Quantitative Spatial Economics 2

Atsushi Yamagishi

October 8, 2024

Commuting and within-city QSE models

- So far, we have implicitly assumed that people live and work in the same location. Is it plausible?

- Maybe not within a city, because **commuting** is a salient feature
  - This allows the separation of residence and workplace
  - Classic issue in urban economics: A central feature of Alonso-Muth-Mills monocentric city models and subsequent studies (e.g., Fujita 1989)

- I introduce Ahlfeldt, Redding, Sturm, Wolf (2015 ECMA) as a primary example of the QSE models with commuting[1]
  - I also discuss a few more applications of the QSE models with commuting.

[1] My lecture on the ARSW paper builds on Adrien Bilal's lecture note
`https://drive.google.com/file/d/1QJKqIF0YPxUrVWKqbBnkuGTQ1aPtNnIx/view`

Model overview: Ahlfeldt, Redding, Sturm, Wolf (2015 ECMA)

- The model is richer than Redding (2016) by including commuting and accommodating commercial land use
  - But the model is simpler in other aspects (e.g., no differentiated goods and trade costs)
  - And many features of the model are similar to each other
- Discrete set of blocks $i = 1...S$ in a city (West Berlin in this paper)
- Workers choose the residential block i and the commuting block j.
- Floor space L_i in block i. Used both for residential and commercial uses
  - Not the land but floor space $\rightarrow$ the floor space is competitively supplied by using numeraire goods and land.
- There is an "outside option" of living outside. The outside utility is exogenous and $\bar{U}$
  - Similar assumption to the Rosen-Roback model.
  - Less problematic than the case of Rosen-Roback model since we are focusing on a single city in a larger economy
- Final goods are homogeneous and freely traded, as the Rosen-Roback model

Workers

- Utility when living in i and working in j is written as follows:

$$U_{ij} = B_i e^{-\kappa \tau_{ij}} z_{ij} \left(\frac{C_{ij}}{\beta} \right)^{\beta} \left(\frac{l_{ij}}{1-\beta} \right)^{1-\beta},$$

  - z_{ij} is the idiosyncratic utility shock for the pair (i, j)
  - τ_{ij} is the commuting time between i and j, κ is the commuting cost per minute
  - B_i is the local amenity of location i (scenic views, shops etc).
  - C_{ij} is the numeraire goods consumption and l_{ij} is the floor space consumption

- After choosing i and j, workers maximize this subject to $C_{ij} + Q_i l_{ij} = w_j$, where w_j is the wage rate and Q_i is the price of residential floor space.

- The indirect utility function:

$$V_{ij} = B_i e^{-\kappa \tau_{ij}} z_{ij} w_j Q_i^{\beta - 1}$$

- Notations are different from Redding (2016), but note the similarities.

Workers

- Assume that z_{ij} follows the Frechet distribution with the level parameter $T_i E_j$ and the dispersion parameter ϵ
  - T_i (E_j) is the "exogenous attractiveness" of residence (workplace)
  - From the result in previous lecture, V_{ij} also follows the Frechet distribution with the level parameter $\Phi_{ij} \equiv T_i E_j \left(B_i e^{-\kappa \tau_{ij}} w_j Q_i^{\beta-1} \right)^{\epsilon}$
- The probability of living in i and working in j is

$$\pi_{ij} = \frac{\Phi_{ij}}{\sum_{i,j} \Phi_{ij}} = \frac{T_i E_j \left(B_i e^{-\kappa \tau_{ij}} w_j Q_i^{\beta-1} \right)^{\epsilon}}{\sum_{i,j} T_i E_j \left(B_i e^{-\kappa \tau_{ij}} w_j Q_i^{\beta-1} \right)^{\epsilon}}$$

- Unconditional residential probability, unconditional workplace probability, and conditional commuting probability are, respectively,

$$\pi_{Ri} = \sum_j \pi_{ij}, \;\; \pi_{Mj} = \sum_i \pi_{ij}, \;\; \pi_{ij|i} = \frac{\pi_{ij}}{\sum_k \pi_{ik}}$$

- Note: The paper also uses H to denote the same object (e.g., $H_{Ri} = \pi_{Ri}$ for residential population, $H_{Mi} = \pi_{Mi}$ for workforce size). Be mindful as this can be confusing

"Gravity structure" of commuting

- Taking the log of the location choice probability, we obtain the "gravity equation of commuting":

$$\ln \pi_{ij} = -\kappa\epsilon\tau_{ij} + \vartheta_i + \varsigma_j,$$

where $\vartheta_i \equiv \ln(T_i B_i^{\epsilon} Q_i^{\epsilon(\beta-1)})$, $\varsigma_j = \ln(E_j w_j^{\epsilon}) - \ln(\sum_{i,j}\Phi_{ij})$.

- The coefficient of the bilateral travel time (τ_{ij}) combines the unit travel cost (κ) and the Frechet dispersion parameter (ϵ)
  - ϵ governs the sensitivity of commuting with respect to utility (including utility loss from commuting)
  - κ governs the sensitivity of utility with respect to commuting cost.
  - Both parameter reduces long-distance commuting

- The fixed effects capture the attractiveness of each location as residence or workplace
  - We can thus use the commuting flow data to infer the attractiveness of each location (Kreindler and Miyauchi 2023 REStat).

“Gravity structure” of commuting

- In estimation, we estimate the following specification that adds an ad-hoc error term (e_{ij})

$$\ln \pi_{ij} = -\kappa \epsilon \tau_{ij} + \vartheta_i + \varsigma_j + e_{ij}.$$

e_{ij} does not have a theoretical foundation, but just comes from outside the model.

- With commuting flow data, we can estimate the above equation by the OLS.
  - For instance, in Japan, the Population Census, the Person-Trip Survey, and the GPS data provide such commuting flow information.
- But we should be careful that many π_{ij} are zero so that $\ln \pi_{ij}$ is undefined
  - A common approach: Use the PPML by Santos Silva and Tenreyro (2006 REStat). Stata/R packages are available (e.g., `ppmlhdfe`)
  - Chen and Roth (2024 QJE forthcoming) provides further discussions about what to do when the outcome variable includes “logs with zeros”

Estimation of commuting gravity equation

- Good fit of the model (high R^2). $\kappa\epsilon \simeq 0.07$, which is called the "commuting elasticity"
- OLS and PPML yield similar results in this data, but they can substantially differ in other datasets

TABLE III

COMMUTING GRAVITY EQUATION[a]

	(1) ln Bilateral Commuting Probability 2008	(2) ln Bilateral Commuting Probability 2008	(3) ln Bilateral Commuting Probability 2008	(4) ln Bilateral Commuting Probability 2008
Travel Time ($-\kappa\varepsilon$)	-0.0697^{***}	-0.0702^{***}	-0.0771^{***}	-0.0706^{***}
	(0.0056)	(0.0034)	(0.0025)	(0.0026)
Estimation	OLS	OLS	Poisson PML	Gamma PML
More than 10 Commuters		Yes	Yes	Yes
Fixed Effects	Yes	Yes	Yes	Yes
Observations	144	122	122	122
R^2	0.8261	0.9059	–	–

[a]Gravity equation estimates based on representative micro survey data on commuting for Greater Berlin for 2008. Observations are bilateral pairs of 12 workplace and residence districts (post 2001 Bezirke boundaries). Travel time is measured in minutes. Fixed effects are workplace district fixed effects and residence district fixed effects. The specifications labelled more than 10 commuters restrict attention to bilateral pairs with 10 or more commuters. Poisson PML is Poisson Pseudo Maximum Likelihood estimator. Gamma PML is Gamma Pseudo Maximum Likelihood Estimator. Standard errors in parentheses are heteroscedasticity robust. * significant at 10%; ** significant at 5%; *** significant at 1%.

Spatial equilibrium condition with the outside location

- In equilibrium, the outside utility ($\bar{U}$) satisfies the following:

$$\bar{U} = \Gamma(\frac{\epsilon - 1}{\epsilon}) \left[\sum_{i,j} T_i E_j \left(B_i w_j e^{-\kappa \tau_{ij}} Q_i^{\beta - 1} \right)^{\epsilon} \right]^{1/\epsilon}$$

- The right hand side is the expected utility from living in the city
  - Given that V_{ij} follows the Frechet distribution, we can derive it following the same strategy as in the previous lecture
- This is called an "open city assumption." We can also consider a "closed city" version in which population level is fixed but the utility level varies
  - Open city model: The city welfare is fixed at $\bar{U}$. Total population of the city $\bar{L}$ adjusts.
  - Closed city model: Total population of the city $\bar{L}$ is fixed. The city welfare can change.

Production

- Cobb-Douglas production function at block j:

$$y_j = A_j \underbrace{H_{Mj}^{\alpha}}_{\text{labor}} \underbrace{L_{Mj}^{1-\alpha}}_{\text{floor space}}$$

- A_j is an endogenous object
  - We later introduce agglomeration forces in A_j

- Solving the cost $(w_j H_{Mj} + q_j L_{Mj})$ minimization problem and imposing the zero profit condition (i.e., the unit goods cost $= 1$), the equilibrium commercial floor space price satisfies

$$q_j \propto A_j^{1/(1-\alpha)} w_j^{-\alpha/(1-\alpha)}$$

(see the production part of Rosen-Roback slides for derivation)

Floor space (housing market)

- No arbitrage condition between residential floor use and production floor use implies

$$q_i = \xi_i Q_i$$

  in mixed locations

- ξ_i is an ad-hoc wedge to match the data
  - If land can be used for either residence or production, then $\xi_i = 1$ must hold in a mixed location.
  - Therefore, ξ_i represents some form of local friction of land use
  - Called "land use regulations" in the paper, but unclear what it represents in Berlin

- All locations are mixed in equilibrium when $A_j, B_i > 0$ due to the Inada condition and the Frechet shock.

Floor space (housing market)

- Floor space supply function:

$$L_i = \underbrace{M_i^{\mu}}_{\text{Freely traded capital}} \times \underbrace{K_i^{1-\mu}}_{\text{Fixed supply of land}} .$$

Let $\mathbb{Q}_i \equiv \max\{q_i, Q_i\}$ be the price of the floor space.

- Solving the cost minimization problem under the common capital price $\mathbb{P}$ and the land price $\mathbb{R}_i$, and then imposing the zero profit condition, we have

$$\mathbb{Q}_i = \chi \mathbb{R}_i^{1-\mu},$$

where χ is a constant.

- Therefore, log floor space price $(\ln \mathbb{Q}_i)$ is proportional to log land price $(\ln \mathbb{R}_i)$, and the slope equals the input share of land in floor space production.
  - We briefly saw this result in the "canonical spatial models" slides

Floor space (housing market)

- Residential floor space market clearing

$$Q_i(1-\theta_i)L_i = (1-\beta)E(w_j|i)H_{Ri},$$

where θ_i is the fraction of floor space used for commerce

- Commercial floor space market clearing

$$\theta_j L_j = \left(\frac{(1-\alpha)A_j}{q_j}\right)^{1/\alpha} H_{Mi}$$

Agglomeration forces in productivity

- Assume

$$A_j = a_j \times \left(\sum_k \underbrace{e^{-\delta \tau_{jk}}}_{\text{“distance” of } j \text{ and } k} \underbrace{\frac{H_{Mk}}{K_k}}_{\text{Employment density at } k} \right)^{\lambda}$$

- The productivity level at block j is determined by
  - Some fundamental productivity a_j, which captures natural conditions (e.g., access to rivers) etc.
  - Employment density in nearby areas. This captures endogenous agglomeration forces (e.g., knowledge spillovers)
    - λ is the strength of agglomeration forces. δ controls its distance decay.

- Note: See Fujita and Ogawa (1982 RSUE) for a seminal work that introduces such productivity spillovers in an urban model.

Agglomeration forces in amenities

- Assume

$$B_i = b_i \times \left(\sum_k \underbrace{e^{-\rho \tau_{ik}}}_{\text{"distance" of } i \text{ and } k} \underbrace{\frac{H_{Rk}}{K_k}}_{\text{Employment density at } k} \right)^{\eta}$$

- The amenity level at block i is determined by
  - Some fundamental amenities b_i, which captures natural conditions (e.g., scenic view) etc.
  - Population density in nearby areas. This captures endogenous agglomeration forces (e.g., shopping environments)
    - η is the strength of agglomeration forces. ρ controls its distance decay.

Existence and uniqueness of equilibrium

- When agglomeration parameters $\lambda = \eta = 0$, the equilibrium exists and is unique.
  - See Ahlfeldt et al. (2015) for the proof. In principle. the proof strategy is analogous to Redding (2016).
- When $\lambda, \eta \neq 0$, there can be multiple equilibria
  - Empirically, this paper finds $\lambda > 0$ and $\eta > 0$. So this is a relevant situation.
  - The equilibrium selection issue arises. This paper chooses an equilibrium that is "numerically closest" to the current one.

Calibration of the model

- Similar three-step structure as Redding, but we want to estimate the agglomeration parameters λ, η (and also the agglomeration decay parameters η, ρ)
    1. Determine the parameter values of α, β, μ, ϵ, κ, $\underbrace{\lambda, \delta, \eta, \rho}_{\text{agglomeration-relevant parameters}}$
    2. Then, from the observed data (e.g., population and employment distribution, floor space price), we can calculate A_i, B_i
    3. Then, from $A_j = a_j \times \left(\sum_k e^{-\delta \tau_{jk}} \frac{H_{Mk}}{K_k} \right)^{\lambda}$, we can recover a_j. Similarly we get b_i.
- But all of the above three steps takes $(\lambda, \delta, \eta, \rho)$ as given $\rightarrow$ we need to find "plausible $(\lambda, \delta, \eta, \rho)$"
- The approach of ARSW paper: Apply the above three steps for *all* values of $(\lambda, \delta, \eta, \rho)$. Then, calculate some moments from the calculated a_i and b_i, and choose $(\lambda, \delta, \eta, \rho)$ that best satisfies the moment conditions
    - That is, the Generalized Method of Moments (GMM) estimation of $(\lambda, \delta, \eta, \rho)$.

Empirical background and Data

- Berlin before and after the war
  - Berlin wall built in 1961 suddenly shuts down the access to East Germany
  - Berlin wall fell down in 1989.
- Data for 1936, 1986, and 2006.
- About 16,000 blocks. Average 274 residents.
- Observes population, employment, and land prices at each block.
- Commuting flow data are also available
  - Travel time is calculated using transportation network data
- How does the changes in the market access, caused by the Berlin wall, affect the distribution of economic activities?
  - And can we use this change in the market access to estimate the strength of agglomeration forces?

Land gradient over time

- Areas near the Berlin wall experienced the decline in land prices, but it re-surged after reunification

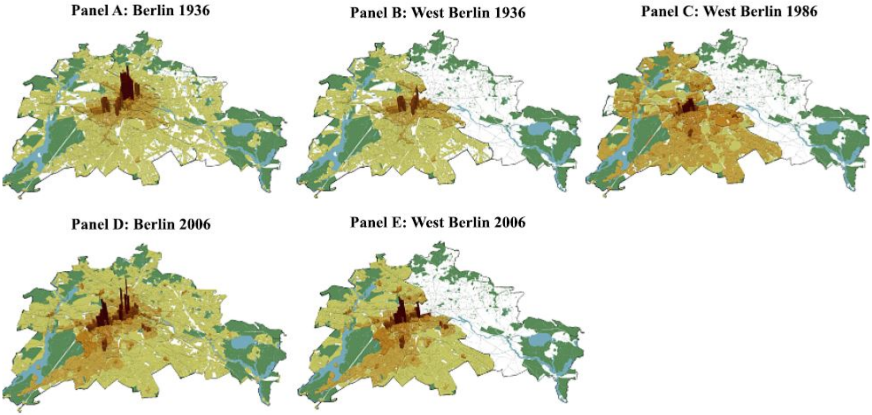


FIGURE 2.—The evolution of land prices in Berlin over time.

Calibration and counterfactual analysis

- We now calibrate/estimate key parameters of the model
  - Commuting time cost κ
  - Frechet dispersion parameter ϵ
  - Agglomeration force parameters $(\lambda, \delta, \eta, \rho)$ and fundamental parameters (a_i, b_i)
  - Other parameters (e.g., α, β) are taken from other studies
- Then, ARSW simulates the counterfactual population and employment distribution *with and without agglomeration forces*
  - If the model still successfully predicts what happens after the division and reunification, then agglomeration forces may not be important
  - If the model now fails to predict what happens after the division and reunification, then this is a structural evidence of the importance of agglomeration forces

Commuting gravity estimation

- Estimate the commuting gravity equation in the OLS and PPML.

$$\ln \pi_{ij} = -\kappa\epsilon\tau_{ij} + \vartheta_i + \varsigma_j + e_{ij}.$$

- $\nu \equiv \kappa\epsilon \simeq 0.07$.

TABLE III

COMMUTING GRAVITY EQUATION[a]

	(1) ln Bilateral Commuting Probability 2008	(2) ln Bilateral Commuting Probability 2008	(3) ln Bilateral Commuting Probability 2008	(4) ln Bilateral Commuting Probability 2008
Travel Time ($-\kappa\varepsilon$)	−0.0697***	−0.0702***	−0.0771***	−0.0706***
	(0.0056)	(0.0034)	(0.0025)	(0.0026)
Estimation	OLS	OLS	Poisson PML	Gamma PML
More than 10 Commuters		Yes	Yes	Yes
Fixed Effects	Yes	Yes	Yes	Yes
Observations	144	122	122	122
R^2	0.8261	0.9059	–	–

[a]Gravity equation estimates based on representative micro survey data on commuting for Greater Berlin for 2008. Observations are bilateral pairs of 12 workplace and residence districts (post 2001 Bezirke boundaries). Travel time is measured in minutes. Fixed effects are workplace district fixed effects and residence district fixed effects. The specifications labelled more than 10 commuters restrict attention to bilateral pairs with 10 or more commuters. Poisson PML is Poisson Pseudo Maximum Likelihood estimator. Gamma PML is Gamma Pseudo Maximum Likelihood Estimator. Standard errors in parentheses are heteroscedasticity robust. * significant at 10%; ** significant at 5%; *** significant at 1%.

Recovering (adjusted) wage, ϵ, and some other parameters

- Recover $\omega_j \equiv w_j^{\epsilon}$ through observed workplace and residential employment:

$$H_{Mj} = \sum_i H_{Ri} \frac{\omega_j e^{-\nu \tau_{ij}}}{\sum_k \omega_k e^{-\nu \tau_{ik}}}$$

  - This involves normalizing $E_j = 1$ for all j. This is okay as A_j and E_j are not separately identified.

- Find ϵ to match the observed wage distribution in data:

$$\epsilon^2 = \frac{Var(\ln \omega_j)}{Var(\ln w_j)}$$

- Then. the commuting time cost $\kappa = \nu / \epsilon$
  - Note: There are alternative approaches used in the literature to determine the value of ϵ. See, for instance, Dingel and Tintelnot (2023) and Takeda and Yamagishi (2024).

- Set other parameters
  - $1 - \beta = 0.25$ to match the observed spending share for housing.
  - $1 - \alpha = 0.2$ is chosen to match the observed share of firm expenditure on commercial floor space.
  - $1 - \mu = 0.25$ to match the share of land in construction costs

Recovering overall productivity and amenities

- Productivity: from firms' optimality conditions:

$$\ln \frac{A_i}{\bar{A}} = (1-\alpha) \ln \frac{Q_i}{\bar{Q}} + \frac{\alpha}{\epsilon} \ln \frac{\omega_i}{\bar{\omega}},$$

  where $\bar{X}$ denotes its geometric average.

- Amenities: from residential choices:

$$\ln \frac{B_i}{\bar{B}} = \frac{1}{\epsilon} \ln \frac{H_{Ri}}{\bar{H}_R} + (1-\beta) \ln \frac{Q_i}{\bar{Q}} - \frac{1}{\epsilon} \ln \frac{W_i}{\bar{W}}, \text{ where } W_i \equiv \sum_j \omega_j e^{-\nu \tau_{ij}}$$

- A_i and B_i combine both exgogenous fundamentals and endogenous agglomeration forces
  - We now decompose them into these two components

Recovering overall productivity and amenities

- *Conditional on the parameter values of λ and δ*, we can decompose A_i into the fundamental a_i and the endogenous agglomeration forces.
- Recall $A_i = a_i \Upsilon_i^{\lambda}$, where $\Upsilon_i = \sum_k e^{-\delta \tau_{ik}} \frac{H_{Mk}}{K_k}$
  - Since we observe the employment distribution and land supply at each location, we can construct Υ_i once we know the value of δ
- Substituting this into the previous equation and taking time differences,

$$\Delta \ln \frac{a_i}{\bar{a}} = (1-\alpha)\Delta \ln \frac{Q_i}{\bar{Q}} + \frac{\alpha}{\epsilon}\Delta \ln \frac{\omega_i}{\bar{\omega}} - \lambda \Delta \ln \frac{\Upsilon_i}{\bar{\Upsilon}},$$

  which we can compute given the value of λ.
- Therefore, given (λ, δ), we can obtain the fundamental productivity a_i
- We can similarly obtain b_i

Recovering overall productivity and amenities

$$\underbrace{\Delta \ln \frac{a_i}{\bar{a}}}_{\text{Unknown}} = \underbrace{(1-\alpha)\Delta \ln \frac{Q_i}{\bar{Q}} + \frac{\alpha}{\epsilon}\Delta \ln \frac{\omega_i}{\bar{\omega}}}_{\text{Known}} - \underbrace{\lambda}_{\text{Unknown}} \underbrace{\Delta \ln \frac{\Upsilon_i}{\bar{\Upsilon}}}_{\text{Known (except for } \delta)} .$$

- Impose the moment conditions $E(\Delta \ln \frac{a_i}{\bar{a}} | i \in \kappa_k) = 0$, where κ_k are distance bins from the CBD.
  - Note that $E(\Delta \ln \frac{a_i}{\bar{a}}) = 0$ *by construction*. So we require that changes in fundamentals is not systematically correlated with distance from the CBD.
  - Intuitively, about the distance from the CBD, there should be no systematic changes in fundamental productivity over time.
  - That is, regardless of distance from the CBD, productivity changes over time should be captured by model's endogenous agglomeration forces, not exogenous fundamentals
- By using the GMM using these moment conditions, we can estimate λ and δ
  - $\lambda > 0$. δ is relatively large (spatial decay is fast).
- We can do the similar for amenities to estimate η and ρ
  - $\eta > 0$. ρ is relatively large (spatial decay is fast).

Counterfactual results

- Column 1 represents the decline pattern of floor prices in the post-division data

TABLE I

BASELINE DIVISION DIFFERENCE-IN-DIFFERENCE RESULTS (1936–1986)[a]

	(1) $\Delta \ln Q$	(2) $\Delta \ln Q$	(3) $\Delta \ln Q$	(4) $\Delta \ln Q$	(5) $\Delta \ln Q$	(6) $\Delta \ln$ EmpR	(7) $\Delta \ln$ EmpR	(8) $\Delta \ln$ EmpW	(9) $\Delta \ln$ EmpW
CBD 1	−0.800***	−0.567***	−0.524***	−0.503***	−0.565***	−1.332***	−0.975***	−0.691*	−0.639*
	(0.071)	(0.071)	(0.071)	(0.071)	(0.077)	(0.383)	(0.311)	(0.408)	(0.338)
CBD 2	−0.655***	−0.422***	−0.392***	−0.360***	−0.400***	−0.715**	−0.361	−1.253***	−1.367***
	(0.042)	(0.047)	(0.046)	(0.043)	(0.050)	(0.299)	(0.280)	(0.293)	(0.243)
CBD 3	−0.543***	−0.306***	−0.294***	−0.258***	−0.247***	−0.911***	−0.460**	−0.341	−0.471**
	(0.034)	(0.039)	(0.037)	(0.032)	(0.034)	(0.239)	(0.206)	(0.241)	(0.190)
CBD 4	−0.436***	−0.207***	−0.193***	−0.166***	−0.176***	−0.356**	−0.259	−0.512***	−0.521***
	(0.022)	(0.033)	(0.033)	(0.030)	(0.026)	(0.145)	(0.159)	(0.199)	(0.169)
CBD 5	−0.353***	−0.139***	−0.123***	−0.098***	−0.100***	−0.301***	−0.143	−0.436***	−0.340***
	(0.016)	(0.024)	(0.024)	(0.023)	(0.020)	(0.110)	(0.113)	(0.151)	(0.124)
CBD 6	−0.291***	−0.125***	−0.094***	−0.077***	−0.090***	−0.360***	−0.135	−0.280**	−0.142
	(0.018)	(0.019)	(0.017)	(0.016)	(0.016)	(0.100)	(0.089)	(0.130)	(0.116)
Inner Boundary 1–6			Yes	Yes	Yes		Yes		Yes
Outer Boundary 1–6			Yes	Yes	Yes		Yes		Yes
Kudamm 1–6				Yes	Yes		Yes		Yes
Block Characteristics					Yes		Yes		Yes
District Fixed Effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	6,260	6,260	6,260	6,260	6,260	5,978	5,978	2,844	2,844
R^2	0.26	0.51	0.63	0.65	0.71	0.19	0.43	0.12	0.33

[a] Q denotes the price of floor space. EmpR denotes employment by residence. EmpW denotes employment by workplace. CBD1–CBD6 are six 500 m distance grid cells for distance from the pre-war CBD. Inner Boundary 1–6 are six 500 m grid cells for distance to the Inner Boundary between East and West Berlin. Outer Boundary 1–6 are six 500 m grid cells for distance to the outer boundary between West Berlin and East Germany. Kudamm 1–6 are six 500 m grid cells for distance to Breitscheid Platz on the Kurfürstendamm. The coefficients on the other distance grid cells are reported in Table A.2 of the Technical Data Appendix. Block characteristics include the log distance to schools, parks and water, the land area of the block, the share of the block's built-up area destroyed during the Second World War, indicators for residential, commercial and industrial land use, and indicators for whether a block includes a government building and urban regeneration policies post-reunification. Heteroscedasticity and Autocorrelation Consistent (HAC) standard errors in parentheses (Conley (1999)). * significant at 10%; ** significant at 5%; *** significant at 1%.

Counterfactual results

- Comparing the model prediction with the data, the model with agglomeration forces fits the data well (column 1)
- However, shutting down production agglomeration forces (column 2) or amenity agglomeration forces (column 3) leads to larger discrepancy with the data
- Importance of agglomeration forces in explaining what happened after division of Berlin

TABLE VII

COUNTERFACTUALS[a]

	(1) $\Delta \ln QC$ 1936–1986	(2) $\Delta \ln QC$ 1936–1986	(3) $\Delta \ln QC$ 1936–1986	(4) $\Delta \ln QC$ 1936–1986	(5) $\Delta \ln QC$ 1986–2006	(6) $\Delta \ln QC$ 1986–2006	(7) $\Delta \ln QC$ 1986–2006
CBD 1	−0.836***	−0.613***	−0.467***	−0.821***	0.363***	1.160***	0.392***
	(0.052)	(0.032)	(0.060)	(0.051)	(0.041)	(0.052)	(0.043)
CBD 2	−0.560***	−0.397***	−0.364***	−0.624***	0.239***	0.779***	0.244***
	(0.034)	(0.025)	(0.019)	(0.029)	(0.028)	(0.044)	(0.027)
CBD 3	−0.455***	−0.312***	−0.336***	−0.530***	0.163***	0.594***	0.179***
	(0.036)	(0.030)	(0.030)	(0.036)	(0.031)	(0.045)	(0.031)
CBD 4	−0.423***	−0.284***	−0.340***	−0.517***	0.140***	0.445***	0.143***
	(0.026)	(0.019)	(0.022)	(0.031)	(0.021)	(0.042)	(0.021)
CBD 5	−0.418***	−0.265***	−0.351***	−0.512***	0.177***	0.403***	0.180***
	(0.032)	(0.022)	(0.027)	(0.039)	(0.032)	(0.038)	(0.032)
CBD 6	−0.349***	−0.222***	−0.304***	−0.430***	0.100***	0.334***	0.103***
	(0.025)	(0.016)	(0.022)	(0.029)	(0.024)	(0.034)	(0.023)
Counterfactuals	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Agglomeration Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	6,260	6,260	6,260	6,260	7,050	6,260	7,050
R^2	0.11	0.13	0.07	0.13	0.12	0.24	0.13

[a]Columns (1)–(6) are based on the parameter estimates pooling division and reunification from Table V. Column (7) is based on the parameter estimates for division from Table V. QC denotes counterfactual floor prices. Column (1) simulates division using our estimates of production and residential externalities and 1936 fundamentals. Column (2) simulates division using our estimates of production externalities and 1936 fundamentals but setting residential externalities to zero. Column (3) simulates division using our estimates of residential externalities and 1936 fundamentals but setting production externalities to zero. Column (4) simulates division using our estimates of production and residential externalities and 1936 fundamentals but halving their rates of spatial decay with travel time. Column (5) simulates reunification using our estimates of production and residential externalities, 1986 fundamentals for West Berlin, and 2006 fundamentals for East Berlin. Column (6) simulates reunification using our estimates of production and residential externalities, 1986 fundamentals for West Berlin and 1936 fundamentals for East Berlin. Column (7) simulates reunification using division rather than pooled parameter estimates, 1986 fundamentals for West Berlin, and 2006 fundamentals for East Berlin. CBD 1–CBD6 are six 500 m distance grid cells for distance from the pre-war CBD. Heteroscedasticity and Autocorrelation Consistent (HAC) standard errors in parentheses (Conley (1999)). * significant at 10%; ** significant at 5%; *** significant at 1%.

Applications of the within-city QSE models

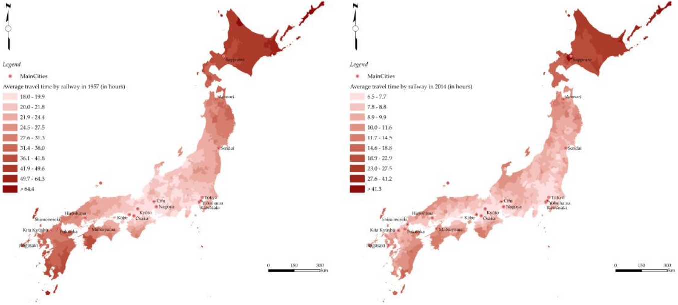
- I now briefly discuss a few more applications of the QSE models with commuting
  - Hayakawa, Koster, Tabuchi, Thisse (2021 wp)
  - Delventhal, Kwon, Parkhomenko (2022 JUE)
  - Dingel and Tintelnot (2023 ECMA R&R)
- The models in these applications are similar to Ahlfeldt, Redding, Sturm, Wolf (2015) but not exactly the same. However, the basic structure of the analysis in ARSW is portable to these applications

Hayakawa, Koster, Tabuchi, Thisse (2021 wp)

- What is the spatial economic impact of high-speed rail (HSR)?
  - Japan's *shinkansen* as the world's first example of the HSR
- Roughly speaking, the model is similar to Redding (2016) but it has commuting.
  - The model is closest to Monte, Redding, Rossi-Hansberg (2018 AER)
- This paper is a good reference for a QSE analysis using Japanese data. You may want to carefully read this paper to see which data are useful, if you plan to use the QSE approach using modern Japanese data.

Hayakawa, Koster, Tabuchi, Thisse (2021 wp)

- Railway access has substantially changed from 1957 (prior to shinkansen) to 2014
- This has changed the transaction costs between firms
  - Using the TSR (Tokyo Shoko Research) data on firm-to-firm transactions, they provide evidence that firm-to-firm transactions responds more to train access rather than road access.
  - Suggests that face-to-face interactions are more important for firm transactions than goods trade

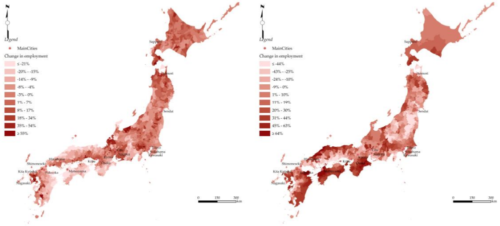


(A) Average travel time in 1957 (B) Average travel time in 2014

Figure 2 – Railway accessibility

Hayakawa, Koster, Tabuchi, Thisse (2021 wp)

- Two counterfactual analyses.
- The planned shinkansen extension would increase the welfare by 5%
  - This also increases the size of Tokyo, Osaka, and Nagoya (especially Nagoya: 11.2% increase)
- Without shinkansen, Tokyo and Osaka would be 6.3% and 4,4% larger but Nagoya would be smaller by 25%.
  - Relative position of municipalities within transport networks and their underlying location fundamentals are important in understanding why the effects of a large infrastructure are positive or negative.



(A) Extension of the Shinkansen network (B) No Shinkansen
Figure 4 – Counterfactual experiments: spatial distribution of employment

Delventhal, Kwon, Parkhomenko (2022 JUE)

- What is the effect of remote work on city structure?
  - Important after the COVID-19 and the increased share of remote work
- Similar model as ARSW, but simple modification to the commuting cost d_{ij}:
  - With probability ψ, a worker becomes "telecommuter" and becomes the traditional commuter otherwise.
  - θ is the fraction of workdays they commute. $\theta = \theta^T$ for telecommuters and $\theta = 1$ for the traditional commuters
  - $d_{ij} = (1 - \theta) + \theta e^{-\kappa t_{ij}}$
- Results:
  - Employment clusters more in the city center, while population move to the periphery
  - Reduced traffic congestion (note: traffic congestion is absent in the ARSW model!)
  - Lower housing costs in the city center, higher in the periphery
- Ongoing topic: See, for instance, Delventhal and Parkhomenko (2023 RES R&R), Monte, Porcher, Rossi-Hansberg (2023 AER R&R) for more recent studies.

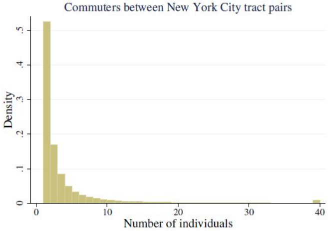
Dingel and Tintelnot (2023 ECMA R&R)

- We have used the "law of large numbers" in determining the number of commuters π_{ij} because workers are assumed to be infinitesimally small.

- Plausible? Maybe not if the spatial unit i and j is small.

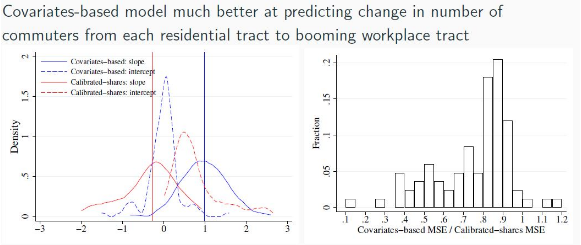
- For instance, New York has 2160 census blocks, so there are $2160 \times 2160 = 4665600$ pairs of workplace and residence.
  - But New York's workforce is just about 2.5 million
  - Many of π_{ij} are zero because of no observation. However, the law of large numbers implies that this is always strictly positive

Figure 1: Number of commuters between pairs of tracts in New York City



Dingel and Tintelnot (2023 ECMA R&R)

- As such, our observation of π_{ij} is actually full of sampling errors.
  - How much does such uncertainty affect our analysis based on QSE models?
- They compare two approaches for estimating commuting costs
  - Parameterize d_{ij} and use the gravity equation. This is the approach in the ARSW paper: $d_{ij} = e^{-\kappa t_{ij}}$
  - We can calculate the commuting costs d_{ij} that exactly match the observed data, and then use the exact-hat algebra for counterfactual analysis. This approach is taken in studies such as Heblich, Redding Sturm (2020 QJE)
- The first approach does good in prediction, the second does poorly
  - The paper proposes some remedies for the second approach, but honestly not very simple
  - My suggestion: Use the first approach for a spatially granular setting.



Taking stock

- We have seen Ahlfeldt, Redding, Sturm, Wold (2015), a seminal QSE model with commuting
  - Also a few other applications of such QSE models
- Gravity commuting equation is implied by these models
  - We can quantify the commuting cost by estimating it (given the Frechet dispersion parameter)
- ARSW highlights that agglomeration forces are crucial in explaining the division and reunification experiences of Berlin.
- We next dig a bit deeper into empirical analyses of agglomeration forces
  - How can we infer the strength of agglomeration forces?
  - What are the microfoundations of agglomeration forces?
- Also, note that the QSE models so far do not have heterogeneity in preferences or income
  - Introducing heterogeneity, the issue of geographical sorting arises. We also discuss them later.