Chapter 4
Quantitative Spatial Economics

In this chapter we introduce a new breed of urban economics models, labelled by Redding and Rossi-Hansberg (2017) as quantitative spatial economics models. The models we have analyzed so far are not very amenable for estimation, starting from their assumptions and their mathematical complexity. Moreover, at least in our study of the empirics of agglomeration, there seems to be a disconnect between the models and the data.

Redding and Rossi-Hansberg (2017) highlight several advantages of this new breed of models, which were developed in international trade and then adapted to spatial economics:

- They easily accommodate many regions and a rich mobility structure
- They can rationalize data as equilibria of the model
- They are usually exactly identified
- They can be used to carry out counterfactual decomposition and welfare analyses

We will start by tackling the question of the extent of agglomeration forces once again. Ahlfeldt et al. (2015) approximate the ideal experiment of density, and use the variation from the experiment to estimate the strength and extent of agglomeration within an urban model. To do this, they use a singular accident of nature: The rise and fall of the Berlin Wall.

To call this paper anything less than a classic would be an understatement. Recently, in 2018, this paper won the Frisch Medal for the best paper published in Econometrica in the last five years. With this rather grandious introduction, let us dig in the basics of this paper.

Ahlfeldt et al. (2015) attack the same question we tackled in the last chapter. How strong are agglomeration and dispersion forces in a city? We already know that this is a difficult question, and they point out two reasons why this is the case. First is the endogeneity: it is hard to disentangle the strength of agglomeration forces from simple differences in location fundamentals. For this you would need exogenous variation in agglomeration. Second, it is hard to bring the traditional urban models to data on cities.

Ahlfeldt et al. (2015) find a fantastic source of exogenous variation to identify the effects of agglomeration: the Berlin Wall. A little historical background is in order. In July of 1945, after the Second World War, Berlin was split into three sectors, for the Americans, the British and the Soviet. A remainder French sector would lately be created from the British sector. These sectors were intended to be approximately equal in population. Although the city was intended to be governed jointly, relationships between the Western and Eastern factions deteriorated and in 1961 East Germany built the Berlin Wall.

Figure 4.1 shows a map of Berlin in 1936 along with the boundary of the Berlin Wall. The wall cut off West Berlin from the CBD, cutting through the subway and rail lines.

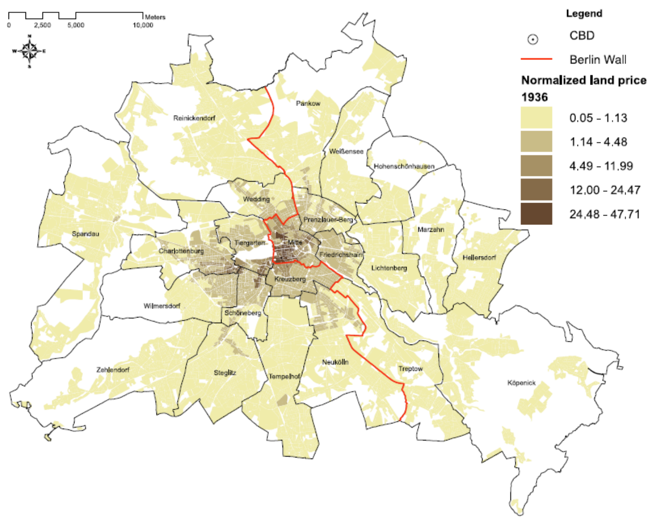

FIGURE 1.—Land prices in Berlin in 1936.

Fig. 4.1 Ahlfeldt et al. (2015) Figure 1

Towards 1989, with the beginning of the fall of the Soviet Union, the Berlin Wall fell on November 9th, 1989, and Germany was reunified in October of 1990.

Ahlfeldt et al. (2015) have data on area, employment, land prices and distances for city blocks in Berlin, for 1936 (pre Berlin-Wall), 1986 (Berlin Wall) and 2006 (Post Berlin Wall). Once the wall goes down, you can think that the potential function jumps in the center of the city.

So how does equilibrium in this city change when the Berlin Wall rises and falls? The issue is that while we can know a lot about equilibrium, we can not know it all from land prices and employment data, because we lack commuting data and wages. Consider the simple example in table 4.

Here we have data on residence and workplace. We do not observe commuting, or amenities. The issue here is that there are many commuting patterns consistent

	Block 1	Block 2
Workers	1	2
Residents	2	1
Workplace amenities	B_1	B_2
Residential amenities	A_1	A_2

with equilibrium, each of which is going to be consistent with a certain pattern of amenities. For example, an equilibrium could feature commuting only from block 1 to block 2, wich would be consistent with high workplace amenities in 2 and high residential amenities in 1. Another equilibrium could have commuting going in both directions. We will model this in a particular way to narrow down these possibilities.

4.1 Model

This section follows Ahlfeldt et al. (2015) closely. If you want a simplified version along with an economic geography model and an overview of the versatility of these models, see Redding and Rossi-Hansberg (2017).

4.1.1 Basic assumptions

We consider an open city. In the agricultural area utility is $\bar{U}$. The city is in discrete space, and there are S blocks indexed by $i = 1, \ldots, S$. There is L_i floor space in every block, and these blocks can be assigned to residential or commercial usage, much as in Fujita and Ogawa (1982). θ_i is the endogenous fraction of every block that is dedicated to commercial use. The city produces a single *numéraire* good that is costlessly traded. The blocks are connected, and there are H (endogenous) workers that can freely move within the city.

Workers get utility from living in a particular block i and working in a particular block j. The utility worker o gets from living in i and working in j is:

$$C_{ijo} = \frac{B_i z_{ijo}}{d_{ij}} \left(\frac{c_{ijo}}{\beta} \right)^{\beta} \left(\frac{\ell_{ijo}}{1-\beta} \right)^{1-\beta} \tag{4.1}$$

.

Here, B_i are residential amenities, d_{ij} is the commuting cost (in terms of utility) from i to j, c is consumption, ℓ is land and β governs the housing share of expenditure. z_{ijo} is a preference shock, that captures the fact that some workers may prefer to live or work in some locations. This is where all the magic happens, so we will explain this in detail.

4.1.2 Fréchet "magic"

We are going to assume a particular extreme-value distribution for this preference shock parameter. This distribution is going to be a Fréchet distribution:

$$F(z_{ijo}) = e^{-T_i E_j z_{ijo}^{-\varepsilon}} \tag{4.2}$$

T_i is the average utility from living in block i, E_j is the average utility from working in block j, and ε is a parameter that measures the dispersion of the distribution.

Eaton and Kortum (2002) show that if worker productivity is the product of inventions that happen over time drawn from a Pareto distribution, then the worker productivity for the most productive technology has a Fréchet distribution. They use this to model worker productivities in a Ricardian model of trade. We are going to use it in a Ricardian model of commuting and migration.

Mathematically, there are two reasons why the Fréchet distribution is useful in this setting. First, the Fréchet is constant under linear transformations. That is, if z is Fréchet, e.g $F(z) = exp(-Tz^{-\theta})$, then $y = kz + b$ is Fréchet, with $F(y) = exp(-Tk^{\varepsilon}(Y-b)^{-\varepsilon})$.

Second, the maximum of Fréchets is Fréchet distributed, and there are simple expressions for the fractions when a Fréchet variable is larger than other. Consider two Fréchet variables z_1 and z_2, with shape parameters T_1 and T_2 and the same dispersion parameter ε. Then

$$\begin{aligned}
Pr(z_1 < z_2) &= \int_0^{\infty} F(z_1) dF(z_2) \\
&= \int_0^{\infty} e^{-T_1 z^{-\varepsilon}} e^{-T_2 z^{-\varepsilon}} T_2 \varepsilon z^{-\varepsilon - 1} dz \\
&= \frac{T_2}{T_1 + T_2} \int_0^{\infty} e^{-(T_1+T_2)z^{-\varepsilon}} \varepsilon (T_1 + T_2) dz \\
&= \frac{T_1}{T_1 + T_2}.
\end{aligned}$$

4.1.3 Household choice

We now turn to solving the household's problem. We adopt the same strategy used in the previous chapters: first, we solve for the optimal consumption and land use given a location, then we examine location choice. If wages in location i are w_i and rents are Q_i, then the indirect utility of living in i and working in j from the first part of the optimal choice is:

$$u_{ijo} = \frac{z_{ijo} B_i w_j Q_i^{\beta - 1}}{d_{ij}}. \tag{4.3}$$

We model the commuting costs as a function of travel times:

$$d_{ij} = e^{\kappa\tau_{ij}} \tag{4.4}$$

Now we'll use our two Fréchet distribution facts. Since z_{ijo} is Fréchet, then u_{ijo} is Fréchet, with shape parameter

$$\Phi_{ij} \equiv T_i E_j \left(d_{ij} Q_i^{1-\beta}\right)^{-\varepsilon} (B_i w_j)^{\varepsilon}. \tag{4.5}$$

The individuals will choose to live in i and work in j with some probability, which is the probability of $u_{ijo} = max_{r,s}\, u_{rso}$. From our second Fréchet useful property, this probability is

$$\pi_{ij} = \frac{\Phi_{ij}}{\sum_{r=1}^{S}\sum_{s=1}^{S}\Phi_{ij}} \equiv \frac{\Phi_{ij}}{\Phi}. \tag{4.6}$$

This completely characterizes the solution of the household's problem and the spatial distribution of households. We can now find the fractions of people who reside in i, π_{Ri} and the fractions of people who work in j, π_{Mj}:

$$\pi_{Ri} = \sum_j \pi_{ij}; \pi_{Mj} = \sum_i \pi_{ij}. \tag{4.7}$$

Notice that the probability of living in a place increases with residential amenities and decreases with rent. The probability of working in a place increases with wages. ε is important in determining these elasticities. Higher commuting costs reduce the probability of choosing a particular pair.

The only remaining condition is that everything adds up. The probability of working on a particular place j conditional on living in i is

$$\pi_{ij|i} = \frac{\pi_{ij}}{\sum_s \pi_{is}} = \frac{E_j (w_j/d_j)^{\varepsilon}}{\sum_s E_s (w_s/d_s)^{\varepsilon}}. \tag{4.8}$$

Adding up then requires that the workers in each place j, H_{Mj}, equal the sum of the residents in each place, H_{Ri}, times the probability that they work in j:

$$H_{Mj} = \sum_i \pi_{ij|i} H_{Ri} \tag{4.9}$$

4.1.4 Firm Choice

Firms are assumed to have *CRS* Cobb-Douglas production functions over workers H_{Mj} and commercially-used land L_{Mj}, for which they pay wages w_j and rents q_j:
$y_j = A_j H_{Mj}^{\alpha} L_{Mj}^{1-\alpha}$

From the first-order conditions of the firm's problem:

$$H_{Mj} = \left(\frac{\alpha A_j}{w_j}\right)^{\frac{1}{1-\alpha}} \tag{4.10}$$

$$q_j = (1-\alpha)\left(\frac{\alpha}{w_j}\right)^{\frac{\alpha}{1-\alpha}} A_j^{\frac{1}{1-\alpha}} \tag{4.11}$$

The first condition determines equilibrium employment in each location, and the second condition determines equilibrium commercial land prices.

4.1.5 Land Markets

Here we simply assume that floor space is supplied proportionally to land K at each location

$$L_i = \varphi_i K_i^{1-\mu} \tag{4.12}$$

Condition (4.11) determines commercial land demand. For residential land demand, use the fact that households spend a constant fraction of their income on commercial land, to get the expected demand of land from each individual l_i:

$$E[l_i]H_{Ri} = (1-\beta)\frac{E[w_s|i]H_{Ri}}{Q_i} \tag{4.13}$$

Together with the supply function 4.12 these determine rents in equilibrium q_i and Q_i. If a block has both residents and firms, $Q_i = q_i$. In all other cases the block is either completely populated by residents or firms.

4.2 Equilibrium and agglomeration

An equilibrium here requires:

- Firms maximize profits and choose optimal locations
- Households maximize utility and choose optimal locations
- Land markets clear
- Labor markets clear

The endogenous variables are population H, the fractions living and working in each place $\pi_{\mathbf{R}}$ and $\pi_{\mathbf{M}}$, land rents $\mathbf{Q}, \mathbf{q}$, wages $\mathbf{w}$ and land use θ. Ahlfeldt et al. show that there is a unique equilibrium in this model. This contrast with the non-uniqueness of equilibria in Fujita and Ogawa (1982), but we have not introduced agglomeration forces yet.

To model agglomeration, Ahlfeldt et al. (2015) make residential and workplace amenities depend on the density of employment and residency:

$$A_j = a_j \left[\sum_s e^{-\delta \tau_{js}} \left(\frac{H_{Ms}}{K_s} \right) \right]^{\lambda} \tag{4.14}$$

$$B_i = b_i \left[\sum_r e^{-\rho \tau_{ir}} \left(\frac{H_{Rr}}{K_r} \right) \right]^{\eta}$$

Here, τ_{ij} is the travel time from i to j. λ and η govern the strength of agglomeration forces (β in Fujita and Ogawa (1982)) and δ and ρ govern their spatial decay (α in Fujita and Ogawa (1982)). In presence of agglomeration forces, the equilibrium is, as expected, not unique.

4.3 Rationalizing data

Now we are in good shape to take our model to data and calculate equilibriums. There is an issue, however. In order to determine the endogenous variables, we need values of the residential amenities, workplace amenities and the density of development. These are unfortunately, unobserved. However, we can assume the data is an equilibrium of the model, and find values of these unobserved variables such that the data is in fact an equilibrium. This is only possible if there is an unique mapping from the data to the unobserved variables. Turns out that this is the case. Given values of the data

- Rent Q $= max(\mathbf{Q}, \mathbf{q})$
- Residents and workers $\mathbf{H_R}, \mathbf{H_M}$
- Land $\mathbf{K}$
- Travel times τ

and values of the parameters: the labor share α, the housing share of expenditure β, the elasticity of land development μ, the dispersion parameter ε, the commuting cost parameter κ, and the agglomeration parameters $\lambda, \delta, \eta, \rho$, there are unique (normalized) A_i, B_i, φ_i that are consistent with the data being an equilibrium of the model. The normalization is necessary because some of these unobserved parameters appear isomorphically in the model.

To fix ideas, we will look at an example (from Matthew Turner) of how all this works. Consider the simplified model in table 4.1. Forget about commercial land rents, and about floor space development. Assume that all places have the same area L, and that $d_{12} = d_{21}$ and $d_{ii} = 1, i = 1, 2$. Commuting is $1/2$ from block 2 to block 1, and we assume there is no cross-commuting from block 1 to block 2.

We observe rents, land areas, and the distribution of workers and residents. Our goal is to get the unobserved amenities and productivities. Notice that we don't observe wages either.

Assume a simpler structure for utility:

$$u_{ijo} = b_i + \frac{w_j z_{ijo}}{d_{ij}} - Q_i, \tag{4.15}$$

Variable	Block 1	Block 2
H_R	1	2
H_M	3/2	3/2
Workplace productivity	a_1	a_2
Residential amenities	b_1	b_2
Rents (observed)	Q_1	Q_2

Table 4.1 Simplified model

where z is Fréchet (T,ε). Also assume a Cobb-Douglas structure for production:

$$Y_i = a_i L^{\alpha} H_{Mi}^{1-\alpha}. \tag{4.16}$$

From the structure of commuting we can recover the conditional probabilities of working in each place.

- $\pi_{11|1} = 1$ since every resident from 1 works in 1
- $\pi_{12|1} = 0$ since we assumed no commuting from 1 to 2
- $\pi_{21|2} = 1/4 = \frac{1/2}{2}$ which is the fraction of block 2 residents that work in block 1
- $\pi_{22|2} = 3/4$

Now to recover wages and productivities. Assuming production is Cobb-Douglas, wages must satisfy

$$w_i = (1-\alpha)\frac{Y_i}{H_{Mi}} = (1-\alpha)\frac{a_i L^{\alpha} H_{Mi}^{1-\alpha}}{H_{Mi}}. \tag{4.17}$$

The only unobserved component here is a_i. From the properties of the Fréchet distribution, we know that the fractions must satisfy:

$$\pi_{11|1} = \frac{T(w_1/d_{11})^{\varepsilon}}{T(w_1/d_{11})^{\varepsilon} + T(w_2/d_{12})^{\varepsilon}}$$
$$\pi_{12|1} = \frac{T(w_2/d_{12})^{\varepsilon}}{T(w_1/d_{11})^{\varepsilon} + T(w_2/d_{12})^{\varepsilon}}$$
$$\pi_{21|2} = \frac{T(w_1/d_{21})^{\varepsilon}}{T(w_1/d_{21})^{\varepsilon} + T(w_2/d_{22})^{\varepsilon}}$$
$$\pi_{22|2} = \frac{T(w_2/d_{22})^{\varepsilon}}{T(w_1/d_{21})^{\varepsilon} + T(w_2/d_{22})^{\varepsilon}}$$

Provided this system of 4 equations has the proper rank, you can solve here for a_1 and a_2. You can be even greedier and try to get α and ε from here, but this is not necessary here as we assumed those were observed. With a_i you also have wages w_i.

With wages in hand, we can solve for expected income $v_i = E\left[\frac{w_i z_{ijo}}{d_{ij}}\right]$.

With expected income in hand, spatial equilibrium implies that utility must be equalized across locations.

$$b_1 + v_1 - Q_1 = b_2 + v_2 - Q_2 \tag{4.18}$$

If we normalize $b_1 = 1$, we can get b_2 from here, and we are done!

4.4 Reduced-form evidence

Before digging into structural estimation of the model, Ahlfeldt et al. (2015) provide some reduced-form evidence of the effects of division and reunification of Berlin. Tables 4.2 provides difference in difference estimates of the effect of division and reunification on land prices and employment in West Berlin, by distance to the CBD. The estimating equation is

$$\Delta ln(O_i) = \alpha + \sum_{k=1}^{K} 1_k \beta_k + lnM_i \gamma + u_i \tag{4.19}$$

Which is the first-difference version of differences-in-difference. k indexes grids of distance to the CBD. Note that any block specific controls in this specification translate to trends that vary by the level of this control in the levels specification.

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%.

Table 4.2 Ahlfeldt et al. (2015) CITE Table 1

The coefficient in the first row of column (1) implies that floor space prices fell by around 55% in the blocks closest to the CBD. Note that the effects become smaller as the distance to the CBD increases, and that they seem robust across specifications. Columns (6) to (9) show decreases in employment. In the paper, they also report results for reunification, which confirm the results here, having the opposite signs and slightly smaller magnitudes.

4.5 Structural Estimation

The last section of Ahlfeldt et al. (2015) estimates the parameters $\lambda, \delta, \eta, \rho$ that govern the strength of agglomeration forces. They use the exogenous variation arising from the change in density from the rise and fall of the berlin wall. The basic idea behind their estimation is the following: any change over time in the location fundamentals a_j and b_i from (4.14) should not be correlated with the changes in density. Changes in density are measured by the distance to the CBD. I will ignore normalization of these fundamentals and write down the identifying conditions as:

$$E\left[1_k \times \Delta ln(a_{it})\right] = 0,\ k = 1, \ldots, K \tag{4.20}$$

$$E\left[1_k \times \Delta ln(b_{it})\right] = 0\ k = 1, \ldots, K \tag{4.21}$$

How to estimate the model now? We use this identifying condition in an analogous way to the identifying conditions in OLS ($X'e = 0$) or IV ($Z'e = 0$). The issue is that while the error term is easy to obtain in a linear model, it is not easy to obtain them here. But it was easy to obtain them in our simpler model in section 4.3.

The estimation proceeds in two loops, it what is called a “nested-fixed point algorithm”. For details on this, see Holmes and Sieg (2015)

1. Fix a starting value of the parameters $\lambda, \delta, \eta, \rho$.
2. For these values, find the values of a_i, b_i that rationalize the data as being an equilibrium of the model. This is the “inner loop”.
3. With estimates of the location fundamentals $\hat{a}_i, \hat{b}_i$ in hand, use the identification condition in (4.20) to obtain estimates $\hat{\lambda}, \hat{\delta}, \hat{\eta}, \hat{\rho}$
4. Find new values of the location fundamentals for the new values of the parameters, $\hat{a}_i', \hat{b}_i'$. This is the “outer loop”
5. Iterate until convergence in $\hat{\lambda}, \hat{\delta}, \hat{\eta}, \hat{\rho}$

Severen (2018) provides of identification strategies for these quantitative spatial economics models. To do the outer loop steps, we will use the generalized method of moments (GMM) to find the estimates.

4.5.1 GMM

This section follows Hayashi (2000). Recall OLS estimation with L variables. The identifying assumption to estimate $\beta_{L\times 1}$ in

$$\mathbf{Y_{n\times 1}} = \mathbf{X_{n\times L}}\beta_{\mathbf{L\times 1}} + \mathbf{e_{n\times 1}}$$

is

$$\mathbf{X'e} = \mathbf{0}.$$

Define the moment condition as

$$\begin{aligned} m(\beta) &= \frac{1}{n}\sum_{i=1}^{n} \mathbf{x_i} \cdot e_i \qquad (4.22) \\ &= \frac{1}{n}\sum_{i=1}^{n} \mathbf{x_i} \cdot (y_i - \mathbf{x_i}\beta) \end{aligned}$$

This is the sample analog of the identifying condition. We would like to set this to 0. This is a system of L equations with L unknowns. If you solve this system, we will get the usual OLS solution for β.

The same logic can be applied to linear instrumental variables estimation. In that case, the identifying condition is the exclusion restriction on the instruments $\mathbf{Z}$. Suppose that we have the same number of instruments as endogenous variables, so $\mathbf{Z}$ is $\mathbf{Z_{n\times L}}$. The identifying condition is $\mathbf{Z'e} = \mathbf{0}$. The moment condition now is:

$$\begin{aligned} m(\beta) &= \frac{1}{n}\sum_{i=1}^{n} \mathbf{z_i} \cdot e_i \qquad (4.23) \\ &= \frac{1}{n}\sum_{i=1}^{n} \mathbf{z_i} \cdot (y_i - \mathbf{x_i}\beta) \end{aligned}$$

And by setting this equal to $\mathbf{0}$ and solving, we will get the usual IV estimator.

Now consider a setting where you have more instruments that endogenous variables, e.g. Z is $n \times M$, X is $n \times L$ and $M > L$. In that case, we can no longer set $m(\beta)$ to $\mathbf{0}$. But we can set it as close to $\mathbf{0}$ as possible.

Consider minimizing the following function with respect to β.

$$J(\beta) = N[m(\beta) - \mathbf{0}]\mathbf{W}[m(\beta) - \mathbf{0}]' \qquad (4.24)$$

What this is doing is choosing the value of β for which the identifying condition is as close to being satisfied as possible in the sample. We minimize a quadratic form in $m(\beta)$, weighting by $\mathbf{W}$.

If you recall the previous chapter, this looks awfully familiar to the Ciccone and Hall setting. Their estimator minimized

$$J(\beta) = Z'[y - f(\beta)]W[y - f(\beta)]'Z \tag{4.25}$$

which as an analogous problem to the previous one. In this case, the identifying condition was $Z'[y - f(\beta)] = 0$. It does not matter that $f(\beta)$ is not linear, as we can still minimize $J(\beta)$ using numerical methods.

Back to Ahlfeldt et al. (2015), they implement this estimator using the moment conditions from (4.20), plus additional moment conditions based on the variance of travel times. In general, you need an instrument for every parameter you want to identify.

4.6 Structural Estimation Results

Table 4.3 shows results of the GMM estimation. The parameters here can be compared to the density elasticity parameters we have seen in previous studies. The parameter λ is the elasticity of productivity to density. It is around 7 %, close to the preferred parameter of 4 % of Ahlfeldt and Pietrostefani (2019). On the other hand, residential externalities seem to be larger: doubling residential density implies an increase in residential amenities of about 15 % according to the estimate of η. The parameters δ and ρ measure how quickly the agglomeration externalities dissipate in space.

TABLE V

GENERALIZED METHOD OF MOMENTS (GMM) ESTIMATION RESULTS[a]

	(1) Division Efficient GMM	(2) Reunification Efficient GMM	(3) Division and Reunification Efficient GMM
Commuting Travel Time Elasticity ($\kappa\varepsilon$)	0.0951***	0.1011***	0.0987***
	(0.0016)	(0.0016)	(0.0016)
Commuting Heterogeneity (ε)	6.6190***	6.7620***	6.6941***
	(0.0939)	(0.1005)	(0.0934)
Productivity Elasticity (λ)	0.0793***	0.0496***	0.0710***
	(0.0064)	(0.0079)	(0.0054)
Productivity Decay (δ)	0.3585***	0.9246***	0.3617***
	(0.1030)	(0.3525)	(0.0782)
Residential Elasticity (η)	0.1548***	0.0757**	0.1553***
	(0.0092)	(0.0313)	(0.0083)
Residential Decay (ρ)	0.9094***	0.5531	0.7595***
	(0.2968)	(0.3979)	(0.1741)

[a]Generalized Method of Moments (GMM) estimates. Heteroscedasticity and Autocorrelation Consistent (HAC) standard errors in parentheses (Conley (1999)). * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 4.3 Ahlfeldt et al. (2015) Table 5

To better understand the spatial decay of agglomeration forces, table 4.4 shows the percentage of the production and residential externalities agents get as they move away from density. These results show that agglomeration externalities are very localized, confirming the results in Arzaghi and Henderson (CITE). After 20 minutes of travel time, production externalities dissipate. Residential externalities dissipate even faster.

TABLE VI

EXTERNALITIES AND COMMUTING COSTS[a]

	(1) Production Externalities $(1 \times e^{-\delta\tau})$	(2) Residential Externalities $(1 \times e^{-\rho\tau})$	(3) Utility After Commuting $(1 \times e^{-\kappa\tau})$
0 minutes	1.000	1.000	1.000
1 minute	0.696	0.468	0.985
2 minutes	0.485	0.219	0.971
3 minutes	0.338	0.102	0.957
5 minutes	0.164	0.022	0.929
7 minutes	0.079	0.005	0.902
10 minutes	0.027	0.001	0.863
15 minutes	0.004	0.000	0.802
20 minutes	0.001	0.000	0.745
30 minutes	0.000	0.000	0.642

[a]Proportional reduction in production and residential externalities with travel time and proportional reduction in utility from commuting with travel time. Travel time is measured in minutes. Results are based on the pooled efficient GMM parameter estimates: $\delta = 0.3617$, $\rho = 0.7595$, $\kappa = 0.0148$.

Table 4.4 Ahlfeldt et al. (2015) Table 6

4.7 Summary

We have introduced state-of-the-art quantitative methods of spatial economics. The combination of transparent reduced-form evidence –from an exogenous source of variation– and structural estimation makes this paper unique. Along the way, we learned about modelling strategies and the generalized method of moments.

The results of Ahlfeldt et al. (2015) are one of the most transparent estimates of the effects of density in the literature. There are many other applications of quantitative spatial models. The discipline is now extending them to analyze many different problems. Here are some examples:

- Monte et al. (2018) analyze how employment responds to local shocks in productivity, accounting for the spatial connectedness across areas through migration and commuting. They conclude that areas that are more open to commuting receive larger increases in employment.
- Severen (2018) examines the effects on commuting of the LA metro rail. Along the way, he surveys estimation methods in quantitative spatial economics.
- Tsivanidis (2018) estimates the welfare and redistributive effects of building a BRT in Colombia.
- Pérez Pérez (2017) extends quantitative spatial models to account for unemployment and uses them to analyze the effects of local minimum wages in spatial equilibrium.
- Pérez-Cervantes (2016) uses these models to show how consumers use commuting to insure against local shocks in Mexico.

In the next section we will move from inside the city to outside the city, and consider the allocation of economic activity across regions.