Growth, Trade, and the Uneven Rise of Port Regions*

Dávid Krisztián Nagy
CREI, UPF, BSE and CEPR

Woan Foong Wong
U of Oregon, CEPR and NBER

March 2026

Abstract

Can spatially neutral productivity improvements have spatially non-neutral consequences? Across three major sources of countrywide productivity growth——increased internet penetration, vaccination campaigns, and improved access to schools——we document a consistent reallocation of economic activity toward port regions within countries. We rationalize these findings in a spatial model in which productivity growth increases export volumes, which in turn triggers a reduction in import costs through the round-trip effect (Wong, 2022), thus allowing port regions to attract economic activity. This prediction contrasts with standard model with exogenous trade costs in which spatially neutral productivity growth shifts economic activity toward interior regions that are more isolated from global markets. We calibrate our model and simulate it with and without the round-trip effect. Preliminary results suggest that the round-trip effect accounts for 15.6% of relative population growth in port regions and amplifies welfare gains from TFP growth by 24.1% in developing countries.

*Contact: dnagy@crei.cat and wfwong@uoregon.edu. We thank Paula Bustos, Giacomo Ponzetto, Stephen Redding and Jaume Ventura, as well as participants at the CREI Faculty Lunch Seminar, the SED and the UPF Applied Lunch Seminar for helpful comments and suggestions. Eduardo Fagundes de Carvalho, Nathan Jones and Alessandro Palucci provided excellent research assistance.

1 Introduction

In recent decades, regional convergence has largely stalled within countries worldwide (Chatterjee, Giannone and Kuno, 2021). In some cases, countries are experiencing rising spatial inequality for the first time in their history — most notably the United States (Eckert, Ganapati and Walsh, 2025) and India (Chatterjee, Lamba and Rai, 2026). In these contexts, disparities are widening between large, prosperous cities and less developed, sparsely populated regions. Given that mobility across space is subject to frictions, such divergence has important implications for interpersonal inequality and may generate political backlash and conflict (Stiglitz, 2012).

What accounts for these recent increases in within-country spatial inequality? A leading explanation points to uneven productivity growth across regions, particularly driven by rising service-sector productivity in large urban centers (Chen et al., 2023; Fan, Peters and Zilibotti, 2023). However, the developing world has also experienced several forms of productivity growth that are spatially neutral — or, if anything, more pronounced in less dense areas. These include vaccination campaigns (Bloom, Canning and Weston, 2005), improvements in school infrastructure (Duflo, 2001; Khanna, 2023), and expanded internet access (Hjort and Poulsen, 2019). This raises a central question: can productivity improvements that are spatially neutral increase economic activity unevenly across regions? If so, why?

In this paper, we show that spatially neutral productivity improvements can interact with international trade costs in ways that generate spatially uneven outcomes, ultimately increasing regional inequality. The key insight is that regions differ in their proximity to global markets. In particular, regions located near ports face lower trade costs and engage more intensively in international trade. If productivity improvements trigger an economic mechanism that reduces the costs of engaging in trade, then port regions become relatively more attractive for both firms and workers. This, in turn, induces a reallocation of population and economic activity toward these areas.

To establish the empirical relevance of this mechanism, we first present reduced-form evidence showing that the three types of productivity improvements above — expanded internet access, vaccination campaigns, and improvements in education — systematically shift economic activity toward port region, using nighttime lights as a proxy. Building on Hjort and Poulsen (2019), we exploit plausibly exogenous variation in the timing of submarine cable arrivals across Africa and find that, following improved internet connectivity, economic activity increases by more than twice as much in port regions compared to non-port regions. For vaccinations, we use a shift-share design based on global vaccination campaigns and show that their impact on economic activity is significantly larger in regions close to ports. For education, we document a stronger correlation between increases in years of schooling and economic activity in port regions relative to non-port regions. These findings are robust across a wide range of specifications. Notably, the spatial extent of the effects is remarkably consistent across contexts, dissipating at approximately

50 kilometers from the nearest port.

We then develop a theoretical framework to rationalize these empirical findings. The central feature of the model is that international trade costs are endogenous to shipping volumes. In particular, we incorporate the *round-trip effect*, whereby the cost of imports declines with the volume of exports. This mechanism is well documented in maritime container trade (Wong, 2022), which accounts for roughly 80% of global trade (UNCTAD, 2024), and also applies to trucks and air freight. The intuition is straightforward: containerships typically operate on round trips between origin and destination. An increase in exports in one direction requires more vessels, which then return with unused capacity that can be used to transport imports at lower cost. As a result, country-wide productivity improvements that boost exports also reduce import costs—an effect that is most pronounced in port regions. As a consequence, population and economic activity reallocate toward these areas.

We embed the round-trip effect within an otherwise standard model of trade and economic geography featuring three regions: a port region, an interior region within the same country, and the rest of the world. Unlike conventional models, trade costs in our framework depend on trade volumes, making them endogenous and potentially asymmetric. That is, the cost of shipping goods from region A to region B may differ from the cost of shipping in the opposite direction. Such cost asymmetries, typically abstracted from in standard models, have been shown to be quantitatively significant in international trade (Waugh, 2010).

As a benchmark, we consider the case in which trade costs are exogenous. In this setting, we prove that a spatially neutral increase in productivity — one that raises productivity equally in port and interior regions — tends to benefit the interior region more. The intuition is that productivity improvements benefit regions in two ways: directly, as well as indirectly due to spillovers through trade with other regions. This indirect effect is more pronounced in the interior as that region primarily trades with itself and the port region. In contrast, the port region trades relatively more with the rest of the world, implying that it experiences a less pronounced indirect effect. While intuitive, this prediction does not align with our empirical findings.

Next, we turn to the model that features endogenous trade costs due to the round-trip effect. We show that when the round-trip effect is sufficiently strong, it can overturn the result that spatially neutral productivity improvements benefit the interior region. This is because the associated decline in international trade costs disproportionately benefits port regions, allowing them to gain more from productivity improvements. In this way, endogenous trade costs driven by the round-trip mechanism can rationalize our empirical findings.

In the final part of the paper, we conduct a preliminary model quantification to measure how the round-trip effect shapes the spatial distribution of population and welfare. To this end, we split the world into two groups of countries: a developing world that has seen positive income growth, which we further split into port and interior regions; and the rest of the world, which

has seen no income growth. Next, we show that, given data on population and income by region as well as trade costs across regions, the model can be inverted to back out unobserved regional fundamentals. To operationalize this inversion procedure, we choose the values of trade costs to match cross-country data on trade values and volumes, and calibrate the key elasticities of the model to existing estimates in the literature.

Armed with the quantified model, we simulate an improvement in the productivity of developing countries that is of the same magnitude as their income growth in the data. We conduct this simulation in the version of the model with endogenous trade costs due to the round-trip effect, but also in the benchmark version with exogenous trade costs. The difference between these two simulations gives us the contribution of the round-trip effect. Our preliminary results suggest that 15.6% of the population growth in the developing world's port regions (relative to interior regions) is attributable to the round-trip effect. Moreover, the round-trip effect has important welfare implications: by lowering trade costs and thus fostering trade, it can amplify the welfare gains from productivity improvements in developing countries by 24.1%.

Related Literature This paper contributes to the growing literature on endogenous trade costs (Coşar and Demir, 2018; Brancaccio, Kalouptsidi and Papageorgiou, 2020; Heiland et al., 2019; Ganapati, Wong and Ziv, 2024; Economides, 2024; Do et al., 2025; Lee, 2026). Waugh (2010) demonstrates that trade costs must be systematically asymmetric, with poorer countries facing higher export costs than richer ones, to match observed trade and price patterns in the data. Wong (2022) empirically identifies the *round-trip effect* in container shipping and studies its impact on international trade. Chen, Ioannides and Rauch (2025) show that asymmetric trade costs help explain the spatial distribution of cities in ancient Greece. Relative to this literature, we focus on how endogenous transport costs — through the round-trip effect — can shape spatial inequality within countries, particularly in developing countries which often face higher export costs.

As such, our work is also related to the literature studying urbanization in developing countries (Glaeser and Henderson, 2017). The developing world is urbanizing at a breathtaking speed, and classical explanations for this phenomenon emphasize the role of "pull" factors in cities related to agglomeration forces (Chauvin et al., 2017) or natural resource endowments (Gollin, Jedwab and Vollrath, 2016) as well as push factors in rural areas such as climate change (Henderson, Storeygard and Deichmann, 2017) and conflict (Fay and Opal, 2000). Within this vein, Faber (2014) and Coşar and Fajgelbaum (2016) point out the role played by domestic transport improvements and proximity to international ports, respectively. We contribute to this literature by highlighting a novel mechanism through which domestic growth and international trade can interact and shape urbanization in developing countries. Our results suggest that this novel mechanism can explain a significant fraction of the growth of port regions, and therefore

contribute to urbanization in the developing world.

Our work also connects to the field of quantitative spatial economics, which utilizes general equilibrium frameworks to analyze the location of economic agents. We build upon the foundational frameworks of Redding (2016) and Redding and Rossi-Hansberg (2017), which provide tractable methods for evaluating how changes in fundamentals have aggregate and heterogeneous spatial impacts. While canonical models by Allen and Arkolakis (2014) and Desmet, Nagy and Rossi-Hansberg (2018) focus on settings with symmetric trade costs, we extend these models to incorporate trade cost asymmetries in a tractable way and provide analytical and quantitative predictions on how productivity growth can increase spatial inequality, diverging from standard models where spatially neutral growth typically benefits isolated, inland regions.

Finally, we contribute to the literature evaluating the impact of productivity-enhancing interventions in developing countries. Empirical evidence has established the large benefits from fast internet (Hjort and Poulsen, 2019), education (Duflo, 2001; Khanna, 2023), and vaccination campaigns (Bloom, Canning and Weston, 2005). We contribute to this literature by studying how these interventions, which improve the productivity of the country as a whole, can have uneven spatial consequences due to endogenous transport costs and international trade.

2 Motivating Evidence

In this section, we provide reduced-form evidence suggesting that expanded internet access, vaccination campaigns, and improvements in education all shift economic activity toward port regions. We first discuss the data sources used in this analysis: data on night lights which serve as the outcome variable (a proxy of GDP), and data on fast internet connectivity, vaccine coverage, and education. Next, we present the results of the three exercises.

2.1 Data

Night Lights at Ports and Surrounding Regions We construct a panel dataset of night lights at port regions from 1992 until 2013 by linking annual Defense Meteorological Program-Operational Line-Scan System (DMSP/OLS) night lights data to a global set of ports and their surrounding regions within a 250km radius. The geocoded port locations are from the World Port Index and for each port, all night-light raster cells within a 250km radius are identified using the 1992 DMSP/OLS grid. Distances from each cell to its associated port are then discretized into 5km distance bins. Following the literature, the luminosity measure is calculated as the inverse hyperbolic sine (asinh) of nighttime light intensity (Hjort and Poulsen, 2019; Ducruet et al., 2024).[1] Annual luminosity measures are then collapsed to the port–distance bin–year level

[1]For grid-cell i in country c in year t, the luminosity measure is $y_{ict} = \ln\left(\text{light} + \sqrt{\text{light}^2 + 1}\right)$.

with top-coded values excluded.

African Submarine Cable Arrivals Between late 2000s and early 2010s, 12 African countries received submarine cables from Europe which significantly increased the speed and capacity of the terrestrial network.[2] The submarine cables reached land at coastal landing stations, usually one per country along the cable route, typically situated near a major city linked to the national backbone network. Data on the arrival of these cables is obtained from Hjort and Poulsen (2019). To study the relationship between fast internet and economic activity spatially, we merge this country-level submarine cable arrival information with our port-region-night-lights dataset.

Vaccine Coverage Immunization data is obtained from the WHO/UNICEF Joint Reporting Form on Immunization (Bloom, Canning and Weston, 2005). This country-level dataset reports percentage of the target population vaccinated for essential antigens like Measles, Tuberculosis, and Whooping cough (check spelling X). To investigate the relationship between vaccine coverage and spatial activity, we merge this country-antigen-level dataset with our port-region-night-lights dataset. Since this dataset captures coverage on different antigens, we use this variation in the construction of an IV to estimate the impact of vaccines on spatial activity of port regions.

Education Attainment Education data is obtained from the Global Data Lab (GDL) Area Database, which provides harmonized subnational development indicators for low- and middle-income countries (Smits and Permanyer, 2019). We focus on two adult education attainment measures: mean years of education for adults aged 20–39 and 40–59. To examine the relationship between education attainment and spatial activity, we merge these subnational education measures to our port-region-night-lights panel. Note that since the GDL dataset is at the subnational level, we are able to observe within-country variation in education relative to the previous two datasets.

2.2 The Spatially Uneven Effects of Productivity Growth

2.2.1 Fast Internet

Here we examine the relationship between countries that receive fast internet as a result of submarine fiber-optic cables and the resultant changes in their economic activity between port and surrounding regions. We start by estimating the difference-in-difference specification as in Hjort and Poulsen (2019). Next, we extend their analysis by studying how this productivity growth differentially impacted ports and their surrounding regions.

We start by running the following difference-in-difference specification:

[2]The countries are Benin, the Democratic Republic of Congo, Nigeria, Namibia, Senegal, Tanzania, Togo, Ghana, Kenya, Madagascar, Mozambique, and South Africa.

$$y_{ict} = \alpha \, \text{Treatment}_{ict} + \lambda_{ct} + \mu_i + \varepsilon_{ict}, \tag{1}$$

where y_{ict} is the inverse hyperbolic sine of light density for grid cell i in country c and year t, λ_{ct} are country–year fixed effects, and μ_i are grid-cell fixed effects. We employ two measures for Treatment_{ict}: (1) the cumulative number of submarine cables that have arrived and become operational in a country c by year t ($\text{Submarine Cables}_{ict}$), and (2) the first country-level measure interacted with a binary indicator equal to one for grid cells that are close (within 0.5km) to the domestic backbone internet network ($\text{Submarine Cables}_{ict} \times \text{Connected}_{ict}$). Standard errors are clustered at the 0.1° grid level.

The first treatment measure, $\text{Submarine Cables}_{ict}$, captures the intensity of a country-level internet infrastructure boost and applies to all grid cells within that country. The second treatment measure captures how intensively a particular grid cell within the country is able to benefit from fast internet by capturing information on how close this cell is to the country's domestic internet network. We present both treatment measures separately because for the purposes of our analysis, most African countries have just one port per country and so we have very few grid cells that are both close to ports and the domestic backbone network. However, even though we have very little variation in the data, we still find that our point estimates are similar in magnitude and within one standard error of the first treatment measure (albeit noisy). Country–year fixed effects control for aggregate growth in the country over time that could influence economic activity proxied by night lights, while grid-cell fixed effects control for the fixed characteristics of different grid cells such as being closer to coastal or urban areas, which might result in different levels of night lights.

Column (1) in Table 1 reports the coefficient from equation (1) using the first treatment measure. Similar to Hjort and Poulsen (2019), we find a positive and significant relationship between internet connectivity and economic activity as proxied by night lights. Next, we look at how this productivity improvement affected ports compared to interior regions.

To do this, we introduce an binary indicator equal to one for grid cells that are within the surrounding region of ports (*port regions*) to the difference-in-difference specification in equation (2):

$$y_{ict} = \alpha \, \text{Treatment}_{ict} \times \text{Port Region}_{ic} + \lambda_{ct} + \mu_i + \varepsilon_{ict}, \tag{2}$$

where Port Region_{ic} is a binary indicator equal to one if the centroid of grid cell i in country c lies within a certain distance threshold of the nearest port. We employ distance thresholds of 40km, 50km, and 60km.

Column (2) in Table 1 reports the coefficients from equation (2) when restricting the grid cells to be 50km from ports. Here we see that most of the increase in economic activity from fast internet is concentrated in the port region. In fact, the cumulative effect on the port region

Table. 1. Fast Internet and Night Lights: Port vs Interior Regions

	(1)	(2)	(3)	(4)	(5)
Submarine Cables	0.029***	0.049***	0.022*	0.049***	
	(0.006)	(0.014)	(0.012)	(0.014)	
Submarine Cables × Port Region		0.069***	0.082***	0.054***	
		(0.015)	(0.017)	(0.013)	
Submarine Cables × Connected					0.031*
					(0.018)
Submarine Cables × Connected × Port Region					0.063
					(0.051)
Distance Threshold from Port	–	50km	40km	60km	50km
Observations	80360	80360	80360	80360	80360
Mean of Outcome	0.442	0.442	0.442	0.442	0.442

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors clustered by 10x10km grid cells in parentheses, as in Hjort and Poulsen (2019). All specifications include country–year and grid-cell fixed effects. Column (1) reports the coefficients from Equation (1). Columns (2), (3), and (4) report the coefficients from Equation (2) for distance thresholds of 50km, 40km, and 60km respectively from ports. Column (5) include an additional interaction term to Column (2) which is whether the grid cell is close (within 0.5km) of the country's domestic internet network.

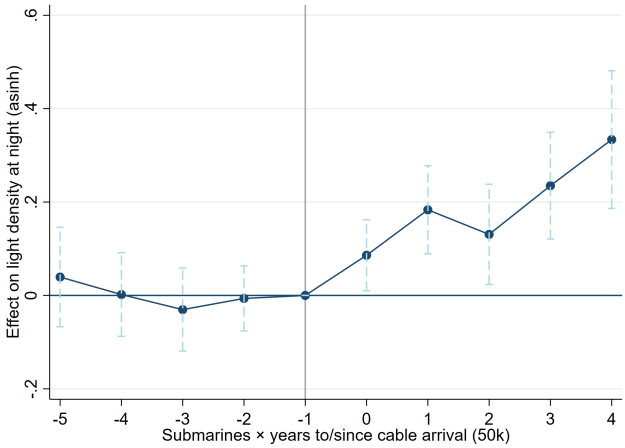
is more than twice the effect in the interior region. When restricting to locations even closer to the ports (40km), we find even larger effects on economic activity (Column (3), Table 1). This effect dissipates the further away locations are from ports: for grid cells 60km away from ports the coefficient decreases (Column (4), Table 1). This decrease continues for larger distances.

As mentioned earlier, we focus on the first treatment measure in Hjort and Poulsen (2019) (country-level treatment measure of when submarine cables arrive in these countries) instead of including whether or not the grid cells are connected to the backbone domestic internet network. This is due to our focus on grid cells near ports versus interior regions. When including the second treatment measure, we only have 23 treated grid cells that are close to a port.[3] This is because while there are 12 African countries that are all coastal, most of their capitals are not coastal (such as Congo, Benin, or Togo) which means that grid cells that are near ports are not necessarily near their domestic internet network. That being said, we include the second treatment measure as a robustness check of our results. Column (5) in Table 1 shows that our port region estimate, albeit noisy, is similar in magnitude and within one standard error of the 50km distance threshold port region estimate in Column (2). This tells us that port regions in fact benefit more from the productivity increase due to fast internet compared to non-port regions.

In order to study the dynamic effects of fast internet on ports compared to interior regions,

[3]Using the first treatment measure, we have 979 treated grid cells.

Figure 1. Event Study Plot of Fast Internet and Night Lights: Port vs Interior Regions



Notes: This figure reports the estimated coefficients from equation (3) for grid cells that are within 50km of the nearest port. These results correspond to the difference-in-difference specification in Column (2) of Table 1.

we estimate the following event-study regression:

$$y_{ict} = \sum_{\substack{\tau=-5 \\ \tau\neq-1}}^{+4} \beta_\tau \left(\text{Treatment}_{ict} \times \text{Port Region}_{ic} \times \mathbf{1}[\text{TimeSince}_t = \tau]\right) + \lambda_{ct} + \mu_i + \varepsilon_{ict}, \tag{3}$$

where $\text{TimeSince}_\tau = t - \text{FirstYear}_c$ is the number of years elapsed since the first submarine cable became operational in country c.[4] Period -1 ($\tau = -1$), corresponding to the year of cable arrival, is omitted and serves as the reference category.

Figure 1 shows the coefficients from equation (3). First, prior to the introduction of the submarine cables, the economic activity at port regions are not statistically different from economic activity in the non-port interior regions. After the arrival of the cables, there is a gradual and significant increase in the economic activity at port regions which continues 4 years after. This shows that spatially neutral productivity improvements such as fast internet have the potential to not only benefit port regions relative to the interior but also to increase the spatial inequality between these regions over time.

2.2.2 Vaccine Coverage

Next, we show that our findings are robust globally and to other settings beyond fast internet in African countries. To study the link between vaccine coverage and economic activity, we estimate the following specification for a global set of 138 countries:

$$y_{ict} = \sum_{d\in\{5,10,\ldots,145\}} \theta_d \left(\text{Treatment}_{ct} \times \mathbf{1}[\text{DistBin}_i = d]\right) + \lambda_{ct} + \mu_i + \varepsilon_{ict} \tag{4}$$

[4]The variable FirstYear_c equals 2009 for Kenya, Mozambique, South Africa, and Tanzania; 2010 for Benin, Ghana, Madagascar, Nigeria, Senegal, and Togo; and 2012 for the Democratic Republic of Congo and Namibia.

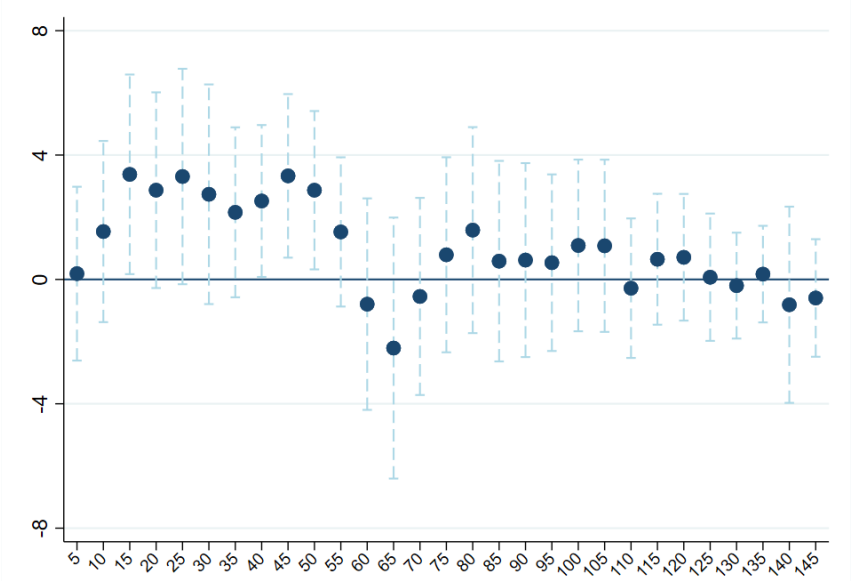
where y_{ict} denotes the outcome for grid cell i in country c and year t, measured as the inverse hyperbolic sine of nighttime light intensity. Treatment_{ct} is a country–year level treatment variable capturing vaccine coverage and is common to all grid cells within country c in year t. Distance to the nearest port is discretized into 5–kilometer bins, and $\mathbf{1}[\text{DistBin}_i = d]$ is an indicator equal to one if grid cell i lies in distance bin d. The coefficients θ_d capture heterogeneous treatment effects by distance to ports: each coefficient measures the differential impact of the country-level treatment for grid cells located d kilometers from the nearest port, relative to the omitted category of grid cells located 150 kilometers from the nearest port.

To address concerns that the observed vaccine coverage may be endogenous to local economic conditions and unobserved shocks correlated with nighttime luminosity, we construct a shift–share IV that exploits differential exposure across locations to global changes in vaccine coverage over time. For each antigen a, we first compute the ratio of global coverage in year t relative to its 1992 level, $g_{at}/g_{a,1992}$, dropping antigens without observed coverage in 1992. We then interact this global growth factor with the local baseline exposure to each antigen, measured by coverage in 1992. The predicted coverage is thus constructed as

$$\widehat{\text{Coverage}}_{ct} = \sum_a \text{Coverage}_{ca,1992} \times \frac{g_{at}}{g_{a,1992}} \tag{5}$$

where $\text{Coverage}_{ca,1992}$ denotes baseline coverage for antigen a in country c in 1992. The identification assumption is that changes in global coverage at the antigen level are exogenous to economic conditions in country c.

Figure 2. Vaccines and Night Lights: Distance from the Nearest Port



Notes: This figure shows the estimated IV coefficients from equation (4) using the IV specified in equation (5).

Figure 2 reports the 2SLS IV coefficients from equation (4). Here we see again that increased

vaccine coverage increases economic activity for distance bins that are close to ports. As we consider grid cells that are farther and farther away from ports, this effect decreases and becomes noisy. In fact, similar to the results in Table 1, we see that the effect from this productivity increase dissipates around 50km from the nearest port.

2.2.3 Education Attainment

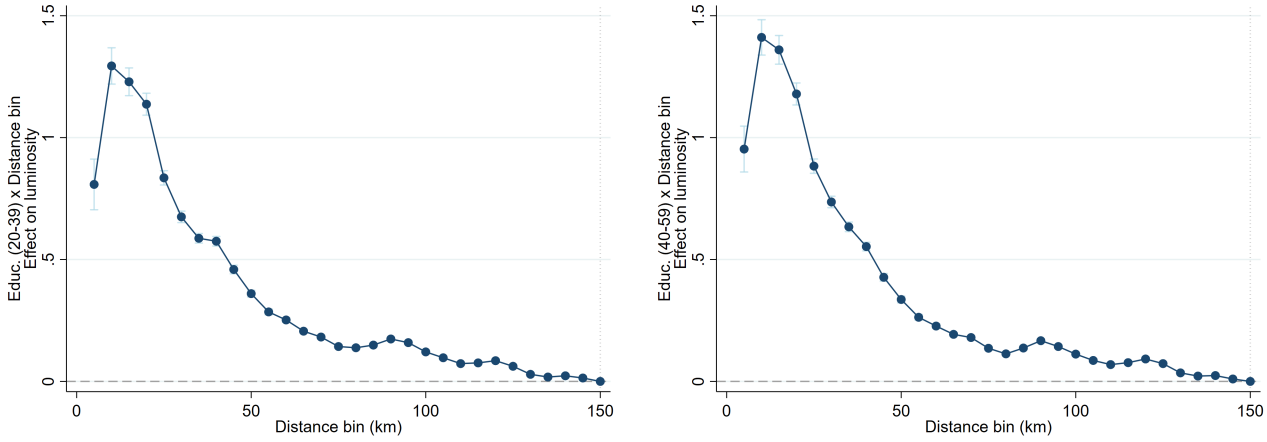
Thirdly, to study the link between education attainment and economic activity, we estimate a similar specification to equation (4),

$$y_{ict} = \sum_{d \in \{5,10,\ldots,145\}} \theta_d \left(\text{Treatment}_{rct} \times \mathbf{1}[\text{DistBin}_i = d] \right) + \lambda_{ct} + \mu_i + \varepsilon_{ict} \tag{6}$$

with one change: as the education attainment variable at the level of subnational regions r, our treatment variable is allowed to vary at that level.

Figure 3 shows that while there is a generally positive association between education attainment and economic activity, this correlation is the strongest for locations closest to ports and dissipates farther away. Yet again, the effect starts to converge near zero around 50km from the port. This result is robust to using both education attainment levels for 20-39 year-olds and for 40-59 year-olds.

Figure 3. Education and Night Lights: Distance from the Nearest Port



(a) Population Ages 20-39 **(b)** Population Ages 40-59

Notes: This figure shows the estimated coefficients from equation (6).

3 Theory

We now develop a theory that can rationalize the empirical results of the previous section. The most important feature of our theory is endogenous trade costs that negatively depend

on the quantity of goods shipped in the opposite direction: the *round-trip effect*. We show that this endogeneity of trade costs is crucial to explain the reallocation of population and economic activity toward port regions in response to spatially neutral productivity improvements, as the effect of such productivity improvements operates in the opposite direction in a standard framework with exogenous trade costs. Finally, we also show that ignoring the above endogeneity and asymmetry of trade costs leads to systematically biased estimates of model fundamentals if one quantifies these fundamentals in the standard way, i.e., using a model inversion.

3.1 Setup

The world consists of three regions that potentially differ in their amenity and productivity levels, land supply, trade costs, and mobility costs. All trade and mobility costs take the 'iceberg' form, as standard in quantitative theories of trade and economic geography. Two of the three regions are in the same country, while the third one represents the rest of the world (ROW). The two regions within the country differ in their geographic location relative to the ROW. One of them, the *port region*, is directly connected to the ROW and faces an export cost of τ_1^X and an import cost of τ_1^M, both of which we endogenize later. The other region of the country, the *interior region*, is only connected to the ROW through the port region, with which it can trade at an exogenous and symmetric cost of τ_0. The interior region's export and import costs with the ROW are thus $\tau_X^1 \tau_0$ and $\tau_M^1 \tau_0$, respectively. Figure 4 illustrates the geography of this three-region world.

An exogenous mass of $\bar{L}$ agents inhabit the world, choose which region they want to live in, and supply one unit of labor and consume goods in the region they chose. We now describe their utility maximization problem, followed by a description of production technologies, shipping technologies, and the definition of the equilibrium.

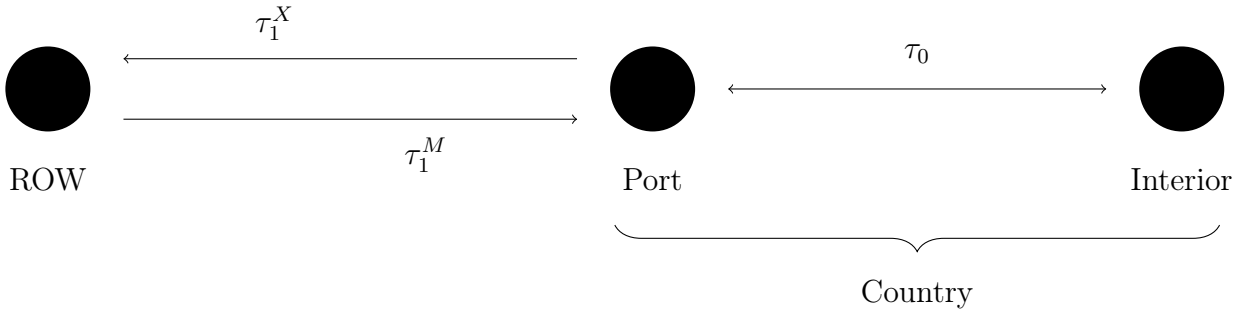
Agents' utility maximization. The utility of an agent who chooses to live in region r is given by

$$u(r) = a(r) \cdot \left[\int_0^1 c_\omega(r)^\rho \, d\omega \right]^{\frac{1}{\rho}} \cdot m(r)^{-1} \tag{7}$$

where $a(r)$ represents the amenity level of region r, $c_\omega(r)$ denotes the agent's consumption level of good ω out of the continuum of goods $\omega \in [0,1]$, and $m(r)$ captures mobility frictions in the form of the cost of 'entering' region r. Such a formulation can be microfounded with bilateral migration costs that enter utility in an iceberg form and are multiplicatively separable in 'entry' and 'exit' costs, as in Desmet, Nagy and Rossi-Hansberg (2018). This formulation is able to capture the key idea that certain regions receive few people despite their high amenity-adjusted real income, as they have a high entry cost $m(r)$.

Agents choose their region of residence and consumption levels to maximize utility (equation (7)) subject to a standard budget constraint, taking their income $y(r)$, goods' prices, amenities

Figure 4. Geography of the world



and mobility costs as given.

Firms' production technology. Goods can be produced in any region by competitive firms. The production function of the representative firm producing good ω in region r is

$$q_\omega(r) = \mu^{-\mu} \cdot (1-\mu)^{-(1-\mu)} \cdot z_\omega(r) \cdot L_\omega(r)^\mu \cdot H_\omega(r)^{1-\mu}$$

where $z_\omega(r)$ denotes the TFP of good ω in region r, and $L_\omega(r)$ and $H_\omega(r)$ are the quantities of labor and land used in production, respectively. Labor is supplied by local agents at a competitive wage $w(r)$. Land is available in fixed positive supply $H(r) > 0$ and is rented out to firms in a competitive market at rent per unit of land $R(r)$. We make the standard assumption that land rents are redistributed to local agents with equal shares, implying that an agent's income in region r is given by

$$y(r) = w(r) + \frac{R(r) H(r)}{L(r)} \tag{8}$$

where $L(r)$ denotes the population of region r.

Following a long tradition starting with Eaton and Kortum (2002), we assume that goods' TFP levels are drawn from a Fréchet distribution

$$\Pr[z_\omega(r) \leq z] = e^{-A(r)^\theta \cdot z^{-\theta}}$$

where $A(r)$ is the average productivity of region r, and θ is a parameter that captures the dispersion of productivity draws and becomes the *trade elasticity*, i.e., the elasticity of trade flows to trade costs, as in Eaton and Kortum (2002). To make the model solvable in closed form and derive analytical predictions, we impose the additional restriction that $\theta = \mu/(1-\mu)$. Existing estimates of the land share around 0.2 (Desmet, Nagy and Rossi-Hansberg, 2018) and the trade elasticity around 4 (Simonovska and Waugh, 2014) support this restriction. Nevertheless, we relax the restriction in Section 3.3.

Shipping technology. A core ingredient of our theory is that trade costs between the port region and the ROW are endogenous. We model this endogeneity by assuming that the export and import costs depend on quantities shipped in both directions:

$$\begin{aligned} \tau_1^X &= \bar{\tau}_1^X \cdot q_X^{\gamma_1} \cdot q_M^{-\gamma_2} \\ \tau_1^M &= \bar{\tau}_1^M \cdot q_M^{\gamma_1} \cdot q_X^{-\gamma_2} \end{aligned} \tag{9}$$

where $\bar{\tau}_1^X$ and $\bar{\tau}_1^M$ represent exogenous components of export and import costs, respectively, and q_X and q_M denote the total quantity of goods exported to and imported from the ROW, respectively. The parameter $\gamma_1 \geq 0$ captures the potential *congestion effect* of shipping a large quantity in the own direction. By contrast, the parameter $\gamma_2 \geq 0$ captures the *round-trip effect*: if a large quantity is shipped in the opposite direction, trade costs go down in the own direction (Wong, 2022).[5] Note that this formulation nests exogenous trade costs as a special case in which $\gamma_1 = \gamma_2 = 0$. It also nests the often-made assumption of exogenous *and* symmetric trade costs under the additional restriction that $\bar{\tau}_1^X = \bar{\tau}_1^M$.

Equilibrium. For given values of structural parameters $\{\rho, \theta, \tau_0, \bar{\tau}_1^X, \bar{\tau}_1^M, \gamma_1, \gamma_2\}$ and a given distribution of fundamentals $\{a(\cdot), m(\cdot), A(\cdot), H(\cdot)\}$, the equilibrium of the model is a distribution of agents $L(\cdot)$, wages $w(\cdot)$, land rents $R(\cdot)$, goods' prices and quantities such that

1. agents choose their region of residence and consumption levels to maximize their utility (7), subject to their income given by equation (8),
2. firms choose their output, labor and land use to maximize their profits,
3. agents source goods from the cheapest source,
4. markets for goods, labor and land clear,
5. export and import costs are determined by equation (9).

In Appendix A, we provide a formal definition of the equilibrium and show how the equilibrium conditions can be combined to solve for equilibrium export costs, import costs, population by region, and income by region.

3.2 The effects of spatially neutral growth in country productivity

We now use the theory to study the effects of a spatially neutral productivity improvement that the country experiences. We introduce this productivity improvement by breaking up the

[5] In principle, the round-trip effect can be strong enough to lead to a cycle of increasing shipping quantities triggering shipping cost reductions, which lead to further increases in shipping quantities, etc., a process ultimately resulting in trade costs collapsing to zero. A regularity condition $\gamma_2 < \gamma_1 + \frac{1}{1+\theta}$ rules out this possibility. This regularity condition is satisfied in our model calibration.

productivity of a region r of the country (i.e., the port region or the interior region) into a common 'country productivity' $\bar{A}$ and a region-specific shifter $\hat{A}(r)$:

$$A(r) = \bar{A} \cdot \hat{A}(r)$$

so that a spatially neutral country-wide productivity improvement can be modeled as an increase in $\bar{A}$. Importantly, we keep everything else fixed in this exercise, including the productivity of the ROW. Our results generalize to a setting in which the ROW also experiences an improvement in its productivity, but with a smaller percentage change than the change in $\bar{A}$.

The following proposition states that the country-wide productivity improvement increases the total quantity exported relative to the total quantity imported, which, in turn, triggers an increase in export costs relative to import costs as long as γ_1 and γ_2 are not both equal to zero.

Proposition 1. *An increase in country productivity $\bar{A}$ leads to*

1. *an increase in relative export quantities q_X/q_M,*
2. *an increase in relative export costs τ_1^X/τ_1^M if and only if $\gamma_1 + \gamma_2 > 0$.*

Proof. See Appendix B. □

The increase in relative exports q_X/q_M is always present in the model, even in the special case of exogenous trade costs ($\gamma_1 = \gamma_2 = 0$). It is a consequence of balanced trade: as $\bar{A}$ increases, the country sees a deterioration in its terms of trade (i.e., its average export price falls relative to its average import price), and the only way to keep trade balanced in value terms is that relative quantities move in the opposite direction.

In the case of exogenous trade costs ($\gamma_1 = \gamma_2 = 0$), this movement in relative quantities obviously does not alter τ_1^X/τ_1^M. By contrast, whenever the congestion effect, the round-trip effect, or both are present, the change in relative quantities makes exports more expensive relative to imports, as can be seen from dividing the two equations in (9):

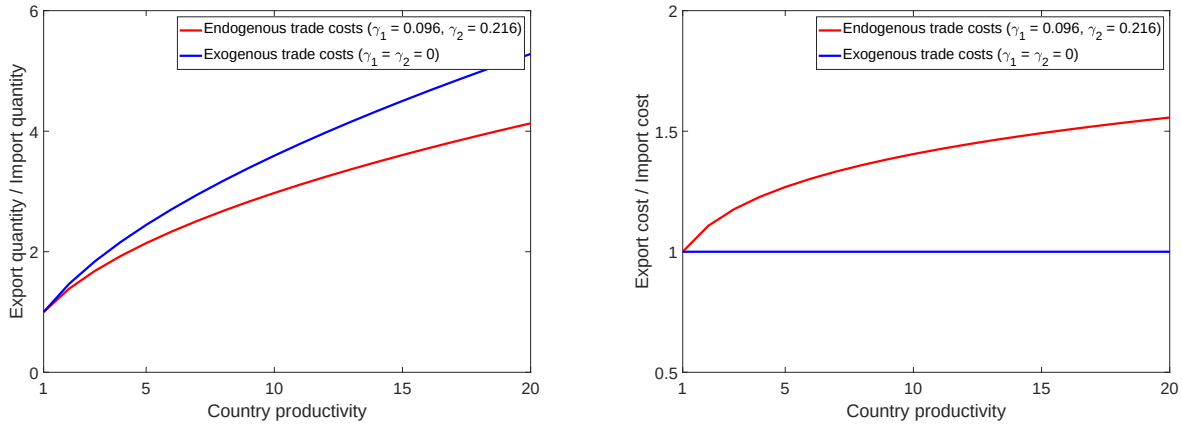
$$\frac{\tau_1^X}{\tau_1^M} = \frac{\bar{\tau}_1^X}{\bar{\tau}_1^M} \left(\frac{q_X}{q_M} \right)^{\gamma_1+\gamma_2}$$

Figure 5 illustrates the results of Proposition 1 by plotting τ_1^X/τ_1^M and q_X/q_M against country productivity $\bar{A}$ in two model calibrations: one with endogenous trade costs ($\gamma_1 = 0.096$ and $\gamma_2 = 0.216$[6], red line) and one with exogenous trade costs ($\gamma_1 = \gamma_2 = 0$, blue line).

How do these changes affect the spatial distribution of population and economic activity within the country? The following proposition shows that, in the case of exogenous trade costs, the improvement in country productivity triggers reallocation of people and income *away* from the port region.

[6] These are the values of γ_1 and γ_2 that we use later in the model calibration.

Figure 5. Export to import quantities and costs as a function of country productivity



Notes: The left figure displays the ratio of export and import quantities q_X/q_M as a function of $\bar{A}$ under endogenous trade costs (red, $\gamma_1 = 0.096, \gamma_2 = 0.216$) vs exogenous trade costs (blue, $\gamma_1 = \gamma_2 = 0$). The right figure displays the ratio of export to import costs τ_1^X/τ_1^M as a function of country productivity $\bar{A}$ under endogenous trade costs (red, $\gamma_1 = 0.096, \gamma_2 = 0.216$) as well as exogenous trade costs (blue, $\gamma_1 = \gamma_2 = 0$). The remaining model parameters are set to: $\theta = 4$ and $\bar{\tau}_1^X = \bar{\tau}_1^M = 1$. The values of fundamentals are set to: $a(r) = m(r) = H(r) = 1$ for all r, $\hat{A}_p = \hat{A}_i = 1$ and $A_{ROW} = 1$.

Proposition 2. *In the case of exogenous trade costs ($\gamma_1 = \gamma_2 = 0$), an increase in country productivity decreases*

1. *the population share of the port region* $\frac{L_p}{L_p+L_i}$,
2. *the income share of the port region* $\frac{L_p y_p}{L_p y_p+L_i y_i}$.

Proof. See Appendix B. □

Proposition 2 stems from the fact that the improvement in country productivity has two effects on a region. One is a direct effect, which is identical between the port region and the interior. The other is an indirect effect, which operates through changes in the region's market access. Market access is increasing in trading partners' productivity as proximity to higher-productivity regions offers better economic opportunities. As the interior region primarily trades with itself and the port region, it experiences a large indirect effect through its market access. By contrast, the port region trades a lot with the ROW, which does not see its productivity improve. As a consequence, its market access does not grow as much as the interior's. All in all, a model with exogenous trade costs predicts a spatially non-neutral effect of increasing $\bar{A}$ that favors the interior region, in contrast to the empirical results of Section 2.

The endogenous response of trade costs due to the round-trip effect has the potential to flip the theoretical predictions of Proposition 2. Whether it does or not depends on the strength of the round-trip effect as well as on its complex interplay with the model's other forces. The following proposition illustrates this in a simple special case of the model in which (i) the port and

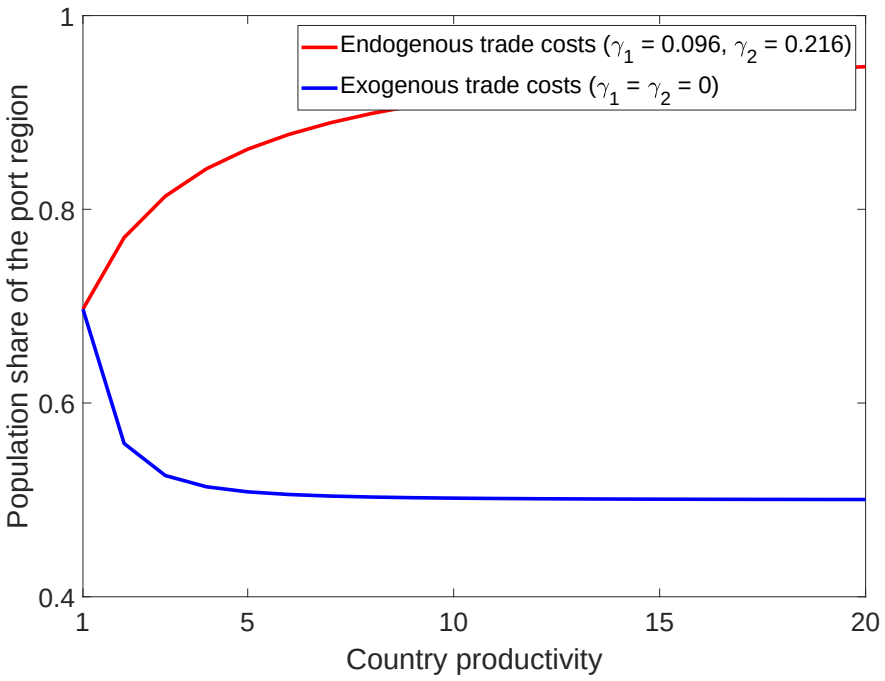
interior regions are symmetric in their fundamentals, (ii) the congestion effect is not too strong, i.e., γ_1 is small, and (iii) the round-trip effect is sufficiently strong, i.e., γ_2 is large enough.

Proposition 3. *Consider the case in which the port and interior regions are identical in their amenities, mobility costs, land supply, and region-specific productivity shifters. If* $\gamma_1 < \frac{\theta}{1+\theta}\gamma_2$ *and* $\gamma_2 > \frac{\sqrt{(1+2\theta)^2+4(1+\theta)\left[(1+\theta)\gamma_1^2+2(1+\theta)\gamma_1+1\right]}-(1+2\theta)}{2(1+\theta)}$*, then an increase in country productivity increases*

1. *the population share of the port region* $\frac{L_p}{L_p+L_i}$,
2. *the income share of the port region* $\frac{L_p y_p}{L_p y_p+L_i y_i}$.

Proof. See Appendix B. □

Figure 6. The port region's population share as a function of country productivity



Notes: Notes: This figure displays the population share of the port region $\frac{L_p}{L_p+L_i}$ as a function of country productivity $\bar{A}$ under endogenous trade costs (red, $\gamma_1 = 0.096, \gamma_2 = 0.216$) as well as exogenous trade costs (blue, $\gamma_1 = \gamma_2 = 0$). The remaining model parameters are set to: $\theta = 4$ and $\bar{\tau}_1^X = \bar{\tau}_1^M = 1$. The values of fundamentals are set to: $a(r) = m(r) = H(r) = 1$ for all r, $\hat{A}_p = \hat{A}_i = 1$ and $A_{ROW} = 1$.

The intuition for Proposition 3 is as follows. Whenever the round-trip effect is strong and the congestion effect is weak, an increase in export quantities triggers a decrease in import costs, which triggers an increase in import quantities through demand, which in turn lowers export costs through the round-trip effect, and so on, ultimately leading to lower export *and* import costs in the new equilibrium. All else fixed, these lower export and import costs make the port

region a more attractive place for people and economic activity. This force that operates through endogenous trade costs can be strong enough to dominate the indirect effect of increasing country productivity through market access, which favors the interior. Thus, the port region can gain more from a spatially neutral productivity improvement, in line with the empirical results of Section 2. Figure 6 illustrates these findings in the same model calibrations as in Figure 5: a calibration with endogenous trade costs ($\gamma_1 = 0.096$ and $\gamma_2 = 0.216$, red line) and one with exogenous trade costs ($\gamma_1 = \gamma_2 = 0$, blue line).

3.3 Implications for the quantification of model fundamentals

The researcher typically does not observe the values of fundamentals such as amenities $a(\cdot)$, mobility costs $m(\cdot)$ and productivities $A(\cdot)$. A standard procedure to quantify these fundamentals is a *model inversion*, i.e., a process in which the structure of the model is used to recover the values of fundamentals that rationalize observed data as an equilibrium (Redding and Rossi-Hansberg, 2017). The following theorem shows that this is also feasible in our model.

Theorem 1. *Given the values of structural parameters as well as data on population $L(\cdot)$, income $y(\cdot)$ and the supply of land $H(\cdot)$ by region as well as bilateral trade costs $\tau(\cdot,\cdot)$ across regions, the model can be inverted to back out (up to scale) unique distributions of amenities relative to mobility costs $a(\cdot)/m(\cdot)$ and productivities $A(\cdot)$ that rationalize the data as an equilibrium.*

Proof. See Appendix B. □

Two remarks are in order regarding Theorem 1. First, amenities and mobility costs cannot be separately identified as higher amenities in region r have observationally identical effects as lower immigration costs to region r: both increase population $L(r)$ through attracting people to r and decrease nominal income through compensating differentials. To tell apart $a(r)$ and $m(r)$, one would need to observe one of the two or rely on subjective well-being data, as in Desmet, Nagy and Rossi-Hansberg (2018). Importantly, however, the model exercises we conduct later in this paper do not require us to separate $a(r)$ from $m(r)$.

Second, and similar to the inversion procedures of standard models, the researcher needs to observe bilateral trade costs to implement Theorem 1. Trade costs, however, are typically not observable. In standard models, this is not a major concern as trade costs are assumed to be exogenous and symmetric, in which case they can be recovered from observed bilateral trade flows $X(\cdot,\cdot)$ in a large class of models (Head and Ries, 2001):

$$\tau(r,s) = \tau(s,r) = \left[\frac{X(r,s)}{X(s,s)}\frac{X(s,r)}{X(r,r)}\right]^{-\frac{1}{2\theta}}$$

However, the round-trip and congestion effects break the exogeneity and symmetry of trade costs. In such cases, the above Head and Ries formula identifies the geometric average of actual trade

costs:

$$\left[\frac{X(r,s)}{X(s,s)}\frac{X(s,r)}{X(r,r)}\right]^{-\frac{1}{2\theta}} = \left[\tau(r,s)\,\tau(s,r)\right]^{1/2}$$

The next theorem presents how abstracting from the endogeneity and asymmetry of trade costs in this way biases the recovered values of model fundamentals.

Theorem 2. *Assume that the researcher thinks that trade costs are exogenous, symmetric and given by the Head and Ries index:*

$$\dot{\tau}_1^X = \dot{\tau}_1^M = \left[\tau_1^X \tau_1^M\right]^{1/2}.$$

Then the model inversion procedure of Theorem 1 falsely identifies fundamentals as

$$\frac{\dot{a}(r)}{\dot{m}(r)} = \frac{a(r)}{m(r)}\left[\tau_1^X \tau_0\right]^{1/2} \quad \textit{and} \quad \dot{A}(r) = A(r)\left[\tau_1^X \tau_0\right]^{-1/2}$$

in the port and interior regions of the country, and as

$$\frac{\dot{a}_{ROW}}{\dot{m}_{ROW}} = \frac{a_{ROW}}{m_{ROW}}\left[\tau_1^M \tau_0\right]^{1/2} \quad \textit{and} \quad \dot{A}_{ROW} = A_{ROW}\left[\tau_1^M \tau_0\right]^{-1/2}$$

in the ROW, where the variables without a dot denote the true values of fundamentals.

Proof. See Appendix B. □

Theorem 2 shows that the bias in model fundamentals caused by ignoring the asymmetry of trade costs is linked to the cost of shipping goods out of the region in consideration. From the perspective of the port and interior regions, this cost of shipping out is the export cost, τ_1^X; from the perspective of the ROW, this cost is the import cost, τ_1^M.

Consider a world in which $\tau_1^X > \tau_1^M$ in equilibrium. In this case, the recovered productivity levels in the port and interior regions are systematically biased downwards, while the recovered productivity of the ROW is biased upwards. The intuition for this result is that high export costs make the country suffer in its competitiveness on the world stage, leading to low nominal income in the country. If we ignore the disproportionately high export cost, the model mistakes this low nominal income for a low productivity level. As poor countries tend to feature a high ratio of export to import costs (Waugh, 2010), their productivity levels might be systematically underestimated in analyses that do not explicitly account for this cost differential. In the case of amenities relative to mobility costs, the bias works in the opposite direction. Low income in the country, if not compensated by disproportionately low import prices, must imply a high level of amenities through the logic of compensating differentials. Thus, poor countries' amenity levels might be systematically overestimated in analyses that ignore the difference between export and import costs.

In Appendix C, we show that Theorems 1 and 2 generalize to an extended model (i) in which we do not impose the restriction $\theta = \mu/(1-\mu)$, and (ii) we allow for agglomeration and congestion externalities, as well as agents having Fréchet-distributed taste draws for regions. Finally, we show in Appendix D that the results of Theorems 1 and 2 further generalize to a model with an arbitrary number of regions as long as these regions constitute a *tree*, i.e., a network without cycles. This is useful for other potential applications of the framework in which the three-region assumption is not a sufficiently good approximation to reality.

4 Quantification (preliminary)

We now take the model to recent worldwide data on population, export and import quantities and values, and night lights. One key pattern that emerges from the data is that population growth is more pronounced in port regions. After taking the model to the data, we conduct a counterfactual exercise to quantify what fraction of this port region growth can be explained by our mechanism, i.e., by an endogenous reaction of trade costs to country productivity growth, due to the round-trip effect.

4.1 Data and time period

For night lights, we use the grid-cell level panel dataset between 1992 and 2013 that we introduce in Section 2.1. Using cells' luminosity levels, we construct proxies of regional income, which correspond to $L(r)\,y(r)$ in the model.

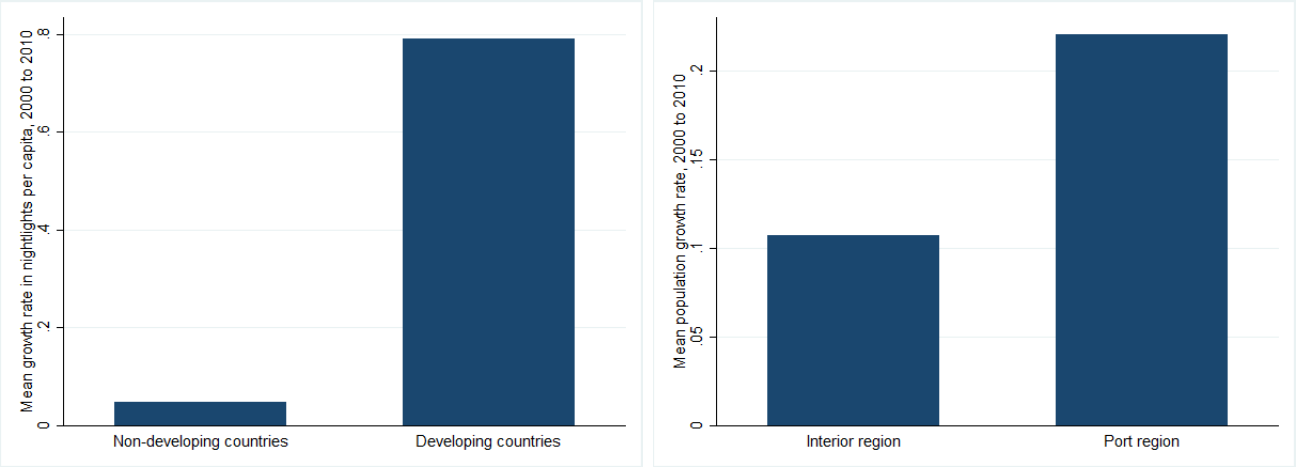
For export and import quantities and values, we use CEPII's BACI dataset (Gaulier and Zignago, 2010), which provides such data at the country pair level between 1995 and 2024.

For population, we use Gridded Population of the World (CIESIN, 2018), which provides estimates of population for the years 2000, 2005, 2010, 2015 and 2020 at the resolution of 1 by 1 arc minute (approximately 1 by 1 km) grid cells — the same spatial resolution as in the night light data. These data will serve as estimates of regional population $L(r)$.

Finally, we also obtain the land area in each grid cell from the Gridded Population of the World dataset, as opposed to the area covered by water. This will allow us to measure each region's land supply $H(r)$.

As the period between 2000 and 2010 is available in each dataset, we focus on this as our study period.[7]

Figure 7. Summary statistics on population and nightlight growth



Notes: The left figure displays the growth rate of mean nightlights per capita between 2000 and 2010 for non-developing countries (left) vs developing countries (right). Developing countries are defined as countries that had above-median nightlight growth rate between 2000 and 2010, while non-developing countries are those with below-median nightlight growth rate. The right figure displays the mean 2000 to 2010 population growth rate of interior regions (1 by 1 arc minute grid cells more than 50km from the nearest port, left) and port regions (less than 50km from the nearest port, right) within developing countries.

4.2 Quantification strategy

Our baseline model features three regions: the port region and the interior (which together constitute the country), as well as the ROW. As we aim to use the model to study the effects of productivity growth in the country while keeping productivity in the ROW constant, we rank countries with respect to their nightlights per capita growth rate between 2000 and 2010 and label the 50% of countries that had above-median growth rate as "developing countries." We call the remaining 50% of countries "non-developing countries." As shown in the left panel of Figure 7, non-developing countries were growing at a rate between 2000 and 2010 that was statistically indistinguishable from zero. This implies that they are a proper ROW according to the model. Developing countries, by contrast, were growing rapidly: their average nightlights per capita growth rate between 2000 and 2010 was as high as 79%. Looking at the list of countries that we classify as "developing," most third-world countries are in fact in this group, while first-world countries are typically not. As an example, China is in the group of developing countries while the U.S. is not; trade between these two countries serves as the prime example of the round-trip effect (Wong, 2022).

Next, we split developing countries into a port region and an interior region. Motivated by the evidence of Section 2 in which we document effects on night lights up to 50 kilometers away from the port, we group grid cells that are less than 50 kilometers from the nearest port in a

[7] We are currently working on expanding the study period to 2020 using more recent night light data.

developing country into a port region, while we label the rest of the cells in these countries as the interior. The port region already features higher population density than the interior region in 2000. However, this gap in population density grows even larger over time: while the population of the interior grows at an average of 10.7% between 2000 and 2010, the population of the port region grows at more than twice this rate, by 22.1% (right panel of Figure 7). That is, the port region indeed saw disproportionately fast growth in its population.

To take the model to the data, we measure each region's population, land supply, and income per capita — proxied by nightlights per capita — in 2000. Armed with these data, we use Theorem 1 to back out the distribution of amenities relative to mobility costs $a(\cdot)/m(\cdot)$ and productivities $A(\cdot)$ that rationalize these data as an equilibrium. This is important as amenities and productivities are not directly observable. To implement Theorem 1, we use the standard model inversion algorithm of Desmet, Nagy and Rossi-Hansberg (2018).

To conduct the model inversion, we need to take a stand on the values of structural parameters and bilateral trade costs. For the trade elasticity θ, we choose a value of four, a standard estimate in the literature (Simonovska and Waugh, 2014). For the congestion and round-trip elasticities, we rely on recent estimates in the literature on maritime trade. In particular, we choose $\gamma_1 = 0.096$ following evidence by Fuchs and Wong (2024) and $\gamma_2 = 0.216$ following evidence by Wong (2022).[8] These values of γ_1, γ_2 and θ satisfy the regularity condition $\gamma_2 < \gamma_1 + \frac{1}{1+\theta}$ as well as the inequality conditions of Proposition 3.

Finally, we choose the values of export costs τ_1^X and import costs τ_1^M to match two moments in the data: (i) the value of total trade between developing and non-developing countries relative to world GDP in 2000 (11.3% in the data), and (ii) the ratio of export and import quantities between the same two groups of countries in 2000. In line with the round-trip mechanism, the value of this second data moment is above one (1.29). That is, developing countries export larger quantities of goods to the ROW than they import from there. In the model, this difference in quantities triggers the round-trip mechanism and thus makes imports cheaper relative to exports. That is, the second data moment allows us to discipline the *ratio* of τ_1^X and τ_1^M. By contrast, the first data moment we match disciplines the *level* of τ_1^X and τ_1^M as higher levels of these costs decrease total trade. This is the intuition for why these two data moments are sufficient to identify τ_1^X and τ_1^M. Once τ_1^X and τ_1^M are known, equation (9) can be rearranged to identify the structural parameters $\bar{\tau}_1^X$ and $\bar{\tau}_1^M$. This step completes the model quantification.

[8] We are currently working on estimating γ_1 and γ_2 within our model.

4.3 Counterfactual: Country productivity growth with and without the round-trip effect

To quantify the contribution of the round-trip effect to the disproportionate growth of port regions seen in the data (right panel of Figure 7), we now conduct the following counterfactual exercise with the quantified model. We first simulate a spatially neutral improvement in country productivity in the full model, i.e., under endogenous trade costs due to the round-trip and congestion effects. Next, we simulate the same productivity improvement in a "benchmark" model with exogenous trade costs, i.e., a version of our model in which we set $\gamma_1 = \gamma_2 = 0$, and therefore $\bar{\tau}_1^X = \tau_1^X$ and $\bar{\tau}_1^M = \tau_1^M$. To make the magnitude of the productivity improvement comparable to the data, we choose the increase in country productivity $\bar{A}$ to be identical in size as the nightlight per capita growth seen in the data (79%, left panel of Figure 7), both in the full model and in the exogenous-cost benchmark. The difference between these two model simulations gives us the contribution of the round-trip effect.

Two important conclusions emerge from our counterfactual exercise. First, we find that the population of the port region relative to the interior is 1.78% higher in the full model than in the exogenous-cost benchmark. As the difference in growth between these two regions was 11.4 percentage points in the data, this implies that the round-trip mechanism can account for $1.78/11.4 = 15.6\%$ of the disproportionate growth in port regions' population in developing countries.

Second, we find a higher value of aggregate welfare $\bar{U}$ in the full model than in the exogenous-cost benchmark. This is not surprising as the productivity improvement triggers a reduction in trade costs in the full model, which leads to additional welfare gains from trade. We find that the welfare gains from the productivity improvement compared to the 2000 equilibrium are 24.1% higher in the full model. Put differently, our results suggest that the round-trip effect amplifies the welfare gains from developing countries' recent productivity growth by 24.1%. While these results are preliminary, they suggest that the round-trip effect has a quantitatively large impact on the distribution of population and welfare in the developing world.

5 Conclusion

This paper examines whether spatially neutral productivity improvements can generate uneven economic outcomes across regions. We provide empirical evidence that productivity improvements from internet expansion, vaccination campaigns, and education systematically shift economic activity toward port regions. We explain this pattern through a model with endogenous and asymmetric trade costs driven by the round-trip effect: increases in exports reduce import costs, disproportionately benefiting regions with direct access to international markets. This

mechanism contrasts with standard models featuring exogenous trade costs and provides a unified explanation for the empirical findings.

Our quantitative results indicate that the round-trip effect plays a meaningful role in shaping spatial outcomes, accounting for a significant share of the relative growth of port regions and amplifying aggregate welfare gains from productivity improvements. These findings suggest that policies aimed at raising national productivity may inadvertently increase spatial inequality by favoring globally connected regions. More broadly, the paper highlights the importance of incorporating endogenous trade costs into analyses of the spatial economy, as well as the need for complementary policies that support more balanced regional growth.

References

Allen, Treb, and Costas Arkolakis. 2014. "Trade and the Topology of the Spatial Economy." *The Quarterly Journal of Economics*, 129(3): 1085–1140.

Bloom, David E, David Canning, and Mark Weston. 2005. "The Value of Vaccination." *World Economics*, 6(3): 15–39.

Brancaccio, Giulia, Myrto Kalouptsidi, and Theodore Papageorgiou. 2020. "Geography, transportation, and endogenous trade costs." *Econometrica*, 88(2): 657–691.

Chatterjee, Shoumitro, Elisa Giannone, and Kan Kuno. 2021. "Unequal global convergence." Mimeo, Pennsylvania State University.

Chatterjee, Shoumitro, Rohit Lamba, and Sonalika Rai. 2026. "Spatial Inequality in India." Working Paper.

Chauvin, Juan Pablo, Edward Glaeser, Yueran Ma, and Kristina Tobio. 2017. "What is different about urbanization in rich and poor countries? Cities in Brazil, China, India and the United States." *Journal of Urban Economics*, 98: 17–49.

Chen, Natalie, Dennis Novy, Carlo Perroni, and Horng Chern Wong. 2023. "Urban-biased structural change." CESifo Working Paper.

Chen, Yuxian, Yannis M Ioannides, and Ferdinand Rauch. 2025. "Asymmetric trading costs and ancient Greek cities." Working Paper.

CIESIN. 2018. "Gridded Population of the World, Version 4 (GPWv4): Population Density, Revision 11." Accessed: 2026-03.

Coşar, A Kerem, and Banu Demir. 2018. "Shipping inside the box: Containerization and trade." *Journal of International Economics*, 114: 331–345.

Coşar, A. Kerem, and Pablo D. Fajgelbaum. 2016. "Internal Geography, International Trade, and Regional Specialization." *American Economic Journal: Microeconomics*, 8(1): 24–56.

Desmet, Klaus, Dávid Krisztián Nagy, and Esteban Rossi-Hansberg. 2018. "The Geographic Spread of a Growth Phenomenon." *Journal of Political Economy*, 126(3): 1015–1077.

Do, Anh D, Sharat Ganapati, Woan Foong Wong, and Oren Ziv. 2025. "Transshipment hubs, trade, and supply chains." National Bureau of Economic Research.

Ducruet, César, Réka Juhász, Dávid Krisztián Nagy, and Claudia Steinwender. 2024. "All aboard: The effects of port development." *Journal of International Economics*, 151: 103963.

Duflo, Esther. 2001. "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment." *American Economic Review*, 91(4): 795–813.

Eaton, Jonathan, and Samuel Kortum. 2002. "Technology, Geography, and Trade." *Econometrica*, 70(5): 1741–1779.

Eckert, Fabian, Sharat Ganapati, and Conor Walsh. 2025. "The Re-emergence of Spatial Inequality in the US." NBER Working Paper.

Economides, Philip. 2024. "Unconventional Protectionism in Containerized Shipping." *Available at SSRN 4930357*.

Faber, Benjamin. 2014. "Trade Integration, Market Size, and Industrialization: Evidence from China's National Trunk Highway System." *The Review of Economic Studies*, 81(3): 1046–1070.

Fan, Tianyu, Michael Peters, and Fabrizio Zilibotti. 2023. "Growing like India—the unequal effects of service-led growth." *Econometrica*, 91(4): 1457–1494.

Fay, Marianne, and Charlotte Opal. 2000. "Urbanization without growth: a not-so-uncommon phenomenon." The World Bank Policy Research Working Paper 2412.

Fuchs, Simon, and Woan Foong Wong. 2024. "Multimodal Transport Networks." CEPR Discussion Paper 19531. CEPR Press, Paris & London.

Ganapati, Sharat, Woan Foong Wong, and Oren Ziv. 2024. "Entrepot: Hubs, scale, and trade costs." *American Economic Journal: Macroeconomics*, 16(4): 239–278.

Gaulier, Guillaume, and Soledad Zignago. 2010. "BACI: International Trade Database at the Product-Level. The 1994-2007 Version." CEPII Working Papers.

Glaeser, Edward, and J. Vernon Henderson. 2017. "Urban economics for the developing World: An introduction." *Journal of Urban Economics*, 98: 1–5.

Gollin, Douglas, Remi Jedwab, and Dietrich Vollrath. 2016. "Urbanization with and without industrialization." *Journal of Economic Growth*, 21(1): 35–70.

Head, Keith, and John Ries. 2001. "Increasing Returns versus National Product Differentiation as an Explanation for the Pattern of U.S.-Canada Trade." *American Economic Review*, 91(4): 858–876.

Head, Keith, and Thierry Mayer. 2021. "The United States of Europe: A Gravity Model Evaluation of the Four Freedoms." *Journal of Economic Perspectives*, 35(2): 23–48.

Heiland, Inga, Andreas Moxnes, Karen-Helene Ulltveit-Moe, and Yuan Zi. 2019. "Trade From Space: Shipping Networks and The Global Implications of Local Shocks." C.E.P.R. Discussion Papers CEPR Discussion Papers 14193.

Henderson, J. Vernon, Adam Storeygard, and Uwe Deichmann. 2017. "Has climate change driven urbanization in Africa?" *Journal of Development Economics*, 124: 60–82.

Hjort, Jonas, and Jonas Poulsen. 2019. "The Arrival of Fast Internet and Employment in Africa." *American Economic Review*, 109(3): 1032–79.

Khanna, Gaurav. 2023. "Large-scale education reform in general equilibrium: Regression discontinuity evidence from india." *Journal of Political Economy*, 131(2): 549–591.

Lee, Sang Min. 2026. "Globalization and structural transformation: The role of tradable services." *Journal of International Economics*, 104220.

Redding, Stephen J. 2016. "Goods Trade, Factor Mobility and Welfare." *Journal of International Economics*, 101: 148–167.

Redding, Stephen J, and Esteban Rossi-Hansberg. 2017. "Quantitative Spatial Economics." *Annual Review of Economics*, 9: 21–58.

Simonovska, Ina, and Michael Waugh. 2014. "The elasticity of trade: Estimates and evidence." *Journal of International Economics*, 92(1): 34–50.

Smits, Jeroen, and Iñaki Permanyer. 2019. "The subnational human development database." *Scientific data*, 6(1): 190038.

Stiglitz, Joseph E. 2012. *The Price of Inequality: How Today's Divided Society Endangers Our Future.* WW Norton & Company.

UNCTAD. 2024. "Review of Maritime Transport 2024." United Nations Conference on Trade and Development.

Waugh, Michael E. 2010. "International Trade and Income Differences." *American Economic Review*, 100(1): 493–527.

Wong, Woan Foong. 2022. "The round trip effect: Endogenous transport costs and international trade." *American Economic Journal: Applied Economics*, 14(4): 127–166.

Appendix

This appendix provides several complements to the theory of Section 3. In Section A, we provide a formal definition of the equilibrium and derive equations that characterize the equilibrium distribution of export costs, import costs, population by region, and income by region. In Section B, we prove Propositions 1 to 3 from Section 3.2 and Theorems 1 and 2 from Section 3.3. In Section C, we show that Theorems 1 and 2 generalize to an extended model in which we allow for agglomeration and congestion forces and relax the restriction $\theta = \mu/(1-\mu)$. Finally, we further generalize the model and Theorems 1 and 2 by allowing for an arbitrary number of regions in Section D.

A Equilibrium of the model

We first provide a formal definition of the equilibrium. Next, we use the equilibrium conditions to obtain expressions for export costs, import costs, population and income.

Definition. *For given values of structural parameters* $\{\rho, \theta, \tau_0, \bar{\tau}_1^X, \bar{\tau}_1^M, \gamma_1, \gamma_2\}$ *and a given distribution of fundamentals* $\{a(\cdot), m(\cdot), A(\cdot), H(\cdot)\}$*, the equilibrium of the model is a distribution of agents* $L(\cdot)$*, wages* $w(\cdot)$*, land rents* $R(\cdot)$*, goods' prices* $p_\omega(r)$ *and quantities* $q(r,s)$ *such that:*

1. *Agents choose their consumption levels to maximize their utility (7) subject to their income (8), implying that the utility level of agents in region* r *is given by*

$$u(r) = \tilde{a}(r) \frac{y(r)}{P(r)} \tag{A.1}$$

where $\tilde{a}(r) = a(r)/m(r)$*, and* $P(r)$ *is the CES price index at* r*.*

2. *Agents choose their region of residence optimally, implying that*

$$\bar{U} = u(r) \tag{A.2}$$

for every region r *with positive population in equilibrium.*

3. *Firms choose their output, labor and land use to maximize their profits under perfectly competitive markets, implying that factory-gate prices become equal to marginal costs,*

$$p_\omega(r) = z_\omega(r)^{-1} w(r)^{\mu} R(r)^{1-\mu}$$

for each good ω *produced in region* r*, and factor payments satisfy*

$$R(r) H(r) = \frac{1-\mu}{\mu} w(r) L(r) \tag{A.3}$$

in each region r.

4. *Agents source goods from the cheapest source, which implies, as in Eaton and Kortum (2002), that trade shares are given by*

$$\pi(r,s) = \frac{A(r)^{\theta} w(r)^{-\mu\theta} R(r)^{-(1-\mu)\theta} \tau(r,s)^{-\theta}}{\sum_v A(v)^{\theta} w(v)^{-\mu\theta} R(v)^{-(1-\mu)\theta} \tau(v,s)^{-\theta}} \tag{A.4}$$

where $\tau(r,s)$ denotes the trade cost from region r to region s, the price index in region r satisfies

$$P(r)^{-\theta} = \sum_s A(s)^{\theta} w(s)^{-\mu\theta} R(s)^{-(1-\mu)\theta} \tau(s,r)^{-\theta} \tag{A.5}$$

and the quantity of goods shipped from region r to region s is given by

$$q(r,s) = A(r)^{1+\theta} w(r)^{-\mu(1+\theta)} R(r)^{-(1-\mu)(1+\theta)} \tau(r,s)^{-(1+\theta)} P(s)^{\theta} y(s) L(s). \tag{A.6}$$

5. *Markets for goods clear, implying*

$$y(r) L(r) = \sum_s \pi(r,s) y(s) L(s). \tag{A.7}$$

6. *Markets for labor clear, implying*

$$\bar{L} = \sum_r L(r). \tag{A.8}$$

7. *Markets for land clear.*

8. *Export and import costs are determined by equation (9).*

Next, we show how the above equilibrium conditions can be used to express equilibrium export costs, import costs, population levels and income levels in each region as a function of fundamentals and $\bar{U}$. First, we combine equations (A.3) and (A.5) to obtain

$$P(r)^{-\theta} = \sum_s A(s)^{\theta} H(s)^{\frac{\theta}{1+\theta}} L(s)^{-\frac{\theta}{1+\theta}} y(s)^{-\theta} \tau(s,r)^{-\theta} \tag{A.9}$$

as well as equations (A.3), (A.4) and (A.7) to obtain

$$y(r) L(r) = \sum_s A(r)^{\theta} H(r)^{\frac{\theta}{1+\theta}} L(r)^{-\frac{\theta}{1+\theta}} y(r)^{-\theta} P(s)^{\theta} y(s) L(s) \tau(r,s)^{-\theta} \tag{A.10}$$

where we have also used the restriction $\theta = \mu/(1-\mu)$. Next, we use equations (A.1) and (A.2) to express the price index as

$$P(r) = \bar{U}^{-1} \tilde{a}(r) y(r). \tag{A.11}$$

Plugging this into equations (A.9) and (A.10) yields the system

$$\begin{aligned}\tilde{a}(r)^{-\theta} y(r)^{-\theta} &= \bar{U}^{-\theta} \sum_s A(s)^{\theta} H(s)^{\frac{\theta}{1+\theta}} L(s)^{-\frac{\theta}{1+\theta}} y(s)^{-\theta} \tau(s,r)^{-\theta} \\ A(r)^{-\theta} H(r)^{-\frac{\theta}{1+\theta}} y(r)^{1+\theta} L(r)^{\frac{1+2\theta}{1+\theta}} &= \bar{U}^{-\theta} \sum_s \tilde{a}(s)^{\theta} y(s)^{1+\theta} L(s) \tau(r,s)^{-\theta}\end{aligned} \tag{A.12}$$

which, under symmetric trade costs $\tau(r,s) = \tau(s,r)$, can be reduced to one equation for $L(\cdot)$ alone (Desmet, Nagy and Rossi-Hansberg, 2018). Trade costs are not symmetric in our setting. However, it is easy to verify that if geography is as in Figure 4, income per capita in region r can be expressed as

$$y(r) = \bar{y} \cdot \tilde{a}(r)^{-\frac{\theta}{1+2\theta}} A(r)^{\frac{\theta}{1+2\theta}} H(r)^{\frac{\theta}{(1+\theta)(1+2\theta)}} \Delta(r)^{-\frac{\theta}{1+2\theta}} L(r)^{-\frac{1}{1+\theta}} \tag{A.13}$$

where $\bar{y}$ can be chosen arbitrarily as a term normalizing prices, and the term $\Delta(r)$ is a 'downstreamness index' that takes the following values: $\Delta(r) = \tau_1^X \tau_0$ if r is the port region or the interior, and $\Delta(r) = \tau_1^M \tau_0$ if r is the ROW. Substituting (A.13) into either equation in system (A.12), we obtain the following equation for region r's population as a function of $\bar{U}$, trade costs, and fundamentals:

$$L(r)^{\frac{\theta}{1+\theta}} = \bar{U}^{-\theta} \tilde{a}(r)^{\frac{\theta(1+\theta)}{1+2\theta}} A(r)^{\frac{\theta^2}{1+2\theta}} H(r)^{\frac{\theta^2}{(1+\theta)(1+2\theta)}} \Delta(r)^{-\frac{\theta^2}{1+2\theta}} \sum_s \tilde{a}(s)^{\frac{\theta^2}{1+2\theta}} A(s)^{\frac{\theta(1+\theta)}{1+2\theta}} H(s)^{\frac{\theta}{1+2\theta}} \Delta(s)^{\frac{\theta^2}{1+2\theta}} \tau(s,r)^{-\theta} \tag{A.14}$$

In the special case in which trade costs are exogenous and symmetric, i.e., $\tau_1^X = \tau_1^M$, $\Delta(r)$ does not vary across space. Therefore, equation (A.14) reduces to a usual equation that determines region r's population as a function of $\bar{U}$ as well as region r's level of amenities, productivity, land supply, and market access:

$$L(r)^{\frac{\theta}{1+\theta}} = \bar{U}^{-\theta} \tilde{a}(r)^{\frac{\theta(1+\theta)}{1+2\theta}} A(r)^{\frac{\theta^2}{1+2\theta}} H(r)^{\frac{\theta^2}{(1+\theta)(1+2\theta)}} \sum_s \tilde{a}(r)^{\frac{\theta^2}{1+2\theta}} A(s)^{\frac{\theta(1+\theta)}{1+2\theta}} H(s)^{\frac{\theta}{1+2\theta}} \tau(s,r)^{-\theta}$$

and $\bar{U}$ can be expressed from combining equation (A.14) with equation (A.8) and rearranging:

$$\bar{U} = \left[\frac{\sum_r \tilde{a}(r)^{\frac{(1+\theta)^2}{1+2\theta}} A(r)^{\frac{\theta(1+\theta)}{1+2\theta}} H(r)^{\frac{\theta}{1+2\theta}} \left[\sum_s \tilde{a}(s)^{\frac{\theta^2}{1+2\theta}} A(s)^{\frac{\theta(1+\theta)}{1+2\theta}} H(s)^{\frac{\theta}{1+2\theta}} \tau(s,r)^{-\theta} \right]^{\frac{1+\theta}{\theta}}}{\bar{L}} \right]^{\frac{1}{1+\theta}} \tag{A.15}$$

In the case of endogenous trade costs, however, equation (A.14) depends on the endogenous objects $\Delta(r)$, $\Delta(s)$ and $\tau(s,r)$. We now turn to solving equations (9) and (A.6) for these objects. In particular, combining equation (A.6) with (A.3), (A.11) and (A.13), the quantities shipped

between the port region and the ROW can be obtained as

$$\begin{aligned} q_X &= \bar{U}^{-\theta}\tilde{a}_p^{\frac{\theta(1+\theta)}{1+2\theta}} A_p^{\frac{(1+\theta)^2}{1+2\theta}} H_p^{\frac{1+\theta}{1+2\theta}} \left(\tau_1^X\right)^{-\frac{(1+\theta)^2}{1+2\theta}} \left(\tau_1^M\right)^{-\frac{\theta(1+\theta)}{1+2\theta}} \\ q_M &= \bar{U}^{-\theta}\tilde{a}_p^{\frac{\theta^2}{1+2\theta}} A_p^{\frac{\theta(1+\theta)}{1+2\theta}} H_p^{\frac{\theta}{1+2\theta}} \left(\tau_1^X\right)^{-\frac{\theta(1+\theta)}{1+2\theta}} \left(\tau_1^M\right)^{-\frac{(1+\theta)^2}{1+2\theta}} \end{aligned} \tag{A.16}$$

where we have normalized $\tilde{a}_{ROW} = A_{ROW} = H_{ROW} = 1$. Plugging these quantities into equation (A.6) and rearranging yields

$$\begin{aligned} \tau_1^X =& \bar{U}^{\frac{\theta(\gamma_2-\gamma_1)}{\chi_2}} \left(\bar{\tau}_1^X\right)^{\frac{1+2\theta-(1+\theta)[\theta\gamma_2-(1+\theta)\gamma_1]}{\chi_1\chi_2}} \left(\bar{\tau}_1^M\right)^{\frac{\chi_3}{\chi_1\chi_2}} \tilde{a}_p^{-\frac{\theta}{(1+2\theta)\chi_2}\left[\theta\gamma_2-(1+\theta)\gamma_1+\frac{(\gamma_1+\gamma_2)\chi_3}{\chi_1}\right]} \cdot \\ & \cdot A_p^{-\frac{1+\theta}{(1+2\theta)\chi_2}\left[\theta\gamma_2-(1+\theta)\gamma_1+\frac{(\gamma_1+\gamma_2)\chi_3}{\chi_1}\right]} H_p^{-\frac{1}{(1+2\theta)\chi_2}\left[\theta\gamma_2-(1+\theta)\gamma_1+\frac{(\gamma_1+\gamma_2)\chi_3}{\chi_1}\right]} \end{aligned} \tag{A.17}$$

for export costs and

$$\begin{aligned} \tau_1^M =& \bar{U}^{\frac{\theta(\gamma_2-\gamma_1)}{\chi_2}} \left(\bar{\tau}_1^X\right)^{\frac{1+2\theta-(1+\theta)[\theta\gamma_2-(1+\theta)\gamma_1]}{\chi_1\chi_2}-\frac{1+2\theta}{\chi_1}} \left(\bar{\tau}_1^M\right)^{\frac{\chi_3}{\chi_1\chi_2}+\frac{1+2\theta}{\chi_1}} \tilde{a}_p^{-\theta\left[\frac{1}{(1+2\theta)\chi_2}\left[\theta\gamma_2-(1+\theta)\gamma_1+\frac{(\gamma_1+\gamma_2)\chi_3}{\chi_1}\right]+\frac{\gamma_1+\gamma_2}{\chi_1}\right]} \cdot \\ & \cdot A_p^{-(1+\theta)\left[\frac{1}{(1+2\theta)\chi_2}\left[\theta\gamma_2-(1+\theta)\gamma_1+\frac{(\gamma_1+\gamma_2)\chi_3}{\chi_1}\right]+\frac{\gamma_1+\gamma_2}{\chi_1}\right]} H_p^{-\left[\frac{1}{(1+2\theta)\chi_2}\left[\theta\gamma_2-(1+\theta)\gamma_1+\frac{(\gamma_1+\gamma_2)\chi_3}{\chi_1}\right]+\frac{\gamma_1+\gamma_2}{\chi_1}\right]} \end{aligned} \tag{A.18}$$

for import costs as a function of fundamentals and $\bar{U}$, where $\chi_1 = 1 + 2\theta + (1+\theta)(\gamma_1+\gamma_2)$, $\chi_2 = 1 - (1+\theta)(\gamma_2-\gamma_1)$ and $\chi_3 = (1+\theta)[(1+\theta)\gamma_2 - \theta\gamma_1]$.

Once export and import costs are obtained from (A.17) and (A.18), the downstreamness indices of regions can be computed as $\Delta(r) = \tau_1^X\tau_0$ (for the port and the interior) and $\Delta(r) = \tau_1^M\tau_0$ (for the ROW). Next, equation (A.14) can be used to recover the population of each region, and $\bar{U}$ can be chosen such that equation (A.8) holds. Finally, equation (A.13) can be used to calculate each region's income per capita.

B Proofs of propositions and theorems

Proof of Proposition 1. Dividing equation (A.17) by (A.18), we can obtain relative export costs as

$$\frac{\tau_1^X}{\tau_1^M} = \left(\frac{\bar{\tau}_1^X}{\bar{\tau}_1^M}\right)^{\frac{1+2\theta}{\chi_1}} \tilde{a}_p^{\frac{\theta(\gamma_1+\gamma_2)}{\chi_1}} \bar{A}^{\frac{(1+\theta)(\gamma_1+\gamma_2)}{\chi_1}} \hat{A}_p^{\frac{(1+\theta)(\gamma_1+\gamma_2)}{\chi_1}} H_p^{\frac{\gamma_1+\gamma_2}{\chi_1}} \tag{B.19}$$

where we have used $A_p = \bar{A}\cdot\hat{A}_p$. As $\chi_1 > 0$, this expression is increasing in country productivity $\bar{A}$ if and only if $\gamma_1 + \gamma_2 > 0$. This proves part 2 of the proposition.

For relative export quantities, combining equation (A.16) with equation (B.19) yields

$$\frac{q_X}{q_M} = \left(\frac{\bar{\tau}_1^X}{\bar{\tau}_1^M}\right)^{-\frac{1+\theta}{\chi_1}} \tilde{a}_p^{\frac{\theta}{\chi_1}} \bar{A}^{\frac{1+\theta}{\chi_1}} \hat{A}_p^{\frac{1+\theta}{\chi_1}} H_p^{\frac{1}{\chi_1}}$$

which is increasing in $\bar{A}$ as $\chi_1 > 0$. This proves part 1 of the proposition.

Proof of Proposition 2. Note that under $\gamma_1 = \gamma_2 = 0$, τ_1^X and τ_1^M are exogenous and therefore do not depend on country productivity $\bar{A}$. From equation (A.14), we can also obtain the relative population of the port region as

$$\frac{L_p}{L_i} = \left(\frac{\tilde{a}_p}{\tilde{a}_i}\right)^{\frac{(1+\theta)^2}{1+2\theta}} \left(\frac{\hat{A}_p}{\hat{A}_i}\right)^{\frac{\theta(1+\theta)}{1+2\theta}} \left(\frac{H_p}{H_i}\right)^{\frac{\theta}{1+2\theta}} \tau_0^{1+\theta} \left[1 - \frac{F_i\left(1-\tau_0^{-2\theta}\right)}{F_{ROW}\tau_0^{-\theta}\bar{A}^{-\frac{\theta(1+\theta)}{1+2\theta}} + F_p\tau_0^{-\theta} + F_i}\right]^{\frac{1+\theta}{\theta}} \tag{B.20}$$

where $F_p = \tilde{a}_p^{\frac{\theta^2}{1+2\theta}} \hat{A}_p^{\frac{\theta(1+\theta)}{1+2\theta}} H_p^{\frac{\theta}{1+2\theta}} \left(\tau_1^X\right)^{\frac{\theta^2}{1+2\theta}}$, $F_i = \tilde{a}_i^{\frac{\theta^2}{1+2\theta}} \hat{A}_i^{\frac{\theta(1+\theta)}{1+2\theta}} H_i^{\frac{\theta}{1+2\theta}} \left(\tau_1^X\right)^{\frac{\theta^2}{1+2\theta}}$ and $F_{ROW} = \left(\tau_1^M\right)^{-\frac{\theta(1+\theta)}{1+2\theta}}$ are all independent of $\bar{A}$. Therefore, it is straightforward to see that (B.20) is decreasing in $\bar{A}$. This proves part 1 of the proposition.

To see that part 2 also holds, note that equation (A.13) implies

$$\frac{L_p y_p}{L_i y_i} = \left(\frac{\tilde{a}_p}{\tilde{a}_i}\right)^{-\frac{\theta}{1+2\theta}} \left(\frac{\hat{A}_p}{\hat{A}_i}\right)^{\frac{\theta}{1+2\theta}} \left(\frac{H_p}{H_i}\right)^{\frac{\theta}{(1+\theta)(1+2\theta)}} \left(\frac{L_p}{L_i}\right)^{\frac{\theta}{1+\theta}}.$$

Therefore, if $\frac{L_p}{L_i}$ is decreasing in $\bar{A}$, $\frac{L_p y_p}{L_i y_i}$ is also decreasing in $\bar{A}$.

Proof of Proposition 3. Under $\tilde{a}_p = \tilde{a}_i = \tilde{a}$, $\hat{A}_p = \hat{A}_i = 1$ and $H_p = H_i = H$, we can rewrite equation (B.20) as

$$\frac{L_p}{L_i} = \left[1 + \frac{1-\tau_0^{-\theta}}{\tau_0^{-\theta} + \tilde{a}^{\frac{\theta^2}{1+2\theta}} H^{\frac{\theta}{1+2\theta}} \left(1+\tau_0^{-\theta}\right) \bar{A}^{\frac{\theta(1+\theta)}{1+2\theta}} \left(\tau_1^X\right)^{\frac{\theta^2}{1+2\theta}} \left(\tau_1^M\right)^{\frac{\theta(1+\theta)}{1+2\theta}}}\right]^{\frac{1+\theta}{\theta}}.$$

Substituting for τ_1^X and τ_1^M from equations (A.17) and (A.18), we obtain

$$\frac{L_p}{L_i} = \left[1 + \frac{1-\tau_0^{-\theta}}{\tau_0^{-\theta} + F \cdot \bar{A}^{\frac{\theta(1+\theta)}{1+2\theta}\left[1 - \frac{\theta\gamma_2 - (1+\theta)\gamma_1 + \frac{(\gamma_1+\gamma_2)\chi_3}{\chi_1}}{\chi_2} - \frac{(1+\theta)(\gamma_1+\gamma_2)}{\chi_1}\right]}}\right]^{\frac{1+\theta}{\theta}}$$

where the constant F is independent of $\bar{A}$. This expression is increasing in $\bar{A}$ as long as $\gamma_2 > \frac{\sqrt{(1+2\theta)^2 + 4(1+\theta)\left[(1+\theta)\gamma_1^2 + 2(1+\theta)\gamma_1 + 1\right]} - (1+2\theta)}{2(1+\theta)}$, where we have also used the fact that $\theta\gamma_2 - (1+\theta)\gamma_1 > 0$ due to $\gamma_1 < \frac{\theta}{1+\theta}\gamma_2$ and, consequently, $\chi_3 = [\theta\gamma_2 - (1+\theta)\gamma_1] + (\gamma_1 + \gamma_2) > 0$.

The proof of part 2 is analogous to the proof of part 2 of Proposition 2.

Proof of Theorem 1. Rearranging equation (A.13) yields

$$A(r) = \bar{y}^{-\frac{1+2\theta}{\theta}} \tilde{a}(r) H(r)^{-\frac{1}{1+\theta}} \Delta(r) L(r)^{\frac{1+2\theta}{\theta(1+\theta)}} y(r)^{\frac{1+2\theta}{\theta}} \tag{B.21}$$

Plugging this into equation (A.14) and rearranging yields

$$\tilde{a}(r)^{-\theta} y(r)^{-\theta} = \bar{U}^{-\theta} \bar{y}^{-(1+2\theta)} \sum_s \tilde{a}(s)^{\theta} \Delta(s)^{\theta} y(s)^{1+\theta} L(s) \tau(s,r)^{-\theta} \tag{B.22}$$

Theorem 2 in Allen and Arkolakis (2014) implies that equation (B.22) yields a unique solution for $\tilde{a}(\cdot)$. This solution can, in turn, be substituted into equation (B.21) to obtain a unique distribution of $A(\cdot)$. Both distributions are only unique up to scale as $\bar{U}$ and $\bar{y}$ are not observed and can be chosen arbitrarily.

Proof of Theorem 2. We can write equation (B.22) for the port region as

$$\tilde{a}_p^{-\theta} y_p^{-\theta} = \tilde{U}^{-\theta} \bar{y}^{-(1+\theta)} \left[\tilde{a}_p^{\theta} \left(\tau_1^X \tau_0\right)^{\theta} y_p^{1+\theta} L_p + \tilde{a}_i^{\theta} \left(\tau_1^X\right)^{\theta} y_i^{1+\theta} L_i + \tilde{a}_{ROW}^{\theta} \tau_0^{\theta} y_{ROW}^{1+\theta} L_{ROW} \right]$$

while under the false assumption of symmetric trade costs $\left[\tau_1^X \tau_1^M\right]^{1/2}$ between the port region and the ROW, equation (B.22) would take the form

$$\dot{\tilde{a}}_p^{-\theta} y_p^{-\theta} = \tilde{U}^{-\theta} \bar{y}^{-(1+\theta)} \left[\dot{\tilde{a}}_p^{\theta} y_p^{1+\theta} L_p + \dot{\tilde{a}}_i^{\theta} y_i^{1+\theta} L_i \tau_0^{-\theta} + \dot{\tilde{a}}_{ROW}^{\theta} y_{ROW}^{1+\theta} L_{ROW} \left[\tau_1^X \tau_1^M\right]^{-\theta/2} \right]$$

It is straightforward to see that the two equations become identical if $\dot{\tilde{a}}_p = \tilde{a}_p \left[\tau_1^X \tau_0\right]^{1/2}$, $\dot{\tilde{a}}_i = \tilde{a}_i \left[\tau_1^X \tau_0\right]^{1/2}$, and $\dot{\tilde{a}}_{ROW} = \tilde{a}_{ROW} \left[\tau_1^M \tau_0\right]^{1/2}$. The proofs for $\tilde{a}_i$ and $\tilde{a}_{ROW}$ are analogous.

For productivities, equation (B.21) for the port region takes the form

$$A_p = \bar{y}^{-\frac{1+2\theta}{\theta}} \tilde{a}_p H_p^{-\frac{1}{1+\theta}} \tau_1^X \tau_0 L_p^{\frac{1+2\theta}{\theta(1+\theta)}} y_p^{\frac{1+2\theta}{\theta}}$$

while under the false assumption of symmetric trade costs, it would be

$$\dot{A}_p = \bar{y}^{-\frac{1+2\theta}{\theta}} \dot{\tilde{a}}_p H_p^{-\frac{1}{1+\theta}} L_p^{\frac{1+2\theta}{\theta(1+\theta)}} y_p^{\frac{1+2\theta}{\theta}}.$$

Substituting for $\dot{\tilde{a}}_p$ and rearranging yields the result that $\dot{A}_p = A_p \left[\tau_1^X \tau_0\right]^{-1/2}$. The proofs for A_i and A_{ROW} are analogous.

C Model with $\theta \neq \mu/(1-\mu)$, agglomeration, and congestion

In this section, we extend the model in four respects and show that Theorems 1 and 2 generalize to the extended model. First, we relax the restriction $\theta = \mu/(1-\mu)$. Second, we introduce the possibility of agglomeration forces by allowing the productivity of a region to depend on its population density with elasticity $\alpha \geq 0$:

$$A(r) = T(r)\left[\frac{L(r)}{H(r)}\right]^{\alpha}$$

Third, we introduce congestion in amenities with elasticity $\lambda \geq 0$:

$$a(r) = b(r)\left[\frac{L(r)}{H(r)}\right]^{-\lambda}$$

Fourth, we allow agents to draw idiosyncratic tastes for regions. That is, we assume that an agent i's utility function takes the form

$$u_i(r) = a(r) \cdot \left[\int_0^1 c_{i,\omega}(r)^{\rho} d\omega\right]^{\frac{1}{\rho}} \cdot m(r)^{-1} \varepsilon_i(r)$$

where $\varepsilon_i(r)$ is distributed Fréchet, independently across agents and regions, with shape parameter $1/\Omega > 0$. This formulation, which is common in the literature, can capture any finite migration elasticity across regions, equal to $1/\Omega$, which has been shown to be the empirically relevant case (Head and Mayer, 2021).

Following the same steps as in Section A and using the properties of the Fréchet distribution, one can show that under these assumptions equations (A.13) and (A.14) become

$$y(r) = \bar{y} \cdot \tilde{b}(r)^{-\frac{\theta}{1+2\theta}} T(r)^{\frac{\theta}{1+2\theta}} H(r)^{\frac{\theta}{1+2\theta}[(1-\mu)-\alpha-\lambda]} \Delta(r)^{-\frac{\theta}{1+2\theta}} L(r)^{\frac{[\alpha+\lambda+\Omega-(1-\mu)]\theta-1}{1+2\theta}} \tag{C.23}$$

and

$$L(r)^{(\lambda+\Omega)\theta-\frac{\theta}{1+2\theta}[[\alpha+\lambda+\Omega-(1-\mu)]\theta-1]} = \bar{U}^{-\theta}\tilde{b}(r)^{\frac{\theta(1+\theta)}{1+2\theta}} T(r)^{\frac{\theta^2}{1+2\theta}} H(r)^{\frac{\theta^2}{1+2\theta}[(1-\mu)-\alpha-\lambda-\Omega]} \Delta(r)^{-\frac{\theta^2}{1+2\theta}}$$
$$\sum_s \tilde{b}(r)^{\frac{\theta^2}{1+2\theta}} T(s)^{\frac{\theta(1+\theta)}{1+2\theta}} H(s)^{\frac{\theta(1+\theta)}{1+2\theta}[(1-\mu)-\alpha]} \Delta(s)^{\frac{\theta^2}{1+2\theta}} L(s)^{1-(\lambda+\Omega)\theta+\frac{1+\theta}{1+2\theta}[[\alpha+\lambda+\Omega-(1-\mu)]\theta-1]} \tau(s,r)^{-\theta} \tag{C.24}$$

where $\tilde{b}(r) = b(r)/m(r)$, and therefore equations (B.21) and (B.22) become

$$T(r) = \bar{y}^{-\frac{1+2\theta}{\theta}} \tilde{b}(r) H(r)^{\alpha+\lambda-(1-\mu)} \Delta(r) L(r)^{\frac{1+\theta}{\theta}-\alpha-\mu-\lambda-\Omega} y(r)^{\frac{1+2\theta}{\theta}} \tag{C.25}$$

and

$$\tilde{b}(r)^{-\theta} H(r)^{-\lambda\theta} L(r)^{(\lambda+\Omega)\theta} y(r)^{-\theta} = \bar{U}^{-(1+\theta)} \bar{y}^{-\frac{(1+\theta)(1+2\theta)}{\theta}} \sum_s \tilde{b}(s)^{\theta} H(s)^{\lambda\theta} \Delta(s)^{\theta} y(s)^{1+\theta} L(s)^{1-(\lambda+\Omega)\theta} \tau(s,r)^{-\theta}. \tag{C.26}$$

These two equations allow us to prove the following generalized versions of Theorems 1 and 2 by following the exact same steps as in Section B:

Theorem 3. *Given the values of structural parameters as well as data on population $L(\cdot)$, income $y(\cdot)$ and the supply of land $H(\cdot)$ by region as well as bilateral trade costs $\tau(\cdot,\cdot)$ across regions, the model can be inverted to back out (up to scale) unique distributions of exogenous amenities relative to mobility costs $b(\cdot)/m(\cdot)$ and exogenous productivities $T(\cdot)$ that rationalize the data as an equilibrium.*

Proof. The proof follows the same steps as the proof of Theorem 1 in Section B. □

Theorem 4. *Assume that the researcher thinks that trade costs are exogenous, symmetric and given by the Head and Ries index:*

$$\dot{\tau}_1^X = \dot{\tau}_1^M = \left[\tau_1^X \tau_1^M\right]^{1/2}.$$

Then the model inversion procedure of Theorem 3 falsely identifies fundamentals as

$$\frac{\dot{b}(r)}{\dot{m}(r)} = \frac{b(r)}{m(r)} \left[\tau_1^X \tau_0\right]^{1/2} \quad \text{and} \quad \dot{T}(r) = T(r) \left[\tau_1^X \tau_0\right]^{-1/2}$$

in the port and interior regions of the country, and as

$$\frac{\dot{b}_{ROW}}{\dot{m}_{ROW}} = \frac{b_{ROW}}{m_{ROW}} \left[\tau_1^M \tau_0\right]^{1/2} \quad \text{and} \quad \dot{T}_{ROW} = T_{ROW} \left[\tau_1^M \tau_0\right]^{-1/2}$$

in the ROW, where the variables without a dot denote the true values of fundamentals.

Proof. The proof follows the same steps as the proof of Theorem 2 in Section B. □

D General model with an arbitrary number of regions

In this section, we further generalize the model of Section C by allowing for an arbitrary number of regions. This is useful for potential applications in which the three-region assumption is not a good enough approximation to reality. That said, the generalization comes at the expense of two restrictions we need to put on geography. First, we need to assume that regions constitute a *tree*. A tree is a graph without cycles. That is, we assume that (i) not necessarily all regions are directly connected to each other, (ii) between any two regions, there is at most one route

through direct connections. The geography of our baseline model is a simple special case of such a tree since the port region can directly trade with both of the other regions but the interior region can only trade with the ROW through the port region.

Second, we need to assume that iceberg trade costs are log-additive over routes.[1] That is, between any two regions r and s that are connected by a route passing through region v, the trade cost between r and s can be decomposed into

$$\tau(r,s) = \tau(r,v)\,\tau(v,s).$$

This assumption is satisfied, for instance, by trade costs that are exponential functions of distance.

Under the log-additivity assumption, we can define the *downstreamness index* of region r as

$$\Delta(r) = \prod_{(s,v)\in\nu(r)} \tau(s,v) \tag{D.27}$$

where $\nu(r)$ denotes the set of region pairs (s,v) that are (i) on a route departing from r, (ii) are directly connected with one another, (iii) s is on the route between r and v. The intuition for the downstreamness index is that it captures the total trade cost that a traveler has to incur if they want to travel from region r to every other region that is reachable from there (but do not have to travel back). Obviously, this total trade cost is higher for more "downstream" regions that face higher costs of exporting.

Given these assumptions, one can follow the same steps as in Section A and show that the equilibrium distributions of income and population still satisfy equations (C.23) and (C.24), except that $\Delta(r)$ is now defined by equation (D.27). As a consequence, exogenous productivities and amenities over migration costs still satisfy equations (C.25) and (C.26). Therefore, Theorem 3 still holds, while Theorem 4 can be further generalized as follows:

Theorem 5. *Assume that the researcher thinks that trade costs are exogenous, symmetric and given by the Head and Ries index:*

$$\dot{\tau}(r,s) = \dot{\tau}(s,r) = \left[\tau(r,s)\,\tau(s,r)\right]^{1/2}.$$

Then the model inversion procedure of Theorem 3 falsely identifies fundamentals as

$$\frac{\dot{b}(r)}{\dot{m}(r)} = \frac{b(r)}{m(r)}\Delta(r)^{1/2} \quad \textit{and} \quad \dot{T}(r) = T(r)\,\Delta(r)^{-1/2}$$

where the variables without a dot denote the true values of fundamentals.

[1] This assumption is commonly made in quantitative geography models that use the Fast Marching algorithm or the Dijkstra algorithm to compute the lowest trade cost between locations. See, for example, Allen and Arkolakis (2014).

Proof. The proof follows the same steps as the proof of Theorem 2 in Section B, except that one needs to use the critical property of the general downstreamness index that

$$\frac{\tau(s,r)}{\tau(r,s)} = \frac{\Delta(s)}{\Delta(r)}.$$

□