

# Location Sorting and Endogenous Amenities: Evidence from Amsterdam\*

Milena Almagro<sup>†</sup>      Tomás Domínguez-Iino<sup>‡</sup>

November 8, 2022  
[Click here for latest version.](#)

## Abstract

This paper argues that the endogeneity of amenities plays a crucial role on the welfare distribution of a city's residents. We quantify this mechanism by constructing a dynamic model of residential choice with heterogeneous households, where urban consumption amenities are the equilibrium outcome of a market for non-tradables. We estimate our model using Dutch administrative microdata and leverage spatial variation in tourism flows and the entry of home-sharing platforms, such as Airbnb, as shifters of location characteristics in Amsterdam. Our results reveal significant heterogeneity across local residents in their valuation of different amenities, as well as in the response of amenities to demographic composition. We then show that the distributional effects of the tourist boom hinge on this heterogeneity: after initial rent increases due to a reduction in the housing supply available to locals, younger groups—the most similar to tourists—are compensated by having amenities tilt in line with their preferences, while older families end up being additionally hurt by this shift in amenities. We show that policies that target amenities can be especially welfare-enhancing when the preferences that residents hold over the amenities tourism brings are sufficiently heterogeneous.

---

\*First version: November 2019. We thank Guillaume Fréchette, Alessandro Lizzeri, Elena Manresa, Tobias Salz, Paul Scott, and Sharon Traiberman for their support at the early stages of this project. For useful feedback, we are grateful to Juan Camilo Castillo, Don Davis, Fabian Eckert, Cecile Gaubert, Jessie Handbury, Allan Hsiao, Anders Humlum, Erik Hurst, Panle Jia-Barwick, and László Tétényi. Marek Bojko, Elias van Emmerick, and Sara Gerstner provided outstanding research assistance. We gratefully acknowledge financial support from the C.V. Starr Center for Applied Economics. Any errors or omissions are our own.

<sup>†</sup>University of Chicago Booth School of Business. E-mail: [milena.almagro@chicagobooth.edu](mailto:milena.almagro@chicagobooth.edu)

<sup>‡</sup>Federal Reserve Board. E-mail: [tdominguezino@gmail.com](mailto:tdominguezino@gmail.com). The views expressed are those of the authors and not necessarily those of the Federal Reserve Board or the Federal Reserve System.

# 1 Introduction

Socioeconomic inequality is tightly linked to residential choice, both across and within-cities (Moretti, 2013). Higher socioeconomic status households can afford to live in more desirable locations, and the locations themselves can in turn have their amenities improve as residential composition changes. This endogenous response of a location's amenities to demographic sorting has been shown to be a quantitatively important mechanism for amplifying welfare inequality (Guerrieri, Hartley and Hurst, 2013; Diamond, 2016). However, relatively little is understood about the nature of these *endogenous amenities*, as they are typically modeled as a one-dimensional object summarizing a wide variety of a location's characteristics.

Within the broad set of welfare-relevant amenities, recent work has focused on the role that *consumption amenities* play in neighborhood change (Couture, Gaubert, Handbury and Hurst, 2019; Hoelzlein, 2019; Miyauchi, Nakajima and Redding, 2021). However, these studies typically restrict household heterogeneity only along income levels. This assumption contrasts with an extensive body of work showing that many other socioeconomic characteristics play a crucial role in a household's residential decisions (Bayer, Ferreira and McMillan, 2007; Couture and Handbury, 2020). Specifically, one may expect different types of households to have diverse tastes across different types of consumption amenities, and the firms providing such amenities to cater to this heterogeneity (George and Waldfogel, 2003). While providing tractability, the aggregation of amenities into a single index does not allow for the *horizontal* differentiation of neighborhoods on the demand side, neither for differential supply-side responses to consumer heterogeneity. Therefore, in this paper, we ask: How does preference heterogeneity over a set of endogenous consumption amenities shape within-city residential sorting and inequality?

To answer our research question, we build and estimate a dynamic spatial equilibrium model of a city with heterogeneity in household preferences over a *bundle of endogenous amenities*, whose supply caters to each neighborhood's demographic composition. To estimate our model, we use restricted access census microdata from the Centraal Bureau voor de Statistiek (CBS), the statistics bureau of the Netherlands. From these data, we construct an annual panel of residential location choices for the universe of residents in the Netherlands. We complement these data with an annual panel of establishment counts for the city of Amsterdam, allowing us to track consumption amenities across time and space. Apart from the availability of high-quality data, we choose Amsterdam as our empirical setting because it has undergone significant change due to the impact of mass tourism on local housing and amenity markets, thus providing an ideal laboratory to study

the link between residential composition and endogenous amenities.

We start by presenting evidence that the expansion of tourism across Amsterdam is sufficiently important to affect both housing and local amenity markets. The number of overnight tourist stays in Amsterdam went from 8 million in 2008 to nearly 16 million in 2017, while commercially operated Airbnb listings grew to nearly 10% of the city's rental stock by 2017.<sup>1</sup> We show the stark increase in tourism flows is not spatially concentrated in the city center, where most hotels are. Instead, short-term rentals expanded the presence of tourists to all areas of the city, even to districts where hotels are non-existent. We then proceed to show this expansion in short-term rentals, namely Airbnb, has had a significant impact on Amsterdam rent prices. We continue by showing this expansion of tourism is correlated with changes in the composition of neighborhood amenities. While amenities catering to tourists increase all over the city, their presence is negatively correlated with amenities catering exclusively to locals, such as private daycare facilities. Next, we show that different demographic groups respond differently to these neighborhood changes through their residential choices, suggesting different valuations for the changes taking place across Amsterdam.

The major obstacle in quantifying the effect of endogenous amenities on within-city inequality is that both amenities and residential choices are equilibrium outcomes and hence simultaneously determined. Therefore, to understand the relationship between them, we model and estimate the demand and supply sides of the amenities market and embed them into a residential choice model. On the demand side, our model features forward-looking households choosing where to live each period, with heterogeneous preferences over different neighborhood amenities. In our context, the dynamic behavior of households should be taken into account for several reasons. First, the literature has shown evidence of forward-looking decisions in a similar panel setting of location choices (Bayer, McMillan, Murphy and Timmins, 2016). Second, moving decisions are infrequent in the short-run, suggesting large moving frictions. Failure to account for these frictions would make agents appear to be less responsive to changes in location characteristics than they actually are, leading to biased estimates toward zero and muted welfare effects from amenity changes. Finally, we provide evidence that these frictions increase the longer households live in the same location. We capture this behavior by allowing agents in our model to accumulate *location capital* that is lost upon moving. Therefore, households face an inter-temporal trade-off between moving to a new location and enjoying the benefits of a more preferred location or staying put

---

<sup>1</sup>This corresponds to 2.5% of the total housing stock. We define commercial listings as entire-home listings that operate year-round, so locals are unlikely to live in them.

and accumulating location capital.

Our model also endogenizes the supply of amenities as well as housing. On the housing supply side, we model atomistic absentee landlords who supply their housing unit to locals on traditional long-term leases or to tourists on short-term leases. On the amenity supply side, we endogenize the supply of different consumption amenities through a market where non-tradable services are provided by monopolistically competitive firms. These firms consider tourists, along with different types of local residents with heterogeneous preferences, as part of their consumer market.<sup>2</sup> Our micro-foundation of amenity supply has several advantages to models where amenities are modelled as a one-dimensional index. First, it provides a clear interpretation of how local amenities respond to demographics. Second, it allows us to study regulations that are targeted to specific amenities. Finally, given our model of demand and supply, the market's equilibrium conditions provide the mapping between the number of firms in each amenity category and the demographic composition of a location, which includes tourists.

We estimate our dynamic location choice model by building upon the Euler Equation in Conditional Choice Probability (ECCP) methodology ([Aguirregabiria and Magesan, 2013](#); [Scott, 2013](#); [Kalouptsidi, Scott and Souza-Rodrigues, 2021](#)) to estimate the preference parameters over neighborhood characteristics. To do so, we use within-neighborhood spatial variation and an instrumental variable approach to deal with the endogeneity of rental prices and consumption amenities. Our estimation results are consistent with intuitive differences in demographics across household types. All households exhibit that moving is costly, with older households having the highest moving costs. In terms of valuation of amenities, households with higher socioeconomic status perceive touristic amenities as dis-amenities. Single middle-aged households without children value restaurants the most, consistent with having the most leisure time among the demographic groups. By contrast, households with children value nurseries the most. On the amenity supply side, we find that the presence of tourists creates positive incentives for the entry of all type of firms but for private day-care facilities. We also find reasonable co-location results between the groups of residents and the entry of different types of services.

Given the estimated parameters, we use the model to run counterfactuals highlighting how preference heterogeneity and the endogeneity of amenities interact to determine spatial sorting and welfare inequality. In our first counterfactual,

---

<sup>2</sup>By non-tradable “service”, we mean a broad sector of amenities, such as restaurants, which may have different “varieties” within it. For example, Italian and Japanese restaurants would be different varieties within the restaurant service.

we compare the equilibrium outcome of a world where amenities are exogenous to one in which they endogenously respond to population composition, finding a significant increase in residential sorting across demographic groups. We find that despite the increased sorting across space, welfare inequality across demographic groups can fall if preferences are heterogenous. Intuitively, if preferences over amenities are misaligned between two demographic groups they sort into different locations, raising the supply of their most preferred amenities while also avoiding competition with each other in the housing market. Thus, there are two mechanisms reducing the welfare gap across locations when preferences are heterogenous: tailored amenities and lower housing prices. Our findings complement the existing literature on residential sorting by introducing heterogeneity in the two-way relationship between households and amenities, which allows us to evaluate the role of preference heterogeneity on welfare inequality.

In our second counterfactual, we evaluate the effect of Airbnb entry on Amsterdam resident welfare. Our main goal is to disentangle the welfare effects for residents into two components: changes in rent and changes in amenities. This decomposition allows us to separate the direct effect of Airbnb entry on rent via the reduction in housing supply, from the indirect effect via the endogenous response of amenities to the increased tourist population. The key insight behind our results is that while all residents lose from higher rent, some lose and some win from the changes in amenities due to preference heterogeneity. In particular, residents whose preferences are correlated with those of tourists will benefit from an increase in the amenities that they bring.

Finally, in our third counterfactual we compare different forms of regulating mass tourism: through housing markets or through amenity markets. Specifically, we compare a tax on short-term rentals to a tax on touristic amenities. First, we show that the distributional impact of both types of policies hinges on preference heterogeneity: households who value touristic amenities lose from them, and vice versa. Second, we find that taxing amenities is preferred to taxing short-term rentals for households with very heterogenous preferences over the amenities tourists bring. The reasoning is that the short-term rental tax reduces the tourist population and all the amenities they bring in an un-targeted manner, including the subset of amenities they bring which are actually desired by locals. By contrast, the amenity tax specifically targets touristic amenities, and in doing so decouples the undesirable amenities brought by tourists, such as souvenir shops, from the desirable ones, such as restaurants. This targeting is therefore especially valuable to groups which value some—but also dislike some—of the amenities tourists bring.

**Related literature.** Spatial equilibrium models date back to Rosen (1979) and

Roback (1982) and have experienced a recent comeback as the benchmark tool to study spatial inequality across and within cities (Moretti, 2013; Diamond, 2016; Couture and Handbury, 2020). A subset of the literature focuses on the within-city margin, developing methods to quantify residential agglomeration and dispersion forces, but typically remains silent on the exact mechanisms through which specific amenities are provided (Bayer et al., 2007; Guerrieri et al., 2013; Ahlfeldt, Redding, Sturm and Wolf, 2015; Davis, Gregory and Hartley, 2018; Su, 2018). Recent work imposes structure on amenity provision by building upon tools from the trade literature, but often lack heterogeneity in residents' preferences over amenities or collapse amenities into a single quality index (Couture et al., 2019; Hoelzlein, 2019; Miyauchi et al., 2021). We contribute to the literature by incorporating preference heterogeneity over amenities into a dynamic model of residential choice, where the provision of consumption amenities is microfounded through a market mechanism. We build upon the notion of "preference externalities" (George and Waldfogel, 2003; Handbury, 2021) that proposes that demand-side preference heterogeneity translates into differences in the variety of products that are supplied in equilibrium. We show how these same insights can be used to interpret neighborhoods as differentiated products where amenities play the role of endogenous product attributes, and we highlight the implications for residential sorting and urban inequality.

Our paper also contributes to the literature examining the recent rise of the short-term rental industry, as well as tourism more broadly. There is extensive reduced-form work on the effects of Airbnb entry on the housing market (Sheppard, Udell et al., 2016; Koster, Van Ommeren and Volkhausen, 2021; Garcia-López, Jofre-Monseny, Martínez-Mazza and Segú, 2020; Barron, Kung and Proserpio, 2021) as well as on hotel revenue (Zervas, Proserpio and Byers, 2017). Additionally, some papers quantify the industry's welfare impact through the lens of a structural model. Farronato and Fradkin (2018) study the effect of Airbnb entry on the competing hotels. Calder-Wang (2019) studies the distributional effects on the New York City rental market, focusing on rent effects but abstracting from amenity effects. Faber and Gaubert (2019) show the importance of tourism in the economic development of the coastline of Mexico. Finally, Allen, Fuchs, Ganapati, Graziano, Madera and Montoriol-Garriga (2021) study the effects of seasonal tourism on prices of goods and amenities borne by residents of Barcelona. We complement their work by simultaneously studying the effects of tourism on both residential and amenity markets, showing how they interact to shape urban inequality.

In terms of methodology, we leverage discrete-choice methods from the empirical industrial organization literature and show how they can be applied to ur-

ban residential markets (McFadden, 1974; Berry, 1994; Berry, Levinsohn and Pakes, 1995; Rust, 1987). Specifically, our dynamic estimation uses the Euler Equation in Conditional Choice Probabilities (ECCP) estimator (Hotz and Miller, 1993; Arcidiacono and Miller, 2011; Aguirregabiria and Magesan, 2013; Scott, 2013; Kalouptsidi et al., 2021). The method has been applied to several contexts where dynamics are first order: agricultural markets (Scott, 2013; Hsiao, 2021), occupational choice (Traiberman, 2019; Humlum, 2019), and residential choice (Diamond, McQuade and Qian, 2018; Davis et al., 2018; Davis, Gregory, Hartley and Tan, 2021).

## 2 Data

**Individual-level data: residential histories and socioeconomic characteristics.** Our restricted-access microdata comes from the statistical bureau of the Netherlands, Centraal Bureau voor de Statistiek (CBS). The key dataset for our dynamic model is the residential cadaster, which allows us to construct a panel of residential history for the universe of individuals in the Netherlands. We also observe household-level socioeconomic characteristics from tax return microdata: income, educational attainment, employment status, household composition, and ethnic background.<sup>3</sup> We classify households as low, medium, or high skill using educational attainment bins. Further details about how we restrict our estimation sample, which spans 2008-2020, are in Appendix S.2.1.

**Housing unit data: rental prices, tax valuations, tenancy status, physical characteristics.** Our restricted-access microdata includes several datasets at the housing unit level, which we combine to obtain rental prices, and housing unit characteristics. First, we obtain property values from a panel of tax appraisal data for the universe of residential housing units for 2006-2020, which also includes geo-coordinates, quality measures, and the occupant's tenancy status (owner-occupied, rental, or social housing). Second, we obtain rental prices from a national rent survey for 2006-2019. Since the survey does not cover the universe of tenants, we impute rental prices by linking it to the universal tax appraisal data. We use a random forest to predict rental prices, which outperforms standard linear hedonic models (Mullainathan and Spiess, 2017). Imputation details are in Appendix S.2.4.

**Neighborhood-level data: amenities, demographics, tourist inflows.** We use two levels of geographic units based on Amsterdam's administrative divisions:

---

<sup>3</sup>Unfortunately, tax returns only allow us to observe household income (pre- and post-tax). We do not have data on work locations, the specific occupations of the household members, or the worked hours. For this reason, our papers focuses on outcomes in the residential market rather than the labor market.

99 “wijken” (neighborhoods) that belong to 22 larger “gebieden” (districts). We observe neighborhood-level outcomes at annual frequency from the publicly available Amsterdam City Data (ACD) from 2008 to 2018. This dataset contains over 700 neighborhood-level variables including sociodemographics (e.g., ethnic, income, and skill composition) as well as a rich set of consumption amenities (e.g., restaurants, bars, grocery stores). We specifically distinguish between consumption amenities targeted to locals versus those targeted to tourists. To do so, we use ACD’s definition of “touristic amenities”, which encompasses lodging, passenger transport, travel agencies, and cultural and recreational retail. We also use ACD as our source for tourist inflows.<sup>4</sup>

**Airbnb listings.** We obtain Airbnb listings data from [Inside Airbnb](#), an independent website that provides monthly web-scraped listings data for many cities around the world. Our web scrapes consist of listing-level observations with detailed information such as geographic coordinates, host identifiers, prices per night, calendar availability, and reviews. We use this information to separately identify “active” from “dormant” Airbnb listings, as well as to flag commercially-operated listings which are likely to be permanently rented to tourists year-round and are therefore removing housing stock away from locals. We define such listings as entire-home listings with sufficient booking activity such that a local cannot plausibly be living in the housing unit permanently. Full details of how we classify listings are in Appendix [S.2.5](#).

### 3 Stylized facts

This section presents the stylized facts that motivate our model’s key features. We show how tourism volume and Airbnb penetration correlate with our outcomes of interest: rental prices, consumption amenities, and residential movements.

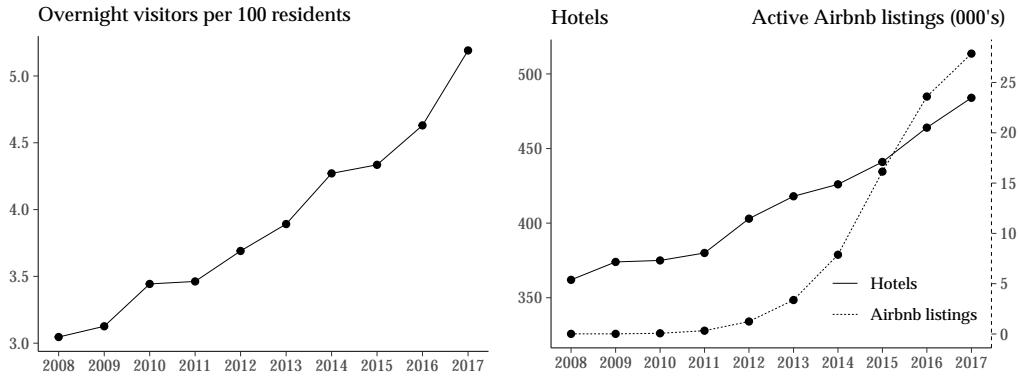
**Fact 1: Tourism inflows and Airbnb listings have grown dramatically and sprawled across Amsterdam.** Amsterdam has one of the highest tourist-to-local ratios in the world, slightly above Florence and below Venice.<sup>5</sup> Figure 1 shows the number of overnight visitors per resident nearly doubled between 2008-2017. To accommodate tourist inflows, the stock of hotels grew from 362 to 484, while active Airbnb listings grew from zero in 2008 to over 25,000 in 2017.

---

<sup>4</sup>The ACD wijk-level data is publicly available at [ACD BBGA](#). Tourism-specific datasets are available at [ACD Tourism](#).

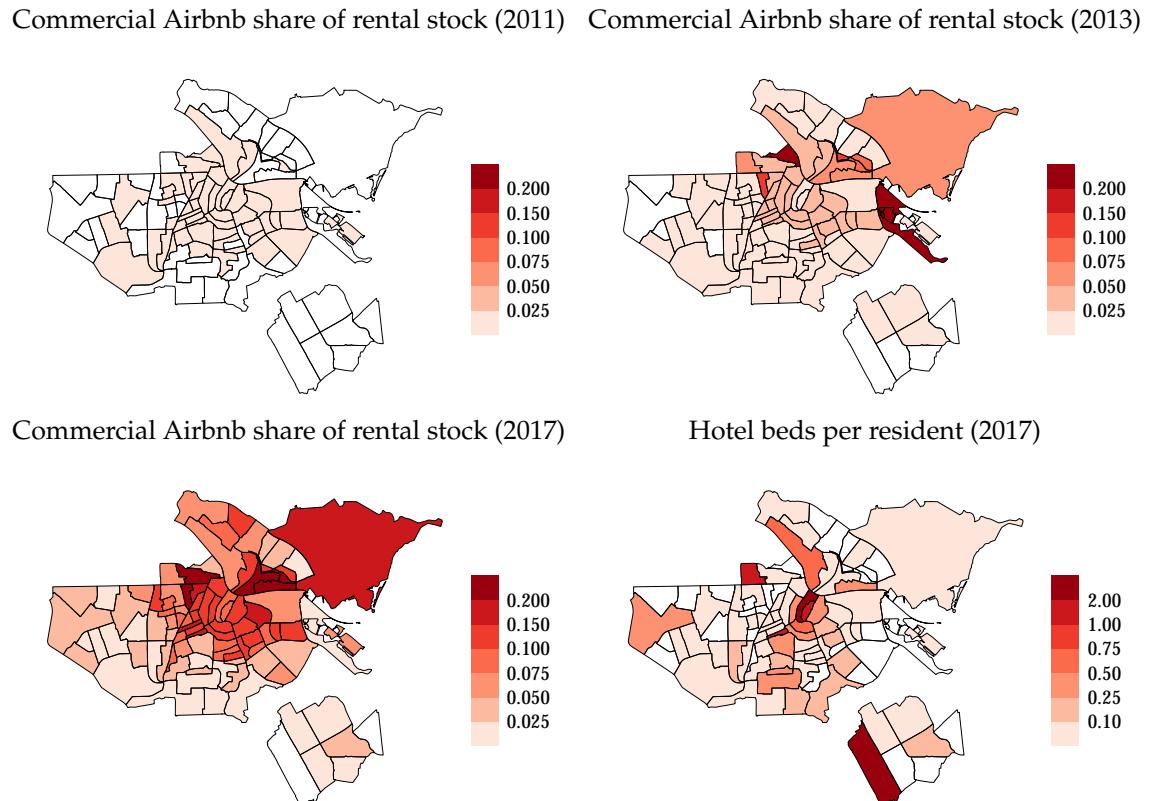
<sup>5</sup>Own calculations based on 2018 tourist arrival and population data from [ESTA](#).

Figure 1: Overnight visitors per resident, hotels, and Airbnb listings (2008-2017).



Notes: Figure shows the increase in overnight visitors and touristic lodgings (data source: [ACD Tourism](#)). Amsterdam population data is from [ACD BBGA](#). We construct active Airbnb listings from Inside Airbnb data (procedure described in Appendix S.2.5).

Figure 2: Airbnb share of rental stock and hotel beds per resident (2011-2017).



Notes: We construct commercial Airbnb listings from Inside Airbnb data (procedure described in Appendix S.2.5). Rental housing stock, hotel beds, and population data is from [ACD BBGA](#).

Figure 2 shows the spatial distribution of Airbnb has sprawled to cover most of the city. By contrast, the current distribution of hotels is concentrated in the city center. At the aggregate level, commercially-operated Airbnb listings represented 7% of Amsterdam's rental market in 2017, exceeding 20% in some central areas.<sup>6</sup> These trends suggest the increasing presence of tourists as part of the city's population is significant enough to alter local housing and amenity markets.

**Fact 2: Rents have increased more in neighborhoods with more Airbnb entry.** Motivated by the heterogeneity in Airbnb growth across the city, Table 1 shows how the intensity of Airbnb penetration is correlated with housing market outcomes: a 1% increase in commercial Airbnb listings is associated with a rent increase between .06-.12%. These magnitudes are sizable given the annual growth rate of rent between 2009-2019 was 1.02%, and are mostly in line with a recent literature estimating the effect of Airbnb on housing market prices.<sup>7</sup>

Table 1: Relationship between housing market outcomes and Airbnb listings

	Ln (rent/m2)			Ln (house sale value)		
	OLS	OLS	FE	OLS	OLS	FE
Ln (commercial Airbnb listings)	0.066*** (0.008)	0.052*** (0.006)	0.115*** (0.018)	0.108*** (0.016)	0.031*** (0.006)	0.045** (0.022)
Ln (housing stock)		-0.056** (0.027)	-0.111*** (0.028)		0.006 (0.027)	-0.045 (0.032)
Ln (average income)		-0.492*** (0.075)	-0.353*** (0.072)		1.013*** (0.071)	0.953*** (0.100)
Ln (high-skill population share)		0.330*** (0.053)	-0.014 (0.100)		0.356*** (0.039)	0.130 (0.090)
District-year FE			X			X
Observations	780	773	773	746	745	745
R2	0.154	0.422	0.579	0.124	0.748	0.885

Notes: Standard errors clustered at the wijk level in parenthesis. We construct commercial Airbnb listings from the Inside Airbnb data, with the exact procedure described in Appendix S.2.5. Rents and house sale values are from a combination of CBS surveys and transaction data, described in section 2. All other variables are from ACD BBGA.

However, this literature does not typically disentangle the underlying mechanisms driving the price effects: Airbnb's direct effect of reducing housing supply

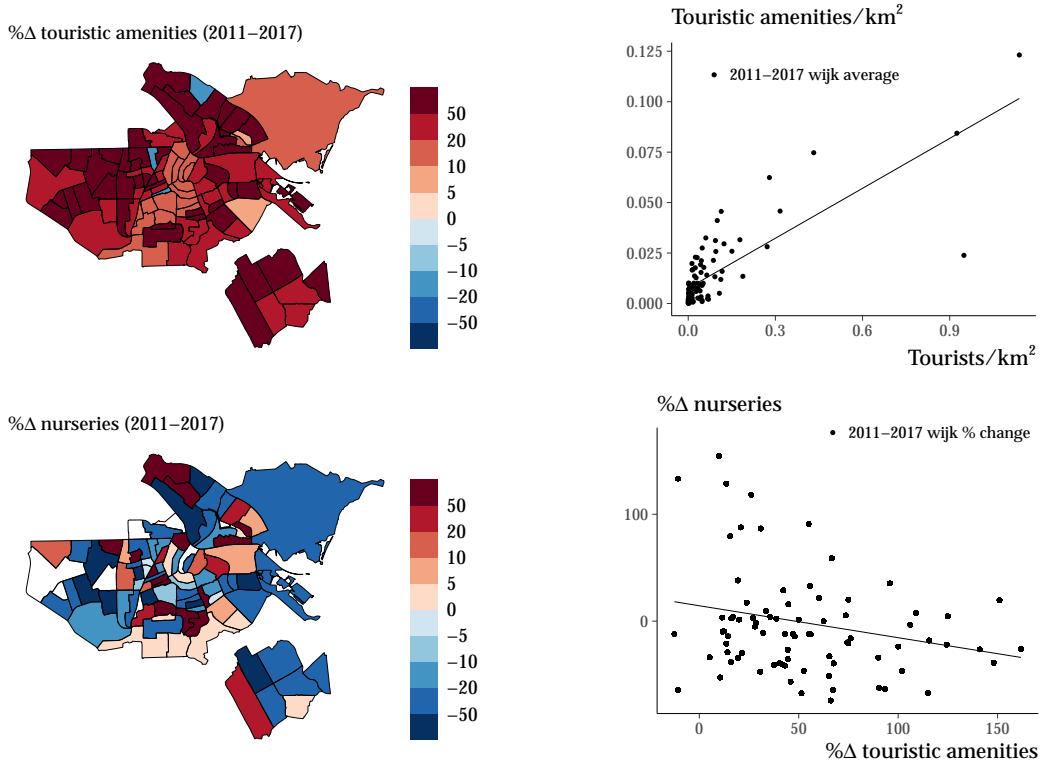
<sup>6</sup>Commercially-operated Airbnb listings amount to 3.2% of the market housing stock. Our definition of market housing stock excludes social housing units because they are not allocated to tenants through a market mechanism. See Appendix S.S.1 for institutional details on how social housing is allocated. For reference, home-owners, renters, and social housing represented 30%, 27%, and 42% of the total housing stock in 2017.

<sup>7</sup>Barron et al. (2021) estimate an Airbnb elasticity of rent of 0.018 and Garcia-López et al. (2020) find a value of 0.0098. In Appendix section S.3.1 we show how our results compare to these studies by using a shift-share identification strategy that is similar to what is typically used in this literature.

is combined with indirect equilibrium effects from changing amenities and resident composition. Given our next facts highlight there have indeed been significant changes in Amsterdam's amenities and resident composition, the goal of our structural model is precisely to disentangle direct and indirect effects.

**Fact 3: Amenities have tilted towards tourists and away from locals.** Beyond the impact of short-term rentals on the housing market, the neighborhood amenities surrounding the housing units have also changed as tourists become an increasing share of the city's population. Figure 3 shows touristic amenities have grown across nearly all neighborhoods, although at different intensities, while amenities that cater exclusively to locals, such as nurseries, have declined. Furthermore, there is a negative relationship between changes in amenities targeted to locals, such as nurseries, and changes in amenities targeted to tourists. This substitution suggests an amenity market with two-sided heterogeneity: firms supplying differentiated amenities to consumers with heterogeneous preferences.

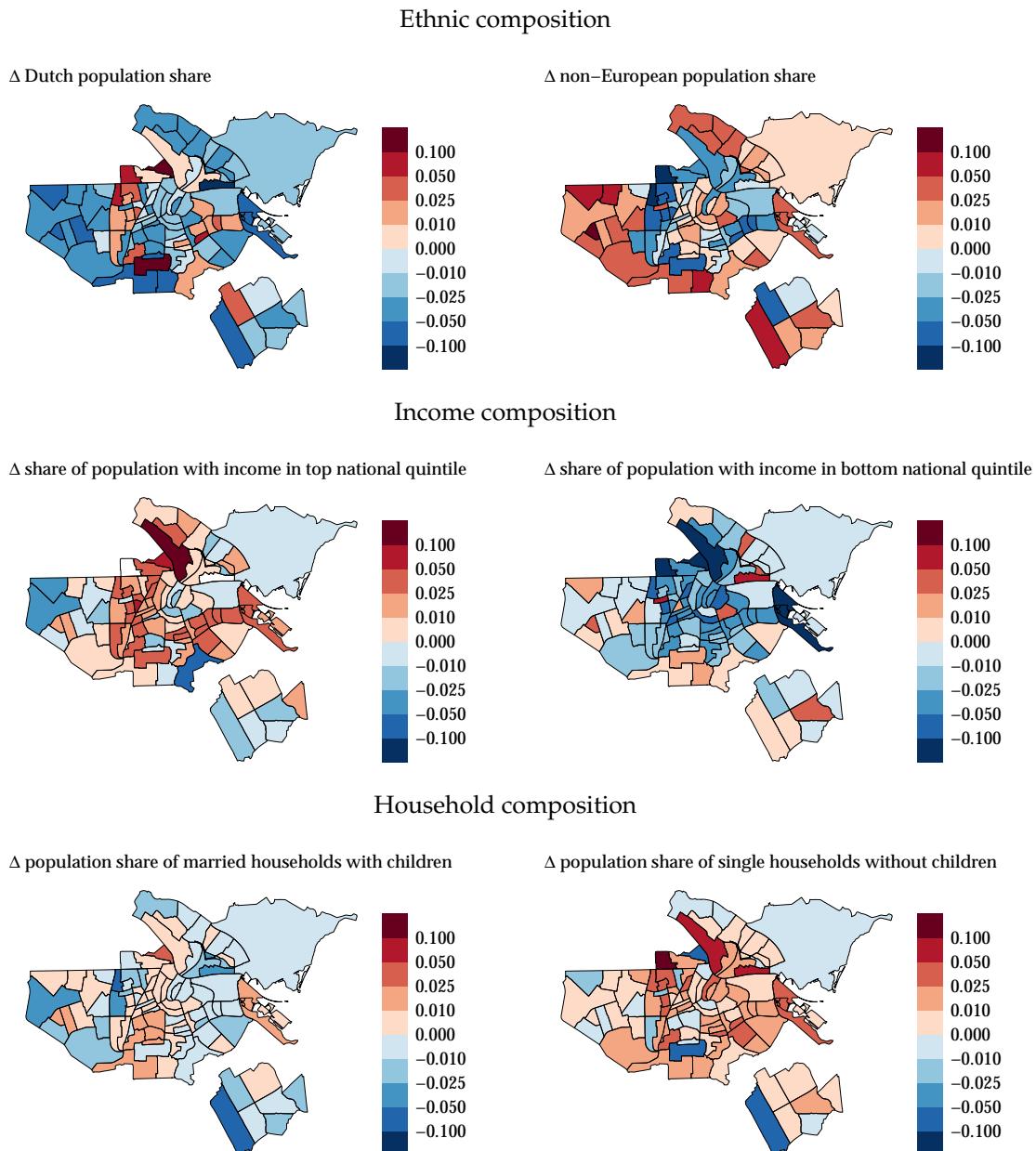
Figure 3: Changes in consumption amenities (2011-2017).



Notes: Data on neighborhood-level consumption amenities is from **ACD BBGA**. ACD has its own definition of “touristic amenities”, which we use directly, and which encompasses lodging, passenger transport, travel agencies, and cultural and recreational retail.

**Fact 4: The composition of residents has changed heterogeneously across neighborhoods.** Figure 4 shows that despite being exposed to the same city-level trends, different types of residents make different moving decisions, suggesting heterogeneity in the valuation of neighborhood characteristics.

Figure 4: Changes in socioeconomic composition of neighborhoods (2011-2017).

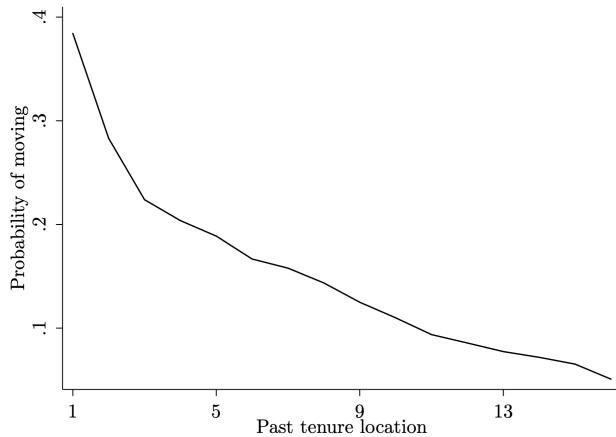


Notes: Neighborhood-level population data by ethnic background, income quantile, marital status, and number of children is all from ACD BBGA.

The top row of Figure 4 shows how moving decisions vary across ethnic groups. The clearest trend is a falling share of residents with Dutch background in most neighborhoods, except those immediately bordering the city center. While the share of immigrants with Western backgrounds has increased nearly everywhere, the share of non-European immigrants has increased in the city's periphery and in a few central neighborhoods. In terms of heterogeneity along income, the middle row of Figure 4 shows the share of residents in the top 20% of the national income distribution has grown in central neighborhoods but not in the outskirts, indicating a rise in income inequality between the city core and periphery. Finally, the bottom row of Figure 4 shows heterogeneity along household composition: married households with children are leaving the city and are being replaced by single households without children. To summarize, the rich heterogeneity in moving decisions across households motivates rich heterogeneity in our model's demand primitives: moving costs, rent elasticities, and valuation of amenities.

**Fact 5: The moving decisions of households are dynamic.** The changes in neighborhood composition we have described are the result of individual moving decisions which occur infrequently and can be expected to have an important dynamic component (Bayer et al., 2016). In line with this view, Figure 5 shows the hazard rate of moving is decreasing in a household's tenure at the prior residence.

Figure 5: Probability of changing residence, conditional on past location tenure.



Notes: Figure shows the probability of moving out of the current location conditional on the number of years lived in the location. We take averages across individuals and across time. Moving probabilities and tenure are constructed using location choice panel derived from the CBS cadaster. More details can be found in sections 2 and S.2.1.

One potential explanation is that households accumulate neighborhood-specific capital over time that is lost upon moving. Moving costs can therefore be inter-

preted to be broader than a fixed switching cost, and to also include the loss of location capital (Diamond et al., 2018). Our structural model explicitly incorporates such dynamics by allowing for bilateral moving costs, forward-looking behavior as well as the accumulation of location capital.

## 4 A dynamic model of an urban rental market

Motivated by the previous facts, we build a dynamic model of a city’s rental market that consists of three parts: i) heterogeneous households with dynamic moving decisions across neighborhoods, ii) landlords who can rent their units to locals or tourists, and iii) a market for amenities that microfounds how the composition of amenities endogenously responds to the composition of locals and tourists.

### 4.1 Endogenous amenities

Consumption amenities are classified into  $S$  sectors, each consisting of a finite number of firms providing differentiated varieties. For example, if the sector is “restaurants”, a firm corresponds to an individual restaurant supplying its own unique variety. Within each sector  $s$  and location  $j$ , there are  $N_{sj}$  firms supplying their varieties in a monopolistically competitive setting with free entry. On the demand side, there are  $K$  types of consumers, each of which holds its own heterogeneous preferences over amenities.

**Demand for amenities.** Conditional on living in location  $j$ , a type  $k$  household chooses how much of its budget net of housing to allocate across the locally available consumption amenities. We microfound the consumer’s problem with *CES preferences over varieties within a sector nested within Cobb-Douglas preferences over amenity sectors* as follows,

$$\max_{\{q_{isjt}^k\}_{is}} \quad \prod_s \left[ \left( \sum_{i=1}^{N_{sjt}} q_{isjt}^k \right)^{\frac{\sigma_s - 1}{\sigma_s}} \right]^{\alpha_s^k} \quad \text{s.t.} \quad \sum_{is} p_{isjt} q_{isjt}^k = I_t^k,$$

where  $q_{isjt}^k$  is the quantity of variety  $i$  in sector  $s$  and location  $j$ ,  $p_{isjt}$  is the price,  $I_t^k$  is the consumer’s after-rent income,  $\alpha_s^k$  is the budget share spent on sector  $s$ , and  $\sigma_s > 1$  is the elasticity of substitution across varieties within the sector. We assume all firms within a sector-location face the same costs, hence in equilibrium  $p_{isjt} = p_{sjt} \forall i \in s, j$ . The demand for a variety by each consumer type,  $q_{isjt}^k$ , and

aggregate demand  $q_{isjt}$ , are respectively,

$$q_{isjt}^k = \frac{\alpha_s^k I_t^k}{p_{sjt} N_{sjt}} \quad \forall i \in sj \implies q_{isjt} = \frac{\sum_k \alpha_s^k I_t^k M_{jt}^k}{p_{sjt} N_{sjt}} \quad \forall i \in sj, \quad (1)$$

where  $M_{jt}^k$  is the number of type  $k$  consumers in location  $j$ . For the full derivation of the consumer problem, we refer the reader to Appendix section A.1.

**Supply of amenities.** Within a sector-location, marginal cost is given by  $c_{sjt}$ . Therefore, optimal prices are given by:

$$p_{isjt} = \frac{c_{sjt}}{1 - \frac{1}{\sigma_s}} \quad \forall i \in sj. \quad (2)$$

To operate in a sector-location, firms must pay a fixed cost each period,

$$F_{sjt} = \Lambda_s \Lambda_j \Lambda_t N_{jt}^\eta \varphi_{sjt},$$

where  $\Lambda_s$ ,  $\Lambda_j$ , and  $\Lambda_t$  are sector-, location-, and time-specific shifters,  $\varphi_{sjt}$  are remaining idiosyncratic cost shifters, and  $N_{jt}^\eta > 0$  is an endogenous entry cost component, which acts as a congestion force aimed to capture competition for commercial real estate between firms in location  $j$ . The number of firms in a sector-location  $N_{sjt}$  is therefore endogenously determined by a zero-profit condition,

$$(p_{isjt} - c_{sjt}) q_{isjt} = F_{sjt} = \Lambda_s \Lambda_j \Lambda_t N_{jt}^\eta \varphi_{sjt} \quad \forall i \in sj. \quad (3)$$

**Equilibrium amenities.** Given all firms in a sector-location  $sj$  make the same pricing decision, the equilibrium is symmetric:  $q_{isjt} = q_{sjt}$  and  $p_{isjt} = p_{sjt} \forall i \in sj$ . Given symmetry, and substituting 1 and 2 into 3 we can derive the equilibrium number of firms in  $sj$ ,

$$N_{sjt} = \frac{1}{\sigma_s F_{sjt}} \sum_k \alpha_s^k I_t^k M_{jt}^k. \quad (4)$$

We define a location's consumption amenities  $a_{jt}$  as the vector of firms in each sector,

$$a_{jt} \equiv [N_{1jt}, N_{2jt}, \dots, N_{Sjt}]' = \mathcal{A}(M_{jt}^1, \dots, M_{jt}^K, M_{jt}^T).$$

The second equality above stresses the role of the amenities model: to microfound a mapping  $\mathcal{A}(\cdot)$  from residential composition  $[M_{jt}^1, \dots, M_{jt}^K, M_{jt}^T]$  to amenities  $a_{jt}$ . We include tourists  $M_{jt}^T$  as a type of "resident" because they are a relevant group of consumers for the firms supplying the amenities.

## 4.2 Housing demand

**Choice set.** At the beginning of every period  $t$ , a household  $i$  chooses a residential location  $j_{it}$  among  $J$  different locations in a city, as well as an outside option of leaving the city altogether,<sup>8</sup>

$$j_{it} = \begin{cases} j & \text{if the household chooses location } j \in \{1, \dots, J\} \\ 0 & \text{if the household chooses a location outside of the city.} \end{cases}$$

Upon moving, households incur a moving cost  $MC^k$  that depends on the distance between the origin and destination location,

$$MC^k(j, j_{it-1}) = \begin{cases} 0 & \text{if } j = j_{it-1} \\ m_0^k + m_1^k \text{dist}(j, j_{it-1}) & \text{if } j \neq j_{it-1} \text{ and } j, j_{it-1} \neq 0 \\ m_2^k & \text{if } j \neq j_{it-1}, \text{ and } j = 0 \text{ or } j_{it-1} = 0. \end{cases}$$

**State variables.** The individual-level state variables are current location  $j_{it}$  and tenure length  $\tau_{it}$ . The latter is key to rationalize the decreasing hazard rate of moving (Figure 5), and evolves as,

$$\tau_{it} = \begin{cases} \min\{\tau_{it-1} + 1, \bar{\tau}\} & \text{if } j_{it} = j_{it-1} \\ 1 & \text{otherwise,} \end{cases}$$

where we have assumed tenure can be accumulated up to a maximum absorbing state  $\bar{\tau}$ . In addition to individual-level state variables, there are aggregate-level state variables: the vector of rental prices across all neighborhoods  $r_t$ , the matrices of consumption amenities  $a_t$  and other non-consumption amenities  $b_t$ , as well as factors that are unobservable to the econometrician  $\xi_t$ . To condense notation, we denote  $x_{it}$  as the vector of individual state variables and  $\omega_t$  as the vector of aggregate state variables,

$$x_{it} \equiv (j_{it-1}, \tau_{it-1}) \in \mathcal{X} \quad \text{and} \quad \omega_t \equiv (r_t, a_t, b_t, \xi_t) \in \Omega.$$

We denote with subscript  $t$  the functions that depend on the aggregate state  $\omega_t$ , in particular the flow utility function and the value function,

$$u^k(j, x_{it}, \omega_t) \equiv u_t^k(j, x_{it}) \quad \text{and} \quad V^k(j, x_{it}, \omega_t) \equiv V_t^k(x_{it}, \epsilon_{it}).$$

---

<sup>8</sup>For simplicity, we assume that homeowners and renters face the same discrete choice problem. Concretely, homeowners are absentee landlords renting to themselves. Under this assumption, it is easy to compute the overall welfare of homeowners by adding up renters' consumer surplus to rental income in our counterfactual simulations.

**Flow utility.** Preference parameters differ by household type. The flow payoff for a household  $i$  of type  $k$  living in location  $j$  is a function of the individual  $i$ 's state,  $x_{it}$ , location characteristics  $\omega_{jt}$ , and location and time fixed effects,  $\delta_j^k$  and  $\delta_t^k$ :

$$u_t^k(j, x_{it}) = \delta_j^k + \delta_t^k + \delta_r^k \log r_{jt} + \delta_a^k \log a_{jt} + \delta_b^k \log b_{jt} + \delta_\tau^k \log \tau_{it} - MC^k(j, j_{it-1}) + \xi_{jt}^k, \quad (5)$$

which is micro-founded by the amenity demand system (see Appendix A.1 for derivations).

**Value function.** Household  $i$ 's value function is defined as,

$$V_t^k(x_{it}, \epsilon_{it}) = \max_D \mathbb{E}_t \left[ \sum_{s \geq t}^{\infty} u_s^k(j, x_{is}) + \epsilon_{idt} \middle| j, x_{it}, \epsilon_{it} \right],$$

where  $\epsilon_{idt}$  is a type I EV idiosyncratic shock and the maximization is taken over policy functions  $D : \mathcal{X} \times \Omega \times \mathbb{R}^J \rightarrow \{0, 1, \dots, J\}$ . Given the recursive nature of the problem, we can write,

$$V_t^k(x_{it}, \epsilon_{it}) = \max_{j \in \{0, 1, \dots, J\}} u_t^k(j, x_{it}) + \epsilon_{it} + \beta \mathbb{E}_t \left[ V_{t+1}^k(x_{it+1}, \epsilon_{it+1}) \middle| j, x_{it}, \epsilon_{it} \right].$$

**Demand for each location.** The probability a type  $k$  household in state  $x_{it}$  chooses location  $j$  is,

$$\mathbb{P}_t^k(j|x_{it}) = \frac{\exp \left( u_t^k(j, x_{it}) + \beta \mathbb{E}_t \left[ V_{t+1}^k(x_{it+1}, \epsilon_{it+1}) \middle| j, x_{it}, \epsilon_{it} \right] \right)}{\sum_{j'} \exp \left( u_t^k(j', x_{it}) + \beta \mathbb{E}_t \left[ V_{t+1}^k(x_{it+1}, \epsilon_{it+1}) \middle| j', x_{it}, \epsilon_{it} \right] \right)}. \quad (6)$$

Demand from all type  $k$  households for location  $j$  is,

$$\mathcal{D}_{jt}^{Lk} = \sum_x \mathbb{P}_t^k(j|x) M_{xt}^k Q_{xt}^k,$$

where  $M_{xt}^k$  is the number of households of type  $k$  with individual state  $x$  at time  $t$  and  $Q_{xt}^k$  is the demanded quantity of housing.<sup>9</sup> Total demand for location  $j$  is obtained by summing over all household types,

$$\mathcal{D}_{jt}^L = \sum_k \sum_x \mathbb{P}_t^k(j|x) M_{xt}^k Q_{xt}^{D,k}. \quad (7)$$

---

<sup>9</sup>Following our microfoundation in Section A.1, we know that  $Q_{xt}^{D,k} = \frac{(1-\phi^k)w_t^k}{r_{jt}}$ , where  $(1-\phi^k)$  is housing expenditure shares that we compute using our microdata.

**Evolution of population distribution.** Denote  $\pi_t^k(j, \tau)$  as a type- $k$  household's probability of living in location  $j$  with tenure  $\tau$ , conditional on the aggregate state at time  $t$ . Denote  $\Pi_t^k$  as the transition matrix across individual states, i.e.,  $\pi_{t+1}^k = \Pi_t^k \pi_t^k$ , where each  $(j, \tau)$  cell evolves as,

$$\pi_{t+1}^k(j, \tau) = \begin{cases} \sum_{\tau'} \sum_{j' \neq j} \mathbb{P}_t^k(j|j', \tau') \pi_t^k(j', \tau') & \tau = 1 \\ \mathbb{P}_t^k(j|j, \tau - 1) \pi_t^k(j, \tau - 1) & \tau \in [2, \bar{\tau}] \\ \mathbb{P}_t^k(j|j, \bar{\tau} - 1) \pi_t^k(j, \bar{\tau} - 1) + \mathbb{P}_t^k(j|j, \bar{\tau}) \pi_t^k(j, \bar{\tau}) & \tau = \bar{\tau}. \end{cases}$$

**Stationary distribution.** We denote  $\Pi^k(\mathbf{r}, \mathbf{a})$  as the transition matrix of type  $k$  households across individual states  $(j, \tau)$ , conditional on the aggregate state (the vectors of rental prices  $\mathbf{r}$  and amenities  $\mathbf{a}$ ). A stationary distribution over individual states,  $\pi^k(\mathbf{r}, \mathbf{a})$ , is therefore defined as  $\pi^k(\mathbf{r}, \mathbf{a}) = \Pi^k(\mathbf{r}, \mathbf{a})\pi^k(\mathbf{r}, \mathbf{a})$ .

### 4.3 Housing supply

We denote by  $\mathcal{H}_{jt}$  the total stock of housing (in units of floor space) in location  $j$  and year  $t$ . We assume that total housing stock is exogenous and determined outside our model. However, we endogenize how  $\mathcal{H}_{jt}$  is split between the long-term rental market (which caters to locals) and the short-term rental market (which caters to tourists) through the following landlord's problem.

**Landlord problem.** Absentee landlords make a binary choice between renting their unit in the long-term market ( $L$ ) or in the short-term market ( $S$ ).<sup>10</sup> Each unit has a given floor space of  $f_j$  square meters. The income obtained from long-term rentals is  $r_{jt}$ , and from short-term rentals is  $p_{jt}$ . In order to capture different matching and managerial costs involved in renting short- versus long-term, we allow for a wedge in landlords' operating costs  $\kappa_{jt}$ , which is unobservable to the econometrician. The landlord's problem is therefore,

$$\max \{ \alpha r_{jt} + \epsilon_L, \quad \alpha p_{jt} - \kappa_{jt} + \epsilon_S \},$$

where  $\alpha$  is the landlord's marginal utility of income and  $\epsilon_L$  and  $\epsilon_S$  are type I EV shocks.

**Supply of housing at each location.** The share of the housing stock allocated to the long- and short-term rental markets are,

---

<sup>10</sup>We can also allow for an outside option, that is, leaving the house empty. However, the number of empty houses in Amsterdam is essentially zero due to regulations that prevent housing units from being vacant for extended periods of time. See [amsterdam.nl/en/housing/obligation-homeowner/](http://amsterdam.nl/en/housing/obligation-homeowner/) for more details.

$$s_{jt}^L = \frac{\exp(\alpha r_{jt})}{\exp(\alpha r_{jt}) + \exp(\alpha p_{jt} - \kappa_{jt})} \quad \text{and} \quad s_{jt}^S = \frac{\exp(\alpha p_{jt} - \kappa_{jt})}{\exp(\alpha r_{jt}) + \exp(\alpha p_{jt} - \kappa_{jt})}.$$

## 4.4 Equilibrium

Our equilibrium definition departs slightly from standard definitions because we not only require market clearing in the housing market, but also in the market for consumption amenities.

**Definition.** A *stationary equilibrium* is,

1. a vector of long-term rental prices  $\mathbf{r} = (r_1, \dots, r_J)$  and a matrix of amenities  $\mathbf{a} = [a_1, \dots, a_J]$ ,
2. a policy function  $h(r_j, p_j; \kappa_j, \epsilon_l)$  for landlords,
3. a policy function  $j^k(j_i, \tau_i, \mathbf{r}, \mathbf{a}; \epsilon_i)$  for each type  $k$  household,
4. a stationary distribution of types over locations and tenure,  $\pi^k(\mathbf{r}, \mathbf{a})$

such that,

1. each landlord  $l$  supplies housing optimally to locals or tourists by choosing  $h(r_j, p_j; \kappa_j, \epsilon_l)$ , so that long- and short-term rental supply in location  $j$  are respectively,<sup>11</sup>

$$\mathcal{H}_j^L(r_j, p_j; \kappa_j) = \frac{\exp(\alpha r_j)}{\exp(\alpha r_j) + \exp(\alpha p_j - \kappa_j)} \mathcal{H}_j \quad \text{and} \quad \mathcal{H}_j^S(r_j, p_j; \kappa_j) = \mathcal{H}_j - \mathcal{H}_j^L(r_j, p_j; \kappa_j).$$

2. each household  $i$  of type  $k$  demands housing optimally by choosing  $j^k(j_i, \tau_i, \mathbf{r}, \mathbf{a}; \epsilon_i)$ , so that demand for long-term rentals in location  $j$  is,

$$\mathcal{Q}_j^{D,L}(\mathbf{r}, \mathbf{a}) = \sum_k M^k \sum_\tau \left[ \pi^k(\mathbf{r}, \mathbf{a}) Q_{j,\tau}^{D,k}(\mathbf{r}) \right]_{j,\tau}$$

where  $M^k$  is the city-wide population of group  $k$ .

3. rental prices  $\mathbf{r}$  clear the long-term rental market,<sup>12</sup>

$$\mathcal{H}_j^L(r_j, p_j; \kappa_j) = \mathcal{Q}_j^{D,L}(\mathbf{r}, \mathbf{a}) \quad \forall j.$$

---

<sup>11</sup>Our assumption of a fixed total housing stock  $\mathcal{H}_j$  implies short-term rental supply is determined as a residual from long-term rental supply.

<sup>12</sup>Equilibrium short-term rental quantities are fully pinned down by supply (tourist demand is perfectly elastic at a vector of exogenously determined short-term prices  $\mathbf{p}$ ).

4. equilibrium amenities are determined by the local composition of residents through the mapping  $\mathcal{A}(\cdot)$ , as microfounded by the amenities model,

$$a_j = \mathcal{A}(M_j^1, \dots, M_j^K, M_j^T),$$

where the total number of tourists  $M_j^T$  is the sum of an exogenously given number of tourists in hotels  $M_j^{HT}$  and an endogenously determined number of tourists in short-term rentals  $M_j^{ST} = \chi_j \mathcal{H}_j^S(r_j, p_j; \kappa_j)$ , where  $\chi_j$  is the mean number of tourists per rental unit in location  $j$ .

Proof of equilibrium existence is in Appendix S.4.1.

## 5 Estimation

Using the data described in Section 2, we construct an annual panel of location choices for 2008-2020 and an annual panel of location characteristics for 2008-2018.

### 5.1 Defining household heterogeneity

We first classify households into three categories based on modal tenancy status: homeowners, private market renters, and social housing renters. This ex-ante classification step is motivated by the fact that the average household belongs to her modal category more than 90% of the time, which suggests this margin of adjustment is minor in our context. It is also useful for several reasons. First, it allows us to abstract away from the homeownership decision. Second, it also allows us to cleanly and separately quantify welfare effects on homeowners and renters in our welfare analysis. Therefore, in what follows, we assume household tenancy status is determined outside our model and constant over time.

After the first classification step, we classify households into “types” using a k-means algorithm on demographics. Existing studies typically classify households into ex-ante groups based on observable demographics, such as race or income (Bayer et al., 2016; Davis et al., 2018). When defining such groups the practitioner faces a variance-bias trade-off. On the one hand, having more groups can capture more heterogeneity, but on the other hand there are fewer observations to estimate choice probabilities, which may lead to noisy estimates. The k-means approach allows us to solve this trade-off in a data-driven manner by reducing the dimension of observed characteristics through exploiting correlations across observables.<sup>13</sup>

---

<sup>13</sup>A clear example is income and skill: income is highly correlated with skill, so a classification including both dimensions separately may be redundant.

Table 2: Summary Statistics by Household Type

Group	Homeowners		Rental		Social Housing Tenants	
	Older Families	Singles	Younger Families	Students	Immigrant Families	Dutch Low Income
Age	44.59	37.84	40.56	28.42	55.12	38.52
Share Children	0.93	0.12	0.65	0.13	0.53	0.43
Share Low-Skilled	3.20%	2.42%	6.09%	5.40%	99.91%	0.02%
Share Medium-Skilled	3.01%	5.87%	2.28%	11.33%	0.09%	16.95%
Share High-Skilled	93.79%	91.71%	91.65%	83.27%	0.00%	83.02
Share Dutch Indies	6.92%	6.59%	4.12%	4.07%	13.22%	12.41%
Share Dutch	64.41%	58.74%	53.13%	61.44%	24.86%	49.36%
Share Non-Western	18.76%	21.43%	21.64%	19.48%	57.96%	30.37%
Share Western	9.91%	13.23%	21.12%	15.01%	3.96%	7.87%
Household Income (€)	62,031.39	30,611.41	47,441.08	16,821.48	21,243.24	27,714.85
Income Pctl.	77.04	45.49	64.64	23.23	33.41	42.17
Per Capita Income (€)	40,155.65	27,609.21	35,058.39	15,162.83	15,167.45	21,178.13
Income Pctl. per Person	73.42	52.84	65.83	26.34	26.69	42.10
Number of Households	106,388	78,561	105,712	124,112	83,117	174,203

Notes: This table presents the groups resulting from k-means classification on mean demographic characteristics over time. We report average characteristics across households in each group. Group names are provided to serve as an easy-to-remember label and are not an outcome of the data.

This dimensionality reduction is particularly useful when there is a large set of observable characteristics, as in our case. Moreover, the method admits heuristics to pin-down the optimal number of groups. The technical details of our classification method are in Appendix S.3.2.

**Results.** Table 2 shows the household types that result from our classification algorithm, along with summary statistics of their average characteristics. We have given each group a label, based on how prominent their demographic characteristics are. For example the "Student" group is characterized by being the youngest and low income. Overall, we consider the groups to be easily distinguishable in terms of age, income, skill level, ethnicity, and household composition.

## 5.2 Amenities

We use equation 4 to derive our estimating equation for amenity  $s$  in neighborhood  $j$  at time  $t$ ,

$$\log N_{sjt} = \lambda_j + \lambda_t + \eta \log N_{jt} + \log \left( \sum_k \beta_s^k X_{jt}^k \right) + \phi_{sjt}, \quad (8)$$

where  $X_{jt}^k$  is the total expenditure of group  $k$  on all amenities,  $\beta_s^k$  determines how such expenditures are allocated to each amenity sector  $s$ ,  $\eta$  is the congestion force in the commercial real estate market,  $\lambda_j$  and  $\lambda_t$  are location- and time-invariant components of firm entry costs, and  $\phi_{sjt}$  captures any remaining unobservables driving such costs.<sup>14</sup> Our main objects of interest are  $\beta_s^k$  and  $\eta$ , which we infer from the correlation between amenity composition (as measured by  $N_{sjt}$ ) and residential composition (as measured by  $X_{jt}^k$ ).

**Identification.** The main identification problem in identifying  $\beta_s^k$  from 8 is simultaneity. For a given location, the distribution of amenity expenditures by household type,  $X_{jt}^k$ , is determined by the local population composition, which is the outcome of residential choices made based on the availability of amenities  $N_{sjt}$ . Hence, any unobservable firm entry cost  $\phi_{sjt}$  affecting  $N_{sjt}$  will be correlated with  $X_{jt}^k$  in equilibrium. Because  $\phi_{sjt}$  is an amenity supply shock, we focus on instruments that are arguably an amenity demand shock. Following this intuition, we compute the total number of available units by tenancy status  $\tau$  in location  $j$ ,  $S_{jt}^\tau$ , from the tax valuation registry and use those counts as shifters of the number of households by type  $k$ . We finally interact the wages of group  $k$  with their corresponding tenancy status counts to construct the following instruments:

$$Z_{jt}^k = w_t^k S_{jt}^{\tau(k)}.$$

The intuition behind our relevance condition is that if there is more availability of social housing units, we should expect higher expenditure on amenities from households living in social housing, and similarly for other tenancy types. Similarly, if group  $k$  has higher income, we should expect higher expenditures in amenity consumption. Because we are incorporating time fixed effects, our exclusion restriction only requires that the tenancy status counts are uncorrelated with

---

<sup>14</sup>More precisely, to get from 4 to 8, we define  $X_{jt}^k \equiv I_t^k M_{jt}^k$ , where  $I_t^k$  is expenditure on total consumption amenities by a type  $k$  household and  $M_{jt}^k$  is the population of type  $k$  households. Under our micro-foundation in Appendix A.1,  $I_t^k = \phi^k w_t^k$ , where  $\phi^k$  is the expenditure share on total consumption amenities and  $w_t^k$  is disposable income, both of which are directly observed in our data. Finally, we define  $\lambda_j \equiv -\log \Lambda_j$ ,  $\lambda_t \equiv -\log \Lambda_t$ ,  $\phi_{sjt} \equiv -\log \varphi_{sjt}$ , and  $\beta_s^k \equiv \alpha_s^k / (\sigma_s \Lambda_s)$ .

unobservable entry costs:

$$\mathbb{E}[S_{jt}^{\tau(k)} \phi_{sjt} | \lambda_j, \lambda_t] = 0.$$

We calibrate  $\eta$  following the estimates of [Eckert, Ganapati and Walsh \(2020\)](#), who also estimate an endogenous entry cost. After the appropriate transformations, we set  $\eta = -0.33$ . We calibrate this parameter for two reasons. First, the model specified in equation 8 contains location  $j$  and time  $t$  fixed effects. When regressing  $\log N_{jt}$  on those fixed effects, we obtain an  $R^2 = 0.99$ , so there is virtually no remaining variation to identify  $\eta$ . Second, observe that the variable  $N_{jt}$  is also endogenous by construction as  $N_{jt} = \sum_s N_{sjt}$ . Therefore, the estimation of  $\eta$  would require an additional instrument.<sup>15</sup>

**Implementation and results.** We choose six consumption amenities: Touristic Amenities, Restaurants, Café Bars, Food Stores, Non-food Stores, and Nurseries.<sup>16</sup> We simultaneously estimate the parameters in equation 8 for all amenities.<sup>17</sup> To do so, we interact each of our instruments with a dummy variable for each amenity  $s$ ,  $Z_{sjt}^k = \mathbb{1}_s Z_{jt}^k$  to construct the following moment conditions,

$$\mathbb{E}\left[g^{sk}(\lambda_j, \lambda_t, \beta_s^k)_{sjt}\right] = \mathbb{E}\left[Z_{sjt}^k \phi_{sjt}\right] = 0.$$

We also impose the following the orthogonality conditions for the location and time fixed effects:

$$\mathbb{E}\left[g^j(\lambda_j, \lambda_t, \beta_s^k)_{sjt}\right] = \mathbb{E}\left[\lambda_j \phi_{sjt}\right] = 0, \quad \text{and} \quad \mathbb{E}\left[g^t(\lambda_j, \lambda_t, \beta_s^k)_{sjt}\right] = \mathbb{E}\left[\lambda_t \phi_{sjt}\right] = 0.$$

We stack all these moments together to form a final vector of moment conditions:

$$\mathbb{E}\left[g(\lambda_j, \lambda_t, \beta_s^k)_{sjt}\right] = \mathbb{E}[Z_{sjt} \phi_{sjt}] = 0.$$

---

<sup>15</sup>For robustness, we also run a model with only time fixed effects,  $\lambda_t$ , and setting the instrument for  $N_{jt}$  as the total number of units that do not have a residential purpose. When doing so, we find  $\eta = -0.11$  which is in line with the range of estimates derived from [Eckert et al. \(2020\)](#), [-0.33, -0.15], after appropriate transformations. We prefer a model that includes location fixed effects because it can capture important attributes that are constant over time but vary across locations, such as land-use restrictions or zoning regulations.

<sup>16</sup>For the first five amenities, we observe the number of establishments. However, for Nurseries we observe the number of available seats in a given location. We assume that the number of seats is constant across nursery establishments. Café bars in Amsterdam should not be confused with coffee-shops. In the former, it is very common to order coffee, tea, or alcoholic beverages. The latter is meant for the consumption of cannabis products.

<sup>17</sup>Observe that the model is not fully identified. First, the full set of dummies corresponding to  $\lambda_j$  and  $\lambda_t$  is collinear. To avoid this collinearity, we set  $\lambda_{j=1} = 0$ . Second, observe that for any given  $c$ , the set of parameters

$$\tilde{\beta}_s^k = c \beta_c^k \quad \text{and} \quad \tilde{\lambda}_t = \lambda_t - c$$

is observationally equivalent. To avoid this issue, we normalize  $\lambda_{t=T} = 0$ .

Table 3: Estimates of Amenity Supply Parameters

Group	Touristic Amenities	Restaurants	Café Bars	Food Stores	Non-Food Stores	Nurseries
Older families	59.944 [0,0,218.18]	0.0 [0,0,16.297]	0.0 [0,0,0,0]	0.0 [0,0,11.998]	2.271 [0,0,25.707]	415.243*** [186.264,837.487]
Singles	364.062 [0,0,833.441]	59.441 [0,0,148.899]	0.0 [0,0,0,0]	52.182 [0,0,167.529]	0.0 [0,0,43.415]	0.0 [0,0,0,0]
Younger families	0.0 [0,0,0,0]	0.0 [0,0,13.121]	3.543 [0,0,21.808]	29.255** [0.729,58.678]	107.138*** [50.957,158.689]	387.489* [0,0,672.534]
Students	488.828* [0,0,1072.092]	199.533*** [76.883,288.674]	21.44 [0,0,40.371]	54.437 [0,0,129.194]	0.0 [0,0,0,0]	0.0 [0,0,729.872]
Immigrant Families	0.0 [0,0,0,0]	0.0 [0,0,9.443]	7.33*** [0.942,29.473]	38.676 [0,0,76.667]	43.796* [0,0,147.762]	153.907 [0,0,663.999]
Dutch Low-Income	0.0 [0,0,137.308]	0.0 [0,0,22.976]	0.0 [0,0,0,0]	0.0 [0,0,36.584]	0.0 [0,0,0,0]	0.0 [0,0,0,0]
Tourists	435.917*** [328.271,582.922]	200.103*** [163.424,240.117]	113.284*** [76.9,130.32]	71.219*** [42.979,93.96]	368.742*** [276.691,430.773]	0.0 [0,0,0,0]

Note: This table presents estimates of coefficients  $\beta_s^k$  from Equation 8 for seven household types and six types of services using a three-way panel of 22 districts in the Amsterdam County, establishment counts for each service, and the time period 2008-2018. The estimates are an output of a constrained GMM estimator, imposing the constraint that each coefficient is positive as implied by the microfoundation of the model in Section A.1. The estimation procedure is outlined in section 5.2. Bayesian-bootstrap with random weights 95% confidence intervals using Dirichlet weights are reported in brackets. Location and time fixed effects are not reported for the ease of exposition.  
 $*p < 0.10$ ,  $**p < 0.05$ ,  $***p < 0.01$ .

To ensure our optimization problem is well-defined, we impose the structural conditions that  $\beta_s^k \geq 0$  so that  $\log \left( \sum_k \beta_s^k X_{jt}^k \right)$  always exists. Concretely, we solve for the following constrained optimization problem:

$$\max_{\lambda_j, \lambda_t, \beta_s^k} \hat{g}(\lambda_j, \lambda_t, \beta_s^k)'_{sjt} \hat{W} \hat{g}(\lambda_j, \lambda_t, \beta_s^k)_{sjt} \quad \text{s.t. } \beta_s^k \geq 0 \quad \forall s, k,$$

where  $\hat{W} = (Z_{sjt} Z_{sjt}')^{-1}$ . Because many of our estimates lie on the boundary, that is  $\hat{\beta}_s^k = 0$ , standard inference does not apply. Therefore, we construct standard errors via a Bayesian bootstrap procedure using random weighting (Shao and Tu, 2012). Results are shown on Table 3.

Our estimates generally align with expected differences in consumption patterns across demographic groups. First, Tourists have a positive parameter and significant parameter for all consumption amenities except Nurseries. Second, Singles, Students and Tourists have the largest estimates for Touristic Amenities, although only significant for the last two groups. Similarly, incentives to enter for Restaurants is highest in the presence of Students and Tourists. Finally, the three groups of families are the only ones with positive estimates on Nurseries, but it is only significant for the first two.

## 5.3 Housing demand

We estimate household preference parameters using the “Euler Equations in Conditional Choice Probabilities” (ECCP) estimator, building on [Aguirregabiria and Mira \(2010\)](#), [Scott \(2013\)](#), and [Kalouptsidi et al. \(2021\)](#). The ECCP estimator is particularly well suited for our application since we can leverage the assumption that location capital is lost whenever a household moves. The method allows us to recover parameters *without* solving value functions and without the need to specify beliefs, thus reducing computational burden. In what follows, we describe the assumptions required for the estimation procedure.

**Assumptions.** We assume the state variables  $\{x, \omega, \epsilon\}$  follow a Markov process, along with the following standard assumptions:

1. **Atomistic agents:** the market-level state  $\omega$  evolves according to a Markov process that is unaffected by individual-level decisions  $j$  or states  $\{x, \epsilon\}$ ,

$$p(\omega'|j, x, \omega, \epsilon) = p(\omega'|\omega).$$

2. **Conditional independence:** the transition density for the Markov process factors as,

$$p(x', \omega', \epsilon' | j, x, \omega, \epsilon) = p_x(x' | j, x, \omega) p_\omega(\omega' | \omega) p_\epsilon(\epsilon').$$

3. **Payoff to the outside option:** The flow payoff of living outside the city, excluding location capita, moving costs, and common time components, is normalized to zero:<sup>18</sup>

$$\delta_0^k + \delta_r^k \log r_{0t} + \delta_a^k \log a_{0t} + \delta_b^k \log b_{0t} + \xi_{0t}^k = 0.$$

The ECCP estimator is a two-step estimator. First, conditional choice probabilities (CCP) are estimated directly from the data. For this first stage, we predict CCPs using a multinomial logit that exploits information about the conditional state. We show in Appendix S.3.4.2 that this approach reduces the finite sample bias relative to a non-parametric approach that estimates CCPs using frequency estimators. Second, the CCPs are plugged into a regression equation that relates differences in the likelihood of two different paths to differences in their flow pay-

---

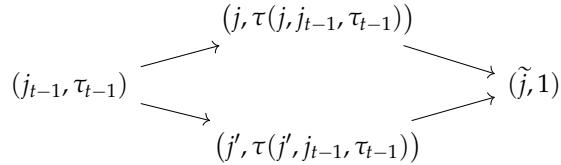
<sup>18</sup>This last assumption allows us to pin down utility levels, as in logit models adding a constant to all choices lead to the same probability. In other words, we can only identify differences in utility  $\delta_j^k - \delta_{j'}^k$ . Following a similar argument, utility components that are common across all choices for a given year  $t$ , captured by  $\delta_t^k$ , cannot be identified. In what follows, we interpret  $\delta_j^k$  as the time invariant component of group  $k$  utility for location  $j$  relative to the outside option.

offs. To formalize how the regression equation is constructed, we first introduce the concept of renewal actions.

**Renewal actions.** Two paths of actions are said to exhibit *finite dependence* if after a finite number of periods, the distribution of future states is the same (Arcidiacono and Miller, 2011). In our model, finite dependence appears whenever two households living in different initial locations,  $j$  and  $j'$ , choose to move to the same new location  $\tilde{j}$ . We call such an action a *renewal action*, because the location tenure is reset, and hence the distribution of future states is the same for both households. Because expectations of future payoffs are unobservable to the econometrician, a key difficulty in estimating dynamic models is disentangling variation in current payoffs from continuation values. Renewal actions separate these two components by equalizing continuation values, thus leaving differences in choice probabilities being solely a function of differences in per-period payoffs.

Concretely, let  $\tau(j, j_{t-1}, \tau_{t-1})$  be the function that maps action  $j$  and state  $x_t = (j_{t-1}, \tau_{t-1})$  to current location capital. Consider the following path represented by Figure 6: let  $j$  and  $j'$  denote actions chosen at state  $x_t = (j_{t-1}, \tau_{t-1})$ , reaching states  $x_{t+1} = (j, \tau(j, j_{t-1}, \tau_{t-1}))$  and  $x'_{t+1} = (j', \tau(j', j_{t-1}, \tau_{t-1}))$ , respectively, and let  $\tilde{j}$  be a renewal action chosen at time  $t + 1$ .

Figure 6: Depiction of path combinations used in the estimation.



From such a path we can derive our main regression equation,

$$Y_{t,j,j',\tilde{j},x_t}^k = u_t^k(j, x_t) - u_t^k(j', x_t) + \beta \left[ u_t^k(\tilde{j}, x_{t+1}) - u_t^k(\tilde{j}, x'_{t+1}) \right] + \tilde{v}_{t,j,j',x_t}^k$$

where,  $Y_{t,j,j',\tilde{j},x_t}^k \equiv \log \left( \frac{\mathbb{P}_t^k(j, x_t)}{\mathbb{P}_t^k(j', x_t)} \right) + \beta \log \left( \frac{\mathbb{P}_{t+1}^k(\tilde{j}, x_{t+1})}{\mathbb{P}_{t+1}^k(\tilde{j}, x'_{t+1})} \right)$ . (9)

On the left hand side,  $Y_{t,j,j',\tilde{j},x_t}^k$  is the likelihood of path  $\{x_t, x_{t+1}\}$  relative to path  $\{x_t, x'_{t+1}\}$ . On the right hand side, we have differences in flow payoffs for the two periods in which the paths diverge, as well as an expectational error  $\tilde{v}_{t,j,j',x_t}^k$ . The key observation is that at time  $t + 1$ , when two agents of the same type  $k$  choose the renewal action  $\tilde{j}$ , they both move to the same individual state and hence their future expected payoffs are the same. Therefore, the value functions from each

path have cancel each other out at  $t + 1$ . Equation 9 conveys that intuition: differences in the likelihood of path  $(j_{t-1}, j, \tilde{j})$  relative to path  $(j_{t-1}, j', \tilde{j})$  are explained solely by differences in utility flows and not in terms of expectations. Finally, if we choose  $j' = 0$ , substitute the functional form for flow utility into 9, and impose assumption 4, we obtain the parameterized version of our regression equation,

$$Y_{t,j,\tilde{j},x_{it}}^k = \delta_j^k + \delta_t^k + \delta_r^k \log r_{jt} + \delta_a^k \log a_{jt} + \delta_b^k \log b_{jt} + \delta_\tau^k \Delta\tau_{it} - \Delta MC_{it}^k + \tilde{\xi}_{t,j,x_{it}}^k, \quad (10)$$

where,

$$\begin{aligned} \Delta\tau_{it} &\equiv \tau'(j, x_{it}) - \tau'(0, x_{it}), \\ \Delta MC_{it}^k &\equiv [MC^k(j, j_{it-1}) - MC^k(0, j_{it-1})] - \beta [MC^k(\tilde{j}, j) - MC^k(\tilde{j}, 0)], \end{aligned}$$

and the last term is the sum of the unobservable time-varying location quality and an expectational error,  $\tilde{\xi}_{t,j,x_{it}}^k = \xi_{jt}^k + \tilde{v}_{t,j,x_{it}}^k$ .

**Implementation.** In practice, neighborhoods in our empirical application are defined as districts (“gebied”), of which there are 22 across Amsterdam.<sup>19</sup> We also define the outside option as any location outside Amsterdam. Our final location choice panel covers 2008 to 2020. We define our market as households that have been observed living in Amsterdam at least once between 2008 and 2020. We set our discount value  $\beta$  equal to 0.85. We discretize the location tenure space similar to Rust (1987). To keep the number of states low, we define two buckets of location capital corresponding to less than three years of tenure or more than four. Appendix S.3.4 shows the technical details of how to deal with such discretization of the state space. Overall, each group has a total of 46 states per year (23 past locations times two location capital states). In our demand estimation for locations, we focus on the first three groups – Older Families, Single Households, and Younger Families—and abstract away from the location decisions of the last three groups—Students and both groups in social housing.<sup>20</sup> The reason to treat students separate from the other households is motivated by evidence that suggests that the housing market for students looks substantially different from the traditional housing market.<sup>21</sup> On the other hand, houses in the social segment of the market are not

---

<sup>19</sup>In 2020, there are 25 districts. However, we combine three districts at the border of the city with their closest district because they have few households living in them.

<sup>20</sup>Even though we do not endogenize the location choice of these three groups, we still keep them as part of locations’ demographic composition as they represent a big part of the population.

<sup>21</sup>Amsterdam is a city with many college students. Some universities offer their own housing options but those tend to be limited. There are multiple popular websites that function as student-specific platforms. Due to the tightness in the housing market, the municipality of Amsterdam recommends securing accommodation well in advance. See <https://www.iamsterdam.com/en/study/steps-to-study-in-amsterdam/finding-housing-in>

allocated through a market but through a centralized application wait list. Hence, we are unable to infer their preferences without further information.

**Identification.** First, the vector of neighborhood characteristics  $b_{jt}$  beyond rent and amenities contains the log of the average apartment size in square meters and the log of social housing units. We assume that the structural error  $\tilde{\xi}_{t,j,x_{it}}^k$  is orthogonal to these housing characteristics, location fixed effects, tenure and moving costs:

$$\mathbb{E}[\tilde{\xi}_{t,j,x_{it}}^k | \delta_t^k, \delta_j^k, \log b_{jt}, \Delta\tau_{it}, \Delta MC_{it}^k] = 0 \quad \forall k.$$

On the other hand, our equilibrium definition already implies that the structural shocks  $\tilde{\xi}_{t,j,x_{it}}^k$  are correlated with neighborhood prices  $r_{jt}$  and amenities  $a_{jt}$ . In the data, we may expect unobservable neighborhood trends, such as gentrification, that correlate with prices, amenities, as well as moving decisions. For that reason, we estimate our structural parameters via optimal two-step GMM using the following moment conditions:

$$\mathbb{E}[Z_{jt}\tilde{\xi}_{t,j,x_{it}}^k] = 0 \quad \forall k,$$

where  $Z_{jt}$  is a vector of instruments. Because both prices and amenities are endogenous, we need  $A + 1$  number of instruments, where  $A$  is the number of amenities. Observe that under rational expectations, it follows that

$$\mathbb{E}[Z_{jt}\tilde{v}_{t,j,x_{it}}^k | \delta_j, \log b_{jt}, \Delta\tau_{it}, \Delta MC_{it}^k] = 0 \quad \forall k,$$

as  $\mathbb{E}[\tilde{v}_{t,j,x_{it}}^k | \mathcal{I}_t] = 0$  for all  $j, t$ , and  $x_{it}$ , where  $\mathcal{I}_t$  is the set of variables that have been realized at time  $t$  or before. Therefore, it suffices to find instruments that are orthogonal to unobservable demand shocks:

$$\mathbb{E}[Z_{jt}\tilde{\xi}_{t,j,x_{it}}^k | \delta_j, \log b_{jt}, \Delta\tau_{it}, \Delta MC_{it}^k] = 0 \quad \forall i, k, j, t.$$

For that reason, we consider  $Z_{jt}$  that can be interpreted as supply shocks. Because we have six amenities, we construct seven instruments in total. Three of those instruments are Bartik-type shocks that leverage three policy changes that can be effectively treated as supply shocks to the tenancy composition of the housing stock. Concretely, new regulations on the rental market were introduced in 2011, 2015, and 2017 that changed the incentives of landlord and housing associations to supply their unit as social housing, a private market unit, or as a short-term rental, respectively. For full details, see Appendix S.1 for the institutional context and Appendix S.1.3 for details on the policy changes. To introduce spatial variation, we interact a dummy that turns one after the introduction of the policy and the log of the units in the housing market segment affected by the shock in the previous

---

amsterdam.

year. Additional instruments are the log number of housing units that are demolished inside the location  $j$  as well as outside the precinct, which we also interpret as supply shocks.<sup>22</sup> Finally, we follow [Bayer et al. \(2007\)](#) and construct two more instruments by using variation in changes of social housing units and the average apartment size in other areas of the city outside the precinct. Using these instruments, we find that the first stage regression of a 2SLS estimation has an F-stat of 139.1. See Appendix [S.3.4](#) for full details on the demand estimation procedure.

**Results.** Table [4](#) shows estimates of the demand parameters for neighborhood characteristics, moving costs, and location capital, for the three groups used in our counterfactual simulations. Overall, our results align with expected differences across demographic groups. All groups exhibit that moving is costly with Older Families having the highest moving costs. All households derive positive utility from the accumulation of location capital. Estimates for rent are negative throughout, as is to be expected. Older families with children are the most sensitive to rent. In terms of amenities, the first two groups perceive a negative payoff from Touristic Amenities. Restaurants show a positive and significant coefficient for Singles, but is not significant for the rest of the groups. Café Bars show a negative and significant coefficient for the first two groups. Non-food stores are positively valued by all groups. Finally, Nurseries are positive and significant for the two groups of families.

## 5.4 Housing supply

Our estimating equation for the supply of long- relative to short-term units is,

$$\log s_{jt}^L - \log s_{jt}^S = \alpha (r_{jt} - p_{jt}) + \kappa_j + \kappa_t + \nu_{jt},$$

where we have parameterized the operating cost wedge  $\kappa_{jt}$  into location- and time-fixed effects, and  $\nu_{jt}$  stands for any remaining unobservables varying at the  $jt$  level.

**Instruments.** OLS estimation will lead to classic simultaneity bias because we are effectively regressing relative prices on relative quantities. The solution is an instrument that shifts relative demand for short- versus long-term units. We use predicted tourist demand from a shift-share research design: the “shift” part of the instrument exploits time variation in worldwide demand for Airbnb as proxied by online search volume ([Barron et al., 2021](#)), while the “share” part constructs neighborhood-level exposure to the shift from the historic spatial distribution of touristic attractions. The relevance condition is testable and straightfor-

---

<sup>22</sup>A precinct (stadsdeel) is a larger geographical unit containing districts (Gebied). There are seven of them in Amsterdam.

Table 4: Preference parameter demand estimation results

	Older Families	Singles	Younger Families
Intra-City Moving Cost	-5.492*** (0.015)	-4.969*** (0.011)	-5.026*** (0.012)
Bilateral Moving Cost	-0.169*** (0.001)	-0.148*** (0.001)	-0.118*** (0.001)
In/Out of City Moving Cost	-4.408*** (0.012)	-4.012*** (0.009)	-4.044*** (0.010)
High Location Capital	0.185*** (0.017)	0.211*** (0.013)	0.263*** (0.013)
Log Rent	-11.769*** (1.201)	-2.523** (0.987)	-2.340** (1.045)
Log Tourism Offices	-1.193*** (0.169)	-0.449*** (0.143)	0.299** (0.144)
Log Restaurants	0.281 (0.284)	0.729*** (0.251)	-0.195 (0.242)
Log Café Bars	-0.822*** (0.092)	-0.547*** (0.079)	-0.081 (0.082)
Log Food Stores	-2.000*** (0.324)	-1.314*** (0.280)	-0.600** (0.289)
Log Nonfood Stores	0.700** (0.341)	1.626*** (0.299)	1.429*** (0.296)
Log Nurseries	1.763*** (0.172)	0.076 (0.141)	0.316** (0.148)
Location FE	✓	✓	✓
Time FE	✓	✓	✓
Neighborhood Controls	✓	✓	✓
N	233772	233772	233772

Notes: This table presents regression results of preference parameters for a dynamic location choice model for 22 districts in Amsterdam for 2008-2019. We estimate preference parameters separately for four demographic groups via two-step optimal GMM. We normalize the utility flow of living outside Amsterdam equal to zero. After taking differences with respect to the outside option, each type has 46 possible states (23 past locations and two location capital categories), 22 possible actions and 21 possible renewal actions over 11 years, which leads to 233,772 possible states and two-step path combinations. We leave out exogenous controls for the ease of exposition but those include the log of social housing units and the log of the average apartment in square meters. Two-step efficient GMM standard errors in parenthesis. \* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ .

ward: higher demand from tourists will raise short- relative to long-term rental prices. The exclusion restriction holds as long as predicted tourist demand is uncorrelated with the evolution of cost shocks affecting landlord decisions to choose the short- over the long-term market.

**Results.** Table 5 presents our estimates for  $\alpha$ , the landlord’s marginal utility of income, which are fairly stable across specifications. OLS estimates are downward-biased, as expected with simultaneity bias. Our preferred specification is the instrumental variable regression with two-way fixed effects.<sup>23</sup>

Table 5: Long-term (LT) relative to short-term (ST) housing supply elasticities

	Dependent variable: $\ln(\text{LT share}) - \ln(\text{ST share})$							
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
LT price - ST price	0.144*	0.354***	0.140*	0.360***	0.096	0.341***	0.020	0.241
	(0.081)	(0.104)	(0.083)	(0.112)	(0.084)	(0.089)	(0.106)	(0.495)
Year FE			X	X			X	X
Wijk FE					X	X	X	X
First stage F-stat		69.22		23.94		14.72		15.82
Observations	271	271	271	271	271	271	271	271

Notes: The table reports estimates of landlords’ marginal utility of income for a static model of housing supply to the market for short- and long-term rentals for 92 wijks in Amsterdam for 2015–2017. Prices are instrumented using a “shift-share” instrument (Barron et al., 2021) that proxies for demand shocks: the “shift” part of the instrument exploits time variation in worldwide demand for Airbnb as measured by online search volume, while the “share” part constructs neighborhood-level exposure to the shift from the historic spatial distribution of touristic attractions. Construction of Airbnb supply and prices is described in Section S.2. Wijk-level clustered standard errors in parenthesis. \* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ .

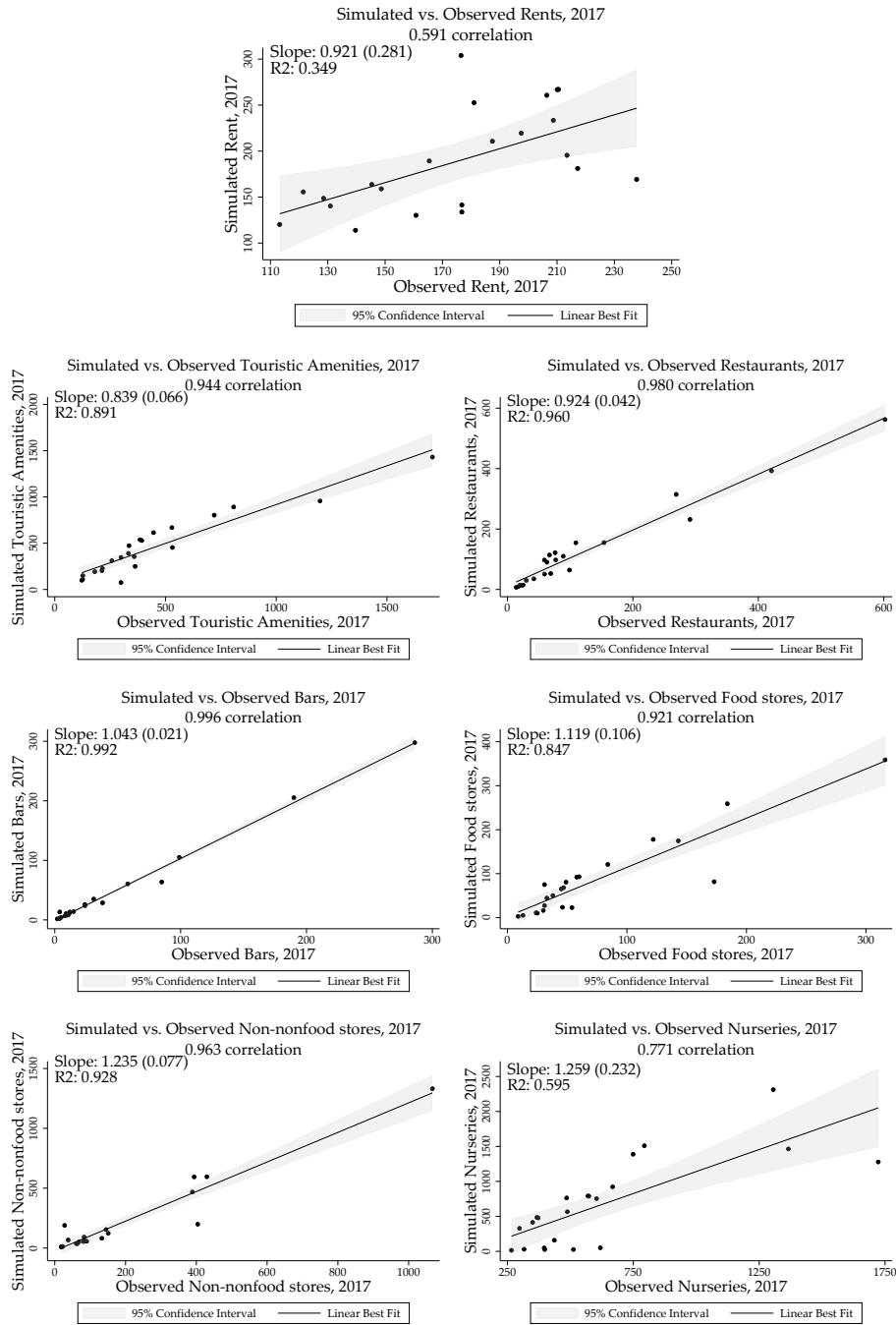
## 5.5 Model fit

In this section we show how our model fits the data and simulate a steady-state equilibrium for 2017. We assume that agents have perfect foresight. We find our equilibrium using a nested fixed-point algorithm outlined in Section S.4.2. We take our preferred housing supply estimate from Section 5.4 and calibrate the landlords’ differential costs to match the number of Airbnb tourists in each location in 2017. Further details are in Section S.5.1. Figure 7 plots the simulated endogenous objects—rents and amenities—against the observed objects in the data. Recall that while we use rent and amenity variation, we are not directly targeting the fit of our simulated vectors against the observed vectors in our estimation.

---

<sup>23</sup>We believe that the lack of significance arises from little within neighborhood variation due the short-panel nature of our data. In any case, the estimates across specifications appear rather stable.

Figure 7: Model fit.



Notes: The figure presents scatter plots, linear fit, and 95% confidence intervals of the simulated objects against the observed objects for a simulated equilibrium of a dynamic model of residential choice with endogenous amenities and Airbnb tourists for 22 districts in Amsterdam County in 2017. The model is defined in Section 4 and estimated details are presented in Section 5. To compute the equilibrium, we used a nested fixed-point algorithm, outlined in Section S.4.2, initiated at the observed prices and amenities.

There are several reasons why our simulated economy may differ from the observed data. First, we assume that 2017 is in steady state and that people have perfect foresight, which is not necessarily true in the data.<sup>24</sup> Second, there could be estimation error. Third, there could also be model mis-specification. However, despite these possible sources of divergence, we generally see that our simulated equilibrium is able to reproduce the observed equilibrium with high precision. We take these results as evidence that our model, estimated parameters, and equilibrium assumptions are a good approximation of the economic forces present in the real world.

### 5.5.1 Multiplicity of equilibria

Given that endogenous amenities act as agglomeration forces, the model may feature multiple equilibria. Computationally, we find multiple equilibria by initiating the equilibrium algorithm solver in subsection S.4.2 from many different starting values. Therefore, we define an equilibrium selection rule as the resulting equilibrium when we set the initial value of our algorithm solver equal to the observed equilibrium. Using this selection rule, we see that our model can reproduce the patterns observed in the data fairly accurately, as shown in subsection 5.5. Moreover, we only find a unique equilibrium when sampling initial values from a neighborhood centered at the observed equilibrium. We take this result as evidence that our selection rules leads to a locally unique equilibrium and therefore we can think of our counterfactuals as locally stable.

## 6 Counterfactuals

In this section we quantify the importance of different model components, such as preference heterogeneity, and evaluate the welfare implications of different counterfactual scenarios. In general, we focus on renter's consumer surplus (CS), in log wages, and measure welfare changes between a baseline and a counterfactual scenario for type  $k$  as its *Consumption Equivalent*,  $CE^k$ . Concretely, we define  $CE^k$  as the monetary compensation required to leave a type- $k$  household on the same ex-ante utility level as the baseline scenario. Hence, a positive  $CE^k$  means that the average type- $k$  household is worse off in the counterfactual relative to the baseline scenario and vice versa. Appendix S.5.6 describes in detail how we construct our different welfare measures.

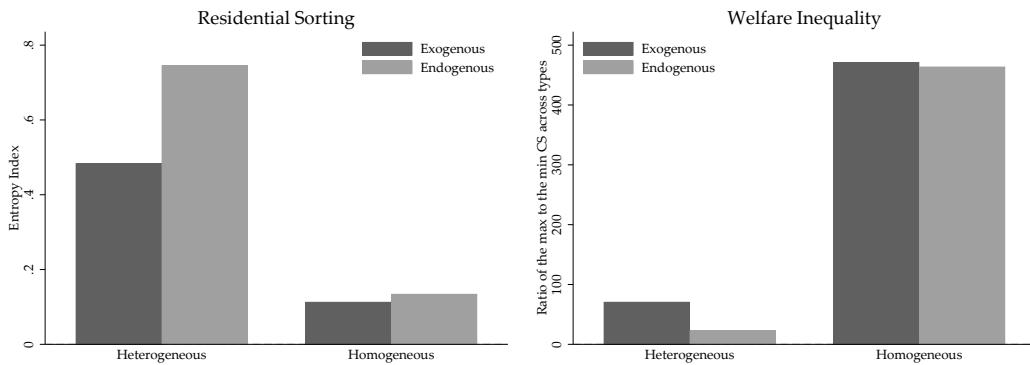
---

<sup>24</sup>Observe that our estimation method does impose any of these assumptions.

## 6.1 Role of preference heterogeneity for sorting and inequality

We first evaluate how preference heterogeneity interacts with the endogeneity of amenities to determine spatial sorting and welfare inequality. We solve the model under our estimated heterogenous preference specification and compare equilibrium outcomes with exogenous amenities to those with endogenous amenities. We then repeat this exercise for a model with homogenous preferences over amenities.

Figure 8: Residential sorting and welfare inequality



Notes: The panel on the left reports the entropy index, a commonly used measure of segregation of household types across districts (see Appendix S.5.4 for a formal definition). The panel on the right reports the welfare gap across household types, measured as the ratio of the consumer surplus in log wages of the highest-welfare household type relative to the lowest-welfare household type.

The left panel of Figure 8 shows residential sorting increases when amenities are endogenous, and more so when preferences are heterogenous. The right panel of Figure 8 shows that despite the increased sorting across space, welfare inequality across household types can *decrease* when amenities are endogenous, especially if preferences are heterogenous. The intuition is that with heterogenous preferences we have more sorting than with homogenous preferences, and as a result neighborhoods become more differentiated in terms of their amenities. Concretely, Table 6 shows the spatial distribution of almost every amenity becomes more concentrated. This misalignment of preferences over amenities implies household types do not compete with each other for the same residential locations, thus keeping rental prices low while obtaining their preferred amenities.

## 6.2 Decomposing welfare effects of the short-term rental industry

Next, we use our model to evaluate the effect of short-term rental entry on Amsterdam resident welfare, which we measure as renter surplus. We model this entry

Table 6: Neighborhood differentiation as spatial dispersion of amenities

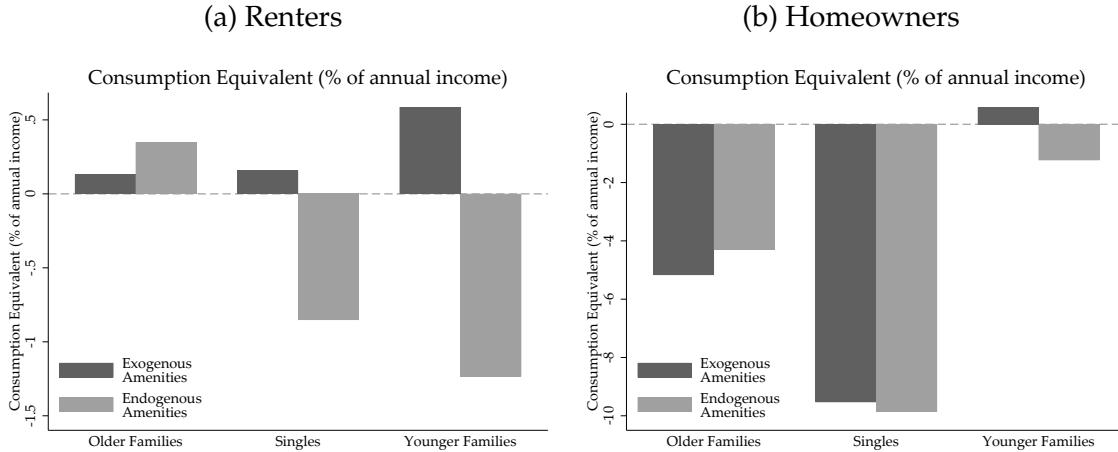
Amenity	Homogeneous preferences			Heterogeneous preferences		
	Exogenous	Endogenous	$\Delta$	Exogenous	Endogenous	$\Delta$
Touristic amenities	0.38	0.44	0.06	0.38	0.38	0
Restaurants	0.54	0.54	0	0.54	0.53	-0.01
Bars	0.66	0.65	-0.02	0.66	0.67	0.01
Food stores	0.46	0.47	0.01	0.46	0.51	0.05
Non-food stores	0.58	0.56	-0.01	0.58	0.65	0.07
Nurseries	0.29	0.43	0.14	0.29	0.49	0.2

Notes: Columns labelled “Exogenous” and “Endogenous” report the Gini index (computed across neighborhoods) for each neighborhood characteristic. “ $\Delta$ ” reports the difference between the “Endogenous” and “Exogenous” columns. A positive  $\Delta$  implies the spatial distribution of the characteristic becomes more unequal.

as a reduction in landlord costs of renting in the short-term market, specifically due to the technology that enables them to match easier with guests. Our goal is to disentangle the welfare effects for residents into two components: changes in rent and changes in amenities. This decomposition allows us to separate the direct effect of Airbnb entry on rent through the reduction in housing supply from the indirect effect on the endogenous adjustment of amenities due to the increased tourist population. To isolate rent effects from amenity effects, we run a restricted version of the counterfactual where amenities are kept exogenously fixed at baseline. The key insight behind our results is that while all residents lose from higher rent, some lose and some win from the changes in amenities due to preference heterogeneity.

Our main results are presented in Figure 9, which reports welfare effects in consumption equivalent terms: how much extra income a household must be given in the counterfactual with short-term rentals in order to be just as well off as in the baseline equilibrium without them. The dark grey bars in Panel (a) of Figure 9 show by how much each household needs to be compensated in the exogenous amenities counterfactual. Every household loses because all Airbnb entry does in this restricted counterfactual is to reduce housing supply and raise rents, as shown in the left panel of Figure 10. However, the effects are regressive: Older Families, which are predominantly Dutch and higher-income, lose the least while Younger Families, which tend to be poorer and of immigrant background, lose the most. Hence, by only taking into account rent effects, we would conclude that the entry of short-term rentals increases inequality.

Figure 9: Decomposition of welfare effects from the entry of short-term rentals.



Notes: The consumption equivalent is computed as how much extra income a household must be given in the counterfactual with short-term rentals in order to obtain the same surplus as in the baseline equilibrium without short-term rentals. Hence, positive values indicate a welfare loss in the counterfactual with short-term rentals. Left panel reports changes in renter's welfare. Right panel reports changes in homeowner's welfare. This final homeownership-adjusted monetary value is then reported as a percentage of the household's income. "Homeownership-adjusted" means that in our simulations, rental income is rebated back to households who are homeowners as a city-wide uniform lump sum transfer. See Appendix for S.5.6 for details.

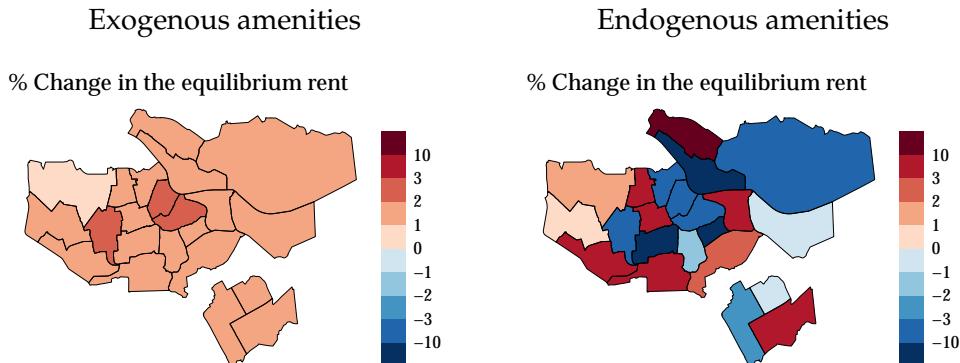
The light grey bars in Figure 9 Panel (a) show the welfare effects when amenities are allowed to endogenously respond to residential composition. Older Families lose more than when amenities were exogenous because on top of facing higher rent they also lose the amenities they value most. Singles and Younger Families now obtain welfare gains because they face an increase in the amenities they like, and this is large enough to outweigh the welfare losses from higher rent. Thus, as we move from exogenous to endogenous amenities our qualitative results are reversed: the entry of short-term rentals now *reduces* welfare inequality.

So far, we have discussed welfare as renter surplus, ignoring whether households gain from an increase in rental income as homeowners. We now relax this restriction by considering the Older Families and Singles to be homeowners while keeping Younger Families as renters. We adjust our simulations by having homeowners pay rent as before, but receiving it back as a lump sum transfer (see Appendix for S.5.6 for details). Panel (b) of Figure 9 shows that homeowners now gain from Airbnb entry. As expected, the entry of short-term rentals increases inequality between homeowners and renters and this increase is mainly driven by gains in rental income.

Figures 10-11 goes into detail on rents, residential movements, amenity changes,

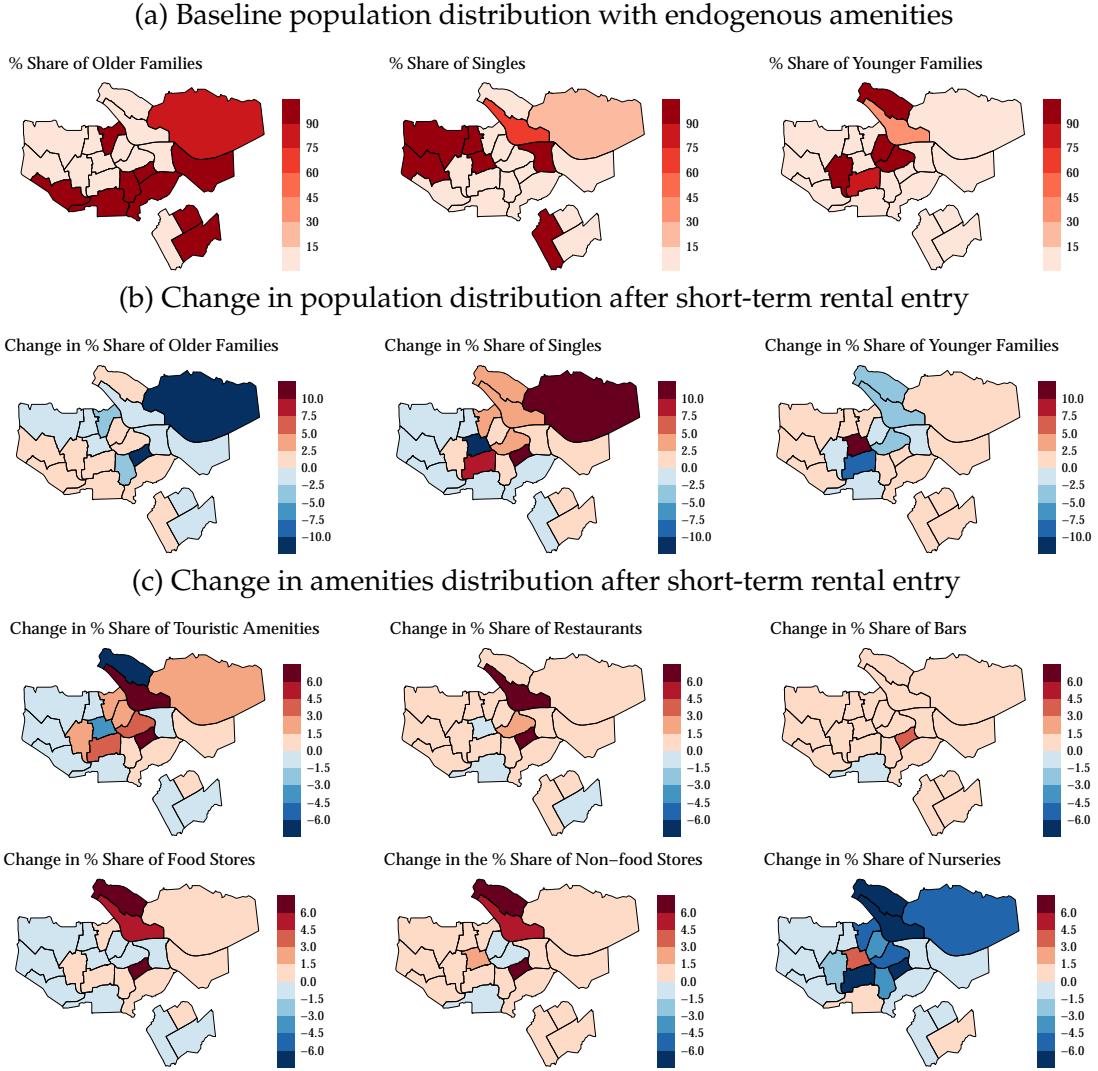
and renter surplus are linked. The top row of the Figure 11 shows that in the baseline equilibrium with endogenous amenities, Older Families tend to live in the south/eastern districts, Singles live mostly in western districts, and Younger Families live in central-western districts. The second row shows that after Airbnb entry the Older Families leave the center-south, Singles leave the west and move towards the center, and Younger Families move west of the center. The third and fourth rows show that these location changes are correlated with amenity changes. Older Families especially leave districts where there is a decline in nurseries (the amenity they value most) and an increase in touristic amenities, which they dislike the most. Hence, on top of higher rent, Older Families lose the amenities they value most in these districts. Singles move to the south-central districts, where there is an increase in the amenities they value most (restaurants and non-food stores). Finally, Younger Families move slightly west, where there is an increase in non-food stores and nurseries, both of which they value.

Figure 10: Rent changes under exogenous and endogenous amenities



Notes: The figures show the percentage change in a neighborhood's equilibrium rent when we simulate short-term rental entry, for the case of exogenous and endogenous amenities. The case of exogenous amenities depicted was computed using amenities from an equilibrium with no short-term rentals and with endogenous amenities.

Figure 11: Spatial distribution at baseline and after short-term rental entry



Notes: All figures correspond to the model with endogenous amenities described in Section 4.4. The top row shows the neighborhood population share of each household type in the baseline without short-term rentals. The second row shows the changes in population shares after short-term rental entry. The third and fourth rows show changes in the neighborhood amenity share of each amenity type, after short-term rental entry. To facilitate comparison between the equilibria, we supplied our equilibrium solver algorithm, described in Section S.4.2, with the same initial values given by the observed vectors of rents and amenities.

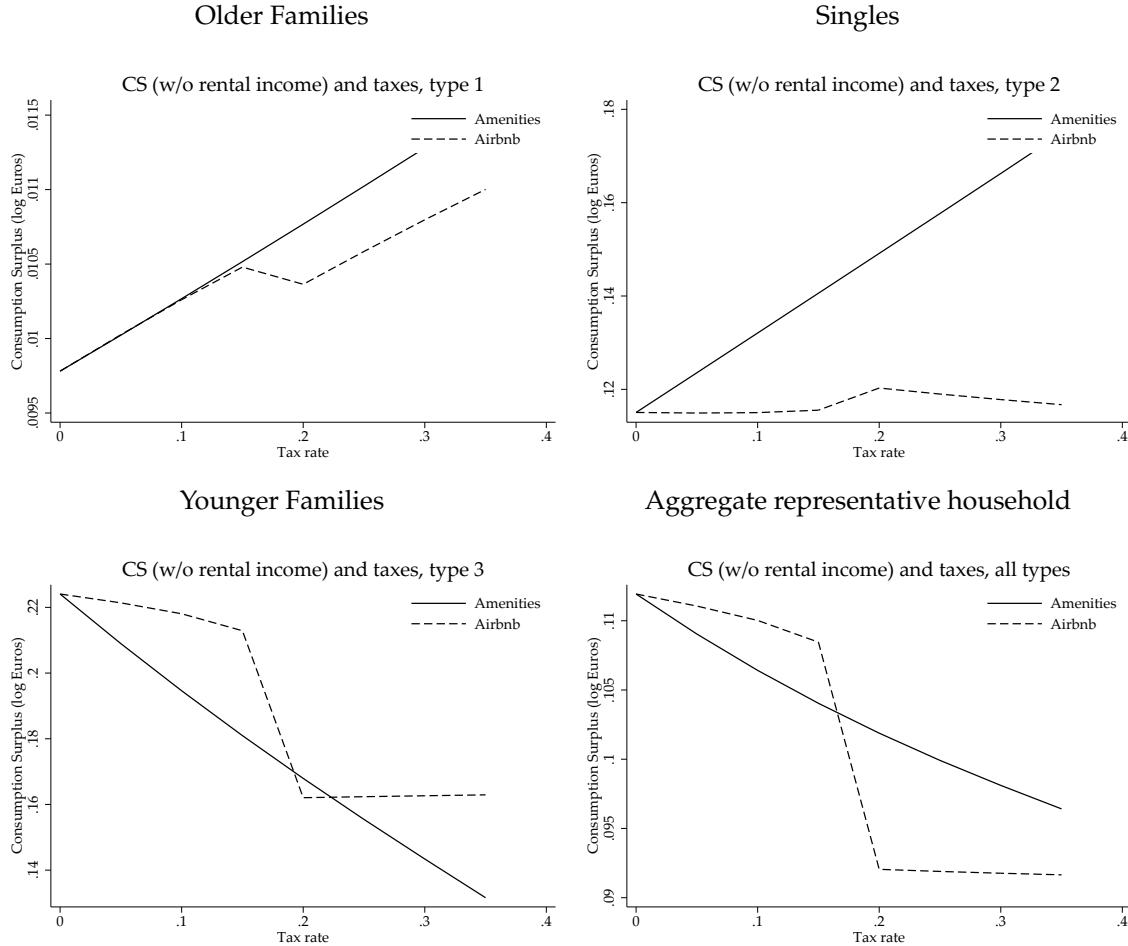
### 6.3 Policy implications for targeting of amenities

Our results from section 6.2 highlight the importance of taking into account preference heterogeneity over both rent and amenities when evaluating the distribu-

tional incidence of shocks to the residential market. In this section, we consider normative implications for policies in both housing and amenity markets. Concretely, we compare a tax on short-term rentals to a tax on touristic amenities. The short-term rental tax is a housing policy: its main goal is to increase the housing supply available for locals and create welfare gains through more affordable rent. However, it may also have indirect equilibrium effects on amenities via the reduction in the tourist population and re-sorting of locals. In particular, it might reduce the types of amenities tourists bring, some of which are not desired by locals, but some of which are. By contrast, the tax on touristic amenities is an amenity-market policy: it is targeted to specifically remove the amenities locals dislike, while keeping the subset of amenities tourists bring in and locals enjoy. It may also have indirect equilibrium effects on the housing market by changing the desirability of certain neighborhoods for locals. Note that it is our microfoundation of the amenities market that allows us to perform such targeted policies, in contrast to models where endogenous amenities are concentrated into one single index.

Figure 12 compares the welfare levels of different household types for a range of tax rates. Older Families and Singles gain from both types of taxes, and their welfare is increasing in the tax rate. The opposite pattern holds for the Younger Families. The reason is that the first two types dislike touristic amenities, while the Younger Families enjoy them. The more interesting observation is that the gap between the two types of taxes is not the same across household types. For Singles, the amenity-tax leaves them much better off than the short-term rental tax, but less so for Older Families. The reason is that although Singles dislike touristic amenities, they enjoy other amenities that tourists bring, such as restaurants. Because the amenity-tax specifically targets touristic amenities, it decouples the undesirable amenities tourists bring from the desirable ones. By contrast, the short-term rental tax reduces the tourist population and all the amenities that come with them in an un-targeted manner. On the other hand, Older Families mainly value Nurseries, which tourists do not bring. Hence, for Older Families there is little welfare gain from differentially targeting the amenities brought by tourists, because all such amenities are relatively undesirable for them. Thus, a short-term rental tax delivers relatively similar welfare gains to a touristic amenity tax for them. To conclude, the desirability of regulating the housing market through a short-term rental tax or the amenity market through a touristic amenity tax depends on the pattern of preference heterogeneity. Concretely, the welfare gains from targeting amenities are highest when households hold very heterogeneous tastes across the various amenities tourists bring because it allows policymakers to specifically target the source of nuisance.

Figure 12: Short-term rental tax vs. Touristic amenity tax (welfare effects)



Notes: The figure reports consumer surplus (in log Euros) for each household type under each type of tax. The exception is the bottom right panel, which reports a representative household aggregated across types, where each type is weighted by population share. Implementation details are in Appendix S.5. Kinks in the Airbnb tax counterfactuals occur due to tipping points in the demographic composition of a few selected neighborhoods, described in Appendix S.5.3.

## 7 Discussion

While our model is rich along many dimensions—dynamic location choices, as well as heterogeneity in both demand for neighborhoods and in the endogenous supply of differentiated amenities—it is tailored to answer a specific set of research questions while remaining silent on others. In this section, we discuss the limitations of our analysis and suggest potential extensions for future work.

**Unified labor market.** Our model assumes households obtain their income from a city-wide integrated labor market rather than neighborhood-level local labor markets. We consider this assumption reasonable for a city as compact as Amsterdam (less than 10 km at its widest) and with first-world public transport infrastructure (which has not experienced major changes during our sample period). Furthermore, our tax return microdata does not specify occupation nor work location, which prevents us from incorporating labor markets into our model in a more explicit way. In that sense, our paper’s main goal is to emphasize the role that housing and amenity markets play for spatial inequality, focusing on location choice within a compact city and abstracting away from employment opportunities as major drivers of location choice.

**Effects of tourism on the labor market.** We have leveraged Amsterdam’s tourism boom as a demand shifter in housing and amenity markets, and quantified the distributional effects on residents in their role as consumers in these markets. However, tourism also affects residents in their role as employees by creating employment opportunities (Faber and Gaubert, 2019). If tourism raises wages mostly in non-tradable service sectors which typically employ lower-skilled workers, then it could have a progressive effect that might counterbalance regressive effects in housing and amenity markets. In a unified labor market, we would expect a tourism shock to create a uniform rise in tourism sector wages across the city, but not to have differential effects across neighborhoods. As mentioned in the prior paragraph, we do not observe household occupations nor work location, so we can only observe wage trends for the household types used in our welfare analysis, which we reassuringly find to be mostly flat during our sample period.<sup>25</sup> Finally, it is worth noting that explicitly incorporating a labor market would not affect our estimation exercise, which relies on across-neighborhood variation under the assumption of a city-wide unified labor market.

**Consuming amenities outside the residential location.** We assume consumers can only access amenities from the location in which they live in. This assumption can be relaxed by allowing for transport costs between residential location and amenity location, but an empirical application would require data on consumption trips across neighborhoods, which we do not have for Amsterdam. To the extent smartphone-based evidence from other cities indeed confirms urban residents tend to consume amenities located near their home (Miyauchi et al., 2021; Allen et al., 2021), our qualitative insights should not significantly change since our

---

<sup>25</sup>See Appendix S.1.2. It is also worth noting that in aggregate data for Amsterdam, tourism has a moderate employment share at around 11% (Fedorova, de Graaf and Sleutjes, 2019), with the bulk of the city’s employment being in financial services and the public sector instead.

mechanisms ultimately depend on amenity consumption choices being correlated with residential location.

## 8 Conclusion

In this paper, we study the role of preference heterogeneity over a set of endogenous location amenities in shaping within-city sorting and welfare inequality. To do so, we build a model of residential choice where heterogeneous forward-looking households consume a set of amenities that are provided by firms in a market for non-tradable services. We leverage increasing tourism flows and the spatial variation in the entry of short-term rentals in Amsterdam as events that shift the demographic composition of neighborhoods, and thus alter local amenities.

First, we show tourism flows and the entry of short-term rental platforms have led to a significant impact on rents, amenities, and within-city migration in Amsterdam. Second, to rationalize our reduced-form findings, disentangle effects on supply from effects on demand, and conduct policy counterfactuals, we build a dynamic spatial equilibrium model of a city’s rental market with heterogeneous forward-looking households, and show how to estimate it using tools from the empirical industrial organization literature. We endogenize and microfound housing supply through landlords’ decisions to rent to locals on traditional leases or full-time to tourists through the short-term rental market. Furthermore, in contrast to most work that collapses amenities into a one-dimensional quality index, we also microfound how different consumption amenities arise in each neighborhood as an equilibrium outcome of a market where firms supplying amenities cater to households that demand them through their perception of neighborhoods as horizontally differentiated products.

We estimate three parts of our structural model using a set of different techniques from the empirical industrial organization literature. On the housing supply side, we find significant heterogeneity of landlords in their operating costs across the long- and short-term rental markets. On the demand side, we estimate location preferences, finding substantial heterogeneity across households in their utility parameters. Furthermore, the preference heterogeneity across groups correlate with sociodemographic status as expected. Finally, the structural parameters of amenity supply indicate important differences in how different services respond to changes in their location demographics.

Armed with our estimated parameters, we explore the role of endogenous amenities in shaping within-city inequality. We find the reinforcement in sorting driven by the endogeneity of amenities can go either way in shaping welfare in-

equality across socioeconomic groups. We find that the sign of this effect depends on how correlated preferences are across groups, with the welfare gap increasing between households whose preferences are substantially aligned and decreasing for those whose preferences are misaligned. Moreover, in quantifying the welfare effects followed by Airbnb entry, we find that while all residents lose from higher rent, some lose and some win from the changes in amenities due to preference heterogeneity, in particular how their preferences are correlated with those of tourists. Finally, we use our model to compare different forms of regulating mass tourism: taxing short-term rentals or taxing touristic amenities. We show that taxing touristic amenities dominates taxing short-term rentals when the preferences of locals are sufficiently heterogeneous over the amenities tourists bring. In such cases of wide heterogeneity, the targeted feature of a tax on touristic amenities is especially welfare-enhancing because it decouples the undesirable amenities brought by tourists from the desirable ones.

## References

- Aguirregabiria, V. and Magesan, A. (2013). Euler equations for the estimation of dynamic discrete choice structural models. In: *Structural Econometric Models*, pp. 3–44. Emerald Group Publishing Limited.
- Aguirregabiria, V. and Mira, P. (2010). Dynamic discrete choice structural models: A survey. *Journal of Econometrics* 156(1):38–67.
- Ahlfeldt, G.M., Redding, S.J., Sturm, D.M. and Wolf, N. (2015). The economics of density: Evidence from the Berlin Wall. *Econometrica* 83(6):2127–2189.
- Allen, T., Fuchs, S., Ganapati, S., Graziano, A., Madera, R. and Montoriol-Garriga, J. (2021). Urban Welfare: Tourism in Barcelona.
- Arcidiacono, P. and Miller, R.A. (2011). Conditional choice probability estimation of dynamic discrete choice models with unobserved heterogeneity. *Econometrica* 79(6):1823–1867.
- Barron, K., Kung, E. and Proserpio, D. (2021). The effect of home-sharing on house prices and rents: Evidence from Airbnb. *Marketing Science* 40(1):23–47.
- Bayer, P., Ferreira, F. and McMillan, R. (2007). A unified framework for measuring preferences for schools and neighborhoods. *Journal of Political Economy* 115(4):588–638.
- Bayer, P., McMillan, R., Murphy, A. and Timmins, C. (2016). A dynamic model of demand for houses and neighborhoods. *Econometrica* 84(3):893–942.
- Berry, S., Levinsohn, J. and Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica* pp. 841–890.

- Berry, S.T. (1994). Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics* pp. 242–262.
- Calder-Wang, S. (2019). The distributional impact of the sharing economy on the housing market. *Job Market Paper*.
- Couture, V., Gaubert, C., Handbury, J. and Hurst, E. (2019). Income growth and the distributional effects of urban spatial sorting. Tech. rep., National Bureau of Economic Research.
- Couture, V. and Handbury, J. (2020). Urban revival in America. *Journal of Urban Economics* 119:103267.
- Davis, M.A., Gregory, J. and Hartley, D.A. (2018). The long run effects of low income housing on neighborhood composition. Tech. rep., Working Paper.
- Davis, M.A., Gregory, J., Hartley, D.A. and Tan, K.T. (2021). Neighborhood effects and housing vouchers. *Quantitative Economics* 12(4):1307–1346.
- Diamond, R. (2016). The determinants and welfare implications of US workers' diverging location choices by skill: 1980-2000. *American Economic Review* 106(3):479–524.
- Diamond, R., McQuade, T. and Qian, F. (2018). The effects of rent control expansion on tenants, landlords, and inequality: Evidence from San Francisco.
- Eckert, F., Ganapati, S. and Walsh, C. (2020). Skilled scalable services: The new urban bias in economic growth. Available at SSRN 3736487 .
- Faber, B. and Gaubert, C. (2019). Tourism and economic development: Evidence from Mexico's coastline. *American Economic Review* 109(6):2245–93.
- Farronato, C. and Fradkin, A. (2018). The welfare effects of peer entry in the accommodation market: The case of Airbnb.
- Fedorova, T., de Graaf, L. and Sleutjes, B. (2019). Tourism trends in Amsterdam Metropolitan Area 2018–2019.
- Garcia-López, M.À., Jofre-Monseny, J., Martínez-Mazza, R. and Segú, M. (2020). Do short-term rental platforms affect housing markets? Evidence from Airbnb in Barcelona. *Journal of Urban Economics* 119:103278.
- George, L. and Waldfogel, J. (2003). Who affects whom in daily newspaper markets? *Journal of Political Economy* 111(4):765–784.
- Guerrieri, V., Hartley, D. and Hurst, E. (2013). Endogenous gentrification and housing price dynamics. *Journal of Public Economics* 100:45–60.
- Handbury, J. (2021). Are poor cities cheap for everyone? Non-homotheticity and the cost of living across US cities. *Econometrica* 89(6):2679–2715.
- Hoelzlein, M. (2019). Two-sided sorting and spatial inequality in cities. *Working paper* .

- Hotz, V.J. and Miller, R.A. (1993). Conditional choice probabilities and the estimation of dynamic models. *The Review of Economic Studies* 60(3):497–529.
- Hsiao, A. (2021). Coordination and commitment in international climate action: Evidence from palm oil. *Unpublished, Department of Economics, MIT*.
- Humlum, A. (2019). Robot adoption and labor market dynamics. *Princeton University*.
- Kalouptsidi, M., Scott, P.T. and Souza-Rodrigues, E. (2021). Linear IV regression estimators for structural dynamic discrete choice models. *Journal of Econometrics* 222(1):778–804.
- Koster, H.R., Van Ommeren, J. and Volkhausen, N. (2021). Short-term rentals and the housing market: Quasi-experimental evidence from Airbnb in Los Angeles. *Journal of Urban Economics* 124:103356.
- McFadden, D. (1974). The measurement of urban travel demand. *Journal of public economics* 3(4):303–328.
- Miyauchi, Y., Nakajima, K. and Redding, S.J. (2021). Consumption access and agglomeration: Evidence from smartphone data. Tech. rep., National Bureau of Economic Research.
- Moretti, E. (2013). Real wage inequality. *American Economic Journal: Applied Economics* 5(1):65–103.
- Mullainathan, S. and Spiess, J. (2017). Machine learning: an applied econometric approach. *Journal of Economic Perspectives* 31(2):87–106.
- Roback, J. (1982). Wages, rents, and the quality of life. *Journal of Political Economy* 90(6):1257–1278.
- Rosen, S. (1979). Wage-based indexes of urban quality of life. *Current issues in urban economics* pp. 74–104.
- Rust, J. (1987). Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher. *Econometrica: Journal of the Econometric Society* pp. 999–1033.
- Scott, P.T. (2013). Dynamic discrete choice estimation of agricultural land use. *Toulouse School of Economics Working Paper* 526.
- Shao, J. and Tu, D. (2012). *The jackknife and bootstrap*. Springer Science & Business Media.
- Sheppard, S., Udell, A. et al. (2016). Do Airbnb properties affect house prices? *Williams College Department of Economics Working Papers* 3.
- Su, Y. (2018). The rising value of time and the origin of urban gentrification. *SSRN Unpublished Manuscript* 12.
- Traiberman, S. (2019). Occupations and import competition: Evidence from Denmark. *American Economic Review* 109(12):4260–4301.
- Zervas, G., Proserpio, D. and Byers, J.W. (2017). The rise of the sharing economy: Estimating the impact of Airbnb on the hotel industry. *Journal of Marketing Research* 54(5):687–705.

# Appendix A. Theory

## A.1 Micro-foundation of the utility function

This section derives the amenity demand equations from section 4.1. In our model, households first choose where to live and then, conditional on that location, how much quantity of housing and amenities to consume. In what follows, we solve the household problem backwards. We suppress time subscripts when unnecessary to simplify notation.

**Housing and overall amenities expenditure.** First, conditional on living in location  $j$ , a type  $k$  household chooses how much of its wage  $w^k$  to spend on housing  $H_j$  and on a bundle of locally available consumption amenities  $C_j$ ,

$$\max_{\{H_j, C_j\}} \quad A_j^k H_j^{1-\phi^k} C_j^{\phi^k} \quad \text{s.t.} \quad r_j H_j + P_{Cj} C_j = w^k.$$

where  $r_j$  is the rental price,  $P_{Cj}$  is the price of the consumption bundle, and  $A_j^k$  is the household's valuation of location attributes, which are taken as given. The optimal choice of housing is  $H_j^* = (1 - \phi^k) \frac{w^k}{r_j}$ , which implies the income left over for amenity consumption is  $I^k = \phi^k w^k$ .

**Individual varieties of amenities.** The term  $C_j$  aggregates varieties of consumption amenities,

$$C_j \equiv \prod_s \left[ \left( \sum_{i=1}^{N_{sj}} q_{isj}^{\frac{\sigma_s-1}{\sigma_s}} \right)^{\frac{\sigma_s}{\sigma_s-1}} \right]^{\alpha_s^k},$$

where  $q_{isj}$  is the quantity demanded for variety  $i$  in sector  $s$  and location  $j$ , and  $N_{sj}$  is the number of firms in the sector-location. The aggregator implies Cobb-Douglas preferences over amenity sectors (with weights  $\alpha_s^k$ ) and CES preferences over varieties within an amenity sector (with substitution elasticity  $\sigma_s > 1$ ). Given  $H_j^*$ , we can redefine the consumer's problem as choosing individual varieties subject to its after-rent income,

$$\max_{\{q_{isj}^k\}_{is}} \quad A_j^k H_j^{*1-\phi^k} \prod_s \left[ \left( \sum_{i=1}^{N_{sj}} q_{isj}^k \right)^{\frac{\sigma_s-1}{\sigma_s}} \right]^{\alpha_s^k \phi^k} \quad \text{s.t.} \quad \sum_{is} p_{isj} q_{isj}^k = I^k. \quad (11)$$

The solution to the variety choice problem above is identical whether we include

the  $A_j^k H_j^{*1-\phi^k}$  term or not. In the main text we therefore omit it. First order conditions with respect to  $q_{isj}^k$  give,

$$A_j^k H_j^{*1-\phi^k} \alpha_s^k \phi^k \left[ \left( \sum_{i=1}^{N_{sj}} q_{isj}^k \right)^{\frac{\sigma_s}{\sigma_s-1}} \right]^{\alpha_s^k \phi^k - 1} \left( \sum_{i=1}^{N_{sj}} q_{isj}^k \right)^{\frac{1}{\sigma_s-1}} q_{isj}^{k-\frac{1}{\sigma_s}} \dots \prod_{x' \neq s} \left[ \left( \sum_{i=1}^{N_{x'j}} q_{ix'j}^k \right)^{\frac{\sigma_{x'}}{\sigma_{x'}-1}} \right]^{\alpha_{x'}^k \phi^k} = \lambda^k p_{isj}.$$

By combining the above for two varieties  $i$  and  $i'$  within the same sector  $s$  we obtain,

$$\frac{q_{isj}^k}{q_{i'sj}^k} = \left( \frac{p_{isj}}{p_{i'sj}} \right)^{-\sigma_s}.$$

Furthermore, the total expenditure on sector  $s$  is  $\alpha_s^k I^k$  is given by  $\sum_{i \in s} p_{isj} q_{isj}^k$ . Therefore, type- $k$ 's demand for variety  $i$  in sector-location  $sj$  is,

$$q_{isj}^k = \frac{\alpha_s^k I^k}{p_{isj}} \left( \frac{p_{isj}}{P_{sj}} \right)^{1-\sigma_s}, \quad \text{with } P_{sj} \equiv \left( \sum_{i=1}^{N_s} p_{isj}^{1-\sigma_s} \right)^{\frac{1}{1-\sigma_s}},$$

where  $P_{sj}$  is the sector's price index. In a symmetric equilibrium, where every firm (variety) within a sector-location faces the same marginal costs we have  $p_{isj} = p_{sj} \forall i \in sj$ . Demand for each individual variety is therefore,

$$q_{isj}^k = q_{sj}^k = \frac{\alpha_s^k I^k}{p_{sj} N_{sj}} \quad \forall i \in sj. \tag{12}$$

Plugging equation 12 into the utility function from 11 gives us,

$$A_j^k H_j^{*1-\phi^k} \prod_s \left[ N_{sj}^{\frac{1}{\sigma_s-1}} \frac{\alpha_s^k I^k}{p_{sj}} \right]^{\alpha_s^k \phi^k} = A_j^k w^k r_j^{-(1-\phi^k)} \varphi^k \prod_s \left[ N_{sj}^{\frac{1}{\sigma_s-1}} \frac{\alpha_s^k}{p_{sj}} \right]^{\alpha_s^k \phi^k},$$

where  $\varphi^k \equiv (1 - \phi^k)^{1-\phi^k} (\phi^k)^{\phi^k}$ . We now add time subscripts, allow for location tenure  $\tau_t$  to affect utility with an elasticity of  $\nu^k$ , and let attributes variable  $A$  to have time-invariant and time-varying components. The indirect utility function is

now,

$$\tau_t^{\nu^k} A_j^k A_t^k A_{jt}^k w_t^k r_{jt}^{-(1-\phi^k)} \varphi^k \prod_s \left[ N_{sjt}^{\frac{1}{\sigma_s - 1}} \frac{\alpha_s^k}{p_{sjt}} \right]^{\alpha_s^k \phi^k}.$$

Taking logs and adding a type I EV error  $\varepsilon_{ijt}$ , we obtain

$$\mu_j^k + \mu_t^k + \nu^k \log \tau_t - (1 - \phi^k) \log r_{jt} + \sum_s \frac{\alpha_s^k \phi^k}{\sigma_s - 1} \log N_{sjt} + \log A_{jt}^k + \psi_{jt}^k + \varepsilon_{ijt}, \quad (13)$$

where  $\mu_j^k \equiv \log A_j^k + \log \varphi^k + \phi^k \sum_s \alpha_s^k \log \alpha_s^k$ ,  $\mu_t^k \equiv \log A_t^k + \log w_t^k$ , and  $\psi_{jt}^k \equiv -\phi^k \sum_s \alpha_s^k \log p_{sjt}$ . Furthermore, we decompose  $A_{jt}^k$  as follows,

$$\log A_{jt}^k = \log \tilde{A}_{jt} + \sum_l \gamma_l^k \log B_{ljt} + \sum_s \gamma_s^k \log N_{sjt},$$

where  $\tilde{A}_{jt}$  is an unobserved location attribute,  $B_{ljt}$  is an observed location attribute (other than consumption amenities), and  $\gamma_s^k \log N_{sjt}$  represents spillovers from observed consumption amenities (which may be positive or negative) that go beyond the value of consumption itself. Finally, we divide log-utility by the standard deviation of  $\varepsilon_{ijt}$ ,  $\sigma_\varepsilon^k$ , in order to normalize the variance of the shock to 1,

$$\delta_j^k + \delta_t^k + \delta_\tau^k \log \tau_t + \delta_r^k \log r_{jt} + \sum_s \delta_s^k \log N_{sjt} + \sum_l \delta_l^k \log B_{ljt} + \xi_{jt}^k + \epsilon_{ijt}, \quad (14)$$

where  $\epsilon_{ijt}$  is a standardized type I EV error and  $\delta$  parameters are the parameters in equation 13 divided by  $\sigma_\varepsilon^k$ .

In the main text, we define the flow utility as 14 net of the type I EV shock, with the addition of the moving cost, and with summations written in vector form. Hence the flow utility of making a location decision  $j$  is,

$$\delta_j^k + \delta_t^k + \delta_\tau^k \log \tau_t + \delta_r^k \log r_{jt} + \delta_a^k \log a_{jt} + \delta_b^k \log b_{jt} - MC^k(j, j_{t-1}) + \xi_{jt}^k.$$

where  $\delta_a^k \equiv [\delta_1^k, \dots, \delta_s^k, \dots, \delta_S^k]$ ,  $\delta_b^k \equiv [\delta_1^k, \dots, \delta_l^k, \dots, \delta_L^k]$ ,  $a_{jt} \equiv [N_{1jt}, \dots, N_{sjt}, \dots, N_{Sjt}]'$ , and  $b_{jt} \equiv [B_{1jt}, \dots, B_{ljt}, \dots, B_{Ljt}]'$ .

## Appendix B. Estimation

### B.1 Technical details of the derivation of the ECCP equation

**Constructing Expected Value Function.** Proceeding similarly as in the main text, the value function is defined as follows.

$$V_t(x, \epsilon) = \max_j \left\{ \mathbb{E}_{x'|j,x} [u_t(x', x)] + \epsilon_j + \beta \mathbb{E}_t [V_{t+1}(x', \epsilon')|j, x, \epsilon] \right\}.$$

Under the conditional independence assumption and the assumption that agents are atomistic, we can integrate over future  $\epsilon$ , defining the ex-ante value function as follows:

$$\begin{aligned} \mathbb{E}_t [V_{t+1}(x', \epsilon')|j, x, \epsilon] &= \int V_{t+1}(x', \epsilon') dF_t(x', \omega_{t+1}, \epsilon'|j, x, \epsilon) \\ &= \int \left( \int V_{t+1}(x', \epsilon') dF(\epsilon') \right) dF_t(x', \omega_{t+1}|j, x) \\ &= \int V_{t+1}(x') dF_t(x', \omega_{t+1}|j, x) \equiv EV_t(j, x). \end{aligned}$$

We next define the conditional value function

$$v_t(j, x) = \sum_{x'} \mathbb{P}_t(x'|j, x) (u_t(x', x) + \beta \bar{V}_t(x')) \equiv \bar{u}_t(j, x) + \beta EV_t(j, x).$$

If idiosyncratic shocks are distributed according to i.i.d. Type I EV errors, choice probabilities and value functions can be written as:

$$\mathbb{P}_t(j|x) = \frac{\exp(v_t(j, x))}{\sum_{j'} \exp(v_t(j', x))}, \quad \text{and} \quad V_t(x) = \log \left( \sum_j \exp v_t(j, x) \right) + \gamma, \quad (15)$$

where  $\gamma$  is Euler's constant. Combining the two previous equations,

$$V_t(x) = v_t(j, x) - \ln(\mathbb{P}_t(j|x)) + \gamma. \quad (16)$$

Observe that the previous equation holds for any state  $x$ , and, more importantly, for any action  $j$ . This will be key to exploit renewal actions.

**Toward a demand regression equation.** Our demand regression equation's starting point follows Hotz and Miller (1993), by taking differences on equation 15:

$$\ln \left( \frac{\mathbb{P}_t(j|x_t)}{\mathbb{P}_t(j'|x_t)} \right) = v_t(j, x_t) - v_t(j', x_t). \quad (17)$$

Substituting for the choice specific value function,

$$\ln \left( \frac{\mathbb{P}_t(j|x_t)}{\mathbb{P}_t(j'|x_t)} \right) = \bar{u}_t(j, x_t) - \bar{u}_t(j', x_t) + \beta(EV_t(j, x_t) - EV_t(j', x_t)). \quad (18)$$

The realized expected value  $V_t(x')$  can be decomposed between its expectation at time  $t$  and its expectational error, where uncertainty is on the aggregate state variable  $\omega_{t+1}$ :

$$V_{t+1}(x') = \bar{V}_t(x') + \nu_t(x').$$

Plugging in everything in equation 18 gives us

$$\begin{aligned} \ln \left( \frac{\mathbb{P}_t(j|x_t)}{\mathbb{P}_t(j'|x_t)} \right) &= \sum_x \mathbb{P}(x|j, x_t) u_t(x, x_t) - \sum_{x'} \mathbb{P}(x'|j', x_t) u_t(x', x_t) \\ &\quad + \beta \left[ \sum_x \mathbb{P}(x|j, x_t) \bar{V}_t(x) - \sum_{x'} \mathbb{P}(x'|j', x_t) \bar{V}_t(x') \right] \\ &= \sum_x \mathbb{P}(x|j, x_t) u_t(x, x_t) - \sum_{x'} \mathbb{P}(x'|j', x_t) u_t(x', x_t) \\ &\quad + \beta \left[ \sum_x \mathbb{P}(x|j, x_t) (V_{t+1}(x) - \nu_{t+1}(x)) - \sum_{x'} \mathbb{P}(x'|j', x_t) (V_{t+1}(x') - \nu_{t+1}(x')) \right] \end{aligned}$$

Using again equation 16 to replace the continuation values  $V_{t+1}$  for choice  $\tilde{j}$  gives us

$$\begin{aligned} \ln \left( \frac{\mathbb{P}_t(j, x_t)}{\mathbb{P}_t(j', x_t)} \right) &= \sum_x \mathbb{P}(x|j, x_t) u_t(x, x_t) - \sum_{x'} \mathbb{P}(x'|j', x_t) u_t(x', x_t) \\ &\quad - \beta \left[ \sum_x \mathbb{P}(x|j, x_t) (v_{t+1}(\tilde{j}, x) - \ln \mathbb{P}_{t+1}(\tilde{j}|x) - \nu_{t+1}(x)) \right. \\ &\quad \left. - \sum_{x'} \mathbb{P}(x'|j', x_t) (v_{t+1}(\tilde{j}, x') - \ln \mathbb{P}_{t+1}(\tilde{j}|x') - \nu_{t+1}(x')) \right] \\ &= \sum_x \mathbb{P}(x|j, x_t) u_t(x, x_t) - \sum_{x'} \mathbb{P}(x'|j', x_t) u_t(x', x_t) \\ &\quad - \beta \left[ \sum_x \mathbb{P}(x|j, x_t) (v_{t+1}(\tilde{j}, x) - \ln \mathbb{P}_{t+1}(\tilde{j}|x)) \right. \\ &\quad \left. - \sum_{x'} \mathbb{P}(x'|j', x_t) (v_{t+1}(\tilde{j}, x') - \ln \mathbb{P}_{t+1}(\tilde{j}|x')) \right] + \tilde{v}_{j, j', x_t}, \end{aligned}$$

where  $\tilde{v}_{j,j',x_t}$  is a sum of expectational errors

$$\tilde{v}_{j,j',x_t} = \beta \left( \sum_x \mathbb{P}(x|j, x_t) v_{t+1}(x) - \sum_{x'} \mathbb{P}(x'|j', x_t) v_{t+1}(x') \right).$$

Observe that if  $\tilde{j}$  is a renewal action then:

$$v_{t+1}(\tilde{j}, x) = \bar{u}_{t+1}(\tilde{j}, x) + EV_t(\tilde{j}, 1) = u_{\tilde{j}, x, t+1} + \delta_\tau \cdot 1 + MC(\tilde{j}, j) + EV_t(\tilde{j}, 1)$$

for all  $x = (j, \tau)$ , regardless of  $\tau$ , where we have decomposed the per-period utility function,  $\bar{u}_{t+1}(\tilde{j}, x)$ , into a location specific component,  $u_{\tilde{j}, x, t+1}$ , a location-tenure component  $\delta_\tau$ , and a moving cost component  $MC(\tilde{j}, j)$ . Substituting and re-arranging,

$$\begin{aligned} & \ln \left( \frac{\mathbb{P}_t(j|x_t)}{\mathbb{P}_t(j'|x_t)} \right) + \beta \left[ \sum_x \mathbb{P}(x|j, x_t) \ln \mathbb{P}_{t+1}(\tilde{j}|x) - \sum_{x'} \mathbb{P}(x'|j', x_t) \ln \mathbb{P}_{t+1}(\tilde{j}|x') \right] \\ &= u_{j, x_t} - u_{j', x_t} + \delta_\tau \left( \sum_x \mathbb{P}(x|j, x_t) \tau(x) - \sum_{x'} \mathbb{P}(x'|j', x_t) \tau(x') \right) \\ &+ MC(j, j_{t-1}) - MC(j', j_{t-1}) + \beta \left( MC(\tilde{j}, j) - MC(\tilde{j}, j') \right) + \tilde{v}_{j,j',x_t}, \end{aligned}$$

which is the generalized version of equation 9 under stochastic transitions of individual state variables.