Reading Notes on HeblichReddingSturm2020

Li Shen and Zeyi Qian
Instructed by Junfu Zhang

Fall 2021

1 Description of the paper

The paper "THE MAKING OF THE MODERN METROPOLIS: EVIDENCE FROM LONDON" was written by STEPHAN HEBLICH, STEPHEN J. REDDING, and DANIEL M. STURM. It was published by The Quarterly Journal of Economics at 2020. The authors use newly constructed spatially disaggregated data for London from 1801 to 1921 to show that the invention of the steam railway led to the first large-scale separation of workplace and residence.

2 Introduction

- Motivation: Huge commuting flows vs. Little knowledge about it's role in sustaining dense concentrations of economic activity;

- Historical events: Steam railways created in the mid-19th century;

- Evidence for the positive effect of residential separation on agglomeration:
  ∗Data set for Greater London during 1801-1921;
  ∗Quantitative urban model.

- Contribution:
  Empirical estimation: Agglomeration effect: change of the relationship between travel time and distance as IV;
  Theoretical models: First, it proves that the urban models established by this paper all have the same land and commuter market clearing conditions, so reliable results can be obtained in such models, and a series of factors that have not been observed to determine the spatial distribution of the economy can be controlled. Second, a new structural estimation method is developed for such urban models that uses bilateral commuter flows from the base year (1921) and performs a comparative static analysis from that base year. This paper shows that the estimation procedure can be used to recover unobserved historical employment and commuting data from bilateral commuting data in the base year as well as historical data on population and land rents. Thirdly, paper shows that the urban model can quantitatively explain the spatial restructuring of economic activities in Greater London with the invention of steam railways.

3 Background

- Three geographical definitions of boundaries:
  * Greater London as defined by the Modern boundaries of the Greater London Authority (GLA): 99 boroughs and 283 parishes
  * Historic County of London: 29 boroughs and 183 parishes;
  * The city of London: 111 parishes.

- BC 43:London has a long history of settlement

- In the early 19th century * Walking: about 3 miles per hour
  * Equestrian: about 6 miles per hour
  * Steam passenger rail: 21 miles per hour (rail development in London was carried out competitively by private companies)

- In the first half of the 19th century * Public goods were largely provided by local parishes and gatekeepers.

4 Data

- Bilateral Commuting: complete matrix of bilateral commuting flows between boroughs in Greater London * Closed commuting market: commuting flows between other parts of England and Wales and Greater London were small in 1921 * Workplace employment-employment by workplace for each borough: summing across rows in the matrix of bilateral commuting flows for Greater London (come to this place)

  * Residence employment-employment by residence for each borough: summing across columns in the matrix of bilateral commuting flows for Greater London (leave this place) * Employment participation rate for each borough in 1921: dividing residence employment by population

- Population by Residence * Population censuses of England and Wales from 1801-1891 was provided by the Cambridge Group for the History of Population and Social Structure (Cambridge Group) 1801, 1811, 1821, 1831, 1841, 1851, 1861, 1871, 1891(no 1881) * Population data from 1901-1921 stem from the Integrated Census Microdata Project (I-CeM). * Assuming that the ratio of residence employment to population is stable for a given borough over time * Use the 1921 value of this ratio and the historical population data to construct residence employment for earlier census years

- Rateable Values * Measure the value of floor space using rateable values, which correspond to the annual flow of rent for the use of land and buildings * Equal the price times the quantity of floor space in the model * These rateable values cover all categories of property (except Crown property occupied by the Crown,Places for divine worship,Concerns listed under No. III Schedule A) * Rateable values data: 1815,1843,1847,1852,1860,1881,1896,1905,1911,1921 * To create consistent spatial units over time, match parishes with the spatial units provided by the CGKO (Cambridge Group Kain Oliver) map * use area weights to create the same mappable units as for the population data-parish-level panel for the years 1815, 1843,

1848, 1852, 1860, 1881 and 1896 ∗ aggregate these parish data to the 1921 boroughs using area weights

- Transport Network ∗ GIS dataset on the transport network in Greater London over time using historical maps of the overground railway network, the underground railway network, and the omnibus and tram network, provided by the Cambridge Group for the History of Population and Social Structure, which based its digitization on Cobb (2003); ∗ Measure bilateral travel times by distinguishing four transport networks based on the historical average travel speeds by transport mode in London reported in London County Council (1907):
(i) overgroundrailways (21 mph)
(ii) underground railways (15 mph)
(iii)omnibuses and trams (6 mph)
(iv) walking (3 mph)
∗ Assumption (i)workers follow the least-cost path in terms of travel time (ii)workers incur a travel time cost of 3 minutes when changing between modes of transport and can only connect to the railway network at railway stations (iii)sets of points connected to each transport network at time t.

$$\tilde{s}_t{}^{OR}, \tilde{s}_t{}^{UR}, \tilde{s}_t{}^{OT}$$

(OR, UR, and OT indicate overground railways, underground railways, and omnibuses and trams,)

(iv) vector of assumed travel time weights for each transport network by

$$\delta = [1 \ \delta^{UR} \ \delta^{OT} \ \delta^{WA}]$$

(normalize the weight for overground railways to 1, and the superscript WA indicates walking)

(v) the bilateral travel times between boroughs n and i at time t as

$$d_{nit}^W = d_{nit}^W(\tilde{s}_t^{OR}, \tilde{s}_t^{UR}, \tilde{s}_t^{OT}, \delta)$$

where the superscript W indicates the weighting by transport mode

(vi) use an instrumental variable based on bilateral travel times in which walking is assumed to be the only mode of transport, so that bilateral travel times depend solely on straight-line distance. (vii) bilateral travel times in the absence of other modes of transport by d_{ni}^S, where the superscript S is a mnemonic for straight-line distance

- Historical Employment by Workplace and Commuting Data ∗ Compare model's historical predictions for workplace employment in Greater London with data on the "day population" of the City of London that are available from the day censuses of 1866, 1881, 1891, and 1911 ∗ Compare model's predictions for commuting distances with historical commuting data based on the residence addresses of the employees of

3

Henry Poole bespoke tailors, which has been located at the same workplace address in Savile Row in the City of Westminster since 1822

- Data for Other Cities * historical data on population, employment,and commuting distances for Berlin, Paris, Boston, Chicago, New York, and Philadelphia

5 Reduced-Form Evidence

5.1:City Size and Structure over Time

- Population is expressed as an index relative to its value in 1801 (such that 1801 = 1).

- City of London: sharp drop in 1851 (Figure II Left)

- Greater London: grew substantially (Figure II Right)

- City of London's residential (or "night") population: day population > night population (Figure III Left)

- Steep decline in its night population from 1851 onward coincides with a sharp rise in its day population (Figure III Left)

- As this expansion occurs, and undeveloped land becomes developed, the share of already-developed land in overall land values tends to fall (Figure III Right)

- In the years after 1851 when the City of London experiences the largest declines in its residential population, its rateable value share increases from 9% to 11%. (Figure III Right)

5.2:Difference-in-Differences Event-Study Specification

Baseline specification:

DID

$$logR_{jt} = \alpha_j + \sum_{\tau=-60}^{\tau=60} \beta_\tau (\mathbb{S}_j \times \mathbb{I}_{j\tau}) + (\mu_j \times Year_t) + d_t + u_{jt} \qquad (1)$$

DDD

$$logR_{jt} = \alpha_j + \sum_{\tau=-60}^{\tau=60} \beta_\tau (\mathbb{S}_j \times \mathbb{I}_{j\tau}) + \sum_{\tau=-30}^{\tau=30} \gamma_\tau (\mathbb{S}_j \times \mathbb{I}_{j\tau} \times \mathbb{I}_j^{Center}) + (\mu_j \times Year_t) + d_t + u_{jt} \qquad (2)$$

Results
DID: For central London and outer areas, the absolute value of the coefficient increases for 60 years after the railway is built. One of the reasons may be that the value of the connection to the railway network increases over time as the railway network expands
DDD: Reduced travel times increase the population of the suburbs of Greater London and decrease the population of the central areas.

6   Theoretical Framework

The authors develop a dynamic model, in which adjustment costs for investments in durable building capital introduce gradual adjustment in response to changes in the transport network. The authors introduce these investments in durable building capital following the standard approach to intertemporal saving and investment decisions in Obstfeld and Rogoff (1996). Time is discrete and is indexed by t.

- Locations
  The authors consider a city (Greater London) that is embedded within a wider economy (Great Britain). The economy as a whole consists of a discrete set of locations (M). Greater London comprises a subset of these locations $N \in M$.
  Locations are allowed to differ from one another in terms of their attractiveness for production and residence, as determined by productivity, amenities, the supply of floor space, and transport connections, where each of these location characteristics can evolve over time.

- Workers
  The economy consists of two types of infinitely-lived agents: workers and landlords. Workers are assumed not to have access to a saving or borrowing technology and are modeled as "hand to mouth," as in Kaplan and Violante (2014). Workers simultaneously choose their preferred residence n and workplace i given their idiosyncratic draws. And each worker chooses a residence-workplace pair either in Greater London or in the rest of the economy.

  Denote an exogenous continuous measure of workers in Great Britain by L_{Mt}.

  Denote the endogenous measure of workers who choose a residence-workplace pair in Greater London by L_{Nt}.

6.1:Preferences(worker)

The intertemporal preferences of worker ω living in residence n and working in workplace i at time t are additive separable and isoelastic:

$$U_{nit}(\omega) = \sum_{s=t}^{\infty} \rho^{s-t} \frac{U_{nit}(\omega)^{1-\frac{1}{\zeta}}}{1 - \frac{1}{\zeta}} \qquad (F1)$$

where, $\zeta > 0$ is the intertemporal elasticity of substitution;
ρ is the subjective rate of time discount;
$U_{nit}(\omega)$ is the worker's instantaneous utility function.

The instantaneous utility for worker ω residing in location n and working in location i is :(in Cobb-Douglas functional form)

$$U_{ni}(\omega) = \frac{B_{ni}b_{ni}(\omega)}{\kappa_{ni}} \left(\frac{C_{ni}(\omega)}{\alpha} \right)^{\alpha} \left(\frac{H_{ni}(\omega)}{1-\alpha} \right)^{1-\alpha} \qquad (F2)$$

where,the authors suppress the time subscript from now on, except where important;
$0 < \alpha < 1$;

B_{ni} captures amenities from the bilateral commute from residence n to workplace i that are common across all workers;

$b_{ni}(\omega)$ is an idiosyncratic amenity draw that captures all the idiosyncratic factors that can cause an individual to live and work in particular locations in the city;

κ_{ni} is an iceberg commuting cost;

$C_{ni}(\omega)$ is consumption of a single tradeable homogeneous final good;

$H_{ni}(\omega)$ is consumption of residential floor space.

The utility maximization problem of workers is:

$$\max_{C_{ni},H_{ni}} U_{ni}(\omega)$$

$$subject\ to$$

$$P_n C_{ni} + Q_n H_{ni} = w_i$$

where, P_n is the price index for consumption goods, which may include both tradeable and nontradeable consumption goods;

Q_n is the price of residential floor space;

w_i is the wage.

$\star$ Note that without access to a savings or borrowing technology, workers choose their residence, workplace, consumption of the final and residential floor space use each period to maximize their instantaneous utility.

Now, the Lagrange function is :

$$\mathcal{L} = \frac{B_{ni}b_{ni}(\omega)}{\kappa_{ni}} \left(\frac{C_{ni}(\omega)}{\alpha}\right)^{\alpha} \left(\frac{H_{ni}(\omega)}{1-\alpha}\right)^{1-\alpha} + \lambda(w_i - P_n C_{ni} - Q_n H_{ni})$$

First Order Conditions:

$$\frac{\partial L}{\partial C_{ni}} = \frac{B_{ni}b_{ni}(\omega)}{\kappa_{ni}}\alpha\left(\frac{C_{ni}}{\alpha}\right)^{\alpha-1}\frac{1}{\alpha}\left(\frac{H_{ni}(\omega)}{1-\alpha}\right)^{1-\alpha} - \lambda P_n = 0$$

$$\frac{\partial L}{\partial H_{ni}} = \frac{B_{ni}b_{ni}(\omega)}{\kappa_{ni}}\left(\frac{C_{ni}}{\alpha}\right)^{\alpha}(1-\alpha)\left(\frac{H_{ni}(\omega)}{1-\alpha}\right)^{-\alpha}\frac{1}{1-\alpha} - \lambda Q_n = 0$$

$$\frac{\partial L}{\partial \lambda} = w_i - P_n C_{ni} - Q_n H_{ni} = 0 \qquad (F2.1)$$

rearrange them to get:

$$\frac{B_{ni}b_{ni}(\omega)}{\kappa_{ni}}\left(\frac{C_{ni}}{\alpha}\right)^{\alpha-1}\left(\frac{H_{ni}(\omega)}{1-\alpha}\right)^{1-\alpha} = \lambda P_n \qquad (F2.2)$$

$$\frac{B_{ni}b_{ni}(\omega)}{\kappa_{ni}}\left(\frac{C_{ni}}{\alpha}\right)^{\alpha}\left(\frac{H_{ni}(\omega)}{1-\alpha}\right)^{-\alpha} = \lambda Q_n \qquad (F2.3)$$

divided $(F2.2)$ by $(F2.3)$ to get:

$$\frac{C_{ni}^{\alpha-1}}{C_{ni}^{\alpha}}\frac{\alpha^{1-\alpha}}{\alpha^{-\alpha}}\frac{H_{ni}^{1-\alpha}}{H_{ni}^{-\alpha}}\frac{(1-\alpha)^{\alpha-1}}{(1-\alpha)^{\alpha}}=\frac{P_n}{Q_n}$$

which can be simplified as:

$$\frac{\alpha}{(1-\alpha)}\frac{H_{ni}}{C_{ni}}=\frac{P_n}{Q_n}\qquad(F2.4)$$

$(F2.4)$ can be rewritten as:

$$H_{ni}=\frac{P_n}{Q_n}\frac{1-\alpha}{\alpha}C_{ni}\qquad(F2.5)$$

substituting $(F2.5)$ to $(F2.1)$ can get:

$$w_i-P_nC_{ni}-Q_n\left[\frac{P_n}{Q_n}\frac{1-\alpha}{\alpha}C_{ni}\right]=0$$

rearrange them to get:

$$C_{ni}=\frac{\alpha w_i}{P_n}\qquad(F2.6)$$

substituting $(F2.6)$ to $(F2.5)$ can get:

$$H_{ni}=\frac{(1-\alpha)w_i}{Q_n}\qquad(F2.7)$$

Substituting $(F2.6)$ and $(F2.7)$ to $(F2)$

$$U_{ni}(\omega)=\frac{B_{ni}b_{ni}(\omega)}{k_{ni}}\left(\frac{\frac{\alpha w_i}{P_n}}{\alpha}\right)^{\alpha}\left(\frac{\frac{(1-\alpha)w_i}{Q_n}}{1-\alpha}\right)^{1-\alpha}$$

$$=\frac{B_{ni}b_{ni}(\omega)}{\kappa_{ni}}\frac{w_i^{\alpha}}{P_n^{\alpha}}\frac{w_i^{1-\alpha}}{Q_n^{1-\alpha}}$$

$$=\frac{B_{ni}b_{ni}(\omega)w_i}{\kappa_{ni}P_n^{\alpha}Q_n^{1-\alpha}}\qquad(3)$$

which is indirect utility for a worker ω residing in n and working in i, and $0<\alpha<1$.

6.1.1:Distribution of Utility

Assume that idiosyncratic amenities $b_{ni}(\omega)$ are drawn from an independent extreme value (Fréchet) distribution for each residence-workplace pair and each worker:

$$G(b)=e^{-b^{-\epsilon}}\qquad(4)$$

where, $\epsilon > 1$ is shape parameter, which controls the sensitivity of worker location decisions to economic variables;and the scale parameter is normalized to 1.

From the indirect utility function in equation (3), the utility(u) also has Fréchet distribution. The *distribution of utility* for residence n and workplace i is:

$$G_{ni}(u) = Pr[u_{ni} \le u]$$

$$= Pr[\frac{B_{ni}b_{ni}(\omega)w_i}{\kappa_{ni}P_n^\alpha Q_n^{1-\alpha}} \le u]$$

$$= Pr[\frac{B_{ni}w_i}{\kappa_{ni}P_n^\alpha Q_n^{1-\alpha}}\frac{1}{u} \le \frac{1}{b_{ni}(\omega)}]$$

$$= Pr[\frac{\kappa_{ni}P_n^\alpha Q_n^{1-\alpha}}{B_{ni}w_i}u \ge b_{ni}(\omega)]$$

$$= e^{-\Psi_{ni}u^{-\epsilon}},$$

where $\Psi_{ni} = (B_{ni}w_i)^\epsilon(\kappa_{ni}P_n^\alpha Q_n^{1-\alpha})^{-\epsilon}$ is the location paremeter; $n \in N, i \in N$.

Since the maximum of a sequence of Fréchet distributed random variables is itself Fréchet distributed, the distribution of utility across all residences k and all workplaces l is:

$$G(u) = Pr[max(u_{ni}) \le u]$$

$$= \prod_k \prod_l e^{-\Psi_{kl}u^{-\epsilon}}$$

$$= e^{\sum_k \sum_l -\Psi_{kl}u^{-\epsilon}}$$

$$= e^{-\Psi_M u^{-\epsilon}}$$

where, $\Psi_M = \sum_k \sum_l \Psi_{kl}, k \in M \ and \ l \in M$

Given this Fréchet distribution for utility,the expected utility is:

$$E(u) = \int_0^\infty uG'(u)du$$

$$= \int_0^\infty u(-\Psi_M)(-\epsilon u^{-\epsilon-1})e^{-\Psi_M u^{-\epsilon}}du$$

$$= \int_0^\infty \epsilon\Psi_M u^{-\epsilon}e^{-\Psi_M u^{-\epsilon}}du$$

define $y = \Psi_M u^{-\epsilon}$, so $dy = -\epsilon\Psi_M u^{-\epsilon-1}du = -\epsilon\Psi_M u^{-\epsilon}u^{-1}du$, and $u = y^{-1/\epsilon}\Psi_M^{1/\epsilon}$ then (notice that when u runs from 0 to ∞, y runs from ∞ to 0):

$$E(u) = \int_0^\infty ue^{-y}dy$$

8

$$= \int_0^\infty y^{-1/\epsilon} \Psi_M^{1/\epsilon} e^{-y} dy$$

which can be written as:

$$E(u) = \nu \Psi_M^{1/\epsilon}$$

$$E(u) = \nu [\sum_{k \in M} \sum_{l \in M} (B_{kl} w_l)^\epsilon (k_{kl} P_k^\alpha Q_k^{1-\alpha})^{-\epsilon}]^{1/\epsilon} \qquad (C9)$$

where $\nu = \Gamma(\frac{\epsilon-1}{\epsilon})$ is the Gamma function. In general, $\Gamma(a) = \int_0^\infty x^{a-1} e^{-x} dx$.

Expected utility conditional on choosing a residence-workplace pair is the same across all residence-workplace pairs in the economy:

$$\bar{U} = \nu [\sum_{k \in N} \sum_{l \in N} (B_{kl} w_l)^\epsilon (k_{kl} P_k^\alpha Q_k^{1-\alpha})^{-\epsilon}]^{1/\epsilon} \qquad (9)$$

which is based on equation (C9).

6.1.2:Bilateral common amenities

$$B_{ni} = B_n^R B_i^L B_{ni}^I \qquad (5)$$

where, $B_n^R, B_i^L, B_{ni}^I > 0$;
B_n^R is residence component common across all workplaces;
B_i^L is workplace component common across all residences;
B_{ni}^I is idiosyncratic component specific to an individual residence-workplace pair.

6.1.3:Residence and Workplace Choices

Using the distribution of utility for pairs of residence and employment locations, the probability that a worker chooses the bilateral commute from n to i out of all possible bilateral commutes is:

$$\lambda_{ni} = Pr[u_{ni} \geq max(u_{kl}); \forall k \neq n, l \neq i]$$

$$= \int_0^\infty [\prod_{k \neq n} \prod_{l \neq i} G_{kl}(u)] dG_{ni}(u)$$

$$= \int_0^\infty [\prod_{k \in M} \prod_{l \in M} e^{-\Psi_{kl} u^{-\epsilon}}] \epsilon \Psi_{ni} u^{-\epsilon-1} du$$

$$= \int_0^\infty [e^{-\Psi_M u^{-\epsilon}}] \epsilon \Psi_{ni} u^{-\epsilon-1} du$$

9

$$= \Psi_{ni} \int_0^\infty \epsilon u^{-\epsilon-1} e^{-\Psi_M u^{-\epsilon}} du$$

$$= \Psi_{ni} \int_0^\infty d(\frac{1}{\Psi_M} e^{-\Psi_M u^{-\epsilon}})$$

$$= \frac{\Psi_{ni}}{\Psi_M} e^{-\Psi_M u^{-\epsilon}} \mid_0^\infty$$

$$= \frac{\Psi_{ni}}{\Psi_M}$$

Note that:

$$\frac{d}{du}[\frac{1}{\Psi_M} e^{-\Psi_M u^{-\epsilon}}] = \epsilon u^{-\epsilon-1} e^{-\Psi_M u^{-\epsilon}}$$

The probability that the worker chooses to live in location n and work in location i is:

$$\frac{L_{ni}}{L_M} = \frac{\Psi_{ni}}{\Psi_M} = \frac{(B_{ni} w_i)^\epsilon (k_{ni} P_n^\alpha Q_n^{1-\alpha})^{-\epsilon}}{\sum_{k\in M} \sum_{l\in M} (B_{kl} w_l)^\epsilon (k_{kl} P_k^\alpha Q_k^{1-\alpha})^{-\epsilon}}$$

where L_{ni} is the measure of commuters from residence n to workplace i; L_M is the measure of workers in the economy as a whole.

Summing across workplaces and residences in Greater London, the probability that a worker chooses a residence-workplace pair in Greater London is given by:

$$\frac{L_N}{L_M} = \frac{\sum_{n\in N} \sum_{i\in N} (B_{ni} w_i)^\epsilon (k_{ni} P_n^\alpha Q_n^{1-\alpha})^{-\epsilon}}{\sum_{k\in M} \sum_{l\in M} (B_{kl} w_l)^\epsilon (k_{kl} P_k^\alpha Q_k^{1-\alpha})^{-\epsilon}}$$

The probability that a worker chooses to live in location n and work in location i conditional on choosing a residence-workplace pair in Greater London is:

$$\lambda_{ni} = \frac{L_{ni}/L_M}{L_N/L_M} = \frac{L_{ni}}{L_N} = \frac{(B_{ni} w_i)^\epsilon (k_{ni} P_n^\alpha Q_n^{1-\alpha})^{-\epsilon}}{\sum_{k\in N} \sum_{l\in N} (B_{kl} w_l)^\epsilon (k_{kl} P_k^\alpha Q_k^{1-\alpha})^{-\epsilon}} \qquad (6) \quad i,n \in N$$

Summing across workplaces $i \in N$ to obtain the probability that a worker lives in residence $n \in N$, conditional on choosing a residence-workplace pair in Greater London:

$$\lambda_n^R = \frac{R_n}{L_N} = \frac{\sum_{i\in N} (B_{ni} w_i)^\epsilon (k_{ni} P_n^\alpha Q_n^{1-\alpha})^{-\epsilon}}{\sum_{k\in N} \sum_{l\in N} (B_{kl} w_l)^\epsilon (k_{kl} P_k^\alpha Q_k^{1-\alpha})^{-\epsilon}} \qquad (7-1)$$

Summing across residences $n \in N$ to obtain the probability that a worker is employed in workplace $i \in N$, conditional on choosing a residence-workplace pair in Greater London:

$$\lambda_n^L = \frac{L_i}{L_N} = \frac{\sum_{n \in N}(B_{ni}w_i)^\epsilon(k_{ni}P_n^\alpha Q_n^{1-\alpha})^{-\epsilon}}{\sum_{k \in N}\sum_{l \in N}(B_{kl}w_l)^\epsilon(k_{kl}P_k^\alpha Q_k^{1-\alpha})^{-\epsilon}} \qquad (7-2)$$

Conditional on working in location i, the conditional probability that a worker commutes from location n is :

$$\lambda_{ni|i}^L = Pr[u_{ni} \geq max(u_{ri}); \forall r \neq n]$$

$$= \int_0^\infty [\prod_{r \neq n} G_{ri}(u)]dG_{ni}(u)$$

$$= \int_0^\infty [e^{-\Psi_i^L u^{-\epsilon}}]\epsilon\Psi_{ni}u^{-\epsilon-1}du$$

$$= \Psi_{ni}\int_0^\infty \epsilon u^{-\epsilon-1}e^{-\Psi_i^L u^{-\epsilon}}du$$

$$= \frac{\Psi_{ni}}{\Psi_i^L}e^{-\Psi_i^L u^{-\epsilon}}\Big|_0^\infty$$

$$= \frac{\Psi_{ni}}{\Psi_i^L}$$

Note that:

$$\frac{d}{du}[\frac{1}{\Psi_i^L}e^{-\Psi_i^L u^{-\epsilon}}] = \epsilon u^{-\epsilon-1}e^{-\Psi_i^L u^{-\epsilon}}$$

where, $\Psi_i^L = \sum_{k \in M}(B_{ki}w_i)^\epsilon(\kappa_{ki}P_k^\alpha Q_k^{1-\alpha})^{-\epsilon}$

Under the assumption of prohibitive commuting costs between Greater London and the wider economy, we have:

$$\sum_{k \in M}(B_{ki}w_i)^\epsilon(\kappa_{ki}P_k^\alpha Q_k^{1-\alpha})^{-\epsilon} = \sum_{k \in N}(B_{ki}w_i)^\epsilon(\kappa_{ki}P_k^\alpha Q_k^{1-\alpha})^{-\epsilon}, \forall i \in N$$

So, the conditional probability that a worker commutes from location n is (conditional on working in location i):

$$\lambda_{ni|i}^L = \frac{(B_{ni}w_i)^\epsilon(\kappa_{ni}P_n^\alpha Q_n^{1-\alpha})^{-\epsilon}}{\sum_{k \in M}(B_{ki}w_i)^\epsilon(\kappa_{ki}P_k^\alpha Q_k^{1-\alpha})^{-\epsilon}}$$

$$= \frac{(B_{ni}w_i)^\epsilon(\kappa_{ni}P_n^\alpha Q_n^{1-\alpha})^{-\epsilon}}{\sum_{k \in N}(B_{ki}w_i)^\epsilon(\kappa_{ki}P_k^\alpha Q_k^{1-\alpha})^{-\epsilon}}$$

$$= \frac{(B_{ni})^\epsilon(\kappa_{ni}P_n^\alpha Q_n^{1-\alpha})^{-\epsilon}}{\sum_{k \in N}(B_{ki})^\epsilon(\kappa_{ki}P_k^\alpha Q_k^{1-\alpha})^{-\epsilon}}$$

11

Conditional on living in location n, the conditional probability that a worker commutes to location i is :

$$\lambda^R_{ni|n} = Pr[u_{ni} \geq max(u_{nl}); \forall l \neq i]$$

$$= \int_0^\infty [\prod_{l \neq i} G_{nl}(u)]dG_{ni}(u)$$

$$= \int_0^\infty [e^{-\Psi^R_n u^{-\epsilon}}]\epsilon\Psi_{ni}u^{-\epsilon-1}du$$

$$= \Psi_{ni}\int_0^\infty \epsilon u^{-\epsilon-1}e^{-\Psi^R_n u^{-\epsilon}}du$$

$$= \frac{\Psi_{ni}}{\Psi^R_n}e^{-\Psi^R_n u^{-\epsilon}}\Big|_0^\infty$$

$$= \frac{\Psi_{ni}}{\Psi^R_n}$$

Note that:

$$\frac{d}{du}[\frac{1}{\Psi^R_n}e^{-\Psi^R_n u^{-\epsilon}}] = \epsilon u^{-\epsilon-1}e^{-\Psi^L_i u^{-\epsilon}}$$

where, $\Psi^R_n = \sum_{l\in M}(B_{nl}w_l)^\epsilon(\kappa_{nl}P^\alpha_n Q^{1-\alpha}_n)^{-\epsilon}$

Under the assumption of prohibitive commuting costs between Greater London and the wider economy, we have:

$$\sum_{l\in M}(B_{nl}w_l)^\epsilon(\kappa_{nl}P^\alpha_n Q^{1-\alpha}_n)^{-\epsilon} = \sum_{l\in N}(B_{nl}w_l)^\epsilon(\kappa_{nl}P^\alpha_n Q^{1-\alpha}_n)^{-\epsilon}, \forall n \in N$$

So, the conditional probability that a worker commutes to location i is (Conditional on living in location n):

$$\lambda^R_{ni|n} = \frac{(B_{ni}w_i)^\epsilon(\kappa_{ni}P^\alpha_n Q^{1-\alpha}_n)^{-\epsilon}}{\sum_{l\in M}(B_{nl}w_l)^\epsilon(\kappa_{nl}P^\alpha_n Q^{1-\alpha}_n)^{-\epsilon}}$$

$$= \frac{(B_{ni}w_i)^\epsilon(\kappa_{ni}P^\alpha_n Q^{1-\alpha}_n)^{-\epsilon}}{\sum_{l\in N}(B_{nl}w_l)^\epsilon(\kappa_{nl}P^\alpha_n Q^{1-\alpha}_n)^{-\epsilon}}$$

$$= \frac{(B_{ni}w_i/\kappa_{ni})^\epsilon(P^\alpha_n Q^{1-\alpha}_n)^{-\epsilon}}{\sum_{l\in N}(B_{nl}w_l/\kappa_{nl})^\epsilon(P^\alpha_n Q^{1-\alpha}_n)^{-\epsilon}}$$

$$= \frac{(B_{ni}w_i/\kappa_{ni})^\epsilon}{\sum_{l\in N}(B_{nl}w_l/\kappa_{nl})^\epsilon} \qquad (14)$$

6.2:Production

The authors assume that consumption goods are produced according to a Cobb-Douglas technology using labor, machinery capital, and commercial floor space.

The homogeneous final good is produced under conditions of perfect competition and constant returns to scale. For simplicity, the authors abstract from the use of machinery capital as a factor of production in this section.

$$Y_i = A_i (\frac{L_i}{\beta^L})^{\beta^L} (\frac{H_i^L}{\beta^H})^{\beta^H}$$

where, $0 < \beta^L, \beta^H < 1$, $\beta^L + \beta^H = 1$;
A_i is final goods productivity and assumed to be exogenous;
Y_i is output of the final good;
L_i is inputs of labor;
H_i^L is commercial floor space.

Cost minimization and zero profits imply that payments for labor, commercial floor space, and machinery are constant shares of revenue (X_i):

$$w_i L_i = \beta^L X_i, \qquad q_i H_i^L = \beta^H X_i, \qquad r M_i = \beta^M X_i \qquad \beta^L + \beta^H + \beta^M = 1 \qquad (10)$$

The authors assume that the final good can be costlessly traded throughout the economy and choose it as our numeraire, such that:$P_{nt} = P_t = 1$. So the profits in workplace i is:

$$\pi_i = Y_i - w_i L_i - q_i H_i^L = A_i (\frac{L_i}{\beta^L})^{\beta^L} (\frac{H_i^L}{\beta^H})^{\beta^H} - w_i L_i - q_i H_i^L$$

where, w_i is the wage;
q_i is the price of commercial floor space, and $q_i = \xi_i Q_i$ (Q_i the price of residential floor space, ξ_i is a location-specific wedge.)

From the first-order conditions for profit maximization, we obtain:

$$\frac{\partial \pi_i}{\partial L_i} = A_i \beta^L (\frac{L_i}{\beta^L})^{\beta^L - 1} \frac{1}{\beta^L} (\frac{H_i^L}{\beta^H})^{\beta^H} - w_i = 0 \qquad (12-1)$$

$$\frac{\partial \pi_i}{\partial H_i^L} = A_i (\frac{L_i}{\beta^L})^{\beta^L} \beta^H \frac{1}{\beta^H} (\frac{H_i^L}{\beta^H})^{\beta^H - 1} - q_i = 0 \qquad (12-2)$$

using (12-1) divided by (12-2) to obtain:

$$(\frac{L_i}{\beta^L})^{-1} (\frac{H_i^L}{\beta^H}) = \frac{w_i}{q_i} \qquad (12-3)$$

rearrange (12-3) to get:

$$q_i H_i^L = \frac{\beta^H}{\beta^L} w_i L_i \qquad (12)$$

13

which means payments for commercial floor space are proportional to workplace income $(w_i L_i)$.

6.3:Commuter market clearing

Commuter market clearing requires that the measure of workers employed in each location i (L_i) equals the sum across all locations n of their measures of residents (R_n) times their conditional probabilities of commuting to i $(\lambda_{ni|n}^R)$:

$$L_i = \sum_{n \in N} \lambda_{ni|n}^R R_n$$

$$= \sum_{n \in N} \frac{(B_{ni} w_i / \kappa_{ni})^\epsilon}{\sum_{l \in N} (B_{nl} w_l / \kappa_{nl})^\epsilon} R_n \qquad (13)$$

Per capita income by residence (v_n) is a weighted average of the wages in all locations:

$$\nu_n = \sum_{i \in N} \lambda_{ni|n}^R w_i$$

$$= \sum_{i \in N} \frac{(B_{ni} w_i / \kappa_{ni})^\epsilon}{\sum_{l \in N} (B_{nl} w_l / \kappa_{nl})^\epsilon} w_i \qquad (15)$$

6.4:Land Market Clearing

The authors assume that floor space is owned by landlords, who receive payments from the residential and commercial use of floor space and consume only consumption goods.

Land market clearing implies that total income from the ownership of floor space equals the sum of payments for residential and commercial floor space use:

$$\mathbb{Q} = Q_n H_n^R + q_n H_n^L = (1 - \alpha)\nu_n R_n + \frac{\beta^H}{\beta^L} w_n L_n \qquad (16)$$

where,α is the overall share of expenditure on consumption goods;
H_n^R is supplies of residential floor space;
H_n^L is supplies of commercial floor space;
R_n is the measure of residents.

7   Quantitative Analysis

7.1:Combined Land and Commuter Market Clearing

- "Exact Hat Algebra" approach:Dekle, Eaton, and Kortum (2007), using rates of change rather than horizontal values to estimate the impact of changes in transport networks.

- In particular, we rewrite our combined land and commuter market-clearing condition (16) for another year τ in terms of the values of variables in a baseline year of t and the relative changes of variables between years τ and t.

$$\tau \neq t$$

because

$$\hat{x}_{n\tau} = \frac{x_{n\tau}}{x_{nt}}$$

and α, β are constant.

So we have (17):

$$\hat{\mathbb{Q}}_{nt}\mathbb{Q}_{nt} = \frac{\mathbb{Q}_{n\tau}}{\mathbb{Q}_{nt}}\mathbb{Q}_{nt}$$

$$= (1-\alpha)\hat{\nu}_{nt}\nu_{nt}\hat{R}_{nt}R_{nt} + \frac{\beta^H}{\beta^L}\hat{w}_{nt}w_{nt}\hat{L}_{nt}L_{nt} \qquad (17)$$

According to (13) and (14),we can simplify $\hat{L}_{it}L_{it}$:

$$\hat{L}_{it}L_{it} = \frac{L_{i\tau}}{L_{it}}L_{it} = L_{i\tau} = \sum_{n\in N}\lambda^R_{ni\tau|n}R_{n\tau} = \sum_{n\in N}\frac{B^\epsilon_{ni\tau}w^\epsilon_{i\tau}\kappa^{-\epsilon}_{ni\tau}}{\sum_{\ell\in N}B^\epsilon_{n\ell\tau}w^\epsilon_{\ell\tau}\kappa^{-\epsilon}_{n\ell\tau}}R_{n\tau}$$

$$= \sum_{n\in N}\frac{B^\epsilon_{ni\tau}w^\epsilon_{i\tau}\kappa^{-\epsilon}_{ni\tau}}{\sum_{\ell\in N}B^\epsilon_{n\ell\tau}w^\epsilon_{\ell\tau}\kappa^{-\epsilon}_{n\ell\tau}}R_{n\tau}\left(\frac{\sum_{\ell\in N}B^\epsilon_{n\ell t}w^\epsilon_{\ell t}\kappa^{-\epsilon}_{n\ell t}}{B^\epsilon_{nit}w^\epsilon_{it}\kappa^{-\epsilon}_{nit}R_{nt}}\right)\left(\frac{B^\epsilon_{nit}w^\epsilon_{it}\kappa^{-\epsilon}_{nit}}{\sum_{\ell\in N}B^\epsilon_{n\ell t}w^\epsilon_{\ell t}\kappa^{-\epsilon}_{n\ell t}}R_{nt}\right)$$

$$= \sum_{n\in N}\frac{B^\epsilon_{ni\tau}w^\epsilon_{i\tau}\kappa^{-\epsilon}_{ni\tau}}{B^\epsilon_{nit}w^\epsilon_{it}\kappa^{-\epsilon}_{nit}}\frac{\sum_{\ell\in N}B^\epsilon_{n\ell t}w^\epsilon_{\ell t}\kappa^{-\epsilon}_{n\ell t}}{\sum_{\ell\in N}B^\epsilon_{n\ell\tau}w^\epsilon_{\ell\tau}\kappa^{-\epsilon}_{n\ell\tau}}\frac{R_{n\tau}}{R_{nt}}\lambda^R_{nit|n}R_{nt}$$

$$= \sum_{n\in N}\frac{B^\epsilon_{ni\tau}w^\epsilon_{i\tau}\kappa^{-\epsilon}_{ni\tau}}{B^\epsilon_{nit}w^\epsilon_{it}\kappa^{-\epsilon}_{nit}}\left(\sum_{\ell\in N}\frac{B^\epsilon_{n\ell\tau}w^\epsilon_{\ell\tau}\kappa^{-\epsilon}_{n\ell\tau}}{\sum_{\ell\in N}B^\epsilon_{n\ell t}w^\epsilon_{\ell t}\kappa^{-\epsilon}_{n\ell t}}\right)^{-1}\frac{R_{n\tau}}{R_{nt}}\lambda^R_{nit|n}R_{nt}$$

$$= \sum_{n\in N}\hat{w}^\epsilon_{it}\hat{\kappa}^{-\epsilon}_{nit}\left(\sum_{\ell\in N}\frac{B^\epsilon_{n\ell t}w^\epsilon_{\ell\tau}\kappa^{-\epsilon}_{n\ell\tau}}{\sum_{\ell\in N}B^\epsilon_{n\ell t}w^\epsilon_{\ell t}\kappa^{-\epsilon}_{n\ell t}}\right)^{-1}\hat{R}_{nt}\lambda^R_{nit|n}R_{nt}$$

$$= \sum_{n\in N}\hat{w}^\epsilon_{it}\hat{\kappa}^{-\epsilon}_{nit}\left(\sum_{\ell\in N}\frac{B^\epsilon_{n\ell t}\hat{w}^\epsilon_{\ell t}w^\epsilon_{\ell t}\hat{\kappa}^{-\epsilon}_{n\ell t}\kappa^{-\epsilon}_{n\ell t}}{\sum_{\ell\in N}B^\epsilon_{n\ell t}w^\epsilon_{\ell t}\kappa^{-\epsilon}_{n\ell t}}\right)^{-1}\hat{R}_{nt}\lambda^R_{nit|n}R_{nt}$$

$$= \sum_{n\in N}\lambda^R_{nit|n}\hat{w}^\epsilon_{it}\hat{\kappa}^{-\epsilon}_{nit}\left(\sum_{\ell\in N}\frac{B^\epsilon_{n\ell t}w^\epsilon_{\ell t}\kappa^{-\epsilon}_{n\ell t}}{\sum_{\ell\in N}B^\epsilon_{n\ell t}w^\epsilon_{\ell t}\kappa^{-\epsilon}_{n\ell t}}\hat{\kappa}^{-\epsilon}_{n\ell t}\hat{w}^\epsilon_{\ell t}\right)^{-1}\hat{R}_{nt}R_{nt}$$

$$= \sum_{n\in N}\lambda^R_{nit|n}\hat{w}^\epsilon_{it}\hat{\kappa}^{-\epsilon}_{nit}\left(\sum_{\ell\in N}\lambda^R_{n\ell t|n}\hat{w}^\epsilon_{\ell t}\hat{\kappa}^{-\epsilon}_{n\ell t}\right)^{-1}\hat{R}_{nt}R_{nt}$$

15

$$= \sum_{n \in N} \frac{\lambda_{nit|n}^R \hat{w}_{it}^\epsilon \hat{\kappa}_{nit}^{-\epsilon}}{\sum_{\ell \in N} \lambda_{n\ell\tau|n}^R \hat{w}_{\ell t}^\epsilon \hat{\kappa}_{n\ell t}^{-\epsilon}} \hat{R}_{nt} R_{nt} \qquad (18)$$

where these equations include terms in changes in wages and commuting costs but not in amenities, because we assume that the workplace and bilateral components of amenities are constant,$\hat{\mathcal{B}}_{it}^L = \hat{\mathcal{B}}_{nit}^I = 1$, and changes in the residential component of amenities, $\hat{\mathcal{B}}_{nt}^R \neq 1$, cancel from the numerator and denominator of the fractions.

According to (14) and (15),we can simplify $\hat{\nu}_{it}\nu_{it}$:

$$\hat{\nu}_{nt}\nu_{nt} = \frac{\nu_{n\tau}}{\nu_{nt}}\nu_{nt} = \sum_{n \in N} \lambda_{ni\tau|n}^R w_{i\tau} = \sum_{n \in N} \frac{B_{ni\tau}^\epsilon w_{i\tau}^\epsilon \kappa_{ni\tau}^{-\epsilon}}{\sum_{\ell \in N} B_{n\ell\tau}^\epsilon w_{\ell\tau}^\epsilon \kappa_{n\ell\tau}^{-\epsilon}} w_{i\tau}$$

$$= \sum_{n \in N} \frac{B_{ni\tau}^\epsilon w_{i\tau}^\epsilon \kappa_{ni\tau}^{-\epsilon}}{\sum_{\ell \in N} B_{n\ell\tau}^\epsilon w_{\ell\tau}^\epsilon \kappa_{n\ell\tau}^{-\epsilon}} w_{i\tau} \left(\frac{\sum_{\ell \in N} B_{n\ell t}^\epsilon w_{\ell t}^\epsilon \kappa_{n\ell t}^{-\epsilon}}{B_{nit}^\epsilon w_{it}^\epsilon \kappa_{nit}^{-\epsilon} w_{it}} \right) \left(\frac{B_{nit}^\epsilon w_{it}^\epsilon \kappa_{nit}^{-\epsilon}}{\sum_{\ell \in N} B_{n\ell t}^\epsilon w_{\ell t}^\epsilon \kappa_{n\ell t}^{-\epsilon}} w_{it} \right)$$

$$= \sum_{n \in N} \frac{B_{ni\tau}^\epsilon w_{i\tau}^\epsilon \kappa_{ni\tau}^{-\epsilon}}{B_{nit}^\epsilon w_{it}^\epsilon \kappa_{nit}^{-\epsilon}} \frac{\sum_{\ell \in N} B_{n\ell t}^\epsilon w_{\ell t}^\epsilon \kappa_{n\ell t}^{-\epsilon}}{\sum_{\ell \in N} B_{n\ell\tau}^\epsilon w_{\ell\tau}^\epsilon \kappa_{n\ell\tau}^{-\epsilon}} \frac{w_{i\tau}}{w_{it}} \lambda_{nit|n}^R w_{it}$$

$$= \sum_{i \in N} \frac{\lambda_{nit|n}^R \hat{w}_{it}^\epsilon \hat{\kappa}_{nit}^{-\epsilon}}{\sum_{\ell \in N} \lambda_{n\ell t|n}^R \hat{w}_{\ell t}^\epsilon \hat{\kappa}_{n\ell t}^{-\epsilon}} \hat{w}_{it} w_{it} \qquad (19)$$

Put (18) and (19) into (17), we have (20):

$$\hat{\mathbb{Q}}_{nt}\mathbb{Q}_{nt} = (1-\alpha) \left[\sum_{i \in N} \frac{\lambda_{nit|n}^R \hat{w}_{it}^\epsilon \hat{\kappa}_{nit}^{-\epsilon}}{\sum_{\ell \in N} \lambda_{n\ell t|n}^R \hat{w}_{\ell t}^\epsilon \hat{\kappa}_{n\ell t}^{-\epsilon}} \hat{w}_{it} w_{it} \right] \hat{R}_{nt} R_{nt}$$

$$+ \frac{\beta^H}{\beta^L} \hat{w}_{nt} w_{nt} \left[\sum_{i \in N} \frac{\lambda_{inti}^R \hat{w}_{nt}^\epsilon \hat{\kappa}_{int}^{-\epsilon}}{\sum_{\ell \in N} \lambda_{i\ell t|i}^R \hat{w}_{\ell t}^\epsilon \hat{\kappa}_{i\ell t}^{-\epsilon}} \hat{R}_{it} R_{it} \right] \qquad (20)$$

Remark 1. • *Suppose that we observe the values of all variables in the initial equilibrium in our baseline year of t:*
the commuting probabilities conditional on residence ($\lambda_{nit|n}^R$);
employment (L_{nt});
residents (R_{nt});
wages (w_{nt});
average per capita income by residence (v_{nt}).

• *Suppose also that we observe relative changes in residents and rateable values between years τ and t: relative changes in residents ($\hat{R}_{nt}$);*
rateable values ($\hat{Q}_{nt}$).

- *Given these observed variables and known values for changes in commuting costs($\hat{\kappa}_{nit}^{-\epsilon}$), this combined land and commuter market-clearing condition (equation (20))provides a system of N equations that determines unique values for the N unknown relative changes in wages in each location($\hat{w}_{nt}$).*

LEMMA 1: Suppose that

$$\left(\hat{\mathbb{Q}}_{nt}, \hat{R}_{nt}, L_{nit}, \lambda_{nit|n}^{R}, \mathbb{Q}, \nu_{nt}, R_{nt}, w_{nt}, L_{nt} \right)$$

are known.

Given known values for model parameters $\left\{ \alpha, \beta^{L}, \beta^{H}, \epsilon \right\}$ and the change in bilateral commuting costs($\hat{\kappa}_{nit}^{-\epsilon}$), the combined land and commuter market-clearing condition (20) determines a unique vector of relative changes in wages ($\hat{w}_{nt}$) in each location.

Proof of LEMMA 1
We first determine the unique vector of relative changes in wages ($\hat{\mathbf{w}}_{\mathbf{t}}$) and then recover the unique vector of relative changes in employment ($\hat{\mathbf{L}}_{\mathbf{t}}$). The combined land and commuter market clearing condition for an earlier year $\tau < t$ can be written as:

$$\hat{\mathbb{Q}}_{nt}\mathbb{Q}_{nt} = T_{n}\left(\hat{\mathbf{w}}_{\mathbf{t}}; \hat{\kappa}_{\mathbf{t}}; \mathbf{Z}_{\mathbf{t}} \right) \qquad (B.1)$$

where $\hat{\mathbb{Q}}_{nt}$ is the observed vector of relative changes in rateable values; $\mathbb{Q}_{nt}$ is the observed vector of rateable values in our baseline year$t = 1921$; $\mathbf{Z}_{\mathbf{t}}$ is a known matrix of relative changes in variables between years τ and t and values for variables in our baseline year t, including $R_{nt}, w_{nt}, \hat{R}_{nt}, \lambda_{nit|n}^{R}$; $\hat{\kappa}_{\mathbf{t}}$ is the matrix of changes in commuting costs; $\hat{\mathbf{w}}_{\mathbf{t}}$ is the vector of relative changes in wages to be determined;

So we have

$$T_{n}\left(\hat{\mathbf{w}}_{\mathbf{t}}; \hat{\kappa}_{\mathbf{t}}; \mathbf{Z}_{\mathbf{t}} \right)$$

$$= (1 - \alpha) \left[\sum_{i \in N} \frac{\lambda_{nit|n}^{R} \hat{w}_{it}^{\epsilon} \hat{\kappa}_{nit}^{-\epsilon}}{\sum_{\ell \in N} \lambda_{n\ell t|n}^{R} \hat{w}_{\ell t}^{\epsilon} \hat{\kappa}_{n\ell t}^{-\epsilon}} \hat{w}_{it} w_{it} \right] \hat{R}_{nt} R_{nt}$$

$$+ \frac{\beta^{H}}{\beta^{L}} \hat{w}_{nt} w_{nt} \left[\sum_{i \in N} \frac{\lambda_{int|i}^{R} \hat{w}_{nt}^{\epsilon} \hat{\kappa}_{int}^{-\epsilon}}{\sum_{l \in N} \lambda_{ilt|i}^{R} \hat{w}_{lt}^{\epsilon} \hat{\kappa}_{ilt}^{-\epsilon}} \hat{R}_{it} R_{it} \right] \qquad (B.2)$$

- Property (i):

$$\hat{\mathbb{Q}}_{t}\mathbb{Q}_{t} > 0 \text{ such that } \hat{\mathbb{Q}}_{nt}\mathbb{Q}_{nt} > 0 \text{ for all } n \in N$$

.

- Property (ii):

$$\lim_{\hat{\mathbf{w}}_{\mathbf{t}} \to 0} T_{n}\left(\hat{\mathbf{w}}_{\mathbf{t}}; \hat{\kappa}_{\mathbf{t}}; \mathbf{Z}_{\mathbf{t}} \right) = 0$$

- Property (iii):

$T_n\left(\hat{\mathbf{w}}_\mathbf{t}; \hat{\kappa}_\mathbf{t}; \mathbf{Z}_\mathbf{t}\right)$ is monotonic in the vector of relative changes in wages, since:

$$\frac{\partial T_n(\cdot)}{\partial \hat{\mathbf{w}}_\mathbf{t}} d\hat{\mathbf{w}}_\mathbf{t} = (1-\alpha)\left[\sum_{i\in N}\frac{\lambda_{nit|n}^R \hat{w}_{it}^\epsilon \hat{\kappa}_{nit}^{-\epsilon}}{\sum_{l\in N}\lambda_{nlt|n}^R \hat{w}_{llt}^\epsilon \hat{\kappa}_{nlt}^{-\epsilon}}\frac{d\hat{w}_{it}}{\hat{w}_{it}}\hat{w}_{it}w_{it}\right]\hat{R}_{nt}R_{nt}$$

$$+ \epsilon(1-\alpha)\left[\sum_{i\in N}\left(1-\frac{\lambda_{nit|n}^R \hat{w}_{it}^\epsilon \hat{\kappa}_{nit}^{-\epsilon}}{\sum_{l\in N}\lambda_{nlt|n}^R \hat{w}_{lt}^\epsilon \hat{\kappa}_{nlt}^{-\epsilon}}\right)\frac{\lambda_{nit|n}^R \hat{w}_{it}^\epsilon \hat{\kappa}_{nit}^{-\epsilon}}{\sum_{l\in N}\lambda_{nlt|n}^R \hat{w}_{it}^\epsilon \hat{\kappa}_{nlt}^{-\epsilon}}\frac{d\hat{w}_{it}}{\hat{w}_{it}}\hat{w}_{it}w_{it}\right]\hat{R}_{nt}R_{nt}$$

$$+ \left(\frac{\beta^H}{\beta^L}\right)\frac{d\hat{w}_{nt}}{\hat{w}_{nt}}\hat{w}_{nt}w_{nt}\left[\sum_{i\in N}\frac{\lambda_{int|i}^R \hat{w}_{nt}^\epsilon \hat{\kappa}_{int}^{-\epsilon}}{\sum_{l\in N}\lambda_{ilt|i}^R \hat{w}_{lt}^\epsilon \hat{\kappa}_{ilt}^{-\epsilon}}\hat{R}_{it}R_{it}\right]$$

$$+ \epsilon\left(\frac{\beta^H}{\beta^L}\right)\hat{w}_{nt}w_{nt}\left[\sum_{i\in N}\left(1-\frac{\lambda_{int|i}^R \hat{w}_{nt}^\epsilon \hat{\kappa}_{int}^{-\epsilon}}{\sum_{l\in N}\lambda_{ilt|i}^R \hat{w}_{lt}^\epsilon \hat{\kappa}_{ilt}^{-\epsilon}}\right)\frac{\lambda_{int|i}^R \hat{w}_{nt}^\epsilon \hat{\kappa}_{int}^{-\epsilon}}{\sum_{l\in N}\lambda_{ilt|i}^R \hat{w}_{lt}^\epsilon \hat{\kappa}_{ilt}^{-\epsilon}}\frac{d\hat{w}_{nt}}{\hat{w}_{nt}}\hat{R}_{it}R_{it}\right] \quad (B.3)$$

where $\frac{\partial T_n(\cdot)}{\partial \hat{\mathbf{w}}_\mathbf{t}} > 0$ for $\hat{\mathbf{w}}_\mathbf{t} > 0$

Proof of Property(iii) :

$\hat{\mathbf{w}}_\mathbf{t} = (\hat{\mathbf{w}}_{1t}, \hat{\mathbf{w}}_{2t}, \cdots, \hat{\mathbf{w}}_{Nt})^T$

Then

$$dT_n = \frac{\partial T_n(\cdot)}{\partial \hat{\mathbf{w}}_{1t}}d\hat{\mathbf{w}}_{1t} + \frac{\partial T_n(\cdot)}{\partial \hat{\mathbf{w}}_{2t}}d\hat{\mathbf{w}}_{2t} + \cdots + \frac{\partial T_n(\cdot)}{\partial \hat{\mathbf{w}}_{Nt}}d\hat{\mathbf{w}}_{Nt} = \frac{\partial T_n(\cdot)}{\partial \hat{\mathbf{w}}_\mathbf{t}}d\hat{\mathbf{w}}_\mathbf{t} \quad (B.3-1)$$

- Property (iv):

$T\left(\hat{\mathbf{w}}_\mathbf{t}; \hat{\kappa}_\mathbf{t}; \mathbf{Z}_\mathbf{t}\right)$ is homogeneous of degree one in the vector of relative changes in wages $(\hat{\mathbf{w}}_\mathbf{t})$ such that

$$T\left(\xi\hat{\mathbf{w}}_\mathbf{t}; \hat{\kappa}_\mathbf{t}; \mathbf{Z}_\mathbf{t}\right)$$

$$= (1-\alpha)\left[\sum_{i\in N}\frac{\lambda_{nit|n}^R \xi\hat{w}_{it}^\epsilon \hat{\kappa}_{nit}^{-\epsilon}}{\sum_{\ell\in N}\lambda_{n\ell t|n}^R \xi\hat{w}_{\ell t}^\epsilon \hat{\kappa}_{n\ell t}^{-\epsilon}}\xi\hat{w}_{it}w_{it}\right]\hat{R}_{nt}R_{nt}$$

$$+ \frac{\beta^H}{\beta^L}\xi\hat{w}_{nt}w_{nt}\left[\sum_{i\in N}\frac{\lambda_{inti}^R \xi\hat{w}_{nt}^\epsilon \hat{\kappa}_{int}^{-\epsilon}}{\sum_{\ell\in N}\lambda_{i\ell t|i}^R \xi\hat{w}_{\ell t}^\epsilon \hat{\kappa}_{i\ell t}^{-\epsilon}}\hat{R}_{it}R_{it}\right]$$

$$= \xi T\left(\hat{\mathbf{w}}_\mathbf{t}; \hat{\kappa}_\mathbf{t}; \mathbf{Z}_\mathbf{t}\right)$$

for any positive scalar ξ.

From properties (i)-(iv), starting from $\hat{w}_{nt} = 0$ for all locations n, and increasing $\hat{w}_{nt}$ for each location n, there exists a unique value for $\hat{w}_{nt}$ for which $\hat{\mathbb{Q}}_t \mathbb{Q}_t = T(\hat{\mathbf{w}}_\mathbf{t}; \hat{\kappa}_\mathbf{t}; \mathbf{Z}_\mathbf{t})$ and equation (B.1) is satisfied.

Using these unique solutions for the relative changes in wages $\hat{w}_{nt}$, the unique vector of relative changes in employment $(\hat{\mathbf{L}}_t)$ can be recovered from the commuter market clearing condition in equation (18), $\hat{L}_{it}L_{it}$, where we have solved for $\hat{w}_{it}^\epsilon$ and we observe $(\hat{R}_{nt}, L_{nt}, R_{nt}, \lambda_{nit|n}^R)$. Similarly,We can solve for the unique relative change inaverage per capita income by residence in equation (19), $\hat{\nu}_{nt}\nu_{nt}$.

Using the conditional commuting probabilities (14),we can obtain the unique relative change in commuting flows:

$$\hat{L}_{nit}L_{nit} = \frac{\lambda_{nit|n}^R \hat{w}_{it}^\epsilon \hat{\kappa}_{nit}^{-\epsilon}}{\sum_{\ell \in N} \lambda_{n\ell t|n}^R \hat{w}_{\ell t}^\epsilon \hat{\kappa}_{n\ell t}^{-\epsilon}} \hat{R}_{nt}R_{nt} \qquad (21)$$

where we have solved for $\hat{w}_{it}^\epsilon$ and we observe $(\hat{R}_{nt}, L_{nt}, R_{nt}, \lambda_{nit|n}^R)$

Remark 2. *Advantages:*

- *Not required to make assumptions about other determinants of economic activity;*

- *Observed probabilities control for unobserved differences in the level of bilateral commuting costs across residence-workplace pairs;*

- *Use $(w_{it}, L_{it}, L_{nit})$ in base line year t and $(\hat{w}_{it}, \hat{L}_{it}, \hat{L}_{nit})$,we can get $(w_{i\tau}, L_{i\tau}, L_{ni\tau})$ and $t \neq \tau$.*

7.2: Commuting Probabilities (Step 1) The authors use the observed data on bilateral commuting flows (L_{nit}) from the population census in our baseline year $t = 1921$ to directly compute the following variables in that baseline year:

- (i) total city employment, $L_t = \sum_{n \in N} \sum_{i \in N} L_{nit}$;

- (ii) the commuting probability conditional on choosing a workplace-residence pair in Greater London,$\lambda_{nit} = L_{nit}/L_{Nt}$;

- (iii) workplace employment, $L_{it} = \sum_{n \in N} L_{nit}$;

- (iv) residence employment, $R_{nt} = \sum_{i \in N} L_{nit}$;

- (v) the commuting probabilities conditional on residence,$\lambda_{nit|n}^R = L_{nit}/R_{nt}$

7.3: Wages in the Initial Equilibrium (Step 2)

The authors solve for wages (w_{nt}) and average per capita income by residence (v_{nt}) in the initial equilibrium in year t = 1921 using the observed workplace employment (L_{nt}), residence employment (R_{nt}), and rateable values $(\tilde{Q}_{nt})$.

19

- First, the authors assume a value for the share of housing in consumer expenditure of $1 - \alpha = 0.25$, which equals the average share of rent in income across occupations in the Registrar General's survey of 30000 workers in Greater London in 1887, as reported to the House of Commons (Parliamentary Papers 1887);

- Second, we assume a value for the share of labor in production costs of $\beta^L = 0.55$, which lies in the middle of the range of 0.43–0.63 considered for the period 1770–1860 in Antras and Voth (2003);

- Third, we assume a share of machinery and equipment in production costs of $\beta^K = 0.20$ and a share of land and building structures in production costs of $\beta^H = 0.25$, which are in line with the data on factor shares for 1856–1913 reported in Matthews, Feinstein, and Odling-Smee (1982);

- $\beta^L + \beta^K + \beta^H = 1$.

7.4: Estimating Commuting Costs (Step 3)

The authors use the model's gravity equation predictions to estimate the relationship between commuting flows and travel times, taking into account the endogeneity of the transport network using instrumental variables estimation.

Assume that bilateral commuting costs are a constant elasticity function of bilateral travel times and a stochastic error for all pairs of boroughs with positive commuting flows:

$$-\epsilon log\left(\kappa_{nit}\right) = -\epsilon\phi log\left(d_{nit}^W\right) + u_{nt}^R + u_{it}^L + u_{nit}^I \qquad (21)$$

where d_{nit}^W is our measure of travel time based on our assumed travel speeds (δ) and the observed transport network ($\tilde{s}_t^{OR}, \tilde{s}_t^{UR}, \tilde{s}_t^{OT}$);
the composite elasticity $-\epsilon\phi$ captures both the elasticity of commuting flows to commuting costs ($-\epsilon$) and the elasticity of commuting costs to travel times (ϕ);
u_{nit}^R is an unobserved residence component of commuting costs;
u_{nit}^L is an unobserved workplace component of commuting costs;
u_{nit}^I is an unobserved idiosyncratic component of commuting costs that is specific to individual residence-workplace pairs.

From equation (5),we can get

$$logB_{nit} = logB_{nt}^R + logB_{it}^L + logB_{nit}^I \qquad (5-1)$$

From equation (9),we can get

$$\frac{\bar{U}}{\nu}\left(\frac{L_{\tilde{N}}}{L_{\tilde{M}}}\right)^{1/\epsilon} = [\textstyle\sum_{k\in\tilde{N}}\sum_{\ell\in\tilde{N}}(B_{k\ell}w_\ell)^\epsilon(k_{k\ell}P_k^\alpha Q_k^{1-\alpha})^{-\epsilon}]^{1/\epsilon} \qquad (9-1)$$

$$\left(\frac{\bar{U}}{\nu}\right)^\epsilon\left(\frac{L_{\tilde{N}}}{L_{\tilde{M}}}\right) = \textstyle\sum_{k\in\tilde{N}}\sum_{\ell\in\tilde{N}}(B_{k\ell}w_\ell)^\epsilon(k_{k\ell}P_k^\alpha Q_k^{1-\alpha})^{-\epsilon} \qquad (9-2)$$

$$\epsilon log\left(\frac{\bar{U}_t}{\nu_t}\right) + log\left(\frac{L_{\tilde{N}t}}{L_{\tilde{M}t}}\right) = log\left(\sum_{k\in\tilde{N}}\sum_{\ell\in\tilde{N}}(B_{k\ell t}w_{\ell t})^\epsilon(k_{k\ell t}P_{kt}^\alpha Q_{kt}^{1-\alpha})^{-\epsilon}\right) \qquad (9-3)$$

From equations (6) and (22), we obtain the following gravity equation for the log commuting probability in our baseline year t:

$$log\lambda_{nit} = log\left(\frac{(B_{nit}w_{it})^\epsilon\left(\kappa_{nit}P_{nt}^\alpha Q_{nt}^{1-\alpha}\right)^{-\epsilon}}{\sum_{k\in\tilde{N}}\sum_{\ell\in\tilde{N}}(B_{k\ell t})^\epsilon\left(\kappa_{k\ell t}P_{kt}^\alpha Q_{kt}^{1-\alpha}\right)^{-\epsilon}}\right)$$

$$= \epsilon log\left(B_{nit}w_{it}\right) - \epsilon log\left(\kappa_{nit}P_{nt}^\alpha Q_{nt}^{1-\alpha}\right) - log\left(\sum_{k\in\tilde{N}}\sum_{\ell\in\tilde{N}}(B_{k\ell t})^\epsilon\left(\kappa_{k\ell t}P_{kt}^\alpha Q_{kt}^{1-\alpha}\right)^{-\epsilon}\right)$$

$$= \epsilon logB_{nt}^R + \epsilon logB_{it}^L + \epsilon logB_{nit}^I + \epsilon log\left(w_{it}\right)$$

$$- \epsilon log\left(\kappa_{nit}P_{nt}^\alpha Q_{nt}^{1-\alpha}\right) - \epsilon log\left(\frac{\bar{U}_t}{\nu_t}\right) - log\left(\frac{L_{\tilde{N}t}}{L_{\tilde{M}t}}\right)$$

$$= \epsilon logB_{nt}^R + \epsilon logB_{it}^L + \epsilon logB_{nit}^I + \epsilon log\left(w_{it}\right)$$

$$- \epsilon\phi log\left(d_{nit}^W\right) + u_{nt}^R + u_{it}^L + u_{nit}^I - \epsilon log\left(P_{nt}^\alpha\right) - \epsilon log\left(Q_{nt}^{1-\alpha}\right) - \epsilon log\left(\frac{\bar{U}_t}{\nu_t}\right) - log\left(\frac{L_{\tilde{N}t}}{L_{\tilde{M}t}}\right)$$

$$= \epsilon log\left(w_{it}\right) + \epsilon logB_{it}^L - log\left(\frac{L_{\tilde{N}t}}{L_{\tilde{M}t}}\right) - \epsilon log\left(\frac{\bar{U}_t}{\nu_t}\right) + u_{it}^L$$

$$+ -\epsilon\alpha log\left(P_{nt}\right) - \epsilon\left(1-\alpha\right)log\left(Q_{nt}\right) + \epsilon logB_{nt}^R + u_{nt}^R$$

$$+ \epsilon logB_{nit}^I + u_{nit}^I - \epsilon\phi log\left(d_{nit}^W\right)$$

$$= \mathbb{W}_{it} + \mathbb{B}_{nt} - \epsilon\phi log\left(d_{nit}^W\right) + e_{nit} \qquad (23)$$

where, the workplace fixed effect includes the wage and any workplace component of amenities or commuting costs, as well as the common expected utility in Greater London ($\mathbb{W}_{it} = \epsilon log\left(w_{it}\right) + \epsilon logB_{it}^L - log\left(\frac{L_{\tilde{N}t}}{L_{\tilde{M}t}}\right) - \epsilon log\left(\frac{\bar{U}_t}{\nu_t}\right) + u_{it}^L$),
the residence fixed effect includes the prices for goods consumption and residential floor space and any residence component of amenities or commuting costs
($\mathbb{B}_{nt} = -\epsilon\alpha log\left(P_{nt}\right) - \epsilon\left(1-\alpha\right)log\left(Q_{nt}\right) + \epsilon logB_{nt}^R + u_{nt}^R$),
the error term captures idiosyncratic shocks to bilateral amenities and commuting costs
($e_{nit} = \epsilon logB_{nit}^I + u_{nit}^I$).

A challenge in estimating equation (23) is that the transport network is endogenous, because overground and underground railways, omnibuses, and trams were operated by private-sector companies.
Therefore, the sets of points connected to the railway and hence the bilateral travel times (d_{nit}^W) are nonrandom. In particular, bilateral pairs that have more commuters for unobserved reasons in the error term (u_{nit}) could have more bilateral transport connections, and hence lower bilateral travel times (d_{nit}^W).

To address this concern, the authors instrument the bilateral travel times based on the transport network (d_{nit}^W) with the bilateral travel times from a world where walking is assumed to be the only mode of transport (d_{ni}^S), in which case bilateral travel times depend solely on straight-line distance. Conditional on the workplace and residence fixed effects, our identifying assumption is that the unobserved factors that affect commuting in the error term (u_{nit}) are orthogonal to the straight-line distance (d_{ni}^S) between locations.

Using the baseline instrumental variables estimates of the elasticity of commuting flows with respect to travel times $(-\epsilon\phi)$ from column (2), the authors construct measures of the change in commuting costs from the removal of the railway network between our baseline year of $t = 1921$ and an earlier year $\tau < t$:

$$\hat{\mathcal{K}}_{nit} = \hat{\kappa}_{nit}^{-\epsilon} = \left(\frac{\kappa_{ni\tau}}{\kappa_{nit}}\right)^{-\epsilon} = \left(\frac{d_{ni\tau}^W\left(\tilde{S}_\tau^{\mathrm{OR}}, \tilde{S}_\tau^{\mathrm{UR}}, \tilde{S}_\tau^{\mathrm{OT}}, \boldsymbol{\delta}\right)}{d_{nit}^W\left(\tilde{S}_t^{\mathrm{OR}}, \tilde{S}_t^{\mathrm{UR}}, \tilde{S}_t^{\mathrm{OT}}, \boldsymbol{\delta}\right)}\right)^{-\epsilon\phi} \quad (24)$$

7.5: Historical Workplace Employment (Step 4)

- Given these estimates of changes in commuting costs $(\hat{\kappa}_{nit}^{-\epsilon})$and a value for the Fr´echet shape parameter $(\epsilon$), we can use ourcombined land and commuter market-clearing condition (20) to solve for unobserved workplace employment back to the early nineteenth century.

- The authors calibrate the Frechet shape parameter$(\epsilon$) by minimizing the sum of squared deviations between the model's predictions for workplace employment and the data on the day population in the City of London for the census years for which these data are available (1881, 1891, and 1911). We obtain a calibrated Fr´echet shape parameter of $\epsilon = 5.25$, which lies in between the estimate of around 6 in Ahlfeldt et al. (2015) and the estimate using migration flows of around 3 in Bryan and Morten (2019).

7.6: Specification Check (Step 5)

- As a specification check, we now compare our model's predictions for commuting patterns during the nineteenth century to historical data that were not used in its calibration. In particular, we use data on commuting distances from the personnel ledgers of Henry Poole bespoke tailors.

- In the left panel of Figure VII, we compare the model's predictions for 1861 with the commuting distances of workers who joined Henry Poole between 1857 and 1877.In the right panel, we compare the model's predictions for 1901 with commuting distances of workers who joined Henry Poole between 1891 and 1911.

- In these comparisons, there are several potential sources ofdiscrepancies between the model and data, including the fact thatthis company is located in a specific site in Westminster, whereas the model covers all of that borough.

8 Productivity, Amenities, and the Supply of Floor Space

In this section, the authors consider one theoretical model to explicitly recover productivity, amenities, and the supply of floor space and to examine the implications of the findings for the strength of agglomeration forces.

In particular, the authors choose an extension of the canonical urban model to incorporate non-traded goods. This framework permits a particularly tractable and transparent approach to recovering productivity and amenities and estimating the strength of agglomeration forces.

8.1: Supply of Floor Space

In this section, the authors separate rateable values into the price and supply of floor space.

Two assumptions:

- 1) the authors assume no arbitrage between commercial and residential floor space: $q_{nt} = Q_{nt}$.
  Hence, the rateable values (from (16)) can be written as:

$$\mathbb{Q}_{nt} = Q_{nt}H_{nt}^R + q_{nt}H_{nt}^L = Q_{nt}H_{nt} \qquad (G.1)$$

- 2) the authors model the supply of floor space (H_{nt}) as depending on geographical land area (K_n) and a constant elasticity function of the price of floor space (Q_{nt}):

$$H_{nt} = hQ_{nt}^{\mu}K_n \qquad (G.2)$$

  where, h is a constant;
  $\mu \geq 0$ is the floor space supply elasticity; and $\mu = 0$ corresponds to the special case of a perfectly inelastic supply of floor space.
  Note that the authors set the floor space supply elasticity in the model to the **median value** of these floor space supply elasticities across this subset of boroughs in the data ($\mu = 1.83$).

Substituting this supply function for floor space (G.2) into our expression for rateable values (G.1), we obtain:

$$\mathbb{Q}_{nt} = Q_{nt}hQ_{nt}^{\mu}K_n = hK_nQ_{nt}^{\mu+1} \qquad (G.3)$$

Re-arranging (G.3), we obtain the closed-form expression for the price of floor space:

$$Q_{nt} = (\frac{\mathbb{Q}_{nt}}{hK_n})^{\frac{1}{1+\mu}} \qquad (26-1)$$

Substituting (26-1) into (G.2), we obtain the closed-form expression for the supply of floor space:

$$H_{nt} = hK_n(\frac{\mathbb{Q}_{nt}}{hK_n})^{\frac{\mu}{1+\mu}} \qquad (26-2)$$

Taking relative changes over time in equations (26-1) and (26-2)can obtain the relative changes in the price of floor space (Q_{nt}) and the supply of floor space (H_{nt}):

$$\hat{Q}_{nt} = \hat{\mathbb{Q}}_{nt}^{\frac{1}{1+\mu}};$$
$$\hat{H}_{nt} = \hat{\mathbb{Q}}_{nt}^{\frac{\mu}{1+\mu}}$$

8.2: Productivity and Amenities In this section, the authors recover productivity and amenities for each location.

To recover productivity and amenities, the authors choose an extension of the canonical urban model to incorporate nontraded goods. This framework permits a particularly tractable and transparent approach to estimating the strength of agglomeration forces. It also allows us to undertake counterfactuals for removing the railway network under a range of alternative assumptions about the floor space supply elasticity and the strength of agglomeration forces.

Starting with preferences in equation (3) in the paper, the authors assume that the consumption goods price index (P_{nt}) is a Cobb-Douglas function of the price of a homogeneous traded good (P_{nt}^T) and of a homogeneous non-traded good (P_{nt}^N), as in equation (27):

$$P_{nt} = (P_{nt}^T)^{\nu}(P_{nt}^N)^{1-\nu} \qquad (27)$$

where, ν is the traded goods expenditure share, and $0 < \nu < 1$.

Within Greater London, we assume that the homogeneous traded good is costlessly traded, such that

$$P_{nt}^T = P_t^T, \forall n \in N$$

Between Greater London and the rest of Great Britain, the authors allow for changes in trade costs for the traded good, which are reflected in changes in the common price of this good within Greater London (P_t^T).

The authors assume that both traded and non-traded goods are produced with labor, machinery capital and commercial floor space under *conditions of perfect competition*; and assume for simplicity the same production technology in both sectors.

For traded production:

$$Y_{nt}^T = A_{nt}^T(\frac{L_{nt}^T}{\beta^L})^{\beta^L}(\frac{H_{nt}^T}{\beta^H})^{\beta^H}(\frac{M_{nt}^T}{\beta^M})^{\beta^M} \qquad (28-1)$$

24

For non-traded production:

$$Y_{nt}^N = A_{nt}^N \left(\frac{L_{nt}^N}{\beta^L}\right)^{\beta^L} \left(\frac{H_{nt}^N}{\beta^H}\right)^{\beta^H} \left(\frac{M_{nt}^N}{\beta^M}\right)^{\beta^M} \qquad (28-2)$$

Cost minimization and zero profits imply that payments for labor, commercial floor space, and machinery are constant shares of revenue ($X_{nt}^T = P_{nt}^T Y_{nt}^T$, and $X_{nt}^N = P_{nt}^N Y_{nt}^N$):

$$w_{nt} L_{nt}^T = \beta^L X_{nt}^T \qquad Q_{nt} H_{nt}^T = \beta^H X_{nt}^T \qquad r_{nt} M_{nt}^T = \beta^M X_{nt}^T \qquad (28-3)$$

$$w_{nt} L_{nt}^N = \beta^L X_{nt}^N \qquad Q_{nt} H_{nt}^N = \beta^H X_{nt}^N \qquad r_{nt} M_{nt}^N = \beta^M X_{nt}^N \qquad (28-4)$$

where, $\beta^L + \beta^H + \beta^M = 1$;
L_{nt}^T and L_{nt}^N denote employment in the traded and non-traded sectors respectively;
X_{nt}^T and X_{nt}^N correspond to revenue in the two sectors;
H_{nt}^T and H_{nt}^N represent commercial floor space use in the two sectors;
M_{nt}^T and M_{nt}^N are machinery inputs in the two sectors;
w_{nt} is wages;
Q_{nt} is the price of floor space;
r_{nt} is the common price of machinery, and $r_{nt} = r_t$.

Using equations in (28-3) to substitute for employment (L_{nt}^T), commercial floor space (H_{nt}^T), and machinery inputs(M_{nt}^T) in traded sector production (28-1):

$$X_{nt}^T = P_{nt}^T Y_{nt}^T$$

$$= P_{nt}^T A_{nt}^T \left(\frac{L_{nt}^T}{\beta^L}\right)^{\beta^L} \left(\frac{H_{nt}^T}{\beta^H}\right)^{\beta^H} \left(\frac{M_{nt}^T}{\beta^M}\right)^{\beta^M}$$

$$= P_{nt}^T A_{nt}^T \left(\frac{w_{nt}^{-1}\beta^L X_{nt}^T}{\beta^L}\right)^{\beta^L} \left(\frac{Q_{nt}^{-1}\beta^H X_{nt}^T}{\beta^H}\right)^{\beta^H} \left(\frac{r_{nt}^{-1}\beta^M X_{nt}^T}{\beta^M}\right)^{\beta^M}$$

$$= P_{nt}^T A_{nt}^T w_{nt}^{-\beta^L} Q_{nt}^{-\beta^H} r_{nt}^{-\beta^M} (X_{nt}^T)^{\beta^L + \beta^H + \beta^M}$$

cancle out X_{nt}^T in both sides, and multiply both sides by $\frac{1}{A_{nt}^T} w_{nt}^{\beta^L} Q_{nt}^{\beta^H} r_{nt}^{\beta^M}$, can get

$$P_{nt}^T = \frac{1}{A_{nt}^T} w_{nt}^{\beta^L} Q_{nt}^{\beta^H} r_{nt}^{\beta^M}$$

also, because $P_{nt}^T = P_t^T$,and $r_{nt} = r_t$,

$$P_t^T = \frac{1}{A_{nt}^T} w_{nt}^{\beta^L} Q_{nt}^{\beta^H} r_t^{\beta^M} \qquad (G.12)$$

Doing the same process using (28-2) and (28-4), we can get the zero-profit condition for Non-traded sector for each location is:

$$P_{nt}^N = \frac{1}{A_{nt}^N} w_{nt}^{\beta^L} Q_{nt}^{\beta^H} r_t^{\beta^M} \qquad (G.13)$$

where, $0 < \beta^L, \beta^H, \beta^M < 1$.

Re-arranging the traded zero-profit condition (G.12) to get expression for composite traded productivity:

$$P_t^T A_{nt}^T r_t^{-\beta^M} = w_{nt}^{\beta^L} Q_{nt}^{\beta^H}$$

composite traded productivity denoted by $\mathbb{A}_{nt}^T$ which captures traded productivity (A_{nt}^T), the common price of the traded good (P_t^T), and the common price of machinery (r_t), so:

$$\mathbb{A}_{nt}^T = w_{nt}^{\beta^L} Q_{nt}^{\beta^H} \qquad (28)$$

where wages (w_{nt}) was determined in the baseline quantitative analysis and the price of floor space (Q_{nt}) in the previous section.

Taking relative changes over time in (28), the relative change in composite traded productivity for each location is:

$$\hat{\mathbb{A}}_{nt}^T = \hat{w}_{nt}^{\beta^L} \hat{Q}_{nt}^{\beta^H} \qquad (G.15)$$

Next, the authors use these additional assumptions on preferences, production technology and market structure in the residence choice probabilities in equation (7),to recover a composite measure of changes in amenities.

Recall equation (7):
the probability that a worker lives in residence $n \in N$, conditional on choosing a residence-workplace pair in Greater London:

$$\lambda_n^R = \frac{R_n}{L_N} = \frac{\sum_{i \in N} (B_{ni} w_i)^\epsilon (\kappa_{ni} P_n^\alpha Q_n^{1-\alpha})^{-\epsilon}}{\sum_{k \in N} \sum_{l \in N} (B_{kl} w_l)^\epsilon (\kappa_{kl} P_k^\alpha Q_k^{1-\alpha})^{-\epsilon}} \qquad (7-1)$$

Substituting the price index for consumption goods (P_{nt}) from equation (27) into these residence choice probabilities:

$$\lambda_{nt}^R = \frac{\sum_{l \in N} (B_{nlt} w_{lt})^\epsilon (k_{nlt} (P_{nt}^T)^{\alpha\nu} (P_{nt}^N)^{\alpha(1-\nu)} Q_n^{1-\alpha})^{-\epsilon}}{\sum_{k \in N} \sum_{l \in N} (B_{klt} w_{lt})^\epsilon (k_{klt} (P_{kt}^T)^{\alpha\nu} (P_{kt}^N)^{\alpha(1-\nu)} Q_{kt}^{1-\alpha})^{-\epsilon}} \qquad (G.16)$$

Using the assumption that the homogeneous traded good is costlessly traded ($P_{nt}^T = P_t^T$), (G.16) can be simplified to:

$$\lambda_{nt}^R = \frac{\sum_{l\in N}(B_{nlt}w_{lt})^\epsilon(k_{nlt}(P_{nt}^N)^{\alpha(1-\nu)}Q_n^{1-\alpha})^{-\epsilon}}{\sum_{k\in N}\sum_{l\in N}(B_{klt}w_{lt})^\epsilon(k_{klt}(P_{kt}^N)^{\alpha(1-\nu)}Q_{kt}^{1-\alpha})^{-\epsilon}} \qquad (G.17)$$

Using (G.13) to substitute for the price of the non-traded good (P_{nt}^N) into (G.17):

$$\lambda_{nt}^R = \frac{\sum_{l\in N}(B_{nlt}w_{lt})^\epsilon(\kappa_{nlt}(\frac{1}{A_{nt}^N}w_{nt}^{\beta^L}Q_{nt}^{\beta^H}r_t^{\beta^M})^{\alpha(1-\nu)}Q_n^{1-\alpha})^{-\epsilon}}{\sum_{k\in N}\sum_{l\in N}(B_{klt}w_{lt})^\epsilon(\kappa_{klt}(\frac{1}{A_{kt}^N}w_{kt}^{\beta^L}Q_{kt}^{\beta^H}r_t^{\beta^M})^{\alpha(1-\nu)}Q_{kt}^{1-\alpha})^{-\epsilon}}$$

$$= \frac{\sum_{l\in N}B_{nlt}^\epsilon w_{lt}^\epsilon\kappa_{nlt}^{-\epsilon}(A_{nt}^N)^{\epsilon\alpha(1-\nu)}w_{nt}^{-\epsilon\alpha(1-\nu)\beta^L}Q_{nt}^{-\epsilon[\alpha(1-\nu)\beta^H+(1-\alpha)]}r_t^{-\epsilon\alpha(1-\nu)\beta^M}}{\sum_{k\in N}\sum_{l\in N}B_{klt}^\epsilon w_{lt}^\epsilon\kappa_{klt}^{-\epsilon}(A_{kt}^N)^{\epsilon\alpha(1-\nu)}w_{kt}^{-\epsilon\alpha(1-\nu)\beta^L}Q_{kt}^{-\epsilon[\alpha(1-\nu)\beta^H)+(1-\alpha)]}r_t^{-\epsilon\alpha(1-\nu)\beta^M}}$$

$$= \frac{\sum_{l\in N}B_{nlt}^\epsilon w_{lt}^\epsilon\kappa_{nlt}^{-\epsilon}(A_{nt}^N)^{\epsilon\alpha(1-\nu)}w_{nt}^{-\epsilon\alpha(1-\nu)\beta^L}Q_{nt}^{-\epsilon[\alpha(1-\nu)\beta^H+(1-\alpha)]}}{\sum_{k\in N}\sum_{l\in N}B_{klt}^\epsilon w_{lt}^\epsilon\kappa_{klt}^{-\epsilon}(A_{kt}^N)^{\epsilon\alpha(1-\nu)}w_{kt}^{-\epsilon\alpha(1-\nu)\beta^L}Q_{kt}^{-\epsilon[\alpha(1-\nu)\beta^H)+(1-\alpha)]}} \qquad (G.18)$$

re-arranging the zero-profit condition for the traded sector (G.12):

$$w_{nt} = [P_t^T A_{nt}^T Q_{nt}^{-\beta^H}r_t^{-\beta^M}]^{\frac{1}{\beta^L}} \qquad (G.12-1)$$

Using (G.12-1) to substitute for the wage by residence (w_{nt} in the numerator and w_{kt} in the denominator) in (G.18):

$$\lambda_{nt}^R = \frac{\sum_{l\in N}B_{nlt}^\epsilon w_{lt}^\epsilon\kappa_{nlt}^{-\epsilon}(A_{nt}^N)^{\epsilon\alpha(1-\nu)}[P_t^T A_{nt}^T Q_{nt}^{-\beta^H}r_t^{-\beta^M}]^{\frac{-\epsilon\alpha(1-\nu)\beta^L}{\beta^L}}Q_{nt}^{-\epsilon[\alpha(1-\nu)\beta^H+(1-\alpha)]}}{\sum_{k\in N}\sum_{l\in N}B_{klt}^\epsilon w_{lt}^\epsilon\kappa_{klt}^{-\epsilon}(A_{kt}^N)^{\epsilon\alpha(1-\nu)}[P_t^T A_{kt}^T Q_{kt}^{-\beta^H}r_t^{-\beta^M}]^{\frac{-\epsilon\alpha(1-\nu)\beta^L}{\beta^L}}Q_{kt}^{-\epsilon[\alpha(1-\nu)\beta^H)+(1-\alpha)]}}$$

$$= \frac{\sum_{l\in N}B_{nlt}^\epsilon w_{lt}^\epsilon\kappa_{nlt}^{-\epsilon}(A_{nt}^N)^{\epsilon\alpha(1-\nu)}(P_t^T)^{-\epsilon\alpha(1-\nu)}(A_{nt}^T)^{-\epsilon\alpha(1-\nu)}Q_{nt}^{\epsilon\alpha(1-\nu)\beta^H}r_t^{\epsilon\alpha(1-\nu)\beta^M}Q_{nt}^{-\epsilon[\alpha(1-\nu)\beta^H+(1-\alpha)]}}{\sum_{k\in N}\sum_{l\in N}B_{klt}^\epsilon w_{lt}^\epsilon\kappa_{klt}^{-\epsilon}(A_{kt}^N)^{\epsilon\alpha(1-\nu)}(P_t^T)^{-\epsilon\alpha(1-\nu)}(A_{kt}^T)^{-\epsilon\alpha(1-\nu)}Q_{kt}^{\epsilon\alpha(1-\nu)\beta^H}r_t^{\epsilon\alpha(1-\nu)\beta^M}Q_{kt}^{-\epsilon[\alpha(1-\nu)\beta^H)+(1-\alpha)]}}$$

$$= \frac{\sum_{l\in N}B_{nlt}^\epsilon w_{lt}^\epsilon\kappa_{nlt}^{-\epsilon}(A_{nt}^N)^{\epsilon\alpha(1-\nu)}(A_{nt}^T)^{-\epsilon\alpha(1-\nu)}Q_{nt}^{-\epsilon(1-\alpha)}}{\sum_{k\in N}\sum_{l\in N}B_{klt}^\epsilon w_{lt}^\epsilon\kappa_{klt}^{-\epsilon}(A_{kt}^N)^{\epsilon\alpha(1-\nu)}(A_{kt}^T)^{-\epsilon\alpha(1-\nu)}Q_{kt}^{-\epsilon(1-\alpha)}} \qquad (G.19)$$

Using equation (5): $B_{nit} = B_{nt}^R B_{it}^L B_{nit}^I$ to decompose bilateral amenities (B_{nit}) in (G.19):

$$\lambda_{nt}^R = \frac{\sum_{l\in N}(B_{nt}^R B_{lt}^L B_{nlt}^I)^\epsilon w_{lt}^\epsilon\kappa_{nlt}^{-\epsilon}(A_{nt}^N)^{\epsilon\alpha(1-\nu)}(A_{nt}^T)^{-\epsilon\alpha(1-\nu)}Q_{nt}^{-\epsilon(1-\alpha)}}{\sum_{k\in N}\sum_{l\in N}(B_{kt}^R B_{lt}^L B_{klt}^I)^\epsilon w_{lt}^\epsilon\kappa_{klt}^{-\epsilon}(A_{kt}^N)^{\epsilon\alpha(1-\nu)}(A_{kt}^T)^{-\epsilon\alpha(1-\nu)}Q_{kt}^{-\epsilon(1-\alpha)}} \qquad (G.20)$$

Defining a measure of residents' commuting market access (RMA_n) as:

$$RMA_{nt} \equiv \Big[\sum_{l\in N}(B_{lt}^L B_{nlt}^I)^\epsilon w_{lt}^\epsilon \kappa_{nlt}^{-\epsilon}\Big]^{\frac{1}{\epsilon}} \qquad (G.21)$$

and a measure of composite residential amenities as:

$$\mathbb{B}_{nt} = B_{nt}^R (A_{nt}^N)^{\alpha(1-\nu)}(A_{nt}^T)^{-\alpha(1-\nu)} \qquad (G.22)$$

We can further re-write the residence choice probabilities in (G.20) as:

$$\lambda_{nt}^R = \frac{(B_{nt}^R)^\epsilon (A_{nt}^N)^{\epsilon\alpha(1-\nu)}(A_{nt}^T)^{-\epsilon\alpha(1-\nu)} Q_{nt}^{-\epsilon(1-\alpha)} \sum_{l\in N}(B_{lt}^L B_{nlt}^I)^\epsilon w_{lt}^\epsilon \kappa_{nlt}^{-\epsilon}}{\sum_{k\in N}(B_{kt}^R)^\epsilon (A_{kt}^N)^{\epsilon\alpha(1-\nu)}(A_{kt}^T)^{-\epsilon\alpha(1-\nu)} Q_{kt}^{-\epsilon(1-\alpha)} \sum_{l\in N}(B_{lt}^L B_{klt}^I)^\epsilon w_{lt}^\epsilon \kappa_{klt}^{-\epsilon}}$$

$$= \frac{\mathbb{B}_{nt}^\epsilon Q_{nt}^{-\epsilon(1-\alpha)} RMA_{nt}^\epsilon}{\sum_{k\in N}\mathbb{B}_{kt}^\epsilon Q_{kt}^{-\epsilon(1-\alpha)} RMA_{kt}^\epsilon} \qquad (G.23)$$

The change in residents' commuting market access is:

$$\widehat{RMA}_{nt} = \frac{RMA_{n\tau}}{RMA_{nt}}$$

$$= \left[\frac{\sum_{l\in N}(B_{l\tau}^L B_{nl\tau}^I)^\epsilon w_{l\tau}^\epsilon \kappa_{nl\tau}^{-\epsilon}}{\sum_{l\in N}(B_{lt}^L B_{nlt}^I)^\epsilon w_{lt}^\epsilon \kappa_{nlt}^{-\epsilon}}\right]^{\frac{1}{\epsilon}}$$

$$= \left[\frac{\sum_{l\in N}(B_{nt}^R B_{lt}^L B_{nlt}^I)^\epsilon w_{l\tau}^\epsilon \kappa_{nl\tau}^{-\epsilon}}{\sum_{l\in N}(B_{nt}^R B_{lt}^L B_{nlt}^I)^\epsilon w_{lt}^\epsilon \kappa_{nlt}^{-\epsilon}}\right]^{\frac{1}{\epsilon}}$$

$$= \left[\sum_{l\in N}\frac{(B_{nt}^R B_{lt}^L B_{nlt}^I)^\epsilon w_{lt}^\epsilon \kappa_{nlt}^{-\epsilon} \hat{w}_{lt}^\epsilon \hat{\kappa}_{nlt}^{-\epsilon}}{\sum_{l\in N}(B_{nt}^R B_{lt}^L B_{nlt}^I)^\epsilon w_{lt}^\epsilon \kappa_{nlt}^{-\epsilon}}\right]^{\frac{1}{\epsilon}}$$

$$= \left[\sum_{l\in N}\frac{(B_{nt}^R B_{lt}^L B_{nlt}^I)^\epsilon w_{lt}^\epsilon \kappa_{nlt}^{-\epsilon}}{\sum_{l\in N}(B_{nt}^R B_{lt}^L B_{nlt}^I)^\epsilon w_{lt}^\epsilon \kappa_{nlt}^{-\epsilon}}\hat{w}_{lt}^\epsilon \hat{\kappa}_{nlt}^{-\epsilon}\right]^{\frac{1}{\epsilon}}$$

$$= \left[\sum_{l\in N}\lambda_{nlt|n}^R \hat{w}_{lt}^\epsilon \hat{\kappa}_{nlt}^{-\epsilon}\right]^{\frac{1}{\epsilon}} \qquad (29-1)$$

where, the authors assume constant workplace and idiosyncratic components of amenities: $\hat{B}_{lt}^L = 1$ and $\hat{B}_{nlt}^I = 1$.

The changes in composite amenities:

$$\hat{\mathbb{B}}_{nt} \equiv \hat{B}_{nt}^{R}(\hat{A}_{nt}^{N})^{\alpha(1-\nu)}(\hat{A}_{nt}^{T})^{-\alpha(1-\nu)} \qquad (29-2)$$

In relative changes, we can re-write the residence choice probabilities as:

$$\hat{\lambda}_{nt}^{R}\lambda_{nt}^{R} = \frac{\lambda_{n\tau}^{R}}{\lambda_{nt}^{R}}\lambda_{nt}^{R} = \lambda_{n\tau}^{R}$$

$$= \frac{\mathbb{B}_{n\tau}^{\epsilon}Q_{n\tau}^{-\epsilon(1-\alpha)}RMA_{n\tau}^{\epsilon}}{\sum_{k\in N}\mathbb{B}_{k\tau}^{\epsilon}Q_{k\tau}^{-\epsilon(1-\alpha)}RMA_{k\tau}^{\epsilon}}$$

$$= [\frac{\mathbb{B}_{n\tau}^{\epsilon}Q_{n\tau}^{-\epsilon(1-\alpha)}RMA_{n\tau}^{\epsilon}}{\sum_{k\in N}\mathbb{B}_{k\tau}^{\epsilon}Q_{k\tau}^{-\epsilon(1-\alpha)}RMA_{k\tau}^{\epsilon}}][\frac{\sum_{k\in N}\mathbb{B}_{kt}^{\epsilon}Q_{kt}^{-\epsilon(1-\alpha)}RMA_{kt}^{\epsilon}}{\mathbb{B}_{nt}^{\epsilon}Q_{nt}^{-\epsilon(1-\alpha)}RMA_{nt}^{\epsilon}}][\frac{\mathbb{B}_{nt}^{\epsilon}Q_{nt}^{-\epsilon(1-\alpha)}RMA_{nt}^{\epsilon}}{\sum_{k\in N}\mathbb{B}_{kt}^{\epsilon}Q_{kt}^{-\epsilon(1-\alpha)}RMA_{kt}^{\epsilon}}]$$

$$= [\frac{\mathbb{B}_{n\tau}^{\epsilon}Q_{n\tau}^{-\epsilon(1-\alpha)}RMA_{n\tau}^{\epsilon}}{\mathbb{B}_{nt}^{\epsilon}Q_{nt}^{-\epsilon(1-\alpha)}RMA_{nt}^{\epsilon}}][\frac{\sum_{k\in N}\mathbb{B}_{kt}^{\epsilon}Q_{kt}^{-\epsilon(1-\alpha)}RMA_{kt}^{\epsilon}}{\sum_{k\in N}\mathbb{B}_{k\tau}^{\epsilon}Q_{k\tau}^{-\epsilon(1-\alpha)}RMA_{k\tau}^{\epsilon}}][\frac{\mathbb{B}_{nt}^{\epsilon}Q_{nt}^{-\epsilon(1-\alpha)}RMA_{nt}^{\epsilon}}{\sum_{k\in N}\mathbb{B}_{kt}^{\epsilon}Q_{kt}^{-\epsilon(1-\alpha)}RMA_{kt}^{\epsilon}}]$$

$$= \hat{\mathbb{B}}_{nt}^{\epsilon}\hat{Q}_{nt}^{-\epsilon(1-\alpha)}\widehat{RMA}_{nt}^{\epsilon}\lambda_{nt}^{R}\underbrace{\frac{\sum_{k\in N}\mathbb{B}_{kt}^{\epsilon}Q_{kt}^{-\epsilon(1-\alpha)}RMA_{kt}^{\epsilon}}{\sum_{k\in N}\mathbb{B}_{k\tau}^{\epsilon}Q_{k\tau}^{-\epsilon(1-\alpha)}RMA_{k\tau}^{\epsilon}}}_{F} \qquad (29-3)$$

To solve part F in (29-3):

$$\frac{\sum_{k\in N}\mathbb{B}_{kt}^{\epsilon}Q_{kt}^{-\epsilon(1-\alpha)}RMA_{kt}^{\epsilon}}{\sum_{k\in N}\mathbb{B}_{k\tau}^{\epsilon}Q_{k\tau}^{-\epsilon(1-\alpha)}RMA_{k\tau}^{\epsilon}}$$

$$= \left[\frac{\sum_{k\in N}\mathbb{B}_{k\tau}^{\epsilon}Q_{k\tau}^{-\epsilon(1-\alpha)}RMA_{k\tau}^{\epsilon}}{\sum_{k\in N}\mathbb{B}_{kt}^{\epsilon}Q_{kt}^{-\epsilon(1-\alpha)}RMA_{kt}^{\epsilon}}\right]^{-1}$$

$$= \left[\sum_{k\in N}\frac{\frac{\mathbb{B}_{k\tau}^{\epsilon}}{\mathbb{B}_{kt}^{\epsilon}}\mathbb{B}_{kt}^{\epsilon}\frac{Q_{k\tau}^{-\epsilon(1-\alpha)}}{Q_{kt}^{-\epsilon(1-\alpha)}}Q_{kt}^{-\epsilon(1-\alpha)}\frac{RMA_{k\tau}^{\epsilon}}{RMA_{kt}^{\epsilon}}RMA_{kt}^{\epsilon}}{\sum_{k\in N}\mathbb{B}_{kt}^{\epsilon}Q_{kt}^{-\epsilon(1-\alpha)}RMA_{kt}^{\epsilon}}\right]^{-1}$$

$$= \left[\sum_{k\in N}\frac{\mathbb{B}_{kt}^{\epsilon}Q_{kt}^{-\epsilon(1-\alpha)}RMA_{kt}^{\epsilon}}{\sum_{k\in N}\mathbb{B}_{kt}^{\epsilon}Q_{kt}^{-\epsilon(1-\alpha)}RMA_{kt}^{\epsilon}}\hat{\mathbb{B}}_{kt}^{\epsilon}\hat{Q}_{kt}^{-\epsilon(1-\alpha)}\widehat{RMA}_{kt}^{\epsilon}\right]^{-1}$$

$$= \left[\sum_{k\in N}\lambda_{kt}^{R}\hat{\mathbb{B}}_{kt}^{\epsilon}\hat{Q}_{kt}^{-\epsilon(1-\alpha)}\widehat{RMA}_{kt}^{\epsilon}\right]^{-1}$$

So:

$$\hat{\lambda}_{nt}^R \lambda_{nt}^R = \frac{\lambda_{nt}^R \hat{\mathbb{B}}_{nt}^\epsilon \hat{Q}_{nt}^{-\epsilon(1-\alpha)} \widehat{RMA}_{nt}^\epsilon}{\sum_{k \in N} \lambda_{kt}^R \hat{\mathbb{B}}_{kt}^\epsilon \hat{Q}_{kt}^{-\epsilon(1-\alpha)} \widehat{RMA}_{kt}^\epsilon} \qquad (29)$$

Notes: The baseline year of $t = 1921$, and $\tau = 1831 < t$;
The authors use equation (29) to solve for a unique vector of changes in composite amenities $\hat{\mathbb{B}}_{nt}$ up to a normalization or choice of units.

8.3: Agglomeration Forces
In this section, the authors separate changes in productivity and amenities into the contributions of agglomeration forces and changes in locational fundamentals.

The authors assume that composite traded productivity $\mathbb{A}_{nt}^T$ depends on production fundamentals and production externalities.

$$\mathbb{A}_{nt}^T = \left(\frac{L_{nt}}{K_n}\right)^{\eta^L} a_{nt} \qquad (G.37)$$

where, a_{nt} is production fundamentals capturing features of physical geography (e.g. access to natural water);
Production externalities are assumed to be a constant elasticity function of workplace employment density ($\frac{L_{nt}}{K_n}$);
K_n is geographical land area;
η^L is production elasticity;
The authors focus on a borough's own workplace employment density.

The authors assume that composite residential amenities $\mathbb{B}_{nt}$ depend on residential fundamentals and residential externalities.

$$\mathbb{B}_{nt} = \left(\frac{R_{nt}}{K_n}\right)^{\eta^R} b_{nt} \qquad (G.38)$$

where, b_{nt} is residential fundamentals capturing features of physical geography (e.g. green areas);
Residential externalities are assumed to be a constant elasticity function of residence employment density ($\frac{R_{nt}}{K_n}$);
η^R is residential elasticity;
The authors focus on a borough's own residence employment density.

Taking logarithms in equation (G.37):

$$ln\mathbb{A}_{nt}^T = \eta^L(lnL_{nt} - lnK_n) + lna_{nt}$$

30

differencing between an earlier year $\tau < t$ and the baseline year of $t = 1921$:

$$ln\mathbb{A}_{n\tau}^{T} - ln\mathbb{A}_{nt}^{T} = \eta^{L}lnL_{n\tau} - \eta^{L}lnK_{n} + lna_{n\tau} - \eta^{L}lnL_{nt} + \eta^{L}lnK_{n} - lna_{nt}$$

$$= \eta^{L}lnL_{n\tau} - \eta^{L}lnL_{nt} + lna_{n\tau} - lna_{nt}$$

Re-arrange it to get:
$$ln\hat{\mathbb{A}}_{nt}^{T} = \eta^{L}ln\hat{L}_{nt} + ln\hat{a}_{nt} \qquad (30-1)$$

in the paper:
$$ln\hat{\mathbb{A}}_{nt}^{T} = \varsigma^{L} + \eta^{L}ln\hat{L}_{nt} + ln\hat{a}_{nt} \qquad (30)$$

where, ς^{L} is constant that control for any factors that are common across all locations within Greater London, such as common changes in productivity, or changes in expected utility in the wider economy;
$ln\hat{a}_{nt}$ captures idiosyncratic shocks to production fundamentals.

Similarly, taking logarithms in equation (G.38):

$$ln\mathbb{B}_{nt} = \eta^{R}(lnR_{nt} - lnK_{n}) + lnb_{nt}$$

differencing between an earlier year $\tau < t$ and the baseline year of $t = 1921$:

$$ln\mathbb{B}_{n\tau} - ln\mathbb{B}_{nt} = \eta^{R}lnR_{n\tau} - \eta^{R}lnK_{n} + lnb_{n\tau} - \eta^{R}lnR_{nt} + \eta^{R}lnK_{n} - lnb_{nt}$$

Re-arrange it to get:
$$ln\hat{\mathbb{B}}_{nt} = \eta^{R}ln\hat{R}_{nt} + ln\hat{b}_{nt} \qquad (31-1)$$

in the paper:
$$ln\hat{\mathbb{B}}_{nt} = \varsigma^{R} + \eta^{R}ln\hat{R}_{nt} + ln\hat{b}_{nt} \qquad (31)$$

where, ς^{R} is constant that control for any factors that are common across all locations within Greater London, such as common changes in amenities, or changes in expected utility in the wider economy;
$ln\hat{b}_{nt}$ captures idiosyncratic shocks to residential fundamentals.

Remark 3. • *A key challenge in estimating the strength of agglomeration forces $(\varsigma^{L}, \varsigma^{R})$ in equations (30) and (31) is that workplace and residence employment are endogenous to productivity and amenities.*

• *To address this challenge, the authors use the quasi-experimental variation from the invention of the steam railway. In particular, the authors estimate equations (30) and (31) using 2SLS, instrumenting the log changes in workplace and residence employment with indicator variables for 5 km distance grid cells from the Guildhall, where the excluded category is locations more than 20 km away.*

31

- *Table II reports the estimation results: columns (1) and (2) report the results using OLS; columns (3) and (4) show the results using 2SLS. IV estimates are smaller than OLS estimates*

- *There is an upward bias in the estimated residential elasticity (η^R) and an upward bias in the estimated production elasticity (η^L) using OLS.*

9 COUNTERFACTUALS

The authors undertake these counterfactuals under a range of alternative assumptions about the floor space supply elasticity (μ) and the strength of agglomeration forces (η^L, η^R).

Given a change in commuting costs ($\hat{\kappa}_{rit}^{-\epsilon}$) from the removal of parts of the railway network and the observed values of variables ($\mathbb{Q}_{nt}, \lambda_{nt}^L, \lambda_{nt}^R, L_{Mt}, \lambda_{nit|n}^R, \lambda_{nit}, w_{it}$) in our initial equilibrium in year $t = 1921$, we solve for counterfactual changes in the following five endogenous variables:

- the workplace choice probability ($\hat{\lambda}_{nt}^L$);

- the residence choice probability ($\hat{\lambda}_{nt}^R$);

- the price of floor space ($\hat{Q}_{nt}$);

- the wage ($\hat{w}_{nt}$);

- total city employment ($\hat{L}_{Nt}$).

These five endogenous variables solve the following system of five equations for the general equilibrium of the model:

- the land market clearing condition (H.1):

$$\hat{\mathbb{Q}}_{nt}\mathbb{Q}_{nt} = \hat{Q}_{nt}^{1+\mu}\mathbb{Q}_{nt} = (1-\alpha)\left[\sum_{i \in N}\frac{\lambda_{nit|n}^R \hat{w}_{it}^\epsilon \hat{\kappa}_{nit}^{-\epsilon}}{\sum_{\ell \in N}\lambda_{n\ell\tau|n}^R \hat{w}_{\ell t}^\epsilon \hat{\kappa}_{n\ell t}^{-\epsilon}}\hat{w}_{it}w_{it}\right]\hat{R}_{nt}R_{nt} + \frac{\beta^H}{\beta^L}\hat{w}_{nt}w_{nt}\hat{L}_{nt}L_{nt}$$

$$= (1-\alpha)\left[\sum_{i \in N}\frac{\lambda_{nit|n}^R \hat{w}_{it}^\epsilon \hat{\kappa}_{nit}^{-\epsilon}}{\sum_{\ell \in N}\lambda_{n\ell\tau|n}^R \hat{w}_{\ell t}^\epsilon \hat{\kappa}_{n\ell t}^{-\epsilon}}\hat{w}_{it}w_{it}\right]\hat{\lambda}_{nt}^R \hat{\bar{L}}\lambda_{nt}^R\bar{L} + \frac{\beta^H}{\beta^L}\hat{w}_{nt}w_{nt}\hat{\lambda}_{nt}^L \hat{\bar{L}}\lambda_{nt}^L\bar{L}$$

$$= \left((1-\alpha)\left[\sum_{i \in N}\frac{\lambda_{nit|n}^R \hat{w}_{it}^\epsilon \hat{\kappa}_{nit}^{-\epsilon}}{\sum_{\ell \in N}\lambda_{n\ell\tau|n}^R \hat{w}_{\ell t}^\epsilon \hat{\kappa}_{n\ell t}^{-\epsilon}}\hat{w}_{it}w_{it}\right]\hat{\lambda}_{nt}^R \lambda_{nt}^R + \frac{\beta^H}{\beta^L}\hat{w}_{nt}w_{nt}\hat{\lambda}_{nt}^L\lambda_{nt}^L\right)\hat{\bar{L}}\bar{L} \qquad (H.1)$$

where recall that $\hat{x}_t = \frac{x_\tau}{x_t}$; we have used $\hat{\mathbb{Q}}_{nt} = \hat{Q}_{nt}\hat{H}_{nt} = \hat{Q}_{nt}^{1+\mu}$; Note that we assume that the homogeneous traded good is costlessly traded:

$$P_{nt}^T = P_t^T, \forall n \in N$$

- the zero-profit condition (H.2):

$$\hat{Q}_{nt} = \frac{Q_{n\tau}}{Q_{nt}} = \left(\frac{\mathbb{A}_{n\tau}^T}{\mathbb{A}_{nt}^T}\right)^{\frac{1}{\beta^H}} \frac{w_{n\tau}}{w_{nt}}^{\frac{-\beta^L}{\beta^H}} = \left(\hat{\mathbb{A}}_{nt}^T\right)^{\frac{1}{\beta^H}} \hat{w}_{nt}^{\frac{-\beta^L}{\beta^H}} \qquad (H.2)$$

- the workplace choice probability (H.3):

$$\hat{\lambda}_{nt}^L \lambda_{nt}^L = \frac{\sum_{i\in N} \lambda_{int} \hat{\mathbb{B}}_{it}^\epsilon \hat{w}_{nt}^\epsilon \hat{\kappa}_{int}^{-\epsilon} \hat{Q}_{it}^{-\epsilon(1-\alpha)}}{\sum_{k\in N} \sum_{\ell\in N} \lambda_{k\ell t} \hat{\mathbb{B}}_{kt}^\epsilon \hat{w}_{\ell t}^\epsilon \hat{\kappa}_{k\ell t}^{-\epsilon} \hat{Q}_{kt}^{-\epsilon(1-\alpha)}} \qquad (H.3)$$

- the residence choice probability (H.4):

$$\hat{\lambda}_{nt}^R \lambda_{nt}^R = \frac{\lambda_{n\ell t} \hat{\mathbb{B}}_{nt}^\epsilon \hat{Q}_{nt}^{-\epsilon(1-\alpha)} \widehat{RMA}^\epsilon}{\sum_{k\in N} \lambda_{k\ell t} \hat{\mathbb{B}}_{kt}^\epsilon \hat{Q}_{kt}^{-\epsilon(1-\alpha)} \widehat{RMA}^\epsilon}$$

$$= \frac{\sum_{\ell\in N} \lambda_{n\ell t} \hat{\mathbb{B}}_{nt}^\epsilon \hat{w}_{\ell t}^\epsilon \hat{\kappa}_{n\ell t}^{-\epsilon} \hat{Q}_{nt}^{-\epsilon(1-\alpha)}}{\sum_{k\in N} \sum_{\ell\in N} \lambda_{k\ell t} \hat{\mathbb{B}}_{kt}^\epsilon \hat{w}_{\ell t}^\epsilon \hat{\kappa}_{k\ell t}^{-\epsilon} \hat{Q}_{kt}^{-\epsilon(1-\alpha)}} \qquad (H.4)$$

- the population mobility condition (H.5):

$$\frac{\hat{L}_{Nt} L_{Nt}}{L_{Mt}} = \frac{L_{N\tau}}{L_{Mt}}$$

$$= \frac{\sum_{k\in N} \sum_{\ell\in N} (B_{k\ell\tau} w_{\ell\tau})^\epsilon \left(\kappa_{k\ell\tau} P_{k\tau}^\alpha Q_{k\tau}^{1-\alpha}\right)^{-\epsilon}}{\sum_{n\in M} \sum_{i\in M} (B_{nit} w_{it})^\epsilon \left(\kappa_{nit} P_{nt}^\alpha Q_{nt}^{1-\alpha}\right)^{-\epsilon}}$$

$$= \frac{\sum_{k\in N} \sum_{\ell\in N} \hat{B}_{k\ell t}^\epsilon \hat{w}_{\ell t}^\epsilon \hat{\kappa}_{k\ell t}^{-\epsilon} \hat{P}_{kt}^{-\epsilon\alpha} \hat{Q}_{kt}^{-\epsilon(1-\alpha)} B_{k\ell t}^\epsilon w_{\ell t}^\epsilon \kappa_{k\ell t}^{-\epsilon} P_{kt}^{-\epsilon\alpha} Q_{kt}^{-\epsilon(1-\alpha)}}{\sum_{n\in M} \sum_{i\in M} (B_{nit} w_{it})^\epsilon \left(\kappa_{nit} P_{nt} Q_{nt}^{1-\alpha}\right)^{-\epsilon}}$$

$$= \sum_{k\in N} \sum_{\ell\in N} \hat{B}_{k\ell t}^\epsilon \hat{w}_{\ell t}^\epsilon \hat{\kappa}_{k\ell t}^{-\epsilon} \hat{P}_{kt}^{-\epsilon\alpha} \hat{Q}_{kt}^{-\epsilon(1-\alpha)} \frac{B_{k\ell t}^\epsilon w_{\ell t}^\epsilon \kappa_{k\ell t}^{-\epsilon} P_{kt}^{-\epsilon\alpha} Q_{kt}^{-\epsilon(1-\alpha)}}{\sum_{n\in M} \sum_{i\in M} (B_{nit} w_{it})^\epsilon \left(\kappa_{nit} P_{nt} Q_{nt}^{1-\alpha}\right)^{-\epsilon}}$$

$$= \sum_{k\in N} \sum_{\ell\in N} \lambda_{k\ell t} \hat{B}_{k\ell t}^\epsilon \hat{w}_{\ell t}^\epsilon \hat{\kappa}_{k\ell t}^{-\epsilon} \hat{P}_{kt}^{-\epsilon\alpha} \hat{Q}_{kt}^{-\epsilon(1-\alpha)}$$

So we have

$$\left(\frac{\hat{L}_{Nt} L_{Nt}}{L_{Mt}}\right)^{\frac{1}{\epsilon}} = \left[\sum_{k\in N} \sum_{\ell\in N} \lambda_{k\ell t} \hat{\mathbb{B}}_{kt}^\epsilon \hat{w}_{\ell t}^\epsilon \hat{\kappa}_{k\ell t}^{-\epsilon} \hat{Q}_{kt}^{-\epsilon(1-\alpha)}\right]^{\frac{1}{\epsilon}} \qquad (H.5)$$

and we hold total employment and expected utility in the economy as a whole constant ($\hat{L}_{Mt} = 1$) and ($\hat{\bar{U}}_t = 1$).

According to

$$ln\hat{\mathbb{B}}_{nt} = \varsigma^R + \eta^R ln\hat{R}_{nt} + ln\hat{b}_{nt}$$

$$ln\hat{\mathbb{A}}^T_{nt} = \varsigma^L + \eta^L ln\hat{L}_{nt} + ln\hat{a}_{nt}$$

the relative changes in composite traded productivity($\hat{\mathbb{A}}^T_{nt}$) and composite amenities ($\hat{\mathbb{B}}_{nt}$) satisfy:

$$\hat{\mathbb{A}}^T_{nt} = \hat{L}_t^{\eta^L} \qquad (H.6)$$

$$\hat{\mathbb{B}}_{nt} = \hat{R}_t^{\eta^R} \qquad (H.7)$$

where we hold production and residential fundamentals constant ($\hat{a}_{nt} = 1$ and $\hat{b}_{nt} = 1$).

Ccounterfactual Predictions for Removing Entire Railway Network and Underground Railway Network

- In Figure IX, they display the model's counterfactual predictions for removing the entire railway network (left panels) and the underground railway network (right panels). Because the first underground line was built in 1863, the counterfactuals in the right panels are flat before 1861. We report results for net commuting into the City of London in the top two panels, the total population of Greater London in the middle two panels, and total rateable values in Greater London in the bottom two panels.

- "Baseline" shows the values of variables from our baseline quantitative analysis from Section VI;

- "Inelastic No Agglom" shows the values of variables from our counterfactual with a perfectly inelastic supply of floor space and exogenous productivity and amenities ($\mu = \eta^L = \eta^R = 0$);

- "Elastic No Agglom" shows the values of variables from our counterfactual using our calibrated floor space supply elasticity and exogenous productivity and amenities ($\mu = 1.83, \eta^L = \eta^R = 0$);

- "Elastic Prod Agglom" shows the values of variables from our counterfactual using our calibrated floor space supply elasticity, estimated production agglomeration forces, and exogenous amenities ($\mu = 1.83, \eta^L = 0.086, and \eta^R = 0$);

- "Elastic Prod + Amen Agglom" shows the values of variables from our counterfactual using our calibrated floor space supply elasticity and estimated production and residential agglomeration forces ($\mu = 1.83, \eta^L = 0.086, and \eta^R = 0.172$).

Results of Comparisons of Economic Impact of Removal of Railway Network with Its Construction Costs

- In Table III, we report the results of these comparisons of the economic impact of the removal of the railway network with its construction costs. In the top panel, we remove the entire railway network. In the middle panel, we eliminate only the underground railway network. In the bottom panel, we examine the extent to which there were diminishing returns to the construction of the railway network by only removing overground and underground railway lines constructed in the final decade of our sample, from 1911–1921.

- As shown in Table III, regardless of whether we assume a 3% or 5% discount rate, and irrespective of whether we consider the entire railway network (top panel) or only the underground railway network (middle panel), we find ratios of reductions in the net present value of rateable values to construction costs that are substantially greater than 1.

10 Conclusions

The authors use the natural experiment of the invention of the steam railway to provide evidence on the role of modern transport technologies in underpinning the concentration of economic activity in large metropolitan areas. The key idea underlying our approach is that a reduction in commuting costs facilitates an increased separation of workplace and residence. In particular, the hub-and spoke structure of the railway disproportionately reduced commuting costs into central locations. If these central locations have high productivity relative to amenities compared with the suburbs, this transport improvement leads them to specialize as a workplace while the suburbs specialize as a residence. In the presence of agglomeration forces, this increased concentration of workers in the center and residents in the suburbs further reinforces these differences in productivity and amenities across locations.

This paper first provides evidence for simplified analysis of the mechanism: 1. With the construction of the railway network, the population of London decreased while that of the suburbs increased; 2. The decline in the population of The City of London combined with the increase in employment; 3. The professional shift from residential to workplace has made the City of London a relatively valuable area. Other metropolitan areas showed the same pattern of results as transportation improved in the 19th century. Using parochial level data for London, this paper found no deviation between population growth and parochial time trends prior to railway arrival, nor evidence of significant deviation from trends after railway arrival.

To further explain the results, the authors established a set of urban spatial structure model which satisfies the gravity equation of commuting flow and has a fixed ratio between labor income and ground space payment. Although only observed in the sample of 1921 bilateral commuting data, the author still show that the framework can be used to estimate all the way back to the early 19th century, the influence of the rail network it has successfully captured since the middle of the 19th century London city population sharply divided between day and night, and copy the early commuting in the data the characteristics of

most people living near their workplace. In addition, this paper combines the traditional city model with this model, and uses a series of identification conditions to estimate the agglomeration effect of production and residence. In a counterfactual analysis of the removal of the rail network, the authors also found that changes in commuting costs alone explained most of the observed segregation. Behind, based on the supply and productivity for the ground space and amenities of different assumptions, further investigates the net for London commuter railway construction to the total population, population changes in the rateable value, finally has carried on the cost-benefit analysis, emphasizes the railway network construction for the important role of London's economic development cannot be ignored.