Regional Science and Urban Economics 71 (2018) 1–24

Contents lists available at ScienceDirect

Regional Science and Urban Economics

journal homepage: www.elsevier.com/locate/regec

Industry location and wages: The role of market size and accessibility in trading networks☆

Javier Barbero[a,*], Kristian Behrens[b,c,d], José L. Zofío[e]

[a] European Commission, Joint Research Centre (JRC), Directorate B – Growth and Innovation, Territorial Development Unit, Seville, Spain
[b] Department of Economics, Université du Québec à Montréal (ESG–UQAM), Canada
[c] National Research University Higher School of Economics, Russian Federation, Russia
[d] CEPR, UK
[e] Department of Economics, Universidad Autónoma de Madrid (UAM), Spain

ARTICLE INFO

JEL Classification:
R12
F12
C63

Keywords:
Market size
Accessibility
Trading networks
Industry location
Wages

ABSTRACT

We analyze the effects of local market size and accessibility on the spatial distribution of economic activity and wages in general equilibrium trade models with many asymmetric countries and costly trade for all goods. In models with a homogeneous sector, local market size is generally more strongly correlated with a country's industry share, whereas accessibility better explains a country's wage. We analytically show that result in a simplified case and then confirm it using simulations with random trading networks. In models with only differentiated sectors, both local market size and accessibility are highly correlated with wages. The impact of local market size on industry location is more robust than the impact of local market size on wages in economic geography models.

1. Introduction

Do market size and accessibility matter for industry location and wages? This question has attracted attention since Krugman's (1980) and Helpman and Krugman's (1985) seminal contributions to new trade theory. The answer is 'yes', at least in simple models: in a world with increasing returns and costly trade, market size and accessibility are locational advantages that influence the geographic distribution of industry and factor prices.[1] Despite its importance, it is fair to say that this result has been derived under a number of restrictive assumptions: (i) the existence of a costlessly tradable good; (ii) a single production factor; (iii) constant elasticity of substitution preferences; (iv) two industries only, with one producing a homogeneous good; and (v) two locations only. Taken together, those assumptions imply that little is known about the robustness of the result and on how it can eventually guide empirical analysis.

Conscious of these limitations, subsequent work has relaxed some of the initial assumptions on preferences, costless trade, and two locations. First, Ottaviano and Thisse (2004), Picard and Zeng (2005), Zeng and Kikuchi (2009), Baldwin et al. (2003), and Head et al. (2002), among others, have shown that the basic insights of 'home market effects' (henceforth, HME) generalize to other preference structures—such as

☆ We thank the editor, Laurent Gobillon, and two anonymous referees for very useful comments. We further thank Jacques Thisse, Gianmarco Ottaviano, Diego Puga, Ted Rosenbaum, Rafael Moner-Colonques, and Inmaculada C. Álvarez, as well as conference participants at the 2014 NARSC Meetings in Atlanta, the 3rd Workshop on Urban Economics in Barcelona, and the 10th Meetings on Economic Integration in Castellón, for useful comments and suggestions. Barbero acknowledges financial support from the Spanish Ministry of Education (AP2010-1401). Behrens acknowledges financial support from the CRC Program of the Social Sciences and Humanities Research Council (SSHRC) of Canada for the funding of the *Canada Research Chair in Regional Impacts of Globalization*. Barbero and Zofío acknowledge financial support from the Spanish Ministry of Science and Innovation (ECO2010-21643, ECO2013-46980-P, and ECO2016-79650-P). This study was funded by the Russian Academic Excellence Project '5-100'. Part of the paper was written while Barbero was visiting UQÀM, the hospitality of which is gratefully acknowledged. The views expressed in this paper and any remaining errors are ours, and may not in any circumstances be regarded as stating an official position of the European Commission.

* Corresponding author.
E-mail addresses: Javier.BARBERO-JIMENEZ@ec.europa.eu (J. Barbero), behrens.kristian@uqam.ca (K. Behrens), jose.zofio@uam.es (J.L. Zofío).

[1] The so-called 'home market effect' quickly became a key building block of New Trade Theory and then of New Economic Geography (see Krugman, 1980; Ottaviano et al., 2002). It also attracted much attention in the empirical trade literature (e.g., Davis and Weinstein, 2003; Hanson and Xiang, 2004), because it allowed to investigate the role of market size in shaping industry structure and trade.

https://doi.org/10.1016/j.regsciurbeco.2018.04.005
Received 3 February 2016; Received in revised form 21 March 2018; Accepted 5 April 2018
Available online 22 April 2018

quadratic-linear preferences—or other market structures—such as oligopolistic competition. Yu (2005) generalizes Davis (1998) by changing the upper-tier utility function from a Cobb-Douglas to the more general CES formulation. He shows that—when expenditure shares are non-constant—a HME or a reversed HME can arise depending on the elasticity of substitution between sectors. Second, Davis (1998) and Takatsuka and Zeng (2012a) have shown that the effect of market size on industry location is strongly dampened or even disappears when the homogeneous good is not costlessly tradable. Davis (1998), in particular, shows that when trading the homogeneous good is as costly as trading the differentiated good, market size has no longer any bearing on country specialization. This is also one basic message of Hanson and Xiang (2004), who argue that—in the absence of a costlessly tradable good—not all increasing returns sectors can display a HME.[2]

While all of the foregoing contributions shed some light on the role of market size and accessibility for industry location and wages, what is missing to date is more systematic evidence for what happens in more 'realistic settings' where several of the basic assumptions are relaxed simultaneously. To the best of our knowledge, there has been no systematic investigation when there are multiple locations, several industries, and costly trade for all goods. This paper addresses precisely these issues.

As a first step, we set up *two trade models* with two sectors, costly trade, and an arbitrary number of countries. We first develop a model with one homogeneous sector and one differentiated sector, both subject to trade costs. This is sufficient to generate very different equilibrium relationships between local market size and wages: depending on the countries' specialization patterns (complete specialization, incomplete specialization, and complete diversification), wages can decrease, be independent of, or increase with local market size. To shed additional light on our analytical results, and to better isolate the effect of accessibility, we solve the model for a limited number of countries and alternative configurations that systematically change network centrality and market size between the extreme ring and star network topologies. Doing so, we confirm the result that changes in market size matter more for changes in industry location, whereas changes in accessibility matter more for changes in wages.

We then develop a model with two differentiated industries subject to increasing returns and costly trade. In that model, local market size and accessibility are associated with higher wages, whereas differences in spending patterns are associated with industry location. In the general case, a mix of the two prevails and the relative effect of market size on the two equilibrium variables—industry location and wages—though positive, depends on the whole structure of the trading network.

As there is little hope to obtain clear-cut analytical results in the general case with an arbitrary number of countries and geographic structures of the trading network, as a second step we resort to systematic numerical simulations. More precisely, we simulate the equilibria of the two models using a large number of randomly generated networks with a large number of countries. We then check how our pencil-and-paper results—and simulations for basic configurations—extend to these higher-dimensional cases and extract the essence of the 'comparative statics' using statistical analysis. Put differently, our research strategy is to combine theory, numerical, and statistical analysis to: (i) first prove some new results in simple models; (ii) then solve larger models by numerical analysis; (iii) then run a statistical analysis of the numerical results, very much like engineers or physicists do; and (iv) finally confront the models with real data in an application to European Union countries.

Our key findings can be summarized as follows. First, in accord with the theoretical results derived in lower-dimensional instances of the models and simulations on simple networks, the effect of local market size on equilibrium wages crucially hinges on the countries' specialization patterns in our numerical simulations. Second, in all models that we simulate, the equilibrium relationship between local market size and industry location is more robust than the relationship between local market size and wages. Although the results vary slightly depending on the type of trading network considered, they are fairly robust. Third, the correlation between equilibrium wages and equilibrium industry shares is rather low, thus suggesting that both variables operate largely independently. Last, when applied to European Union country-level data, we find that in both cases the models generally predict well the distribution of industries, yet predict less well wages. A formal test does not allow to reject the null hypothesis that the industry distribution predicted by the models is the same than that observed in the data. The test does, however, reject the predicted wage distributions, because the stylized models cannot replicate the observed dispersion in wages across countries. This again shows that the models do a good job at predicting industry location, but are less useful for predicting wages.

The remainder of the paper is organized as follows. Section 2 develops our two simple trade models. We derive a number of comparative static results using specific instances of those models and illustrate several key economic properties and specialization patterns for simple network configurations. These results serve to guide the numerical analysis in Section 3. There, we extend the models to a larger scale and analyze a set of numerical results obtained from simulating those two models for a large number of random networks, generated by using two alternative attachment algorithms for network growth. We then present, in Section 4, an application to the case of European Union country data. Finally, Section 5 concludes. Most technical details are relegated to a set of appendices.

2. Two models with costly trade

We develop two models within which we analyze the geographic distribution of economic activity and wages.[3] In both models, there are $M \geq 2$ countries subscripted by $i = 1, 2, \ldots, M$. Each country is endowed with L_i immobile workers-consumers. The total population in the economy is fixed at $L \equiv \sum_i L_i$. Labor is the only production factor, i.e., we abstract from comparative advantages across countries.

2.1. Model 1: one differentiated sector and one homogeneous sector

Our first model builds on Helpman and Krugman (1985) and its multi-location extensions by Behrens et al. (2007, 2009). There is one increasing returns to scale (IRS) sector that operates under monopolistic competition and produces a continuum of varieties of a horizontally differentiated good; and one constant returns to scale (CRS) sector that

[2] Takatsuka and Zeng (2012b) propose another model with capital mobility, trade cost, and two countries. They find that the HME always appears for transport costs in both the homogeneous and differentiated sector. Turning to multi-country extensions of those models, Behrens et al. (2007, 2009) derive results when there are more than two countries. They show that the topology of the trading network matters for several of the results, and that the impact of market size on industry location arises only when differences in factor costs and in accessibility to markets are controlled for. While empirically relevant, multi-location extensions of new trade models to arbitrary geographic structures have been quite rare in the literature until now (see, e.g., Bosker et al., 2010; Stelder, 2016). Finally, Takahashi et al. (2013) derive analytical results in the case without factor price equalization with two countries, while Zeng and Kikuchi (2009) and Zeng and Uchikawa (2014) provide results without factor price equalization for many countries. Behrens et al. (2009) use a 'hybrid' approach, where trading the homogeneous good is costless, but where exogenous Ricardian differences in labor productivity in the homogeneous sector across countries create exogenous wage differences.

[3] The two models are not nested. We could develop a three sector model that nests our two models, but there is little gain from doing so. Indeed, as we will see later, it is the presence or absence of the homogeneous sector that is important for the key results of the model, not the presence of several differentiated sectors.

operates under perfect competition and produces a homogeneous good. As is standard in the literature, the sectors producing the differentiated and homogeneous goods are also referred to as industrial (or modern) and agricultural (or traditional), respectively. In the differentiated sector, the combination of IRS, costless product differentiation, and the absence of scope economies yields a one-to-one equilibrium relationship between firms and varieties.

2.1.1. Preferences and demands

Preferences of a representative consumer in country j are given by:

$$U_j = H_j^{1-\mu} D_j^{\mu}, \tag{1}$$

where H_j stands for the consumption of the homogeneous good; where D_j is an aggregate of the varieties of the differentiated good; and where $0 < \mu < 1$ is the share of income spent on the differentiated good. We assume that D_j is given by a CES subutility function

$$D_j = \left[\sum_i \int_{\Omega_i} d_{ij}(\omega)^{(\sigma-1)/\sigma} \mathrm{d}\omega \right]^{\frac{\sigma}{\sigma-1}},$$

where $d_{ij}(\omega)$ is the individual consumption in country j of variety ω produced in country i; and where Ω_i is the set of varieties produced in i. The parameter $\sigma > 1$ measures the elasticity of substitution between any two varieties. Let p_j^H denote the price of the homogeneous good in country j and $p_{ij}(\omega)$ the price of variety ω produced in country i and consumed in country j. Let w_j denote the wage in country j. Maximizing (1) subject to the budget constraint $p_j^H H_j + \sum_i \int_{\Omega_i} p_{ij}(\omega) d_{ij}(\omega) \mathrm{d}\omega = w_j$ yields the following individual demands:

$$d_{ij}(\omega) = \frac{p_{ij}(\omega)^{-\sigma}}{\mathbb{P}_j^{1-\sigma}} \mu w_j \quad \text{and} \quad H_j = \frac{(1-\mu) w_j}{p_j^H}, \tag{2}$$

where $\mathbb{P}_j$ is the CES price index in country j, given by

$$\mathbb{P}_j = \left[\sum_i \int_{\Omega_i} p_{ij}(\omega)^{1-\sigma} \mathrm{d}\omega \right]^{\frac{1}{1-\sigma}}. \tag{3}$$

2.1.2. Differentiated good

We first explain the workings of the sector operating under increasing returns to scale. The technology is assumed to be identical across firms and countries, therefore implying that firms differ only by the variety they produce and the country they are located in. Since varieties enter preferences in a symmetric way, we henceforth suppress the variety index ω to alleviate notation. Production of any variety involves a fixed labor requirement, F, and a constant marginal labor requirement, c. Denote by x_{ij} the amount of a variety produced in i and shipped to j. The total labor requirement for producing output $x_i \equiv \sum_j x_{ij}$ is given by $l_i = F + c x_i$.

Trade in the differentiated good is costly. Following standard practice we assume that trade cost are of the *iceberg* form: $\tau_{ij} \geq 1$ units must be dispatched from country i in order for one unit to arrive in country j. We further assume that trade costs are symmetric, i.e., $\tau_{ij} = \tau_{ji}$.[4] Using the demands (2), each firm in i maximizes its profit

$$\pi_i = \sum_j (p_{ij} - c w_i \tau_{ij}) L_j \frac{p_{ij}^{-\sigma}}{\mathbb{P}_j^{1-\sigma}} \mu w_j - F w_i \tag{4}$$

with respect to the prices p_{ij}, taking the price indices $\mathbb{P}_j$ and the wages w_j as given. Because of CES preferences, profit-maximizing prices display constant markups and are given by

$$p_{ij} = \frac{\sigma}{\sigma - 1} c w_i \tau_{ij}. \tag{5}$$

In what follows, we denote by n_i the endogenously determined mass of firms located in country i, and by $N \equiv \sum_i n_i$ the total mass of firms in the economy. We also denote by $\lambda_i \equiv n_i / N$ the share of firms located in country i.

Because of iceberg trade costs, a firm in country i has to produce $x_{ij} \equiv L_j d_{ij} \tau_{ij}$ units to satisfy aggregate demand in country j. Free entry and exit imply that profits are non-positive in equilibrium which, using (4) and the pricing rule (5), yields the condition

$$x_i \equiv \sum_j L_j d_{ij} \tau_{ij} \leq \frac{F(\sigma - 1)}{c}. \tag{6}$$

When (6) holds with equality, the firm located in country i makes zero profits, whereas it makes losses should the inequality be strict. Note that we may have strict inequalities since countries can specialize in the traditional sector and have no industrial activity.

Let $\phi_{ij} \equiv \tau_{ij}^{1-\sigma} \in [0, 1]$ denote the 'freeness of trade' in the differentiated good between countries i and j. Inserting the demand (2) and the price index (3) into (6), multiplying both sides by p_{ij}, and using the prices (5), we get:

$$\sum_j \frac{w_i^{-\sigma} w_j \phi_{ij} L_j}{\sum_k w_k^{1-\sigma} \phi_{kj} n_k} \leq \frac{\sigma F}{\mu}. \tag{7}$$

Dividing both sides by the total population, L, letting $\theta_j \equiv L_j / L$, and choosing—without loss of generality—units for F such that $F \equiv \mu L / \sigma$, we can rewrite (7) as follows[5]:

$$\mathrm{RMP}_i \equiv \sum_j \frac{w_i^{-\sigma} w_j \phi_{ij} \theta_j}{\sum_k w_k^{1-\sigma} \phi_{kj} n_k} \leq 1, \tag{8}$$

where RMP_i stands for the *real market potential* of country i (Head and Mayer, 2004). The mass of workers employed in the differentiated industry of country i, when it has n_i firms, is

$$L_i^D \equiv n_i l_i = n_i (F + c x_i) = n_i \mu L, \tag{9}$$

where we have made use of our normalization of F.

Equation (8) is crucial for determining the equilibrium allocation of firms across countries. Roughly speaking, this condition subsumes how many firms n_i can be located in each country i while making zero profits conditional on the different wages w_i, market sizes θ_i, and trade costs ϕ_{ij} across countries. If $\mathrm{RMP}_i < 1$, firms cannot make positive profits in country i and $n_i = 0$ must hold. If $n_i > 0$, i.e., there are firms operating in country i, then $\mathrm{RMP}_i = 1$ must hold because of the free entry zero profit condition.

2.1.3. Homogeneous good

We next explain the workings of the perfectly competitive sector that operates under constant returns to scale. We again assume that technology is the same in all countries. Without loss of generality, we normalize the unit labor requirement to one. Perfect competition implies marginal cost pricing. Given L_i^D workers employed in the differentiated good industry, the number of workers employed in the homogeneous sector equals $L_i^H \equiv L_i - L_i^D$. Plugging (9) into that expression, we can rewrite the number of workers in the homogeneous sector as

$$L_i^H = L_i - n_i \mu L. \tag{10}$$

[4] This assumption is not crucial but relatively standard. We relax it later in Section 4 when applying our model to European Union countries.

[5] This normalization is innocuous since we do not conduct any comparative static exercises with respect to those variables in what follows.

Note that (10) need not be strictly positive, i.e., some countries may specialize in the production of the differentiated good only.

We assume that trading the homogeneous good is costly.[6] Hence, factor price equalization (FPE) need not hold and the world mass of firms in the differentiated industry need no longer be constant.[7] The price of the homogeneous good produced in i and delivered to j equals its marginal cost of production, the wage w_i, times the trade cost τ_{ij}^H between countries i and j: $p_{ij}^H = w_i\tau_{ij}^H \equiv w_i\xi\tau_{ij}$, where $\xi > 0$ is a parameter that captures the *relative cost of trading the homogeneous good compared to the differentiated good*. If $\xi = 1$, there are no cost differences. When $\xi > 1$, trading the homogeneous good is more costly than trading the differentiated good, and vice versa when $\xi < 1$. In what follows, we set $\xi < 1$ because in the opposite case there is no trade in the homogeneous good so that the only equilibrium is one where industry shares are proportional to the size of the local market (see Davis, 1998).[8]

Because good H is homogeneous and can be produced in, and imported from, any country, its price in country i equals the lowest one that can be secured from any source:

$$p_i^H = \min_k\{w_k\xi\tau_{ki}\} \tag{11}$$

Let X_{ji} denote the imports of the homogeneous good from country j. Demand for the homogeneous good is given by (2), while supply is determined by the domestic production for the local market, X_{ii}, and the sum of imports X_{ji} from all sources. Market clearing for the homogeneous good in country i hence requires that:

$$\frac{(1-\mu)w_iL_i}{p_i^H} = X_{ii} + \sum_{j\neq i} X_{ji}. \tag{12}$$

Dividing the foregoing expression by the total population, L, and using the price (11), we can write (12) in terms of population shares, production, and per capita imports:

$$\frac{(1-\mu)w_i\theta_i}{\min_k\{w_k\xi\tau_{ki}\}} = \widetilde{X}_{ii} + \sum_{j\neq i}\widetilde{X}_{ji}, \tag{13}$$

where $\widetilde{X}_{ii} \equiv X_{ii}/L$, and $\widetilde{X}_{ji} \equiv X_{ji}/L$ denote per capita variables. Labor market clearing in country i then requires that $L_i^H = L_i - n_i\mu L = \xi\left(\tau_{ii}X_{ii} + \sum_{j\neq i}\tau_{ij}X_{ij}\right)$. Since $L_i = \theta_i L$, we can rewrite the foregoing condition in per capita terms as follows:

$$\theta_i - n_i\mu = \xi\left(\tau_{ii}\widetilde{X}_{ii} + \sum_{j\neq i}\tau_{ij}\widetilde{X}_{ij}\right) \tag{14}$$

Because of perfect competition, the homogeneous good will not be simultaneously imported and exported by the same country. Hence, it must be that

$$\widetilde{X}_{ij} = \begin{cases} > 0 & \text{if } w_i\tau_{ij} \le \min_k\{w_k\tau_{kj}\}, \\ = 0 & \text{otherwise.}\end{cases}$$

This latter condition can be expressed equivalently using complementary slackness as follows:

$$\widetilde{X}_{ij}\cdot\left[w_i\tau_{ij} - \min_k\{w_k\tau_{kj}\}\right] = 0 \quad \text{and} \quad \widetilde{X}_{ij} \ge 0, \quad \forall j = 1,2,\ldots,M. \tag{15}$$

2.1.4. Equilibrium

An equilibrium is such that the real market potential (8) is equal to one in all countries with a positive mass of IRS firms, and less than one for countries devoid of such firms. If all countries have a positive mass of IRS firms, we have an *interior equilibrium*, whereas if there are some countries without differentiated firms we get a *corner equilibrium*. Following Behrens et al. (2007, 2009), and as previously explained, an equilibrium is formally given by:

$$\begin{aligned} &\mathrm{RMP}_i = 1 \quad \text{if} \quad n_i^* > 0. \\ &\mathrm{RMP}_i \le 1 \quad \text{if} \quad n_i^* = 0. \end{aligned} \tag{16}$$

Using complementary slackness notation, this implies that $n_i^*\cdot(\mathrm{RMP}_i - 1) = 0$ and $n_i^* \ge 0$ for all countries. In addition to the zero profit free entry condition (16), the market clearing conditions (14) for the homogeneous good must hold for all countries at the equilibrium wages w_i. Expressions (13), (14), and (16), with M conditions each, and (15), with $M(M-1)$ conditions, yield a system of 3 M + $M(M-1)$ equations in that many unknowns—the M firm masses n_i, the M wages w_i, the M per capita domestic supplies $\widetilde{X}_{ii}$, and the $M(M-1)$ per capita imports $\widetilde{X}_{ij}$.

2.2. Model 2: two differentiated sectors

Our second model builds on Krugman (1980) and Behrens and Ottaviano (2011). There are two IRS sectors with CES monopolistic competition.[9] Countries' market sizes differ both because of the numbers of consumers and because consumers have different spending patterns for the two goods. In such a setting, we can look at how differences in *absolute market sizes*—the population shares θ_i—and differences in *relative market sizes*—the expenditure shares μ_i—affect wages and the location patterns of industries.

2.2.1. Preferences and demands

The basic setup is the same as in Section 2.1, except that there are now two CES sectors and no homogeneous sector. Preferences of a representative consumer in country j are given by:

$$U_j = D_{1j}^{\mu_{1j}}D_{2j}^{\mu_{2j}}, \tag{17}$$

where D_{sj} is the CES consumption aggregate in sector s and country j; and $0 < \mu_{sj} < 1$ are the *country-specific* income shares for sector s. With two sectors, μ_{sj} is equal to μ_j in sector 1 and to $1-\mu_j$ in sector 2. Since expenditure shares are country specific, the relative consumption patterns differ across countries. Hence, market sizes differ due to spending patterns on top of differences in countries' population sizes.

The aggregator for consumption of the differentiated good, D_{sj}, is as follows:

$$D_{sj} = \left[\sum_i\int_{\Omega_{si}} d_{sij}(\omega)^{(\sigma-1)/\sigma}\mathrm{d}\omega\right]^{\frac{\sigma}{\sigma-1}},$$

[6] See Appendix A.1 for a discussion of the case with costless trade of the homogeneous good. There we also explain why we disregard that case in what follows.

[7] The total mass of firms, N, varies with the spatial structure of the economy when there is costly trade in the homogeneous good (see, e.g., Takatsuka and Zeng, 2012a, b). Hence, (8) cannot be generally expressed in the usual share notation λ_i with respect to firms, which explains the presence of n_k in that expression.

[8] There is, of course, still two-way trade in the differentiated good and the wages adjust to balance that trade. However, our focus is on industry structure and wages. The former cannot be meaningfully analyzed when we assume that $\xi \ge 1$, whereas the latter cannot be meaningfully analyzed if we assume that there is free trade in the homogeneous good (see Appendix A.1).

[9] Hanson and Xiang (2004) develop a model with a continuum of sectors, but their focus is on two countries only. In this section, we take a complementary approach: we focus on two sectors only, but consider a large number of countries to look at industry location and wages with a more complex geography.

where $d_{sij}(\omega)$ is the individual consumption in country j of sector-s variety ω produced in country i; and where Ω_{si} is the set of sector-s varieties produced in i. For simplicity, we assume that the elasticity of substitution between any two varieties, σ, is the same in both sectors.[10] Let $p_{sij}(\omega)$ denote the price of sector-s variety ω produced in i and consumed in j; and let w_j denote the wage in country j. Maximizing (17) subject to the budget constraint $\sum_i \left[\int_{\Omega_{1i}} p_{1ij}(\omega) d_{1ij}(\omega)\mathrm{d}\omega + \int_{\Omega_{2i}} p_{2ij}(\omega) d_{2ij}(\omega)\mathrm{d}\omega\right] = w_j$ yields the following individual demands:

$$d_{sij}(\omega) = \frac{p_{sij}(\omega)^{-\sigma}}{\mathbb{P}_{sj}^{1-\sigma}}\mu_{sj}w_j, \quad \text{where} \quad \mathbb{P}_{sj} = \left[\sum_i \int_{\Omega_{si}} p_{sij}(\omega)^{1-\sigma}\mathrm{d}\omega\right]^{\frac{1}{1-\sigma}} \tag{18}$$

is the CES price index in sector s and country j.

2.2.2. Technology and trade

For simplicity, we assume that technology and transport costs are the same in both sectors. As in Section 2.1, the total labor requirement for producing the output $x_{si} \equiv \sum_j x_{sij}$ is given by $l_{si} = F + cx_{si}$. Trade in both differentiated goods is costly and trade cost are symmetric and of the iceberg form: $\tau_{ij} = \tau_{ji} \geq 1$ units must be dispatched from country i in order for one unit of a variety of any sector to arrive in country j. Using (18), a sector-s firm in i maximizes profit

$$\pi_{si} = \sum_j \left(p_{sij} - cw_i\tau_{ij}\right) L_j \frac{p_{sij}^{-\sigma}}{\mathbb{P}_{sj}^{1-\sigma}}\mu_{sj}w_j - Fw_i, \tag{19}$$

with respect to all its prices p_{sij}, taking the price indices $\mathbb{P}_{sj}$ and the wages w_j as given. As before, profit-maximizing prices display constant markups:

$$p_{sij} = \frac{\sigma}{\sigma-1} cw_i\tau_{ij}. \tag{20}$$

We denote by n_{si} the endogenously determined mass of sector-s firms located in i, and by $N_s \equiv \sum_i n_{si}$ the total mass of sector-s firms in the economy. Last, $\lambda_{si} \equiv n_{si}/N_s$ denotes the share of sector-s firms in country i.

A firm in country i and sector s has to produce $x_{sij} \equiv L_j d_{sij}\tau_{ij}$ units to satisfy aggregate demand in country j. Free entry and exit imply that profits are non-positive in equilibrium which, using the prices (20), yields again the standard free entry zero profit condition (6). Inserting the demands and the price index (18) into that expression, using the prices (20), and letting $\phi_{ij} \equiv \tau_{ij}^{1-\sigma} \in [0,1]$ denote the 'freeness of trade' between countries i and j, we get:

$$\sum_j \frac{w_i^{-\sigma} w_j \phi_{ij} L_j \mu_{sj}}{\sum_k w_k^{1-\sigma}\phi_{kj} n_{sk}} \leq \sigma F. \tag{21}$$

Dividing both sides by world population, L, letting $\theta_j \equiv L_j/L$ as before, and choosing without loss of generality units of F such that $F = L/\sigma$, we obtain the real market potential for sector-s firms in country i as follows:

$$\mathrm{RMP}_{si} \equiv \sum_j \frac{w_i^{-\sigma} w_j \phi_{ij} \theta_j \mu_{sj}}{\sum_k w_k^{1-\sigma}\phi_{kj} n_{sk}} \leq 1. \tag{22}$$

As before, condition (22) subsumes how many firms n_{si} in sector s can be located in each country i while making zero profits conditional on the different wages w_i, market sizes θ_i, spending patterns μ_{si}, and the freeness of trade ϕ_{ij} across countries.

[10] We could relax that assumption, but there is not much to be learned from that exercise. The same holds true for relaxing the assumption of identical technologies in the two sectors. Nevertheless, as explained in footnote [26] below, we have also studied the effects of alternative values of σ and expenditure patterns μ_{sj} on industry shares, λ_{si}, and wages, w_i.

2.2.3. Equilibrium

Expressions (22) define 2 M conditions in the 3 M unknowns $\{n_{1i}, n_{2i}, w_i\}$, for $i = 1, 2, \ldots, M$. To pin down the wages, we can impose either the labor market clearing conditions or the trade balance conditions. In what follows, we use the former as they are easier to handle given our choices of normalization. Labor market clearing in i requires that $L_i = n_{1i}(F + cx_{1i}) + n_{2i}(F + cx_{2i}) = L(n_{1i} + n_{2i})$, where we have used the normalization of F. Hence,

$$\theta_i = n_{1i} + n_{2i}. \tag{23}$$

Conditions (22) and (23) can be solved for the equilibrium wages and industry shares. The total masses of firms in the two sectors in the economy, $N_1 = \sum_i n_{1i}$ and $N_2 = \sum_i n_{2i}$ are not constant and vary with the spatial distribution of demand and with the structure of the trading network. Note, of course, that the total mass of firms in both sectors in the world economy is equal to one: $\sum_i (n_{1i} + n_{2i}) = \sum_i \theta_i = 1$ from (23).

2.3. Analytical results for simple networks

Although our primary objective is to simulate the two models using more complex spatial structures involving many countries, we first derive a number of results using simplified versions of those models. Doing so will provide guidance for the interpretation of the numerical results that we derive later. We proceed in two steps. First, we establish several analytical results on the relationship between market size, wages, and industry location using two or three countries only. This setup facilitates the exposition and the algebra while allowing us to understand a number of key properties. Second, we present some numerical simulations using simple networks that we vary smoothly between two extreme configurations, namely the ring (circle) and star networks. Doing so allows us to control the network structure to distill insights into the importance of accessibility and provides intuition about the impact of the shape of trade networks on economic geography.

2.3.1. Analytical results

Model 1: The importance of the homogeneous good. We first show that the comparative statics of industry shares and wages with respect to the size of the local market depend on the specialization and trade patterns. Let us start with three countries. For simplicity, we assume that countries 2 and 3 have the same size, $\theta_2 = \theta_3 \equiv (1-\theta)/2$, whereas the size of country 1 is $\theta_1 \equiv \theta$. Consider a pattern involving *complete specialization*, i.e., country 1 is the 'manufacturing core' for the differentiated good whereas countries 2 and 3 are 'agricultural peripheries' specialized in the production of the traditional good. Using share notation, we thus have $n_1 = \theta/\mu$ and $n_2 = n_3 = 0$. Hence, θ parametrizes the size of the core compared to the size of the (symmetric) peripheries, and we naturally have $\partial n_1/\partial\theta > 0$. Let the wage in country 2 be chosen as the numeraire, i.e., $w_2 \equiv 1$. As shown in Appendix A.2, the wage in country 1 can then be expressed as follows:

$$w_1 = \frac{\mu}{1-\mu}\frac{(1-\theta)}{2\theta}\left(1 + \frac{\tau_{21}}{\tau_{31}}\right) \tag{24}$$

The foregoing expression reveals two important properties. First, $\partial w_1/\partial\theta < 0$, i.e., the wage in country 1 is decreasing with the local market size. The intuition for this result is a classical terms-of-trade effect: as country 1 becomes larger—and its trading partners become smaller—its relative wage falls (see equation (13), recalling that $\widetilde{X}_{ii} = 0$ and that μ is fixed). The reason is that the shift in sizes must reduce demand for the traditional good in country 1 and increase it in countries 2 and 3. Hence, with costly trade in the homogeneous good and with complete specialization, we no longer necessarily have a positive relationship between local market size and wages. Second, assume that country 2 becomes more remote from country 1 (i.e., τ_{21} increases). As can be seen from (A-3) in the appendix, $\partial w_1/\partial\tau_{21} > 0$. In other words, more central regions—specialized in the modern good—have higher wages in that configuration.[11]

Consider next the case of *incomplete specialization* with two countries. In a first regime (type 1), country 1 produces both the differentiated and the homogeneous good, and country 2 only the homogeneous good. In that case, it must be that $w_1 = w_2\xi\tau = \xi\tau > 1$ for the traditional good to be produced in country 1 and simultaneously imported from country 2 to 1, where $\tau \equiv \tau_{12}$ denotes the symmetric trade cost between the two countries. This directly implies that $\partial w_1/\partial\theta_1 = 0$, i.e., the wage in country 1 is independent of the size of the local market. The reason is that costly trade in the homogeneous good imposes strong restrictions on relative wages, and those restrictions can destroy the positive link between market size and equilibrium wages. Over the range of incomplete specialization, the 'law of one price' breaks the link between local market size and wages. As shown in Appendix A.3, the equilibrium mass of firms in this equilibrium with incomplete specialization is given by

$$n_1 = \theta + \frac{1-\theta}{\xi\tau}, \tag{25}$$

which reveals that $\partial n_1/\partial\theta > 0$. Hence, although wages are independent of market size, the latter is reflected in industry structure.

A second regime of incomplete specialization (type 2) arises when the modern sector is active in both countries, $n_1 > 0$ and $n_2 > 0$, as well as the traditional sector, $L_1^H > 0$ and $L_2^H > 0$; whereas country 1 imports some of the homogeneous good. In that case, we still have $w_1 = \xi\tau$ so that $\partial w_1/\partial\theta_1 = 0$, i.e., the wage in country 1 is independent of the size of the local market for the same reason as before. Furthermore, as shown in Appendix A.3, we again have $\partial n_1/\partial\theta > 0$.

Finally, there is the case of *complete diversification*, i.e., when both the modern and the traditional sectors are active in both countries ($n_1 > 0$, $n_2 > 0$, and $L_1^H > 0$, $L_2^H > 0$). Under complete diversification, each country serves its own demand for the traditional good locally: $X_{21} = X_{12} = 0$. This directly implies that $X_{11} = (1-\mu)L_1$ and $X_{22} = (1-\mu)L_2$. Since $L_1^D = L_1 - L_1^H = \mu L n_1$, we have

$$n_1 = \frac{L_1}{L} = \theta \quad \text{and} \quad n_2 = \frac{L_2}{L} = 1-\theta, \tag{26}$$

i.e., industry location is proportional to market size, which implies that $\partial n_1/\partial\theta > 0$. As there is no trade in the traditional good, this configuration also requires that $1/w_1 < \xi\tau < w_1$, and the wage w_1 adjusts so that firms in the two countries make zero profits. As shown in Appendix A.4, in an equilibrium with complete diversification, there is a positive relationship between market size and wages: $\partial w_1/\partial\theta > 0$. The reason is that a larger local market provides a locational advantage for the increasing returns sector, and in order to guarantee that this sector operates in both countries the wage in country 1 must increase to offset the advantage of a larger local market size. Since there is no trade in the traditional good, there are no strong constraints on how wages can change.

To summarize, the key message from the foregoing developments is that *there is no clear relationship between wages and local market size in the model with costly trade in the homogeneous good.* Depending on the trade and specialization patterns, this relationship can be positive (complete diversification), zero (incomplete specialization), or even negative (complete specialization). As should be clear—and as we will show in the simulations—with multiple countries we will have different configurations for different sets of countries. Some countries will be completely diversified, some will be completely specialized, and some will be somewhere in between. Hence, we expect that the results on the link between local market size and wages will be fuzzy. However, we should see a clearer relationship between local market size and the share of industry, as shown by the foregoing developments.

Model 2: The importance of absolute and relative size. We next summarize some analytical results for the model with two differentiated sectors. To solve the model, we let $w_1 \equiv 1$ by choice of numeraire. Focusing on two countries with symmetric trade costs and free intra-country trade ($\phi_{ii} = 1$ and $\phi_{ij} = \phi$ for all $i \neq j$), Behrens and Ottaviano (2011) have investigated two opposite special cases: *absolute advantage*, i.e., when the spending patterns of the two countries are the same but they differ by population size ($\mu_{11} = \mu_{12}$ and $\mu_{21} = \mu_{22}$, but $\theta_1 > \theta_2$); and *comparative advantage*, i.e., when spending patterns are anti-symmetric but countries have the same population size ($\mu_{11} = \mu_{22}$ and $\mu_{21} = \mu_{12}$, but $\theta_1 = \theta_2$). General results with $N > 2$ countries are hard to come by and they are not required for the subsequent analysis.

Starting with pure 'comparative advantage', assume that preferences are anti-symmetric across countries ($\mu_{11} = \mu_{22}$ and $\mu_{21} = \mu_{12}$), and that both countries are of the same size ($\theta_1 = \theta_2$). As shown by Behrens and Ottaviano (2011), the equilibrium is such that

$$n_{11}^* = n_{22}^* = \frac{\mu(1+\phi)-\phi}{2(1-\phi)} \quad \text{and} \quad n_{21}^* = n_{12}^* = \frac{1-\mu(1+\phi)}{2(1-\phi)}, \tag{27}$$

and the equilibrium relative wage satisfies $w_2^* = 1$. In this case, each country is the larger market for one of the two goods. Hence, each country specializes in the production of the good for which it has a relatively larger local demand. In other words, relative differences in market sizes lead to different specialization patterns but do not affect factor prices.

Consider next the polar case of pure 'absolute advantage'. Assume that preferences are symmetric across countries ($\mu_{11} = \mu_{12}$ and $\mu_{21} = \mu_{22}$), and that country 1 has the larger market ($\theta_1 > \theta_2$). The equilibrium is then such that

$$n_{1i}^* = \mu\theta_i \quad \text{and} \quad n_{2i}^* = (1-\mu)\theta_i, \tag{28}$$

for $i = 1, 2$, whereas the equilibrium relative wage satisfies $0 < w_2^* < 1$. In this case, one country is the larger market for both goods. Hence, the wage in the larger country must be higher because it offers a locational advantage for *both* industries. Clearly, this is akin to absolute advantage in a Ricardian sense and it is, therefore, capitalized into factor prices.

To summarize, both industry location and wages are positively related to local market size in the model with two differentiated industries, but the exact extent depends on the relative importance of absolute and of comparative advantage. The two cases discussed above are 'pure' ones to illustrate the key findings, but intermediate cases where both absolute and comparative advantage play a role should be considered. Unfortunately, clear results on the impacts of accessibility are not easy to derive in this model, even with a small number of countries. Hence, it will be of interest to relax the assumption of just two countries and of symmetric trade costs to investigate also the interactions with 'geography' using numerical methods. This is what we do using simulations in the next sections and European Union data in Section 4.

2.3.2. Numerical results for simple networks

While the foregoing developments allow us to understand the existence or absence of a link between market size and wages, they provide less information on the role played by the structure of the trading network. To understand the latter, we now provide results using 'controlled' networks. Following Barbero and Zofío (2016), we solve the first model for a large number of networks comprised between two extreme topologies: the ring and the star configurations. The ring topology characterizes a homogeneous space where all countries lie on a circle—hence the nickname 'racetrack' economy—so that no country enjoys a locational advantage. On the contrary, the star topology—also

[11] If countries 2 and 3 are located symmetrically with respect to country 1 (i.e., $\tau_{21} = \tau_{31}$), the model reduces to the two-country case. Indeed, it is well known that in this type of model symmetric configurations with $N > 2$ countries can be reduced to two-country cases with different market sizes (see, e.g., Behrens et al., 2007, 2009). This shows that asymmetric trade cost structures are fundamental to the investigation of the multi-country models.

known as 'hub-and-spoke'—represents the most extreme heterogeneous space where the country situated in the center enjoys the most privileged central position. We generate a large number of intermediate networks and study how the equilibrium industry shares and wages behave depending on a measure of network centrality (closeness) and market size. To keep the dimensionality at a manageable level, we set $M = 4$.[12]

The freeness of trade matrices corresponding to the ring and the star networks, with the second country being the center of the star, are the following:

$$\phi_{ring} = \begin{bmatrix} 1.00 & 0.20 & 0.04 & 0.20 \\ 0.20 & 1.00 & 0.20 & 0.04 \\ 0.04 & 0.20 & 1.00 & 0.20 \\ 0.20 & 0.04 & 0.20 & 1.00 \end{bmatrix} \quad \text{and} \quad \phi_{star} = \begin{bmatrix} 1.00 & 0.20 & 0.04 & 0.04 \\ 0.20 & 1.00 & 0.20 & 0.20 \\ 0.04 & 0.20 & 1.00 & 0.04 \\ 0.04 & 0.20 & 0.04 & 1.00 \end{bmatrix} \tag{29}$$

On the one hand, at the level of an individual node, the closeness measure reflects how central a country is in a given network, i.e., it can be interpreted as a measure of the country's locational advantage. The closeness centrality of country i is defined as

$$c_i = \left[\frac{\sum_j d_{ij}}{\min_k \left\{ \sum_j d_{kj} \right\}} \right]^{-1}, \tag{30}$$

where d_{ij} denotes the length of the link—the distance—between countries i and j. By definition, closeness varies between 0 and 1. On the other hand, at the level of the network and following Freeman (1978), a measure of network centrality is computed as the sum of the centrality differences between the location with the highest centrality and all remaining locations, divided by the maximum sum of the differences that can exist in a network with the same number of countries. This measure of network centrality ranges from 0, when no country has a locational advantage, to 1, when there is only one country—the central one—with a locational advantage. It is given by:

$$C(h) = \frac{\sum_{i=1}^{N} \left(c_{i^*}^h - c_i^h \right)}{\max \left[\sum_{i=1}^{N} \left(c_{i^*}^h - c_i^h \right) \right]} = \frac{\sum_{i=1}^{N} \left(c_{i^*}^h - c_i^h \right)}{\frac{(N-1)(N-2)}{(2N-3)}}, \tag{31}$$

where h corresponds to the network being measured and $c_{i^*}^h = 1$ corresponds to the location(s) with maximum accessibility in the network. In our controlled networks, the ring topology has a network centrality of 0, whereas the network centrality is 1 for the star.

We generate 100 configurations with increasing centralities between the ring and the star. The difference in the freeness of trade matrix between each network is computed as:

$$\phi_{diff} = \frac{\phi_{ring} - \phi_{star}}{100 - 1}. \tag{32}$$

Consequently, the freeness of trade matrix for an intermediate network h is given by:

$$\phi_h = \phi_{star} + (h-1)\phi_{diff}, \quad \forall h = 1, 2, \ldots, 100. \tag{33}$$

For each network h, we solve the first trade model for different population shares taking the most central node as reference benchmark—i.e., the second node in the star configuration—with θ_2 ranging from 0.1 to 0.9, and setting $\theta_i = (1 - \theta_2)/3$ for the remaining three countries. Taking successive increments in the population share of the central region of 0.05, and given the 100 networks generated with different topologies, we evaluate a total of 1, 600 networks.

The results for the spatial equilibria are shown in Figs. 1 and 2. The surface plots relate the equilibrium industry shares (λ_2^*) or wages (w_2^*) of the differentiated sector in the central node (on the z-axis) to the network centrality, $C(h)$ (on the y-axis) and to the population shares θ_2 (on the x-axis). For the equilibrium shares, although a nonlinear relation is observed in Fig. 1, it corresponds to a concave function that displays a monotonic and positive relationship with both population and node centrality. As shown, market size θ_2 is the most influential variable: for a given population, increasing the centrality of the central country marginally increases its share of the differentiated sector, whereas keeping accessibility constant while increasing its population share leads to a steep increase in industry shares.

We thus confirm the existence of a positive and strong (even if non-linear) relationship between market size and equilibrium industry shares. Moreover, examining the trade and specialization patterns of the four countries, the outcomes of our simulations for the central region reveal that type-2 incomplete specialization characterizes the world economy for $\theta_2 > 0$, as both the central and peripheral countries produce both the manufacturing and agricultural goods, with the latter being traded among them (thereby excluding the case of complete diversification). However, when the population share in the most accessible country 2 increases along with network centrality, the structure of the world economy evolves to the first type of incomplete specialization. Indeed, for $\theta_2 > 0.65$ and $C(h) > 0.5$, manufacturing tends to agglomerate in the central country $\lambda_2 \approx 1$, depleting the peripheral countries of activity in that sector. The latter now specialize in the homogeneous good and export it to the core.

Turning to the equilibrium wages, whose relationship with market size can take several forms as shown analytically in the foregoing section, our simulations confirm the two distinct alternatives for incomplete specialization of type 1 and 2. On the one hand, market size strongly drives up wages, with $w_2^* > 1$ except for the case where θ_2 is small. As population increases, and regardless of centrality, equilibrium wages w_2^* increase and reach the maximum value $w_2^* \approx 1.05$ for $\theta_2 > 0.35$, with $\partial w_2^* / \partial \theta_2 = 0$, as predicted by the theory. Comparing the gradient of λ_2^* and w_2^* between Figs. 1 and 2 with respect to market size θ_2, we see that differential accessibility plays a larger role for wages than for industry location. Hence, market size seems to matter more (and strongly) for

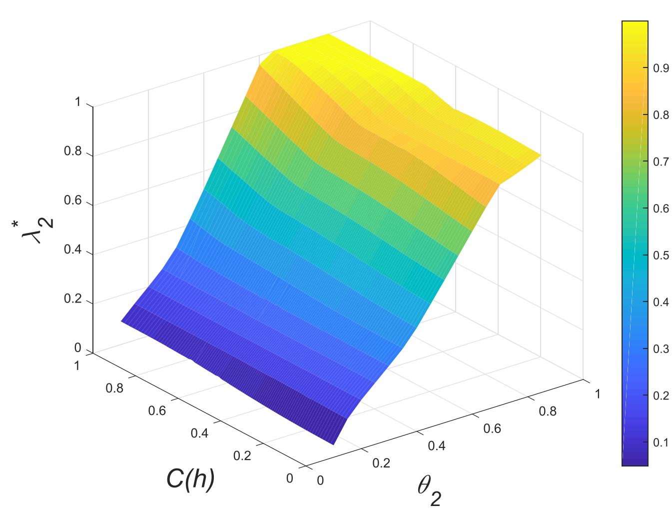

Fig. 1. Equilibrium shares of the most central region (λ_2^*) for alternative values of network centrality ($C(h)$) and population (θ_2).

[12] We could use a larger number of countries, but this makes little difference since the setup if still fairly 'symmetric' in our controlled approach.

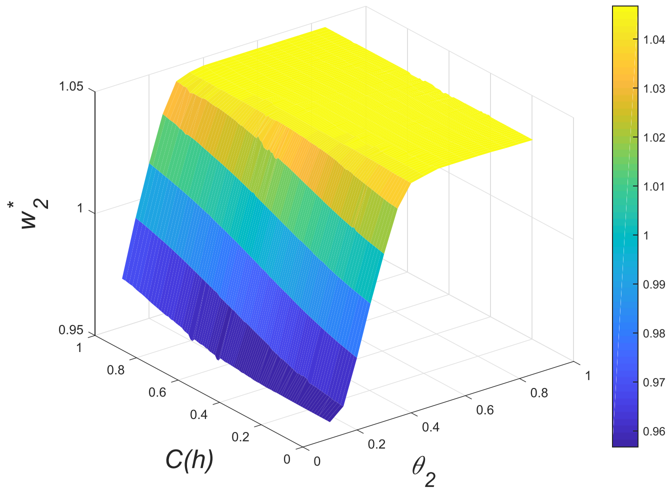

Fig. 2. Equilibrium wages (w_2^*) for alternative values of network centrality ($C(h)$) and market size (θ_2).

industry location, whereas accessibility matters more (but less strongly) for wages.

Although the foregoing results confirm our analytical findings for basic network topologies and alternative market size distributions in a systematic way, it remains to be seen whether they will hold for larger and more complex networks where all trade and economic specialization regimes can emerge and coexist.[13] They provide intuition for—and tell us what to expect from—the simulations, regression analyses, and numerical checks that we undertake in the following sections.

3. Size and accessibility in random tree networks

Despite our analytical results and the insights derived from 'controlled' networks, it is virtually impossible to derive general (analytical) results concerning the impact of market size on wages and industry location in an arbitrary multi-country setting. The reason is that the impact depends on the equilibrium patterns of trade and specialization, which are determined by a complex trade-off between a country's market size and its accessibility in the trading network. Although we have partially explored the role of accessibility and market size for industry location and wages in controlled networks with a small number of countries, in this section we consistently explore it in larger and less restricted networks that are generated following two alternative algorithms. This allows us to gain further insight into how accessibility and market size—as well as the whole structure of the trading network—influence the equilibrium.

We proceed as follows. First, we generate large random tree networks with a random number of nodes (see Appendix B.1 for details). Again, the nodes are the countries, and the links between nodes represent the connections for shipping goods. Networks are generated incrementally either by having equal attachment probabilities for new nodes, or by using the Barabási and Albert (1999; henceforth BA) preferential attachment algorithm that generates networks which exhibit a 'hub-and-spoke' structure. Second, we assign a random population share, θ_i, to each node i of the network.[14] In the case with two differentiated industries, we also randomly assign a country-specific expenditure share for each industry. Third, we solve the two models for their equilibria (see Appendix A for the equilibrium conditions). We repeat this three-step process for a large number of randomly generated networks and then relate selected characteristics of the equilibria thus obtained to underlying networks characteristics. Doing so allows us to gain insights into how size and accessibility interact to determine the country allocation of firms and wages. We describe the numerical implementation in detail in Appendix B.2. In the following sections, we explore the results obtained for the two models.

Before proceeding with the analysis, two important comments are in order. First, one may wonder why we look at local size and accessibility separately. Indeed, as shown in the literature, there is a theoretical link between a measure of 'market potential' and the location of industry and wages. Hence, we should use market potential as a theory-based determinant of the equilibrium allocation. Yet, as is well known, the market potential conflates size and accessibility (Head and Mayer, 2004), and thus does not allow to *separately* investigate the contribution of each to the equilibrium allocation. Since our objective is to disentangle the impacts of size and of accessibility on the equilibrium, we cannot simply use market potential in our subsequent analysis. Furthermore, as shown by Behrens et al. (2007), there is a theoretical link between industry shares and a measure of network centrality in some versions of this type of model. Hence, looking at the impact of centrality is of theoretical interest in its own right.

Second, to simplify matters we run non-linear regressions of equilibrium outcomes on the exogenous measures of accessibility and local size. We view these regressions as 'comparative static' exercises that allow us to approximate the non-linear relations characterizing the models. A natural option is to consider a flexible functional form that, being twice continuously differentiable, can approximate any function to the second order at an arbitrary point (constituting a specific Taylor approximation).[15] These regressions are a natural starting point in the absence of any knowledge about the non-linear equilibrium relationships and the way that accessibility and size can be theoretically separated. When there are non-linear structural relationships in the theoretical model and accessibility and size cannot be clearly separated, estimation errors capture those aspects but have no other structural interpretation.

Our set of results is complemented with the calculation of average marginal effects of accessibility and size on equilibrium wages and industry locations. As shown in Section 2.3, since the comparative statics depend on the equilibrium trade and specialization regimes, we provide results for both the aggregate model and for the different types of nodes depending on their patterns of specialization and trade. Finally, we depict the graphs of two representative networks—one with preferential attachment and one with equal probabilities—which allows us to highlight the role that accessibility and size play in shaping specialization and trade patterns.

[13] Note that the regime with complete specialization—where some countries specialize in manufacturing only and some in agriculture only—does not arise in the controlled networks. Hence, the regime where wages decrease in market size is absent.

[14] Choosing 'totally random' networks—though providing an interesting benchmark case—is not fully satisfying because transportation networks are endogenous and obey certain rules. This is why we also derive results using networks that display a 'hub-and-spoke' structure to capture the empirical fact that some places are very well connected while others are very poorly connected (see, e.g., Xie and Levinson, 2008, for the case of the road network in Indiana). Observe that we assign θ_i randomly, i.e., there is no systematic correlation between size and accessibility. The reason for that choice is that we want to study the distribution of industry as a function of size and accessibility separately. Introducing a systematic correlation between the two (though empirically relevant since larger places that are better connected tend to grow larger; see Duranton and Turner, 2012) is not required for our analysis.

[15] Diewert (1971) formalized this notion of flexibility. Among the alternative candidates allowing for a second-order approximation are the quadratic, the generalized Leontief, or the translog functional forms (see, e.g., Thompson, 1988). We run regressions for these different functional forms but only report the quadratic results based on goodness-of-fit criteria.

Table 1
Simple correlations for Model 1.

	λ_i^*	n_i^*	w_i^*	θ_i	closeness$_i$	degree$_i$
λ_i^*	1					
n_i^*	0.9987	1				
w_i^*	0.0849	0.0806	1			
θ_i	0.8119	0.8065	0.0899	1		
closeness$_i$	0.2680	0.2693	0.1316	0.0134	1	
degree$_i$	0.3972	0.4023	0.1799	0.0135	0.7075	1

Notes: We set $\sigma = 5$, $\mu = 0.4$, and $\xi = 0.7$. See Section 4.1 for more details on those choices. Simple correlations for 100 random tree networks with a random number of 20–30 nodes. The table gives correlations at the level of individual nodes (pooled across all 100 networks). The shares λ_i^* are given by $\lambda_i^* = n_i^*/(\sum_j n_j^*)$.

3.1. Model 1: numerical results

We first compute simple correlations between the equilibrium masses of firms in the different countries (n_i^*), their population shares (θ_i), and their centrality (c_i). The latter is measured either by the closeness centrality in expression (30)—henceforth ‘closeness’, for short—or by the node's degree—henceforth ‘degree’, for short. ‘Degree’ is simply measured by the number of links of the node. Centrally located countries have both a high value for closeness and for degree. This can be seen from the correlations in Table 1. As expected, size (θ_i) and accessibility (closeness$_i$ and degree$_i$) are on average positively linked to a country's equilibrium industry share (λ_i^* or, alternatively, n_i^*). They are also on average positively linked to a country's wage, w_i^*, although this link is much weaker (as suggested by our results in Section 2.3). Observe that size is (relatively) more strongly linked to industry location, whereas accessibility is (relatively) more strongly linked to wages. Put differently, size differences map more strongly into differences in industry structure, whereas accessibility differences translate more strongly into factor price differences. It is finally of interest to note that the correlations between the equilibrium industry shares, λ_i^* (or the equilibrium masses of firms, n_i^*) and the equilibrium wages—though positive—are fairly small (0.080 and 0.085, respectively). This suggests that wages and industry location shape the equilibrium outcome differently, depending crucially on the observed patterns of specialization and trade.

To go beyond simple univariate correlations, we now run several regressions to gauge the partial effect of increasing market size or centrality of nodes on the equilibrium shares of manufacturing activity and the equilibrium wages, controlling for accessibility and for size. In Model 1, there are two endogenous variables that can be analyzed in the regressions: the equilibrium allocation of firms, λ_i^*, and the equilibrium wages, w_i^*.[16] We regress these two equilibrium outcomes on measures of: (i) the node's centrality, as given by either closeness or degree; and (ii) the node's local market size.[17] We perform a pooled analysis with both types of networks (based on preferential attachment, BA, or equal probabilities)—in which case we include a network dummy indicating the network type—and separate regressions for each type of network. Formally, we estimate the following quadratic specifications

$$\begin{aligned}\lambda_i^* \;&=\; \beta_0 + \beta_1 \text{centrality}_i + \beta_2 \text{centrality}_i^2 + \beta_3\theta_i + \beta_4\theta_i^2 \\ &\quad + \beta_5(\text{centrality}_i \times \theta_i) + \text{network\_dummy}_i + \varepsilon_i\end{aligned} \tag{34}$$

$$\begin{aligned}w_i^* \;&=\; \gamma_0 + \gamma_1 \text{centrality}_i + \gamma_2 \text{centrality}_i^2 + \gamma_3\theta_i + \gamma_4\theta_i^2 \\ &\quad + \gamma_5(\text{centrality}_i \times \theta_i) + \text{network\_dummy}_i + \varepsilon_i\end{aligned} \tag{35}$$

for all the nodes of the networks that we have generated.

Table 2 summarizes our results for the estimations of (34) and (35). As can be seen from that table, both centrality and market size have nonlinear effects on a node's equilibrium share of firms and its equilibrium wage. Observe that the linear and the quadratic effects usually differ in sign, with the value of the latter markedly exceeding that of the former. Observe further that the cross effects are always positive for λ_i^*, whereas they are always insignificant for w_i^*. Hence, local size has a stronger effect on industry location for nodes with high accessibility, whereas such an effect does not arise for equilibrium wages. The average marginal effects reported in the upper panel of Table 3 confirm that the empirical definition of the HME—defined here as a more than proportional increase in industry shares in response to an increase in local market size—always arises in both types of networks: $\partial\lambda_i^*/\partial\theta_i > 1$. As explained before, the cross effect with accessibility reinforces this pattern. The HME thus generally seems to hold in models without FPE and a large number of locations (see also Zeng and Uchikawa, 2014). We have calculated equivalent average marginal effects by the alternative regimes of node specialization (see Table 4), and the HME always holds in its derivative formulation. It is particularly strong for the central—completely specialized—nodes, and a robust result.

We now study the theoretical definition of the HME at the level of the whole network. Following Behrens et al. (2009) and Zeng and Uchikawa (2014), a network exhibits this effect if the following sequence of inequalities holds once countries are ordered by decreasing size—θ_1 being the largest country:

$$\theta_1 > \theta_2 > \ldots > \theta_M \quad \Rightarrow \quad \frac{\lambda_1^*}{\theta_1} > \frac{\lambda_2^*}{\theta_2} > \ldots > \frac{\lambda_M^*}{\theta_M} \tag{36}$$

Although (36) can theoretically hold for networks of any size under very specific assumptions, it is expected that it is not verified in large multi-country settings like those corresponding to the random networks that we generate. Indeed, it does not hold in a single of our 100 networks. However, a way to test whether there is a relevant ordering that might hold statistically is to check if there exists some correlation between θ_i and λ_i^*/θ_i. Out of the 100 networks, 90 exhibit positive and statistically significant correlations, whose average coefficient is $\overline{\rho} = 0.543$. This shows that while (36) does not hold strictly, a positive relationship exists between market size and the relative share of production to demand in the differentiated good.[18] Fig. 3 graphs the relationship between these two variables, where the value of λ_i^*/θ_i (on the y-axis) is plotted against the value of θ_i (on the x-axis). A very distinct cluster of nodes with $\lambda_i^*/\theta_i > 2$ is clearly visible. These are the countries enjoying a privileged position in the world production and trading network and, from Figs. 4 and 5 below, we see that they correspond to nodes that completely specialize in the differentiated sector because of the their central location in the network.

We can further explore and check the robustness of these results by focusing on the values of λ_i^*/θ_i by node specialization, and selecting as reference threshold a more than proportional share of production to demand, i.e. $\lambda_i^*/\theta_i > 1$. Note that, although (36) generally will not hold, $\lambda_i^*/\theta_i > 1$ must hold for at least some nodes in the network, i.e., some nodes must have a disproportionate share of production. Fig. 4 shows these distributions. The results for the completely specialized nodes, presented in the left panels, are clear-cut. For this regime, 82.2% of the nodes exhibit a more than proportional share of production to demand. Furthermore, it is possible to identify a threshold value $\underline{\theta}_i \geq 0.001$ (0.1%), above which all nodes exhibit $\lambda_i^*/\theta_i > 1$. Both the high percentage of nodes displaying this characteristic and the existence of a threshold value are interesting results. As completely specialized nodes emerge in the most central locations (hubs) of the trading networks

[16] Due to the high correlation between λ_i^* and n_i^* (see Table 1), there is no reason to look at the latter separately.

[17] We do not include both measures of centrality simultaneously, because of their high correlation (see Table 1).

[18] These results are confirmed using Spearman's rank correlation between both sets of variables. In this case, 98 out of 100 networks exhibit a positive correlation with an average coefficient of 0.615. The remaining two correlations are not statistically significant.

Table 2
Regression results for Model 1.

	(i)	(ii)	(iii)	(iv)	(v)	(vi)
			Dependent variable: λ_i^*			
Closeness_i	-0.2505^a		-0.3653^a		-0.1795^a	
	(−12.677)		(−12.586)		(−7.769)	
Closeness_i^2	0.1895^a		0.2805^a		0.1260^a	
	(13.892)		(14.078)		(7.883)	
Degree_i		0.0022^a		0.0084^a		-0.0096^a
		(3.516)		(9.294)		(−9.882)
Degree_i^2		-0.0001^c		-0.0007^a		0.0015^a
		(−1.786)		(−8.349)		(10.732)
θ_i	0.2006^b	0.6493^a	-0.2756^b	0.6625^a	0.5413^a	0.6614^a
	(2.429)	(15.604)	(−2.188)	(10.122)	(5.844)	(14.107)
θ_i^2	1.9280^a	2.1178^a	2.4084^a	1.8730^a	1.1918^b	1.8009^a
	(3.625)	(5.015)	(2.866)	(2.684)	(2.079)	(4.067)
$\text{Closeness}_i \times \theta_i$	1.3266^a		2.0438^a		0.8932^a	
	(12.688)		(12.317)		(7.893)	
$\text{Degree}_i \times \theta_i$		0.1989^a		0.1934^a		0.2182^a
		(25.077)		(17.982)		(20.223)
Constant	0.0717^a	-0.0097^a	0.1080^a	-0.0167^a	0.0531^a	0.0035^b
	(9.963)	(−7.757)	(10.225)	(−9.340)	(6.336)	(2.313)
Network type	Both	Both	BA	BA	Equal	Equal
Network dummy	Yes	Yes	No	No	No	No
Observations	2, 498	2, 498	1, 274	1, 274	1, 224	1, 224
Adjusted R^2	0.760	0.849	0.754	0.831	0.830	0.899
	Dependent variable: w_i^*					
Closeness_i	-0.2118^a		-0.3139^a		−0.0491	
	(−4.018)		(−5.245)		(−0.557)	
Closeness_i^2	0.1934^a		0.2680^a		0.0857	
	(5.317)		(6.525)		(1.406)	
Degree_i		0.0038^c		0.0012		0.0029
		(1.838)		(0.544)		(0.613)
Degree_i^2		0.0001		0.0004^c		0.0004
		(0.665)		(1.799)		(0.538)
θ_i	−0.1219	-0.4275^a	−0.3716	-0.7246^a	0.4239	−0.1227
	(−0.553)	(−3.068)	(−1.431)	(−4.457)	(1.199)	(−0.531)
θ_i^2	6.4050^a	6.5170^a	6.5982^a	6.5911^a	5.2136^b	5.7051^a
	(4.515)	(4.609)	(3.808)	(3.803)	(2.384)	(2.616)
$\text{Closeness}_i \times \theta_i$	−0.3996		−0.5192		−0.6072	
	(−1.433)		(−1.517)		(−1.406)	
$\text{Degree}_i \times \theta_i$		0.0194		0.0105		0.0552
		(0.732)		(0.395)		(1.039)
Constant	1.0524^a	0.9926^a	1.0890^a	0.9979^a	0.9804^a	0.9800^a
	(54.786)	(237.347)	(49.996)	(224.726)	(30.645)	(129.737)
Network type	Both	Both	BA	BA	Equal	Equal
Network dummy	Yes	Yes	No	No	No	No
Observations	2, 498	2, 498	1, 274	1, 274	1, 224	1, 224
Adjusted R^2	0.052	0.058	0.063	0.065	0.066	0.074

Notes: We set $\sigma = 5$, $\mu = 0.4$, and $\xi = 0.7$. OLS regressions. BA denotes networks generated using the Barabási and Albert (1999) algorithm. *T*-stats in parentheses. [a], [b], and [c] denote coefficients significant at 1%, 5%, and 10%, respectively.

(see Fig. 5), these results imply that a more than proportional share of production to that of population is a distinctive feature of key nodes within a network regardless of their market size (i.e., those with the highest accessibility). This is not, however, a feature that is generally observed for the remaining specialization regimes and majority of nodes. Across the whole network—for the alternative regimes of incomplete specialization and complete diversification—the percentage of nodes exhibiting more than proportional values, $\lambda_i^*/\theta_i > 1$, falls substantially, and there is no market size threshold separating the spatial equilibria according to this value.[19,20]

Turning to Table 3, as predicted by theory, both measures of centrality—closeness and degree—have a significant positive association with the equilibrium allocation of firms across countries, as captured by their average marginal impact. The results pertaining to the equilibrium wages in the bottom panels of Tables 2 and 3 deserve special attention. First, as can be seen, the two measures of centrality are positively linked to a country's equilibrium wage in a non-linear way given their linear and quadratic coefficients, with the predominant effect, corresponding to the marginal effects, being positive. In other words, more centrally located countries with better market access command higher wages, which is in line with predictions of new economic geography models and with empirical evidence (see, e.g., Mion, 2004, for Italy; and Hanson, 2005, for the US).

Second, the correlation between w_i^* and θ_i is quite low—though still positive—as reported in Table 1. A larger local market is weakly associated with higher wages, except in hub-and-spoke type BA networks where the effect is on average negative (see columns (iii) and (iv) of Tables 2 and 3). This latter result is surprising and requires some further

[19] Other relevant results are the magnitudes of λ_i^*/θ_i, much larger in the completely specialized nodes, as well as the non-linear positive relationship between this ratio and market size, clearly visible in the lower panel for the nodes where $\lambda_i^*/\theta_i < 1$.

[20] The results are identical when using the mass of firms, n_i^*, instead of the share of firms, λ_i^*.

Table 3
Marginal effects for Model 1.

	(i)	(ii)	(iii)	(iv)	(v)	(vi)
	Average marginal effect on λ_i^*					
Closeness$_i$	0.0461^a		0.0697^a		0.0217^a	
	(15.56)		(15.12)		(6.77)	
Degree$_i$		0.0097^a		0.0135^a		0.0048^a
		(31.78)		(28.07)		(13.11)
θ_i	1.2074^a	1.2007^a	1.2038^a	1.1838^a	1.2228^a	1.2246^a
	(79.02)	(98.89)	(52.25)	(61.97)	(71.27)	(92.39)
Network type	Both	Both	BA	BA	Equal	Equal
	Average marginal effect on w_i^*					
Closeness$_i$	0.0208^a		0.0028		0.0391^a	
	(2.63)		(0.30)		(3.20)	
Degree$_i$		0.0051^a		0.0030^b		0.0065^a
		(4.96)		(2.55)		(3.64)
θ_i	0.1341^a	0.1316^a	−0.1701^a	−0.1767^a	0.4429^a	0.4401^a
	(3.29)	(3.24)	(−3.58)	(−3.72)	(6.77)	(6.74)
Network type	Both	Both	BA	BA	Equal	Equal

Notes: We set $\sigma = 5$, $\mu = 0.4$, and $\xi = 0.7$. *T*-statistics are given in parentheses. ^{a}, ^{b}, and ^{c} denote coefficients significant at 1%, 5%, and 10%, respectively.

Table 4
Number of occurrences of each specialization pattern for the different nodes.

Node type	# nodes	% nodes	$\overline{\theta}_i$	$\overline{w}_i^*$
Barabási and Albert				
Complete specialization	122	10%	0.0314	1.0001
Incomplete specialization	815	64%	0.0472	0.9848
Complete diversification	126	10%	0.0467	0.9885
Only homogeneous good	211	17%	0.0132	0.9978
Equal probability				
Complete specialization	41	3%	0.0247	1.0113
Incomplete specialization	902	74%	0.0463	0.9990
Complete diversification	91	7%	0.0479	1.0044
Only homogeneous good	190	16%	0.0099	0.9958

Notes: Breakdown of individual nodes by specialization type. The sample is the same than that used for the regression analysis. $\overline{\theta}_i$ and $\overline{w}_i^*$ denote the average market size and the average equilibrium wages of the types of nodes.

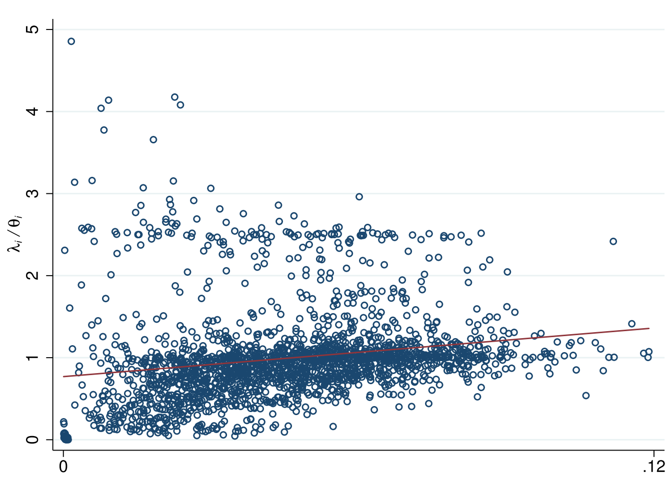

Fig. 3. Ratio of relative shares of production to market size, λ_i^* / θ_i, all nodes.

explanation relating the type of network (preferential attachment or equal probability) to the alternative regimes of node specialization or diversification. As can be seen from Table 4, the 'hub-and-spoke' topology mostly generated by the BA algorithm supports more completely specialized nodes—with more of the differentiated sector at the core of the trading network and more of the homogeneous sector as we move to the periphery. As shown in Section 2.3.1, complete specialization tends to lead to a negative association between equilibrium wages and local market size, and this effect seems to be strong enough on average in the BA networks to lead to negative coefficient estimates.[21] Indeed, focusing on all networks, 7% of nodes are completely specialized in the differentiated sector, 16% in the homogenous sector, and 69% are not fully specialized producing both goods and importing the homogeneous good. Finally, only 9% of the nodes are fully diversified in both sectors. For BA networks, these percentages are higher for the case of complete specialization, and smaller for incompletely specialized nodes. They are similar in the remaining cases.

To better see the difference in the patterns induced by the two network structures, we depict the different equilibrium types of nodes in Fig. 5. As one can see, in both types of networks the most central nodes specialize in the differentiated good—regardless of their size—while the most peripheral regions specialize in the homogeneous good. The latter holds particularly true in BA networks. As one can further see from Fig. 5, the case of incomplete specialization is the most frequent in both network types, while complete diversification can hardly be observed in the two networks. Since only the latter type is clearly associated with a positive link between local market size and equilibrium wages, this may explain why we find on average no strong correlation between those two variables in our simulations.

Based on the alternative equilibrium specialization regimes, we run separate regressions corresponding to (35) and report their marginal effects in Table 5. For nodes with complete specialization in the homogeneous good (bottom panel of Table 5), the negative effect of size on wages is not statistically different from zero (or at best marginally significant). It is, however, strongly significant for nodes specialized in the differentiated sector in BA networks (top panel), which drives the highly significant average effect.[22] For nodes not specialized in either good (middle panel), the effect is positive in the equal probability networks, whereas it is insignificant in the BA networks. This result is the only one that does not seem in line with the analytical results derived in a

[21] Recall that if country i imports some of the homogeneous good from country j, the relative wage w_i/w_j in the two countries just depends on the relative trade costs τ_{ji}/τ_{ii}, but it is independent of market sizes θ_i and θ_j. In other words, it is just the structure of the trading network that matters, but not the distribution of market sizes.

[22] Recall that there are less fully specialized nodes in the case of equal probability networks. Hence, sample sizes are smaller there, which may explain the lack of precision in the estimates.

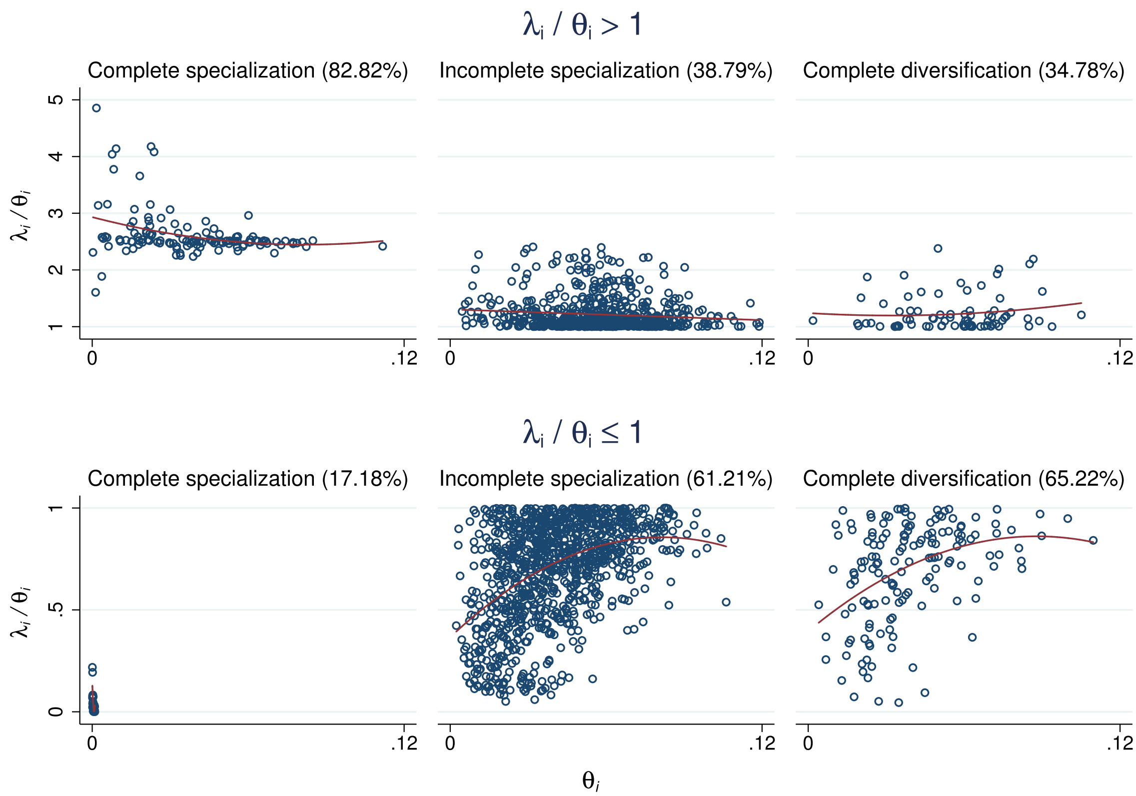

Fig. 4. Ratio of relative shares of production to market size, λ_i^*/θ_i, by node specialization pattern. Notes: Regimes are vertically aligned in decreasing order of the percentage of nodes exhibiting $\lambda_i^*/\theta_i > 1$ (reported in the headings). Nodes exhibiting $\lambda_i^*/\theta_i > 1$ are shown in the top panels, while the remaining nodes are shown in the bottom panels ($0 < \lambda_i^*/\theta_i \leq 1$).

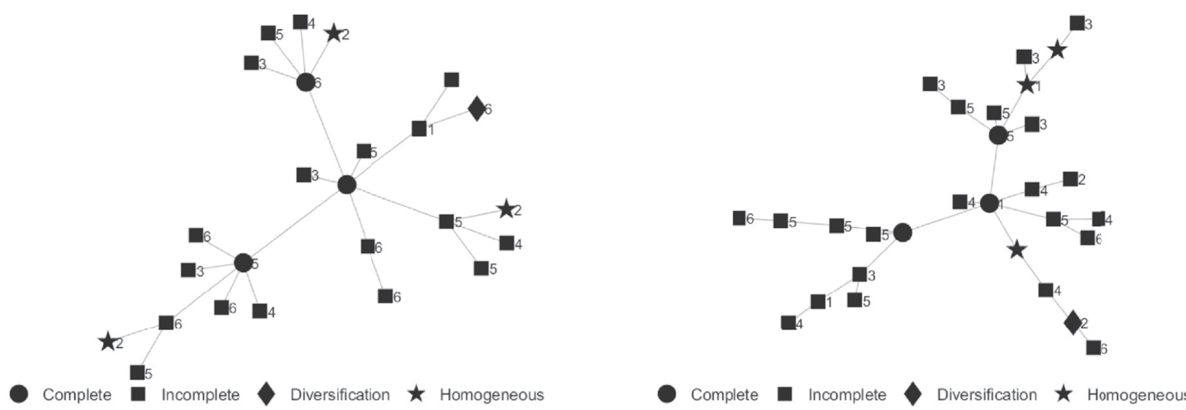

Fig. 5. Node specialization patterns, BA (left panel) and equal probability (right panel). Note: Numbers represent rounded population shares (in percentages). Missing numbers indicate small shares that would round to zero. The length of the links between nodes is not kept equal to illustrate clusters in the network topology and prevent visual cluttering.

simplified version of the model in Section 2.3.1, where we have shown that wages were independent from local market size. A positive relation emerges in the general case, and that relation is driven by the equal probability networks.

As explained before, pooling across all types of nodes allows only to compute average effects, and those average effects can go either way depending on the shares of node types. This explains the weak link between size and wages. To see how the theoretical result derived with three countries extends to the general case with many countries, we compute the correlation between θ_i and w_i^* for the countries that are specialized in the differentiated good. In that case the correlation increases from 0.09 to about 0.4. This result clearly shows that costly trade in the homogeneous good imposes strong restrictions on relative wages, and those restrictions partly destroy the positive link between market size and equilibrium wages.[23]

To sum up, our findings suggest that any analysis focusing on two countries only or disregarding the spatial structure of the trading network is likely to miss an important part of the story. Depending on the trading network, countries will display different specialization and trade patterns, and those patterns yield different relationships between local market size and equilibrium wages. Hence, the relationship between market size and wages is necessarily weaker than the relationship

[23] This result probably explains why the model cannot generate too much dispersion in relative wages—since those are constrained by trade costs differences—when applied to real data, which are not that large between European Union countries (see Section 4 below).

Table 5
Average marginal effect on w_i^* by type of node specialization.

	(i)	(ii)	(iii)	(iv)	(v)	(vi)
	Nodes specialized in the differentiated good					
Closeness_i	0.1696^a		0.2048^a		−0.0177	
	(6.03)		(7.71)		(−0.16)	
Degree_i		0.0118^a		0.0120^a		0.0124^c
		(8.33)		(10.71)		(1.71)
θ_i	-0.5901^a	-0.6486^a	-0.4514^a	-0.5443^a	−0.6510	-1.2198^b
	(−3.10)	(−3.88)	(−2.60)	(−3.66)	(−0.93)	(−2.13)
Network type	Both	Both	BA	BA	Equal	Equal
Observations	163	163	122	122	41	41
	Nodes incompletely specialized or completely diversified					
Closeness_i	0.0009		-0.0288^a		0.0245^b	
	(0.11)		(−2.88)		(2.01)	
Degree_i		0.0041^a		0.0011		0.0044^b
		(3.08)		(0.65)		(2.24)
θ_i	0.3996^a	0.3895^a	−0.0065	−0.0157	0.7829^a	0.7871^a
	(8.83)	(8.64)	(−0.13)	(−0.31)	(10.86)	(11.09)
Network type	Both	Both	BA	BA	Equal	Equal
Observations	1934	1934	941	941	993	993
	Nodes specialized in the homogeneous good					
Closeness_i	0.1101^a		0.1164^a		0.0784	
	(3.38)		(2.67)		(1.62)	
Degree_i		0.0119^b		0.0158^c		0.0083
		(2.06)		(1.91)		(1.04)
θ_i	−0.5713	−0.4646	−0.6430	−0.6001	−1.0125	−1.0505
	(−1.14)	(−0.92)	(−1.24)	(−1.16)	(−0.98)	(−0.99)
Network type	Both	Both	BA	BA	Equal	Equal
Observations	401	401	211	211	190	190

Notes: We set $\sigma = 5$, $\mu = 0.4$, and $\xi = 0.7$. *T*-statistics are given in parentheses. ^{a}, ^{b}, and ^{c} denote coefficients significant at 1%, 5%, and 10%, respectively.

between market size and industry location. We have shown this result theoretically in simple versions of the model, and we confirmed it numerically using a large set of simulations of more complex versions of the model. In our regressions, we generally obtain high values for the R^2s for the distribution of economic activity, whereas they are substantially lower for the wage relationship. Since the R^2s for equilibrium industry location are high, but the R^2s for wages are low—while the regressors are the same in both models—this suggests that size and accessibility can be relatively well separated from one another, whereas equilibrium wages are much more non-linear than equilibrium industry locations.[24]

3.2. Model 2: numerical results

We now look at the case with two differentiated CES industries. As shown in Section 2.3.1, we can obtain sharp theoretical results in special cases, involving especially two countries and specific symmetries in expenditure shares. Unfortunately, we have not been able to derive sharp results with less restrictive assumptions. In particular, it is not clear whether or not we can cleanly separate the effects of ‘absolute’ and of ‘comparative’ size differences. We know that the former leads to a positive relationship between local market size and wages, whereas the latter leads to a positive relationship between local spending patterns and industry location. We thus expect that the equilibrium results in more complex cases are a combination of both, but we do not know to what degree. Given that we have no theoretical guidance, we hence resort to numerical simulations which are the best direction to infer the possible average relationships between equilibrium industry shares and wages and the centrality and size of countries. To the best of our knowledge, this case has not been investigated until now with multiple countries. With two differentiated sectors, we have to examine the spatial distribution of firms in both sectors, $\lambda_{1i}^* \equiv n_{1i}^*/(\sum_j n_{1j}^*)$ and $\lambda_{2i}^* \equiv n_{2i}^*/(\sum_j n_{2j}^*)$, as well as the equilibrium wages w_i^*. Simple correlations among the equilibrium values are presented in Table 6.

As can be seen from Table 6, as expected size and accessibility are strongly positively linked to the equilibrium industry shares and to the equilibrium wages, respectively. Although market size still positively influences wages, there is almost no correlation between our measures of centrality and the shares of firms in the two industries. As can further be seen, there is country specialization, as shown by the negative correlation between the equilibrium shares in both industries, as well as the positive correlation with the own expenditure share, and the negative correlation with the other industry's expenditure share. In words, this specialization is strongly driven by differences in local spending patterns, as shown in the last two lines of Table 6. Our finding thus extends the result on ‘comparative advantage’ to a multi-country setting. Note, finally, that market size has roughly the same positive impact on industry location in both industries *conditional* on expenditure shares. This is the manifestation of market size as ‘absolute advantage’, which states that more centrally located countries should have, ceteris paribus, higher wages.

As for Model 1 in the previous section, we run the same regressions (34) and (35). We now run them separately for the equilibrium shares of firms in each of the two sectors, λ_{1i}^* and λ_{2i}^*. We also control for the country-specific share of expenditure on the two differentiated sectors, μ_{1i} and μ_{2i}.[25] Table 7

[24] We also estimated linear versions of the models, and the results in terms of the relative R^2s are similar. Hence, the relationship between local size and industry location seems to be ‘fairly linear.’ As shown by Behrens et al. (2009, p.262), theory predicts a linear relationship between the distribution of economic activity (λ_i) and market size (θ_i) at an interior equilibrium in the model with FPE, regardless of the number of countries and the trade cost matrix. Our results suggest that this results continues to hold approximately true and that the relationship remains fairly linear even without FPE.

[25] These shares are also randomly assigned to countries. Since these expenditure shares sum by definition to one, only one of the shares is included in the regressions.

Table 6
Simple correlations for Model 2.

	λ^*_{1i}	λ^*_{2i}	w^*_i	θ_i	closeness$_i$	degree$_i$	μ_{1i}	μ_{2i}
λ^*_{1i}	1							
λ^*_{2i}	−0.2194	1						
w^*_i	0.3202	0.3397	1					
θ_i	0.6134	0.6261	0.5314	1				
closeness$_i$	0.0008	0.0199	0.2916	0.0134	1			
degree$_i$	0.0004	0.0182	0.3295	0.0135	0.7075	1		
μ_{1i}	0.6470	−0.6587	−0.0183	−0.0157	0.0072	0.0004	1	
μ_{2i}	−0.6470	0.6587	0.0183	0.0157	−0.0072	−0.0004	−1.0000	1

Notes: We set $\sigma = 5$. Simple correlations for 100 random tree networks with a random number of 20–30 nodes. The table gives correlations at the level of individual nodes (pooled across all 100 networks). The shares λ^*_{si} are computed as $\lambda^*_{si} = n^*_{si}/(\sum_j n^*_{sj})$, for $s = 1, 2$.

Table 7
Regression results for Model 2 (industry shares).

	(i)	(ii)	(iii)	(iv)	(v)	(vi)
	Dependent variable: λ^*_{1i}					
Closeness$_i$	0.0426[b]		0.0427[c]		0.0446	
	(2.297)		(1.706)		(1.595)	
Closeness$_i^2$	−0.0300[b]		−0.0292[c]		−0.0322[c]	
	(−2.342)		(−1.701)		(−1.664)	
Degree$_i$		−0.0003		−0.0007		0.0003
		(−0.450)		(−0.728)		(0.188)
Degree$_i^2$		0.0001		0.0001		−0.0001
		(0.840)		(1.249)		(−0.313)
θ_i	1.1836[a]	1.1353[a]	1.2301[a]	1.1383[a]	1.1354[a]	1.1199[a]
	(15.263)	(23.053)	(11.331)	(16.719)	(10.117)	(15.193)
θ_i^2	−1.4622[a]	−1.4805[a]	−1.3027[c]	−1.2748[c]	−1.5630[b]	−1.5877[b]
	(−2.930)	(−2.963)	(−1.799)	(−1.756)	(−2.251)	(−2.281)
Closeness$_i \times \theta_i$	−0.1055		−0.1835		−0.0337	
	(−1.074)		(−1.283)		(−0.246)	
Degree$_i \times \theta_i$		−0.0089		−0.0123		−0.0028
		(−0.953)		(−1.103)		(−0.163)
μ_{1i}	0.1067[a]	0.1067[a]	0.1089[a]	0.1091[a]	0.1045[a]	0.1044[a]
	(74.845)	(74.771)	(55.248)	(55.263)	(50.553)	(50.461)
Constant	−0.0694[a]	−0.0548[a]	−0.0714[a]	−0.0562[a]	−0.0683[a]	−0.0537[a]
	(−10.230)	(−33.546)	(−7.823)	(−26.732)	(−6.696)	(−20.686)
Network type	Both	Both	BA	BA	Equal	Equal
Network dummy	Yes	Yes	No	No	No	No
Observations	2, 498	2, 498	1, 274	1, 274	1, 224	1, 224
Adjusted R^2	0.808	0.808	0.814	0.814	0.802	0.802
	Dependent variable: λ^*_{2i}					
Closeness$_i$	−0.0466[b]		−0.0474[c]		−0.0475[c]	
	(−2.509)		(−1.913)		(−1.676)	
Closeness$_i^2$	0.0328[b]		0.0332[c]		0.0338[c]	
	(2.560)		(1.953)		(1.721)	
Degree$_i$		−0.0000		0.0000		0.0001
		(−0.032)		(0.037)		(0.095)
Degree$_i^2$		−0.0000		−0.0000		−0.0000
		(−0.292)		(−0.386)		(−0.160)
θ_i	0.8199[a]	0.8818[a]	0.7807[a]	0.8725[a]	0.8643[a]	0.8996[a]
	(10.548)	(17.852)	(7.262)	(12.925)	(7.591)	(12.026)
θ_i^2	0.9869[b]	1.0122[b]	0.8775	0.8891	1.0313	1.0658
	(1.973)	(2.020)	(1.223)	(1.235)	(1.464)	(1.509)
Closeness$_i \times \theta_i$	0.1343		0.1917		0.0787	
	(1.365)		(1.353)		(0.566)	
Degree$_i \times \theta_i$		0.0114		0.0136		0.0073
		(1.214)		(1.222)		(0.423)
μ_{2i}	0.1073[a]	0.1073[a]	0.1086[a]	0.1087[a]	0.1059[a]	0.1058[a]
	(75.053)	(74.929)	(55.626)	(55.564)	(50.517)	(50.416)
Constant	−0.0361[a]	−0.0516[a]	−0.0358[a]	−0.0517[a]	−0.0356[a]	−0.0516[a]
	(−5.286)	(−31.176)	(−3.931)	(−24.717)	(−3.443)	(−19.233)
Network type	Both	Both	BA	BA	Equal	Equal
Network dummy	Yes	Yes	No	No	No	No
Observations	2, 498	2, 498	1, 274	1, 274	1, 224	1, 224
Adjusted R^2	0.814	0.813	0.821	0.821	0.806	0.805

Notes: We set $\sigma = 5$. OLS regressions. BA denotes networks generated using the Barabási and Albert (1999) algorithm. *T*-stats in parentheses. [a], [b], and [c] denote coefficients significant at 1%, 5%, and 10%, respectively.

shows that market size, θ_i, and the expenditure share for the two differentiated sectors, μ_{1i} and μ_{2i}, are the key variables that explain in a (mostly) linear way the spatial distribution λ_{1i}^* and λ_{2i}^* of firms in the two sectors (note that the quadratic terms for the expenditure shares are never significant and we hence drop them from the regressions). The positive average effect of market size, as shown by its associated marginal effect in Table 9, is clearly driven by the labor market clearing condition (23) which requires that the number of firms in the two sectors sums to the population share. Once we control for local market size and the spending patterns, the centrality of a country is no longer associated with its industry share. The reason is that centrality affects both industries in the *same way*, i.e., accessibility is akin to an absolute Ricardian advantage and is, therefore, capitalized into factor prices.[26]

Turning next to wages, both market size, θ_i, and centrality are positively linked to wages, w_i^*, as shown by Table 8. Countries with better access to markets and/or more trading links tend to have higher wages. Note also that the expenditure shares μ_{si} are nowhere near statistical significance in our wage regressions. In words, different expenditure shares affect industries differentially and, therefore, have no strong effect on country wages. This is again in line with our analytical results on comparative and absolute advantage.[27] When taken together, our findings suggest that the key analytical properties of the two-country model extend to multiple asymmetric countries. Finally, observe that the R^2s are much higher for the wage equation in the model with two differentiated sectors as compared to the model with a homogeneous good. This confirms that the non-linear relationships between size and wages or accessibility and wages can be better captured in the second model by the quadratic specification than in the former one. The absence of a homogeneous sector—and thus the absence of the strong restriction that sector imposes on cross-country wage differences—allows for a much cleaner average effect of local market size on wages.

To summarize, expenditure patterns (and local market size) determine the structure of country specialization in the two industries in Model 2, whereas local market size (and accessibility) have a strong impact on wages. Observe that a home market effect—when defined as a more than proportional increase in industry shares in response to an increase in local market size, i.e., $\partial\lambda_i^*/\partial\theta_i > 1$—generally does not arise, as shown in Table 9. The reason is that when all sectors are operating under increasing returns and face trade costs, not all of them can—by definition—exhibit home market effects (see also Hanson and Xiang, 2004). In that case, an alternative definition of the HME, involving both the size θ_i and the expenditure share μ_{si}, would be required. To the best of our knowledge, such a definition has not been used to date in the literature.

3.3. Summary of the simulation results

A number of findings emerge from the foregoing analyses of the two models. Let us briefly summarize the key insights. First, starting from Model 1 with a single CES sector, we have seen that accessibility has a strong impact on industry location and, to a lesser extent, on wages. This suggests that any analysis involving trade in homogeneous goods and focusing on two countries only—or disregarding the spatial structure of the trading network entirely—is likely to miss an important part of the story. Secondly, we have shown that the correlations between w_i^* and either λ_i^* or θ_i are quite low, i.e., there is no strong correlation between either market size or the equilibrium industry shares and the equilibrium wages. As we have explained and theoretically shown, this unexpected result is most likely due to the fact that different specialization patterns impose strong restrictions on the relative wages of the trading partners, which breaks the clear link between market size and wages. Although this result needs to be partly qualified—see the results in the middle panel of Table 5—we believe that it is largely driven by these differences in node types: for many pairs of nodes, relative wages depend on relative trade costs only but are (partly or fully) independent of the countries' market sizes. This finding once more suggests that going beyond the two-country case is very important. More work is called for here. Last, the home market effect—when defined as a more than one-for-one increase in industry shares in response to changes in local market size—generally holds even when trading the homogeneous good is costly, provided that it is less costly than trading the differentiated good (as in Davis, 1998).

Turning to Model 2 with two CES sectors, both absolute market size—as captured by θ_i—and centrality—as measured by either closeness or the degree distribution—are capitalized into factor prices, thus showing that they constitute absolute advantage affecting all industries in the same way. Differences in spending patterns—as captured by the μ_{si}—are however capitalized into industry structure, thus showing that they constitute comparative advantage affecting industries differently. Our findings, therefore, numerically extend the theoretical results of Behrens and Ottaviano (2011) to a multi-country setting.

Last, it is worth pointing out that the effects of accessibility and market size on wages are an order of magnitude smaller in Model 1 than in Model 2 (compare Tables 3 and 9). As explained, the reason is that the equalization of prices in the traded homogeneous sector imposes strong restrictions on the determination of wages among trading partners when specialization is incomplete (a very frequent case). This in turn breaks the link between accessibility and market size in the wage determination. In a nutshell, as expected market size and centrality matter more the more industries are subject to trade costs and increasing returns to scale, as is the case in Model 2 of our analysis. This suggests that any empirical analysis needs to focus on the 'right set' of industries.

4. Numerical application to EU countries

Our foregoing simulations highlight regularities of multi-country trade models without FPE, yet they cannot provide a sense of how well those models perform when confronted with data. The aim of this section is hence to use 'calibrated' versions of the models to check their fit with the data and to run some counterfactuals.

4.1. Data

We compute the equilibria of the two models using country-level data for 25 European Union (EU) countries in 2001 and 2012.[28,29] Table 12 in Appendix C provides details on the variables used to solve the different models. For both models, these include the population shares (θ_i) and 'trade costs' (τ_{ij}), which we proxy using population and minimum travel

[26] This result would be weakened if accessibility affected industries in different ways (as in, e.g., Hanson and Xiang, 2004). In that case, accessibility would also be in part a 'comparative advantage' and would, therefore, have a much stronger impact on industry location and not only wages.

[27] We performed extensive sensitivity analyses with respect to σ and μ_{1i} and μ_{2i} in Model 2. Holding the network structure constant, we study the behavior of industry location (λ_{1i} and λ_{2i}) and country wages, w_i, when these basic parameters change. The results consistently show that increasing σ in the range $(1, 10]$ leads to higher nominal wages, while industry shares remain largely unaffected. In accord with the regression results that we report in the main text, the income shares map into the specialization. Solving a hundred times the model for each specific network and assigning random values of μ_{1i} and μ_{2i}, we find that industry shares present a strong correlation with income shares ($\rho = 0.7$). Income shares, however, are basically uncorrelated with nominal wages.

[28] We initially have 28 EU countries, but we exclude Lithuania due to missing data, and Cyprus and Malta because of their insular nature that affects the computation of minimum-time road travel distance.

[29] In a supplemental online appendix, we also provide an application to Spanish regions. The advantage of the Spanish case is that we focus on one country and, therefore, have less heterogeneity that may affect our results. We also have a better (and more direct) measure for transport costs. However, the assumption of immobile population is obviously much less realistic for the Spanish case as compared to the EU case.

Table 8
Regression results for Model 2 (wages).

	(i)	(ii)	(iii)	(iv)	(v)	(vi)
			Dependent variable: w_i^*			
Closeness_i	−0.0827		−0.0687		−0.1076	
	(−1.040)		(−0.576)		(−1.038)	
Closeness_i^2	0.2613[a]		0.2888[a]		0.2489[a]	
	(4.763)		(3.524)		(3.468)	
Degree_i		0.0373[a]		0.0342[a]		0.0373[a]
		(12.176)		(7.696)		(6.863)
Degree_i^2		−0.0011[a]		−0.0009[b]		−0.0006
		(−3.573)		(−2.114)		(−0.828)
θ_i	5.0415[a]	3.9362[a]	4.8117[a]	3.5859[a]	5.4439[a]	4.3179[a]
	(15.169)	(19.053)	(9.288)	(11.109)	(13.081)	(16.392)
θ_i^2	−16.3637[a]	−16.5993[a]	−13.9718[a]	−15.0340[a]	−19.2477[a]	−18.4263[a]
	(−7.649)	(−7.919)	(−4.043)	(−4.369)	(−7.475)	(−7.408)
$\text{Closeness}_i \times \theta_i$	−2.4894[a]		−2.8056[a]		−2.3729[a]	
	(−5.917)		(−4.110)		(−4.669)	
$\text{Degree}_i \times \theta_i$		−0.2533[a]		−0.2405[a]		−0.2709[a]
		(−6.432)		(−4.531)		(−4.472)
μ_{1i}	−0.0061	−0.0055	−0.0071	−0.0085	−0.0061	−0.0029
	(−0.997)	(−0.922)	(−0.759)	(−0.909)	(−0.795)	(−0.389)
Constant	0.7982[a]	0.7943[a]	0.7801[a]	0.7967[a]	0.8074[a]	0.7808[a]
	(27.445)	(115.967)	(17.907)	(79.922)	(21.334)	(84.090)
Network type	Both	Both	BA	BA	Equal	Equal
Network dummy	Yes	Yes	No	No	No	No
Observations	2, 498	2, 498	1, 274	1, 274	1, 224	1, 224
Adjusted R^2	0.393	0.417	0.330	0.337	0.477	0.513

Notes: We set $\sigma = 5$. OLS regressions. BA denotes networks generated using the Barabási and Albert (1999) algorithm. *T*-stats in parentheses. [a], [b], and [c] denote coefficients significant at 1%, 5%, and 10%, respectively.

Table 9
Marginal effects for Model 2.

	(i)	(ii)	(iii)	(iv)	(v)	(vi)
	Average marginal effect for λ_{1i}^*					
Closeness_i	−0.0001		−0.0014		0.0010	
	(−0.05)		(−0.36)		(0.26)	
Degree_i		−0.0005		−0.0008		−0.0001
		(−1.27)		(−1.54)		(−0.14)
θ_i	0.9988[a]	0.9996[a]	1.0102[a]	1.0126[a]	0.9881[a]	0.9874[a]
	(69.65)	(69.57)	(50.87)	(50.94)	(47.58)	(47.42)
Network type	Both	Both	BA	BA	Equal	Equal
	Average marginal effect for λ_{2i}^*					
Closeness_i	0.0010		0.0021		−0.0001	
	(0.34)		(0.53)		(−0.01)	
Degree_i		0.0004		0.0005		0.0003
		(0.98)		(0.91)		(0.53)
θ_i	0.9852[a]	0.9848[a]	0.9716[a]	0.9697[a]	0.9985[a]	0.9989[a]
	(68.54)	(68.33)	(49.40)	(49.20)	(47.39)	(47.27)
Network type	Both	Both	BA	BA	Equal	Equal
	Average marginal effect for w_i^*					
Closeness_i	0.1535[a]		0.1825[a]	0.0213[a]	0.1240[a]	0.0240[a]
	(12.89)		(9.64)	(8.98)	(8.63)	(11.80)
Degree_i		0.0231[a]		0.0213[a]		0.0240[a]
		(15.20)		(8.98)		(11.80)
θ_i	2.1318[a]	2.1210[a]	1.9269[a]	1.9206[a]	2.3459[a]	2.3224[a]
	(34.68)	(35.19)	(20.33)	(20.38)	(30.46)	(31.21)
Network type	Both	Both	BA	BA	Equal	Equal

Notes: *T*-statistics are given in parentheses. [a], [b], and [c] denote coefficients significant at 1%, 5%, and 10%, respectively.

time by road, respectively. For the latter, following the definition of the freeness of trade, ϕ_{ij}, transport costs are computed as follows: $\phi_{ij} = \tau_{ij}^{1-\sigma} = \left(\text{Time}_{ij}/\min\{\text{Time}_{ij}\}\right)^{1-\sigma} \in [0, 1]$. For Model 1, we associate the homogeneous sector with agriculture, while the differentiated sector corresponds to the manufacturing industry. The share of gross value added in the differentiated sector is used to proxy the observed n_i or λ_i, which we will contrast with the equilibrium values n_i^* or λ_i^*, backed out

Table 10
Simulation results for Model 1 and Model 2.

Country	Model 1				Model 2					
	λ_i^*		w_i^*		λ_{1i}^*		λ_{2i}^*		w_i^*	
	2001	2012	2001	2012	2001	2012	2001	2012	2001	2012
(AT) Austria	0.021	0.016	1.000	1.000	0.015	0.016	0.021	0.022	1.000	1.000
(BE) Belgium	0.000	0.008	1.163	1.016	0.019	0.019	0.021	0.023	1.065	1.053
(BG) Bulgaria	0.015	0.017	0.941	0.533	0.020	0.020	0.009	0.009	0.963	0.963
(CZ) Czech Republic	0.022	0.021	1.011	0.981	0.022	0.022	0.023	0.023	1.027	1.022
(DE) Germany	0.185	0.188	1.246	1.182	0.175	0.178	0.200	0.202	1.274	1.273
(DK) Denmark	0.013	0.010	0.916	1.16	0.013	0.012	0.013	0.012	0.949	0.94
(EE) Estonia	0.003	0.001	0.811	2.084	0.003	0.003	0.002	0.002	0.808	0.806
(EL) Greece	0.020	0.018	0.963	0.612	0.02	0.018	0.021	0.019	0.993	0.978
(ES) Spain	0.080	0.082	1.123	1.016	0.083	0.086	0.077	0.079	1.158	1.159
(FI) Finland	0.011	0.012	0.975	0.862	0.010	0.010	0.013	0.013	0.954	0.96
(FR) France	0.123	0.124	1.223	1.052	0.116	0.117	0.133	0.133	1.214	1.212
(HR) Croatia	0.007	0.007	0.917	0.864	0.008	0.009	0.005	0.005	0.883	0.887
(HU) Hungary	0.018	0.016	0.993	0.867	0.019	0.018	0.017	0.016	0.997	0.988
(IE) Ireland	0.008	0.007	0.884	1.630	0.008	0.008	0.009	0.009	0.900	0.898
(IT) Italy	0.107	0.110	1.200	0.910	0.114	0.117	0.097	0.099	1.194	1.195
(LU) Luxembourg	0.000	0.000	1.318	1.353	0.001	0.001	0.001	0.001	1.009	1.013
(LV) Latvia	0.004	0.004	0.85	1.453	0.005	0.004	0.004	0.003	0.848	0.836
(NL) Netherlands	0.062	0.054	1.036	0.98	0.033	0.034	0.047	0.046	1.100	1.094
(PL) Poland	0.066	0.071	1.113	1.049	0.078	0.082	0.051	0.053	1.135	1.139
(PT) Portugal	0.024	0.021	0.984	0.86	0.022	0.019	0.027	0.023	1.014	0.997
(RO) Romania	0.051	0.041	1.078	0.687	0.069	0.054	0.025	0.02	1.101	1.068
(SE) Sweden	0.020	0.021	0.987	0.755	0.020	0.020	0.021	0.021	0.996	0.995
(SI) Slovenia	0.000	0.005	1.012	0.832	0.004	0.004	0.005	0.005	0.882	0.871
(SK) Slovakia	0.010	0.013	0.938	0.807	0.012	0.012	0.007	0.008	0.944	0.946
(UK) United Kingdom	0.129	0.133	1.209	1.214	0.113	0.116	0.152	0.155	1.220	1.221
Mean	0.040	0.040	1.036	1.030	0.040	0.040	0.040	0.040	1.025	1.020
Std. Dev	0.049	0.050	0.132	0.333	0.047	0.048	0.052	0.053	0.124	0.124
Max	0.185	0.188	1.318	2.084	0.175	0.178	0.200	0.202	1.274	1.273
Min	0.000	0.000	0.811	0.533	0.001	0.001	0.001	0.001	0.808	0.806

Notes: Simulations use the following values. For Model 1, we let $\sigma = 5$, $\mu = 0.4$, and $\xi = 0.7$. For Model 2, we let $\sigma = 5$. See Appendix C for the raw data used in the simulations.

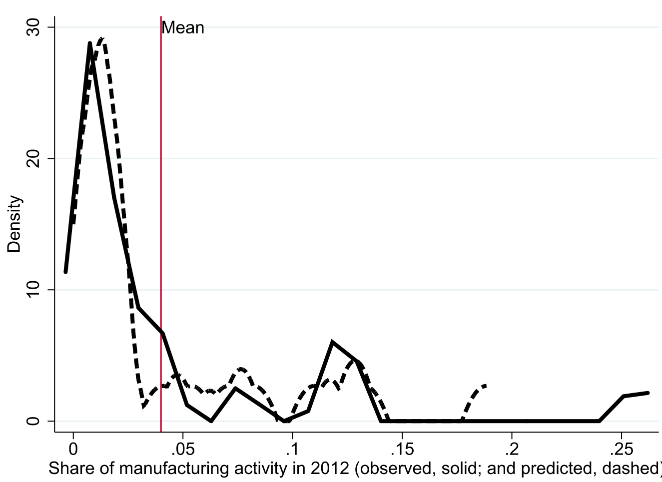

Fig. 6. Observed and predicted manufacturing distributions for Model 1.

from the models. For Model 2, λ_{1i} and λ_{2i} are associated with manufacturing and with services gross value added, respectively.[30] Population and industrial gross value added for 2012 are obtained from the European Union Open Data Portal.[31] Bilateral trade costs are measured as the time needed for traveling between origin i and destination j, assuming cost minimizing behavior. Using data provided by Stelder et al. (2013) and ArcGIS, we compute proxies for bilateral transport costs between EU countries using digitized road and maritime transportation networks. These trade cost proxies are based on shortest-route travel time between country centroids (see also Stelder, 2016). Consistent with the results presented in Spiekermann and Wegener (2008), Germany has the highest closeness centrality within the original members of the EU, followed by the most central countries in the current EU, in particular the Czech Republic (0.99) and Austria (0.98). The lowest centrality is for peripheral countries such as Finland (0.41), Portugal (0.45), and Ireland (0.48). In Table 12, we report the value of the centrality index for each country computed using this particular time-cost proxy for trade costs.

As for the remaining structural parameters, μ and σ, of the models we use key estimates reported in the literature that have proved robust in several country studies (e.g., Table 5 in Head and Mayer, 2004). More recently, Broda and Weinstein (2006) estimate the elasticities of substitution for traded goods imports to the US using SITC rev2 for 1972–1988, and SITC rev3 for 1990–2001 at the 3-, 4-, and 5-digit levels, respectively. At the 3-digit level and across all goods, they find a mean elasticity of 6.8 from 1972–1988 and of 4.0 from 1990–2001, respectively. Looking only at differentiated goods—as defined using the Rauch (1999) classification—at the 4-digit level, they find a mean elasticity of 5.2 from 1972–1988 and of 4.7 from 1990–2001, respectively. Therefore we take a midpoint value of $\sigma = 5$ (as we also assumed in the numerical simulations performed in the previous sections). Turning to the expenditure shares on the differentiated products, μ_{1i} and μ_{2i}, we use the expenditure shares for manufacturing goods and services in total domestic demand from the European Union Open Data Portal for each country in 2012. For Model 1 we round the value to $\mu = 0.4$ (as we also assumed in the numerical simulations).

[30] We leave out agriculture from the analysis. We determine expenditure shares to match the production side from the overall structure of consumption expenditure by detailed COICOP level, with the first share corresponding to manufacturing and utilities (processed food, clothing, water, electricity, …) and the second one to services (health, communication, leisure, education, accommodation, …).

[31] Available at https://open-data.europa.eu/data/dataset.

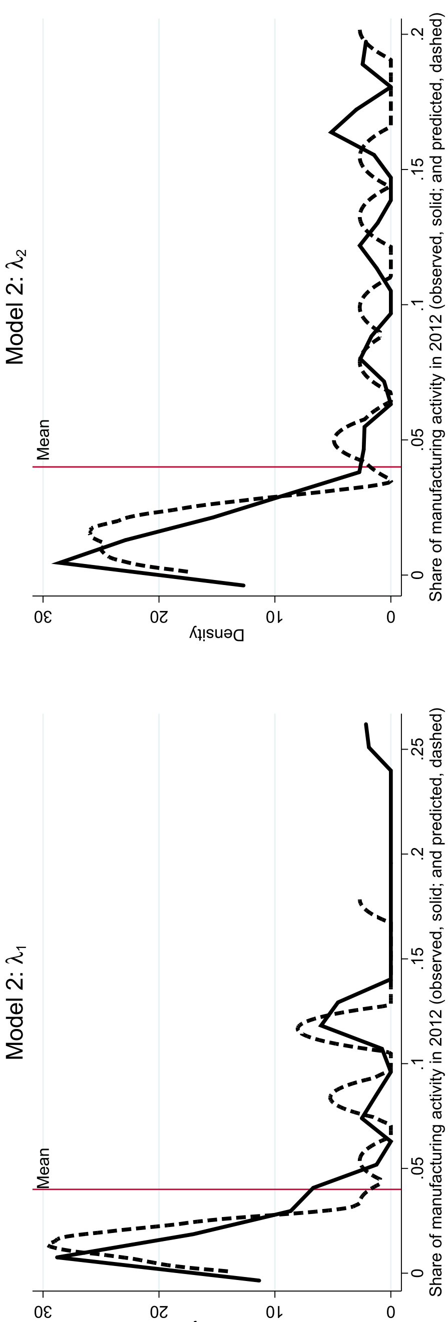

Fig. 7. Observed and predicted distributions of sectors for Model 2.

As for the parameter ξ capturing the relative level of trade cost of the homogeneous good compared to the differentiated good, we adopt a value 0.7. Based on data from the 'Ongoing Survey on Freight Road Transportation', carried out by the Ministry of Transport in Spain under European Union guidelines (see Ministerio de Fomento, MFOM, 2007a), we can calculate a comparative range of relative freight costs in terms of tons-kilometer.[32] The difference in the cost of shipping homogeneous and differentiated products ranges from 0.7 to 1, with an average around 0.8. To keep consistency with the values adopted in the previous section, we take the lower bound for ξ.

Last, since wages are endogenous, we require additional data to test whether the results of the calibrated model match the observed values. In particular, we need information on wages. The latter are obtained, as in many previous studies, by dividing aggregate gross value added by total employment (see the literature review in Head and Mayer, 2004).

4.2. Equilibrium distributions vs. observed distributions

We compare the equilibrium distribution of economic activity and the wages predicted by our models with the data, taking country sizes (i.e., the θ_i's) and the transportation network (i.e., the τ_{ij}'s) as primitives. Doing so will allow us to assess to what extent the models can 'replicate' the observed distributions of wages and industries.

The equilibrium distributions of firms in 2001 and 2012, as well as the equilibrium wages, are summarized in Table 10. Our results show large disparities in the distribution of manufacturing across countries. The distribution of firms varies from almost 0% to about 18%. Unsurprisingly, the largest EU countries have the largest shares of firms in both models. To tentatively gauge the predictive power of the models, we first compare the actual and the predicted distributions of the manufacturing shares, and then check the statistical significance of the differences between them for the alternative equilibria of the models: λ_i^* (Model 1) and λ_{1i}^* and λ_{2i}^* (Model 2). A visual inspection of the distribution of the manufacturing shares of the two models (Fig. 6 for Model 1, and Fig. 7 for Model 2) corresponding to the most recent year show that the models are capable of replicating the observed distributions fairly well, with the exception of the largest EU country (Germany), for which the actual share is 0.257 but the predicted share in Model 1 is only 0.190. A similar pattern is observed in Model 2.

Turning to more formal statistical tests, besides Pearson's r and Spearman's ρ coefficients of correlation, we also test the equality of the distributions by way of a Kolmogorov-Smirnoff test. Table 11 summarizes our results.

As can be seen from Table 11, both for linear (Pearson) and rank (Spearman) correlations the coefficients are large and highly significant. The Pearson standard correlation is 0.94 for λ_{1i}^* in Model 2 and a remarkable 0.980 for λ_{2i}^*. The values for the Spearman correlations confirm that the rank-orders observed in the data are mirrored by the solution values. Additionally, the hypothesis of equality of distributions cannot be generally rejected. Our results show that solving the models using real data yields model equilibrium distributions of economic activity that are in many cases statistically hard to distinguish from those observed in the real economy. Of course, population and economic activity are highly correlated in the data (0.96), so those results do not come as a big surprise.

We hence next turn to wages.[33] Here, the correlation is weaker—as expected, given our theoretical results—yet still positive. The

[32] The 'Ongoing Survey on Freight Road Transportation' classifies shipments of manufactured goods according to Council Regulation (EC) No 1172/98, and the prevalent type of vehicle used to transport each type of good, along with the information provided by the Observatory of Road Freight Transport Costs on each type of vehicle (MFOM, 2007b).

[33] Without loss of generality we choose the wage of the first country, Austria, as the numeraire.

Table 11
Differences between observed and model distributions.

Test		Pearson's[1] r		Spearman's[2] ρ		Kolmogorov-Smirnov[3]	
Model	Share	2001	2012	2001	2012	2001	2012
1	λ_i^*	0.9508	0.9585	0.7438	0.8562	0.1600	0.1600
		(0.000)	(0.000)	(0.000)	(0.000)	(0.877)	(0.877)
2	λ_{1i}^*	0.9249	0.9430	0.7608	0.8646	0.2400	0.2000
		(0.000)	(0.000)	(0.000)	(0.000)	(0.414)	(0.649)
	λ_{2i}^*	0.9802	0.9798	0.8754	0.8908	0.2400	0.1600
		(0.000)	(0.000)	(0.000)	(0.000)	(0.414)	(0.877)
1	w_i^*	0.5541	0.2642	0.3231	0.3377	0.5200	0.4400
		(0.004)	(0.202)	(0.115)	(0.099)	(0.001)	(0.010)
2	w_i^*	0.2871	0.1884	0.2215	0.1738	0.5200	0.5200
		(0.164)	(0.367)	(0.286)	(0.404)	(0.001)	(0.001)

Notes: [1,2]The null hypothesis is that both variables are independent; [3]The null hypothesis is that both variables come from the same continuous distribution. p-values for all tests in parenthesis.

correlations range from 0.26 to 0.55 for Model 1 in 2001, and are almost equal to 0.33 in 2012, whereas they are fairly low (from 0.17 to 0.28 depending on the year) and not significant in Model 2. For both models, a Kolmogorov-Smirnoff test fails to reject the null that the two distributions are different, although Model 1 globally performs better than Model 2. In a nutshell, we do not find large and statistically strong correlations between wages as proxied by GVA per employee (our empirical counterpart for 'wages' at the aggregate level) and the equilibrium relative wages implied by the two models. Hence, while the models perform well in terms of their spatial predictions of economic activity, they perform much worse in terms of their predictions for prices. There are at least four possible reasons for this. First, as shown in the previous section, the multi-country simulated models do not deliver clear results as to the roles of market size and centrality on wages. It is thus not surprising that their empirical fit to wage data in also fairly weak. Second, GVA per employee—though widely used in the literature (see Head and Mayer, 2004)—is only a crude proxy for wages. Third, it is well known that cross-country simulations suffer significantly from unobserved country-level heterogeneity that is not embedded into the model. We provide an application to regions within a country (Spain) in the supplemental online appendix. Yet, even in that case, the model does not do a fairly good job in predicting relative wages, especially their dispersion (which is much higher in the data than in the model). Last, there are many factors outside the models or not concurring with the theoretical assumptions, such as the existence of unemployment and the absence of labor market clearing—particularly in southern European countries. These factors may also explain the divergences between the predictions of the model and the observed data.

5. Conclusions

Combining both theoretical results and systematic numerical simulations for two different trade models—one with a homogeneous and a differentiated sector, and one with two differentiated sectors—we have studied whether and how size and accessibility are linked to the equilibrium industry shares and to wages when there is no factor price equalization. Our key findings can be summarized as follows.

First, in accord with the theoretical results derived in lower-dimensional instances of the models, the effect of local market size on equilibrium wages crucially hinges on the countries' specialization patterns. This finding generalizes to our numerical simulations.

Second, in all models that we simulate, the equilibrium relationship between local market size and industry location is more robust than the relationship between local market size and wages. Although the results vary slightly depending on the type of trading network considered, they are quite systematic. This suggests that equilibrium industry location is fairly robust to the structure of the trade networks, but not wages and the terms of trade. While two-country results can be partly extrapolated to make statements about industry location in a multicountry world, this definitively holds not true for two-country results on wages.

Third, the correlation between equilibrium wages and equilibrium industry shares is rather low, thus suggesting that both variables operate largely independently. Empirical tests and formal definitions of the home market effect should therefore take into account both dimensions—industry location and wages—in order to be relevant. To the best of our knowledge, tests looking simultaneously at industry location and factor prices have not yet been devised.

Last, when applied to European Union country-level data, we find that in both cases the models generally predict well the distribution of industries, yet predict less well wages. A formal test does not allow to reject the null hypothesis that the industry distribution predicted by the models is the same as that observed in the data. The test does, however, reject the equality of observed and predicted wage distributions, because the stylized models cannot replicate the observed dispersion in wages across countries. This again shows that the models do a good job at predicting industry location, but are less robust in terms of predicted wages. Empirical work using wages to test for new economic geography effects should thus be considered with some caution.

Appendix

A.1. Factor price equalization

Assume that the homogeneous good can be costlessly traded across all countries. This is the case usually considered in the literature (e.g., Helpman and Krugman, 1985). Marginal cost pricing then implies that the price of the homogeneous good is equal to the wage, which must be the same everywhere. In other words, factor price equalization (FPE) holds.

In a multi-country world, the assumption of FPE has a major technical drawback. To see this, ask under what conditions FPE will hold? Clearly, FPE will hold if and only if some homogeneous good is produced in every country. Following Behrens et al. (2007, 2009), a *sufficient condition* is that

$$\theta_i > \mu, \quad \forall i = 1, 2, \ldots, M. \tag{A-1}$$

When (A-1) holds for all countries, and when trade in the homogeneous good is free, we have $w_i = 1$ for all $i = 1, 2, \ldots, M$. Observe that condition (A-1) is extremely restrictive. Consider, e.g., a world with 30 countries. If market sizes θ_i were identical across countries, we must have $\mu < 1/30$. This is already very restrictive. But in our case, since we randomly assign the shares θ_i to countries, we may have very small shares in some cases. In those cases, the foregoing restriction can never be met for 'reasonable values' of μ.

Although condition (A-1) is technically speaking only a *sufficient* condition – i.e., we may still have FPE even when it is violated – it seems still very unlikely to be met in general. Another potential problem in the FPE version of the model is that it displays a much larger share of 'corner equilibria', i.e., equilibria in which some countries are deindustrialized and do not host any of the differentiated sector. We have simulated the model with FPE and find that the number of nodes with a zero industry share is 920 out of 2498, i.e. 36.82%. In a nutshell, the FPE model does not make much sense in a world with many countries, neither theoretically nor empirically, and it is difficult to implement consistently for reasonable values of μ. We thus disregard it in the remainder of this paper.

A.2. Complete specialization

In the case of complete specialization with three countries, we have by definition, $X_{22} = (1-\mu)L_2$ and $X_{33} = (1-\mu)L_3$. We further have

$$X_{21} + X_{31} = \frac{(1-\mu)w_1 L_1}{\xi\tau_{21}}, \tag{A-2}$$

with $w_2 \equiv 1$, i.e., the wage in country 2 is the numeraire, and $w_3 = \tau_{21}/\tau_{31}$ since both countries supply country 1 in the traditional good. Hence, $w_2\xi\tau_{21} = w_3\xi\tau_{31}$ must hold. The labor market clearing conditions in countries 2 and 3 are $(1-\mu)L_2 + \tau_{21}\xi X_{21} = L_2$ and $(1-\mu)L_3 + \tau_{31}\xi X_{31} = L_3$, which directly yields

$$X_{21} = \frac{\mu L_2}{\tau_{21}\xi} \quad \text{and} \quad X_{31} = \frac{\mu L_3}{\tau_{31}\xi}.$$

Substituting into (A-2) allows us to solve for the wage in country 1 as follows:

$$w_1 = \frac{\mu}{1-\mu}\frac{(1-\theta)}{2\theta}\left(1 + \frac{\tau_{21}}{\tau_{31}}\right) \tag{A-3}$$

One can verify that $\text{RMP}_1 = 1$ in that configuration, whereas

$$\text{RMP}_2 = w_1^{\sigma}\left[\mu\phi_{21} + \frac{1-\mu}{1+\tau_{21}/\tau_{31}}\left(\frac{1}{\phi_{12}} + \frac{\tau_{21}}{\tau_{31}}\frac{\phi_{32}}{\phi_{31}}\right)\right]$$

is a complicated function of all trade barriers. An analogous expression holds for RMP_3. Depending on the values of trade costs, $\text{RMP}_2 < 1$ and $\text{RMP}_3 < 1$ can hold, as required for equilibria with complete specialization. For example, if we set $\phi_{23} \approx 0$ and $\tau_{21} = \tau_{31}$, we need $\mu > \theta$ so that $w_1 > 1$, and this can be made compatible with $\text{RMP}_2 < 1$ and $\text{RMP}_3 < 1$. Generally, equilibria with complete specialization seem to be more likely to occur when the number of regions increases. [34]

A.3. Incomplete specialization

Incomplete specialization of type 1. The equilibrium with incomplete specialization is determined as follows. We have $X_{12} = 0$ and $X_{22} = (1-\mu)L_2$ and

$$X_{11} > 0 \quad \text{and} \quad X_{21} = \frac{(1-\mu)w_1 L_1}{\xi\tau} - X_{11}. \tag{A-5}$$

Equating labor supply and demand in country 2, including the traditional good lost in shipping, we thus have $L_2 = X_{22} + \xi\tau X_{21} = (1-\mu)L_2 + (1-\mu)w_1L_1 - \xi\tau X_{11}$. Solving for L_2, we have

$$L_2 = \frac{\xi\tau}{\mu}[(1-\mu)L_1 - X_{11}] \tag{A-6}$$

since $w_1 = \xi\tau$. Dividing by L and rewriting yields:

[34] For the case of two countries, where 1 is the core and 2 the periphery, we have.

$$\text{RMP}_2 = w_1^{\sigma}\left[\mu\phi + \frac{1-\mu}{\phi}\right], \tag{A-4}$$

which needs to be smaller than 1 for an equilibrium with complete specialization to exist. Although given for the case of two countries, this could be given for any symmetric star configuration (since the model can be reduced to the 2-country case under the assumption of symmetry; see footnote [12]). Hence, this result holds for any symmetric setting with more than two countries. Observe that for the 3-country case, the denominator in expression (A-4) would be 2ϕ. Hence, more countries make it harder that the bracketed term is greater than one, thus making equilibria with complete specialization more likely.

$$1-\theta=\frac{\xi\tau}{\mu}\left[(1-\mu)\theta-\widetilde{X}_{11}\right] \quad \Rightarrow \quad \widetilde{X}_{11}=(1-\mu)\theta-\frac{(1-\theta)\mu}{\xi\tau}. \tag{A-7}$$

The latter expression is positive if and only if

$$\xi\tau>\frac{\mu}{1-\mu}\frac{1-\theta}{\theta}. \tag{A-8}$$

Since $\xi\tau>1$, a sufficient condition for (A-8) to hold is that $\mu>\theta$. Since $L_1^H=X_{11}=L_1-n_1\mu L$, using (A-7) then yields

$$(1-\mu)\theta-\frac{(1-\theta)\mu}{\xi\tau}=\theta-\mu n_1 \quad \Rightarrow \quad n_1=\theta+\frac{1-\theta}{\xi\tau},$$

which reveals that $n_1>\theta$ and $0<\partial n_1/\partial\theta<1$. It can further readily be verified that $\mathrm{RMP}_1=1$ and that

$$\mathrm{RMP}_2=\frac{w_1^\sigma}{n_1}\left(\theta\phi+\frac{1-\theta}{\phi\xi\tau}\right) \tag{A-9}$$

Expression (A-9) needs to be strictly smaller than 1 for an equilibrium with incomplete specialization to occur. To see when this can be the case, assume that $\xi<1$, i.e., trading the homogeneous good is less costly than trading the differentiated good (Davis, 1998). Assume also, for the sake of the argument, that $\tau\approx 1$, i.e., trade costs for manufactures are low. In that case, we have

$$\mathrm{RMP}_2=\frac{\xi^\sigma}{\theta+(1-\theta)/\xi}\left(\theta+\frac{1-\theta}{\xi}\right)\Big\langle 1 \quad \text{iff} \quad \xi<1.$$

and this type of equilibrium can occur.

Incomplete specialization of type 2. Proceeding as in the case of type 1 incomplete specialization, and equating labor supply and demand in country 2 we now have $L_2=X_{22}+\xi\tau X_{21}+L_2^D=(1-\mu)L_2+(1-\mu)w_1L_1-\xi\tau X_{11}+L_2^D$, with $L_2^D=n_2\mu L$ (recall that country 2 has also some of the modern sector). Solving for $\widetilde{X}_{11}$, we obtain

$$\widetilde{X}_{11}=(1-\mu)\theta-\frac{(1-\theta-n_2)\mu}{\xi\tau}, \tag{A-10}$$

since $w_1=\xi\tau$. Since $L_1^H=X_{11}=L_1-n_1\mu L$, using (A-10) then yields

$$(1-\mu)\theta-\frac{(1-\theta-n_2)\mu}{\xi\tau}=\theta-\mu n_1 \quad \Rightarrow \quad n_1=\theta+\frac{1-\theta-n_2}{\xi\tau}. \tag{A-11}$$

For $n_2<1-\theta$, we again have $n_1>\theta$ and $0<\partial n_1/\partial\theta<1$. To solve explicitly for n_1 and n_2, we can use either $\mathrm{RMP}_1(n_1,n_2)=1$ or $\mathrm{RMP}_2(n_1,n_2)=1$ and (A-11).

A.4. Complete diversification

In the case of complete diversification, there are modern firms in each country. The two real market potentials are given by

$$\begin{aligned}
\mathrm{RMP}_1 &= \left[\frac{w_1\theta}{w_1^{1-\sigma}\theta+(1-\theta)\phi}+\frac{(1-\theta)\phi}{w_1^{1-\sigma}\theta\phi+(1-\theta)}\right]w_1^{-\sigma}\\
\mathrm{RMP}_2 &= \left[\frac{w_1\theta\phi}{w_1^{1-\sigma}\theta+(1-\theta)\phi}+\frac{1-\theta}{w_1^{1-\sigma}\theta\phi+(1-\theta)}\right]
\end{aligned}$$

Letting $\mathrm{RMP}_1=\mathrm{RMP}_2=1$ and totally differentiating, we can use the implicit function theorem to show that $\mathrm{d}w_1/\mathrm{d}\theta>0$, i.e., local wages are increasing in local market size. Indeed, we have

$$\frac{\partial\mathrm{RMP}_1}{\partial\theta} = \frac{w_1^{1-\sigma}\phi}{\left[w_1^{1-\sigma}\theta+\phi(1-\theta)\right]^2}-\phi^2\frac{w_1^{1-\sigma}(2\theta-1)}{\left[w_1^{1-\sigma}\theta\phi+(1-\theta)\right]^2} \tag{A-12}$$

$$\begin{aligned}
\frac{\partial\mathrm{RMP}_1}{\partial w_1} = & -\frac{(\sigma-1)\theta w_1^{-\sigma}\phi(1-\theta)}{\left[w_1^{1-\sigma}\theta+\phi(1-\theta)\right]^2}\\
& -\frac{\sigma(1-\theta)^2\phi w_1^{-\sigma-1}+(1-\theta)\theta\phi^2w_1^{-2\sigma}}{\left[w_1^{1-\sigma}\theta\phi+(1-\theta)\right]^2},
\end{aligned} \tag{A-13}$$

with (A-12) being always negative, and (A-13) being positive around $\theta = 1/2$. Hence, by the implicit function theorem, we have

$$\frac{\mathrm{d}w_1}{\mathrm{d}\theta}|_{\theta=1/2} = -\frac{\partial \mathrm{RMP}_1/\partial\theta}{\partial \mathrm{RMP}_1/\partial w_1}|_{\theta=1/2} > 0.$$

More generally, this relationship holds for all $\theta > 1/2$ since $\partial \mathrm{RMP}_1/\partial\theta > 0$ at $\theta = 1/2$ and at $\theta = 1$ and since it is monotonic in θ.

B.1. Generating random tree networks

We use two different algorithms for generating random tree networks. The first one is based on Barabási and Albert (1999). This algorithm starts with a network having M_0 linked nodes. Then, it adds new nodes one by one, up to M_T nodes in total, where M_T is the number of nodes of the network (i.e., the number of countries in the model). Each time a new node is added to the network at iteration t, it is connected to M_{t-1} pre-existing nodes. The probability of being linked to an existing node during iteration t depends on the number of links– degree– of the node in the following way: $p_{it} = \deg(i_{t-1})/[\sum_j \deg(j_{t-1})]$, where p_{it} is the probability of being linked to node i at iteration t, and where $\deg(i_{t-1})$ is the degree of node i at iteration $t-1$. The Barabási and Albert (1999) preferential attachment algorithm tends to create networks with some nodes that have a high degree, who are very well connected, and other nodes with a very low degree, who are badly connected. Put differently, the resulting network tends to have hub-and-spoke characteristics. By setting the initial number of nodes to $M_0 = 2$, and by setting the number of links for new nodes to $m = 1$, we ensure that the resulting network is a connected tree with $M_T - 1$ links.

In the second algorithm we use, new nodes are added to preexisting nodes with equal attachment probability, which means that the probability of being linked to node i at iteration t does not depend on the degree of node i. Formally, we have $p_{it} = 1/M_{t-1}$, where M_{t-1} is the number of nodes in the network when adding the new node at iteration t.

Observe that the average degree of the tree network is equal to $2(M_T - 1)/M_T$, independently of the algorithm used to generate it. The reason is that in an undirected graph, the degree sum formula is $\sum_j \deg(j) = 2|E|$, where $|E|$ is the number of links in the network. Since in the generated tree networks there are $M_T - 1$ links, the degree sum formula becomes $2(M_T - 1)$. Then, the average degree of the network, defined as the degree sum over the number of nodes in the network, is equal to $2(M_T - 1)/M_T$.

Observe further that the standard deviation of the degree of the nodes in the network will usually be higher in networks using the Barabási and Albert (1999) algorithm than in totally random tree networks. The reason is that this algorithm tends to generate a few nodes with a high degree, and a lot of nodes with a very low degree.

Last, when generating random links in the networks we follow Behrens et al. (2007) and assume that the freeness of trade, ϕ_{ij}, between adjacent nodes i and j is given by 1/5. Hence, the freeness of trade between two nodes i and k, linked by a path $\mathscr{P} = \{i, j_1, j_2, \ldots, j_{n-1}, k\}$ of length n, is given by

$$\phi_{ik} = \prod_{(j,l)\in\mathscr{P}} \phi_{jl}. \tag{B-1}$$

We use only shortest paths in the network, which are computed using the Floyd-Warshall algorithm. Because we work with trees, the shortest path is uniquely determined.

B.2. Details on the numerical implementation

We first use the algorithms described in Appendix B.1 to generate random networks. In all cases, we compute the equilibria of the two models for the *same* set of networks. Hence, the results are directly comparable across models. For computational reasons, we generate random networks with between 20 and 30 nodes, the number of nodes being itself random (and drawn from a uniform distribution). Larger networks require too long to solve in the case with a homogeneous good.

To solve the model, we transform the spatial equilibrium conditions (16) into complementary slackness conditions as follows:

$$[\mathrm{RMP}_i(\mathbf{n}) - 1]n_i = 0, \quad i = 1, 2, \ldots, M, \tag{B-2}$$

where we make explicit the dependence of the real market potential on the whole distribution of firms $\mathbf{n} = (n_1, n_2, \ldots, n_M)$.

Model 1: One differentiated sector and one homogeneous sector. We add as nonlinear inequality constraints the equilibrium conditions (13) in the homogeneous good market, the labor market clearing conditions (14), and the complementary slackness conditions (15) for exports of the homogeneous good:

$$\begin{aligned}
\frac{(1-\mu)w_i\theta_i}{\min_k\{w_k\xi\tau_{ki}\}} - \left(\widetilde{X}_{ii} + \sum_{j\neq i}\widetilde{X}_{ji}\right) &= 0, \quad \forall i \\
\theta - n_i\mu - \left(\tau_{ii}\widetilde{X}_{ii} + \sum_{j\neq i}\xi\tau_{ij}\widetilde{X}_{ij}\right) &= 0, \quad \forall i \\
\widetilde{X}_{ij}\left[w_i\xi\tau_{ij} - \min_k\{w_k\xi\tau_{kj}\}\right] &= 0, \quad \forall i
\end{aligned}$$

Furthermore, the following bounds for the variables are imposed: $w_i > 0$ for all i and $\widetilde{X}_{ji} \geq 0$ for all i and j. We also have the constraints that $n_i \geq 0$ for all i. Note that the presence of the min function, which is not differentiable, makes it more difficult to solve the problem. To overcome this problem, we replace all occurrences of the min function with a new variable, z_i. To make sure that this new variable z_i will be equal to the minimum, we substract it from the objective function (i.e., it works as a penalty). Thus, the solver will maximize it. We add the constraint that it should not exceed the delivered price of the homogeneous good: $z_i \leq w_j\xi\tau_{ji}$, $\forall i, j$. In doing so, we make sure that – in the final iteration – z_i is equal to the minimum delivered price of the good.

We transform (B-2) into an equivalent problem that consists in minimizing the sum of squared residuals subject to the set of equilibrium constraints. The numerical implementation of the minimization problem – when substituting out the min operator – is as follows:

$$
(\mathscr{P}_1)\begin{cases}
\min\limits_{\mathbf{n},\mathbf{w},\widetilde{\mathbf{X}}} \sum_{i=1}^{M}\{[\mathrm{RMP}_i(\mathbf{n})-1]n_i\}^2-\sum_{i=1}^{M} z_i \\
\mathrm{RMP}_i(\mathbf{n})\le 1, \quad i=1,2,\ldots,M \\
\dfrac{(1-\mu)w_i\theta_i}{z_i}-\left(\widetilde{X}_{ii}+\sum_{j\neq i}\widetilde{X}_{ji}\right) \qquad = 0, \quad \forall i \\
\theta_i-n_i\mu-\left(\tau_{ii}\widetilde{X}_{ii}+\sum_{j\neq i}\xi\tau_{ij}\widetilde{X}_{ij}\right) \qquad = 0, \quad \forall i \\
z_i\le w_j\xi\tau_{ji}, \quad \forall i,j \\
\theta_i-n_i\mu\ge 0, \quad \forall i \\
\widetilde{X}_{ij}\left[w_i\xi\tau_{ij}-z_j\right] \qquad = 0, \quad \forall i \\
n_i\ge 0, \quad \forall i, \quad w_i>0, \quad \forall i \\
\widetilde{X}_{ji}\ge 0, \quad \forall i,j
\end{cases} \tag{B-3}
$$

As starting values for the solver, we use the population share, θ_i, for the mass of firms, i.e., $n_i^0=\theta_i$. For the wages, we use $w_i^0=1$ for all i. Last, we start with zeros for trade in the homogeneous good, $\widetilde{X}_{ki}^0=0$, and $\widetilde{X}_{ii}^0=(1-\mu)\theta_i$ for the domestic supply of the homogeneous good to the local market.

Model 2: Two differentiated sectors. For the model with two differentiated sectors, we minimize the sum of the squared residuals of the two complementary slackness conditions of the real market potential for each sector:

$$[\mathrm{RMP}_{si}(\mathbf{n}_s,\mathbf{w})-1]n_{si}=0, \quad i=1,2,\ldots,M, \quad s=1,2.$$

The minimization problem is similar to the one in the case with a homogeneous good, but with two real market potential functions with the number of firms, n_{si}, in each sector, the inclusion of the wages, and the constraint on the number of firms and the population shares:

$$
(\mathscr{P}_2)\begin{cases}
\min\limits_{\mathbf{n}_1,\mathbf{n}_2,\mathbf{w}} \sum_{i=1}^{M}\{[\mathrm{RMP}_{1i}(\mathbf{n}_1,\mathbf{w})-1]n_{1i}\}^2 \\
+\sum_{i=1}^{M}\{[\mathrm{RMP}_{2i}(\mathbf{n}_2,\mathbf{w})-1]n_{2i}\}^2 \\
\mathrm{RMP}_{1i}(\mathbf{n}_1,\mathbf{w})\le 1, \quad \forall i \\
\mathrm{RMP}_{2i}(\mathbf{n}_2,\mathbf{w})\le 1, \quad \forall i \\
\theta_i=n_{1i}+n_{2i}, \quad \forall i \\
n_{1i}>0, \quad \forall i, \quad n_{2i}>0, \quad \forall i, \quad w_i>0, \quad \forall i,
\end{cases} \tag{B-4}
$$

We solve the problems $(\mathscr{P}_1)$ and $(\mathscr{P}_2)$ for their equilibria $\{n_i^*,w_i^*\}$, and $\{n_{1i}^*,n_{2i}^*,w_i^*\}$, respectively. We use the MATLAB function fmincon with the *interior-point* algorithm. The code is available upon request.

C. Additional data tables

Table 12 below summarizes the data underlying the applications in Section 4. Additional references to data sources and the choice of parameter values is also provided in that section.

Table 12
Data for the 25 European Union countries.

Country	Data 2001				Data 2012				μ_1
	Labor %	Closeness Centrality	G.V.A. Industry %	G.V.A. Services %	Labor %	Closeness Centrality	G.V.A. Industry %	G.V.A. Services %	
(AT) Austria	1.769	0.977	2.526	2.094	1.867	0.977	2.736	2.230	0.520
(BE) Belgium	1.979	0.933	2.634	2.726	2.078	0.928	2.697	3.004	0.530
(BG) Bulgaria	1.502	0.617	0.153	0.139	1.541	0.603	0.357	0.258	0.760
(CZ) Czech Republic	2.252	0.991	1.054	0.650	2.246	0.980	1.804	0.992	0.570
(DE) Germany	18.531	1.000	24.207	21.641	18.832	1.000	25.697	19.229	0.550
(DK) Denmark	1.286	0.792	1.719	1.791	1.216	0.782	1.683	1.842	0.570
(EE) Estonia	0.277	0.541	0.072	0.066	0.277	0.551	0.150	0.117	0.670
(EL) Greece	2.022	0.542	1.190	1.556	1.808	0.546	0.798	1.603	0.570
(ES) Spain	8.041	0.524	7.963	6.651	8.275	0.537	7.685	8.148	0.600
(FI) Finland	1.089	0.406	1.826	1.252	1.138	0.398	1.572	1.375	0.520
(FR) France	12.302	0.804	12.999	16.611	12.364	0.793	12.436	16.694	0.550
(HR) Croatia	0.657	0.774	0.255	0.223	0.706	0.774	0.34	0.290	0.700

(continued on next page)

Table 12 (*continued*)

Country	Data 2001				Data 2012				μ_1
	Labor %	Closeness Centrality	G.V.A. Industry %	G.V.A. Services %	Labor %	Closeness Centrality	G.V.A. Industry %	G.V.A. Services %	
(HU) Hungary	1.807	0.900	0.664	0.519	1.716	0.901	0.850	0.612	0.610
(IE) Ireland	0.817	0.485	1.657	1.036	0.824	0.478	1.450	1.292	0.560
(IT) Italy	10.692	0.726	12.815	13.200	10.932	0.719	11.657	12.175	0.620
(LU) Luxembourg	0.089	0.959	0.159	0.275	0.103	0.949	0.161	0.388	0.550
(LV) Latvia	0.442	0.617	0.086	0.092	0.394	0.623	0.155	0.160	0.650
(NL) Netherlands	3.896	0.920	4.308	4.934	3.900	0.912	4.358	5.044	0.510
(PL) Poland	6.637	0.853	2.381	1.975	6.989	0.868	3.789	2.484	0.680
(PT) Portugal	2.403	0.450	1.338	1.316	2.066	0.474	1.084	1.271	0.530
(RO) Romania	5.091	0.659	0.596	0.326	3.952	0.661	1.481	0.764	0.790
(SE) Sweden	2.051	0.550	2.887	2.574	2.074	0.539	3.388	3.040	0.570
(SI) Slovenia	0.430	0.932	0.290	0.203	0.423	0.922	0.334	0.235	0.550
(SK) Slovakia	0.992	0.890	0.307	0.208	1.044	0.898	0.786	0.459	0.670
(UK) United Kingdom	12.945	0.671	15.914	17.939	13.236	0.662	12.551	16.298	0.510
Mean	4.000	0.741	4.000	4.000	4.000	0.739	4.000	4.000	0.596
Std. Dev	4.846	0.191	6.111	6.266	4.941	0.188	5.928	5.763	0.078
Max	18.531	1.000	24.207	21.641	18.832	1.000	25.697	19.229	0.790
Min	0.089	0.406	0.072	0.066	0.103	0.398	0.150	0.117	0.510

Notes: For Model 2, we have $\mu_2 = 1 - \mu_1$ by definition. See Section 4 for additional information on data sources and the choice of parameter values.

Appendix D. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.regsciurbeco.2018.04.005

References

Baldwin, Richard, Forslid, Rikard, Martin, Philippe, Ottaviano, Gianmarco I.P., Robert-Nicoud, Frédéric, 2003. Economic Geography and Public Policy. Princeton University Press, Princeton, NJ.

Barabási, Albert-László, Albert, Réka, 1999. Emergence of scaling in random networks. Science 286 (5439), 509–512.

Barbero, Javier, Zofío, José L., 2016. The multiregional core-periphery model: the role of the spatial topology. Network. Spatial Econ. 16 (2), 469–496.

Behrens, Kristian, Ottaviano, Gianmarco I.P., 2011. General equilibrium trade theory and firm behavior. In: Bernhofen, Daniel, Greenaway, David, Falvey, Rod, Kreikemeier, Udo (Eds.), Palgrave Handbook of International Trade. Palgrave MacMillan, Basingstoke, UK, pp. 119–159.

Behrens, Kristian, Lamorgese, Andrea R., Ottaviano, Gianmarco I.P., Tabuchi, Takatoshi, 2007. Changes in transport and non-transport costs: local vs global impacts in a spatial network. Reg. Sci. Urban Econ. 37 (6), 625–648.

Behrens, Kristian, Lamorgese, Andrea R., Ottaviano, Gianmarco I.P., Tabuchi, Takatoshi, 2009. Beyond the Home Market Effect: market size and specialization in a multi-country world. J. Int. Econ. 79 (2), 259–265.

Bosker, Maarten, Brakman, Steven, Garretsen, Harry, Schramm, Marc, 2010. Adding geography to the new economic geography. J. Econ. Geogr. 10 (6), 793–823.

Broda, Christian, Weinstein, David E., 2006. Globalization and the gains from variety. Q. J. Econ. 121 (2), 541–585.

Davis, Donald R., 1998. The home market, trade, and industrial structure. Am. Econ. Rev. 88 (5), 1264–1276.

Davis, Donald R., Weinstein, David E., 2003. Market access, economic geography and comparative advantage: an empirical test. J. Int. Econ. 59 (1), 1–23.

Diewert, Walter E., 1971. An application of the Shephard duality theorem: a generalized Leontief production function. J. Polit. Econ. 79 (3), 481–507.

Duranton, Gilles, Turner, Matthew A., 2012. Urban growth and transportation. Rev. Econ. Stud. 79 (4), 1407–1440.

Freeman, Linton C., 1978. Centrality in social networks: conceptual clarification. Soc. Network. 1 (3), 215–239.

Hanson, Gordon H., 2005. Market potential, increasing returns and geographic concentration. J. Int. Econ. 67 (1), 1–24.

Hanson, Gordon H., Xiang, Chong, 2004. The Home-Market Effect and bilateral trade patterns. Am. Econ. Rev. 94 (4), 1108–1129.

Head, Keith, Mayer, Thierry, 2004. The empirics of agglomeration and trade. In: Henderson, J. Vernon, Thisse, Jacques-François (Eds.), Handbook of Regional and Urban Economics, vol. 4. North Holland, Amsterdam, pp. 2609–2669.

Head, Keith, Mayer, Thierry, Ries, John, 2002. On the pervasiveness of home market effects. Economica 69 (275), 371–390.

Helpman, Elhanan, Krugman, Paul R., 1985. Market Structure and Foreign Trade. MIT Press, Cambridge, MA.

Krugman, Paul R., 1980. Scale economies, product differentiation and the pattern of trade. Am. Econ. Rev. 70 (5), 950–959.

MFOM, 2007a. Encuesta Permanente de Transporte de Mercancías por Carretera. Año 2007. Ministerio de Fomento, Madrid.

MFOM, 2007b. Observatorio de Costes de Transporte de Mercancías por Carretera. Año 2007. Ministerio de Fomento, Madrid.

Mion, Giordano, 2004. Spatial externalities and empirical analysis: the case of Italy. J. Urban Econ. 56 (1), 97–118.

Ottaviano, Gianmarco I.P., Tabuchi, Takatoshi, Thisse, Jacques-François, 2002. Agglomeration and trade revisited. Int. Econ. Rev. 43 (2), 409–436.

Ottaviano, Gianmarco I.P., Thisse, Jacques-François, 2004. Agglomeration and economic geography. In: Henderson, J.V., Thisse, J.-F. (Eds.), Handbook of Regional and Urban Economics, vol. 4. Elsevier, North-Holland, pp. 2563–2608.

Picard, Pierre M., Zeng, Dao-Zhi, 2005. Agricultural sector and industrial agglomeration. J. Dev. Econ. 77 (1), 75–106.

Rauch, James E., 1999. Networks versus markets in international trade. J. Int. Econ. 48 (1), 7–35.

Spiekermann, Klaus, Wegener, Michael, 2008. Shrinking continent: accessibility, competitiveness, and cohesion. In: Faludi, A. (Ed.), European Spatial Research and Planning. Lincoln Institute of Land Policy, Cambridge, MA, pp. 115–140.

Stelder, Dirk, 2016. Regional accessibility in Europe: the impact of road infrastructure 1957-2012. Reg. Stud. 50 (6), 983–995.

Stelder, Dirk, Groote, Peter, de Bakker, Marien, 2013. Changes in Road Infrastructure and Accessibility in Europe since 1960. Final Report, European Commission, Directorate-General Regional Policy, Policy development Economic and quantitative analysis, 2012. CE.16.BAT.040.

Takahashi, Toshiaki, Takatsuka, Hajime, Zeng, Dao-Zhi, 2013. Spatial inequality, globalization, and footloose capital. Econ. Theor. 53 (1), 213–238.

Takatsuka, Hajime, Zeng, Dao-Zhi, 2012a. Trade liberalization and welfare: differentiated-good versus homogeneous-good markets. J. Jpn. Int. Econ. 26 (3), 308–325.

Takatsuka, Hajime, Zeng, Dao-Zhi, 2012b. Mobile capital and the home market effect. Can. J. Econ. 45 (3), 1062–1082.

Thompson, Gary D., 1988. Choice of flexible functional forms: review and appraisal. J. Agric. Resource Econ. 13 (2), 169–183.

Xie, Feng, Levinson, David, 2008. Topological evolution of surface transportation networks. Comput. Environ. Urban Syst. 33 (3), 211–223.

Yu, Zhihao, 2005. Trade, market size, and industrial structure: revisiting the home-market effect. Can. J. Econ. 38 (1), 255–272.

Zeng, Dao-Zhi, Kikuchi, Toru, 2009. Home market effect and trade costs. Jpn. Econ. Rev. 60 (2), 253–270.

Zeng, Dao-Zhi, Uchikawa, Tomohiro, 2014. Ubiquitous inequality: the home market effect in a multicountry space. J. Math. Econ. 50, 225–233.