reading group summer 24
Location Sorting and Endogenous Amenities: Evidence from Amsterdam (2024)

Milena Almagro and Tomas Domınguez-Iino

August 2024

Outline

Introduction

Data

Model - brace yourselves

Empirical Strategy and Results

Counterfactuals

Introduction

- build a model of residential choice with hetereogenous households - amenities improve in response to location sorting
- but nature of amenities has not been explored
- estimate the model using several data
- run counterfactual exercises

Context

- massive expansion in tourism in Amsterdam
- increased supply of private rentals, increased supply of STRs
- New regulation in Amsterdam severely restricting STR supply (hotels and Airbnbs)

Related Literature

- spatial equilirium models
- effects of STR entry on the housing market and hotel revenue
- discrete-choice tools from the empirical io literature applied to urban residential markets

Outline

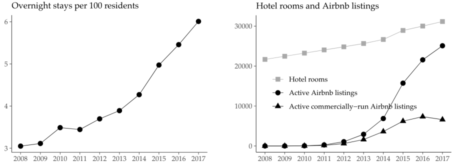
Data

- individual-level microdata: Centraal Bureau voor de Statistiek, Netherlands they have complete residential histories, demographics
- housing data: panel on physical characteristics, occupancies, and values. impute rents
- neighbourhood data: Amsterdam City Data consists of demographics, tourist flows, and consumption amenities: restaurants, bars, food stores, non-food stores, nurseries, and "touristic amenities"
- STR data: Inside Airbnb. monthly data on listings, geo-coordinates, prices, reviews. need to identify preoperties that are permanently rented to tourists.

Stylised Facts

Fact 1: Tourists and STR listings have grown dramatically and sprawled across Amsterdam

Figure 1: Overnight stays per resident, hotel rooms, and STR listings (2008-2017).



Notes: On the left, "overnight stays per 100 residents" is constructed as annual overnight stays (in hotels and STR) divided by population, and multiplied by 100—a value of 5 means that on an average night there are 5 tourists per 100 residents. On the right, active and commercial Airbnb listings are constructed from Inside Airbnb data using the procedure described in Appendix A.2.7. Hotel, stay and population data are from ACD Tourism and ACD BBGA.

Figure 1:

Stylised Facts

Fact 2: Rents have increased more in neighborhoods with more STR entry using OLS (likely biased) and IV (shift-share)

Table 1: Airbnb intensity and housing market outcomes

	Ln (rent/m2)					
	OLS	IV	OLS	IV	OLS	IV
Ln (commercial Airbnb listings)	**0.065*** **	**0.091*** **	**0.051*** **	**0.114*** **	**0.109*** **	**0.205* **
	(0.008)	(0.021)	(0.006)	(0.021)	(0.018)	(0.093)
Control variables			X	X	X	X
District-year FE					X	X
First stage F-stat		586.89		384.21		69.66
Observations	770	770	763	763	763	763

	Ln (house sale price)					
	OLS	IV	OLS	IV	OLS	IV
Ln (commercial Airbnb listings)	**0.109*** **	**0.290*** **	**0.034*** **	**0.149*** **	**0.037* **	**0.326** **
	(0.016)	(0.030)	(0.006)	(0.016)	(0.022)	(0.102)
Control variables			X	X	X	X
District-year FE					X	X
First stage F-stat		572.02		370.87		65.9
Observations	738	738	737	737	737	737

Notes: Observations are at the wijk (neighborhood) level. A "district" is a larger spatial unit than a neighborhood. Rent prices are neighborhood-average long-term rental prices constructed from CBS rent surveys. House sale prices are neighborhood average transaction values, constructed from CBS data covering the universe of housing transactions. Commercial Airbnb listings are constructed from the Inside Airbnb data (see Appendix A.2.7 for construction details). Neighborhood-level control variables are: housing stock, average income, high-skill population share, all from ACD BBGA. Standard errors are clustered at the wijk level in parenthesis.

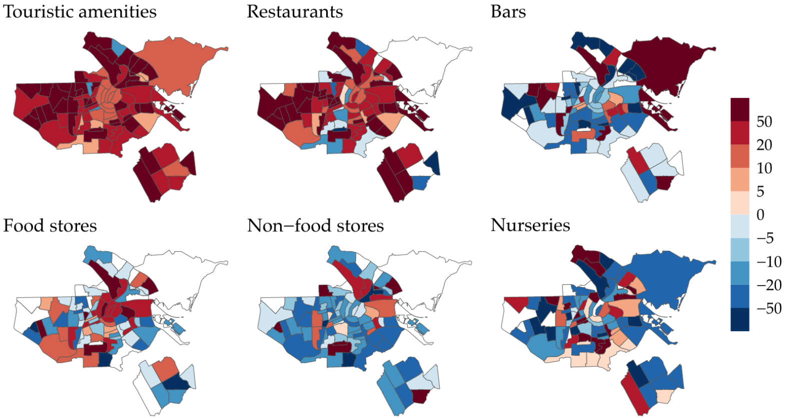
Figure 2:

a 1% increase in a neighborhood's commercial STR listings is associated with a rent increase between .06-.11% and a house sale price increase between .04-.11%

Stylised Facts

Fact 3: Amenities have tilted towards tourists and away from locals

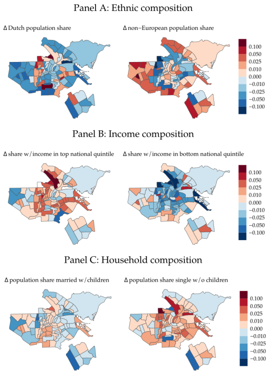
Figure 3: Evolution of consumption amenities (2011-2017 pp changes).



Notes: Maps show percentage point changes between 2011-2017 for each amenity sector. Data is from ACD BBGA.

Fact 4: The composition of residents has changed heterogeneously across neighborhoods

Figure 5: Changes in socioeconomic composition of neighborhoods (2011-2017).



Notes: Maps shows changes in neighborhood population share of each group. Data is from ACD BBGA.

Figure 4:

Outline

Notation

- J locations $+1$ outside option
- M^k_{jt}: number of type k households in location j
- consumption amenities s in S sectors
- N_{sjt}: the number of varieties in sector s and location j at time t
- amenities $a_{jt} = [N_{1jt}, N_{2jt}, ..., N_{Sjt}]'$

Amenities Demand

- Demand: Cobb-Douglas over bundles H and C
- also Cobb-Douglas preferences over amenity sectors, CES over firms/varieties within sector

$$q^k_{isjt} = \frac{\alpha^k_s \phi^k w^k_t}{p_{isjt}} \left(\frac{p_{isjt}}{P_{sjt}}\right)^{1-\sigma_s}, \text{ with } P_{sjt} \equiv \left(\sum_{i=1}^{N_{sjt}} p_{isjt}^{1-\sigma_s}\right)^{\frac{1}{1-\sigma_s}} \quad q_{isjt} = \sum_k q^k_{isjt} M^k_{jt}.$$

Figure 5:

Amenities Supply

- Firms within s, j in monopolistic competition - same marginal cost, free entry
- fixed cost F_{sjt} - assumed to be increasing in N_{jt} (congestion costs)
- implies that firms choose same price $p_{sjt} = p_{ijst}$ and quantity $q_{sjt} = q_{isjt}$

$$(p_{sjt} - c_{sjt})q_{sjt} = F_{sjt}(N_{jt}), \quad \text{where } N_{jt} = \sum_s N_{sjt}.$$

Figure 6:

Amenities Market Clearing

- usual market clearing delivers equilibrium number of varieties/firm
- and a mapping from population composition to amenities

Housing Supply

- total housing stock inelastic in the short run
- Landlords face a binary choice: long-term rentals or short-term rentals:

$$\max\left\{\alpha r_{jt} + \epsilon_{LT}, \quad \alpha p_{jt} - \kappa_{jt} + \epsilon_{ST}\right\},$$

Figure 7:

Housing Supply

- total housing stock inelastic in the short run
- Landlords face a binary choice: long-term rentals or short-term rentals:

$$\mathcal{H}_{jt}^{LT,S}(r_{jt}, p_{jt}) = \frac{\exp(\alpha r_{jt})}{\exp(\alpha r_{jt}) + \exp(\alpha p_{jt} - \kappa_{jt})} \mathcal{H}_{jt},$$
$$\mathcal{H}_{jt}^{ST,S}(r_{jt}, p_{jt}) = \mathcal{H}_{jt} - \mathcal{H}_{jt}^{LT,S}(r_{jt}, p_{jt}).$$

Figure 8:

Housing Demand

- moving costs and location capital

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

Figure 9:

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

Figure 10:

- some macro...

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

Figure 11:

Housing Demand

- choice probabilities

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

Figure 12:

- transition matrix

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

Figure 13:

- (population) demand for STRs

$$\mathcal{H}_{jt}^{LT,D}(r_t, a_t) = \sum_{k=1}^{K} M_{jt}^{k}(r_t, a_t) f_{jt}^{k}.$$

Figure 14:

$$\mathcal{H}_{jt}^{LT,D}(r_t, a_t) = \sum_{k=1}^{K} M_{jt}^{k}(r_t, a_t) f_{jt}^{k}.$$

Figure 15:

Housing Demand: tourists

- toursists' flow payoff from staying in STR

$$u_{jt}^{ST} = \delta_j^{ST} + \delta_t^{ST} + \delta_p^{ST} \log p_{jt} + \delta_a^{ST} \log a_{jt} + \xi_{jt}^{ST},$$

Figure 16:

Housing Demand: tourists

- toursists' demand for STR

Figure 17:

Housing Demand: tourists

- hotels are an outside option
- but no price or bookings data
- so use data on hotel capacity across locations

$$M_{jt}^{H}(p_t, a_t) \; = \; s_{jt}^{beds} \times M_t^H(p_t, a_t)$$

Figure 18:

Housing Equilibrium

- a stationary equilibrium
- a bunch of market clearing equations
- long-term rental markets clear, short-term rental markets clear, amenities markets clear

Outline

Amenity supply

- estimating equation:

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

Figure 19:

- X_{jt}^k total expenditure of the type k population in location j on consumption amenities
- β_s^k describes how this expenditure is allocated to each amenity sector s
- IV: something that shifts amenities demand

Amenity Supply

- IV: $Z_{jt}^k = w_t^k S_{jt}^{\gamma(k)}$, where S is housing stock by tenancy status
- GMM moment equation to idenfity βs: $E[\omega_{sjt} Z_{sjt}^k] = 0$

Amenity Supply

- results:

Table 3: Estimates of amenity supply parameters.

	Touristic Amenities	Restaurants	Bars	Food Stores	Non-Food Stores	Nurseries
Older Families	186.3	7.374	0.0	4.469	11.359	980.803***
	[0.0,431.929]	[0.0,30.401]	[0.0,0.0]	[0.0,29.098]	[0.0,50.577]	[368.357,1684.881]
Singles	403.022	90.723	0.0	102.631	11.347	0.0
	[0.0,1816.258]	[0.0,317.227]	[0.0,0.0]	[0.0,327.09]	[0.0,176.969]	[0.0,0.0]
Younger Families	0.0	1.077	11.365	52.846**	194.655***	637.116
	[0.0,0.0]	[0.0,16.47]	[0.0,38.983]	[0.0,123.07]	[83.304,331.513]	[0.0,1326.254]
Students	984.639*	402.153***	22.562	123.078	2.355	221.204
	[0.0,2014.696]	[198.125,673.928]	[0.0,101.561]	[0.0,319.632]	[0.0,1.365]	[0.0,1759.481]
Immigrant Families	0.122	5.687	25.705**	90.549	127.724**	540.228
	[0.0,0.0]	[0.0,49.542]	[1.126,63.178]	[0.0,210.37]	[3.251,331.895]	[0.0,1740.204]
Dutch Low Income	110.617	9.908	0.0	9.077	0.0	0.0
	[0.0,371.65]	[0.0,56.277]	[0.0,0.0]	[0.0,78.283]	[0.0,0.0]	[0.0,0.0]
Tourists	749.072***	397.274***	211.571***	136.337***	724.223***	0.0
	[522.649,974.412]	[316.498,477.965]	[156.406,269.374]	[80.554,189.6]	[579.264,892.963]	[0.0,0.0]

Notes: This table reports bootstrap results for coefficients β_s^k from Equation 24 for using a three-way panel of 22 districts in Amsterdam for 2008-2018 over 500 draws. Parameters β_s^k and fixed effects λ_j and λ_t are estimated via GMM, where we restrict parameters to be weakly positive as implied by the microfoundation of the amenity model in Appendix A.3.1. The estimation procedure is outlined in section 5.2 following a Bayesian-bootstrap with random Dirichlet weights. Total expenditure X_{jt}^k is measured in thousands of Euros. Top rows indicate average estimates of the bootstrap samples. Results inside square brackets indicate 95% confidence intervals. We omit estimates of the location and time fixed effects. $^*p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Housing Demand: by locals

- "Euler Equations in Conditional Choice Probabilities" (ECCP) estimator: which I didn't understand and didn't have time to study in detail

Table 4: Preference parameter demand estimation results

	Dependent variable: Relative likelihood of renewal paths		
	Older Families	Singles	Younger Families
High Location Capital	0.187***	0.210***	0.264***
	(0.017)	(0.013)	(0.014)
Intra-City Moving Cost	-5.916***	-5.337***	-5.384***
	(0.015)	(0.011)	(0.012)
Bilateral Moving Cost	-0.067***	-0.059***	-0.041***
	(0.000)	(0.000)	(0.000)
In/Out of City Moving Cost	-4.407***	-4.012***	-4.043***
	(0.012)	(0.009)	(0.010)
Log Rent	-10.886***	-2.310**	-1.964*
	(1.205)	(0.999)	(1.027)
Log Touristic Amenities	-1.319***	-0.496***	0.317*
	(0.215)	(0.182)	(0.177)
Log Restaurants	0.288	0.735**	-0.280
	(0.346)	(0.305)	(0.286)
Log Bars	-0.757***	-0.528***	-0.104
	(0.099)	(0.085)	(0.086)
Log Food Stores	-1.695***	-1.216***	-0.540*
	(0.327)	(0.281)	(0.282)
Log Nonfood Stores	0.427	1.533***	1.383***
	(0.356)	(0.311)	(0.302)
Log Nurseries	1.631***	0.044	0.246*
	(0.173)	(0.143)	(0.147)
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 three groups via two-step optimal GMM. The dependent variable is differences in path likelihoods, after normalizing with respect to the outside option. After this normalization, 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 omit exogenous controls—the log of social housing units and the log of the average apartment in square meters—for the ease of exposition. Two-step efficient GMM standard errors in parenthesis. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Figure 21:

- interpret signs

Housing Demand: by tourists

- no IV here: they use reveiews of Airbnbs to create a score variable for every location

Table 5: Tourist demand across locations.

	Dependent Variable: $\log \mathbb{P}^{ST}_{jt} - \log \mathbb{P}^{H}_{t}$			
	Baseline		Controlling for reviews	
Log Price Per Guest	-2.725***	(0.820)	-2.660***	(0.759)
Log Touristic Amenities	1.009***	(0.376)	0.838**	(0.394)
Log Restaurants	0.048	(0.259)	0.017	(0.243)
Log Bars	0.051	(0.155)	0.056	(0.164)
Log Food Stores	-0.001	(0.300)	0.037	(0.323)
Log Nonfood Stores	-0.228	(0.417)	-0.185	(0.407)
Log Nurseries	-0.234*	(0.137)	-0.231*	(0.136)
Log Review Scores			4.768	(3.699)
N	371.000		370.000	
R^2	0.529		0.537	

Notes: Table reports estimates of tourists' preference for neighborhood (wijk-level) characteristics for a static model of location choice, using neighborhood-level data for 2015-2018. Construction of Airbnb supply and prices is described in Appendix A.2. Wijk-level clustered standard errors in parenthesis. $^*p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Figure 22:

Housing supply

- estimating equation:

$$\log \mathcal{H}_{jt}^{LT,S} - \log \mathcal{H}_{jt}^{ST,S} = \alpha \left(r_{jt} - p_{jt} \right) + \kappa_j + \kappa_t + \nu_{jt},$$

Figure 23:

- OLS is biased - use shift-share IV that shifts demand
- shift: worldwide change in STR demand
- share: neighborhood-level exposure to the shift from the historic spatial distribution of touristic attractions
- relevance and exclusion are satisifed. (note: not mention of monotonicty, CALL IVAN)

Housing supply

- IV results: increase in the gap between STR prices and LTR prices of one standard deviation (29%) would raise the market share of the ST relative to the LT segment by 13.6%

Table 6: Long-term (LT) relative to short-term (ST) housing supply elasticities

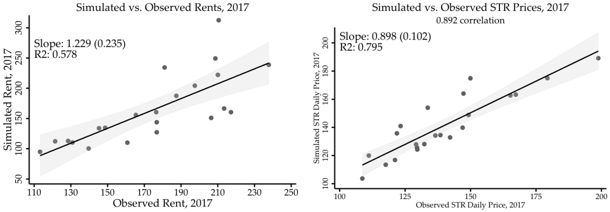
	Dependent variable: ln (LT share) - ln (ST share)			
	OLS	IV	IV	IV
LT price-ST price	**0.242***	**0.287****	**0.309****	**0.385**
	(0.099)	(0.086)	(0.091)	(0.639)
Year FE			X	X
Wijk FE				X
First stage F-stat		65.68	61.62	3.24
Observations	275	275	275	275

Notes: Table reports estimates of landlords' marginal utility of income for a discrete choice model between the short- and long-term rental markets. Data are a panel with 92 locations 2015-2017. Prices are instrumented using a shift-share instrument (Barron et al., 2021) that proxies for demand shocks. Wijk-level clustered standard errors in parenthesis. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Figure 24:

- pretty well!

Figure 7: Model fit: Rents and STR prices



Notes: The figure presents scatter plots, linear fit, and 95% confidence intervals of simulated rents and STR prices, against observed rents and prices for 22 districts. Rents are in $Euros/m^2$ per year. STR prices are average daily prices.

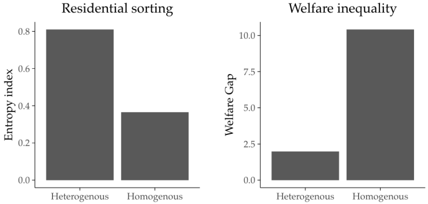
Figure 25:

Outline

Counterfactual 1: Preference heterogeneity vs homogeneity

- set preference parameters for consumption amenities to the average value across all household types, weighted by the size of groups
- segregation is higher when households have heterogeneous preferences for amenities

Figure 9: Role of preference heterogeneity for spatial sorting and inequality across households.



Notes: The left panel reports the entropy index, a measure of spatial segregation of household types: higher values indicate more segregation (see Appendix A.5.6 for a formal definition). The right panel reports the ratio of the highest consumer surplus household (in Euros) to that of the lowest household: higher values indicate more inequality.

Figure 26:

- ... but inequality is lower when preferences are heterogeneous

Counterfactual 1: Preference heterogeneity vs homogeneity

- ... but inequality is lower when preferences are heterogeneous

Table 7: Neighborhood differentiation as spatial dispersion of amenities.

	Gini index for each preference specification		
Amenity	Homogenous (HO)	Heterogenous (HE)	HE-HO
Touristic amenities	0.34	0.37	0.03
Restaurants	0.43	0.56	0.13
Bars	0.59	0.66	0.07
Food stores	0.32	0.57	0.25
Non-food stores	0.53	0.67	0.14
Nurseries	0.51	0.43	-0.08

Notes: Columns "Homogeneous" and "Heterogeneous" report the Gini index for each amenity sector: how concentrated the number of establishments in each sector is across locations. Higher values indicate most of the sector's establishments are clustered in a few locations. Column HE-HO reports the difference between the "Heterogeneous" and "Homogeneous" columns. Positive values in the HE-HO column indicate the spatial distribution of the amenity becomes more clustered across space when preferences are heterogeneous.

Figure 27:

- because high income groups do not compete with low income groups for the same locations, allowing low income groups to obtain their preferred amenities without having the high income groups bid up their rents

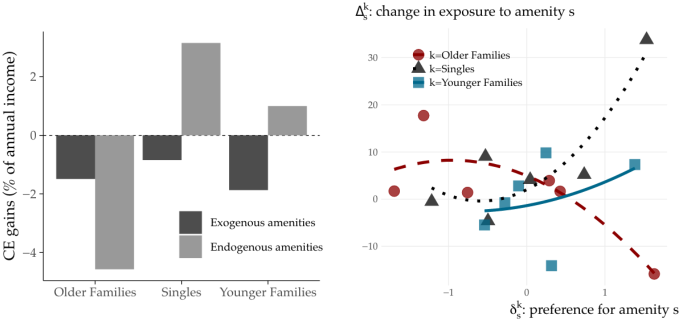
Counterfactual 2: Decomposing welfare effects of the STR industry

- STR entry reduces rent & change amenity composition
- disentangle: pre-entry baseline, allow STR keeping amenities fixed, then allow amenities to adjust

Counterfactual 2: Decomposing welfare effects of the STR industry

- touristic amenities grow the most in areas populated by old people

Figure 10: Decomposition of welfare effects from STR entry.

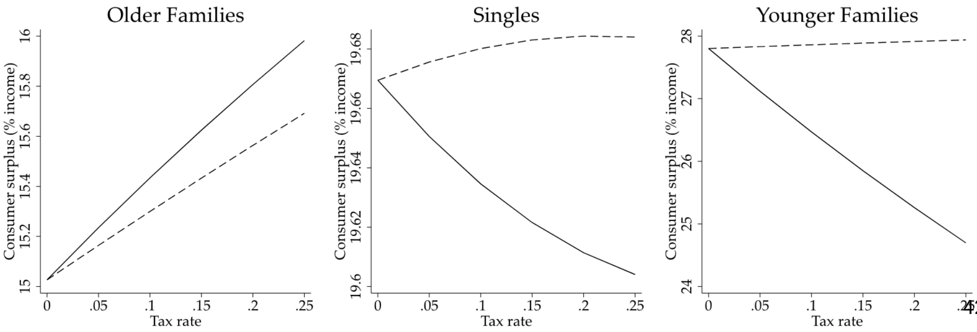


Notes: On the left panel, the consumption equivalent (CE) gains on the vertical axes are computed as how much extra income a household must be given in the baseline equilibrium to obtain the same utility as in the counterfactual equilibrium. Therefore, positive values indicate welfare gains due to STR entry. Details in Appendix for A.5.5. On the right panel, the horizontal axis shows preference parameters for amenity sectors. The vertical axis shows the change in exposure to amenity s after STR entry for a type k household, defined as $\Delta_s^k \equiv \sum_j \Delta N_{sj} \times \omega_j^k$, where ΔN_{sj} is the change in sector s amenities in location j after STR entry, weighted by $\omega_j^k = M_j^k / M^k$, location j's share of the city-wide population of type k before STR entry. Hence, ω_j^k is type k's exposure to location j.

Counterfactual 3: Taxing STRs and/or Touristic amenities

- tax short-term rentals - directly reduces rent
- vs tax touristic amenities
- taxing STR has monotonically increasing effects (in tax rates) on welfare

Figure 12: Welfare effects: short-term rental tax vs. touristic amenity tax.



Conclusion

- studied the role of preference heterogeneity over a set of endogenous location amenities in shaping within-city sorting and welfare inequality
- there exists hetereogenity preference over amenities
- leads to increased sorting but welfare effects are ambiguous

See ya