Commuting, Migration and Local Employment Elasticities[*]

Ferdinando Monte[†]
Georgetown University

Stephen J. Redding[‡]
Princeton University

Esteban Rossi-Hansberg[§]
Princeton University

April 24, 2015

Abstract

Many changes in the economic environment are local, including policy changes and infrastructure investments. The effect of these changes depends crucially on the ability of factors to move in response. Therefore a key object of interest for policy evaluation and design is the *elasticity of local employment* to these changes in the economic environment. We develop a quantitative general equilibrium model that incorporates spatial linkages between locations in goods markets (trade) and factor markets (commuting and migration). We find substantial heterogeneity across locations in local employment elasticities. We show that this heterogeneity can be well explained with theoretically motivated measures of commuting flows. Without taking into account this dependence, estimates of the local employment elasticity lack external validity. We also find that commuting flows and their importance cannot be accounted for with standard measures of size or wages at the country or commuting zone levels.

[†]Carey Business School, 100 International Drive, Baltimore, MD 21202. Email: ferdinando.monte@jhu.edu.

[‡]Dept. Economics and WWS, Fisher Hall, Princeton, NJ 08544. Tel: 609 258 4016. Email: reddings@princeton.edu.

[§]Dept. Economics and WWS, Fisher Hall, Princeton, NJ 08544. Tel: 609 258 4024. Email: erossi@princeton.edu.

1 Introduction

Agents spend about 8% of their workday commuting to and from work.[1] They make this significant daily investment, to live and work in different locations, so as to balance their living costs and residential amenities with the wage they can obtain at their place of employment. The ability of firms in a location to attract workers depends, therefore, not only on the ability to attract local residents through migration, but also on the ability to attract commuters from other, nearby, locations. Together, the response of migration and commuting to any local shock, including regulatory changes and infrastructure investments, determine the *local employment elasticity*. This elasticity is of great policy interest since it determines the effectiveness of local policy. Estimating its magnitude as a response to a variety of shocks (such as aggregate industry shocks, the discovery of natural resources, financial crises and regression discontinuities associated with state policy interventions) has been the main concern of a large empirical literature. In this paper we explore its determinants and characteristics using a detailed quantitative spatial equilibrium theory.

We develop a quantitative spatial general equilibrium model that incorporates spatial linkages between locations in both goods markets (trade) and factor markets (commuting and migration). We show that there is no single local employment elasticity. Instead the local employment elasticity is an endogenous variable that differs across locations depending on their linkages to one another in goods and factor markets. Calibrating our model to county-level data for the United States, we find that the elasticity of local employment with respect to local productivity shocks varies from close to zero to more than three. Therefore an average local employment elasticity estimated from cross-section data can be quite misleading when used to predict the impact of a local shock on any individual county and can lead to substantial under or overprediction of the effect of the shock. We use our quantitative model to understand the systematic determinants of the local employment elasticity and argue that a large part of the variation results from differences in commuting links between a location and its neighbors. This allows us to propose variables that can be included in reduced-form regressions to improve their ability to predict the heterogeneity in local employment responses. Of course, one needs to use the full general equilibrium model to quantify the total impact of a shock and its implications for welfare, using counterfactuals that use only observed values of variables in an initial equilibrium.

Our theoretical framework allows for an arbitrary number of locations that can differ in productivity, amenities and geographical relationship to one another. The spatial distribution of economic activity is driven by a tension between productivity differences and home market effects (forces for the concentration of economic activity) and an inelastic supply of land and commuting costs (dispersion forces). Commuting allows workers to access high productivity employment locations without having to live there and hence alleviates the congestion effect in such high productivity locations. We show that these commuting flows between locations exhibit a gravity equation relationship with

---

[1]See Redding and Turner (2014).

a much higher distance elasticity than for goods flows, suggesting that moving people is more costly than moving goods across geographic space. We discipline our quantitative spatial model to match these observed gravity equation relationships for goods and commuting flows as well as the observed cross-section distributions of employment, residents and wages across U.S. counties. Given observed data on wages, employment by workplace, commuting flows and land area, and a parameterization of trade and commuting costs, we show that the model can be used to recover unique values of the unobserved location fundamentals (productivity and amenities) that exactly rationalize the observed data as an equilibrium of the model.

We use our model to undertake three quantitative exercises that shed light on spatial linkages in goods and factor markets. First, we provide evidence on the distribution of local employment elasticities across counties. We do so by calculating 3,111 counterfactual exercises, each shocking one county in the U.S. with a 5 percent productivity shock. We find substantial heterogeneity in local employment elasticities across counties. This heterogeneity remains even when we aggregate counties into Commuting Zones (CZ's), that are intended to better approximate local labor market areas. We show that the main determinant of this large heterogeneity is the observed patterns of commuting in the equilibrium prior to the productivity shock. That is, the response of a county to a local shock depends crucially on its observed commuting ties to other counties. Local employment elasticities estimated in a particular location, that do not account for this dependence, will in general not apply in another region or context. Taken together, these results highlight that locations within countries are not independent units in cross-section regression relationships but are rather systematically linked to one another in goods and factor markets in a quantitatively significant way.

Given that the heterogeneity in local employment elasticity depends crucially on commuting links, our next step is to explore these links more thoroughly. We find that commuting intensities cannot easily be accounted for by standard measures of the size of economic activity in a county or its neighbors. So to gauge further the importance and determination of commuting, in a second exercise we use our quantitative model to evaluate the welfare gains from commuting by considering a counterfactual with prohibitive commuting costs. The commuting technology facilitates a separation of workplace and residence, enabling people to work in relatively high productivity locations and live in relatively high amenity locations. Therefore eliminating commuting leads to a dispersal of employment from locations that were net importers of commuters in the initial equilibrium (e.g. Manhattan) towards those locations that were initially net exporters (e.g. parts of New Jersey). We find that these reallocations are substantial, with Manhattan experiencing a decline in employment of over 70 percent, while parts of New Jersey experience increases in employment of over 90 percent.

The above logic seems to suggest that commuting might be important only for areas that encompass the larger cities in the U.S. This is, however, not the case. Although the changes in employment as a result of eliminating commuting are well explained by initial commuting intensity, this intensity is not well accounted for by either county or commuting zone size. The results underscore again the relevant information embedded in commuting links. Overall, we find a welfare cost of eliminat-

ing commuting of around 5 percent of real GDP, which is comparable to standard estimates of the welfare gains from international trade for a country of a similar size to the United States.

The previous literature in trade and economic geography has mostly abstracted from commuting when analyzing the gains from trade within or across countries. However, given the importance of commuting links in shaping the distribution of economic activity across counties, it is natural to expect that these links also determine the magnitude of the impact of reductions in trade costs. Hence, in a third exercise we investigate the interaction between trade and commuting costs. We compare the counterfactual effects of a 20 percent reduction of trade costs in the actual world with commuting to the effects in a hypothetical world without commuting. Here we find a subtle pattern of results. On the one hand, trade and commuting are weak substitutes in terms of aggregate welfare, in the sense that the reduction in trade costs has a larger effect on welfare in the hypothetical world without commuting (4.7 percent) than in the actual world with commuting (4.4 percent). The intuition is that commuting costs are a force for the dispersion of economic activity. Therefore increasing commuting costs to prohibitive levels leads to greater dispersion of economic activity. In such a hypothetical world with a more dispersed distribution of economic activity, reductions in trade costs have a larger effect on welfare than in the actual world with commuting.

On the other hand, lower trade costs increase the access of high productivity locations to the U.S. market. In a world without commuting, the ability of these locations to increase employment is limited by their ability to attract residents in the face of congestion costs from an inelastic supply of land. In a world with commuting, these locations can increase employment by attracting residents to neighboring locations, thereby alleviating congestion costs. Therefore, for the most congested locations, we find that reductions in trade costs lead to larger increases in employment with commuting than without commuting. Indeed, this effect can be so strong that these locations experience increases in employment in the actual world with commuting but reductions in employment in the hypothetical world without commuting. These results demonstrate how abstracting from commuting might lead to distorted conclusions on the impact of trade cost reductions and highlight the richness of the spatial linkages between locations.

Our paper is related to a number of empirical literatures. In international trade, our work relates to quantitative theoretical models of costly trade in goods following Eaton and Kortum (2002) and extensions. In economic geography, our research contributes to the literature on goods trade and factor mobility, including Krugman (1991), Hanson (1996, 2005), Helpman (1998), Fujita et al. (1999), Rossi-Hansberg (2005), Redding and Sturm (2008), Redding (2012), Moretti and Klein (2014), Allen and Arkolakis (2014a), Caliendo et al. (2014) and Desmet and Rossi-Hansberg (2014).

In urban economics, our analysis builds on a long line of research on costly trade in people (commuting), including Alonso (1964), Mills (1967), Muth (1969), Lucas and Rossi-Hansberg (2002), Desmet and Rossi-Hansberg (2013), Ahlfeldt, Redding Sturm and Wolf (2015), Monte (2015) and Allen, Arkolakis and Li (2015). In public finance and labor economics, our quantitative analysis connects with an empirical literature that has estimated the local incidence of labor demand shocks,

including Bartik (1991), Blanchard and Katz (1992), Bound and Holzer (2000), Greenstone, Hornbeck and Moretti (2010), Michaels (2011), Moretti (2011), Busso, Gregory and Kline (2013), Autor, Dorn and Hanson (2013), Diamond (2013), Notowidigdo (2013) and Yagan (2014) among others.

Unlike most of the above studies, we develop a quantitative general equilibrium spatial model that incorporates both costly trade in goods and costly mobility and commuting of people between locations. Through incorporating these spatial linkages in both goods and factor markets, our theoretical and empirical framework unifies research on systems of cities (with trade and migration between cities) and the internal organization of economic activity within cities (with commuting within cities). We discipline our quantitative model to match the observed gravity equation relationships for goods trade and commuting flows and to rationalize the observed cross-section distribution of employment, residents and wages across U.S. counties. The result is a quantitative framework that can be used to evaluate both local effects of shocks and their spillovers to other counties through general equilibrium linkages in goods and factor markets.

The remainder of the paper is structured as follows. Section 2 develops our theoretical framework. Section 3 discusses data and the quantification of the model using the U.S. data and reports summary statistics on commuting between counties. Section 4 uses the model to undertake three sets of counterfactuals that highlight the importance of spatial linkages in goods and factor markets. We examine the local incidence of productivity shocks (local labor demand shocks), the effects of changes in commuting technology, and the interaction between changes in trade and commuting costs. Section 5 summarizes our conclusions. An appendix at the end contains the detailed derivations of theoretical results and the proofs of propositions.

2 The Model

We consider an economy that consists of a set of locations $n, i \in N$. These locations are linked in goods markets through costly trade and in factor markets through migration and costly commuting. The economy as a whole is populated by a measure $\bar{L}$ of workers who are endowed with one unit of labor that is supplied inelastically. Each location n is endowed with an inelastic supply of land (H_n).

2.1 Preferences and Endowments

Workers are geographical mobile and have heterogeneous preferences for locations. Each worker chooses a pair of residence and workplace locations to maximize their utility taking as given the choices of other firms and workers.[2] The preferences of a worker ω who lives and consumes in region n and works in region i are defined over final goods consumption ($C_{n\omega}$), residential land use ($H_{n\omega}$), an idiosyncratic amenities shock ($b_{ni\omega}$) and commuting costs (κ_{ni}), according to the Cobb-Douglas

---

[2]Throughout the following we use n to denote a worker's location of residence and consumption and i to denote a worker's location of employment and production, unless otherwise indicated.

form,[3]

$$U_{ni\omega} = \frac{b_{ni\omega}}{\kappa_{ni}} \left(\frac{C_{n\omega}}{\alpha}\right)^{\alpha} \left(\frac{H_{n\omega}}{1-\alpha}\right)^{1-\alpha}, \tag{1}$$

where $\kappa_{ni} \in [1, \infty)$ is an iceberg commuting cost in terms of utility. The idiosyncratic amenities shock ($b_{ni\omega}$) captures the idea that individual workers can have idiosyncratic reasons for living and working in different locations. We model this heterogeneity in amenities following McFadden (1974) and Eaton and Kortum (2002).[4] For each worker ω living in location n and working in location i, idiosyncratic amenities ($b_{ni\omega}$) are drawn from an independent Fréchet distribution,

$$G_{ni}(b) = e^{-B_{ni}b^{-\epsilon}}, \qquad B_{ni} > 0, \epsilon > 1, \tag{2}$$

where the scale parameter B_{ni} determines the average amenities from living in location n and working in location i, and the shape parameter $\epsilon > 1$ controls the dispersion of amenities. This idiosyncratic amenities shock implies that different workers make different choices about their workplace and residence locations when faced with the same prices and wages. All workers ω residing in region n and working in region i receive the same wage and make the same consumption and residential land choices. Hence we suppress the implicit dependence on ω except where important.

The goods consumption index in location n takes the constant elasticity of substitution (CES) or Dixit-Stiglitz form and is defined over a continuum of varieties sourced from each location i,

$$C_n = \left[\sum_{i \in N} \int_0^{M_i} c_{ni}(j)^{\rho} dj\right]^{\frac{1}{\rho}}, \qquad \sigma = \frac{1}{1-\rho} > 1. \tag{3}$$

Equilibrium consumption in location n of each variety sourced from location i is

$$c_{ni}(j) = \alpha X_n P_n^{\sigma-1} p_{ni}(j)^{-\sigma}, \tag{4}$$

where X_n is aggregate expenditure in location n; P_n is the price index dual to (3), and $p_{ni}(j)$ is the "cost inclusive of freight" price of a variety produced in location i and consumed in location n.

We assume that land is owned by absentee landlords and land income is not spent within the economy.[5] Land market clearing implies that the equilibrium land price (Q_n) can be determined from the requirement that land income equals total expenditure on land

$$Q_n = (1-\alpha) \frac{\bar{v}_n L_{Rn}}{H_n}, \tag{5}$$

where $\bar{v}_n$ is the average income of residents; H_n is the inelastic supply of land; and L_{Rn} is the measure of residents.

---

[3]For empirical evidence using U.S. data in support of the constant housing expenditure share implied by the Cobb-Douglas functional form, see Davis and Ortalo-Magne (2011).

[4]A long line of research models location decisions using preference heterogeneity, as in Artuc, Chaudhuri and McClaren (2010), Kennan and Walker (2011), Grogger and Hanson (2011), Moretti (2011) and Busso, Gregory and Kline (2013).

[5]Although we consider absentee landlords for simplicity, it is straightforward to consider alternative specifications where (i) landlords are a separate factor and consume where they live, (ii) expenditure on land is redistributed lump-sum to the residents of a location, (iii) expenditure on land accrues to a global fund and workers hold shares in this global fund. Although these alternatives are straightforward to consider, they have different normative implications, because they imply different pecuniary externalities from workers' location choices (see Caliendo et al. 2014).

2.2 Production

Production is modelled as in the new economic geography literature following Krugman (1991) and Helpman (1998). Varieties are produced under conditions of monopolistic competition. To produce a variety, a firm must incur a fixed cost of F units of labor and a constant variable cost in terms of labor that depends on a region's productivity A_i.[6] Therefore the total amount of labor ($l_i(j)$) required to produce $x_i(j)$ units of a variety j in region i is[7]

$$l_i(j) = F + \frac{x_i(j)}{A_i}. \tag{6}$$

Profit maximization implies that equilibrium prices are a constant mark-up over marginal cost, namely,

$$p_{ni}(j) = \left(\frac{\sigma}{\sigma - 1}\right) \frac{d_{ni} w_i}{A_i}, \tag{7}$$

where w_i is the wage in region i. Profit maximization and zero profits imply that equilibrium output of each variety is equal to

$$x_i(j) = A_i F (\sigma - 1). \tag{8}$$

A constant equilibrium output of each variety and labor market clearing then imply that the total measure of produced varieties (M_i) is proportional to its measure of employed workers (L_{Mi}),

$$M_i = \frac{L_{Mi}}{\sigma F}. \tag{9}$$

2.3 Goods Trade

The model implies a gravity equation for bilateral trade between regions. Using the CES expenditure function, equilibrium prices (7) and the measure of firms in (9), the share of region n's expenditure on goods produced in region i is

$$\pi_{ni} = \frac{M_i p_{ni}^{1-\sigma}}{\sum_{k \in N} M_k p_{nk}^{1-\sigma}} = \frac{L_{Mi} (d_{ni} w_i / A_i)^{1-\sigma}}{\sum_{k \in N} L_{Mk} (d_{nk} w_k / A_k)^{1-\sigma}}. \tag{10}$$

Therefore trade between regions n and i depends on bilateral trade costs (d_{ni}) in the numerator ("bilateral resistance") and on trade costs to all possible sources of supply k in the denominator ("multilateral resistance"). Trade balance then implies that total workplace income in each location equals total expenditure on goods produced in that location, namely,[8]

$$w_i L_{Mi} = \sum_{n \in N} \pi_{ni} \bar{v}_n L_{Rn}. \tag{11}$$

---

[6]We assume a representative firm within each location. However, it is straightforward to generalize the analysis to introduce firm heterogeneity following Melitz (2003), where all firms entering in location i draw a productivity φ from a Pareto distribution $g(\varphi)$ that can be used to produce varieties in location i.

[7]We abstract from commercial land use, although it is straightforward to extend the model to consider the case where land is used both for residential and commercial purposes.

[8]Although we refer to trade balance as total workplace income equals total expenditure on goods produced in a location, total residential income need not equal total workplace income (because of commuting). Therefore total workplace income need not equal total residential expenditure, which implies that total exports need not equal total imports.

Using equilibrium prices (7) and labor market clearing (9), the price index dual to the consumption index (3) can be expressed as

$$P_n = \frac{\sigma}{\sigma - 1} \left(\frac{1}{\sigma F}\right)^{\frac{1}{1-\sigma}} \left[\sum_{i \in N} L_{Mi} \left(d_{ni} w_i / A_i\right)^{1-\sigma}\right]^{\frac{1}{1-\sigma}} = \frac{\sigma}{\sigma - 1} \left(\frac{L_{Mn}}{\sigma F \pi_{nn}}\right)^{\frac{1}{1-\sigma}} \frac{d_{nn} w_n}{A_n}. \tag{12}$$

where the second equality uses (10) to write the price index (12) as in the class of models considered by Arkolakis and Allen (2014b). This class of models includes increasing returns models such as Krugman (1991) and constant returns models such as versions of Armington (1969) and Eaton and Kortum (2002).

One difference between the increasing returns model of new economic geography considered here and the constant returns versions of Armington (1969) and Eaton and Kortum (2002) is that here the measure of employed workers in each region (L_{Mi}) enters the trade share (π_{ni}) and price index (P_n). This role for the measure of employed workers reflects consumer love of variety and the endogenous measure of varieties. As more agents choose to work in a region, this increases the measure of varieties produced in that region, which increases the share of consumer expenditure allocated to that region and reduces the consumer price index (this is the pecuniary externality highlighted by the New Economic Geography Literature). A similar role for the measure of employed workers arises in versions of Armington (1969) and Eaton and Kortum (2002) augmented to include external economies of scale, as considered further below.

2.4 Labor Mobility and Commuting

Workers are geographically mobile and choose their pair of residence and workplace locations to maximize their utility. Given our specification of preferences (1), the indirect utility function for a worker ω residing in region n and working in region i is

$$U_{ni\omega} = \frac{b_{ni\omega} w_i}{\kappa_{ni} P_n^\alpha Q_n^{1-\alpha}}. \tag{13}$$

Indirect utility is a monotonic function of idiosyncratic amenities ($b_{ni\omega}$) and these amenities have a Fréchet distribution. Therefore, the indirect utility for a worker living in region n and working in region i also has a Fréchet distribution given by

$$G_{ni}(U) = e^{-\Psi_{ni} U^{-\epsilon}}, \qquad \Psi_{ni} = B_{ni} \left(\kappa_{ni} P_n^\alpha Q_n^{1-\alpha}\right)^{-\epsilon} w_i^\epsilon. \tag{14}$$

Each worker selects the bilateral commute that offers her the maximum utility, where the maximum of Fréchet distributed random variables is itself Fréchet distributed. Using these distributions of utility, the probability that a worker chooses to live in location n and work in location i is

$$\lambda_{ni} = \frac{B_{ni} \left(\kappa_{ni} P_n^\alpha Q_n^{1-\alpha}\right)^{-\epsilon} w_i^\epsilon}{\sum_{r \in N} \sum_{s \in N} B_{rs} \left(\kappa_{rs} P_r^\alpha Q_r^{1-\alpha}\right)^{-\epsilon} w_s^\epsilon} = \frac{\Phi_{ni}}{\Phi}. \tag{15}$$

The idiosyncratic shock to preferences $b_{ni\omega}$ implies that individual workers choose different bilateral commutes when faced with the same prices (P_n, Q_n, w_i), commuting costs (κ_{ni}) and location characteristics (B_{ni}). Other things equal, workers are more likely to live in location n and work in location i, the lower the consumption goods price index (P_n) and land prices (Q_n) in n, the higher the wages (w_i) in i , the more attractive the amenities for this bilateral commute (B_{ni}), and the lower the commuting costs (κ_{ni}).

Summing these probabilities across workplaces i for a given residence n, we obtain the overall probability that a worker resides in location n (λ_{Rn}). Similarly, summing across residences n for a given workplace i, we obtain the overall probability that a worker works in location i (λ_{Mi}). So,

$$\lambda_{Rn} = \frac{L_{Rn}}{\bar{L}} = \sum_{i\in N} \lambda_{ni} = \sum_{i\in N} \frac{\Phi_{ni}}{\Phi}, \qquad \text{and} \qquad \lambda_{Mi} = \frac{L_{Mn}}{\bar{L}} = \sum_{n\in N} \lambda_{ni} = \sum_{n\in N} \frac{\Phi_{ni}}{\Phi}, \tag{16}$$

where labor market clearing corresponds to $\sum_n \lambda_{Rn} = \sum_i \lambda_{Mi} = 1$.

The expected income of a worker living in n is a combination of the wages in all the nearby locations. To construct it, note first that the probability that a worker commutes to location i conditional on living in location n is

$$\lambda_{ni|n} = \frac{B_{ni} (w_i/\kappa_{ni})^\epsilon}{\sum_{s\in N} B_{ns} (w_s/\kappa_{ns})^\epsilon}. \tag{17}$$

Note that equation (17) implies that, as with trade flows, commuting flows exhibit a gravity equation with an elasticity of commuting flows with respect to commuting cost (κ_{ni}) given by $-\epsilon$. The probability of commuting to location i conditional on living in location n depends on the wage (w_i), amenities (B_{ni}) and commuting costs for workplace i in the numerator ("bilateral resistance") as well as the wage (w_s), amenities (B_{ns}) and commuting costs (κ_{ns}) for all other possible workplaces s in the denominator ("multilateral resistance").

Using these conditional commuting probabilities, we obtain the following labor market clearing condition that equates the measure of workers employed in location i (L_{Mi}) with the measure of workers choosing to commute to location i, namely,

$$L_{Mi} = \sum_{n\in N} \lambda_{ni|n} L_{Rn}, \tag{18}$$

where L_{Rn} is the measure of residents in location n. Expected worker income conditional on living in location n is then equal to the wages in all possible workplaces weighted by the probabilities of commuting to those workplaces conditional on living in n, or

$$\bar{v}_n = \sum_{i\in N} \lambda_{ni|n} w_i. \tag{19}$$

Expected worker income ($\bar{v}_n$) is high in locations that have low commuting costs (low κ_{ni}) to high-wage employment locations.

Finally, population mobility implies that expected utility is the same for all pairs of residence and workplace and equal to expected utility for the economy as a whole. That is,

$$\bar{U} = \mathbb{E}[U_{ni\omega}] = \Gamma\left(\frac{\epsilon-1}{\epsilon}\right) \left[\sum_{r\in N}\sum_{s\in N} B_{rs} \left(\kappa_{rs} P_r^\alpha Q_r^{1-\alpha}\right)^{-\epsilon} w_s^\epsilon\right]^{\frac{1}{\epsilon}} \quad \text{all } n, i \in N, \tag{20}$$

9

where $\mathbb{E}$ is the expectations operator and the expectation is taken over the distribution for the idiosyncratic component of utility and $\Gamma(\cdot)$ is the Gamma function.

Although expected utility is equalized across all pairs of residence and workplace locations, real wages differ as a result of preference heterogeneity. Workplaces and residences face upward-sloping supply functions for workers and residents respectively (the choice probabilities (16)). Each workplace must pay higher wages to increase commuters' real income and attract additional workers with lower idiosyncratic amenities for that workplace. Similarly, each residential location must offer a lower cost of living to increase commuters' real income and attract additional residents with lower idiosyncratic amenities for that residence. Bilateral commutes with attractive characteristics (high workplace wages and low residence cost of living) attract additional commuters with lower idiosyncratic amenities until expected utility (taking into account idiosyncratic amenities) is the same across all bilateral commutes.

Our quantitative framework yields a simple expression for counterfactual changes in welfare in terms of observable sufficient statistics, including each location's domestic trade and commuting shares. From expected utility (20) and commuting probabilities (15), welfare in any given location n can be expressed as

$$\bar{U} = \Gamma\left(\frac{\epsilon - 1}{\epsilon}\right) \frac{w_n}{\kappa_{nn} \cdot P_n^\alpha Q_n^{1-\alpha}} \left(\frac{B_{nn}}{\lambda_{nn}}\right)^{\frac{1}{\epsilon}}. \tag{21}$$

Hence, the relative change in this common level of expected utility following a counterfactual change in the model's exogenous variables is given by

$$\widehat{U} = \frac{1}{\widehat{\kappa}_{nn}} \left(\frac{\widehat{A}_n}{\widehat{d}_{nn}}\right)^\alpha \left(\frac{\widehat{B}_{nn}}{\widehat{\lambda}_{nn}}\right)^{\frac{1}{\epsilon}} \left(\frac{1}{\widehat{\pi}_{nn}}\right)^{\frac{\alpha}{\sigma-1}} \left(\frac{\widehat{w}_n}{\widehat{v}_n}\right)^{1-\alpha} \frac{\widehat{L}_{Mn}^{\frac{\alpha}{\sigma-1}}}{\widehat{L}_{Rn}^{1-\alpha}}. \tag{22}$$

This equation identifies the mechanisms through which welfare can change in spatial equilibrium through linkages in goods and factor markets.

2.5 General Equilibrium

The general equilibrium of the model can be referenced by the following vector of six variables $\{w_n, \bar{v}_n, Q_n, L_{Mn}, L_{Rn}, P_n\}_{n=1}^N$ and a scalar $\bar{U}$. Given this equilibrium vector and scalar, all other endogenous variables of the model can be determined. This equilibrium vector solves the following six sets of equations: trade balance in each location (11), average residential income (19), land market clearing (9), workplace choice probabilities ((16) for L_{Mn}), residence choice probabilities ((16) for L_{Rn}), and price indices (12). The last condition needed to determine the scalar $\bar{U}$ is the labor market clearing condition, $\bar{L} = \sum_{n \in N} L_{Rn} = \sum_{n \in N} L_{Mn}$.

Proposition 1 (*Existence and Uniqueness*) *If*

$$\frac{1 + \varepsilon}{1 + (1 - \alpha)\varepsilon} < \sigma$$

there exists a unique general equilibrium of this economy.

10

Proof. See appendix. ∎

The condition for the existence of a unique general equilibrium in Proposition 1 is a generalization of the condition in the Helpman (1998) model to incorporate commuting and heterogeneity in worker preferences over locations. Defining $\tilde{\alpha} = \alpha / (1 + 1/\epsilon)$, this condition for a unique general equilibrium can be written as $\sigma(1 - \tilde{\alpha}) > 1$. Assuming prohibitive commuting costs ($\kappa_{ni} \to \infty$ for $n \neq i$) and taking the limit of no heterogeneity in worker preference over locations ($\epsilon \to \infty$), this reduces to the Helpman (1998) condition for a unique general equilibrium of $\sigma(1 - \alpha) > 1$.

We follow the new economic geography literature in modeling agglomeration forces through love of variety and increasing returns to scale. But the system of equations for general equilibrium in our new economic geography model is isomorphic to that in an Eaton and Kortum (2002) model with external economies of scale (as in Redding 2012) or an Armington (1969) model with external economies of scale (as in Arkolakis and Allen 2014a and Arkolakis, Allen and Li 2015).

Proposition 2 *(Isomorphisms) The system of equations for general equilibrium in our new economic geography model with agglomeration forces through love of variety and increasing returns to scale is isomorphic to that in an Eaton and Kortum (2002) model with external economies of scale and an Armington (1969) model with external economies of scale.*

Proof. See appendix. ∎

We focus on versions of these models that follow a long line of research in spatial economics in modelling congestion forces through residential land use and an inelastic supply of land. But the same system of equations for general equilibrium can be generated by models that instead directly assume a congestion force (e.g. by assuming that utility in each location depends on the measure of residents in that location).

2.6 Computing Counterfactuals

We use our quantitative framework to solve for the counterfactual effects of changes in the exogenous variables of the model (productivity A_n, amenities B_{ni}, commuting costs κ_{ni}, and trade costs d_{ni}) without having to necessarily determine the unobserved values of these exogenous variables. Instead in the Appendix we show that the system of equations for the counterfactual changes in the endogenous variables of the model can be written solely in terms of the observed values of variables in an initial equilibrium (employment L_{Mi}, residents L_{Ri}, workplace wages w_n, average residential income $\bar{v}_n$, trade shares π_{ni}, and commuting probabilities λ_{ni}). This approach uses observed bilateral commuting probabilities to capture unobserved bilateral commuting costs and amenities. Similarly, if bilateral trade shares between locations are available, they can be used to capture unobserved bilateral trade frictions (as in Dekle, Eaton and Kortum 2007). However, since bilateral trade data are only available at a higher level of aggregation than the counties we consider in our data, we make

some additional parametric assumptions to solve for implied bilateral trade shares between counties, as discussed below.

In the model, we assume that trade is balanced so income equals expenditure. However, when taking the model to the data, we allow for intertemporal trade deficits that are treated as exogenous in our counterfactuals, as in Dekle, Eaton and Kortum (2007) and Caliendo and Parro (2015).

3   Data and Measurement

Our empirical analysis combines data from a number of different sources for the United States. From the Commodity Flow Survey (CFS), we use data on bilateral trade and distances shipped for 123 CFS regions. From the American Community Survey (ACS), we use data on commuting probabilities between counties. From the Bureau of Economic Analysis (BEA), we use data on employment and wages by workplace. We combine these data sources with a variety of other Geographical Information Systems (GIS) data.

We use our data on employment and commuting to calculate the implied number of residents and their average income by county. First, from commuting market clearing (18), we obtain the number of residents (L_{Rn}) using data on the number of workers (L_{Mn}) and commuting probabilities conditional on living in each location ($\lambda_{ni|n}$). Second, we use these conditional commuting probabilities, together with county wages, to obtain average residential income ($\bar{v}_n$) as defined in (19).

3.1   Gravity in Goods Trade

One of the key predictions of the model is a gravity equation for goods trade. We observe bilateral trade between 123 CFS regions but not bilateral trade between counties.[9] To implement the model quantitatively at the county level, we use the trade balance condition (11),

$$w_i L_{Mi} - \sum_{n \in N} \frac{L_{Mi} \left(d_{ni} w_i / A_i \right)^{1-\sigma}}{\sum_{k \in N} L_{Mk} \left(d_{nk} w_k / A_k \right)^{1-\sigma}} \bar{v}_n L_{Rn} = 0, \tag{23}$$

where we observe (or have solved for) wages (w_i), employment (L_{Mi}), average residential income ($\bar{v}_i$) and residents (L_{Ri}).

Given the elasticity of substitution (σ), the observed data (w_i, L_{Mi}, $\bar{v}_i$, L_{Ri}) and a parameterization of trade costs ($d_{ni}^{1-\sigma}$), equation (23) provides a system of N equations that can be solved for a unique vector of N unobserved productivities (A_i). We prove this formally in the next proposition.

Proposition 3 *(Productivity Inversion) Given the elasticity of substitution (σ), the observed data on wages, employment, average residential income and residents $\{w_i, L_{Mi}, \bar{v}_i, L_{Ri}\}$, and a parameterization of trade costs ($d_{ni}^{1-\sigma}$), there exist unique values of the unobserved productivities (A_i) for each location i that are consistent with the data being an equilibrium of the model.*

---

[9]Other recent studies using the CFS data include Caliendo et. al (2014), Duranton, Morrow and Turner (2014).and Dingel (2015). The CFS is a random sample of plant shipments within the United States (foreign trade shipments are not included). CFS regions are the smallest geographical units for which this random sample is representative, which precludes constructing bilateral trade flows between smaller geographical units using the sampled shipments.

Proof. See the appendix. ∎

The resulting solutions for productivities (A_i) capture unobserved factors that make a location more or less attractive for employment conditional on the observed distribution of wages, residents and average residential income and the parameterized values of trade costs (e.g. natural resources). Having recovered these unique unobserved productivities (A_i), the model generates predictions for bilateral trade between counties from (10), which we use in our counterfactuals for changes in the model's exogenous variables as discussed above. Aggregating trade between counties within pairs of CFS regions, we can compare the model's predictions for trade between CFS to the observed trade between CFS regions in the data. We undertake this comparison below.

To solve the trade balance condition, we use a central estimate for the elasticity of substitution from the existing empirical literature: we assume $\sigma = 4$ based on the estimates using U.S. data in Bernard, Eaton, Jensen and Kortum (2003). We model bilateral trade costs as a function of distance. For bilateral pairs with positive trade, we assume that bilateral trade costs are a constant elasticity function of distance and a stochastic error ($d_{ni} = dist_{ni}^{\psi} \tilde{e}_{ni}$). For bilateral pairs with zero trade, the model implies prohibitive trade costs ($d_{ni} \rightarrow \infty$). Taking logarithms in the trade share (10) for pairs with positive trade, we obtain that the value of bilateral trade between source i and destination n (X_{ni}) is given by

$$\log X_{ni} = \zeta_n + \chi_i - (\sigma - 1)\,\psi \log dist_{ni} + \log e_{ni}, \tag{24}$$

where the source fixed effect (χ_i) controls for employment, wages and productivity (L_{Mi}, w_i, A_i) in the source i; the destination fixed effect (ζ_n) controls for average income, $\bar{v}_n$, residents, L_{Rn}, and multilateral resistance (as captured in the denominator of equation (10)) in the destination n; and $\log e_{ni} = (1 - \sigma) \log \tilde{e}_{ni}$.

Estimating the gravity equation (24) for all bilateral pairs with positive trade using OLS, we find a regression R-squared of 0.83. In Figure 1, we display the conditional relationship between the log value of trade and log distance, after removing source and destination fixed effects from both log trade and log distance. We find that the log linear functional form provides a good approximation to the data, with a tight and approximately linear relationship between the two variables. We estimate a coefficient on log distance of $-(\sigma - 1)\,\psi = -1.29$. For our assumed value of $\sigma = 4$, this implies an elasticity of trade costs with respect to distance of $\psi = 0.43$. The tight linear relationship in Figure 1, makes us confident in the parametrization of trade costs in terms of distance as a way of solving the trade balance condition (23) for unobserved productivities (A_i).

As an alternative check on our parameterization of trade costs, Figure 2 displays the model's predictions for bilateral trade between CFS regions against the corresponding values in the data. The only way in which we used the data on trade between CFS regions was to estimate the distance elasticity $-(\sigma - 1)\,\psi = -1.29$. Given this distance elasticity, we use the goods market clearing condition (23) to solve for productivities and generate the model's predictions for bilateral trade between countries and hence CFS regions. Therefore, the model's predictions and the data can differ
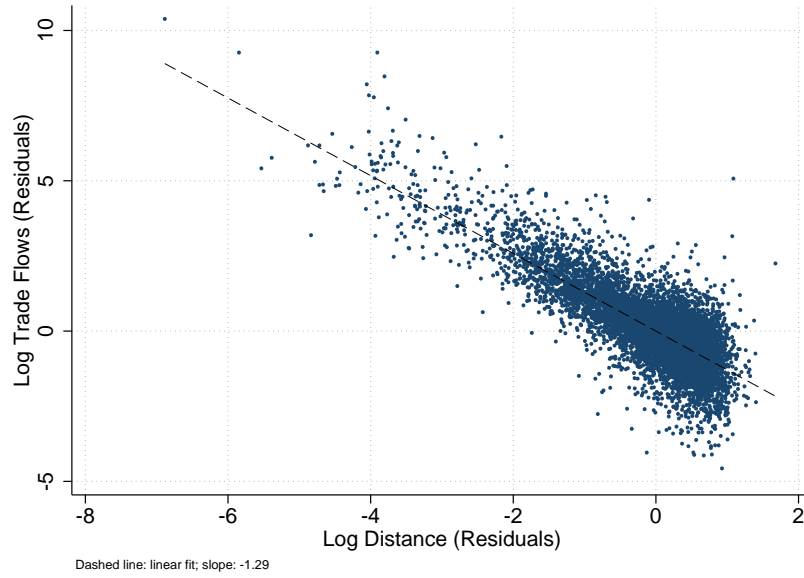
13

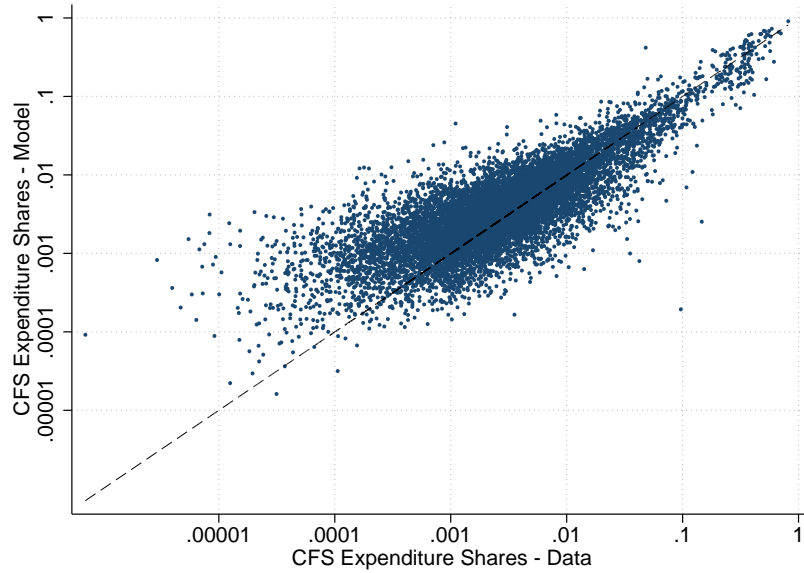Figure 1: Gains in Goods Trade Between CFS Regions

Figure 2: Bilateral Trade Shares in the Model and Data

from one another. Nonetheless, we find a strong and approximately log linear relationship between the model's predictions and the data, which is tighter for larger trade values.

3.2 The Magnitude and Gravity of Commuting Flows

We start by providing evidence on the quantitative relevance of commuting as a source of spatial linkages between counties and CZ's. To do so we use data from the American Community Survey

(ACS), which reports county-to-county worker flows for 2006-2010. To abstract from business trips that are not between a worker's usual place of residence and workplace, we define commuting flows as those of less than 120 kilometers in each direction (a round trip of 240 kilometers).[10] Table 1 reports some descriptive statistics for these commuting flows.

We find that commuting beyond county boundaries is substantial and varies in importance across locations. For the median county, around 17 percent of its residents work outside the county (first row, third column) and around 11 percent of its workers live outside the county (second row, third column). For the county at the 90th percentile, these two figures rise to 43 and 32 percent respectively (fifth column, first and second rows respectively).

One might think that using commuting zones circumvents the need to incorporate commuting in the analysis since these areas are designed to encompass most commuting flows between counties. Nevertheless, we find that CZ's provide an imperfect measure of local labor markets, with substantial commuting beyond CZ boundaries that again varies in importance across locations. For the median county, around 34 percent of the workers who commute outside their county of residence also commute outside their CZ of residence (third row, third column), while around 40 percent of the workers who commute outside their county of workplace also commute outside their CZ of workplace (fourth row, third column). For the CZ at the 90th percentile, these two figures rise to 90 and 81 percent respectively (fifth column). Taken together, these results highlight the quantitative relevance of commuting as a source of spatial linkages between counties and CZ's within the United States.

	p10	p25	p50	p75	p90	Max	Mean
Commuters from Residence	0.03	0.08	0.17	0.30	0.43	0.96	0.20
Commuters to Workplace	0.01	0.03	0.11	0.21	0.32	0.82	0.14
Outside CZ — Total (Res)	0.03	0.11	0.34	0.64	0.90	1.00	0.40
Outside CZ — Total (Work)	0.07	0.19	0.40	0.61	0.81	1.00	0.42

Note: Tabulations on 3,111 counties and 709 commuting zones; minimum is 0 for all. First row shows fraction of residents that work outside county. Second row shows fraction of workers who live outside county. Third row shows fraction of residents that work outside county who also work outside the county's commuting zone. Fourth row shows fraction of workers that live outside county who also live outside the county's commuting zone. p10, p25 etc refer to the 10th, 25th etc percentiles of the distribution.

Table 1: Commuting Across Counties and Commuting Zones

A key prediction of the model is a gravity equation for commuting. Using land market clearing

---

[10]The majority of commutes are less than 45 minutes in each direction (Duranton and Turner 2011), with commutes of 120 minutes in each direction rare. In our analysis, we measure distance between counties' centroids. We choose the 120 kilometers threshold based on a change in slope of the relationship between log commuters and log distance at this distance threshold. See the web appendix for further discussion.

(5) and the price index (12), the commuting probability (15) can be written as

$$\lambda_{ni} - \frac{B_{ni}\kappa_{ni}^{-\epsilon}\left(\frac{L_{Mn}}{\pi_{nn}}\right)^{-\frac{\alpha\epsilon}{\sigma-1}}A_n^{\alpha\epsilon}w_n^{-\alpha\epsilon}\bar{v}_n^{-\epsilon(1-\alpha)}\left(\frac{L_{Rn}}{H_n}\right)^{-\epsilon(1-\alpha)}w_i^{\epsilon}}{\sum_{r\in N}\sum_{s\in N}B_{rs}\kappa_{rs}^{-\epsilon}\left(\frac{L_{Mr}}{\pi_{rr}}\right)^{-\frac{\alpha\epsilon}{\sigma-1}}A_r^{\alpha\epsilon}w_r^{-\alpha\epsilon}\bar{v}_r^{-\epsilon(1-\alpha)}\left(\frac{L_{Rr}}{H_r}\right)^{-\epsilon(1-\alpha)}w_s^{\epsilon}} = 0. \qquad (25)$$

The commuting probabilities (25) provide a system of $N \times N$ equations that can be solved for a unique matrix of $N \times N$ unobserved amenities (B_{ni}). The next proposition shows this formally.

Proposition 4 *(Amenities Inversion) Given the share of consumption goods in expenditure (α), the hetero-geneity in location preferences (ϵ), the observed data on wages, employment, trade shares, average residential income, residents and land area $\{w_i, L_{Mi}, \pi_{ii}, \bar{v}_i, L_{Ri}, H_i\}$, and a parameterization of commuting costs ($\kappa_{ni}^{-\epsilon}$), there exist unique values of the unobserved amenities (B_{ni}) for each location n that are consistent with the data being an equilibrium of the model.*

Proof. See the appendix. ∎

The resulting solutions for amenities (B_{ni}) capture unobserved factors that make a pair of residence and workplace locations more or less attractive conditional on the observed wages, employment, trade shares, average residential income, residents, and land area, as well as the parameterized commuting costs (e.g. attractive scenery and differences in transport infrastructure that are not captured in the parameterized commuting costs). Together unobserved productivity (A_i) and amenities (B_{ni}) correspond to structural residuals that ensure that the model exactly replicates the observed data given the parameters and our assumptions for trade and commuting costs.

We model bilateral commuting costs as a function of distance. For bilateral pairs with positive commuting, we assume that bilateral commuting costs depend on distance with a unitary elasticity and a stochastic error ($\kappa_{ni} = dist_{ni}\tilde{g}_{ni}$). For bilateral pairs with zero commuting, the model implies negligible amenities ($B_{ni} \to 0$) and/or prohibitive commuting costs ($\kappa_{ni} \to \infty$). Taking logarithms in the commuting probability (15) for pairs with positive commuting, we obtain:

$$\log \lambda_{ni} = \eta_n + \mu_i - \epsilon \log dist_{ni} + \log g_{ni}, \qquad (26)$$

where we allow amenities to have a residence component ($\tilde{B}_n$), a workplace component ($\tilde{B}_i$) and a bilateral component that is idiosyncratic to a residence-workplace pair ($\tilde{B}_{ni}$) such that $B_{ni} = \tilde{B}_n\tilde{B}_i\tilde{B}_{ni}$. The origin fixed effect (η_n) controls for residence characteristics (P_n, Q_n, $\tilde{B}_n$); the destination fixed effect (μ_i) controls for workplace characteristics (w_i, $\tilde{B}_i$); and $\log g_{ni} = \log \tilde{B}_{ni} - \epsilon \log \tilde{g}_{ni}$.

Estimating the gravity equation (26) for all bilateral pairs with positive commuters using OLS, we find a regression R-squared of 0.76. In Figure 3, we display the conditional relationship between log commuters and log distance, after removing residence and workplace fixed effects from both log commuters and log distance. We find that the log linear functional form provides a good approximation to the data, with a tight and approximately linear relationship between the two variables. We find an estimated coefficient on log distance of $-\epsilon = -4.34$, which determines the Fréchet shape
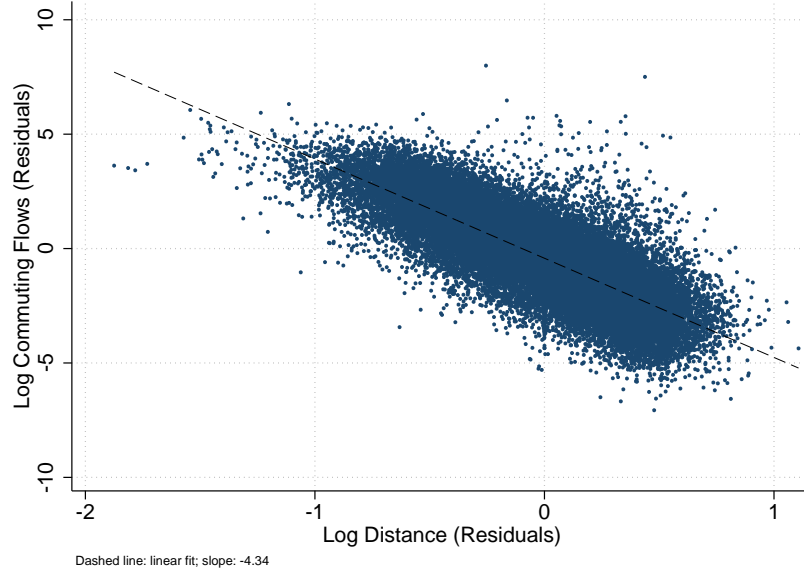
Figure 3: Gravity in Commuting Between Counties

parameter for the heterogeneity of worker preferences. We assume a central value for the share of housing in consumer expenditure from the Bureau of Economic Analysis of $1 - \alpha = 0.40$ percent. The tight fit shown in Figure 3, makes us confident that our parametrization of commuting costs and amenities in terms of distance fits the data quite well.[11]

3.3 House Prices

As discussed above, we calibrate productivities (A_i) and amenities (B_{ni}) such that the model exactly replicates the observed data on commuting flows, wages, employment and land area and the implied values of residents, and average residential income for the assumed parameters. As a check on our use of the Cobb-Douglas functional form to model residential land use, Figure 4 displays the model's predictions for land prices against median house prices in the data[12]. Although the model is necessarily an abstraction, and there are a number of potential sources of differences between land prices and house prices, we find a strong and approximately log linear relationship across counties between the model's predictions and the data. The slope of the relationship is 2.12 and the R-squared is 0.27, which is quite large given the simplicity of equation (5).

---

[11]Alternatively, we could use the commuting probability (15), to measure a composite of relative commuting costs and bilateral amenities with the Head and Ries (2001) index:

$$\left(\frac{\lambda_{ni}}{\lambda_{nn}} \frac{\lambda_{in}}{\lambda_{ii}} \right)^{\frac{1}{2}} = \left(\frac{B_{ni}\kappa_{ni}^{-\epsilon}}{B_{nn}\kappa_{nn}^{-\epsilon}} \frac{B_{in}\kappa_{in}^{-\epsilon}}{B_{ii}\kappa_{ii}^{-\epsilon}} \right)^{\frac{1}{2}}.$$

[12]We use the county median housing value from the ACS (2010).

Figure 4: Land Prices in the Model and House Prices in the Data

4 Quantitative Exercises

We have now quantified the model to be consistent with gravity in goods trade, gravity in commuting flows and the observed cross-section distribution of employment and wages. We next use the model to undertake three quantitative exercises that shed light on spatial linkages in goods and factor markets. First, we provide evidence on the local employment elasticity and demonstrate its heterogeneity across counties. We show that commuting links are essential to explain this heterogeneity, so in a second exercise we set out to explain the role of commuting in determining the spatial distribution of economic activity and welfare by considering a counterfactual with prohibitive commuting costs. Given the importance of commuting links, in a third exercise we investigate their interaction with trade costs. We compare the counterfactual effects of a reduction in trade costs in the actual world with commuting to the effects in a hypothetical world without commuting.

4.1 Local Employment Elasticities

To provide evidence on local employment elasticities, we compute 3,111 counterfactual exercises where we shock each county with a 5 percent productivity shock (holding productivity in all other counties and holding all other exogenous variables constant). Figure 5 shows the estimated kernel density for the distribution of the general equilibrium elasticity of employment with respect to the productivity shock across these treated counties (black line). We also show the 95 percent confidence intervals around this estimated kernel density (gray shading).

The mean estimated local employment elasticity of around 1.4 is greater than one because of the home market effects in the model. Around this mean, we find substantial heterogeneity in the pre-
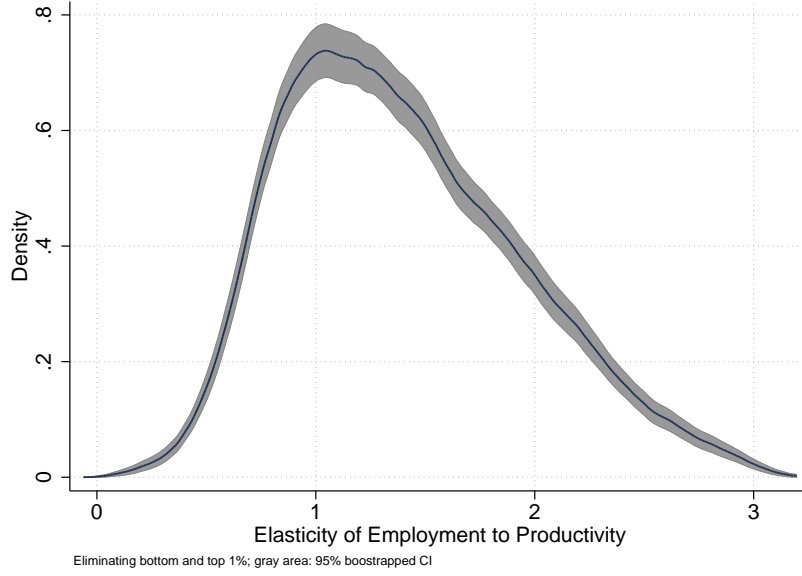
18

Figure 5: Kernel density for the distribution of employment elasticities in response to a productivity shock across counties

dicted effects of the productivity shock, which vary from close to 0 to more than 3. This variation is surprisingly large. It implies that taking a local employment elasticity estimated for one group of counties and applying that elasticity to another group of counties can lead to substantial discrepancies between the true and predicted impacts of a productivity shock.

To provide a point of comparison, Figure 6 also includes the general equilibrium elasticity of residents in a county with respect to the same 5 percent productivity shock in that county (again holding other parameters constant). Again we show the estimated kernel density for the distribution of the general equilibrium elasticity of residents with respect to the productivity shock across the 3,111 treated counties (black line) and the 95 percent confidence intervals (gray shading). We find far more heterogeneity across counties in the employment elasticity than in the residents elasticity (which ranges from close to 0 to around 1). Since employment and residents can only differ through commuting, this by itself suggests that the heterogeneity in the local employment elasticity in response to the productivity shock is largely driven by commuting links between counties.

4.1.1 Explaining the Heterogeneity in Local Employment Elasticities

We want to use the model to provide intuition on the determinants of local employment elasticities. Furthermore, we want the theory to help us determine a set of variables that can be used empirically to account for the estimated heterogeneity in the distribution of local employment elasticities. To do so, we compute partial equilibrium elasticities of own wages and own employment with respect to the productivity shock. These partial equilibrium elasticities capture the direct effect of a productivity shock on wages, employment and residents in the treated location, holding constant all other
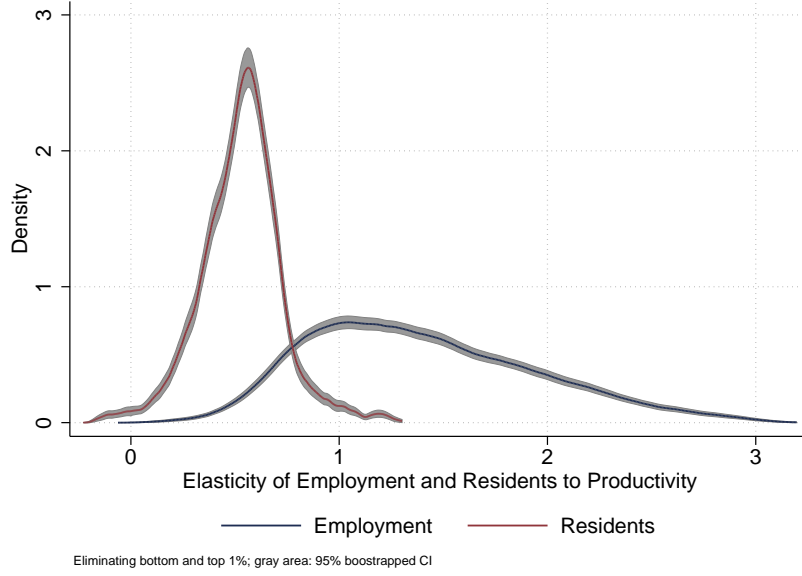
Figure 6: Kernel density for the distribution of employment and residents elasticities in response to a productivity shock across counties

endogenous variables at their values in the initial equilibrium.[13] Hence, although potentially useful to provide intuition, or as empirical controls, they do not account for all the rich set of interactions in the model captured by the general equilibrium elasticities presented in Figures 5.

If we hold constant all variables except for w_n, L_{M_n}, and L_{Rn} in the treated county n, the local employment elasticity as a result of a productivity shocks is

$$\frac{\partial L_{Mn}}{\partial A_n} \frac{A_n}{L_{Mn}} = \frac{\partial L_{Mn}}{\partial w_n} \frac{w_n}{L_{Mn}} \cdot \frac{\partial w_n}{\partial A_n} \frac{A_n}{w_n}. \tag{27}$$

From the trade balance condition (11), the partial elasticity of wages with respect to the productivity shock is given by

$$\frac{\partial w_n}{\partial A_n} \frac{A_n}{w_n} = \frac{(\sigma - 1) \sum_{i \in N} (1 - \pi_{in}) \xi_{in}}{[1 + (\sigma - 1) \sum_{i \in N} (1 - \pi_{in}) \xi_{in}] + [1 - \sum_{i \in N} (1 - \pi_{in}) \xi_{in}] \frac{\partial L_{Mn}}{\partial w_n} \frac{w_n}{L_{Mn}} - \xi_{nn} \frac{\partial L_{Rn}}{\partial w_n} \frac{w_n}{L_{Rn}}}. \tag{28}$$

where $\xi_{in} = \pi_{in} \alpha \bar{v}_i L_{Ri} / w_n L_n$ is the share of expenditure from market i in location n's income. The intuition for this expression can be seen most clearly by evaluating it for small changes in employment and residents ($\frac{\partial L_{Mn}}{\partial w_n} \frac{w_n}{L_{Mn}} = \frac{\partial L_{Rn}}{\partial w_n} \frac{w_n}{L_{Rn}} = 0$). From the remaining terms in $(\sigma - 1) \sum_{i \in N} (1 - \pi_{in}) \xi_{in}$, the elasticity of wages with respect to productivity is high when location n accounts for a small share of expenditure (small π_{in}) in markets i that account for a large share of its income (high ξ_{in}). In these circumstances, the productivity shock reduces the prices of location n's goods and results in only a small reduction in the consumption goods price indices (small π_{in}) in its main markets (high ξ_{in}).[14] Therefore the productivity shock leads to a large increase in the demand for location n's goods

---

[13]See Section A.8 in the appendix for the derivation of these partial equilibrium elasticities.

[14]These price indices summarize the price of competing varieties in each market. Note that the elasticity of the price index (12) in location n with respect to wages in location i is $(\partial P_n / \partial w_i)(w_i / P_n) = \pi_{ni}$.

and hence in its wages. Thus $\sum_{i \in N} (1 - \pi_{in}) \xi_{in}$ provides a measure of location i's linkages to other locations in goods markets.

From the commuting market clearing condition (18), the partial elasticity of employment respect to wages is

$$\frac{\partial L_{Mn}}{\partial w_n} \frac{w_n}{L_{Mn}} = \epsilon \sum_{i \in N} (1 - \lambda_{in|i}) \vartheta_{in} + \vartheta_{nn} \left(\frac{\partial L_{Rn}}{\partial w_n} \frac{w_n}{L_{Rn}} \right), \tag{29}$$

where $\vartheta_{in} = \lambda_{in|i} L_{Ri} / L_{Mn}$ is the share of commuters from source i in location n's employment. The intuition for this expression again can be seen most clearly by evaluating it for small changes in residents ($\frac{\partial L_{Rn}}{\partial w_n} \frac{w_n}{L_{Rn}} = 0$). From the remaining term in $\epsilon \sum_{i \in N} (1 - \lambda_{in|i}) \vartheta_{in}$, the elasticity of employment with respect to wages is high when location n accounts for a small share of commuters (small $\lambda_{in|i}$) from sources i that supply a large fraction of its employment (high ϑ_{in}). In these circumstances, location n's wage increase makes it a more attractive workplace and results in only a small increase in commuting market access (small $\lambda_{in|i}$) in its main sources of commuters (high ϑ_{in}).[15] Therefore the increase in wages leads to a large increase in commuters to location n and hence in its employment. Thus $\sum_{i \in N} (1 - \lambda_{in|i}) \vartheta_{in}$ provides a measure of location i's linkages to other locations through commuting markets.

From the residential choice probabilities (16), the partial elasticity of residents with respect to wages is

$$\frac{\partial L_{Rn}}{\partial w_n} \frac{w_n}{L_{Rn}} = \epsilon \left(\frac{\lambda_{nn}}{\lambda_{Rn}} - \lambda_{Mn} \right), \tag{30}$$

which also has an intuitive interpretation. A higher wage in location n makes it a more attractive workplace and increases its employment. Whether this increase in location n's employment leads to an increase in its share of residents depends on the fraction of residents who work locally ($\lambda_{nn}/\lambda_{Rn}$) relative to location n's overall share of employment (λ_{Mn}). Thus ($\lambda_{nn}/\lambda_{Rn} - \lambda_{Mn}$) provides a measure of location n's linkages to other locations through migration.

Using (30) in (29), and substituting the result into (28), we obtain closed-form solutions for the partial equilibrium elasticities of wages, employment and residents to a productivity shock. These partial equilibrium elasticities depend solely on the observed values of variables in the initial equilibrium: residential employment shares (λ_{Rn}), conditional commuting probabilities ($\lambda_{in|i}$), employment shares (ϑ_{in}), trade shares (π_{in}) and income shares (ξ_{in}).

Clearly, when we undertake our counterfactuals, we solve for the full general equilibrium effect of the productivity shock to each county. But these partial equilibrium elasticities in terms of observed variables have substantial explanatory power in predicting the impact of the productivity shock across locations.

In Table 2, we examine the determinants of the heterogeneity in the elasticity of employment with respect to productivity across the treated counties. In Column (1) we regress these elasticities

---

[15]Commuting market access appears in the numerator of the residential choice probabilities (λ_{Ri} in (16)) and summarizes access to employment opportunities: $W_n = \left[\sum_{s \in N} B_{ns} (w_s / \kappa_{ns})^\epsilon \right]^{1/\epsilon}$. Note that the elasticity of commuting market access in location n with respect to wages in location i is $(\partial W_n / \partial w_i) (w_i / W_n) = \lambda_{ni|n}$.

on a constant, which captures the mean employment elasticity across the 3,111 treated counties. In Columns (2) through (4) we attempt to explain the heterogeneity in local employment elasticities using standard county controls. In Column (2) we include log county employment as a control for the size of economic activity in a county. In Column (3) we also include log county wages and log county land area. In Column (4) we also include the average wage and total employment in neighboring counties. Although these controls are all typically statistically significant, we find that they are not particularly successful in explaining the variation in employment elasticities. Adding a constant and all these controls yields an R-squared of only 0.35 in Column (4). Clearly, there is substantial variation not captured by these controls.

In the remaining columns of the table we attempt to explain the heterogeneity in local employment elasticities using the partial equilibrium elasticities derived above. In Column (5) we first use the intuition (obtained by comparing the distributions in Figure 6) that commuting is essential to explain this elasticity. As a summary statistic of the lack of commuting links of a county we use $\lambda_{nn|n}$, namely, the probability that a worker lives and works in n. The weaker the commuting links of a county the higher $\lambda_{nn|n}$, which should reduce the local employment elasticity of that county. This is exactly what we find in Column (5). Furthermore, this variable alone yields an R-squared of 0.67, nearly double the R-squared in the regression where we include all the standard controls. This result underscores the importance of commuting links in explaining local employment elasticities.

The partial local elasticities computed above allow us to do better than just adding a summary measure of commuting links as the explanatory variable. In Column (6) we relate the variation in local employment elasticities to the measure of commuting linkages suggested by the model, $\sum_{i \in N} \left(1 - \lambda_{in|i}\right) \vartheta_{in}$. We also add the measures of migration and trade linkages suggested by the model, $(\lambda_{nn}/\lambda_{Rn} - \lambda_{Mn})$ and $\frac{\partial w_n}{\partial A_n} \frac{A_n}{w_n}$. Unsurprisingly, since these data are generated from the model, counties that account for a small share of commuters (small $\lambda_{in|i}$) from their main suppliers of commuters (high ϑ_{in}) have higher employment elasticities. More strikingly, this partial equilibrium measures of linkages explains around 81 percent of the variation in the general equilibrium elasticity. In fact, equation (27) implies that we should use the product of $\frac{\partial w_n}{\partial A_n} \frac{A_n}{w_n}$ with the first two terms rather than each term separately. This restriction yields similar results in Column (7) and confirms the importance of commuting linkages and, to a lesser extent, the interaction between migration and goods linkages. Finally, in the last two columns we combine these partial equilibrium elasticities with the standard controls we used in the first four columns. Clearly, although all variables are significant, these standard controls add little once we control for the partial equilibrium elasticities.

In sum, Table 2 shows that the heterogeneity in partial equilibrium elasticities is not well explained by standard county controls. In contrast, adding a summary statistic of commuting, or the partial equilibrium elasticities we propose above, can go a long way in explaining the heterogenous response of counties to productivity shocks. Still, the general equilibrium interactions between counties are complex, and cannot be fully captured by any of these variables (with R-squareds that never go beyond 0.83). The next subsection connects these results more tightly with the prevalent estima-

tion of treatment effects in the empirical literature.

	1	2	3	4	5	6	7	8	9
Dependent Variable:					Elasticity of Employment				
$\log L_{Mn}$		0.082**	0.079**	0.015				0.048**	0.053**
		(0.007)	(0.008)	(0.009)				(0.005)	(0.006)
$\log w_n$			-0.073	-0.104*				-0.228**	-0.222**
			(0.050)	(0.051)				(0.026)	(0.027)
$\log H_n$			-0.384**	-0.270**				-0.070**	-0.077**
			(0.011)	(0.016)				(0.009)	(0.009)
$\log L_{M,-n}$				0.061**				-0.046**	-0.050**
				(0.014)				(0.007)	(0.008)
$\log \bar{w}_{-n}$				0.877**				0.437**	0.497**
				(0.088)				(0.045)	(0.046)
$\lambda_{nn|n}$					-3.196**				
					(0.040)				
$\sum_{r\in N}\left(1-\lambda_{rn|r}\right)\vartheta_{rn}$						2.713**		3.535**	
						(0.219)		(0.218)	
$\vartheta_{nn}\left(\frac{\lambda_{nn}}{\lambda_{Rn}}-\lambda_{Mn}\right)$						-0.246		0.741**	
						(0.211)		(0.209)	
$\frac{\partial w_n}{\partial A_n}\frac{A_n}{w_n}$						-1.037**		-1.880**	
						(0.127)		(0.174)	
$\frac{\partial w_n}{\partial A_n}\frac{A_n}{w_n}\cdot\sum_{r\in N}\left(1-\lambda_{rn|r}\right)\vartheta_{rn}$							1.759**		1.525**
							(0.102)		(0.138)
$\frac{\partial w_n}{\partial A_n}\frac{A_n}{w_n}\cdot\vartheta_{nn}\left(\frac{\lambda_{nn}}{\lambda_{Rn}}-\lambda_{Mn}\right)$							-1.185**		-1.208**
							(0.097)		(0.133)
Constant	1.406**	0.620**	4.285**	-5.770**	3.951**	1.804**	1.707**	0.117	-0.443
	(0.011)	(0.070)	(0.484)	(0.809)	(0.033)	(0.209)	(0.096)	(0.491)	(0.453)
R^2	0.00	0.04	0.30	0.35	0.67	0.81	0.80	0.83	0.82
N	3,111	3,111	3,111	3,076	3,111	3,111	3,111	3,076	3,076

In this table, $L_{M,-n}\equiv\sum_{\{r:\text{dist}_{rn}\leq 120,r\neq n\}}L_{Mr}$ is the total employment in n neighbors whose centroid is no more than 120km away;

Table 2: Explaining the general equilibrium local employment elasticities to a 5 percent productivity shock

4.1.2 Measuring the Incidence of Local Labor Demand Shocks

A large empirical literature is concerned with estimating the elasticity of local employment with respect to local demand shocks. The central specification in this empirical literature is a "differences-in-differences" specification across locations i and time t given by

$$\Delta \ln Y_{it} = a_0 + a_1 \mathbb{I}_{it} + a_2 X_{it} + u_{it},\tag{31}$$

where Y_{it} is the outcome of interest (e.g. employment by workplace); $\mathbb{I}_{it}$ is a measure of the local demand shock (treatment); X_{it} are controls; and u_{it} is a stochastic error. The coefficient on the treatment a_1 corresponds to the elasticity of the outcome of interest with respect to the local demand shock.

This specification has a differences-in-differences interpretation, because the first difference is over time (before and after the shock), and the second difference is between treated and control counties.

This literature considers many sources of local labor demand shocks. Blanchard and Katz (1992) and Bound and Holzer (2000) directly measure labor demand shocks using innovations in employment and total hours worked respectively. To address the endogeneity of local labor input, many papers follow Bartik (1991) and use instruments based on interactions between changes in aggregate industry composition and the importance of industries in local employment. For example, Notowidigdo (2013) examines the impact of such local labor demand shocks across metropolitan areas in the United States. In other related research, Greenstone, Hornbeck and Moretti (2010) exploit regression discontinuities between winning and runner-up locations for million-dollar plants, while Michaels (2011) uses the discovery of oil in the Southern United States. Although both of these papers use county-level data, a number of other studies have used Commuting Zones (CZ's), which are aggregations of counties that are intended to better approximate local labor markets. For example, Autor and Dorn (2013) examines the automation of routine tasks and the polarization of U.S. employment and wages; Autor, Dorn and Hanson (2013) investigate rising Chinese import competition; and Yagan (2014) considers migration as insurance for idiosyncratic shocks during the Great Recession. Moretti (2011) reviews this burgeoning literature on local labor markets.

This empirical literature has discussed a number of potential concerns faced by the reduced-form specification (31). To address endogenous assignment to the treatment, recent research has sought exogenous sources of variation in local labor demand. Emphasis has been placed on measurement of the local labor demand shock (for interpreting the magnitude of the treatment effect). A widely-discussed limitation of the regression specification (31) is that it abstracts from general equilibrium effects, because it focuses on relative comparisons between the treatment and control group. More generally, the interpretation of the coefficient on the treatment is complicated if the control group is affected by the treatment through spatial linkages in goods and factor markets. Finally, the treatment effect can be heterogeneous across locations, because of different linkages in goods and factor markets. In this case, the estimates have an interpretation as a local average treatment effect (LATE), and a coefficient estimated for one group of counties need not apply for another group of counties.

We now compare the general equilibrium elasticities of employment with respect to the productivity shock in the model to the results of this type of reduced-form "differences-in-differences" estimates of the local average treatment effects of the productivity shock. In particular, we construct a regression sample including both treated and untreated countries from our 3,111 counterfactuals in which we shock each county in turn with a 5 percent productivity shock ($3,111^2 = 9,678,321$ observations). We use these data to estimate a "differences-in-differences" specification of a similar form to (31):

$$\Delta \ln Y_{it} = a_0 + a_1 \mathbb{I}_{it} + a_2 X_{it} + a_3 \left(\mathbb{I}_{it} \times X_{it} \right) + u_{it},$$

where i denotes the 3,111 counties and t indexes the 3,111 counterfactuals; $\Delta \ln Y_{it}$ is the change in log employment between the counterfactual and actual equilibria; $\mathbb{I}_{it}$ is a (0,1) indicator for whether

a county is treated with a productivity shock; and X_{it} are controls. We again consider two sets of controls (X_{it}): the model-suggested measures of linkages in goods and factor markets and more standard econometric controls (log employment, log wages and land area). We include both the main effects of these controls (captured by a_2) and their interactions with the treatment indicator to capture heterogeneity in the treatment effects (captured by a_3).

The key difference between this regression specification and the above results for the general equilibrium elasticities for the treated counties is that this regression specification *differences* relative to the untreated counties. The empirical literature has argued for the need to difference relative to untreated counties given the likely presence of other shocks or events, beyond the treatment, that can affect the treated counties simultaneously and can confound the true treatment effect. Of course, our synthetic dataset was generated without including any of these alternative shocks and so differencing is not needed in order to calculate the correct treatment effects.

In a specification without the controls ($a_2 = a_3 = 0$), the average effect of the productivity shock on the untreated counties is captured in the regression constant (a_0), and the local average treatment effect (a_1) corresponds to the difference in means between the treated and untreated counties. We compare estimating this regression specification including (i) a random untreated county in the control group, (ii) only the nearest untreated county in the control group, (iii) only neighboring counties within 120 kilometers of the treated county in the control group, (iv) only non-neighboring counties located from 120-240 kilometers from the treated county, and (v) all untreated counties in the control group.

In Figure 7 we show the predicted treatment effects ($a_1 + a_3 X_{it}$) from specifications including the model suggested controls. For comparability with the results reported above, we scale the predicted effects by the size of the productivity shock (we multiply by 1/0.05=20). The black thick line shows the general equilibrium elasticity of employment with respect to the productivity shock across the 3,111 treated counties (as reported above). The red and dashed, orange, green, blue and brown medium lines show the predicted treatment effects for specifications including closest, random, neighboring, non-neighboring and all untreated counties in the control group respectively. Across all specifications, we again find that the model suggested controls capture the heterogeneity in treatment effects. We also find that this treatment heterogeneity is large relative to the discrepancy between the general equilibrium elasticities and the predicted treatment effects after differencing relative to the control group (a discrepancy of up to 0.5 between the thick black line and the other lines compared to a range of about 3 in the treatment effects). However, the general equilibrium effect of about 0.5 is still economically quite relevant, as it represents more than 1/3 of the mean treatment effect of 1.4.

Finally, we find similar results across the different control groups, except for the specification using only the nearest counties to treated counties as controls. In this case, we find a larger discrepancy between the general equilibrium elasticities and the predicted treatment effects, because employment in the nearest untreated county is typically negatively affected by the increase in productivity
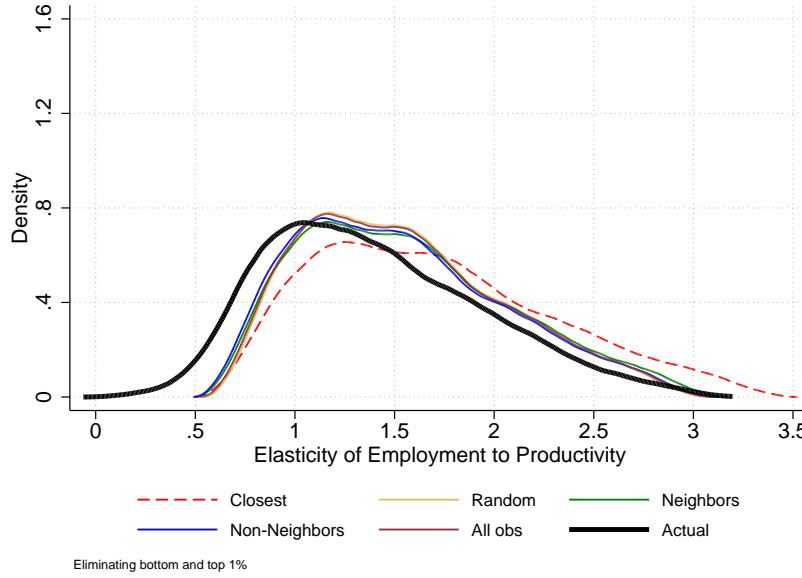
Figure 7: "Difference-in-difference" estimates of heterogeneous treatment effects (model-suggested controls)

in the treated county. While the use of contiguous locations as controls is often motivated based on similar unobservables (as in regression discontinuity designs), this pattern of results highlights that contiguous locations are also likely to be the most severely affected by general equilibrium linkages in goods and factor markets.

In Figure 8 we show the predicted treatment effects ($a_1 + a_3 X_{it}$) from specifications including log employment, log wages and land area as controls. Again for comparability, we scale the predicted effects by the size of the productivity shock. We use the same color scheme to indicate the general equilibrium elasticities for the treated counties and the predicted treatment effects from the "difference-in-difference" regressions. Across all specifications, we again find that these standard econometric controls are far less successful in capturing the heterogeneity in the treatment effects than the model suggested controls. As a result, we find a much larger discrepancy between the general equilibrium elasticities and the predicted treatment effects after differencing relative to the control group (a discrepancy of up to 0.75 between the thick black line and the other colored lines).

Taken together, the results of this section show that our quantitative spatial general equilibrium model implies substantial heterogeneity in local employment elasticities across counties; this heterogeneity is primarily explained by the commuting linkages in the model; partial equilibrium measures of commuting linkages suggested by the model are successful in capturing the heterogeneity in the general equilibrium local employment elasticities; and these model suggested measures of commuting linkages substantially outperform more standard econometric controls in explaining the heterogeneity in local employment elasticities. While capturing the full general equilibrium effects of the productivity shocks requires solving the model-based counterfactuals, we find that augmenting
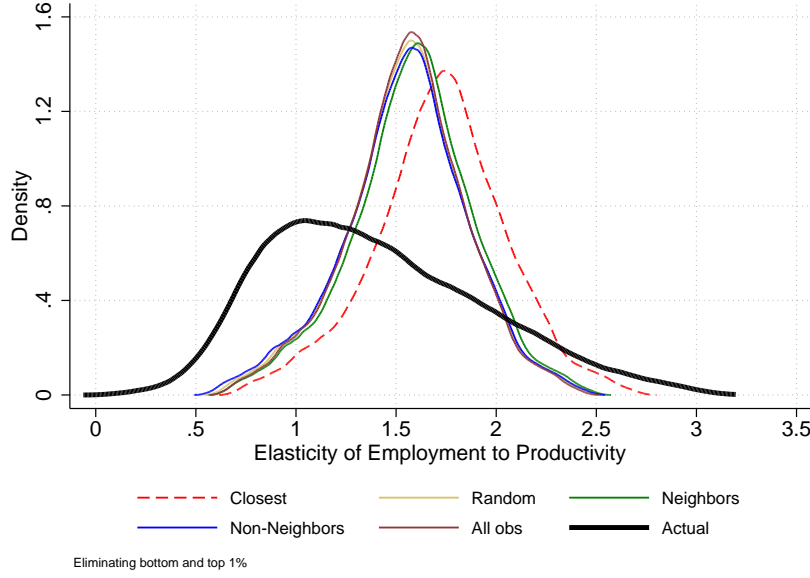
26

Figure 8: "Difference-in-difference" estimates of heterogeneous treatment effects (employment, wage and area controls)

"difference-in-difference" regressions with the model suggested measures of commuting linkages substantially improves their ability to predict the heterogeneity in the estimated treatment effects.

4.2 The Role of Commuting Costs

The results in the previous section underscore the importance of commuting links to explain the measured heterogeneity in local employment elasticities. They also highlight that commuting patterns have information about these links that cannot be accounted for with other county measures like employment, wages, number of residents, or even measures of employment and wages in surrounding counties. To provide further evidence on the quantitative relevance of commuting as a source of spatial linkages as well as to try to understand better its role in the spatial distribution of economic activity, we next undertake a counterfactual for prohibitive commuting costs to other locations. That is, we let $\kappa_{ni} \to \infty$ for $n \neq i$ and leave κ_{nn} and all other exogenous parameters unchanged.

Commuting enables workers to access high productivity locations without having to pay the high cost of living in those locations. Removing this technology restricts the opportunity set available to firms and workers and hence reduces welfare. Locations that were previously net exporters of commuters in the initial equilibrium become less attractive residences, while locations that were previously net importers of commuters in the initial equilibrium become less attractive workplaces. The result is a decline in the specialization of counties as residential or business locations. As the menu of potential pairs of workplace and employment locations changes, agents relocate.

In the equilibrium with commuting, expected utility conditional on choosing a pair of workplace and residence locations is the same across all bilateral pairs with positive commuting. In the equi-

27

librium without commuting, only those bilateral pairs in which workers live and work in the same location are feasible. Therefore, expected utility conditional on living and working in the same location is equalized throughout the economy. From (22), the change in the common level of expected utility from prohibitive commuting costs can be decomposed into the contributions of changes in the domestic commuting share (which equals one without commuting) and real residential income, where the latter change can be further decomposed into the contributions of changes in the domestic trade share, wages, expected residential income, residents and workers. Namely,

$$\widehat{U} = \left(\frac{1}{\widehat{\lambda}_{nn}}\right)^{\frac{1}{\epsilon}} \left(\frac{1}{\widehat{\pi}_{nn}}\right)^{\frac{\alpha}{\sigma-1}} \left(\frac{\widehat{w}_n}{\widehat{v}_n}\right)^{1-\alpha} \frac{\widehat{L}_{Mn}^{\frac{\alpha}{\sigma-1}}}{\widehat{L}_{Rn}^{1-\alpha}}, \tag{32}$$

where we have used the fact that $\{\kappa_{nn}, B_{nn}, A_n, d_{nn}\}$ are unchanged. Although the contributions of these individual components of welfare can differ across locations, their net effect must be such as to deliver the same common change in expected utility for the economy as a whole.

We find that increasing commuting costs to prohibitive levels reduces aggregate welfare by around 5.11 percent. This effect is comparable in magnitude to benchmark estimates of the welfare gains from international trade for an economy of the size of the United States (as for example in Eaton and Kortum 2002 and Arkolakis, Costinot and Rodriguez-Clare 2012). These results suggest that observed commuting flows are not only large and relevant for understanding the local effects of labor demand shocks, as shown above, but also have important implications for aggregate welfare. Smaller values for the Fréchet shape parameter (ϵ) imply more heterogeneity in preferences for pairs of residence and workplace locations and hence greater welfare losses from eliminating commuting conditional on the same observed initial equilibrium in the data. For example, in a world with a 50% lower value of ϵ the welfare loses from eliminating commuting amount to 10.16 percent.

We also find substantial effects of commuting on the spatial distribution of economic activity. With commuting, workers can live in suburban counties close to high productivity locations. In contrast, without commuting, workers must either live in those high productivity locations or disperse to other locations. As an illustration, Figures 9-12 show the impact of prohibitive commuting costs on economic activity in the New York area. In each figure, deeper green colors show lower values, while deeper brown colors indicate higher values. To provide some context, Figure 9 shows the ratio of employment to residents in the initial equilibrium in the observed data. Locations with ratios higher than one are net importers of commuters, while locations with ratios lower than one are net exporters of commuters. As apparent from the figure, Manhattan is initially the largest net importer, whereas other parts of the tri-state area are initially large net exporters.

Figures 10, 11 and 12 show the counterfactual changes in employment, residents and real income, respectively, from prohibitive commuting costs. Unsurprisingly, Manhattan experiences a substantial reduction in employment in Figure 10, as it can no longer attract commuters from the surrounding tri-state area. More notably, Manhattan also experiences a reduction in residents in Figure 11, and the tri-state area as a whole experiences a reduction in real income in Figure 12. This part of the country is one of the most intensive users of the commuting technology in the initial equilibrium.
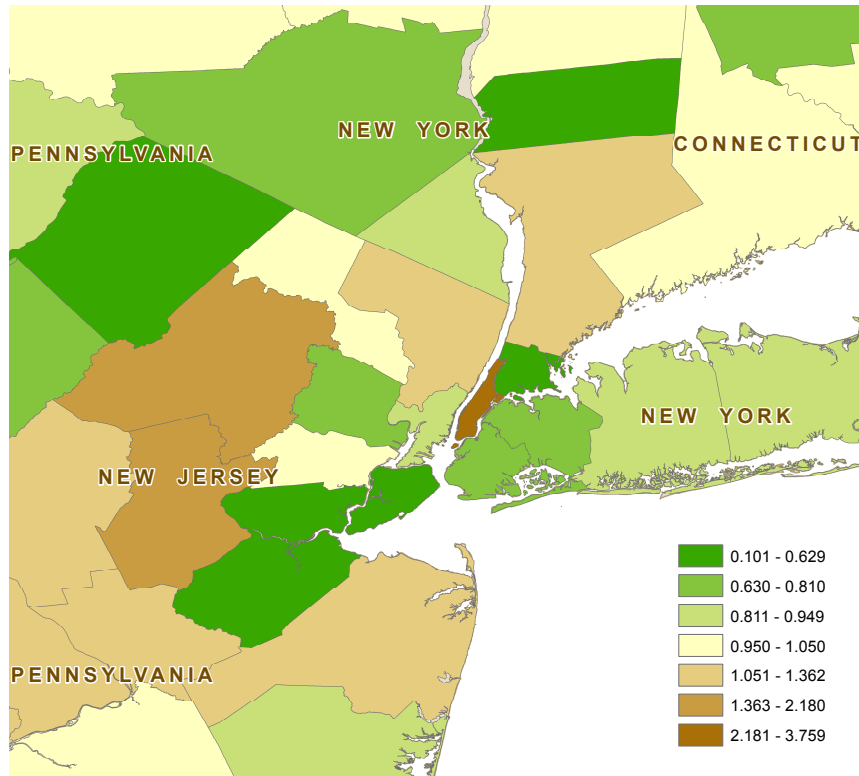
Figure 9: Ratio of employment to residents in the New York area in the initial equilibrium

Therefore, the loss of access to this technology reduces its attractiveness as a location of workplace and residence, and leads to a dispersion of economic activity towards other locations that were less intensive users of the commuting technology in the initial equilibrium.

In Figure 13, we show the counterfactual change in employment from increasing commuting costs to prohibitive levels for each county in our data. We display these counterfactual changes against each county's ratio of employment to residents in the initial equilibrium in the observed data. As employment and residents converge towards one another, locations that were large importers of commuters (greater than one on the horizontal axis) experience reductions in employment, while locations that were large exporters of commuters (less than one on the horizontal axis) experience increases in employment. This illustrates that adjustments happen mostly through employment changes, rather than migration as agents change their residence.

Figure 13 shows that the implications of the model for the change in employment as a result of shutting down commuting are well explained by the commuting intensity L_{Mn}/L_{Rn}. Commuting intensity is, however, not easy to explain with standard county-based measures of employment, residents, wages, or surrounding county characteristics. Table 3 presents the results of regressions that try to account for the cross-section of employment, residents, and the employment to residents ratio (a measure of county commuting intensity). The first column shows that one can account for most of the variation in county employment using the number of residents and wages. Column (2) shows a similar result for the number of residents and Columns (3) and (4) show that the results
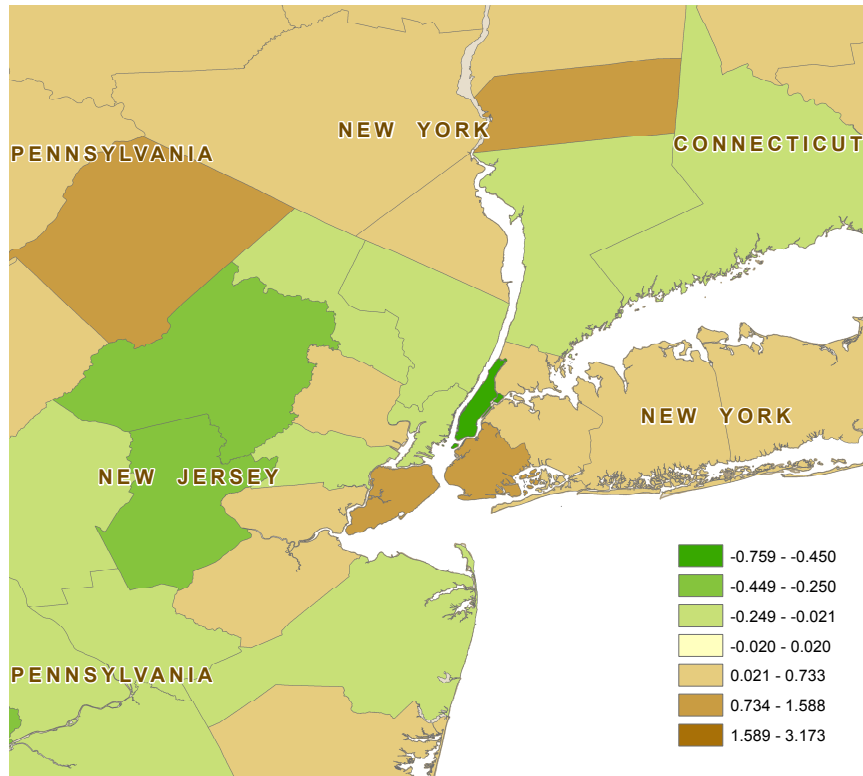
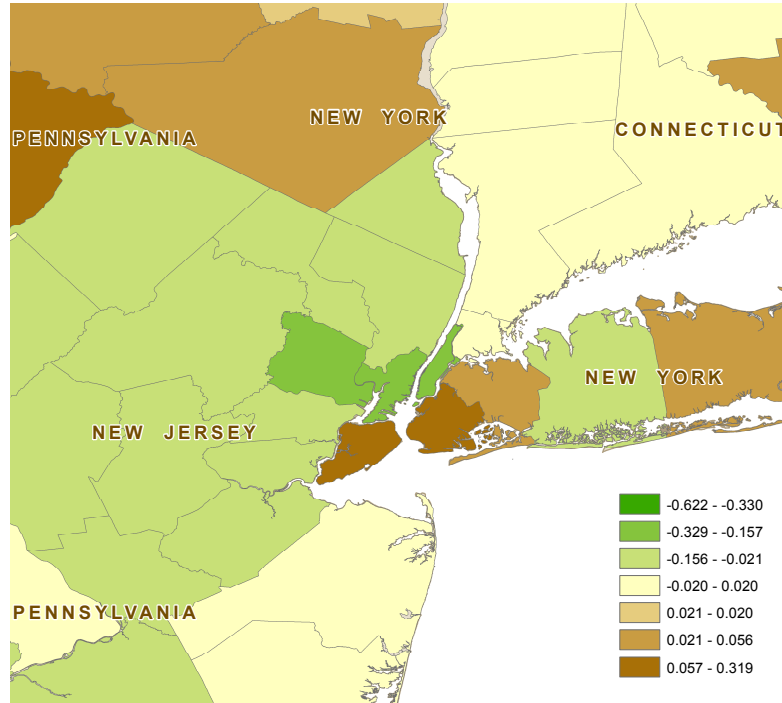Figure 10: Counterfactual relative change in employment ($\hat{L}_M$) from prohibitive commuting costs

Figure 11: Counterfactual relative change in residents ($\hat{L}_R$) from prohibitive commuting costs
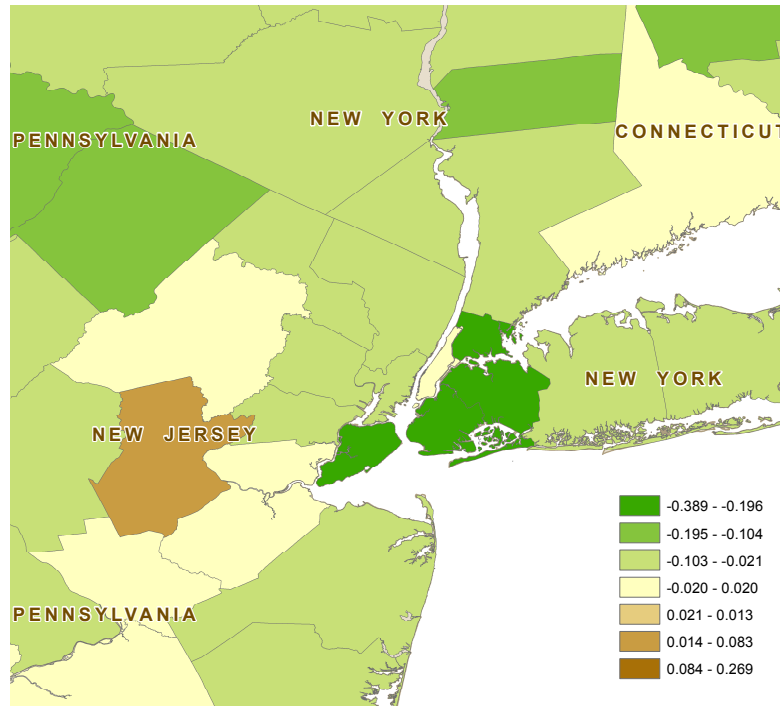
Figure 12: Counterfactual relative change in real income from prohibitive commuting costs
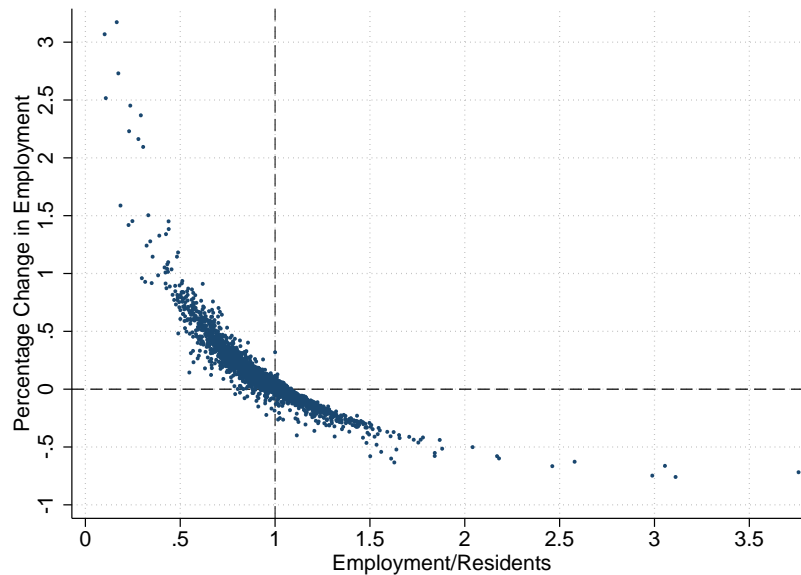
Figure 13: Counterfactual relative change in county employment ($\hat{L}_M$) from prohibitive commuting costs against initial employment to residents ratio (L_M / L_R)

are not affected when we add employment and wages in surrounding counties. So the first four columns show that it is relatively easy to explain the variation in employment and residents with standard variables. The following four columns demonstrate that this is not the case for commuting intensity. The level of residents, measures of wages, and measures employment, residents and wages

in surrounding counties do a poor job in accounting for the variation in commuting intensity. None of the R-squared in the last four columns of Table 3 amounts to more than 0.18. Once again, we confirm that the variation in commuting patterns provides information that cannot be easily replicated by other more standard variables. The commuting links revealed by this information are, to a large extent, responsible for the impact of changes in commuting costs, as well as for the heterogeneity in local employment elasticities discussed above.

	1	2	3	4	5	6	7	8
Dependent Variable:	$\log L_{Mn}$	$\log L_{Rn}$	$\log L_{Mn}$	$\log L_{Rn}$	L_{Mn}/L_{Rn}	L_{Mn}/L_{Rn}	L_{Mn}/L_{Rn}	L_{Mn}/L_{Rn}
$\log L_{Rn}$	0.965** (0.003)		0.978** (0.004)		-0.012** (0.004)		0.000 (0.004)	
$\log w_n$	0.409** (0.021)		0.452** (0.023)			0.276** (0.020)		0.264** (0.021)
$\log L_{Mn}$		0.962** (0.004)		0.929** (0.004)		0.002 (0.003)		0.026** (0.004)
$\log \bar{v}_n$		0.079** (0.026)		0.074** (0.029)	0.204** (0.024)		0.246** (0.027)	
$\log H_n$			0.021** (0.007)	0.043** (0.008)			-0.012 (0.007)	-0.034** (0.007)
$\log L_{R,-n}$			0.005 (0.006)				0.862** (0.113)	0.686** (0.107)
$\log \bar{w}_{-n}$			-0.343** (0.039)				0.714** (0.256)	0.602* (0.243)
$\log L_{M,-n}$				0.073** (0.007)			-0.884** (0.112)	-0.725** (0.107)
$\log \bar{v}_{-n}$				0.031 (0.046)			-0.826** (0.260)	-0.735** (0.247)
Constant	-4.026** (0.206)	-0.383 (0.254)	-1.178** (0.370)	-1.635** (0.398)	-1.084** (0.232)	-1.969** (0.194)	-0.074 (0.378)	0.117 (0.359)
R^2	0.98	0.97	0.98	0.98	0.02	0.09	0.09	0.18
N	3,111	3,111	3,076	3,076	3,111	3,111	3,076	3,076

In this table, $L_{M,-n} \equiv \sum_{r:d_{rn}\leq 120,r\neq n} L_{Mr}$ is the total employment in n neighbors whose centroid is no more than 120km away; $\bar{w}_{-n} \equiv \sum_{r:d_{rn}\leq 120,r\neq n} \frac{L_{Mr}}{L_{M,-n}} w_r$ is the weigthed average of their workplace wage. Analogous definitions apply to $L_{R,-n}$ and $\bar{v}_{-n}$. * $p < 0.05$; ** $p < 0.01$

Table 3: Explaining employment levels and commuting intensity

Even though Table 3 shows that commuting intensity is not closely related to county size, the results for the New York area described above together with those in Figure 13, suggest that maybe commuting is particularly important in large cities (or large commuting zones). Thus, eliminating commuting might affect particularly commuting zones with a large number of employees. In Figures 14-15, we examine the implications of these counterfactual changes for economic statistics at the commuting zone (CZ) level. We aggregate our county data in both the initial equilibrium and the counterfactual equilibrium to the CZ level and compute counterfactual changes for each CZ in our data. In Figure 14, we show the counterfactual change in CZ employment from prohibitive commuting costs against a measure of CZ dependence on commuting. This measure is the average share of

workers in a county within the CZ that live in the same county in which they work, which provides an inverse measure of dependence on commuting.

We find that aggregating up to the CZ level does not eliminate the effects of the change in commuting technology. Rather we find substantial changes in relative CZ employment from prohibitive commuting costs, ranging from declines of around 0.2 log points to rises of around 0.4 log points. Consistent with our earlier results for counties, we find that CZs that are more intensive users of commuting (smaller values on the horizontal axis) experience larger declines in employment, whereas CZs that are less intensive users of commuting (larger values on the horizontal axis) experience larger rises in employment.

Even though eliminating commuting affects the CZs that use commuting intensively, perhaps surprisingly, the change in employment is not related to the initial employment size of the CZ. Figure 15 displays the same change in employment from prohibitive commuting costs (same vertical axis) against initial CZ employment size (different horizontal axis). In line with our earlier findings for local employment elasticities, we find that employment size is a imperfect measure for commuting linkages in factor markets. We find no clear relationship between relative changes in employment and initial employment size in the CZ in Figure 15. Hence, the importance of commuting is by no means restricted to large cities. This explains why commuting could not be simply approximated by measures of county (or surrounding area) employment size in the previous section.

Taken together, the results in this section have shown that commuting linkages in factor markets are not only important for understanding the incidence of local labor demand shocks but also matter for the spatial distribution of economic activity and aggregate welfare. Increasing commuting costs to prohibitive levels involves a deterioration in the technology for organizing the spatial distribution of economic activity. This deterioration in technology implies less specialization across locations (in residents versus employment) and hence a less efficient sorting of workers and residents across locations in response to spatial variation in productivity and amenities. Furthermore, commuting links are not easily explained by measures of employment size either at the county or the commuting zone level. A result that, we believe, highlights further the importance of incorporating commuting data in regional analysis.

4.3 Interaction Between Trade and Commuting Costs

Most equilibrium models that feature geography, trade, and migration abstract from commuting (e.g. Allen and Arkolakis, 2014a, Caliendo et al., 2014, or Redding, 2014). Given the demonstrated importance of commuting for the determination of employment levels across regions, it is natural to expect that the effect of reductions in transport costs on welfare are affected by commuting as well. To provide evidence on the interaction of spatial linkages in goods and factor markets, we compare the effects of reductions in trade costs, both with and without commuting. We first undertake a counterfactual for a 20 percent reduction in trade costs between locations ($d_{ni} = 0.8$ for $n \neq i$ and $d_{nn} = 1$) starting from the observed initial equilibrium with commuting (using the observed bilateral com-
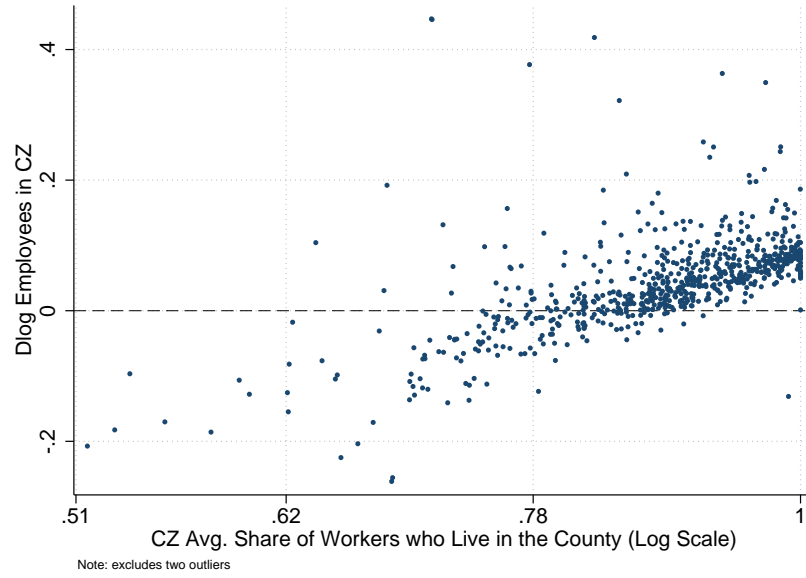
Figure 14: Counterfactual relative change in commuting zone employment ($\hat{L}_M$) from prohibitive commuting costs and initial dependence on commuting
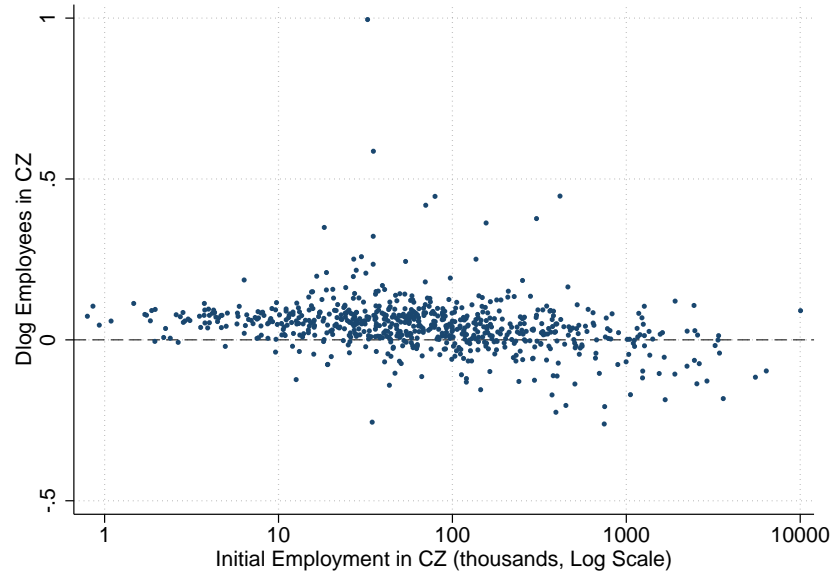
Figure 15: Counterfactual relative change in commuting zone employment ($\hat{L}_M$) from prohibitive commuting costs and initial employment size

muting shares to implicitly reveal the magnitude of bilateral commuting costs). We next undertake a counterfactual for the same 20 percent reduction in trade costs between locations from a counterfactual equilibrium with no commuting (starting from the observed equilibrium, we first undertake a counterfactual for no commuting, before then undertaking the counterfactual for the reduction in trade costs).

Population mobility again implies that expected utility is equalized across locations. From (22), the change in the common level of expected utility from the trade cost reduction can be decomposed as before into the contributions of changes in the domestic commuting share and real residential income, where the latter change can be further decomposed into the contributions of changes in the domestic trade share, wages, expected residential income, residents and workers:

$$\widehat{U} = \left(\frac{1}{\widehat{\lambda}_{nn}}\right)^{\frac{1}{\epsilon}} \left(\frac{1}{\widehat{\pi}_{nn}}\right)^{\frac{\alpha}{\sigma-1}} \left(\frac{\widehat{w}_n}{\widehat{\overline{v}}_n}\right)^{1-\alpha} \frac{\widehat{L}_{Mn}^{\frac{\alpha}{\sigma-1}}}{\widehat{L}_{Rn}^{1-\alpha}}, \tag{33}$$

where we have used the fact that $\{\kappa_{nn}, B_{nn}, A_n, d_{nn}\}$ are unchanged.

We find that trade and commuting are weak substitutes in terms of aggregate welfare. Starting from the observed equilibrium, we find aggregate welfare gains from the trade cost reduction of 4.4 percent. In contrast, starting from the counterfactual equilibrium without commuting, we find aggregate welfare gains from the same trade cost reduction of 4.7 percent. The reason is that employment and production are more dispersed without commuting, as employment and residents converge towards to one another. Therefore reductions in trade costs become more valuable in improving access to the traded varieties produced in these dispersed locations.

Trade and commuting, however, are complements for the spatial concentration of economic activity. The reason is again that commuting facilitates a separation of employment and residents and hence provides a technology for specialization. Reductions in trade costs increase the scope for specialization using this technology. As trade costs fall, and the most productive locations are better able to access the national market, commuting helps these locations to expand, by drawing workers from a suburban hinterland, without bidding up land prices as much as would otherwise occur.

To illustrate this complementarity, Figure 16 shows the relative change in employment from a 20 percent reduction in trade costs in the New York region (with and without commuting). Without commuting (left panel), the reduction in trade costs decreases employment in Manhattan, which reflects a general tendency in this class of economic geography models for lower trade costs to lead to a more dispersed spatial distribution of economic activity. In contrast, with commuting (right panel), the reduction in trade costs increases employment in Manhattan, as it enables Manhattan to increasingly take advantage of the commuting technology to service the national market. This increase in employment in Manhattan goes hand in hand with an increase in residents in the commuter belt in the surrounding tri-state area (not shown).

Therefore both lower trade costs and higher commuting costs are forces for the dispersion of economic activity. On the one hand, lower trade costs weaken agglomeration forces by reducing the incentive for firms and workers to locate close to one another. On the other hand, higher commuting costs increase congestion forces by forcing workers to live where they work, thereby bidding up land prices in congested locations. However, these two sets of forces interact with one another, so that the impact of a reduction in trade costs depends on the level of commuting costs. This interaction can be so strong as to reverse the sign of the effect of the trade cost reduction. While lower trade costs necessarily redistribute employment away from the most congested locations without commuting,

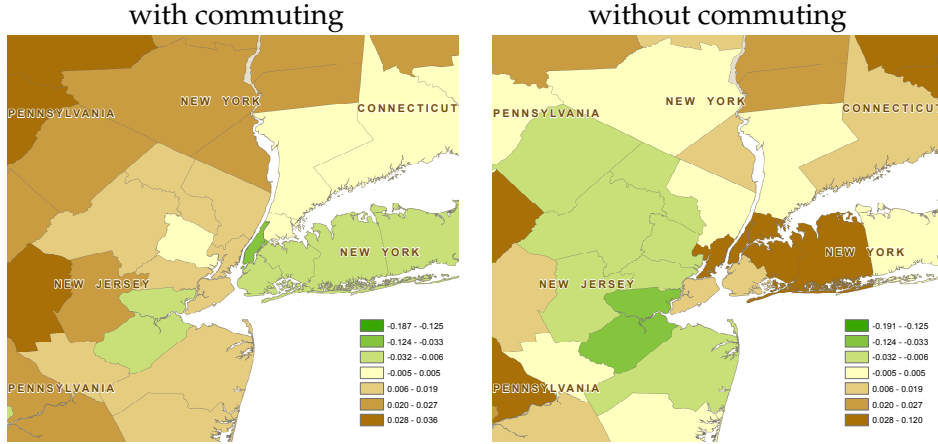they can increase employment in these locations with commuting.

Figure 16: Relative change in employment ($\hat{L}_M$) from a 20 percent reduction in trade costs (with and without commuting) in the New York area

This exercise also illustrates more generally the role of commuting linkages in shaping the consequences of a reduction in trade costs. Figure 17 shows changes in county employment following a reduction in trade costs in worlds without commuting (vertical axis) and with commuting (horizontal axis), alongside a 45-degree line. We find a low correlation between changes in employment with and without commuting, highlighting the importance of commuting linkages in shaping the impact of a reduction in trade costs. We find similar differences in predictions with and without commuting for other outcomes, such as residents, real income, consumption goods prices and land prices (not shown here). These results further underscore the prominence of commuting linkages in shaping the equilibrium spatial distribution of economic activity, and the necessity of incorporating them in models of economic geography.

5   Conclusions

Local technology, regulation, or infrastructure shocks can have far reaching economic effects through spatial linkages between locations. We have developed a spatial general equilibrium model that quantifies these spatial linkages in both goods markets (trade) and factor markets (commuting and migration). Our quantitative model uses the observed gravity equation relationships for goods and commuting flows to estimate the two key parameters of the model and matches exactly the observed cross-section distributions of employment, residents and incomes across U.S. counties.

We show that commuting flows are large and heterogeneous across counties and that commuting zones are imperfect in capturing these flows. The resulting differences in patterns of commuting lead to substantial variation across locations in elasticities of employment to productivity shocks, which have a mean of 1.4, but range from close to zero to more than three. These results question
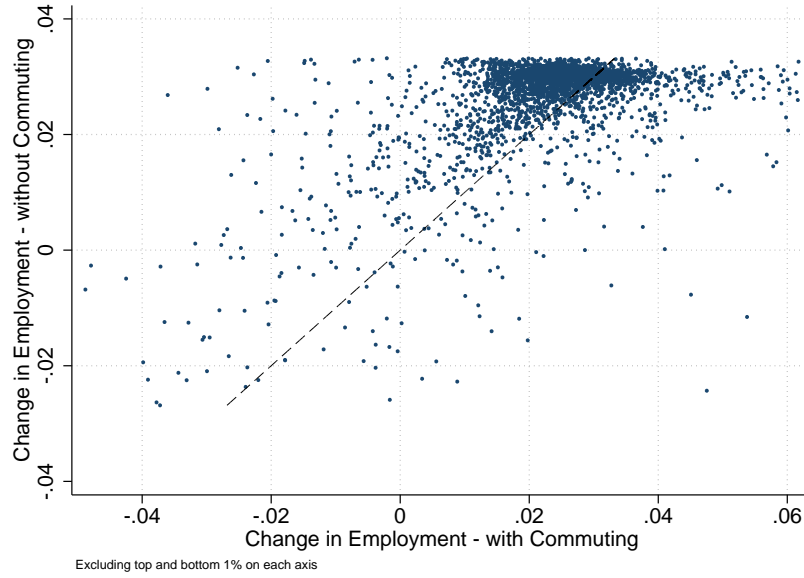
Figure 17: Relative change in employment ($\hat{L}_M$) from a 20 percent reduction in trade costs (with and without commuting) across all counties

the external validity of estimates of this elasticity that do not account for its large spatial heterogeneity. Furthermore, as our theoretical model incorporates multiple spatial linkages between locations (trade in goods and migration as well as commuting), it becomes an empirical question which of these linkages is more important given the observed patterns in the data in the initial equilibrium. We show that simple measures of commuting, or partial equilibrium measures of the spatial linkages derived from the model, are empirically successful in accounting for this variation in general equilibrium local employment elasticities.

We hope that our results are used to motivate the inclusion of these measures of commuting, and the other partial equilibrium elasticities, in future empirical estimations of local employment elasticities. Their inclusion is simple, as these measures depend only on observables in the initial equilibrium. Furthermore, these terms are not well accounted for by the inclusion of other variables like measures of employment or wages in the treated or neighboring counties. So including observed measures of these linkages is essential. Our results also question the use of empirical difference-in-differences strategies that use neighboring untreated counties to measure the treatment effect of a policy or shock. In our counterfactual exercises, untreated counties tend to be significantly affected by the shocks to the treated counties.

We find that commuting matters not only for the incidence of local shocks but also for aggregate welfare and the spatial distribution of economic activity. Increasing commuting costs to prohibitive levels reduces aggregate welfare by around 5.1 percent, which is comparable in magnitude to some estimates of the welfare gains from international trade. Commuting enables workers to live close to high productivity locations without having to pay the high land prices in those locations. Therefore

increasing commuting costs to prohibitive levels redistributes economic activity away from areas that use commuting intensively (e.g. the New York region) towards areas that use commuting less intensively. At the commuting zone level, these reallocations can lead to decreases and increases in employment of up to 0.2 and 0.4 log points respectively.

The theory we propose in this paper is, we believe, quite ambitious and rich. We view this theory as a reasonable framework to quantify the heterogeneity in local employment elasticities, because it is able to account for key observed relationships in the data (gravity in goods trade and commuting flows) and rationalizes the observed cross-section distribution of employment, wages, commuting and international trade flows in the initial equilibrium. As we calibrate to the initial equilibrium, the response to the economy to subsequent shocks in principle could be different from its behavior in the initial equilibrium. But since this initial equilibrium is itself the result of the accumulation of past shocks, we believe that it provides a reasonable benchmark against which to calibrate the model. Furthermore, this approach enables us to undertake counterfactuals using bilateral patterns of goods trade and commuting flows in the initial equilibrium to capture the rich unobserved patterns of trade and commuting costs between locations.

Finally, we have extended quantitative general equilibrium models of economic geography to incorporate an additional margin that is empirically relevant (commuting between locations). But we acknowledge that there remain other margins that can also matter for local employment elasticities in response to shocks (e.g. labor force participation and unemployment). We view the incorporation of these additional margins into spatial general equilibrium models as an important part of the future research agenda. However, given the quantitative magnitudes of commuting flows in the data, we believe that commuting will continue to play an important role in shaping local employment elasticities even in such future models incorporating a wider range of adjustment margins.

A Appendix

A.1 Commuting Decisions

A.1.1 Distribution of Utility

From all possible pairs of residence and employment locations, each worker chooses the bilateral commute that offers the maximum utility. Since the maximum of a sequence of Fréchet distributed random variables is itself Fréchet distributed, the distribution of utility across all possible pairs of residence and employment locations is:

$$1 - G(u) = 1 - \prod_{r=1}^{S} \prod_{s=1}^{S} e^{-\Psi_{rs} u^{-\epsilon}},$$

where the left-hand side is the probability that a worker has a utility greater than u, and the right-hand side is one minus the probability that the worker has a utility less than u for all possible pairs of residence and employment locations. Therefore we have:

$$G(u) = e^{-\Phi u^{-\epsilon}}, \qquad \Psi = \sum_{r=1}^{S} \sum_{s=1}^{S} \Psi_{rs}. \tag{34}$$

Given this Fréchet distribution for utility, expected utility is:

$$\mathbb{E}\left[u\right] = \int_{0}^{\infty} \epsilon \Psi u^{-\epsilon} e^{-\Psi u^{-\epsilon}} du. \tag{35}$$

Now define the following change of variables:

$$y = \Phi u^{-\epsilon}, \qquad dy = -\epsilon \Psi u^{-\epsilon+1}. \tag{36}$$

Using this change of variables, expected utility can be written as:

$$\mathbb{E}\left[u\right] = \int_{0}^{\infty} \Psi^{1/\epsilon} y^{-1/\epsilon} e^{-y} dy, \tag{37}$$

which can be in turn written as:

$$\mathbb{E}\left[u\right] = \delta \Psi^{1/\epsilon}, \qquad \delta = \Gamma\left(\frac{\epsilon - 1}{\epsilon}\right), \tag{38}$$

where $\Gamma(\cdot)$ is the Gamma function. Therefore we have the expression in the main text above:

$$\mathbb{E}\left[u\right] = \delta \Psi^{1/\epsilon} = \delta \left[\sum_{r=1}^{S} \sum_{s=1}^{S} B_{rs} \left(\kappa_{rs} P_r^{\alpha} Q_r^{1-\alpha}\right)^{-\epsilon} w_s^{\epsilon}\right]^{1/\epsilon}. \tag{39}$$

A.1.2 Residence and Workplace Choices

Using the distribution of utility for pairs of residence and employment locations (14), the probability that a worker chooses the bilateral commute from n to i out of all possible bilateral commutes is:

$$\pi_{ni} = \Pr\left[u_{ni} \geq \max\{u_{rs}\}; \forall r, s\right],$$

$$= \int_0^\infty \prod_{s \neq i} G_{ns}(u) \left[\prod_{r \neq n} \prod_s G_{rs}(u)\right] g_{ni}(u) du,$$

$$= \int_0^\infty \prod_{r=1}^S \prod_{s=1}^S \epsilon \Psi_{ni} u^{-(\epsilon+1)} e^{-\Psi_{rs} u^{-\epsilon}} du.$$

$$= \int_0^\infty \epsilon \Psi_{ni} u^{-(\epsilon+1)} e^{-\Psi u^{-\epsilon}} du.$$

Note that:

$$\frac{d}{du}\left[-\frac{1}{\Psi} e^{-\Psi u^{-\epsilon}}\right] = \epsilon u^{-(\epsilon+1)} e^{-\Psi u^{-\epsilon}}. \tag{40}$$

Using this result to evaluate the integral above, the probability that the worker chooses to live in location n and commute to work in location i is:

$$\lambda_{ni} = \frac{\Psi_{ni}}{\Psi} = \frac{B_{ni}\left(\kappa_{ni} P_n^\alpha Q_n^{1-\alpha}\right)^{-\epsilon}(w_i)^\epsilon}{\sum_{r=1}^S \sum_{s=1}^S B_{rs}\left(\kappa_{rs} P_r^\alpha Q_r^{1-\alpha}\right)^{-\epsilon}(w_s)^\epsilon}. \tag{41}$$

Summing across all possible workplaces s, we obtain the probability that a worker chooses to live in location n out of all possible locations is:

$$\lambda_n = \frac{L_{Rn}}{\bar{L}} = \frac{\Psi_n}{\Psi} = \frac{\sum_{s=1}^S B_{ns}\left(\kappa_{ns} P_n^\alpha Q_n^{1-\alpha}\right)^{-\epsilon}(w_s)^\epsilon}{\sum_{r=1}^S \sum_{s=1}^S B_{rs}\left(\kappa_{rs} P_r^\alpha Q_r^{1-\alpha}\right)^{-\epsilon}(w_s)^\epsilon}. \tag{42}$$

Similarly, summing across all possible residence locations r, we obtain the probability that a worker chooses to work in location i out of all possible locations is:

$$\lambda_i = \frac{L_{Mi}}{\bar{L}} = \frac{\Psi_i}{\Psi} = \frac{\sum_{r=1}^S B_{ri}\left(\kappa_{ri} P_r^\alpha Q_r^{1-\alpha}\right)^{-\epsilon}(w_i)^\epsilon}{\sum_{r=1}^S \sum_{s=1}^S B_{rs}\left(\kappa_{rs} P_r^\alpha Q_r^{1-\alpha}\right)^{-\epsilon}(w_s)^\epsilon}. \tag{43}$$

For the measure of workers in location i (L_{Mi}), we can evaluate the conditional probability that they commute from location n (conditional on having chosen to work in location i):

$$\lambda_{ni|i} = \Pr\left[u_{ni} \geq \max\{u_{ri}\}; \forall r\right],$$

$$= \int_0^\infty \prod_{r \neq n} G_{ri}(u) g_{ni}(u) du,$$

$$= \int_0^\infty e^{-\Psi_i u^{-\epsilon}} \epsilon \Psi_{ni} u^{-(\epsilon+1)} du.$$

Using the result (40) to evaluate the integral above, the probability that a worker commutes from location n conditional on having chosen to work in location i is:

$$\lambda_{ni|i} = \frac{\Psi_{ni}}{\Psi_i} = \frac{B_{ni}\left(\kappa_{ni} P_n^\alpha Q_n^{1-\alpha}\right)^{-\epsilon}(w_i)^\epsilon}{\sum_{r=1}^S B_{ri}\left(\kappa_{ri} P_r^\alpha Q_r^{1-\alpha}\right)^{-\epsilon}(w_i)^\epsilon},$$

which simplifies to:

$$\lambda_{ni|i} = \frac{B_{ni} \left(\kappa_{ni} P_n^\alpha Q_n^{1-\alpha} \right)^{-\epsilon}}{\sum_{r=1}^{S} B_{ri} \left(\kappa_{ri} P_r^\alpha Q_r^{1-\alpha} \right)^{-\epsilon}}. \tag{44}$$

For the measure of residents of location n (L_{Rn}), we can evaluate the conditional probability that they commute to location i (conditional on having chosen to live in location n):

$$\lambda_{ni|n} = \Pr \left[u_{ni} \geq \max\{u_{ns}\}; \forall s \right],$$

$$= \int_0^\infty \prod_{s \neq i} G_{ns}(u) g_{ni}(u) du,$$

$$= \int_0^\infty e^{-\Psi_n u^{-\epsilon}} \epsilon \Psi_{ni} u^{-(\epsilon+1)} du.$$

Using the result (40) to evaluate the integral above, the probability that a worker commutes to location i conditional on having chosen to live in location n is:

$$\lambda_{ni|n} = \frac{\Psi_{ni}}{\Psi_n} = \frac{B_{ni} \left(\kappa_{ni} P_n^\alpha Q_n^{1-\alpha} \right)^{-\epsilon} (w_i)^\epsilon}{\sum_{s=1}^{S} B_{ns} \left(\kappa_{ns} P_n^\alpha Q_n^{1-\alpha} \right)^{-\epsilon} (w_s)^\epsilon},$$

which simplifies to:

$$\lambda_{ni|n} = \frac{B_{ni} (w_i / \kappa_{ni})^\epsilon}{\sum_{s=1}^{S} B_{ns} (w_s / \kappa_{ns})^\epsilon}. \tag{45}$$

These conditional commuting probabilities provide microeconomic foundations for the reduced-form gravity equations estimated in the empirical literature on commuting patterns.[16] The probability that a resident of location n commutes to location i depends on the wage at i and the amenities and commuting costs from living in n and working in i in the numerator ("bilateral resistance"). But it also depends on the wage at all other workplaces s and the amenities and commuting costs from living in n and commuting to all other workplaces s in the denominator ("multilateral resistance").

Labor market clearing requires that the measure of workers employed in each location i (L_{Mi}) equals the sum across all locations n of their measures of residents (L_{Rn}) times their conditional probabilities of commuting to i (λ_{ni}):

$$L_{Mi} = \sum_{n=1}^{S} \lambda_{ni|n} L_{Rn} \tag{46}$$

$$= \sum_{n=1}^{S} \frac{B_{ni} (w_i / \kappa_{ni})^\epsilon}{\sum_{s=1}^{S} B_{ns} (w_s / \kappa_{ns})^\epsilon} L_{Rn},$$

where, since there is a continuous measure of workers residing in each location, there is no uncertainty in the supply of workers to each employment location.

Expected worker income conditional on living in location n is equal to the wages in all possible workplace locations weighted by the probabilities of commuting to those locations conditional on

---

[16]See also McFadden (1975). For reduced-form evidence of the explanatory power of a gravity equation for commuting flows, see for example Erlander and Stewart (1990) and Sen and Smith (1995).

living in n:

$$\bar{v}_n = \mathbb{E}\left[w|n\right] \tag{47}$$

$$= \sum_{i=1}^{S} \lambda_{ni|n} w_i,$$

$$= \sum_{i=1}^{S} \frac{B_{ni}\left(w_i/\kappa_{ni}\right)^{\epsilon}}{\sum_{s=1}^{S} B_{ns}\left(w_s/\kappa_{ns}\right)^{\epsilon}} w_i,$$

where $\mathbb{E}$ denotes the expectations operator and the expectation is taken over the distribution for idiosyncratic amenities. Intuitively, expected worker income is high in locations that have low commuting costs (low κ_{ns}) to high-wage employment locations.

Finally, another implication of the Fréchet distribution of utility is that the distribution of utility conditional on residing in location n and commuting to location i is the same across all bilateral pairs of locations with positive residents and employment, and is equal to the distribution of utility for the economy as a whole. To establish this result, note that the distribution of utility conditional on residing in location n and commuting to location i is given by:

$$= \frac{1}{\lambda_{ni}} \int_0^u \prod_{s \neq i} G_{ns}(u) \left[\prod_{r \neq n} \prod_s G_{rs}(u)\right] g_{ni}(u) du, \tag{48}$$

$$= \frac{1}{\lambda_{ni}} \int_0^u \left[\prod_{r=1}^{S} \prod_{s=1}^{S} e^{-\Psi_{rs} u^{-\epsilon}}\right] \epsilon \Psi_{ni} u^{-(\epsilon+1)} du,$$

$$= \frac{\Psi}{\Psi_{ni}} \int_0^u e^{-\Psi u^{-\epsilon}} \epsilon \Psi_{ni} u^{-(\epsilon+1)} du,$$

$$= e^{-\Psi u^{\epsilon}}.$$

On the one hand, lower land prices in location n or a higher wage in location i raise the utility of a worker with a given realization of idiosyncratic amenities b, and hence increase the expected utility of residing in n and working in i. On the other hand, lower land prices or a higher wage induce workers with lower realizations of idiosyncratic amenities b to reside in n and work in i, which reduces the expected utility of residing in n and working in i. With a Fréchet distribution of utility, these two effects exactly offset one another. Pairs of residence and employment locations with more attractive characteristics attract more commuters on the extensive margin until expected utility is the same across all pairs of residence and employment locations within the economy.

A.2 Computing Counterfactuals Using Changes

We denote the value of variables in the counterfactual equilibrium by a prime (x') and the relative change of a variable between the initial and the counterfactual equilibrium by a hat $(\hat{x} = x'/x)$. Given the model's parameters $\{\alpha, \sigma, \epsilon, \delta, \kappa\}$ and counterfactual changes in the model's exogenous variables $\{\hat{A}_n, \hat{B}_n, \hat{\kappa}_{ni}, \hat{d}_{ni}\}$, we can solve for the counterfactual changes in the model's endogenous

variables $\{\hat{w}_n, \widehat{\bar{v}}_n, \hat{Q}_n, \hat{\pi}_{ni}, \hat{\lambda}_{ni}, \hat{P}_n, \hat{L}_{Rn}, \hat{L}_{Mn}\}$ from the following system of eight equations (an iterative algorithm is described in the Appendix):

$$\hat{w}_i \hat{L}_{Mi} w_i L_{Mi} = \sum_{n \in N} \pi_{ni} \hat{\pi}_{ni} \widehat{\bar{v}}_n \hat{L}_{Rn} \bar{v}_n L_{Rn}, \tag{49}$$

$$\widehat{\bar{v}}_n \bar{v}_n = \sum_{i \in N} \frac{\lambda_{ni} \hat{B}_{ni} (\hat{w}_i / \hat{\kappa}_{ni})^{\epsilon}}{\sum_{s \in N} \lambda_{ns} \hat{B}_{ns} (\hat{w}_s / \hat{\kappa}_{ns})^{\epsilon}} \hat{w}_i w_i, \tag{50}$$

$$\hat{Q}_n = \widehat{\bar{v}}_n \hat{L}_{Rn}, \tag{51}$$

$$\hat{\pi}_{ni} \pi_{ni} = \frac{\pi_{ni} \hat{L}_{Mi} \left(\hat{d}_{ni} \hat{w}_i / \hat{A}_i \right)^{1-\sigma}}{\sum_{k \in N} \pi_{nk} \hat{L}_{Mk} \left(\hat{d}_{nk} \hat{w}_k / \hat{A}_k \right)^{1-\sigma}}, \tag{52}$$

$$\hat{\lambda}_{ni} \lambda_{ni} = \frac{\lambda_{ni} \hat{B}_{ni} \left(\hat{P}_n^{\alpha} \hat{Q}_n^{1-\alpha} \right)^{-\epsilon} (\hat{w}_i / \hat{\kappa}_{ni})^{\epsilon}}{\sum_{r \in N} \sum_{s \in N} \lambda_{rs} \hat{B}_{rs} \left(\hat{P}_r^{\alpha} \hat{Q}_r^{1-\alpha} \right)^{-\epsilon} (\hat{w}_s / \hat{\kappa}_{rs})^{\epsilon}}, \tag{53}$$

$$\hat{P}_n = \left(\frac{\hat{L}_{Mn}}{\hat{\pi}_{nn}} \right)^{\frac{1}{1-\sigma}} \frac{\hat{d}_{nn} \hat{w}_n}{\hat{A}_n}, \tag{54}$$

$$\hat{L}_{Rn} = \frac{\bar{L}}{L_{Rn}} \sum_i \lambda_{ni} \hat{\lambda}_{ni}, \tag{55}$$

$$\hat{L}_{Mi} = \frac{\bar{L}}{L_{Mi}} \sum_n \lambda_{ni} \hat{\lambda}_{ni}, \tag{56}$$

where these equations correspond to trade balance (49), expected worker income (50), the land market clearing condition (51), trade shares (52), commuting probabilities (53), price indices (54), residential choice probabilities (55) and workplace choice probabilities (56).

A.3 Proof of Proposition 1

Assume balanced trade so $w_i L_{Mi} = \sum_{n \in N} X_{ni}$, and so

$$w_i L_{Mi} = \sum_{n \in N} \frac{\frac{L_{Mi}}{\sigma F} \left(\frac{\sigma}{\sigma-1} \frac{d_{ni} w_i}{A_i} \right)^{1-\sigma}}{P_n^{1-\sigma}} \bar{v}_n L_{Rn}. \tag{57}$$

Using (12) to substitute for the price index, this balanced trade condition can be written as

$$w_i^{\sigma} A_i^{1-\sigma} = \sum_{n \in N} \pi_{nn} d_{ni}^{1-\sigma} \frac{L_{Rn}}{L_{Mn}} w_n^{\sigma-1} A_n^{1-\sigma} \bar{v}_n. \tag{58}$$

Note that balanced trade (57) also can be written as

$$w_i^{\sigma} A_i^{1-\sigma} = \sum_{n \in N} \frac{1}{\sigma F} \left(\frac{\sigma}{\sigma-1} \right)^{1-\sigma} d_{ni}^{1-\sigma} P_n^{\sigma-1} \bar{v}_n L_{Rn}. \tag{59}$$

Note that the residential choice probabilities (16) can be written as

$$\frac{L_{Rn}}{\bar{L}} = \frac{\bar{w}_n^{\epsilon}}{(\bar{U}/\delta)^{\epsilon} P_n^{\alpha\epsilon} Q_n^{(1-\alpha)\epsilon}}, \tag{60}$$

where $\bar{w}_n$ is a measure of commuting market access

$$\bar{w}_n = \left[\sum_{s \in N} B_{ns} \left(w_s / \kappa_{ns} \right)^\epsilon \right]^{\frac{1}{\epsilon}}. \tag{61}$$

Rearranging the residential choice probabilities (60), we obtain the following expression for the price index

$$P_n^{\alpha\epsilon} = \frac{\bar{w}_n^\epsilon}{(\bar{U}/\delta)^\epsilon \, Q_n^{(1-\alpha)\epsilon} \, (L_{Rn}/\bar{L})}. \tag{62}$$

Using land market clearing (5), this expression for the price index can be re-written as

$$P_n^{1-\sigma} = \left(\frac{\bar{w}_n}{\bar{W}} \right)^{\frac{1-\sigma}{\alpha}} \bar{v}_n^{-(1-\sigma)\left(\frac{1-\alpha}{\alpha} \right)} H_n^{(1-\sigma)\left(\frac{1-\alpha}{\alpha} \right)} L_{Rn}^{-(1-\sigma)\left(\frac{1}{\alpha\epsilon} + \frac{1-\alpha}{\alpha} \right)}, \tag{63}$$

where

$$\bar{W} = (\bar{U}/\delta) \left(1 - \alpha \right)^{1-\alpha} \bar{L}^{-1/\epsilon}.$$

Substituting price index expression (63) into (12), we obtain the following expression for the domestic trade share

$$\pi_{nn} = \bar{W}^{\frac{1-\sigma}{\alpha}} \frac{\left(\frac{\sigma}{\sigma-1} \right)^{1-\sigma}}{\sigma F} \bar{w}_n^{\frac{\sigma-1}{\alpha}} w_n^{1-\sigma} L_{Mn} L_{Rn}^{(1-\sigma)\left(\frac{1}{\alpha\epsilon} + \frac{1-\alpha}{\alpha} \right)} \bar{v}_n^{(1-\sigma)\left(\frac{1-\alpha}{\alpha} \right)} H_n^{(\sigma-1)\left(\frac{1-\alpha}{\alpha} \right)} A_n^{\sigma-1}. \tag{64}$$

Now substitute this expression for the domestic trade share in the wage equation (58) to obtain

$$w_n^\sigma A_n^{1-\sigma} = \bar{W}^{\frac{1-\sigma}{\alpha}} \frac{\left(\frac{\sigma}{\sigma-1} \right)^{1-\sigma}}{\sigma F} \left[\sum_{k \in N} \bar{w}_k^{\frac{\sigma-1}{\alpha}} L_{Rk}^{1-(\sigma-1)\left(\frac{1}{\alpha\epsilon} + \frac{1-\alpha}{\alpha} \right)} \bar{v}_k^{1-(\sigma-1)\left(\frac{1-\alpha}{\alpha} \right)} H_k^{(\sigma-1)\left(\frac{1-\alpha}{\alpha} \right)} d_{kn}^{1-\sigma} \right]. \tag{65}$$

Now return to the price index (63) and use the definition of the price index from (12) and commuting market clearing (18)

$$\bar{w}_n^{\frac{1-\sigma}{\alpha}} \bar{v}_n^{-(1-\sigma)\left(\frac{1-\alpha}{\alpha} \right)} H_n^{(1-\sigma)\left(\frac{1-\alpha}{\alpha} \right)} L_{Rn}^{-(1-\sigma)\left(\frac{1}{\alpha\epsilon} + \frac{1-\alpha}{\alpha} \right)} \tag{66}$$

$$= \bar{W}^{\frac{1-\sigma}{\alpha}} \frac{\left(\frac{\sigma}{\sigma-1} \right)^{1-\sigma}}{\sigma F} \left\{ \sum_{k \in N} \left[\sum_{r \in N} \frac{B_{rn} \left(w_n / \kappa_{rn} \right)^\epsilon}{\sum_{s \in N} B_{rs} \left(w_s / \kappa_{rs} \right)^\epsilon} L_{Rr} \right] \left(\frac{d_{nk} w_k}{A_k} \right)^{1-\sigma} \right\}.$$

Together (65) and (66) provide systems of $2N$ equations that determine the $2N$ unknown values of $\{ \mathbf{w}, \mathbf{L}_R \}$. Consider first the system (65). Define the operator T_w by

$$T_w \left(\mathbf{w}; \mathbf{L}_R \right) \equiv \left(\frac{\bar{W}^{\frac{1-\sigma}{\alpha}} \frac{\left(\frac{\sigma}{\sigma-1} \right)^{1-\sigma}}{\sigma F} \left[\sum_{k \in N} \bar{w}_k^{\frac{\sigma-1}{\alpha}} L_{Rk}^{1-(\sigma-1)\left(\frac{1}{\alpha\epsilon} + \frac{1-\alpha}{\alpha} \right)} \bar{v}_k^{1-(\sigma-1)\left(\frac{1-\alpha}{\alpha} \right)} H_k^{(\sigma-1)\left(\frac{1-\alpha}{\alpha} \right)} d_{kn}^{1-\sigma} \right]}{A_i^{1-\sigma}} \right)^{\frac{1}{\sigma}}.$$

Note that for any $\lambda \geq 0$

$$\bar{w}_n \left(\lambda w_n \right) = \lambda \bar{w}_n \left(w_n \right),$$

$$\bar{v}_n \left(\lambda w_n \right) = \lambda \bar{v}_n \left(w_n \right).$$

Hence,

$$T_w\left(\lambda\mathbf{w};\mathbf{L}_R\right) = \lambda^{\left[\frac{\sigma-1}{\alpha}+1-(\sigma-1)\left(\frac{1-\alpha}{\alpha}\right)\right]\frac{1}{\sigma}}T_w\left(\mathbf{w};\mathbf{L}_R\right).$$

Hence, since,

$$\left[\frac{\sigma-1}{\alpha}+1-(\sigma-1)\left(\frac{1-\alpha}{\alpha}\right)\right]\frac{1}{\sigma} = 1 \in (0,1]$$

the results in Fujimoto and Krause (1985) guarantee that there exists a unique fixed point for $\mathbf{w}$ conditional on $\mathbf{L}_R$

$$\mathbf{w} = T_w\left(\mathbf{w};\mathbf{L}_R\right).$$

Denote this fixed point by $\mathbf{w}^{FP}\left(\mathbf{L}_R\right)$. Note also that

$$T_w\left(\mathbf{w};\lambda\mathbf{L}_R\right) = \lambda^{\left[1-(\sigma-1)\left(\frac{1}{\alpha\epsilon}+\frac{1-\alpha}{\alpha}\right)\right]\frac{1}{\sigma}}T_w\left(\mathbf{w};\mathbf{L}_R\right)$$

and so

$$\mathbf{w}^{FP}\left(\lambda\mathbf{L}_R\right) = \lambda^{\left[1-(\sigma-1)\left(\frac{1}{\alpha\epsilon}+\frac{1-\alpha}{\alpha}\right)\right]\frac{1}{\sigma}}\mathbf{w}^{FP}\left(\mathbf{L}_R\right).$$

Now consider the system (66). Define an operator $\tilde{T}_L$ by

$$\tilde{T}_L\left(\mathbf{L}_R;\mathbf{w}\right) \equiv \left(\frac{\overline{W}^{\frac{1-\sigma}{\alpha}}\left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma}}{\sigma F}\left[\sum_{k\in N}\left[\sum_{r\in N}\frac{B_{rn}(w_n/\kappa_{rn})^{\epsilon}}{\sum_{s\in N}B_{rs}(w_s/\kappa_{rs})^{\epsilon}}L_{Rr}\right]\left(\frac{d_{nk}w_k}{A_k}\right)^{1-\sigma}\right] \times \overline{w}_n^{-\frac{1-\sigma}{\alpha}}\overline{v}_n^{(1-\sigma)\left(\frac{1-\alpha}{\alpha}\right)}H_n^{-(1-\sigma)\left(\frac{1-\alpha}{\alpha}\right)}\right)^{\frac{1}{-(1-\sigma)\left(\frac{1}{\alpha\epsilon}+\frac{1-\alpha}{\alpha}\right)}}.$$

Note that

$$\tilde{T}_L\left(\lambda\mathbf{L}_R;\mathbf{w}\right) = \lambda^{\frac{1}{-(1-\sigma)\left(\frac{1}{\alpha\epsilon}+\frac{1-\alpha}{\alpha}\right)}}\tilde{T}_L\left(\mathbf{L}_R;\mathbf{w}\right)$$

and

$$\tilde{T}_L\left(\mathbf{L}_R;\lambda\mathbf{w}\right) = \lambda^{\frac{1-\sigma-\frac{1-\sigma}{\alpha}+(1-\sigma)\left(\frac{1-\alpha}{\alpha}\right)}{-(1-\sigma)\left(\frac{1}{\alpha\epsilon}+\frac{1-\alpha}{\alpha}\right)}}\tilde{T}_L\left(\mathbf{L}_R;\mathbf{w}\right).$$

Combining these results, we can define the operator for the two systems (65) and (66) together as

$$T_L\left(\mathbf{L}_R\right) \equiv \tilde{T}_L\left(\mathbf{L}_R;\mathbf{w}^{FP}\left(\mathbf{L}_R\right)\right).$$

where

$$T_L\left(\lambda\mathbf{L}_R\right) = \lambda^{\frac{1}{-(1-\sigma)\left(\frac{1}{\alpha\epsilon}+\frac{1-\alpha}{\alpha}\right)}+\frac{1-\sigma-\frac{1-\sigma}{\alpha}+(1-\sigma)\left(\frac{1-\alpha}{\alpha}\right)}{-(1-\sigma)\left(\frac{1}{\alpha\epsilon}+\frac{1-\alpha}{\alpha}\right)}\left(1-(\sigma-1)\left(\frac{1}{\alpha\epsilon}+\frac{1-\alpha}{\alpha}\right)\right)\frac{1}{\sigma}}T_L\left(\mathbf{L}_R\right).$$

The results in Fujimoto and Krause (1985) guarantee that there is a unique fixed point $\mathbf{L}_R^{FP}$

$$\mathbf{L}_R^{FP} = T_L\left(\mathbf{L}_R^{FP}\right)$$

to the system of equations (65) and (66) if

$$\frac{1}{-(1-\sigma)\left(\frac{1}{\alpha\epsilon}+\frac{1-\alpha}{\alpha}\right)}+\left[\frac{1-\sigma-\frac{1-\sigma}{\alpha}+(1-\sigma)\left(\frac{1-\alpha}{\alpha}\right)}{-(1-\sigma)\left(\frac{1}{\alpha\epsilon}+\frac{1-\alpha}{\alpha}\right)}\right]\left(1-(\sigma-1)\left(\frac{1}{\alpha\epsilon}+\frac{1-\alpha}{\alpha}\right)\right)\frac{1}{\sigma} \in (0,1].$$

Note that

$$\left[\frac{1-\sigma-\frac{1-\sigma}{\alpha}+(1-\sigma)\left(\frac{1-\alpha}{\alpha}\right)}{-(1-\sigma)\left(\frac{1}{\alpha\epsilon}+\frac{1-\alpha}{\alpha}\right)}\right]\frac{1}{\sigma} = 0.$$

Therefore the above condition reduces to

$$\frac{1}{-(1-\sigma)\left(\frac{1}{\alpha\epsilon}+\frac{1-\alpha}{\alpha}\right)} \in (0,1]$$

Since $-(1-\sigma)\left(\frac{1}{\alpha\epsilon}+\frac{1-\alpha}{\alpha}\right) > 0$ by assumption, the binding constraint is that

$$-(1-\sigma)\left(\frac{1}{\alpha\epsilon}+\frac{1-\alpha}{\alpha}\right) > 1,$$

or

$$\sigma > \frac{1+\epsilon}{1+\epsilon-\alpha\epsilon}. \tag{67}$$

Hence, under the condition imposed in the statement of the proposition, there exists a unique solution to the system of equations (65) and (66).

The only remaining task is to use the labor market clearing condition to guarantee that there exists a unique scalar $\bar{U}$ such that $\sum_{n\in N} L_{Rn} = \bar{L}$. First note that the above result and the definition of the operator $T_L(\mathbf{L}_R)$ guarantees that for any $\lambda > 0$ and any $\bar{U}$

$$\mathbf{L}_{R;\lambda\bar{U}} = T_L\left(\mathbf{L}_{R;\lambda\bar{U}};\lambda\bar{U}\right) = T_L\left(\mathbf{L}_{R;\lambda\bar{U}};\bar{U}\right)\lambda^{\frac{\frac{1-\sigma}{\alpha\sigma}}{-(1-\sigma)\left(\frac{1}{\alpha\epsilon}+\frac{1-\alpha}{\alpha}\right)}} = T_L\left(\mathbf{L}_{R;\lambda\bar{U}};\bar{U}\right)\lambda^{\frac{1}{-\sigma\left(\frac{1}{\epsilon}+1-\alpha\right)}}$$

where we have used the definition of $\bar{W}$. Then,

$$\Gamma\mathbf{L}_{R;\lambda\bar{U}} = T_L\left(\Gamma\mathbf{L}_{R;\lambda\bar{U}};\lambda\bar{U}\right)\Gamma^{1-\frac{1}{-(1-\sigma)\left(\frac{1}{\alpha\epsilon}+\frac{1-\alpha}{\alpha}\right)}} = T\left(\Gamma\mathbf{L}_{R;\lambda\bar{U}};\bar{U}\right)\Gamma^{1-\frac{1}{-(1-\sigma)\left(\frac{1}{\alpha\epsilon}+\frac{1-\alpha}{\alpha}\right)}}\lambda^{\frac{\frac{1-\sigma}{\alpha\sigma}}{-(1-\sigma)\left(\frac{1}{\alpha\epsilon}+\frac{1-\alpha}{\alpha}\right)}}$$

and so $\mathbf{L}_{R;\bar{U}} = \Gamma\mathbf{L}_{R;\lambda\bar{U}}$ if

$$\Gamma = \lambda^{\frac{\sigma-1}{\alpha\sigma}\left(1-(\sigma-1)\left(\frac{1}{\alpha\epsilon}+\frac{1-\alpha}{\alpha}\right)\right)}.$$

Note that since $1-(\sigma-1)\left(\frac{1}{\alpha\epsilon}+\frac{1-\alpha}{\alpha}\right) < 0$, $\mathbf{L}_{R;\bar{U}}$ is monotone decreasing in $\bar{U}$ through this power function. Hence, there exists a unique $\bar{U}$ such that $\sum_{n\in N} L_{Rn} = \bar{L}$.

A.4 Proof of Proposition 2

A.4.1 New Economic Geography Model

We begin by considering our new economic geography model with agglomeration forces through love of variety and increasing returns to scale. The general equilibrium vector $\{w_n, \bar{v}_n, Q_n, L_{Mn}, L_{Rn}, P_n\}$ and scalar $\bar{U}$ solve the following system of equations. First, income equals expenditure on goods produced in each location:

$$w_i L_{Mi} = \sum_{n\in N}\frac{L_{Mi}\left(d_{ni}w_i/A_i\right)^{1-\sigma}}{\sum_{k\in N} L_{Mk}\left(d_{nk}w_k/A_k\right)^{1-\sigma}}\bar{v}_n L_{Rn}. \tag{68}$$

Second, expected worker income depends on wages:

$$\bar{v}_n = \sum_{i\in N}\frac{B_{ni}\left(w_i/\kappa_{ni}\right)^{\epsilon}}{\sum_{s\in N} B_{ns}\left(w_s/\kappa_{ns}\right)^{\epsilon}}w_i. \tag{69}$$

Third, land prices depend on expected worker income and the measure of residents:

$$Q_n = (1 - \alpha) \frac{\bar{v}_n L_{Rn}}{H_n}. \tag{70}$$

Fourth, workplace choice probabilities solve:

$$\frac{L_{Mn}}{\bar{L}} = \frac{\sum_{r \in N} B_{rn} \left(\kappa_{rn} P_r^\alpha Q_r^{1-\alpha} \right)^{-\epsilon} w_n^\epsilon}{\sum_{r \in N} \sum_{s \in N} B_{rs} \left(\kappa_{rs} P_r^\alpha Q_r^{1-\alpha} \right)^{-\epsilon} w_s^\epsilon}. \tag{71}$$

Fifth, residential choice probabilities solve:

$$\frac{L_{Rn}}{\bar{L}} = \frac{\sum_{s \in N} B_{ns} \left(\kappa_{ns} P_n^\alpha Q_n^{1-\alpha} \right)^{-\epsilon} w_s^\epsilon}{\sum_{r \in N} \sum_{s \in N} B_{rs} \left(\kappa_{rs} P_r^\alpha Q_r^{1-\alpha} \right)^{-\epsilon} w_s^\epsilon}. \tag{72}$$

Sixth, price indices solve:

$$P_n = \frac{\sigma}{\sigma - 1} \left(\frac{1}{\sigma F} \right)^{\frac{1}{1-\sigma}} \left[\sum_{i \in N} L_{Mi} \left(d_{ni} w_i / A_i \right)^{1-\sigma} \right]^{\frac{1}{1-\sigma}}. \tag{73}$$

Seventh, expected utility satisfies:

$$\bar{U} = \delta \left[\sum_{r \in N} \sum_{s \in N} B_{rs} \left(\kappa_{rs} P_r^\alpha Q_r^{1-\alpha} \right)^{-\epsilon} w_s^\epsilon \right]^{\frac{1}{\epsilon}}, \tag{74}$$

where $\delta = \Gamma \left(\frac{\epsilon - 1}{\epsilon} \right)$ and $\Gamma \left(\cdot \right)$ is the Gamma function.

A.4.2 Eaton and Kortum (2002) with External Economies of Scale

We consider an Eaton and Kortum (2002) with external economies of scale augmented to incorporate heterogeneity in worker preferences over workplace and residence locations. Utility remains as specified in (1), except that the consumption index (C_n) is defined over a fixed interval of goods $j \in [0, 1]$:

$$C_n = \left[\int_0^1 c_n \left(j \right)^\rho dj \right]^{\frac{1}{\rho}}.$$

Productivity for each good j in each location i is drawn from an independent Fréchet distribution:

$$F_i \left(z \right) = e^{-A_i z^{-\theta}}, \qquad A_i = \tilde{A}_i L_{Mi}^\eta, \qquad \theta > 1,$$

where the scale parameter of this distribution (A_i) depends on the measure of workers (L_{Mi}) and η parameterizes the strength of external economies of scale. The general equilibrium vector $\{w_n, \bar{v}_n, Q_n, L_{Mn}, L_{Rn}, P_n\}$ and scalar $\bar{U}$ solve the following system of equations. First, income equals expenditure on goods produced in each location:

$$w_i L_{Mi} = \sum_{n \in N} \frac{\tilde{A}_i L_{Mi}^\eta \left(d_{ni} w_i \right)^{-\theta}}{\sum_{k \in N} \tilde{A}_k L_{Mk}^\eta \left(d_{nk} w_k \right)^{-\theta}} \bar{v}_n L_{Rn}. \tag{75}$$

Second, expected worker income depends on wages:

$$\bar{v}_n = \sum_{i \in N} \frac{B_{ni} (w_i / \kappa_{ni})^\epsilon}{\sum_{s \in N} B_{ns} (w_s / \kappa_{ns})^\epsilon} w_i. \tag{76}$$

Third, land prices depend on expected worker income and the measure of residents:

$$Q_n = (1 - \alpha) \frac{\bar{v}_n L_{Rn}}{H_n}. \tag{77}$$

Fourth, workplace choice probabilities solve:

$$\frac{L_{Mn}}{\bar{L}} = \frac{\sum_{r \in N} B_{rn} \left(\kappa_{rn} P_r^\alpha Q_r^{1-\alpha}\right)^{-\epsilon} w_n^\epsilon}{\sum_{r \in N} \sum_{s \in N} B_{rs} \left(\kappa_{rs} P_r^\alpha Q_r^{1-\alpha}\right)^{-\epsilon} w_s^\epsilon}. \tag{78}$$

Fifth, residential choice probabilities solve:

$$\frac{L_{Rn}}{\bar{L}} = \frac{\sum_{s \in N} B_{ns} \left(\kappa_{ns} P_n^\alpha Q_n^{1-\alpha}\right)^{-\epsilon} w_s^\epsilon}{\sum_{r \in N} \sum_{s \in N} B_{rs} \left(\kappa_{rs} P_r^\alpha Q_r^{1-\alpha}\right)^{-\epsilon} w_s^\epsilon}. \tag{79}$$

Sixth, price indices solve:

$$P_n = \gamma \left[\sum_{i \in N} \tilde{A}_i L_{Mi}^\eta (d_{ni} w_i)^{-\theta} \right]^{-\frac{1}{\theta}}, \tag{80}$$

where $\gamma = \left[\Gamma \left(\frac{\theta - (\sigma - 1)}{\theta} \right) \right]^{\frac{1}{1-\sigma}}$ and $\Gamma (\cdot)$ denotes the Gamma function. Seventh, expected utility satisfies:

$$\bar{U} = \delta \left[\sum_{r \in N} \sum_{s \in N} B_{rs} \left(\kappa_{rs} P_r^\alpha Q_r^{1-\alpha}\right)^{-\epsilon} w_s^\epsilon \right]^{\frac{1}{\epsilon}}. \tag{81}$$

The system of equations (75)-(81) is isomorphic to the system of equations (68)-(74) under the following restrictions:

$$
\begin{aligned}
\theta^{\text{EK}} &= \sigma^{\text{NEG}} - 1, \\
\eta^{\text{EK}} &= 1, \\
A_i^{\text{EK}} &= \left(A_i^{\text{NEG}} \right)^{\sigma^{\text{NEG}} - 1}, \\
\gamma^{\text{EK}} &= \frac{\sigma^{\text{NEG}}}{\sigma^{\text{NEG}} - 1} \left(\frac{1}{\sigma^{\text{NEG}} F^{\text{NEG}}} \right)^{\frac{1}{1 - \sigma^{\text{NEG}}}}.
\end{aligned}
$$

Under these parameter restrictions, both models generate the same general equilibrium vector $\{w_n, \bar{v}_n, Q_n, L_{Mn}, L_{Rn}, P_n\}$ and scalar $\bar{U}$.

A.4.3 Armington (1969) with External Economies of Scale

We consider an Armington (1969) model with external economies of scale augmented to incorporate heterogeneity in worker preferences over workplace and residence locations. Utility remains as

specified in (1), except that the consumption index (C_n) is defined over goods that are horizontally differentiated by location of origin:

$$C_n = \left[\sum_{i \in N} C_i^\rho \right]^{\frac{1}{\rho}}.$$

The goods supplied by each location are produced under conditions of perfect competition and external economies of scale such that the "cost inclusive of freight" (cif) price of good produced in location i and consumed in location n is:

$$P_{ni} = \frac{d_{ni} w_i}{A_i}, \qquad A_i = \tilde{A}_i L_{Mi}^\eta.$$

The general equilibrium vector $\{w_n, \bar{v}_n, Q_n, L_{Mn}, L_{Rn}, P_n\}$ and scalar $\bar{U}$ solve the following system of equations. First, income equals expenditure on goods produced in each location:

$$w_i L_{Mi} = \sum_{n \in N} \frac{A_i^{\sigma-1} L_{Mi}^{\eta(\sigma-1)} (d_{ni} w_i)^{1-\sigma}}{\sum_{k \in N} A_k^{\sigma-1} L_{Mk}^{\eta(\sigma-1)} (d_{nk} w_k)^{1-\sigma}} \bar{v}_n L_{Rn}. \tag{82}$$

Second, expected worker income depends on wages:

$$\bar{v}_n = \sum_{i \in N} \frac{B_{ni} (w_i / \kappa_{ni})^\epsilon}{\sum_{s \in N} B_{ns} (w_s / \kappa_{ns})^\epsilon} w_i. \tag{83}$$

Third, land prices depend on expected worker income and the measure of residents:

$$Q_n = (1 - \alpha) \frac{\bar{v}_n L_{Rn}}{H_n}. \tag{84}$$

Fourth, workplace choice probabilities solve:

$$\frac{L_{Mn}}{\bar{L}} = \frac{\sum_{r \in N} B_{rn} \left(\kappa_{rn} P_r^\alpha Q_r^{1-\alpha} \right)^{-\epsilon} w_n^\epsilon}{\sum_{r \in N} \sum_{s \in N} B_{rs} \left(\kappa_{rs} P_r^\alpha Q_r^{1-\alpha} \right)^{-\epsilon} w_s^\epsilon}. \tag{85}$$

Fifth, residential choice probabilities solve:

$$\frac{L_{Rn}}{\bar{L}} = \frac{\sum_{s \in N} B_{ns} \left(\kappa_{ns} P_n^\alpha Q_n^{1-\alpha} \right)^{-\epsilon} w_s^\epsilon}{\sum_{r \in N} \sum_{s \in N} B_{rs} \left(\kappa_{rs} P_r^\alpha Q_r^{1-\alpha} \right)^{-\epsilon} w_s^\epsilon}. \tag{86}$$

Sixth, price indices solve:

$$P_n = \left[\sum_{i \in N} A_i^{\sigma-1} L_{Mi}^{\eta(\sigma-1)} (d_{ni} w_i)^{1-\sigma} \right]^{\frac{1}{1-\sigma}}. \tag{87}$$

Seventh, expected utility satisfies:

$$\bar{U} = \delta \left[\sum_{r \in N} \sum_{s \in N} B_{rs} \left(\kappa_{rs} P_r^\alpha Q_r^{1-\alpha} \right)^{-\epsilon} w_s^\epsilon \right]^{\frac{1}{\epsilon}}. \tag{88}$$

The system of equations (82)-(88) is isomorphic to the system of equations (68)-(74) under the following restrictions:

$$
\begin{aligned}
\sigma^{\text{AR}} &= \sigma^{\text{NEG}}, \\
\eta^{\text{EK}} &= \frac{1}{\sigma^{\text{NEG}} - 1}, \\
A_i^{\text{AR}} &= A_i^{\text{NEG}}, \\
1 &= \frac{\sigma^{\text{NEG}}}{\sigma^{\text{NEG}} - 1} \left(\frac{1}{\sigma^{\text{NEG}} F^{\text{NEG}}} \right)^{\frac{1}{1 - \sigma^{\text{NEG}}}}.
\end{aligned}
$$

Under these parameter restrictions, both models generate the same general equilibrium vector $\{w_n, \bar{v}_n, Q_n, L_{Mn}, L_{Rn}, P_n\}$ and scalar $\bar{U}$.

A.5 Proof of Proposition 3

Note that the goods market clearing condition (23) can be written as the following excess demand system:

$$
D_i(\tilde{\mathbf{A}}) = w_i L_{Mi} - \sum_{n \in N} \frac{\tilde{A}_i L_{Mi} (d_{ni} w_i)^{1-\sigma}}{\sum_{k \in N} \tilde{A}_k L_{Mk} (d_{nk} w_k)^{1-\sigma}} \bar{v}_n L_{Rn} = 0, \tag{89}
$$

where $\tilde{A}_i = A_i^{\sigma-1}$ and $\{w_i, L_{Mi}, \bar{v}_n, L_{Rn}, d_{ni}\}$ have already been determined from the observed data or our parameterization of trade costs. This excess demand system exhibits the following properties in $\tilde{A}_i$:

Property (i): $D(\tilde{\mathbf{A}})$ is continuous, as follows immediately from inspection of (89).

Property (ii): $D(\tilde{\mathbf{A}})$ is homogenous of degree zero, as follows immediately from inspection of (89).

Property (iii): $\sum_{i \in N} D_i(\tilde{\mathbf{A}}) = 0$ for all $\tilde{\mathbf{A}} \in \Re_+^N$. This property can be established by noting:

$$
\begin{aligned}
\sum_{i \in N} D_i(\tilde{\mathbf{A}}) &= \sum_{i \in N} w_i L_{Mi} - \sum_{n \in N} \frac{\sum_{i \in N} \tilde{A}_i L_{Mi} (d_{ni} w_i)^{1-\sigma}}{\sum_{k \in N} \tilde{A}_k L_{Mk} (d_{nk} w_k)^{1-\sigma}} \bar{v}_n L_{Rn}, \\
&= \sum_{i \in N} w_i L_{Mi} - \sum_{n \in N} \bar{v}_n L_{Rn}, \\
&= 0.
\end{aligned}
$$

Property (iv): $D(\tilde{\mathbf{A}})$ exhibits gross substitution:

$$
\frac{\partial D_i(\tilde{\mathbf{A}})}{\partial \tilde{A}_r} > 0 \qquad \text{for all } i, r, \neq i, \qquad\qquad\qquad \text{for all } \tilde{\mathbf{A}} \in \Re_+^N,
$$

$$
\frac{\partial D_i(\tilde{\mathbf{A}})}{\partial \tilde{A}_i} < 0 \qquad \text{for all } i, \qquad\qquad\qquad\qquad \text{for all } \tilde{\mathbf{A}} \in \Re_+^N.
$$

This property can be established by noting:

$$
\frac{\partial D_i(\tilde{\mathbf{A}})}{\partial \tilde{A}_r} = \sum_{n \in N} \frac{L_{Mr} (d_{nr} w_r)^{1-\sigma} \tilde{A}_i L_{Mi} (d_{ni} w_i)^{1-\sigma}}{\left[\sum_{k \in N} \tilde{A}_k L_{Mk} (d_{nk} w_k)^{1-\sigma} \right]^2} \bar{v}_n L_{Rn} > 0.
$$

and using homogeneity of degree zero, which implies:

$$
\nabla D(\tilde{\mathbf{A}}) \tilde{\mathbf{A}} = 0,
$$

50

and hence:

$$\frac{\partial D_i\left(\tilde{\mathbf{A}}\right)}{\partial \tilde{A}_i} < 0 \qquad \text{for all } \tilde{\mathbf{A}} \in \Re_+^N.$$

Therefore we have established gross substitution. We now use these five properties to establish that the system of equations (89) has at most one (normalized) solution. Gross substitution implies that $D\left(\tilde{\mathbf{A}}\right) = D\left(\tilde{\mathbf{A}}'\right)$ cannot occur whenever $\tilde{\mathbf{A}}$ and $\tilde{\mathbf{A}}'$ are two technology vectors that are not colinear. By homogeneity of degree zero, we can assume $\tilde{\mathbf{A}}' \geq \tilde{\mathbf{A}}$ and $\tilde{A}_i = \tilde{A}_i'$ for some i. Now consider altering the productivity vector $\tilde{\mathbf{A}}'$ to obtain the productivity vector $\tilde{\mathbf{A}}$ in $N-1$ steps, lowering (or keeping unaltered) the productivity of all the other $N-1$ locations $n \neq i$ one at a time. By gross substitution, the excess demand in location i cannot decrease in any step, and because $\tilde{\mathbf{A}} \neq \tilde{\mathbf{A}}'$, it will actually increase in at least one step. Hence $D\left(\tilde{\mathbf{A}}\right) > D\left(\tilde{\mathbf{A}}'\right)$ and we have a contradiction.

We next establish that there exists a productivity vector $\tilde{\mathbf{A}}^* \in \Re_+^N$ such that $D(\tilde{\mathbf{A}}^*) = 0$. By homogeneity of degree zero, we can restrict our search for this productivity vector to the unit simplex $\Delta = \left\{\tilde{\mathbf{A}} \in \Re_+^N : \sum_{i \in N} \tilde{A}_i = 1\right\}$. Define on Δ the function $D^+\left(\cdot\right)$ by $D_i^+\left(\tilde{\mathbf{A}}\right) = \max\left\{D_i\left(\tilde{\mathbf{A}}\right), 0\right\}$. Note that $D^+\left(\cdot\right)$ is continuous. Denote $\alpha\left(\tilde{\mathbf{A}}\right) = \sum_{i \in N}\left[\tilde{A}_i + D_i^+\left(\tilde{A}_i\right)\right]$. We have $\alpha\left(\tilde{\mathbf{A}}\right) \geq 1$ for all $\tilde{\mathbf{A}}$.

Define a continuous function $f\left(\cdot\right)$ from the closed convex set Δ into itself by:

$$f\left(\tilde{\mathbf{A}}\right) = \left[1/\alpha\left(\tilde{\mathbf{A}}\right)\right]\left[\tilde{\mathbf{A}} + D^+\left(\tilde{\mathbf{A}}\right)\right].$$

Note that this fixed-point function tends to increase the productivities of locations with excess demand. By Brouwer's Fixed-point Theorem, there exists $\tilde{\mathbf{A}}^* \in \Delta$ such that $\tilde{\mathbf{A}}^* = f\left(\tilde{\mathbf{A}}^*\right)$.

Since $\sum_{i \in N} D_i\left(\tilde{\mathbf{A}}\right) = 0$, it cannot be the case that $D_i\left(\tilde{\mathbf{A}}\right) > 0$ for all $i \in N$ or $D_i\left(\tilde{\mathbf{A}}\right) < 0$ for all $i \in N$. Additionally, if $D_i\left(\tilde{\mathbf{A}}\right) > 0$ for some i and $D_r\left(\tilde{\mathbf{A}}\right) < 0$ for some $r \neq i$, $\tilde{\mathbf{A}} \neq f\left(\tilde{\mathbf{A}}\right)$. It follows that at the fixed point for productivity, $\tilde{\mathbf{A}}^* = f\left(\tilde{\mathbf{A}}^*\right)$, and $D_i\left(\tilde{\mathbf{A}}^*\right) = 0$ for all i. It follows that there exists a unique vector of unobserved productivities $(\tilde{\mathbf{A}})$ that solves the excess demand system (89).

A.6 Proof of Proposition 4

Note that the commuting probability (25) can be written as the following excess demand system:

$$D_i(\mathbf{B}) = \lambda_{ni} - \frac{B_{ni}\kappa_{ni}^{-\epsilon}\left(\frac{L_{Mn}}{\pi_{nn}}\right)^{-\frac{\alpha\epsilon}{\sigma-1}}A_n^{\alpha\epsilon}w_n^{-\alpha\epsilon}\bar{v}_n^{-\epsilon(1-\alpha)}\left(\frac{L_{Rn}}{H_n}\right)^{-\epsilon(1-\alpha)}w_i^{\epsilon}}{\sum_{r \in N}\sum_{s \in N}B_{rs}\kappa_{rs}^{-\epsilon}\left(\frac{L_{Mr}}{\pi_{rr}}\right)^{-\frac{\alpha\epsilon}{\sigma-1}}A_r^{\alpha\epsilon}w_r^{-\alpha\epsilon}\bar{v}_r^{-\epsilon(1-\alpha)}\left(\frac{L_{Rr}}{H_r}\right)^{-\epsilon(1-\alpha)}w_s^{\epsilon}} = 0, \qquad (90)$$

where $\{w_i, L_{Mi}, \bar{v}_n, L_{Rn}, \kappa_{ni}, \pi_{nn}, A_n, H_n\}$ have already been determined from the observed data or our parameterization of commuting costs. Note that the excess demand system (90) exhibits the same properties in $\mathbf{B}$ as the excess demand system (89) exhibits in $\tilde{\mathbf{A}}$. It follows that there exists a unique vector of unobserved productivities $(\tilde{\mathbf{A}})$ that solves the excess demand system (89).

A.7 An iterative algorithm for counterfactual equilibrium

Given the model's parameters $\{\alpha, \sigma, \epsilon, \delta, \kappa\}$ and changes in the exogenous variables of the model $\{\hat{A}_n, \hat{B}_n, \hat{\kappa}_{ni}, \hat{d}_{ni}\}$, we can solve for the resulting counterfactual changes in the endogenous variables

of the model $\{\hat{w}_n, \widehat{\overline{v}}_n, \hat{Q}_n, \hat{\pi}_{ni}, \hat{\lambda}_{ni}, \hat{P}_n, \hat{L}_{Rn}, \hat{L}_{Mn}\}$ from the system of eight equations (49)-(56). We solve this system of equations using the following iterative algorithm. We first conjecture changes in workplace wages and commuting probabilities at iteration t, $\hat{w}_i^{(t)}$ and $\hat{\lambda}_{ni}^{(t)}$. We next update these conjectures to $\hat{w}_i^{(t+1)}$ and $\hat{\lambda}_{ni}^{(t+1)}$ using the current guesses and data. We start by computing:

$$\widehat{\overline{v}}_n^{(t)} = \frac{1}{\overline{v}_n} \sum_{i \in N} \frac{\hat{B}_{ni} \lambda_{ni} \left(\hat{w}_i^{(t)}/\hat{\kappa}_{ni} \right)^{\epsilon}}{\sum_{s \in N} \hat{B}_{ns} \lambda_{ns} \left(\hat{w}_s^{(t)}/\hat{\kappa}_{ns} \right)^{\epsilon}} \hat{w}_i^{(t)} w_i, \tag{91}$$

$$\hat{L}_{Mi}^{(t)} = \frac{\bar{L}}{L_{Mi}} \sum_{n} \lambda_{ni} \hat{\lambda}_{ni}^{(t)}, \tag{92}$$

$$\hat{L}_{Rn}^{(t)} = \frac{\bar{L}}{L_{Rn}} \sum_{i} \lambda_{ni} \hat{\lambda}_{ni}^{(t)}, \tag{93}$$

which are only a function of data and current guesses. We use (91) and (93) in (51) to compute:

$$\hat{Q}_n^{(t)} = \widehat{\overline{v}}_n^{(t)} \hat{L}_{Rn}^{(t)}. \tag{94}$$

We use (92) and (52) to compute:

$$\hat{\pi}_{ni}^{(t)} = \frac{\hat{L}_{Mi}^{(t)} \left(\hat{d}_{ni} \hat{w}_i^{(t)}/\hat{A}_i \right)^{1-\sigma}}{\sum_{k \in N} \pi_{nk} \hat{L}_{Mk}^{(t)} \left(\hat{d}_{nk} \hat{w}_k^{(t)}/\hat{A}_k \right)^{1-\sigma}}. \tag{95}$$

We use (92), (95) and (54) to compute:

$$\hat{P}_n^{(t)} = \left(\frac{\hat{L}_{Mn}^{(t)}}{\hat{\pi}_{nn}^{(t)}} \right)^{\frac{1}{1-\sigma}} \frac{\hat{w}_n^{(t)}}{\hat{A}_n}. \tag{96}$$

We use (91)-(96) to rewrite (49) and (53) as:

$$\tilde{w}_i^{(t+1)} = \frac{1}{Y_i \hat{L}_{Mi}^{(t)}} \sum_{n \in N} \pi_{ni} \hat{\pi}_{ni}^{(t)} \widehat{\overline{v}}_n^{(t)} \hat{L}_{Rn}^{(t)} Y_n, \tag{97}$$

$$\tilde{\lambda}_{ni}^{(t+1)} = \frac{\hat{B}_{ni} \left(\hat{P}_n^{(t)\alpha} \hat{Q}_n^{(t)1-\alpha} \right)^{-\epsilon} \left(\hat{w}_i^{(t)}/\hat{\kappa}_{ni} \right)^{\epsilon}}{\sum_{r \in N} \sum_{s \in N} \hat{B}_{rs} \lambda_{rs} \left(\hat{P}_r^{(t)\alpha} \hat{Q}_r^{(t)1-\alpha} \right)^{-\epsilon} \left(\hat{w}_s^{(t)}/\hat{\kappa}_{rs} \right)^{\epsilon}}. \tag{98}$$

Finally, we update our conjectures for wages and commuting probabilities using:

$$\hat{w}_i^{(t+1)} = \zeta \hat{w}_i^{(t)} + (1-\zeta) \tilde{w}_i^{(t+1)}, \tag{99}$$

$$\hat{\lambda}_i^{(t+1)} = \zeta \hat{\lambda}_i^{(t)} + (1-\zeta) \tilde{\lambda}_i^{(t+1)}, \tag{100}$$

where $\zeta \in (0,1)$ is an adjustment factor.

A.8 Partial Equilibrium Elasticities

Wage Elasticity: Totally differentiating the goods market clearing condition (11), we have:

$$\frac{dw_n}{w_n} w_n L_{Mn} + \frac{dL_{Mn}}{L_{Mn}} w_n L_{Mn} = \sum_{r \in N} (1 - \pi_{rn}) \pi_{rn} \bar{v}_r L_{Rr} \frac{dL_{Mn}}{L_{Mn}} - \sum_{r \in N} \sum_{s \in N} \pi_{rs} \pi_{rn} \bar{v}_r L_{Rr} \frac{dL_{Ms}}{L_{Ms}}$$
$$- (\sigma - 1) \sum_{r \in N} (1 - \pi_{rn}) \pi_{rn} \bar{v}_r L_{Rr} \frac{dw_n}{w_n} + (\sigma - 1) \sum_{r \in N} \sum_{s \in N} \pi_{rs} \pi_{rn} \bar{v}_r L_{Rr} \frac{dw_s}{w_s}$$
$$+ (\sigma - 1) \sum_{r \in N} (1 - \pi_{rn}) \pi_{rn} \bar{v}_r L_{Rr} \frac{dA_n}{A_n} - (\sigma - 1) \sum_{r \in N} \sum_{s \in N} \pi_{rs} \pi_{rn} \bar{v}_r L_{Rr} \frac{dA_s}{A_s}$$
$$+ \sum_{r \in N} \pi_{rn} \bar{v}_r L_{Rr} \frac{d\bar{v}_r}{\bar{v}_r} + \sum_{r \in N} \pi_{rn} \bar{v}_r L_{Rr} \frac{dL_{Rr}}{L_{Rr}}.$$

To consider the direct effect of a productivity shock in location n on wages, employment and residents in that location, holding constant all other endogenous variables at their values in the initial equilibrium, we set $dA_s = dw_s = dL_{Ms} = dL_{Rs} = 0$ for $s \neq n$ and $d\bar{v}_r = 0$ for all r, which yields:

$$\frac{dw_n}{w_n} w_n L_{Mn} + \frac{dL_{Mn}}{L_{Mn}} w_n L_{Mn} = \sum_{r \in N} (1 - \pi_{rn}) \pi_{rn} \bar{v}_r L_{Rr} \frac{dL_{Mn}}{L_{Mn}} - (\sigma - 1) \sum_{r \in N} (1 - \pi_{rn}) \pi_{rn} \bar{v}_r L_{Rr} \frac{dw_n}{w_n}$$
$$+ (\sigma - 1) \sum_{r \in N} (1 - \pi_{rn}) \pi_{rn} \bar{v}_r L_{Rr} \frac{dA_n}{A_n} + \pi_{nn} \bar{v}_n L_{Rn} \frac{dL_{Rn}}{L_{Rn}}.$$

This implies:

$$\frac{dw_n}{dA_n} \frac{A_n}{w_n} + \frac{dL_{Mn}}{dA_n} \frac{A_n}{L_{Mn}} = \sum_{r \in N} (1 - \pi_{rn}) \frac{\pi_{rn} \bar{v}_r L_{Rr}}{w_n L_{Mn}} \left(\frac{dL_{Mn}}{dA_n} \frac{A_n}{L_{Mn}} \right) - (\sigma - 1) \sum_{r \in N} (1 - \pi_{rn}) \frac{\pi_{rn} \bar{v}_r L_{Rr}}{w_n L_{Mn}} \left(\frac{dw_n}{dA_n} \frac{A_n}{w_n} \right)$$
$$+ (\sigma - 1) \sum_{r \in N} (1 - \pi_{rn}) \frac{\pi_{rn} \bar{v}_r L_{Rr}}{w_n L_{Mn}} + \frac{\pi_{nn} \bar{v}_n L_{Rn}}{w_n L_{Mn}} \frac{dL_{Rn}}{dA_n} \frac{A_n}{L_{Rn}},$$

which can be re-written as:

$$\frac{dw_n}{dA_n} \frac{A_n}{w_n} + \left(\frac{dL_{Mn}}{dw_n} \frac{w_n}{L_{Mn}} \right) \left(\frac{dw_n}{dA_n} \frac{A_n}{w_n} \right) = \sum_{r \in N} (1 - \pi_{rn}) \frac{\pi_{rn} \bar{v}_r L_{Rr}}{w_n L_{Mn}} \left(\frac{dL_{Mn}}{dw_n} \frac{w_n}{L_{Mn}} \right) \left(\frac{dw_n}{dA_n} \frac{A_n}{w_n} \right)$$
$$- (\sigma - 1) \sum_{r \in N} (1 - \pi_{rn}) \frac{\pi_{rn} \bar{v}_r L_{Rr}}{w_n L_{Mn}} \left(\frac{dw_n}{dA_n} \frac{A_n}{w_n} \right) + (\sigma - 1) \sum_{r \in N} (1 - \pi_{rn}) \frac{\pi_{rn} \bar{v}_r L_{Rr}}{w_n L_{Mn}}$$
$$+ \frac{\pi_{nn} \bar{v}_n L_{Rn}}{w_n L_{Mn}} \left(\frac{dL_{Rn}}{dw_n} \frac{w_n}{L_{Rn}} \right) \left(\frac{dw_n}{dA_n} \frac{A_n}{w_n} \right),$$

where we have used the fact that productivity does not directly enter the commuting market clearing condition (18) and the residential choice probabilities (16) and hence employment and residents only change to the extent that wages change as a result of the productivity shock. Rearranging this expression, we obtain the partial equilibrium elasticity in the main text above:

$$\frac{dw_n}{dA_n} \frac{A_n}{w_n} = \frac{(\sigma - 1) \sum_{r \in N} (1 - \pi_{rn}) \xi_{rn}}{[1 + (\sigma - 1) \sum_{r \in N} (1 - \pi_{rn}) \xi_{rn}] + [1 - \sum_{r \in N} (1 - \pi_{rn}) \xi_{rn}] \frac{dL_{Mn}}{dw_n} \frac{w_n}{L_{Mn}} - \xi_{nn} \frac{dL_{Rn}}{dw_n} \frac{w_n}{L_{Rn}}}, \quad (101)$$

where $\xi_{rn} = \pi_{rn} \bar{v}_r L_{Rr} / w_n L_{Mn}$ is the share of expenditure from location r in location n's income.

Employment Elasticity: Totally differentiating the commuting market clearing condition (18), we have:

$$\frac{dL_{Mn}}{L_{Mn}} = \epsilon \sum_{r \in N} (1 - \lambda_{rn|r}) \frac{dw_n}{w_n} \frac{\lambda_{rn|r} L_{Rr}}{L_{Mn}} - \epsilon \sum_{r \in N} \sum_{s \neq n} \lambda_{rs|r} \frac{dw_s}{w_s} \frac{L_{Mrn}}{L_{Mn}}$$
$$+ \sum_r \frac{dL_{Rr}}{L_{Rr}} \frac{L_{Mrn}}{L_{Mn}}.$$

To consider the direct effect of a productivity shock in location n on its employment and residents through a higher wage in that location, holding constant all other endogenous variables at their values in the initial equilibrium, we set $dw_s = dL_{Ms} = dL_{Rs} = 0$ for $s \neq n$, which yields:

$$\frac{dL_{Mn}}{L_{Mn}} = \epsilon \sum_{r \in N} (1 - \lambda_{rn|r}) \frac{\lambda_{rn|r} L_{Rr}}{L_{Mn}} \frac{dw_n}{w_n} + \frac{\lambda_{nn|n} L_{Rn}}{L_{Mn}} \frac{dL_{Rn}}{L_{Rn}}.$$

Rearranging this expression, we obtain the partial equilibrium elasticity in the main text above:

$$\frac{dL_{Mn}}{dw_n}\frac{w_n}{L_{Mn}} = \epsilon \sum_{r \in N} \left(1 - \lambda_{rn|r}\right) \vartheta_{rn} + \vartheta_{nn} \left(\frac{dL_{Rn}}{dw_n}\frac{w_n}{L_{Rn}}\right), \tag{102}$$

where $\vartheta_{rn} = \lambda_{rn|r} L_{Rr} / L_{Mn}$ is the share of commuters from source r in location n's employment.

Residents Elasticity: Totally differentiating the residential choice probability (λ_{Rn} in (16)), we have:

$$\begin{aligned}
\frac{dL_{Rn}}{L_{Rn}}\frac{L_{Rn}}{\bar{L}} &= -\epsilon\alpha\left(1 - \lambda_{Rn}\right)\lambda_{Rn}\frac{dP_n}{P_n} + \epsilon\alpha\sum_{r \neq n}\lambda_{Rr}\lambda_{Rn}\frac{dP_r}{P_r} \\
&\quad -\epsilon\left(1 - \alpha\right)\left(1 - \lambda_{Rn}\right)\lambda_{Rn}\frac{dQ_n}{Q_n} + \epsilon\left(1 - \alpha\right)\sum_{r \neq n}\lambda_{Rr}\lambda_{Rn}\frac{dQ_r}{Q_r} \\
&\quad +\epsilon\lambda_{nn}\frac{dw_n}{w_n} - \epsilon\lambda_{Mn}\lambda_{Rn}\frac{dw_n}{w_n} - \epsilon\sum_{s \neq n}\lambda_{Ms}\lambda_{Rn}\frac{dw_s}{w_s}.
\end{aligned}$$

To consider the direct effect of a productivity shock in location n on its residents through a higher wage in that location, holding constant all other endogenous variables at their values in the initial equilibrium, we set $dP_r = dQ_r = 0$ for all r and $dw_s = 0$ for $s \neq n$, which yields:

$$\frac{dL_{Rn}}{L_{Rn}}\frac{L_{Rn}}{\bar{L}} = \epsilon\left(\lambda_{nn} - \lambda_{Mn}\lambda_{Rn}\right)\frac{dw_n}{w_n}.$$

This implies:

$$\frac{dL_{Rn}}{dw_n}\frac{w_n}{L_{Rn}} = \epsilon\left(\frac{\lambda_{nn}}{\lambda_{Rn}} - \lambda_{Mn}\right), \tag{103}$$

which corresponds to the expression in the main text above. Using the residents elasticity (103) in the employment elasticity (102), and using the residents and employment elasticities ((103) and (102) respectively) in the wage elasticity (101), we obtain the following partial equilibrium elasticities for the productivity shock:

$$\frac{dL_{Rn}}{dw_n}\frac{w_n}{L_{Rn}} = \epsilon\left(\frac{\lambda_{nn}}{\lambda_{Rn}} - \lambda_{Mn}\right),$$

$$\frac{dL_{Mn}}{dw_n}\frac{w_n}{L_{Mn}} = \epsilon\sum_{r \in N}\left(1 - \lambda_{rn|r}\right)\vartheta_{rn} + \vartheta_{nn}\epsilon\left(\frac{\lambda_{nn}}{\lambda_{Rn}} - \lambda_{Mn}\right),$$

$$\frac{dw_n}{dA_n}\frac{A_n}{w_n} = \frac{(\sigma-1)\sum_{r \in N}(1-\pi_{rn})\xi_{rn}}{\left[1+(\sigma-1)\sum_{r \in N}(1-\pi_{rn})\xi_{rn}\right]+\left[1-\sum_{r \in N}(1-\pi_{rn})\xi_{rn}\right]\left[\epsilon\sum_{r \in N}\left(1-\lambda_{rn|r}\right)\vartheta_{rn}+\epsilon\vartheta_{nn}\left(\frac{\lambda_{nn}}{\lambda_{Rn}}-\lambda_{Mn}\right)\right]-\xi_{nn}\epsilon\left(\frac{\lambda_{nn}}{\lambda_{Rn}}-\lambda_{Mn}\right)}.$$

References

[1] Ahlfeldt, Gabriel, Stephen Redding, Daniel Sturm and Nikolaus Wolf (2015) "The Economics of Density: Evidence from the Berlin Wall," *Econometrica*, forthcoming.

[2] Allen, T., and C. Arkolakis (2014a) "Trade and the Topography of the Spatial Economy," *Quarterly Journal of Economics*, 129(3), 1085-1140.

[3] Allen, T., and C. Arkolakis (2014b) "Universal Gravity," *NBER Working Paper*, 20787.

[4] Allen, T., C. Arkolakis and X. Li (2015) "Optimal City Structure," Yale University, mimeograph.

[5] Alonso, William (1964) *Location and Land Use*, Cambridge MA: Harvard.

[6] Arkolakis, Costas, Arnaud Costinot and Andres Rodriguez-Clare (2012) "New Trade Models, Same Old Gains," *American Economic Review*, 102(1), 94-130.

[7] Armington, Paul S. (1969) "A Theory of Demand for Products Distinguished by Place of Production," *IMF Staff Papers*, 16(1), 159-178.

[8] Artuc, E., S. Chaudhuri, and J. McLaren (2010) "Trade Shocks and Labor Adjustment: A Structural Empirical Approach," *American Economic Review*, 100(3), 1008-1045.

[9] Autor David and David Dorn (2013) "The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market," *American Economic Review*, 103(5), 1553-1597.

[10] Autor David, David Dorn and Gordon Hanson (2013) "The China Syndrome: Local Labor Market Effects of Import Competition in the United States," *American Economic Review*, 103(6), 2121-2168.

[11] Bartik, Timothy J. (1991) "Who Benefits from State and Local Economic Development Policies?" W.E. Upjohn Institute for Employment Research, Upjohn Press.

[12] Blanchard, Olivier Jean, and Lawrence F. Katz (1992) "Regional Evolutions," *Brookings Papers on Economic Activity*, 23(1), 1-76.

[13] Bernard, Andrew B., Jonathan Eaton, J. Bradford Jensen, and Samuel S. Kortum (2003) "Plants and Productivity in International Trade," *American Economic Review*, 93(4), 1268-1290.

[14] Bound, John and Holzer, Harry J (2000) "Demand Shifts, Population Adjustments, and Labor Market Outcomes during the 1980s," *Journal of Labor Economics*, 18(1), 20-54.

[15] Busso, M., J. Gregory, and P. Kline (2013) "Assessing the Incidence and Efficiency of a Prominent Place Based Policy," *American Economic Review*, 103(2), 897-947.

[16] Caliendo, L. and F. Parro (2015) "Estimates of the Trade and Welfare Effects of NAFTA," *Review of Economic Studies*, 82(1), 1-44.

[17] Caliendo, L., F. Parro, E. Rossi-Hansberg, and P.-D. Sarte (2014) "The Impact of Regional and Sectoral Productivity Changes on the U.S. Economy," *NBER Working Paper*, 20168.

[18] Dekle, R., J. Eaton, and S. Kortum (2007) "Unbalanced Trade," *American Economic Review*, 97(2), 351-355.

[19] Desmet, Klaus and Esteban Rossi-Hansberg (2013) "Urban Accounting and Welfare," *American Economic Review*, 103:6, 2296-2327.

[20] Desmet, Klaus and Esteban Rossi-Hansberg (2014) "Spatial Development," *American Economic Review*, 104(4), 1211-1243.

[21] Diamond, Rebecca (2013) "The Determinants and Welfare Implications of US Workers? Diverging Location Choices by Skill: 1980-2000," Stanford Graduate School of Business, mimeograph.

[22] Dingel, Jonathan (2015) "The Determinants of Quality Specialization," Chicago Booth School of Business, mimeograph.

[23] Duranton, Gilles and Matthew A. Turner (2011) "The Fundamental Law of Road Congestion: Evidence from U.S. Cities," *American Economic Review*, 101(6), 2616-2652.

[24] Duranton, Gilles, Peter Morrow and Matthew A. Turner (2014) "Roads and Trade: Evidence from the U.S.," *Review of Economic Studies*, 81(2), 681-724.

[25] Eaton, Jonathan and Kortum, Samuel (2002) "Technology, Geography, and Trade," *Econometrica*, 70(5), 1741-1779.

[26] Fujimoto, Takao and Krause, Ulrich (1985) "Strong Ergodicity for Strictly Increasing Nonlinear Operators," *Linear Algebra and its Applications*, 71(1), 101-112

[27] Fujita, Masahisa, Paul Krugman and Anthony J. Venables (1999) *The Spatial Economy: Cities, Regions and International Trade*, Cambridge: MIT Press.

[28] Greenstone, M., R. Hornbeck, and E. Moretti (2010) "Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plant Openings," *Journal of Political Economy*, 118(3), 536-598.

[29] Grogger, J., and G. Hanson (2011) "Income Maximization and the Selection and Sorting of International Migrants," *Journal of Development Economics*, 95(1), 42-57.

[30] Hanson, G. H. (1996) "Localization Economies, Vertical Organization, and Trade," *American Economic Review*, 86(5), 1266-1278.

[31] Hanson, Gordon H. (2005) "Market Potential, Increasing Returns, and Geographic Concentration," *Journal of International Economics*, 67(1), 1-24.

[32] Head, Keith and John Ries (2001) "Increasing Returns versus National Product Differentiation," *American Economic Review*, 91, 858-876.

[33] Helpman, Elhanan (1998) "The Size of Regions," In *Topics in Public Economics: Theoretical and Applied Analysis*, ed. David Pines, Efraim Sadka, and Itzhak Zilcha, Cambridge: Cambridge University Press.

[34] Kennan, J., and J. R. Walker (2011) "The Effect of Expected Income on Individual Migration Decisions," *Econometrica*, 79(1), 211-251.

[35] Krugman, Paul (1991) "Increasing Returns and Economic Geography," *Journal of Political Economy*, 99(3), 483-99.

[36] Lucas, Robert E. and Esteban Rossi-Hansberg (2002) "On the Internal Structure of Cities", *Econometrica*, 70(4), 1445-76.

[37] McFadden, D. (1974) "The Measurement of Urban Travel Demand," *Journal of Public Economics*, 3(4), 303-328.

[38] Melitz, Marc J. (2003) "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, 71(6), 1695-1725.

[39] Michaels, Guy (2011) "The Long Term Consequences of Resource-Based Specialisation," *Economic Journal*, 121(551), 31-57.,

[40] Mills, Edwin S. (1967) "An Aggregative Model of Resource Allocation in a Metropolitan Centre," *American Economic Review*, 57(2), 197-210.

[41] Monte, Ferdinando (2015) "The Local Incidence of Trade Shocks," Georgetown University, mimeograph.

[42] Moretti, Enrico (2011) "Local Labor Markets," in *Handbook of Labor Economics*, (ed.) D. Card, and O. Ashenfelter, Volume 4b, Amsterdam: Elsevier North Holland, 1238-1303.

[43] Moretti, Enrico and Patrick Klein (2011) "Local Economic Development, Agglomeration Economies, and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority," *Quarterly Journal of Economics*, 129(1), 275-331.

[44] Muth, Richard (1969) *Cities and Housing*, Chicago: University of Chicago Press.

[45] Notowidigdo, Matthew (2013) "The Incidence of Local Labor Demand Shocks," Northwestern University, mimeograph.

[46] Redding, Stephen J. and Daniel Sturm (2008) "The Costs of Remoteness: Evidence from German Division and Reunification," *American Economic Review*, 98(5), 1766-1797.

[47] Redding, Stephen J. (2012) "Goods Trade, Factor Mobility and Welfare," *NBER Working Paper*, 18008.

[48] Rossi-Hansberg, Esteban (2005) "A Spatial Theory of Trade," *American Economic Review*, 95(5), 1464-1491.

[49] Tolbert, Charles M. and Molly Sizer (1996) "U.S. Commuting Zones and Labor Market Areas: a 1990 Update," *U.S. Department of Agriculture Staff Paper*, AGES-9614.

[50] Yagan, Danny (2014) "Moving to Opportunity? Migratory Insurance over the Great Recession," University of California, Berkeley.