Codebook for:
The economics of density: Evidence from the Berlin Wall *

Gabriel M. Ahlfeldt† Stephen J. Redding‡ Daniel M. Sturm§ Nikolaus Wolf¶

March 18, 2024

Abstract

This codebook is part of the toolkit for Ahlfeldt, Redding, Sturm, Wolf (2015): The economics of density: Evidence from the Berlin Wall, *Econometrica*, 83(6), pp. 2127-89. The toolkit does not cover all stages of the analysis presented in the article. Instead, it covers a subset of codes that are crucial for the quantification and simulation of the model using 2006 data. The codebook summarizes the primitives and endogenous objects of the models and introduces selected numerical algorithms in pseudo-code. The focus is on algorithms that are essential for the quantification and simulation of the respective quantitative models.

*We thank Alexander Hansen and Andrea Herrera for comments and suggestions.

†Humboldt University. Email: g.ahlfeldt@hu-berlin.de

‡Princeton University. Email: reddings@princeton.edu

§London School of Economics. Email: d.m.sturm@lse.ac.uk

¶Humboldt University. Email: nikolaus.wolf@wiwi.hu-berlin.de

A Ahlfeldt, Redding, Sturm, and, Wolf (2015)

This section covers *Ahlfeldt, Redding, Sturm, and Wolf (2015): The economics of density: Evidence from the Berlin Wall. Econometrica, 83*, https://doi.org/10.3982/ECTA10876. The model consists of exogenous parameters exogenous location characteristics, and endogenous objects tabulated in Table 1.

A.1 Equilibrium

Given the model's exogenous parameters $\{\alpha, \beta, \mu, \varepsilon, \kappa\}$, the exogenous reservation utility level $\bar{U}$, exogenous location characteristics $\{T_i, E_i, A_i, B_i, \varphi_i, K_i, \xi_i, \tau_{ij}\}$, the general equilibrium of the model without endogenous agglomeration forces is referenced by $\{\pi_{Mi}, \pi_{Ri}, Q_i, q_i, w_i, \theta_i, H\}$. These seven components of the equilibrium are determined by the following seven equations:

1. Population mobility (9): $\mathbb{E}[u] = \gamma \left[\sum_{r=1}^{S} \sum_{s=1}^{S} T_{r,s} E_s \left(d_{rs} Q_r^{-\beta} \right)^{-\varepsilon} \left(B_{r,w_s} \right)^{\varepsilon} \right]^{1/\varepsilon} = \bar{U}$
2. Residential choice probability (5): $\pi_{R_i} = \sum_{j=1}^{S} \pi_{ij} = \frac{\sum_{j=1}^{S} \Phi_{ij}}{\Phi}$
3. Workplace choice probability (5): $\pi_{M_j} = \sum_{i=1}^{S} \pi_{ij} = \frac{\sum_{i=1}^{S} \Phi_{ij}}{\Phi}$
4. Commercial land market clearing (18): $\left(\frac{(1-\alpha) A_j}{q_j} \right)^{1/\alpha} H_{M_j} = \theta_j L_j$
5. Residential land market clearing (19): $\left(\frac{(1-\alpha) A_j}{q_j} \right)^{1/\alpha} H_{R_j} = \theta_j L_j$
6. Profit maximization and zero profits (12): $q_j = (1-\alpha) \left(\frac{\alpha}{w_j} \right)^{\alpha/(1-\alpha)} A_j^{1/(1-\alpha)}$.
7. No-Arbitrage between alternative uses of land (13): $\theta_i = \begin{cases} 1 & \text{if } q_i > \xi_i Q_i, \\ [0,1] & \text{if } q_i = \xi_i Q_i, \\ 0 & \text{if } q_i < \xi_i Q_i. \end{cases}$

To quantify the model, we use the mapping created via the above equations to invert adjusted productivities, amenities, and density of development $\{\tilde{A}_i, \tilde{B}_i, \tilde{\varphi}_i\}$ which incorporate $\{T_i, E_i, A_i, B_i, \varphi_i, K_i, \xi_i\}$. When incorporating endogenous agglomeration forces, we decompose $\{\tilde{A}_i, \tilde{B}_i\}$ into adjusted fundamental productivity and adjusted fundamental amenity $\{\tilde{a}_i, \tilde{b}_i\}$ and endogenous productivity and amenity, which depend on nearby workplace and residence employment density. We observe τ_{ij} (through GIS modelling), but travel costs could, in principle, be inverted from bilateral commuting data. Once the model is fully quantified, we solve it numerically. To this end, we use the above equations to pin down the values of the target variables $\{\tilde{w}_i, q_i, Q_i, \theta_i\}$. For given values of model primitives and these target variables, we can solve for all other endogenous objects.

Notice that in the case with endogenous agglomeration forces, $\{A_i, B_i\}$ are endogenous objects and $\{a_i, a_i\}$ are exogenous characteristics. We have four more exogenous parameters $\{\lambda, \delta, \eta, \rho\}$. For the two additional endogenous outcomes, we have two more equations.

8. Productivity spillovers (20): $A_j = a_j \times \Upsilon_j^{\lambda}, \quad \Upsilon_j = \left[\sum_{s=1}^{S} e^{-\delta \tau_{is}} \left(\frac{H_{Ms}}{K_s}\right)\right]$

9. Residential spillovers (21): $B_i = b_i \times \Omega_i^{\eta}, \quad \Omega_i = \left[\sum_{s=1}^{S} e^{-\rho \tau_{is}} \left(\frac{H_{Rs}}{K_s}\right)\right]$

Therefore, we remain exactly identified.

Table 1: Codebook

Object	Description
	Structural parameters
α	Input share of floor space in production
β	Expenditure share on non-housing consumption
μ	Share of non-land inputs in floor space production
ε	Preference heterogeneity (labour supply elasticity)
κ	Iceberg commuting cost parameter
$\nu = \kappa\epsilon$	commuting decay
λ	Density elasticity or productivity
δ	Productivity decay
η	Density elasticity of residential amenity
ρ	Residential decay
χ	Scale parameter in relationship between floor space prices and land prices
$\gamma = \Gamma\frac{\varepsilon-1}{\varepsilon}$	Scale parameter in expected utility
Δ	Parameter vector including $\{\nu, \epsilon, \lambda, \delta, \nu, \rho\}$
	Exogenous characteristics
a_j	Production fundamentals
A_j	Productivity
Υ_j	Endogenous productivity
b_i	Residential fundamentals
B_i	Amenities
Ω_i	Endogenous amenity
ξ_i	Tax-equivalent of land use regulations
E_j	Commuting destination-specific mean preference (from idiosyncratic shock)
$\tilde{A}_i = A_i E_i^{\alpha/\epsilon}$	Adjusted productivity
$\tilde{a}_i = a_i E_i^{\alpha/\epsilon}$	Adjusted production fundamentals
$\tilde{B}_i = B_i T_i^{1/\epsilon} / I_{R_i}^{1-\beta}$	Adjusted amenities
$\tilde{b}_i = b_i T_i^{1/\epsilon} / I_{R_i}^{1-\beta}$	Adjusted residential fundamentals
K_i	Land endowment
$\varphi_i = M_i^{\mu} = \frac{L_i}{K_i^{1-\mu}}$	Density of development
$\tilde{\varphi}_i = \tilde{\varphi}_i(\varphi_i, E_i^{1/\varepsilon}, \xi_i)$	Adjusted density of development
$L_i = \varphi_i K_i^{1-\mu}$	Total floor space in block, subscripts M,R index use
τ_{ij}	Bilateral travel times
$\bar{U}$	Reservation utility level
	Endogenous variables and scalars

Continued on next page

Table 1 Continued from previous page

Object	Description
H	Total city employment
π_{ij}	Commuting probabilities
H_{Mj}	Workplace employment (observed)
H_{Ri}	Residence employment (observed)
w_j	Nominal wages (solved within the model)
$\mathbb{E}[w_s \mid i] = \sum_{s=1}^{S} \pi_{is\mid j} w_s$	Expected worker income
$\tilde{w}_j$	Adjusted wages $\tilde{w}_j = E_j^{\frac{1}{\epsilon}} w_j$
ω_j	Transformed wages $\omega_j = \tilde{w}_j^{\epsilon} = E_j w_j^{\epsilon}$
$W_{it} = \sum_{s=1}^{S} \omega_{st}/e^{\nu' \tau_{ist}}$	(Residential) Commuting market access
l_i	Floor space per resident
M_i	Capital
θ_i	Share of commercial floor space at total floor space
q_i	Commercial floor space rent
Q_i	Residential floor space rent
θ_i	Share of floor space use commercially
	1 if $q_i > \xi_i Qi$, [0,1] if $q_i = \xi_i Qi$, 0 if $q_i < \xi_i Qi$
$\mathbb{Q}_i = \max\{q_i, Q_i\}$	Observed floor space price, q_i if $q_i > \xi_i Q_i$, q_i if $q_i = \xi_i Q_i$, Q_i if $q_i < \xi_i Q_i$
$\mathbb{R}_i$	Land price
y_i	output

A.2 Algorithms

We introduce a subset of algorithms used by ARSW. We abstract from the computationally demanding structural estimation and focus on algorithms used for the quantification and simulation of the model. We refer to objects whose values we solve numerically as target objects. Within the solver, we refer to guessed values when values of target objects are input into the construction of other objects. We refer to predicted values when the values of target objects are computed as functions of other objects (which typically depend on guesses). Typically, we solve for the values of target objects using iterative procedures in which we update our guesses to weighted combinations of guessed and predicted values.

A.2.1 One-step estimation of ω

While estimation is not the focus of this toolkit, we cover the one-step estimation of the taste heterogeneity parameter ε. The procedure described below is not particularly computationally demanding. ε is less consensual than most of the other parameters and it is likely context-dependent. Therefore, it can be useful to estimate ε even though it might be justifiable to borrow values of many other parameters from the literature.

To estimate ε, we combine three algorithms. Algorithm 1 solves for transformed wages, ω_i. Algorithm 2 uses transformed wages to compute the difference between the Bezirke-level variance in adjusted wages, $\tilde{w}_i$, in model and data, our objective function. Algorithm 3 nests Algorithm 2 and searches for the value of ε that minimizes the value of the objective function.

Algorithm 1: Solving for transformed wages (ω_j): `comegaopt0.m` (used in estimation)

Data: Given values for structural parameters $\{\alpha, \beta, \kappa\varepsilon\}$, bilateral travel times τ_{ij}, observed workplace employment H_{Mj} and residence employment H_{Ri}
Guesses of transformed wages $\tilde{\omega}_j^0$

while *Predicted workplace employment* $\hat{H}_{Mj} \neq H_{Mj}$ **do**
- Use guessed values of transformed wages $\tilde{\omega}_j^0$ in Eq. (S.44) to predict $\hat{H}_{Mj}$
- Generate new guesses $\tilde{\omega}_j^1 = \frac{H_{Mj}}{\hat{H}_{Mj}}\hat{\omega}_j^0$
- Update guesses to weighted combination of new and old guesses
- Normalize guesses by geometric mean

Result: Transformed wages $\tilde{\omega}_j$

Algorithm 2: Compute value of objective function $f(\varepsilon)$: `cdensityoptren.m`

Data: Current guess of ε

Model solutions of transformed wages ω from Algorithm 1 (`comegaopt.m`)

Variance of log wages across Bezirke $\hat{\sigma}^2(\log(w_B))$ in data

Compute adjusted wages $\tilde{w}_j = \omega_i^{\frac{1}{\varepsilon}}$ using ω_i and current guess of ε

Compute Bezirke adjusted wages $\tilde{w}_B$

Compute variance of log wages across Bezirke, $\sigma^2(\log(\tilde{w}_B))$ in model

Compute residual $\mathbf{ftD} = \hat{\sigma}^2(\log(w_B)) - \sigma^2(\log(\tilde{w}_B))$

Compute current value of objective function $f(\varepsilon) = \mathbf{ftD}'\mathbf{ftD}$ (the residual sum of squares)

Result: Value of objective function $f(\varepsilon)$

A.2.2 Quantification with exogenous fundamentals (sequential procedure)

This sub-section summarizes a sequential procedure during which Eqs. S.47 and S.48 are used to recover adjusted productivity, $\tilde{A}_i$ and adjusted amenity, $\tilde{B}_i$. Algorithm 4 recovers adjusted wages, $\tilde{w}_i$, from the commuting market clearing condition. Given adjusted wages, $\tilde{w}_i$ and observed floor space prices, $\mathbb{Q}_i$, it recovers adjusted productivity, $\tilde{A}_i$. Using adjusted wages, $\tilde{w}_i$, Algorithm 5 then recovers adjusted amenity $\tilde{B}_i$. Using adjusted amenity $\tilde{B}_i$ along with adjusted wages, $\tilde{w}_i$, Algorithm 7 computes expected total income, $\mathbb{E}(w_i)H_{Ri}$. Income, adjusted productivity and adjusted amenity are then used by Algorithm 8 to recover total floor space, L_i and the share of commercial floor space, θ_i.

Going through these algorithms and the sequential quantification procedure is particularly useful for didactic purposes since there is a close link to the analytical solutions (up to scale) for $\{\tilde{A}_i, \tilde{B}_i\}$ in Eqs. (27) and (28). Moreover, each of the algorithms individually is arguably more intuitive than the simultaneous quantification procedure introduced in Section A.3. Finally, the algorithms introduced here can be useful to understand how to solve for selected variables (e.g. adjusted wages) in other contexts.

Notice, however, that the algorithmic implementation of this sequential procedure in the original replication directory does not ensure that adjusted amenity and productivity are in the right scale to be input into the equilibrium solver in Algorithm 11. Algorithm 9 introduced in the next section, chooses the scales of $\{\tilde{A}_i, \tilde{B}_i\}$ such that aggregate employment matches observed total employment. This is why Algorithm 9 is used for the quantification preceding counterfactuals in the paper. Adjusted productivity, $\tilde{A}_i$ and amenity, $\tilde{B}_i$, recovered by Algorithms 1 and 5 are, of course, still entirely suitable if the objective is to analyze relative differences in amenity and productivity across locations. However, they should not be used as a starting point for counterfactuals without further adjustment.

To this end, this toolkit contains Algorithm 6, which is not part of the original replication directory. When executed after Algorithms 1 and 5, it rescales adjusted amenity, $\tilde{A}_i$, and productivity, $\tilde{B}_i$, to match observed total employment. As a result, the values of $\{\tilde{A}_i, \tilde{B}_i\}$ are identical to the values recovered by Algorithm 9. This means that Algorithm 11 correctly recovers the initial equilibrium using these values of $\{\tilde{A}_i, \tilde{B}_i\}$. So, after executing Algorithm 6, $\{\tilde{A}_i, \tilde{B}_i\}$ from the sequential procedure can be used to solve for counterfactual equilibria.

Algorithm 3: Estimating ε using the MATLAB `patternsearch` algorithm

Data: Define initial guess of parameter value (ε)
Objective function $f(\varepsilon)$ defined in Algorithm 2 (`cdensityoptren.m`)
Define initial_step_size
Define step_size_threshold

while *current_step_size* $>$ *step_size_threshold* **do**
  pattern_points ← generate_pattern(current_point, current_step_size);
  (generate a pattern of test points [parameter values] around the current point)
  for *point* **to** *pattern_points* **do**
    point_value ← objective_function(point) using Algorithm 2;
    (evaluate the objective function at each point [parameter value] in the pattern)
    if *point_value* $<$ *objective_function(current_point)* **then**
      current_point ← point;
      (If a point with a lower value is found, update the current point [paramter value])
  if *current_point was not updated* **then**
    current_step_size ← reduce_step_size(current_step_size);
    (If no better points were found, reduce the step size)

Result: Estimate of parameter value for ε that minimizes objective function $f(\varepsilon)$

Algorithm 4: Solving for adjusted wages ($\tilde{w}_j$) and productivities ($\tilde{A}_j$): `comegaoptC.m` (used in calibration)

Data: Solver for transformed wages (ω_j) `comegaopt0.m` Algorithm 1
Given values for structural parameters $\{\alpha, \beta, \kappa\varepsilon, \varepsilon\}$, bilateral travel times τ_{ij}, observed workplace employment H_{Mj} and residence employment H_{Ri}, and floor space prices $\mathbb{Q}_j$
Guesses of adjusted wages $\tilde{w}_i^0$

1. Compute guesses of transformed wages $\omega_i^0 = (\tilde{w}_i^0)^\varepsilon$
2. Use solver `comegaopt0.m` Algorithm 1 to solve for transformed wages ω_j
3. Compute equilibrium adjusted wages $\tilde{w}_j = \omega_j^{1/\epsilon}$
4. Use adjusted wages $\tilde{w}_j$ and floor space prices $\mathbb{Q}_j$ in Eq. S48 to solve for adjusted productivities $\tilde{A}_j$

Result: Adjusted wages and productivities $\{\tilde{w}_j, \tilde{A}_j\}$

Note: In the actual code directory, `comegaoptC.m` replicats much of the code from `comegaopt0.m` (instead of calling `comegaopt0.m`).

Algorithm 5: Solving for adjusted amenities ($\tilde{B}_i$): `camen.m`

Data: Adjusted wages, $\tilde{w}_j$ solved by Algorithm 4 (`comegaoptC.m`)
Observed residence employment H_{Ri} and floor space prices $\mathbb{Q}_i$,
bilateral travel times τ_{ij}, estimated preference heterogeneity ε and
commuting decay $\varepsilon\kappa$

1. Use adjusted wages $\tilde{w}_i$, bilateral travel times τ_{ij}, and commuting decay $\kappa\varepsilon$ to compute commuting market access W_i using the commuting market access equation in Section S.3.1.2.
2. Use solved commuting market access W_i, observed residence employment H_{Ri}, floor space prices $\mathbb{Q}_j$ and estimated ε in Eq. S47 to solve for adjusted amenities $\tilde{B}_j$

Result: Adjusted amenities, residential commuting market access $\{\tilde{B}_i, W_i\}$

Algorithm 6: Rescaling adjusted amenities $\tilde{B}_i$ and productivities $\tilde{A}_i$ to rationalize observed population H: `calcal_adj_TD.m`

Data: Given values for structural parameters $\{\beta, \kappa\varepsilon, \varepsilon\}$, bilateral travel times τ_{ij}, adjusted productivities $\tilde{A}_i$ solved by Algorithm 4, adjusted amenities $\tilde{B}_i$ solved by Algorithm 5, observed floor space prices $\mathbb{Q}_j$

1. Normalize adjusted productivities $\tilde{A}_i$ by the geometric mean
2. Use rescaled adjusted productivities $\tilde{A}_i$ and observed floor space prices $\mathbb{Q}_j$ in Eq. (12) to compute rescaled adjusted wages $\tilde{w}_j$
3. Use bilateral travel times τ_{ij}, adjusted amenities $\tilde{B}_i$, observed floor space prices $\mathbb{Q}_j$, and rescaled adjusted wages $\tilde{w}_j$ to compute Φ (the denominator in Eq. (12) and the total employment in the model)
4. Rescale adjusted amenities $\tilde{B}_i$ by multiplying them by the adjustment factor $\left(\frac{H}{\Phi}\right)^{\frac{1}{\varepsilon}}$ (see supplement p. 18 to see that H scales in any spatially invariant component of B_i at an elasticity of ε)

Result: Rescaled adjusted productivity and amenity $\{\tilde{A}_i, \tilde{B}_i\}$

Algorithm 7: Solving for total expected income ($\mathbb{E}(w_i)H_{Ri}$) and adjusted productivity $\tilde{A}_i$: `expincome.m`

Data: Given values for structural parameters $\{\beta, \kappa\varepsilon, \varepsilon\}$, bilateral travel times τ_{ij}, adjusted wages $\tilde{w}_j$ solved by Algorithm 4, adjusted amenities $\tilde{B}_i$ solved by Algorithm 5 (`camen.m`), and observed residence employment H_{R_i} floor space prices $\mathbb{Q}_i$

1. Compute bilateral commuting probabilities π_{ij} using Eq. (5)
2. Compute conditional commuting probabilities $\pi_{ij|i}$ using Eq. (6)
3. Compute expected worker income $\mathbb{E}(w_i)$ using Eq. (S.20)
4. Use expected worker income $\mathbb{E}(w_i)$ and residential employment to compute total expected income $\mathbb{E}(w_i)H_{Ri}$

Result: Total expected income $\mathbb{E}(w_i)H_{Ri}$

Algorithm 8: Solving for density of development (φ_i), total floor space L_i, and commercial floor space share θ_i: `cdenisty.m`

Data: Given values for structural parameters $\{\alpha, \beta\}$, land endowment K_i, observed floor space prices $\mathbb{Q}_j$, adjusted productivity $\tilde{A}_j$ solved by Algorithm 4 (`comegaoptC.m`), adjusted amenities $\tilde{B}_i$ solved by Algorithm 5 (`camen.m`)

1. Compute commercial floor space demand $\theta_i L_i$ using Eq. (S.29)
2. Compute residential floor space demand $(1-\theta_i)L_i$ using Eq. (S.30)
3. Compute total floor space demand $L_i = \theta_i L_i + (1-\theta_i)L_i$
4. Use expected worker income $\mathbb{E}(w_i)$ and residence employment to compute total expected income $\mathbb{E}(w_i)H_{Ri}$
5. Compute density of development φ_i using L_i, land area K_i and Eq. (S.31)
6. Compute commercial floor space share $\theta_i = \frac{\theta L_i}{L_i}$

Result: Density of development, total floor space, and commercial floor space share $\{\varphi_i, L_i, \theta_i\}$

A.3 Counterfactuals with exogenous fundamentals

A simple two-step procedure can be used to conduct flexible counterfactuals. In the first step, we use Algorithms 9 and 10 as well as observed values of the endogenous variables residence employment, H_{Ri}, workplace employment, H_{Mi}, and floor space prices $\mathbb{Q}_i$, to invert adjusted productivity $\tilde{A}_i$ and adjusted amenity $\tilde{B}_i$. In the second step, we use the inverted values of $\{\tilde{A}_i, \tilde{B}_i\}$ and Algorithm 11 to solve for all endogenous outcomes. We obtain solutions to a counterfactual instead of the observed equilibrium by changing any of the model's primitives.

Algorithm 9 recovers $\{\tilde{A}, \tilde{B}\}$ from one iterative procedure. The algorithm starts from guessed values of $\{\tilde{A}, \tilde{B}\}$ and exploits the unique mapping from primitives to endogenous outcomes to compute residence and workplace employment $\{\hat{H}_{Mi}, \hat{H}_{Ri}\}$. It keeps adjusting the guesses of $\{\tilde{A}, \tilde{B}\}$ until $\{\hat{H}_{Mi}, \hat{H}_{Ri}\}$ correspond to the values observed in data. For the final $\{\tilde{A}, \tilde{B}\}$ values, it further generates endogenous objects not observed in data, such as adjusted wages

$\tilde{w}_j$, commuting probabilities π_{ij}, and expected income $\mathbb{E}(\tilde{w}_s)$. Using the values of $\{\tilde{A}_i, \mathbb{E}(\tilde{w}_i)\}$ solved by Algorithm 9, Algorithm 10 recovers the exogenous density of development and total floor space $\{\varphi_i, L_i\}$ as well as the endogenous commercial floor space shares θ_i, which completes the quantification of the model.

Algorithm 11 exploits the unique mapping from primitives to endogenous outcomes to solve for the endogenous variables. It uses the values of $\{\tilde{A}, \tilde{B}\}$ recovered using Algorithm 9. To this end, it uses an iterative procedure to solve for the values of the target variables adjusted wages, floor space prices, and commercial floor space shares $\{\tilde{w}_i, q_i, Q_i, \theta_i\}$. For the solved values of these target variables, the algorithm generates a range of further endogenous outcomes (all endogenous variables can be recovered from if the values of the target variables and primitives are known).

Notice that Algorithm 11 solves for the spatial equilibrium holding total employment, H constant. This corresponds to the closed-city case and implies the expected utility changes in the counterfactual. We cover the open-city case with endogenous H and fixed exogenous $\bar{U}$ in the next section, where we incorporate endogenous agglomeration forces.

Algorithm 9: Solving for adjusted wage $\tilde{w}_i$, adjusted productivity $\tilde{A}_i$ and adjusted amenity $(\tilde{B}_i)$: `cmodexog.m`

Data: Given values of endogenous variables floor space prices, workplace employment, residence employment, $\{\mathbb{Q}_i, H_{Mj}, H_{Ri}\}$, structural parameters $\{\alpha, \beta, \kappa\varepsilon, \varepsilon\}$, bilateral travel times τ_{ij}, guesses of adjusted productivity and adjusted amenity, $\{\tilde{A}_j^0, \tilde{B}_i^0\}$

1 **while** *guesses of* $\{\tilde{A}_j^0, \tilde{B}_i^0\}$ *change* **do**

2 Compute adjusted wages $\tilde{w}_j$ using guesses of $\tilde{A}_j^0$, observed $\mathbb{Q}_j$, and Eq. (12)

3 Compute commuting probabilities π_{ij} using Eq. (4) using guesses of $\{\tilde{A}_j^0, \tilde{B}_i^0\}$ and $\tilde{w}_j$ using Eq. (4)

4 Use π_{ij} to compute predicted workplace and residence employment $\{\hat{H}_{Mj}, \hat{H}_{Ri}\}$ using Eq. (5)

5 Generate new guesses of $\tilde{A}_i^1$ by inflating old guesses by the ratio of observed over predicted workplace employment $H_{Mi}/\hat{H}_{Mi}$ (we increase guesses if we underpredict employment)

6 Generate new guesses of $\tilde{B}_i^1$ by inflating old guess by the ratio of observed over predicted residence employment $H_{Ri}/\hat{H}_{Ri}$ (we increase guesses if we underpredict employment)

7 Update guesses of $\tilde{A}_i^0$ to the weighted combination of new guess $\tilde{A}_i^1$ and old guesses $\tilde{A}_i^0$

8 Update guesses of $\tilde{B}_i^0$ to the weighted combination of new guess $\tilde{B}_i^1$ and old guesses $\tilde{B}_i^0$

9 Normalize guesses $\tilde{A}_i^0$ by the geometric mean to ensure a unit mean

10 Inflate guesses of $\tilde{B}_i^0$ by the ratio of total employment in data over total predicted employment in model (we make the city more attractive if we underpredict total employment)

11 Use solved $\tilde{w}_j$, τ_{ij}, and $\kappa\epsilon$ to compute commuting market access (see p. 40 in supplement for equation)

Result: Adjusted productivities, adjusted amenities, adjusted wages, commuting probabilities, expected income, predicted workplace emplyoment, predicted residence employment, predicted total employment $\{\tilde{A}_j, \tilde{B}_j, \tilde{w}_j, \pi_{ij}, \mathbb{E}(\tilde{w}_s|i), \hat{H}_{Mi}, \hat{H}_{Ri}\}$

Algorithm 10: Solving for density of development φ_i, total floor space stock L_i, and commercial floor space share θ_i: `cdenisityE.m`

Data: Given values for structural parameters $\{\alpha, \beta\}$, land endowment K_i, observed floor space prices $\mathbb{Q}_j$ adjusted productivity $\tilde{A}_j$, adjusted amenity $\tilde{B}_i$, and total income $\mathbb{E}(w_i)$ solved by Algorithm 9 (`cmodexog.m`)

1 Compute commercial floor space demand $\theta_i L_i$ using Eq. (S.29)

2 Compute residential floor space demand $(1-\theta_i)L_i$ using Eq. (S.30)

3 Compute total floor space demand $L_i = \theta_i L_i + (1-\theta_i)L_I$

4 Use expected worker income $\mathbb{E}(w_i)$ and residence employment to compute total expected income $\mathbb{E}(w_i)H_{Ri}$

5 Compute density of development φ_i using L_i, land area K_i and Eq. (S.31)

6 Compute commercial floor space share $\theta_i = \frac{\theta L_i}{L_i}$

7 Compute commercial commercial and residential floor space using θ_i and L_i

Result: Density of development, total floor space, and commercial floor space share $\{\varphi_i, L_i, \theta_i\}$

Note: This is essentially the same code as in `cdenisity.m`), except that the algorithm also generates the stock of commercial and residential floor space, which serve as guesses in `smodex.m` to speed up convergence.

Algorithm 11: Solving for the equilibrium for given primitives: `smodexog.m`

Data: Given values of structural parameters $\{\alpha, \beta, \kappa, \varepsilon\}$, bilateral travel times τ_{ij}, inverted adjusted productivity, adjusted amenity, and floor space stock $\{\tilde{A}_j, \tilde{B}_i, L_i\}$; guesses of the target variables adjusted wages, commercial and residential floor space prices, commercial floor space shares $\{\tilde{w}_i^0, q_i^0, Q_i^0, \theta_i^0\}$; total employment H.

while *guesses of target variables* $\{\tilde{w}_i^0, q_i^0, Q_i^0, \theta_i^0\}$ *change* **do**

- Compute location choice probabilities π_{ij} using guesses of $\{\tilde{w}_i^0, q_i^0, Q_i^0\}$ and $\{\tilde{A}_j, \tilde{B}_i\}$ in in Eq. (4)
- Compute residence and workplace employment $\{\hat{H}_{Mi}, \hat{H}_{Ri}\}$ using π_{ij}, H and residence and Eq. (5).
- Compute output $\hat{Y}_i$ using the production function in Eq. (10), workplace employment $\hat{H}_{Mi}$, inverted total floor space L_i, and guesses of θ_i^0.
- Compute predicted adjusted wage $\tilde{w}_i^1 = \alpha \frac{\hat{Y}_i}{\hat{H}_{Mi}}$ using the input demand function derived from F.O.C. of Eq. (10) and $\{\hat{Y}_i$, and $\hat{H}_{Mi}\}$
- Compute total income $\mathbb{E}(\hat{w}_i \times \hat{H}_{Mi}\}$ using Eq. (S.20), predicted $\tilde{w}_i^1$, $\hat{H}_{Mi}$, and conditional commuting probabilities $\pi_{ij|i} = \frac{\pi_{ij}}{\sum_j \pi_{ij}}$
- Compute predicted commercial and residential floor space prices $\{q_i^1, Q_i^1\}$ using $\hat{Y}$, guesses of θ_i^0, and Marshallian demand and input demand based on Eqs. (1) and (10)
- Compute predicted values of θ_i^0 using commercial floor space input $L_{Mi} = (1-\alpha)\frac{\hat{Y}_i}{q_i^1}$ recovered from input demand function based on Eq. (10) and L_i in Eq. (S.53)
- Update guesses of target variables $\{\tilde{w}_i^0, q_i^0, Q_i^0, \theta_i^0\}$ to weighted average of old guesses $\{\tilde{w}_i^0, q_i^0, Q_i^0, \theta_i^0\}$ and predicted values $\{w_i^1, q_i^1, Q_i^1, \theta_i^1\}$

Result: Predicted values of adjusted wage, total income, commercial floor space shares, output, commercial and residential floor space prices, workplace employment and residence, total employment, unconditional commuting probabilities $\{\tilde{w}_i, \mathbb{E}(\hat{w}_i \times \hat{H}_{Mi}), \theta_i, Y_i, q_i, Q_i, H_{Mi}, H_{Ri}, \pi_{ij}\}$

A.4 Counterfactuals with endogenous agglomeration forces

Conducting counterfactuals with endogenous agglomeration forces follows the two steps—quantification and simulation—as in the case with exogenous fundamentals. To quantify the model with endogenous agglomeration forces, we first recover adjusted productivity and adjusted amenity $\{\tilde{A}_i, \tilde{B}_i\}$ using either the sequential procedure in Algorithms 4 and 5, plus the rescaling Algorithm 9, or the simultaneous procedure in Algorithm 9. We then break down $\{\tilde{A}_i, \tilde{B}_i\}$ into endogenous components that depend on nearby density, $\{\Upsilon_i, \Omega_i\}$, and exogenous components $\{a_i, b_i\}$ using Algorithms 12 and 13. We then use Algorithm 10 to recover density of development, total floor space, and commercial floor space share $\{\varphi_i, L_i, \theta_i\}$. Using Algorithm 14 to recover the reservation utility level $\bar{U}$ completes the quantification.

For given primitives, Algorithm 15 solves for the equilibrium values of the endogenous outcomes in the closed-city case where total employment H is exogenous and expected utility $\bar{U}$ is endogenous. To this end, it uses an iterative procedure to solve for the values of the target variables adjusted wages, floor space prices, and commercial floor space shares $\{\tilde{w}_i, q_i, Q_i, \theta_i\}$. For the solved values of these target variables, the algorithm generates a range of further endogenous outcomes (all endogenous variables can be recovered if the values of the target variables and primitives are known).

Algorithm 16 similarly solves for the equilibrium values of the endogenous outcomes in the open-city case where $\bar{U}$ is exogenous and total employment H is endogenous. The algorithm is very similar to Algorithm 15. It nests an additional adjustment within the loop in line 17 (of the pseudo code). If the predicted expected utility in the city exceeds the target, total employment is increased, which reduces the expected utility via the model's endogenous congestion force. To ensure that the algorithm converges to the target expected utility, the stopping rule is extended in line 5.

Algorithm 12: Decomposing adjusted productivity $\tilde{A}_i$: `cprod.m`

Data: Given values for structural parameters $\{\lambda, \delta\}$, land endowment K_i, travel times τ_{ij}, adjusted productivities $\tilde{A}_j$ solved by Algorithm 9 and observed workplace employment H_{Mi}

1 Compute endogenous productivity Υ_i using $\{\delta, \tau_{ij}, E_{Mi}, K_i\}$ in Eq. (20)

2 Compute exogenous productivity a_i using $\{\lambda, \tilde{A}_i, \Upsilon_i\}$ in Eq. (20)

Result: Endogenous and exogenous productivity Υ_i, a_i

Algorithm 13: Decomposing adjusted amenity $\tilde{B}_i$: `cres.m`

Data: Given values for structural parameters $\{\eta, \rho\}$, land endowment K_i, travel times τ_{ij}, adjusted amenity $\tilde{B}_j$ solved by Algorithm 9 and observed residence employment H_{Ri}

1 Compute endogemous amenity Ω_i using $\{\rho, \tau_{ij}, E_{Ri}, K_i\}$ in Eq. (21)

2 Compute exogenous amenity b_i using $\{\eta, \tilde{B}_i, \Omega_i\}$ in Eq. (21)

Result: Endogenous and exogenous amenity Ω_i, b_i

Algorithm 14: Computing reservation utility level $\bar{U}$: `ubar.m`

Data: Given values for structural parameters $\{\kappa, \epsilon\}$, travel times τ_{ij}, adjusted amenity $\tilde{B}_j$ solved by Algorithm 9 and floor space prices and wages $\{\mathbb{Q}_i, \tilde{w}_i\}$

1 Compute $\bar{U}$ using all inputs in Eq. (9)

Result: Endogenous and exogenous amenity Ω_i, b_i

Algorithm 15: Solving for the equilibrium with endogenous agglomeration forces and exogenous total employment H: `smodendog.m`

Data: Given values of structural parameters $\{\alpha, \beta, \kappa, \varepsilon, \mu\}$, bilateral travel times τ_{ij}, land area K_i; inverted fundamental productivity and amenity $\{a_i, b_i\}$, adjusted density of development $\tilde{\varphi}_i$; guesses of the target variables: adjusted wages, commercial and residential floor space prices, commercial floor space shares $\{\tilde{w}_i^0, q_i^0, Q_i^0, \theta_i^0\}$ and initial values of workplace and residence employment $\{H_{Mi}, H_{Ri}\}$; total employment H

Compute floor space L_i using $\tilde{\varphi}_i$ and K_i in Eq. (S.52)
Compute adjusted productivity $\tilde{A}_i$ using τ_{ij} and initial values of H_{Mi} and Eq. (20)
Compute adjusted amenity $\tilde{B}_i$ using τ_{ij} and initial values of H_{Ri} and Eq. (21)
Compute total employment $H = \sum H_{Mi}$
while *guesses of target variables* $\{\tilde{w}_i^0, q_i^0, Q_i^0, \theta_i^0\}$ *change* **do**
    Compute choice probabilities π_{ij} using guesses of $\{\tilde{w}_i^0, q_i^0, Q_i^0, \theta_i^0\}$ and $\{\tilde{A}_i, \tilde{B}_i\}$ in Eq. (4)
    Compute residence and workplace employment $\{\hat{H}_{Mi}, \hat{H}_{Ri}\}$ using $\hat{\pi}_{ij}$, H and Eq. (5)
    Compute utility $\hat{U}$ using guesses of $\{\tilde{w}_i^0, q_i^0, Q_i^0, \theta_i^0\}$ and $\{\tilde{A}_i, \tilde{B}_i\}$ in Eq. (9)
    Compute $\tilde{A}_i$ using τ_{ij} and $\hat{H}_{Mi}$ in Eq. (20)
    Compute $\tilde{B}_i$ using τ_{ij} and $\hat{H}_{Ri}$ in Eq. (21)
    Compute output $\hat{Y}_i$ using the production function in Eq. (10), $\hat{H}_{Mi}$, total floor space L_i, and guesses of θ_i^0
    Compute predicted adjusted wage $\tilde{w}_i^1 = \alpha \frac{\hat{Y}_i}{\hat{H}_{Mi}}$ using the input demand function derived from F.O.C. of Eq. (10), and $\{\hat{Y}_i, \hat{H}_{Mi}\}$
    Compute total income $\mathbb{E}(\hat{w}_i \times \hat{H}_{Ri}\}$ using Eq. (S.20), $\{\tilde{w}_i^1, \hat{H}_{Mi}\}$, and conditional commuting probabilities $\pi_{ij|i} = \frac{\Phi_{ij}}{\sum_j \Phi_{ij}}$
    Compute predicted commercial and residential floor space prices $\{q_i^1, Q_i^1\}$ using $\hat{Y}$, guesses of θ_i^0, and Marshallian demand functions and input demand functions based on Eqs. (1) and (10)
    Compute predicted values of commercial floor space θ_i^1 using commercial floor space input $L_{Mi} = (1-\alpha)\frac{\hat{Y}_i}{q_i^1}$ recovered from input demand function based on Eqs. (10) and (S.53)
    Update guesses of target variables $\{\tilde{w}_i^0, q_i^0, Q_i^0, \theta_i^0\}$ to weighted average of old guesses $\{\tilde{w}_i^0, q_i^0, Q_i^0, \theta_i^0\}$ and predicted values $\{w_i^1, \hat{q}_i^1, \hat{Q}_i^1, \hat{\theta}_i^1\}$

Result: Predicted values of adjusted wage, total income, commercial floor space shares, output, commercial and residential floor space prices, workplace employment and residence, total employment, unconditional commuting probabilities $\{\tilde{w}_i, \mathbb{E}(\hat{w}_i \times \hat{H}_{Mi}, \theta_i, Y_i, q_i, Q_i, H_{Mi}, H_{Ri}, \tilde{A}i, \tilde{B}_i, \bar{U}, \pi_{ij}\}$

Algorithm 16: Solving for the equilibrium with endogenous agglomeration forces and exogenous reservation utility: `ussmodendog.m`

Data: Given values of structural parameters $\{\alpha, \beta, \kappa, \varepsilon, \mu\}$, bilateral travel times τ_{ij}, land area K_i; inverted fundamental productivity and amenity $\{a_i, b_i\}$, adjusted density of development $\tilde{\varphi}_i$; guesses of the target variables: adjusted wages, commercial and residential floor space prices, commercial floor space shares $\{\tilde{w}_i^0, q_i^0, Q_i^0, \theta_i^0\}$ and initial values of workplace and residence employment $\{H_{Mi}, H_{Ri}\}$; calibrated value of exogenous reservation utility level $\bar{U}$

- Compute floor space L_i using $\tilde{\varphi}_i$ and K_i in Eq. (S.52)
- Compute adjusted productivity $\tilde{A}_i$ using τ_{ij} and initial values of H_{Mi} and Eq. (20)
- Compute adjusted amenity $\tilde{B}_i$ using τ_{ij} and initial values of H_{Ri} and Eq. (21)
- Compute total employment $H = \sum H_{Mi}$
- **while** *guesses of target variables* $\{\tilde{w}_i^0, q_i^0, Q_i^0, \theta_i^0\}$ *change or* $\mathbb{E}(U) \neq \bar{U}$ **do**
    - Compute choice probabilities π_{ij} using guesses of $\{\tilde{w}_i^0, q_i^0, Q_i^0, \theta_i^0\}$ and $\{\tilde{A}_i, \tilde{B}_i\}$ in Eq. (4)
    - Compute residence and workplace employment $\{\hat{H}_{Mi}, \hat{H}_{Ri}\}$ using $\hat{\pi}_{ij}$, H and Eq. (5)
    - Compute utility $\hat{U}$ using guesses of $\{\tilde{w}_i^0, q_i^0, Q_i^0, \theta_i^0\}$ and $\{\tilde{A}_i, \tilde{B}_i\}$ in Eq. (9)
    - Compute $\tilde{A}_i$ using τ_{ij} and $\hat{H}_{Mi}$ in Eq. (20)
    - Compute $\tilde{B}_i$ using τ_{ij} and $\hat{H}_{Ri}$ in Eq. (21)
    - Compute output $\hat{Y}_i$ using the production function in Eq. (10), $\hat{H}_{Mi}$, total floor space L_i, and guesses of θ_i^0
    - Compute predicted adjusted wage $\tilde{w}_i^1 = \alpha \frac{\hat{Y}_i}{\hat{H}_{Mi}}$ using the input demand function derived from F.O.C. of Eq. (10), and $\{\hat{Y}_i, \hat{H}_{Mi}\}$
    - Compute total income $\mathbb{E}(\hat{w}_i \times \hat{H}_{Ri}\}$ using Eq. (S.20), $\{\tilde{w}_i^1, \hat{H}_{Mi}\}$, and conditional commuting probabilities $\pi_{ij|i} = \frac{\Phi_{ij}}{\sum_j \Phi_{ij}}$
    - Compute predicted commercial and residential floor space prices $\{q_i^1, Q_i^1\}$ using $\hat{Y}$, guesses of θ_i^0, and Marshallian demand functions and input demand functions based on Eqs. (1) and (10)
    - Compute predicted values of commercial floor space θ_i^1 using commercial floor space input $L_{Mi} = (1-\alpha)\frac{\hat{Y}_i}{q_i^1}$ recovered from input demand function based on Eqs. (10) and (S.53)
    - Update guesses of target variables $\{\tilde{w}_i^0, q_i^0, Q_i^0, \theta_i^0\}$ to weighted average of old guesses $\{\tilde{w}_i^0, q_i^0, Q_i^0, \theta_i^0\}$ and predicted values $\{w_i^1, \hat{q}_i^1, \hat{Q}_i^1, \hat{\theta}_i^1\}$
    - Update H using adjustment factor $\frac{\hat{U}}{\bar{U}}$ (we increase total employment if expected utility exceeds reservation utility)

Result: Predicted values of adjusted wage, total income, commercial floor space shares, output, commercial and residential floor space prices, workplace employment and residence, total employment, unconditional commuting probabilities $\{\tilde{w}_i, \mathbb{E}(\hat{w}_i) \times \hat{H}_{Mi}, \theta_i, Y_i, q_i, Q_i, H_{Mi}, H_{Ri}, \tilde{A}i, \tilde{B}_i, H, \pi_{ij}\}$