Online Appendix for "Location Sorting and Endogenous Amenities: Evidence from Amsterdam"

Milena Almagro* Tomás Domínguez-Iino†

November 20, 2022
Click here for latest version.

S.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.

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

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

$$\begin{aligned}
\mathbb{E}_t\left[V_{t+1}(x',\epsilon')|j,x,\epsilon\right] &= \int V_{t+1}(x',\epsilon') dF_t(x',\omega_{t+1},\epsilon'|j,x,\epsilon) && (1)\\
&= \int \Big(\int V_{t+1}(x',\epsilon') dF(\epsilon') \Big) dF_t(x',\omega_{t+1}|j,x) && (2)\\
&= \int V_{t+1}(x') dF_t(x',\omega_{t+1}|j,x) \equiv EV_t(j,x). && (3)
\end{aligned}$$

We next define the conditional value function

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

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

$$\mathbb{P}_t(j|x) = \frac{\exp(v_t(j,x))}{\sum_{j'} \exp(v_t(j',x))}, \quad \text{and} \quad V_t(x) = \log\left(\sum_j \exp v_t(j,x) \right) + \gamma, \tag{4}$$

where γ is Euler's constant. Combining the two previous equations,

$$V_t(x) = v_t(j,x) - \ln(\mathbb{P}_t(j|x)) + \gamma. \tag{5}$$

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.

*University of Chicago Booth School of Business. E-mail: milena.almagro@chicagobooth.edu
†Federal Reserve Board. E-mail: tdomingueziino@gmail.com. The views expressed are those of the authors and not necessarily those of the Federal Reserve Board or the Federal Reserve System.

S.1.0.1 Toward a demand regression equation.

Our demand regression equation's starting point follows Hotz and Miller (1993), by taking differences on equation 4:

$$\ln\Big(\frac{\mathbb{P}_t(j|x_t)}{\mathbb{P}_t(j'|x_t)}\Big) = v_t(j,x_t) - v_t(j',x_t). \tag{6}$$

Substituting for the choice specific value function,

$$\ln\Big(\frac{\mathbb{P}_t(j|x_t)}{\mathbb{P}_t(j'|x_t)}\Big) = \bar{u}_t(j,x_t) - \bar{u}_t(j',x_t) + \beta\big(EV_t(j,x_t) - EV_t(j',x_t)\big). \tag{7}$$

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 ω_{t+1}:

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

Plugging in everything in equation 7 gives us

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

Using again equation 5 to replace the continuation values V_{t+1} for choice $\tilde{j}$ gives us

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

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

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

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 τ, 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 δ_τ, and a moving cost component $MC(\tilde{j}, j)$. Substituting and re-arranging,

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

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

S.1.0.2 First-stage estimation of Conditional Choice Probabilities

We follow a similar procedure as in Traiberman (2019) and Humlum (2021), estimating CCPs in a first stage projecting empirical frequencies on a linear model of state characteristics. We depart from their approach and use a multinomial logit on individual decisions for three reasons. First, this approach allows us to directly estimate the CCPs probability using individual data. Second, our data reveal that the likelihood of not moving is approximately 85%, while the probability of moving to any other location remains close to zero. This bimodal nature of our data suggests that an exponential relationship should be better suited to fit individual decisions compared to a linear model. Third, we find that many predicted probabilities of the linear model lie below zero or above one, a feature that requires an ad-hoc extra censoring step.

Observe that our microdata allows us to observe individual moving decisions as well as individual states. That is, for individual i we observe her state at time t $x_{it} = (j_{t-1}, \tau_{t-1})$, where j_{t-1} is the previous location, τ_{t-1} and type $k(i)$, as well as the moving decision variables for all j: $j_{it} = \mathbb{1}\{d(i)_t = j\}$. Therefore, defining some base outcome 0, we estimate the following multinomial logit model across all groups k:

$$\mathbb{P}(j_{it} = j) = \frac{\exp(\beta_j^{k(i)} x_{it})}{1 + \sum_{j'=1}^{J} \exp(\beta_{j'}^{k(i)} x_{it})}.$$

Concretely, we define the model as follows:[1]

$$\beta_j^k s_t = \lambda_j^k + \lambda_t^k + \alpha_1^k \tau_{t-1} + \alpha_2^k \tau_{t-1}^2.$$

In what follows, we show that Monte Carlo simulations reveal that our multinomial predicted CCPs lead to a lower finite sample bias in second-stage estimates compared to the the standard approach where CCPs are estimated using a empirical frequencies.

S.1.0.3 Computation

Let θ be the set of coefficients to be estimated excluding fixed effects and let λ be the set of location fixed effects, with $dim(\theta) = 13$ and $dim(\lambda) = 21$. Also note that we have $J = 22$ locations (gebieds), $T = 10$ time

[1] For computational simplicity, we estimate the following model group by group. In principle, it would be preferred to estimate all groups jointly. Unfortunately this was impractical due to the large nature of our individual data and the limited computational power in the remote environment provided by the Dutch Microcensus facilities.

periods and the maximum location tenure is fixed to $\bar{\tau} = 2$. In total, this yields $(J+1) \times \bar{\tau} \times J \times T \times (J-1) = L = 212,520$ equations.

In matrix notation, if we define by $X_{t,j,\tilde{j},x}$ a vector the right-hand-side variables in Equation 9 excluding $\tilde{\xi}_{t,j,x_t}$ and the fixed effects, and by $Z_{t,j,\tilde{j},x}$ a vector of their corresponding instruments, we stack the variables as follows. We define

$$y \equiv [Y_{t,j,\tilde{j},x}]_{\{t,j',\tilde{j},x\}} \qquad X \equiv [X_{t,j,\tilde{j},x}]_{\{t,j',\tilde{j},x\}} \qquad Z \equiv [Z_{t,j,\tilde{j},x}]_{\{t,j',\tilde{j},x\}} \tag{8}$$

Further, define $g(\theta, \lambda) = Z'(y - X\theta - H\lambda)$, where by H we denote the corresponding matrix of location fixed-effect dummies. For a suitable weighting matrix W, GMM solves the following problem

$$\min_{(\theta,\lambda)\in\mathbb{R}^{34}} g(\theta,\lambda)' W g(\theta,\lambda) \tag{9}$$

Optimization over such a large space is computationally demanding. We use the procedure of Somaini and Wolak (2016) and X and y on the orthogonal complement of the image of the projection matrix of H. However, to define the annihilator matrix of H, we need to store a matrix with more than 45 billion entries, which is a demanding task for most computers. We instead exploit the block structure of the algebraic objects as follows. If we define by D the $((J-1)\cdot T) \times dim(\lambda)$ matrix of dummies for a single location and period, the projection matrix of H is given by

$$H(H'H)^{-1}H' = \frac{1}{(J+1)\cdot\bar{\tau}\cdot(J-1)} \begin{bmatrix} D(D'D)^{-1}D' & \dots & D(D'D)^{-1}D' \\ \vdots & & \vdots \\ D(D'D)^{-1}D' & \dots & D(D'D)^{-1}D' \end{bmatrix} \tag{10}$$

Hence, we obtain

$$\left(I - H(H'H)^{-1}H'\right)Y = Y - \begin{bmatrix} \frac{1}{(J+1)\cdot\bar{\tau}\cdot(J-1)} \sum_{i=1}^{(J+1)\cdot\bar{\tau}\cdot(J-1)} D(D'D)^{-1}D'Y_i \\ \vdots \\ \frac{1}{(J+1)\cdot\bar{\tau}\cdot(J-1)} \sum_{i=1}^{(J+1)\cdot\bar{\tau}\cdot(J-1)} D(D'D)^{-1}D'Y_i \end{bmatrix} \tag{11}$$

Thus, we only need to store the projection matrix of D.

S.1.1 Monte Carlo simulations

We validate our demand estimation strategy using Monte Carlo studies. We perform the exercise for a model with agents' beliefs given by (i) rational expectation and (ii) perfect foresight. Since the main estimation procedure is performed independently for each type, we conduct the exercise for a single type.

The period flow utility function used in the simulations is as follows (with rational expectations, we set $\eta_t = 0, \forall t$)

$$u_t((d,\tau),x_t) = \alpha \log(r_{dt}) + \sum_s \beta_s \log N_{dst} + \xi_{dt} + \eta_t + \lambda_d + MC(d, j_{t-1}) + \delta_\tau \tau.$$

We first construct the unobservable term ξ_{dt}, which contains the sum of two i.i.d. terms, u and v, representing unobservable characteristics of locations that vary across time and an expectational error component, respectively. Next, we construct the exogenous parts of household budget terms and amenities for each location-period pair i.i.d. from the Log-Normal distribution. The terms are a weighted sum of the exogenous terms and v, depending on whether we consider a model with endogeneity. We also draw the distance between inner-city locations and location and time fixed effects, constructing the flow utility function for all states. The full set of parameters is in Table 1.

We compute the EV function for each time period as follows. Starting in the last period T, we assume that the economy is in steady-state. We define EV_T as the unique solution to the following fixed-point

Table 1: Parameters used in simulations

Variables		Parameters	
Name	Distribution (i.i.d.) / Value	Name	Value
u	$N(0, 0.05)$	α	- 0.05
v	$N(0, 0.05)$	β_1, β_2	0.1
ξ	$u + v$	γ_0, γ_1	-0.0025
b_{exo}	$LogN(0.5, 0.1)$	γ_2	-0.5
r	$0.75 \cdot b_{\text{exo}} + 0.25 \cdot v$	δ	0.1
a_{exo}	$LogN(1.5, 0.5)$	$N_{\text{households}}$	$\in [5 \cdot 10^5, 10^6]$
a	$0.75 \cdot a_{\text{exo}} + 0.25 \cdot v$	J	24
$\text{dist}(j, j'), j, j' \neq 0; \rho_d$	$LogN(1, 0.5)$	S	2
λ_j	$N(0, 0.1)$	$\bar{\tau}$	2
λ_t (Perfect foresight)	$N(0, 0.1)$	T	10
λ_t (Rational expectations)	0	tol in EV iteration	10^{-10}

equation

$$EV_T(j_T, \tau_T) = \log\left(\sum_d \exp\left(\sum_{x'} \mathbb{P}_T(x'|d, x_T)\left[u_T(x', x_T) + \beta EV_T(d, x_T)\right]\right)\right) \tag{12}$$

For periods $t = 1, \ldots, T-1$, we compute EV_t using backward substitution as follows

$$EV_t(j_t, \tau_t) = \log\left(\sum_d \exp\left(\sum_{\tau'} \mathbb{P}_{t+1}(x'|d, x_t)\left[u_{t+1}(d, x_t) + \beta EV_{t+1}(d, x')\right]\right)\right) \tag{13}$$

Assuming a uniform initial distribution of individuals across states, we simulate each individual forward for 10 time periods using the the conditional choice probabilities computed using the flow utilities and EV functions presented above. We repeat each procedure 10 times and average the resulting output. The selected population sizes are 50 thousand and 1 million; the former roughly corresponding to the cardinality of each type of households, and the latter simulates large samples and provides insight into convergence properties. We test two models correcting for zero observed conditional choice probabilities: (i) we predict conditional choice probabilities using a multi-nomial logit model and (ii) we use observed frequencies and replace zero probabilities with a small $\epsilon = 10^{-5}$.

The results are presented below for the models with perfect foresight and rational expectations in Tables 2 and 3, respectively. In both cases, models with transition probabilities estimated using the multinomial logit model yield a strictly dominant finite sample performance. The gap is most pronounced in small samples where the likelihood of observing zero flows between states in the data is higher, where the frequency-based estimator uses small but arbitrary values imputed by the researcher, which can be far from the true transition probabilities. The multi-nomial logit approximates the true probabilities well, reducing finite-sample bias in the final estimation stage. As we increase the sample size, the number of observed zero flows diminishes, and we observe convergence of both estimators to the performance of the first-best estimator using the true transition probabilities.

Table 2: Monte Carlo simulations with fixed effects and an indicator for high location capital

			Mean of the absolute value of bias						
ξ	Pop (in 10^3)	Prob.	α	β_1	β_2	γ_0	γ_1	γ_2	δ
zero	50	T	7.6E-16	1.6E-16	1.8E-16	1.7E-15	4.8E-19	1.1E-15	1.7E-15
		L	1.7E-02	3.5E-03	1.7E-03	4.5E-02	2.0E-04	4.8E-02	1.0E-01
		F	3.2E-01	2.1E-01	1.8E-01	1.1E+00	2.5E-03	8.0E-01	8.8E-01
	1000	T	9.1E-16	1.3E-16	1.9E-16	1.6E-15	1.1E-18	1.2E-15	1.9E-15
		L	3.4E-03	1.1E-03	5.9E-04	3.8E-02	2.2E-05	3.4E-02	1.0E-01
		F	7.7E-03	2.3E-03	1.6E-03	3.3E-02	6.7E-05	1.6E-02	3.9E-02
exogenous	50	T	3.0E-02	8.5E-03	6.1E-03	1.7E-15	1.0E-18	9.9E-16	1.8E-15
		L	3.3E-02	8.2E-03	7.5E-03	4.9E-02	2.0E-04	3.4E-02	1.0E-01
		F	3.0E-01	1.5E-01	1.3E-01	1.2E+00	2.4E-03	8.9E-01	7.6E-01
	1000	T	3.8E-02	5.5E-03	5.6E-03	1.8E-15	1.4E-18	1.6E-15	2.0E-15
		L	3.7E-02	5.7E-03	5.4E-03	4.5E-02	3.6E-05	3.6E-02	1.0E-01
		F	4.0E-02	5.7E-03	6.6E-03	2.5E-02	1.0E-04	7.3E-03	3.0E-02
endogenous	50	T	4.2E-02	9.8E-03	8.4E-03	1.8E-15	1.3E-18	1.2E-15	1.8E-15
		L	3.9E-02	1.0E-02	1.1E-02	4.7E-02	1.6E-04	3.5E-02	1.0E-01
		F	2.7E-01	1.7E-01	1.2E-01	1.1E+00	2.8E-03	8.6E-01	6.6E-01
	1000	T	5.0E-02	1.0E-02	6.7E-03	1.7E-15	1.6E-18	1.3E-15	1.8E-15
		L	5.4E-02	1.0E-02	6.6E-03	3.7E-02	2.6E-05	3.4E-02	1.0E-01
		F	6.0E-02	1.0E-02	6.8E-03	2.3E-02	8.5E-05	9.3E-03	2.6E-02

Notes: The table presents averaged absolute distance between the estimated parameter and the true parameter, over 10 random draws of datasets. T represents estimation using the true transition probabilities; L represents estimation using predicted probabilities by a multi-nomial logit model; and F represents estimation using transition probabilities computed based on empirical shares.

Table 3: Monte Carlo simulations with location fixed effects only and an indicator for high location capital

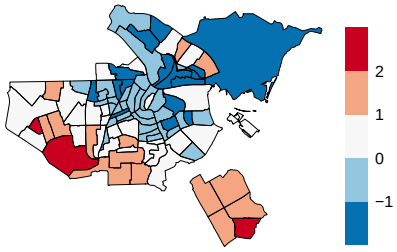
			Mean of the absolute value of bias						
ξ	Pop (in 10^3)	Prob.	α	β_1	β_2	γ_0	γ_1	γ_2	δ
zero	50	T	1.2E-15	1.9E-16	3.2E-16	1.3E-15	8.2E-19	1.3E-15	1.8E-15
		L	2.3E-02	2.3E-03	4.5E-03	5.7E-02	1.4E-04	3.9E-02	1.0E-01
		F	6.1E-01	1.7E-01	1.7E-01	1.3E+00	2.9E-03	7.8E-01	1.3E+00
	1000	T	8.1E-16	2.5E-16	2.6E-16	1.2E-15	6.5E-19	9.4E-16	1.6E-15
		L	5.6E-03	1.2E-03	8.0E-04	3.3E-02	2.8E-05	3.6E-02	1.0E-01
		F	1.4E-02	3.9E-03	2.6E-03	1.3E-02	5.0E-05	1.4E-02	3.8E-02
exogenous	50	T	2.7E-02	3.8E-03	3.5E-03	7.2E-03	5.0E-06	1.4E-15	2.0E-15
		L	2.7E-02	3.5E-03	5.7E-03	5.9E-02	2.2E-04	2.8E-02	1.0E-01
		F	4.6E-01	2.2E-01	2.7E-01	7.7E-01	2.9E-03	5.2E-01	1.5E+00
	1000	T	4.2E-02	1.5E-02	1.4E-02	7.9E-02	3.6E-04	1.3E-15	2.0E-15
		L	4.4E-02	1.5E-02	1.4E-02	1.0E-01	3.7E-04	3.4E-02	1.0E-01
		F	4.8E-02	1.6E-02	1.4E-02	9.7E-02	4.7E-04	1.3E-02	4.3E-02
endogenous	50	T	2.5E-02	9.9E-03	1.1E-02	9.1E-03	1.3E-05	1.2E-15	2.0E-15
		L	2.8E-02	9.5E-03	1.0E-02	5.6E-02	1.6E-04	4.4E-02	1.0E-01
		F	4.3E-01	2.5E-01	2.4E-01	7.9E-01	2.8E-03	6.5E-01	1.0E+00
	1000	T	3.0E-02	6.8E-03	3.2E-03	1.1E-02	1.0E-05	1.1E-15	1.9E-15
		L	2.8E-02	6.9E-03	3.0E-03	2.5E-02	5.5E-05	3.7E-02	1.0E-01
		F	4.2E-02	8.3E-03	5.5E-03	2.8E-02	1.0E-04	9.2E-03	3.3E-02

Notes: The table presents averaged absolute distance between the estimated parameter and the true parameter, over 10 random draws of datasets. T represents estimation using the true transition probabilities; L represents estimation using predicted probabilities by a multi-nomial logit model; and F represents estimation using transition probabilities computed based on empirical shares.

S.1.2 Housing supply estimation

Figure 1 shows the spatial distribution of the estimates for κ_j, the landlord's differential operating costs between short- and long-term markets. The results suggest landlords in central locations face lower costs of renting short-term relative to long-term. This is consistent with the notion that they match with guests easier and face lower vacancy risk.

Figure 1: Spatial distribution of location fixed effects κ_j.



Notes: estimates of κ_j are from the elasticity IV specification with wijk and year fixed effects. The values have been standardized, so positive values are above average, negative are below, and a value of 1 indicates a 1 standard deviation above the mean κ_j.

S.2 Simulations and Counterfactuals

S.2.1 Segregation measure

We follow White (1986) and use the entropy index for the city as our measure of segregation. We construct the index as follows. First, we define the entropy index for a single location. Let d_j^k be the fraction of households of type k in location j, i.e. if the demand for location j from household type k is D_j^k, we have $d_{jk} \equiv D_j^k / \sum_k D_j^k$. For location j, the entropy index v_j is defined as

$$v_j \equiv -\sum_{k=1}^{K} d_j^k \log(d_j^k)$$

We proceed to defining the entropy index for the whole city. First, we define the following objects: $D_j \equiv \sum_{k=1}^{K} D_j^k$, $D^k \equiv \sum_{j=1}^{J} D_j^k$, and $D \equiv \sum_{j=1}^{J}\sum_{k=1}^{K} D_j^k$. Pooling all locations together and a weighted average of the location-wise indices, weighted by the total demand for a location, respectively, we obtain

$$\widehat{v} \equiv -\sum_{k=1}^{K} \frac{D^k}{D} \log\left(\frac{D^k}{D}\right) \qquad \text{and} \qquad \overline{v} \equiv \sum_{j=1}^{J} v_j \frac{D_j}{D}.$$

Finally, we normalize the above indices to obtain an index in the interval $[0, 1]$. Our measure of segregation then is

$$v \equiv \frac{\widehat{v} - \overline{v}}{\widehat{v}},$$

Observe that higher v means higher segregation, v equals 0 if and only if the share of each type in each location is given by its population share in the whole population, and v equals 1 if and only if each location is occupied by exactly one type.

The entropy index is 0.04 in the observed data and 0.7 in the simulated equilibrium. While the difference might seem concerning, we argue that a high level of segregation is natural for a stationary equilibrium, as we take the observed dynamics to the infinite limit. Indeed, such patterns in steady-state equilibrium is consistent with the trends described in section 3. These more extreme segregation patterns are a product of the combination of large moving costs and endogenous amenities. A simple comparative static exercise presented in section S.2.2 reveals that segregation in the steady-state equilibrium is increasing in the moving cost. Finally, focusing on a stationary equilibrium allows us to decrease the computational burden of computing the full equilibrium path while extracting the core economic lessons.[2]

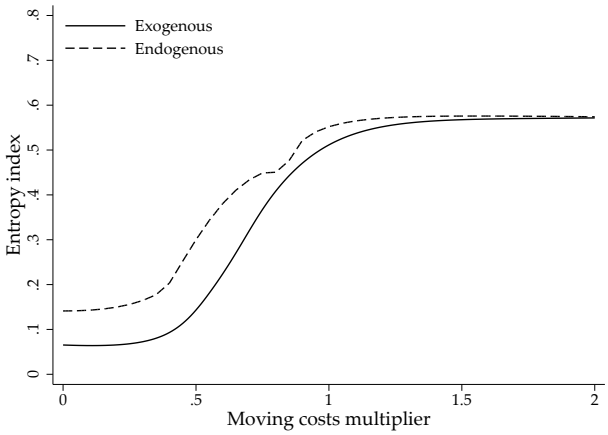
[2]One caveat is that by focusing only on steady-state, we are not able to say anything about welfare along the transition paths.

S.2.2 Relationship between segregation in steady-state and moving costs

In this section, we shed light on the relationship between the level of moving costs and the level of segregation in equilibrium. Due to the complexity of the non-linear dynamic system, we cannot construct a closed-form solution of this mapping and perform instead a simulation exercise.

We use the same specification outlined in Section **??**. We vary moving costs by scaling each parameter of the moving costs and location capital by the same scalar, and compute the entropy index in equilibrium for an equally-spaced grid of scalars on the interval $[0, 2]$.[3] We perform this exercise both with exogenous and endogenous amenities. The resulting mappings are presented in Figure 2.

Figure 2: The relationship between moving costs and entropy index for the city.



Notes: The figure presents a plot from a simulation exercise of recomputing equilibrium for a set of moving cost multipliers taken from an equally-spaced grid on the interval $[0, 2]$. For the model with exogenous amenities, we used 201 points, with an increment of 0.01. For the model with endogenous amenities, due to the computational burden of repeatedly solving for equilibrium, we used 41 points, with an increment of 0.05.

Several patterns stand out. First, the function mapping moving costs to the entropy index follows a sigmoidal curve. The curve is essentially flat for moving costs below a third of the estimated level, increases at an increasing rate between one and two thirds of the estimated level of moving costs, which is followed by a concave region on the rest of the domain. As moving costs increase beyond the estimated level, we observe convergence to an upper bound below 0.6. This pattern is robust for both models.

Intuitively, higher moving costs imply both greater persistence and path dependence in the dynamics, decreasing the likelihood of moving to a different location for each state. Agglomeration forces attract peers to move to the same set of locations, increasing homogeneity of the type composition of neighborhoods. The interaction of these two forces leads to a higher degree of segregation in the steady state equilibrium. The shape of the relationship is closely related to the shape of the distribution of the error term in flow utility and the definition of our segregation index. An increase in moving costs for a small base level has a higher impact on dampening the likelihood of moving, generating the convex relationship. The marginal impact of an increase in moving costs is low when moving costs are already high.

Second, for each level of moving costs, the entropy index for the full equilibrium with endogenous amenities is strictly higher than for the model with exogenous amenities. This follows because amenities act as an agglomeration force. Third, the entropy index is strictly positive for each level of moving costs. Intuitively, heterogeneity among households leads to heterogeneous sorting patterns even in the absence of moving costs.

[3] We also scale the location capital as it serves an analogous role to moving costs. As we increase location capital, the opportunity cost of moving to a different location also increases.

References

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.

Humlum, A. (2021). Robot adoption and labor market dynamics. *Working Paper* .

Somaini, P. and Wolak, F.A. (2016). An algorithm to estimate the two-way fixed effects model. *Journal of Econometric Methods* 5(1):143–152.

Traiberman, S. (2019). Occupations and import competition: Evidence from Denmark. *American Economic Review* 109(12):4260–4301.

White, M.J. (1986). Segregation and diversity measures in population distribution. *Population index* pp. 198–221.