177/00420980231217389*

Sagner, P. & Voigtländer, M. (2023), 'Mismatch im Wohnungsmarkt', *IW-Kurzbericht* (5).

Schröder, C., König, J., Fedorets, A., Goebel, J., Grabka, M. M., Lüthen, H., Metzing, M., Schikora, F. & Liebig, S. (2020), 'The economic research potentials of the German Socio-Economic Panel study', *German Economic Review* **21**(3), 335–371.

Schwab, J. J., Nadeau, S. E. & Warheit, G. J. (1979), 'Crowding and mental health', *The Pavlovian Journal of Biological Science* **14**(4), 226–233.

Schwartz, H. & Seabrooke, L. (2008), 'Varieties of residential capitalism in the international economy: Old welfare states and the new politics of housing', *Comparative European Politics* **6**, 237–261.

Siedentop, S., Schiller, G., Koziol, M., Walther, J. & Gutsche, J.-M. (2006), 'Siedlungsentwicklung und Infrastrukturfolgekosten – Bilanzierung und Strategieentwicklung'.

SOEP (2024), 'Socio-economic panel (SOEP), version 39, data for years 1984-2022 (soep-core v39, EU edition)'. DOI: 10.5684/soep.core.v39eu.

Solari, C. D. & Mare, R. D. (2012), 'Housing crowding effects on children's wellbeing', *Social Science Research* **41**(2), 464–476.

Sunega, P. & Lux, M. (2016), 'Subjective perception versus objective indicators of overcrowding and housing affordability', *Journal of Housing and the Built Environment* **31**, 695–717.

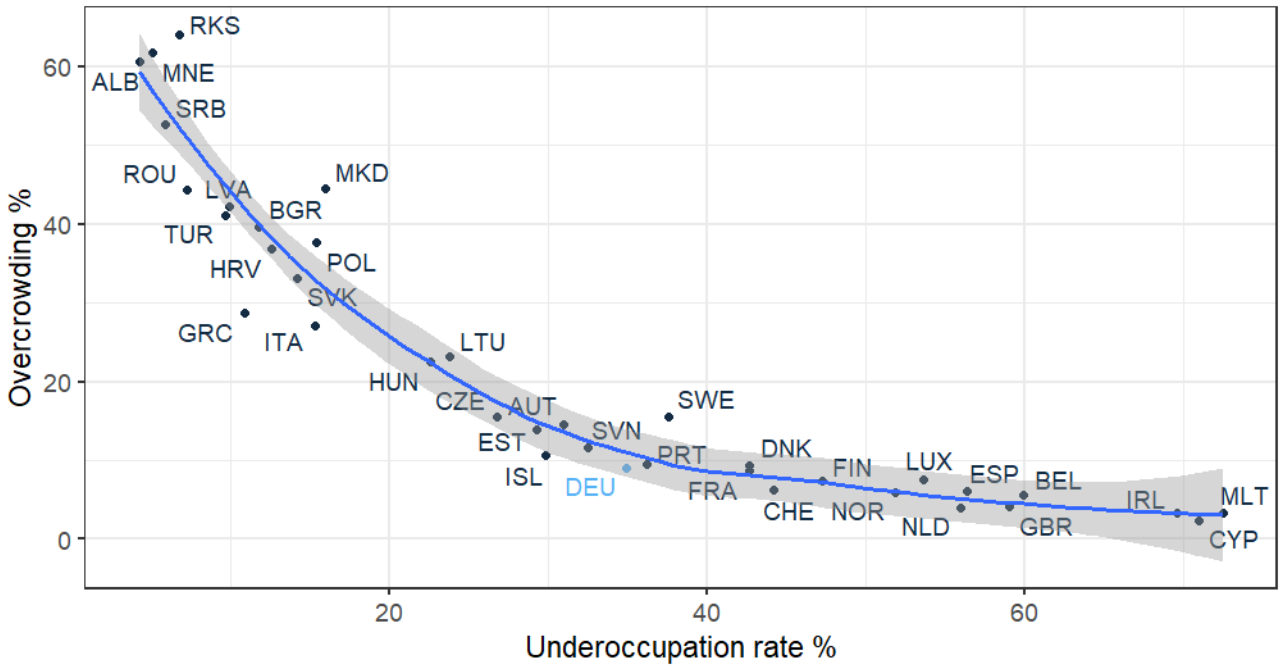
Waltersbacher, M. (2022), 'Empirische Analysen zur Lage am Wohnungsmarkt in Deutschland', *WSI Mitteilungen* **75**(3).

Wang, X. & Liu, T. (2023), 'Home-made blues: Residential crowding and mental health in Beijing, China', *Urban Studies* **60**(3), 461–482.

Wells, N. M. & Harris, J. D. (2007), 'Housing quality, psychological distress, and the mediating role of social withdrawal: A longitudinal study of low-income women', *Journal of Environmental Psychology* **27**(1), 69–78.

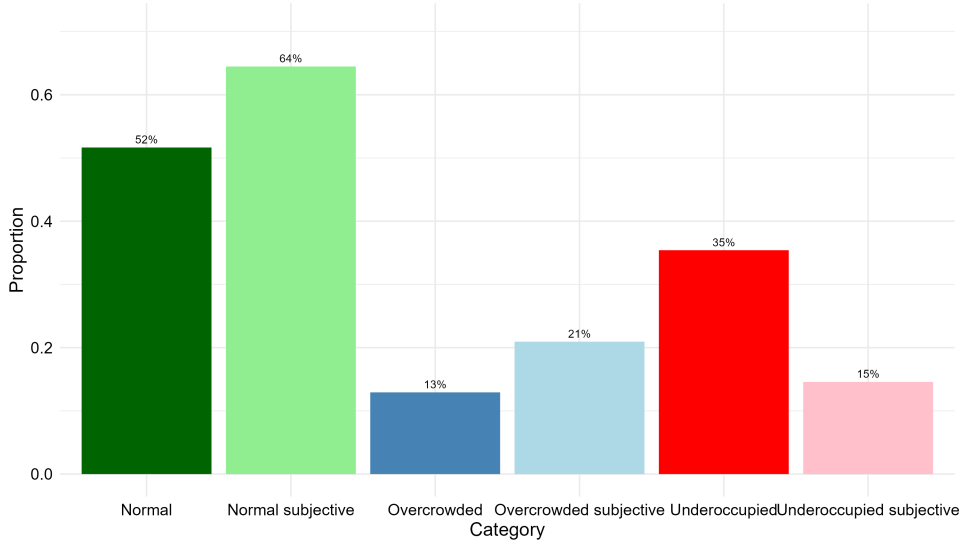
Appendix A: Descriptive Statistics

Figure A.1: Percent Under- and Overcrowded Population in the EU 2022



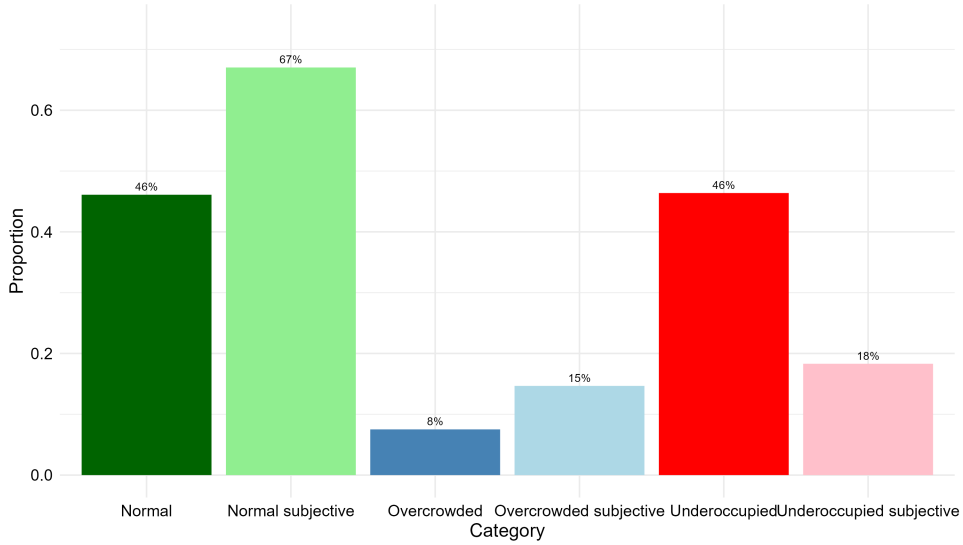
Eurostat data, https://ec.europa.eu/eurostat/databrowser/view/ilc_lvho05b/default/table?lang=en

Figure A.2: Share of Different Housing Conditions in 2022 – Urban Areas



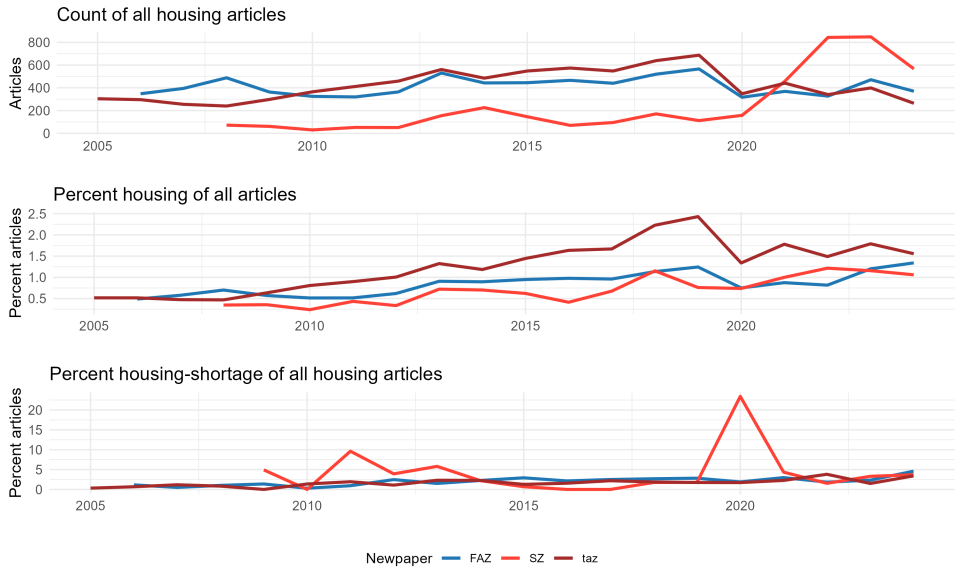
Notes - Data are drawn from the SOEP version 39. Survey year: 2022.

Figure A.3: Share of Different Housing Conditions in 2022 – Rural Areas



Notes - Data are drawn from the SOEP version 39. Survey year: 2022.

Figure A.4: Analysis of housing topic salience in German main newspapers



Notes - Sourced from Nexis - Housing articles were identified by the disjunctive set of (Wohnungspolitik OR Wohnungsbau OR Hauspreis* OR Wohnungsnot OR Wohnungskrise OR Wohnungsmangel OR Wohnungsknappheit OR Wohnungsmiet* OR Mietpreis* OR Immobilienpreis* OR Wohnungsleerstand)′ and scarcity-related articles by a search of (*Mangel OR zu wenig gebaut OR zu wenig Raum OR zu wenig Wohnung* gebaut OR zu wenig Wohnraum OR zu wenig günstigen Wohnraum OR Wohnraumknappheit OR beengt* OR überfüllt* OR überbeleg*) BUT NOT (Mängel)

Table A.1: Descriptive Statistics

	Full Sample	
	Mean	Std. dev.
Overcrowding	0.11	0.31
Overcrowding (a bit/much too small)	0.18	0.39
Underoccupation	0.33	0.47
Underoccupation (a bit/much too big)	0.11	0.31
Switcher	0.29	0.45
Square meters per capita	42.33	23.21
Female	0.49	0.50
Age	42.01	13.17
Married	0.53	0.50
Immigrant	0.13	0.33
Number of children	0.54	0.89
Single parent	0.07	0.26
Medium educated	0.65	0.48
High educated	0.18	0.38
Housing burden	21.61	18.14
Excessive housing burden	0.11	0.31
Monthly HH income	3195.87	2050.16
Homeownership	0.46	0.50
Urban area	0.69	0.46
Dwelling: House	0.52	0.50
Observations	550,848	

Note: Overcrowding is defined following the measure used by Eurostat (for an explanation, see Section 3). For the subjective measure (*overcrowding (a bit/much too small)*), respondents in the SOEP are asked to evaluate the size of their apartment. Similarly, *Underoccupation* is defined following the measure used by Eurostat (for an explanation, see Section 3), and *Underoccupation (a bit/much too big)* corresponds to the subjective evaluation of the apartment size by SOEP respondents. *Switcher* is a dummy coded as one if an individual has changed the overcrowding status —either entering or exiting overcrowding— at least once during the sample period. For *Square meters per capita*, we consider the housing unit's size in square meters divided by the number of persons living in the household. The variable *Immigrant* refers to first-generation immigrants and therefore takes the value 1 if an individual was born outside of Germany. *Children* refers to the number of individuals aged younger than 18 that are currently living in the household. *Single parent* is a categorical variable taking the value 1 for persons in our sample that are part of a single-parent household. An individual is considered *Medium educated* if they hold a middle vocational degree, a vocational degree and Abitur, or a higher vocational degree. They are considered *High educated* if they have a higher educational degree. *Housing burden* is defined according to the Eurostat measure, as explained in Section 3. *Excessive housing burden* is a categorical variable that equals 1 if a household spends more than 40% of its available income on housing costs. As a measure of *Monthly HH income*, we use the current monthly net household income. *Homeownership* is a categorical variable that takes the value 1 if the dwelling is owned by the household (including those with a mortgage) and 0 otherwise. *Urban area* indicates whether a household is living in a predominantly urban area. Finally, *Dwelling: House* is a binary variable that takes the value 1 if the household is living in a house instead of an apartment. Individuals aged 18-64.

Approximately 11% of the sample reports living in overcrowded households, and 18% believes their accommodation is a bit or much too small. The average number of squared meters is 42 and the average household monthly income is approximately 3,195 euros. Approximately 11% of the sample suffers an excessive housing-cost burden, i.e., individuals spend more than 40% of their income on housing. The average age of the sample is 42 years old, approximately 53% are married, and approximately 13% are

foreign-born. Approximately 46% of the sample are home-owners, 65% have received a high school education, and about 18% have obtained a college degree. On average, 69% of the sample resides in a urban area, close to 7% is part of a single-parent household, and about 52% lives in a house.

Appendix B: Robustness Checks

We begin by assessing the robustness of our main results to the use of different samples. Since there have been a number of substantial changes to the SOEP since its beginning–especially the inclusion of households from former East Germany in 1990–we restrict the observation period to 2000-2022. As shown in column 2 of Table B.1, this does not substantially change any of our coefficients. To check whether our main findings are driven by households that switch into or out of overcrowding over the period 1985-2022, we divide the sample into two groups: individuals who changed their overcrowding status (see column 3 of Table B.1); and individuals who did not experience any change (see column 4 of Table B.1). Overall, we find that, although some coefficients differ in magnitude, the results remain largely consistent with the benchmark specification (see Table 1). When limiting our sample to individuals residing in urban areas (see column 5), we find that the coefficients are more pronounced than for those living in rural areas (see column 6). Similarly, the coefficients are of a higher magnitude when focusing on persons living for rent (see column 7) compared to homeowners (see column 8). These findings are consistent to those from our main analysis (see Table 1).

The same holds true when running separate regressions by age groups (see Table B.2) and time periods (see Table B.3).

Table B.1: Alternative Sample Specifications

Dep. var.: Overcrowding	(1) **Benchmark**	(2) **2000-2022**	(3) **Switchers**	(4) **Non-Switchers**	(5) **Urban Areas**	(6) **Rural Areas**	(7) **Tenants**	(8) **Homeowners**
Female	-0.020***	-0.018***	-0.047***	-0.011***	-0.025***	-0.010***	-0.036***	-0.005***
	(0.002)	(0.003)	(0.006)	(0.002)	(0.003)	(0.003)	(0.004)	(0.002)
Age	-0.005***	-0.005***	-0.005***	-0.002***	-0.005***	-0.004***	-0.008***	-0.002**
	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Age^2	0.000***	0.000***	0.000	0.000***	0.000***	0.000**	0.000***	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Immigrant	0.103***	0.085***	0.093***	0.082***	0.110***	0.065***	0.128***	0.037***
	(0.006)	(0.007)	(0.011)	(0.006)	(0.007)	(0.012)	(0.007)	(0.009)
Married	-0.039***	-0.023***	-0.047***	-0.027***	-0.045***	-0.030***	-0.076***	-0.013***
	(0.004)	(0.004)	(0.010)	(0.003)	(0.005)	(0.006)	(0.005)	(0.004)
Children	0.057***	0.045***	0.069***	0.028***	0.062***	0.046***	0.088***	0.027***
	(0.002)	(0.002)	(0.004)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
Single parent	0.088***	0.097***	0.124***	0.028***	0.093***	0.074***	0.102***	0.025**
	(0.008)	(0.010)	(0.013)	(0.007)	(0.010)	(0.013)	(0.011)	(0.011)
Medium educated	-0.047***	-0.049***	-0.041***	-0.028***	-0.051***	-0.034***	-0.054***	-0.029***
	(0.005)	(0.006)	(0.009)	(0.004)	(0.006)	(0.007)	(0.006)	(0.006)
High educated	-0.057***	-0.058***	-0.053***	-0.035***	-0.064***	-0.041***	-0.061***	-0.046***
	(0.006)	(0.007)	(0.013)	(0.005)	(0.007)	(0.008)	(0.008)	(0.007)
Excessive housing burden	-0.026***	-0.025***	-0.068***	-0.006*	-0.029***	-0.025***	-0.055***	-0.008
	(0.004)	(0.004)	(0.009)	(0.003)	(0.005)	(0.005)	(0.006)	(0.005)
Monthly HH income	-0.061***	-0.067***	-0.082***	-0.031***	-0.068***	-0.042***	-0.095***	-0.009**
	(0.003)	(0.004)	(0.008)	(0.003)	(0.004)	(0.005)	(0.005)	(0.003)
Homeownership	-0.041***	-0.031***	-0.060***	-0.012***	-0.044***	-0.036***		
	(0.004)	(0.005)	(0.012)	(0.003)	(0.006)	(0.007)		
Urban area	0.023***	0.027***	0.028**	0.014***			0.049***	-0.003
	(0.005)	(0.005)	(0.011)	(0.003)			(0.007)	(0.005)
Dwelling: House	-0.058***	-0.055***	-0.076***	-0.032***	-0.067***	-0.036***	-0.057***	-0.066***
	(0.005)	(0.006)	(0.011)	(0.003)	(0.006)	(0.008)	(0.006)	(0.009)
Observations	550,848	420,078	166,619	384,229	366,613	184,235	296,429	251,270
R-squared	0.134	0.129	0.103	0.100	0.146	0.098	0.136	0.052
Mean of Dep. Var.	0.106	0.106	0.106	0.106	0.106	0.106	0.106	0.106
Std.Err. of Dep. Var.	0.308	0.308	0.308	0.308	0.308	0.308	0.308	0.308

Notes - Data are drawn from the SOEP version 39. Standard errors are reported in parentheses and are clustered at the household level. Individuals aged 18-64. All specifications include year FE, and state FE. All estimates are weighted using survey weights.

*Significant at 10 per cent; **Significant at 5 per cent; ***Significant at 1 per cent.

Table B.2: Different Age Groups

Dep. var.: Overcrowding	(1) **Benchmark**	(2) **18-25**	(3) **26-30**	(4) **31-35**	(5) **36-40**	(6) **41-45**	(7) **46-50**	(8) **51-55**	(9) **56-60**	(10) **61-64**
Female	-0.020***	-0.025***	-0.021***	-0.012**	0.005	-0.006	-0.019***	-0.028***	-0.026***	-0.024***
	(0.002)	(0.006)	(0.007)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.004)	(0.005)
Age	-0.005***									
	(0.001)									
Age^2	0.000***									
	(0.000)									
Immigrant	0.103***	0.143***	0.102***	0.110***	0.113***	0.140***	0.135***	0.087***	0.054***	0.029***
	(0.006)	(0.015)	(0.014)	(0.011)	(0.011)	(0.013)	(0.012)	(0.011)	(0.009)	(0.010)
Married	-0.039***	-0.178***	-0.097***	-0.060***	-0.045***	-0.007	-0.025***	-0.039***	-0.049***	-0.045***
	(0.004)	(0.013)	(0.008)	(0.008)	(0.008)	(0.008)	(0.007)	(0.007)	(0.006)	(0.006)
Children	0.057***	0.117***	0.061***	0.055***	0.058***	0.049***	0.050***	0.072***	0.089***	0.077***
	(0.002)	(0.005)	(0.005)	(0.004)	(0.003)	(0.003)	(0.004)	(0.005)	(0.009)	(0.013)
Single parent	0.088***	0.072***	0.081***	0.090***	0.051***	0.083***	0.097***	0.080***	0.040***	0.042**
	(0.008)	(0.012)	(0.018)	(0.018)	(0.015)	(0.014)	(0.013)	(0.016)	(0.015)	(0.018)
Medium educated	-0.047***	-0.018***	-0.045***	-0.068***	-0.117***	-0.087***	-0.056***	-0.047***	-0.015**	-0.015**
	(0.005)	(0.006)	(0.013)	(0.013)	(0.014)	(0.013)	(0.011)	(0.009)	(0.008)	(0.008)
High educated	-0.057***	-0.017	-0.021	-0.051***	-0.132***	-0.123***	-0.094***	-0.058***	-0.024***	-0.015*
	(0.006)	(0.015)	(0.015)	(0.015)	(0.015)	(0.014)	(0.012)	(0.011)	(0.008)	(0.009)
Excessive housing burden	-0.026***	-0.029***	-0.016*	-0.021***	-0.035***	-0.049***	-0.030***	-0.032***	-0.018**	-0.018***
	(0.004)	(0.008)	(0.009)	(0.008)	(0.010)	(0.008)	(0.008)	(0.008)	(0.008)	(0.007)
Monthly HH income	-0.061***	-0.088***	-0.121***	-0.102***	-0.084***	-0.065***	-0.032***	-0.035***	-0.017***	-0.021***
	(0.003)	(0.007)	(0.008)	(0.008)	(0.008)	(0.008)	(0.007)	(0.006)	(0.005)	(0.005)
Homeownership	-0.041***	-0.090***	-0.023**	-0.039***	-0.058***	-0.059***	-0.035***	-0.035***	-0.028***	-0.019***
	(0.004)	(0.011)	(0.009)	(0.007)	(0.008)	(0.009)	(0.008)	(0.006)	(0.005)	(0.005)
Urban area	0.023***	0.026***	0.030***	0.029***	0.025***	0.021**	0.020***	0.022***	0.013**	0.010
	(0.005)	(0.009)	(0.010)	(0.009)	(0.009)	(0.008)	(0.007)	(0.007)	(0.006)	(0.007)
Dwelling: House	-0.058***	-0.115***	-0.056***	-0.050***	-0.075***	-0.073***	-0.056***	-0.032***	-0.023***	-0.018***
	(0.005)	(0.011)	(0.009)	(0.008)	(0.008)	(0.010)	(0.008)	(0.007)	(0.006)	(0.006)
Observations	550,848	80,605	49,928	58,652	66,439	70,028	68,999	62,224	54,493	39,480
R-squared	0.134	0.174	0.122	0.138	0.163	0.149	0.123	0.118	0.089	0.072
Mean of Dep. Var.	0.106	0.187	0.140	0.115	0.127	0.121	0.0900	0.0670	0.0442	0.0322
Std.Err. of Dep. Var.	0.308	0.390	0.347	0.319	0.333	0.326	0.286	0.250	0.205	0.176

Notes - Data are drawn from the SOEP version 39. Standard errors are reported in parentheses and are clustered at the household level. Individuals aged 18-64. Column 1 includes year FE, and state FE. Columns 2-10 include year FE and state FE. All estimates are weighted using survey weights.

*Significant at 10 per cent; **Significant at 5 per cent; ***Significant at 1 per cent.

Table B.3: Different Time Periods

Dep. var.: Overcrowding	(1) **Benchmark**	(2) **1985-1989**	(3) **1990-2000**	(4) **2001-2010**	(5) **2011-2022**
Female	-0.021***	-0.017***	-0.026***	-0.019***	-0.017***
	(0.002)	(0.006)	(0.004)	(0.003)	(0.003)
Age	-0.005***	0.001	-0.004***	-0.004***	-0.007***
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)
Age^2	0.000***	-0.000	0.000	0.000**	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Immigrant	0.099***	0.181***	0.128***	0.079***	0.088***
	(0.006)	(0.016)	(0.012)	(0.012)	(0.007)
Married	-0.034***	-0.125***	-0.048***	-0.029***	-0.019***
	(0.004)	(0.011)	(0.008)	(0.005)	(0.005)
Children	0.057***	0.077***	0.072***	0.043***	0.046***
	(0.002)	(0.006)	(0.005)	(0.003)	(0.003)
Single parent	0.089***	0.025	0.092***	0.108***	0.088***
	(0.008)	(0.026)	(0.015)	(0.014)	(0.011)
Medium educated	-0.051***	-0.045***	-0.049***	-0.045***	-0.053***
	(0.005)	(0.009)	(0.007)	(0.008)	(0.008)
High educated	-0.063***	-0.068***	-0.060***	-0.052***	-0.063***
	(0.006)	(0.014)	(0.011)	(0.010)	(0.009)
Excessive housing burden	-0.027***	-0.027**	-0.025***	-0.030***	-0.019***
	(0.004)	(0.011)	(0.009)	(0.006)	(0.005)
Monthly HH income	-0.062***	-0.037***	-0.058***	-0.063***	-0.066***
	(0.003)	(0.010)	(0.007)	(0.006)	(0.005)
Homeownership	-0.041***	-0.058***	-0.059***	-0.023***	-0.036***
	(0.004)	(0.011)	(0.008)	(0.007)	(0.006)
Urban area	0.025***	0.009	0.019**	0.031***	0.024***
	(0.005)	(0.013)	(0.008)	(0.007)	(0.006)
Dwelling: House	-0.058***	-0.063***	-0.060***	-0.055***	-0.056***
	(0.005)	(0.011)	(0.009)	(0.008)	(0.007)
Observations	550,848	38,364	109,389	153,127	249,968
R-squared	0.130	0.151	0.142	0.118	0.137
Mean of Dep. Var.	0.106	0.128	0.129	0.0908	0.0931
Std.Err. of Dep. Var.	0.308	0.334	0.336	0.287	0.291

Notes - Data are drawn from the SOEP version 38. Standard errors are reported in parentheses and are clustered at the household level. Individuals aged 18-64. Column 1 includes year FE, and state FE. Columns 2-10 include state FE. Note that during the time period covered in column 2 (1985-1989), East Germany was not yet included in the SOEP. All estimates are weighted using survey weights.

*Significant at 10 per cent; **Significant at 5 per cent; ***Significant at 1 per cent.

We further probe the robustness of the results by modifying the model specification. To account for the dynamic nature of household income, in column 2 of Table B.4 we re-estimate our baseline specification using one-year lagged monthly household income among the set of control variables. This change has very little effect on our results. In order to rule out that the results are driven by unobserved time-invariant heterogeneity, we test the robustness of our analysis to the inclusion of individual fixed effects in Model (1) (see column 3 of Table B.4). Compared to our benchmark specification, the coefficient on household income remains negative and significantly related to overcrowding; however, its magnitude becomes much smaller when individual fixed effects are included. In addition, the coefficient on living in an urban area is close to zero and no longer significant. However, our findings still indicate that socio-demographic factors account for the majority of the variation in overcrowding. Finally, in column 4 of Table B.4 we show that our results are unchanged to using sampling weights in our analysis.

Table B.4: Alternative Model Specifications

	(1)	(2)	(3)
Dep. var.: Overcrowding	**Benchmark**	**Lagged HH Income**	**Individual Fixed Effects**
Female	-0.020***	-0.020***	
	(0.002)	(0.002)	
Age	-0.005***	-0.005***	
	(0.001)	(0.001)	
Age^2	0.000***	0.000***	
	(0.000)	(0.000)	
Immigrant	0.103***	0.106***	
	(0.006)	(0.006)	
Married	-0.039***	-0.045***	-0.030***
	(0.004)	(0.004)	(0.005)
Children	0.057***	0.056***	0.033***
	(0.002)	(0.002)	(0.003)
Single parent	0.088***	0.090***	0.098***
	(0.008)	(0.008)	(0.009)
Medium educated	-0.047***	-0.048***	-0.015***
	(0.005)	(0.005)	(0.005)
High educated	-0.057***	-0.064***	-0.030***
	(0.006)	(0.006)	(0.011)
Excessive housing burden	-0.026***	-0.016***	-0.020***
	(0.004)	(0.004)	(0.003)
Monthly HH income	-0.061***	-0.045***	-0.026***
	(0.003)	(0.003)	(0.003)
Homeownership	-0.041***	-0.047***	-0.042***
	(0.004)	(0.004)	(0.005)
Urban area	0.023***	0.021***	0.011
	(0.005)	(0.005)	(0.011)
Dwelling: House	-0.058***	-0.060***	-0.043***
	(0.005)	(0.005)	(0.006)
Observations	550,848	550,848	530,470
R-squared	0.134	0.130	0.601
Mean of Dep. Var.	0.106	0.106	0.106
Std.Err. of Dep. Var.	0.308	0.308	0.308

Notes - Data are drawn from the SOEP version 39. Standard errors are reported in parentheses and are clustered at the household level. Individuals aged 18-64. All specifications include year FE and state FE. All estimates are weighted using survey weights.

*Significant at 10 per cent; **Significant at 5 per cent; ***Significant at 1 per cent.

In Table B.5, we explore alternative outcomes, starting with the duration of overcrowding, given its potential to be a temporary phenomenon (see column 2). In general, our results are in line with those of our main analysis. Factors that influence the likelihood of experiencing overcrowding are also linked to spending more years in overcrowded housing conditions. Since the Eurostat overcrowding measure considers only rooms and not the square meter size of a dwelling, our results could theoretically be driven by changes in the number of rooms that do not affect the square meter size. To control for this, we use a household's square meters per person as an alternative outcome in column 3 of Table B.5. Our results are largely in line with the previous findings: Variables associated with a higher likelihood of overcrowding in our main analysis are now linked to a lower dwelling size per capita. The only exception is marital status, which is negatively associated with overcrowding but also negatively correlated with square meters per person. This may be explained by the fact that the Eurostat measure does not consider couples sharing a single bedroom as overcrowded. Therefore, moving from a studio apartment or flatshare into a new apartment with a spouse can result in no longer being classified as overcrowded, even if the square meter size per capita decreases. The finding that paying an excessive share of household income on housing costs is related to a lower square meter size per person is a bit surprising. As in our main analysis, this could partly be driven by the fact that not all components of housing costs are included in our measure (see Section 3).

Table B.5: Alternative Outcomes

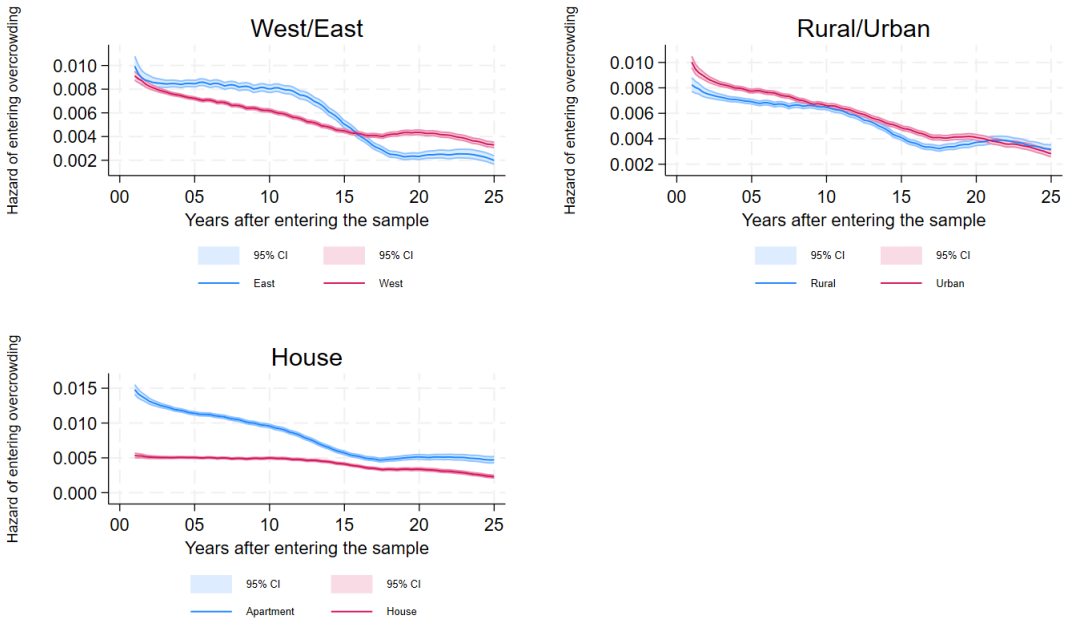
	(1)	(2)	(3)
Dep. var.: Overcrowding	**Benchmark**	**Duration of Overcrowding**	**Square Meters per Person**
Female	-0.020***	-0.095***	0.029***
	(0.002)	(0.022)	(0.003)
Age	-0.005***	0.117***	0.026***
	(0.001)	(0.006)	(0.001)
Age^2	0.000***	-0.001***	-0.000***
	(0.000)	(0.000)	(0.000)
Immigrant	0.103***	0.438***	-0.162***
	(0.006)	(0.036)	(0.007)
Married	-0.039***	-0.291***	-0.246***
	(0.004)	(0.031)	(0.006)
Children	0.057***	0.374***	-0.198***
	(0.002)	(0.013)	(0.003)
Single parent	0.088***	0.404***	-0.123***
	(0.008)	(0.050)	(0.008)
Medium educated	-0.047***	-0.343***	0.109***
	(0.005)	(0.033)	(0.005)
High educated	-0.057***	-0.540***	0.236***
	(0.006)	(0.040)	(0.008)
Excessive housing burden	-0.026***	-0.449***	0.143***
	(0.004)	(0.036)	(0.005)
Monthly HH income	-0.061***	-0.504***	-0.017***
	(0.003)	(0.027)	(0.005)
Homeownership	-0.041***	-0.224***	0.133***
	(0.004)	(0.029)	(0.007)
Urban area	0.023***	0.219***	-0.025***
	(0.005)	(0.029)	(0.007)
Dwelling: House	-0.058***	-0.430***	0.142***
	(0.005)	(0.031)	(0.007)
Observations	550,848	86,264	550,386
R-squared	0.134	0.102	0.429
Mean of Dep. Var.	0.106	0.812	42.33
Std.Err. of Dep. Var.	0.308	2.244	23.21

Notes - Data are drawn from the SOEP version 39. Standard errors are reported in parentheses and are clustered at the household level. Individuals aged 18-64. All specifications include state FE. Columns 1 and 3 also include survey year fixed effects. All estimates are weighted using survey weights.

*Significant at 10 per cent; **Significant at 5 per cent; ***Significant at 1 per cent.

Appendix C: Additional Evidence from Hazard Models

Figure C.1: Hazard of Entering Overcrowding - by Regional Characteristics



Notes - Data are drawn from the SOEP version 39. Each figure shows the unconditional pattern of entering overcrowding. Individuals aged 18-64. Survey years: 1985-2022. Figures are displayed without survey weights in order to include confidence intervals. Results are very similar when survey weights are used.

Table C.1: Potential Determinants of the Hazard of Entering Overcrowding - 1985-2022

Dep. var.: Hazard of entering overcrowding	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
	Socio-demographic characteristics							**Economic characteristics**			**Regional characteristics**		
Female	0.126***	0.158***	0.143***	0.162***	0.122***	0.092***	0.087***	0.084***	0.062*	0.040	0.040	0.039	0.037
	(0.031)	(0.031)	(0.031)	(0.032)	(0.032)	(0.032)	(0.032)	(0.032)	(0.032)	(0.032)	(0.032)	(0.032)	(0.032)
Age		0.075***	0.058***	0.084***	0.043***	0.038***	0.073***	0.071***	0.041***	0.009	0.011	0.011	0.009
		(0.013)	(0.013)	(0.014)	(0.013)	(0.013)	(0.015)	(0.015)	(0.015)	(0.014)	(0.014)	(0.014)	(0.014)
Age^2		-0.002***	-0.001***	-0.002***	-0.001***	-0.001***	-0.001***	-0.001***	-0.001***	-0.001***	-0.001***	-0.001***	-0.001***
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Immigrant			0.784***	0.818***	0.781***	0.781***	0.752***	0.758***	0.691***	0.569***	0.592***	0.589***	0.568***
			(0.057)	(0.058)	(0.058)	(0.058)	(0.059)	(0.059)	(0.059)	(0.059)	(0.060)	(0.060)	(0.061)
Married				-0.308***	-0.689***	-0.568***	-0.574***	-0.580***	-0.435***	-0.380***	-0.376***	-0.376***	-0.369***
				(0.056)	(0.063)	(0.066)	(0.066)	(0.066)	(0.067)	(0.065)	(0.065)	(0.065)	(0.065)
Children					0.421***	0.402***	0.386***	0.387***	0.438***	0.470***	0.471***	0.473***	0.483***
					(0.024)	(0.025)	(0.025)	(0.025)	(0.025)	(0.025)	(0.025)	(0.025)	(0.025)
Single parent						0.457***	0.427***	0.414***	0.309***	0.301***	0.308***	0.310***	0.313***
						(0.069)	(0.069)	(0.069)	(0.070)	(0.071)	(0.071)	(0.071)	(0.071)
Medium educated							-0.284***	-0.294***	-0.241***	-0.214***	-0.229***	-0.231***	-0.232***
							(0.055)	(0.054)	(0.054)	(0.054)	(0.054)	(0.054)	(0.054)
High educated							-0.510***	-0.524***	-0.306***	-0.324***	-0.367***	-0.376***	-0.388***
							(0.080)	(0.080)	(0.079)	(0.079)	(0.080)	(0.080)	(0.079)
Excessive housing cost								0.141**	-0.085	-0.042	-0.034	-0.039	-0.044
								(0.065)	(0.068)	(0.069)	(0.069)	(0.069)	(0.069)
Monthly HH income									-0.641***	-0.446***	-0.441***	-0.445***	-0.436***
									(0.046)	(0.049)	(0.049)	(0.049)	(0.050)
Homeownership										-0.755***	-0.737***	-0.727***	-0.633***
										(0.057)	(0.058)	(0.058)	(0.070)
Urban area												0.117*	0.096
												(0.062)	(0.062)
Dwelling: House													-0.183***
													(0.066)
Observations	404,957	404,957	404,957	404,957	404,957	404,957	404,957	404,957	404,957	404,957	404,957	404,957	404,957
Mean of Dep. Var.	0.0142	0.0142	0.0142	0.0142	0.0142	0.0142	0.0142	0.0142	0.0142	0.0142	0.0142	0.0142	0.0142
Std.Err. of Dep. Var.	0.118	0.118	0.118	0.118	0.118	0.118	0.118	0.118	0.118	0.118	0.118	0.118	0.118

Notes - Data are drawn from the SOEP version 39. Logit estimations; average marginal effects reported. Standard errors are reported in parentheses and are clustered at the household level. Individuals aged 18-64. In columns 11 to 13 state FE are included. All estimates are weighted using survey weights.
*Significant at 10 per cent; ** Significant at 5 per cent; ***Significant at 1 per cent.

Table C.2: Potential Determinants of the Hazard of Leaving Overcrowding - 1985-2022

Dep. var.: Hazard of leaving overcrowding	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
	Socio-demographic characteristics							**Economic characteristics**			**Regional characteristics**		
Female	0.212***	0.199***	0.200***	0.196***	0.226***	0.207***	0.215***	0.220***	0.227***	0.229***	0.225***	0.225***	0.229***
	(0.039)	(0.039)	(0.039)	(0.038)	(0.039)	(0.038)	(0.039)	(0.039)	(0.039)	(0.038)	(0.038)	(0.038)	(0.038)
Age		0.044***	0.051***	0.027**	0.034***	0.034***	0.016	0.018	0.033**	0.033**	0.037***	0.038***	0.041***
		(0.011)	(0.012)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
Age^2		-0.001***	-0.001***	-0.001***	-0.001***	-0.001***	-0.000***	-0.000***	-0.001***	-0.001***	-0.001***	-0.001***	-0.001***
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Immigrant			-0.235***	-0.304***	-0.253***	-0.252***	-0.230***	-0.233***	-0.216***	-0.230***	-0.202***	-0.198***	-0.182***
			(0.056)	(0.058)	(0.057)	(0.057)	(0.058)	(0.059)	(0.059)	(0.060)	(0.061)	(0.061)	(0.061)
Married				0.311***	0.506***	0.554***	0.574***	0.570***	0.464***	0.470***	0.433***	0.424***	0.409***
				(0.061)	(0.066)	(0.070)	(0.070)	(0.071)	(0.073)	(0.073)	(0.073)	(0.073)	(0.072)
Children					-0.194***	-0.202***	-0.190***	-0.193***	-0.214***	-0.214***	-0.218***	-0.221***	-0.230***
					(0.026)	(0.026)	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)
Single parent						0.142**	0.160**	0.167**	0.180**	0.180**	0.180**	0.173**	0.176**
						(0.071)	(0.071)	(0.071)	(0.072)	(0.072)	(0.072)	(0.072)	(0.072)
Medium educated							0.147***	0.156***	0.150***	0.153***	0.152***	0.154***	0.150***
							(0.051)	(0.051)	(0.051)	(0.051)	(0.051)	(0.051)	(0.051)
High educated							0.266***	0.290***	0.266***	0.266***	0.277***	0.291***	0.292***
							(0.083)	(0.084)	(0.083)	(0.083)	(0.085)	(0.084)	(0.084)
Excessive housing burden								-0.245***	-0.115	-0.111	-0.098	-0.092	-0.092
								(0.089)	(0.092)	(0.092)	(0.092)	(0.092)	(0.092)
Monthly HH income									0.267***	0.276***	0.290***	0.295***	0.288***
									(0.056)	(0.057)	(0.057)	(0.057)	(0.057)
Homeownership										-0.026	-0.064	-0.102	-0.222**
										(0.088)	(0.087)	(0.088)	(0.098)
Urban area												-0.226***	-0.169**
												(0.076)	(0.077)
Dwelling: House													0.275***
													(0.078)
Observations	47,203	47,203	47,203	47,203	47,203	47,165	47,165	47,165	47,165	47,165	47,165	47,165	47,165
Mean of Dep. Var.	0.150	0.150	0.150	0.150	0.150	0.150	0.150	0.150	0.150	0.150	0.150	0.150	0.150
Std.Err. of Dep. Var.	0.357	0.357	0.357	0.357	0.357	0.357	0.357	0.357	0.357	0.357	0.357	0.357	0.357

Notes - Data are drawn from the SOEP version 39. Logit estimations; average marginal effects reported. Standard errors are reported in parentheses and are clustered at the household level. Individuals aged 18-64. In columns 11 to 13 state FE are included. All estimates are weighted using survey weights.
*Significant at 10 per cent; ** Significant at 5 per cent; ***Significant at 1 per cent.