# Grounded by Gravity: A Well-Behaved Trade Model with Industry-Level Economies of Scale*


Konstantin Kucheryavyy — U Tokyo

Gary Lyn — Iowa State

Andrés Rodríguez-Clare† — UC Berkeley and NBER


December 1, 2017


**Abstract**

For over a century trade economists have struggled to understand the positive and normative implications of localized industry-level external economies of scale. We develop a multi-country and multi-industry model with scale economies that admits a unique equilibrium and a standard gravity equation, and show that if the conditions for uniqueness are satisfied then all countries gain from trade. The presence of scale economies lowers the gains from trade (conditional on trade flows) except if the country specializes in industries with high scale economies, and they amplify the gains from trade liberalization except if it leads to specialization in industries with low scale economies. Our quantitative analysis reveals that the presence of scale economies implies on average larger gains from removing all tariffs observed in the data, with gains more than doubling for some countries, although gains fall for some countries and even become negative for one of the countries in our sample.



*We thank James Anderson, Costas Arkolakis, Gaurab Aryal, Backhoon Song, Lorenzo Caliendo, Arnaud Costinot, Richard Cottle, Svetlana Demidova, Dave Donaldson, Jonathan Eaton, Michal Fabinger, Cecile Gaubert, Gene Grossman, Nail Kashaev, Tim Kehoe, Hideo Konishi, Sam Kortum, Beresford Parlett, Donald Richards, Steve Redding, Michael Tsatsomeros, Guang Yang, Xi Yang, and Stephen Yeaple for valuable discussions, and Kala Krishna for pointing us back to perfect competition. We thank Mauricio Ulate and Piyush Panigrahi for valuable research assistance. All errors are our own.



†549 Evans Hall No. 3880, Berkeley, CA 94720-3880. Tel: (510) 642-0822 Fax: (510) 642-6615 Email: andres@econ.berkeley.edu

# 1. Introduction

For over a century trade economists have struggled to understand the positive and normative implications of localized industry-level external economies of scale — see Marshall (1890), Graham (1923), Chipman (1965) and Ethier (1982). Standard models yielded some discomforting results, including "a bewildering variety of equilibria" (Krugman, 1995) — even ones in which trade patterns do not conform to comparative advantage — along with the "paradoxical implication that trade motivated by the gains from concentrating production need not benefit the participating countries" (Grossman and Rossi-Hansberg, 2010). The source of these results is a circularity whereby the scale of an industry affects its productivity, while an industry's productivity affects its scale through the impact on the pattern of trade and specialization. In the standard analysis, this leads to multiple equilibria.

Grossman and Rossi-Hansberg (2010) proposed a two-country Ricardian model with national industry-level external economies of scale (or *Marshallian externalities*) that avoids this circularity by moving away from perfect competition and assuming Bertrand competition instead. Since firms in each industry understand the implications of their decisions on industry output and productivity, the equilibrium is unique and similar to that in Dornbusch, Fischer and Samuelson (1977). While the framework successfully eliminates the multiplicity of equilibria in a world free of trade costs, Lyn and Rodríguez-Clare (2013a,b) illustrate circumstances under which multiplicity arises in the presence of trade costs. Coupled with the fact that the equilibrium has mixed strategies for some levels of trade costs, the framework quickly becomes intractable.

In this paper we present a multi-country and multi-industry Ricardian model with Marshallian externalities that admits a unique equilibrium under intuitive parameter restrictions. We retain the assumption of perfect competition, as in the standard Ricardian model, but relax the implicit assumption in that model that firms within each industry are producing a homogenous good. In particular, we allow for intra-industry heterogeneity as in Eaton and Kortum (2002, henceforth EK) and find that this adds some "curvature" that helps in establishing uniqueness of equilibrium as long as the strength of Marshallian externalities is not too high. The framework yields the standard gravity-type equation and so provides a platform to assess quantitatively the importance of these externalities for the welfare effects of trade.[1]

The system of equilibrium equations of the multi-industry Ricardian model with Marshallian externalities turns out to be isomorphic to the equilibrium system of a more general version of the multi-industry Krugman (1980) model of product differentiation with internal economies of scale.[2] The existence and uniqueness result that we prove for our Ricardian setting can then be seamlessly applied to the multi-industry Krugman model. Not surprisingly, the isomorphism extends also to the multi-industry Melitz (2003) model if the productivity distribution is Pareto as in Chaney (2008) and the fixed exporting costs are paid in units of labor of the destination country.

The common mathematical structure that characterizes the equilibrium in all these multi-industry gravity models is governed by two elasticities that can vary across industries: the elasticity of bilateral trade flows to bilateral trade costs, commonly referred to as the *trade elasticity*; and the elasticity of productivity with respect to industry size, which we will refer to as the *scale elasticity*. We show that a necessary condition for uniqueness is that in all industries the product of these two elasticities is not higher than one. We establish that this condition is also sufficient for uniqueness in the case of a small open economy, if we have two countries or if there are more than two countries but trade is frictionless or wages are exogenous. The theoretical difficulty in establishing a more general sufficiency result lies in the fact that, with trade frictions and more than two countries, the excess labor demand system does not satisfy the gross substitutes property — a sufficient condition for unique wages that is often satisfied in similar environments. Alternative approaches to prove equilibrium uniqueness such as those in Kehoe (1980) or Allen, Arkolakis and Li (2016) turn out to be either impractical or to require a much more stringent condition than appears necessary.

In the second half of the paper we use our unified framework to study the implications of scale economies for the welfare effects of trade. We first establish that all countries gain from trade as long as the product of the trade and scale elasticities is weakly lower than one in all industries. This is so even if the scale elasticity differs

[1]Our analysis restricts to the case of Marshallian externalities, which operate inside each industry. An alternative case is the one in which some of the externalities operate across industries. Yatsynovich (2014) has recently shown conditions under which a model with such cross-industry externalities exhibits a unique equilibrium for the case with frictionless trade.

[2]Abdel-Rahman and Fujita (1990), Allen, Arkolakis and Takahashi (2014) and Redding (2016) explore similar isomorphisms for spatial equilibrium models in the economic geography literature.

across industries — for example, because of cross-industry variation in the strength of Marshallian externalities in the Ricardian model. This is an important finding in light of previous results with this type of model where countries could lose from trade.[3]

We extend the "sufficient statistics approach" to the quantification of the gains from trade in Arkolakis, Costinot and Rodríguez-Clare (2012) (henceforth ACR) to multi-industry models with scale economies. The isomorphism that ACR establish across single-industry models now applies to multi-industry models as well: for the same industry-level trade and scale elasticities, the models deliver the same gains from trade and the same counterfactual implications given industry-level data.[4] Given trade and scale elasticities, we now have a general formula for the gains from trade in terms of industry-level trade, expenditure and revenue shares. This formula can be used to explore the way in which scale economies affect the gains from trade.

For the simple case in which the scale elasticity is the same across industries, we show that the gains from trade are *lower* with scale economies than without. This result may seem counterintuitive, but the reader should keep in mind that the gains from trade are defined as *conditional on trade shares*, so that — as in Costinot and Rodríguez-Clare (2014) — we can compare the gains from trade implied by different models that are consistent with the same data. Thus, the intuition that scale economies should lead to larger gains from trade through deeper industry-level specialization and larger trade flows is simply not operative here, although as we explain below this intuition is relevant for the *gains from trade liberalization*.

So why do scale economies lead to lower gains from trade? The move back to autarky implies a reallocation of labor across industries that, in the presence of scale economies, leads to productivity gains in expanding industries and productivity losses in contracting industries. Since the industries that expand are those where the country has positive net imports, it must be that they have a high expenditure share or a low employment share. A higher expenditure share implies that a given productivity gain matters more for welfare, whereas a low employment share implies that a given absolute increase in employment leads to a higher proportional expansion and a higher productivity gain. As a result, a move back to autarky generates an expenditure-weighted

[3] See, for example, Ethier (1982).

[4] The gains from trade are defined as in ACR as the negative of the welfare change caused by a move back to autarky.

average productivity improvement, implying lower welfare losses. A corollary of this reasoning is that the decline in the gains from trade from the presence of economies of scale is stronger for economies that exhibit a higher degree of industry-level specialization.

The previous results are specific to the case in which the scale elasticity is the same across industries. If the scale elasticity varies across industries, the implication of economies of scale depends not only on the *degree* of industry-level specialization, but also on its *pattern*. Everything else equal, countries that happen to specialize in industries with high scale elasticities will gain more from trade than countries that specialize in sectors with low scale elasticities. Countries that specialize in industries with high scale elasticities may even gain more from trade in the presence of economies of scale compared to the standard model without them.

We next consider two simple cases for which we can analytically characterize the impact of scale economies on the gains from trade liberalization (i.e., the welfare effect from a decline in trade costs): first, a case of two mirror-image countries, and, second, a case with exogenous wages. In the first case we find that the gains from trade liberalization are higher with scale economies than without, a reflection of the magnified response of industry-level specialization and trade to the decline in trade costs in the presence of economies of scale. In the second case we find that countries lose from unilateral trade liberalization if the product of the trade and scale elasticities is above a threshold value that is a function of industry-level import and export shares. We think of this as a generalization of the results in Venables (1987), which were specific to the Krugman (1980) model with an outside good.

We complement our exploration of the effect of scale economies on the gains from trade and the gains from trade liberalization by applying our framework to data on 31 industries and 33 countries (and a rest of the world aggregate) from the World Input Output Database (WIOD, Timmer *et al.*, 2015) in 2008. The effects of scale economies on the gains from trade discussed above turn out to be quantitatively important. For example, moving from a model with no scale economies to one with a scale elasticity that is common across sectors and equal to that estimated in recent papers, the gains from trade for South Korea fall from 6.6% to 4.4%; however, allowing scale elasticities to vary across manufacturing sub-sectors as in Bartelme *et al.* (2017) and assuming no scale economies outside of manufacturing brings those gains back up to 7.3%. We also use the model to quantify the welfare implications of removing all tariffs observed in the data. We find that the presence of scale economies implies on average larger gains from the removal of all tariffs, with gains more than doubling for some countries. Still, gains could be lower for countries that are pushed to specialize in industries with relatively low economies of scale, even becoming negative in the case of Brazil.

We end with a quantitative exploration of the effect of scale economies on industry specialization and trade flows. We find that the removal of economies of scale implies a decline in the degree of specialization and total trade, but the effects are small. Without economies of scale amplifying Ricardian comparative advantage, industry-level specialization would be 1-3% less than what we observe in the data, implying that scale economies are much less important than Ricardian comparative advantage in driving industry-level specialization. In addition, without economies of scale world trade would be between 2.5% and 7% lower than what we see in the data.

There are two papers that study the question of uniqueness of equilibrium in the multi-industry Krugman model: Hanson and Xiang (2004) consider the case of two countries and a continuum of industries, while Behrens *et al.* (2009) consider the case of many countries, one industry and exogenous wages. They both show uniqueness under the assumption that there are no corner allocations. We extend their result to a more general environment, introduce the key condition that the product of the trade and scale elasticities is weakly lower than one in all industries, and allow for corner allocations.

Our welfare analysis is related to that in Costinot and Rodríguez-Clare (2014), who compute gains from trade and gains from trade liberalization for multi-industry economies under perfect and monopolistic competition. Compared to that paper, we further establish analytically that all countries gain from trade as long as the product of the trade and scale elasticities is weakly lower than one in all industries, we connect a country's decline in the gains from trade to its degree of industry specialization, we analyze how varying scale elasticities across industries interact with a country's inter-industry trade pattern to affect its gains from trade, and we connect the results for the gains from trade liberalization to the insights in Venables (1987).

Our paper is also related to Somale (2017), who introduces sector-specific innovation into a multi-sector Eaton and Kortum (2002) model (via mechanisms from Eaton and Kortum, 2001) to quantify its implications for welfare. Interestingly, although the model in Somale (2017) is dynamic, the balanced growth path is also characterized by the same system of equations as all the models that we consider in this paper, and so our results extend to this case as well. Somale (2017) also studies the quantitative importance of scale economies in determining industry-level specialization, but whereas he focuses on the variance of comparative advantage, we compare direct measures of trade and specialization between the data and those that would arise in a counterfactual world where everything is the same except that there are no economies of scale.

## 2. A Ricardian Model with Marshallian Externalities

We extend the multi-industry EK model as in Costinot, Donaldson and Komunjer (2012, henceforth CDK) to allow for Marshallian externalities. Formally, there are $N$ countries indexed by $n$, $i$ and $l$, and $K$ industries or sectors indexed by $k$. Each industry is composed of a continuum of goods or varieties $\omega \in \Omega \equiv [0,1]$. The only factor of production is labor, which is immobile across countries and perfectly mobile across industries within a country. We use $\bar{L}_i$ and $w_i$ to denote the inelastic labor supply and the wage level in country $i$, respectively. Each country has a representative consumer with upper-tier Cobb-Douglas preferences with industry-level expenditure shares $\beta_{i,k} \in (0,1)$ for all $(i,k)$ with $\sum_{k=1}^{K} \beta_{i,k} = 1$ for all $i$, and lower-tier CES across varieties within an industry with elasticity of substitution $\sigma_k$. Trade costs are of the standard iceberg type, so that delivering a unit of any industry $k$ good from country $i$ to country $n$ requires shipping $\tau_{ni,k} \geq 1$ units of the good, with $\tau_{ii,k} = 1$ for all $i$ and all $k$ and $\tau_{nl,k} \leq \tau_{ni,k}\tau_{il,k}$ for all $n$, $l$, $i$ and $k$ (triangular inequality).

The production technology exhibits constant or increasing returns to scale due to national external economies of scale at the industry level (i.e., Marshallian externalities). In particular, labor productivity for good $\omega$ in industry $(i,k)$ is $z_{i,k}(\omega)L_{i,k}^{\phi_k}$, where $z_{i,k}(\omega)$ is an exogenous productivity parameter, $L_{i,k}$ is the total labor allocated to industry $(i,k)$, and $\phi_k$ is the industry-specific parameter that governs the strength of Marshallian externalities (i.e., the elasticity of industry productivity with respect to industry size), which we refer to as the *scale elasticity* in industry $k$. We model $z_{i,k}(\omega)$ as in EK:

$z_{i,k}(\omega)$ is independently drawn from a Fréchet distribution with shape parameter $\theta_k$ and scale parameter $T_{i,k}$.

There is perfect competition, and the positive effect of industry size on productivity, $L_{i,k}^{\phi_k}$, is external to the firm. Thus, firms take as given both prices and unit costs, which are given by $c_{ni,k}(\omega) = \frac{\tau_{ni,k} w_i}{z_{i,k}(\omega) L_{i,k}^{\phi_k}}$. This implies that $p_{ni,k}(\omega) = c_{ni,k}(\omega)$. Since consumers can shop for the best deal around the world, prices must satisfy $p_{n,k}(\omega) = \min_{1 \le i \le N} \{p_{ni,k}(\omega)\}$.

Let $X_{ni,k}$ denote the total expenditure of country $n$ on goods in industry $(i,k)$, and let $\lambda_{ni,k} \equiv X_{ni,k} \Big/ \sum_{l=1}^{N} X_{nl.k}$ denote industry-level bilateral trade shares. Following the procedure in EK yields

$$\lambda_{ni,k}(\boldsymbol{w}, \boldsymbol{L}_k) = \frac{T_{i,k} L_{i,k}^{\alpha_k} \left(w_i \tau_{ni,k}\right)^{-\theta_k}}{\sum_l T_{l,k} L_{l,k}^{\alpha_k} \left(w_l \tau_{nl,k}\right)^{-\theta_k}}, \tag{1}$$

where $\boldsymbol{w} \equiv (w_1, ..., w_N)$ is the vector of wages, $\boldsymbol{L}_k \equiv \left(L_{1,k}, \dots, L_{N,k}\right)$ is the vector of labor allocations to industry $k$ across all countries, and $\alpha_k \equiv \theta_k \phi_k$. In turn, the price index for industry $k$ in country $n$ is

$$P_{n,k} = \gamma_k \left( \sum_l T_{l,k} L_{l,k}^{\alpha_k} \left(w_l \tau_{nl,k}\right)^{-\theta_k} \right)^{-1/\theta_k}, \tag{2}$$

and the aggregate price index is $P_n = \widetilde{\beta}_n \prod_{k=1}^{K} P_{n,k}^{\beta_{n,k}}$, where $\gamma_k$ and $\widetilde{\beta}_n$ are some constants.[5]

To write down the equilibrium condition for each industry $(i,k)$, it is important to note that Marshallian externalities imply that — in contrast to the CDK model — some industries could have zero labor allocations (i.e., $L_{i,k} = 0$). Therefore, we need to be careful when formulating the market equilibrium conditions in our model. By analogy with the standard Ricardian model, the complementary slackness conditions require the price to be weakly lower than the unit cost, with equality if there is positive production in the industry. Multiplying both the price and the unit cost by labor productivity (adjusted by trade costs), this is the same as requiring that revenue per worker be weakly lower than the wage, with equality if there is positive employment in the industry, i.e.,

[5]Specifically, $\gamma_k \equiv \Gamma\left(\frac{1-\sigma_k+\theta_k}{\theta_k}\right)^{\frac{1}{1-\sigma_k}}$, where $\Gamma$ is the Gamma function, and $\widetilde{\beta}_n$ is the standard Cobb-Douglas term $\widetilde{\beta}_n \equiv \prod_k \beta_{n,k}^{-\beta_{n,k}}$.

$$G_{i,k}\left(\boldsymbol{w},\boldsymbol{L}_k\right) \equiv w_i - \frac{1}{L_{i,k}}\sum_n \lambda_{ni,k}(\boldsymbol{w},\boldsymbol{L}_k)\beta_{n,k} w_n \bar{L}_n \geq 0. \tag{3}$$

In Appendix A we provide a formal definition of function $G_{i,k}\left(\boldsymbol{w},\boldsymbol{L}_k\right)$ that guarantees that $G_{i,k}$ is well-defined even for $(\boldsymbol{w},\boldsymbol{L}_k)$ with $L_{i,k}=0$. The equilibrium labor allocation in industry $(i,k)$ must satisfy the following complementary slackness condition:

$$L_{i,k} \geq 0, \quad G_{i,k}\left(\boldsymbol{w},\boldsymbol{L}_k\right) \geq 0, \quad L_{i,k} G_{i,k}\left(\boldsymbol{w},\boldsymbol{L}_k\right) = 0. \tag{4}$$

Note that the last equality is the standard industry clearing condition from CDK. And since in that model the equilibrium labor allocations are strictly positive for all $(i,k)$, the first and second inequalities above are automatically satisfied in any equilibrium.

The labor-market clearing condition for any country $i$ is just the same as in CDK:

$$\sum_k L_{i,k} = \bar{L}_i. \tag{5}$$

Denote by $\boldsymbol{L} \equiv (\boldsymbol{L}_1,\ldots,\boldsymbol{L}_k)$ the vector of labor allocations across industries. The equilibrium of the economy is a wage vector and labor allocation $(\boldsymbol{w},\boldsymbol{L}) \in \mathbb{R}^N_{++} \times \left(\mathbb{R}^{NK}_{+} \setminus \mathbb{Z}^{NK}_0\right)$ such that (4) holds for all $(i,k)$ and (5) holds for all $i$, where

$$\mathbb{Z}^{NK}_0 \equiv \left\{(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_K) \in \mathbb{R}^{NK}_{+} : \boldsymbol{x}_k = \boldsymbol{0} \text{ for some } k\right\}.$$

Restricting to $\boldsymbol{L} \notin \mathbb{Z}^{NK}_0$ is justified by the fact that there is no equilibrium in which no country allocates labor to some industry (i.e., $L_{i,k}=0$ for all $i$).

Before establishing existence and uniqueness for our model with Marshallian externalities, we first show isomorphisms between our model and other workhorse models in trade. A consequence is that our existence and uniqueness results immediately apply to these settings as well.

## 3. A Common Framework

In this section we discuss how generalized multi-sector versions of the models in Krugman (1980) and Melitz (2003) lead to the same equilibrium conditions as the model of the previous section, except of course that the trade and scale elasticities correspond to different structural parameters. At the end of the section we introduce new notation for the trade and scale elasticities ($\varepsilon_k$ and $\psi_k$ respectively) that is valid across models with different micro-foundations.[6]

As we discuss below, a multi-sector version of Krugman (1980) would imply that $\alpha_k = 1$ for all $k$. To add flexibility and allow for $\alpha_k \neq 1$, we assume that the elasticity of substitution across varieties from different countries can differ from the elasticity of substitution across varieties from the same country. Specifically, there is a continuum of differentiated varieties within each industry, and preferences are multi-tiered: Cobb-Douglas across industries with weights $\beta_{i,k}$, CES across country bundles within an industry with elasticity of substitution $\eta_k$, and CES across varieties within a country bundle with elasticity of substitution $\sigma_k > 1$.[7] Everything else is as in the standard Krugman model: labor is the only factor of production and the total labor cost of producing quantity $q$ of any variety in industry $(i,k)$ is $F_{i,k} + q/A_{i,k}$, where labor productivity $A_{i,k}$ and fixed cost $F_{i,k}$ are exogenous but can vary across industry-country pairs.

In Appendix B we show that the expressions for trade shares and industry price indices in this generalized-Krugman model collapse to those in equations (1) and (2) but now with trade and scale elasticities given by $\eta_k - 1$ and $(\sigma_k - 1)^{-1}$, respectively, and $\alpha_k \equiv \frac{\eta_k - 1}{\sigma_k - 1}$. If $\sigma_k = \eta_k$ for all $k$ then this is just the standard multi-industry Krugman model and $\alpha_k = 1$.[8] See rows 2 and 4 of Table 1.[9]

Now consider the Melitz (2003) model with Pareto distributed productivity and the same preferences as in the Krugman model above. After paying a fixed "entry" cost $F_{i,k}$ in units of labor in country $i$, firms are able to produce a variety in industry $(i,k)$ with labor productivity drawn from a Pareto distribution with shape parameter $\theta_k > \sigma_k - 1$ and location parameter $b_{i,k}$. Firms from $i$ can then pay a fixed "marketing" cost $f_{n,k}$ in

[6] In a similar spirit, Helpman and Krugman (1985) provide an integrated framework for international trade in the presence of different returns to scale and market structures. While their analysis is theoretical and limited to the case of frictionless trade, we focus on gravity models and the quantitative implications of scale economies for trade flows and welfare.

[7] Feenstra *et al.* (2014) also consider a multi-industry Melitz-Pareto model with possibly different elasticities of substitution across varieties from different countries and across varieties from the same country.

[8] Note also that if $\sigma_k \to \infty$, then $(\eta_k - 1)/(\sigma_k - 1) \to 0$ and we obtain the multi-industry Armington model.

[9] Is straightforward to incorporate Marshallian externalities into the multi-industry Krugman model presented above. For instance, letting $A_{i,k} \equiv \tilde{A}_{i,k} L_{i,k}^{\phi_k}$ leads to scale and trade elasticities $(\sigma_k - 1)^{-1} + \phi_k$ and $\eta_k - 1$, respectively.

units of labor of $n$ to serve that market.[10,11] As shown in Appendix B, the expressions for trade shares and industry price indices in this model collapse to those in equations (1) and (2) but now with trade and scale elasticities given by $\frac{\theta_k}{1+\theta_k\left(\frac{1}{\eta_k-1}-\frac{1}{\sigma_k-1}\right)}$ and $1/\theta_k$, respectively. Again, if $\sigma_k = \eta_k$ for all $k$ then this is just the standard multi-sector Melitz-Pareto model and $\alpha_k = 1$ for all $k$. See rows 3 and 5 in Table 1.

In Appendix B we show that both the generalized multi-sector Krugman and Melitz models lead to an equilibrium system that is equivalent to the one for the model presented in Section 2, except for constants that play no role for equilibrium analysis and comparative statics.[12] We can then think of a common framework that nests the multi-sector EK model with Marshallian externalities of Section 2 as well as the Krugman and Melitz models presented in the current section. We refer to this common framework in the remainder of the paper, and use notation $\varepsilon_k$ and $\psi_k$ to denote the trade and scale elasticities, $S_{i,k}$ to capture industry productivity, and $\mu_k$ for the constant affecting the price index. With $\alpha_k \equiv \varepsilon_k \psi_k$, the expressions for trade shares and price indices are now

$$\lambda_{ni,k}(\boldsymbol{w}, \boldsymbol{L}_k) = \frac{S_{i,k} L_{i,k}^{\alpha_k} \left(w_i \tau_{ni,k}\right)^{-\varepsilon_k}}{\sum_l S_{l,k} L_{l,k}^{\alpha_k} \left(w_l \tau_{nl,k}\right)^{-\varepsilon_k}} \tag{6}$$

and

$$P_{n,k} = \mu_k \left( \sum_l S_{l,k} L_{l,k}^{\alpha_k} \left(w_l \tau_{nl,k}\right)^{-\varepsilon_k} \right)^{-1/\varepsilon_k}. \tag{7}$$

## 4. Characterizing Equilibrium

We now turn to characterizing the equilibrium of the multi-sector gravity model with industry-level economies of scale. The analysis proceeds in two steps: we first charac-

[10] To simplify the analysis, we assume that the fixed marketing cost to serve destination $n$ does not vary across origins $i$. Allowing these fixed costs to vary across country pairs would not change any of our main conclusions below.

[11] The assumption that fixed marketing costs are paid in units of labor of the destination country is critical for the result that this model collapses to the general structure introduced above. This is related to the discussion in ACR about how their macro-level restriction R3' obtains in the Melitz-Pareto model if and only if the fixed cost is paid in units of labor of the destination country.

[12] This is true in the Melitz model under the balanced trade assumption. If a comparative statics exercise involves changing trade deficits, then the outcome of the Melitz model would not be equivalent to the outcome of the other two models. We ignore this issue in the analysis that follows and recognize that, whenever our comparative static exercises involve changing trade deficits, the results apply only to the Ricardian and Krugman models.

| Model | Trade elasticity, $\varepsilon_k$ | Scale elasticity, $\psi_k$ | $\alpha_k = \varepsilon_k \psi_k$ |
|---|---|---|---|
| CDK with ME | $\theta_k$ | $\phi_k$ | $\theta_k \phi_k$ |
| Multi-Sector Krugman | $\sigma_k - 1$ | $\frac{1}{\sigma_k - 1}$ | 1 |
| Multi-Sector Melitz-Pareto Model | $\theta_k$ | $\frac{1}{\theta_k}$ | 1 |
| Generalized Multi-Sector Krugman | $\eta_k - 1$ | $\frac{1}{\sigma_k - 1}$ | $\frac{\eta_k - 1}{\sigma_k - 1}$ |
| Generalized Multi-Sector Melitz-Pareto | $\frac{\theta_k}{1+\theta_k\left(\frac{1}{\eta_k-1}-\frac{1}{\sigma_k-1}\right)}$ | $\frac{1}{\theta_k}$ | $\frac{1}{1+\theta_k\left(\frac{1}{\eta_k-1}-\frac{1}{\sigma_k-1}\right)}$ |

Table 1: Mapping to Different Models

terize the equilibrium labor allocations given wages, and then we characterize wages that satisfy labor market clearing given the corresponding equilibrium labor allocations.

**Two-Step Equilibrium Definition**. The equilibrium labor allocations for some wage vector $\boldsymbol{w} \in \mathbb{R}^N_{++}$ are given by $\boldsymbol{L} \in \mathbb{R}^{NK}_{+} \setminus \mathbb{Z}^{NK}_0$ that satisfy (4) for all $(i,k)$. Let $\mathscr{L}(\boldsymbol{w})$ be the set of such equilibrium allocations. A wage vector $\boldsymbol{w} \in \mathbb{R}^N_{++}$ is an equilibrium wage vector if there exists an element $\boldsymbol{L} \in \mathscr{L}(\boldsymbol{w})$ such that $\boldsymbol{L}$ also satisfies (5) for all $i$.

Note that given wages, for each industry $k$ we have a system of $N$ nonlinear complementary slackness conditions in $L_{i,k}$ for $i = 1,...,N$ specified by (4). For the first step we exploit the fact that this system is independent across $k$. We now introduce some additional notation and definitions.

**Interior, Corner and Complete Specialization Allocations**. An allocation $\boldsymbol{L}_k$ is an *interior* allocation if $L_{i,k} > 0$ for all $i$; an allocation $\boldsymbol{L}_k$ is a *corner* allocation if $L_{i,k} = 0$ for at least one $i$; and an allocation $\boldsymbol{L}_k$ is a *complete specialization* allocation if there is a unique $i^*(k)$ such that $L_{i,k} = 0$ for all $i \neq i^*(k)$.

**Industry-Level Equilibrium Labor Allocations**. Given wage $\boldsymbol{w}$, $\mathscr{L}_k(\boldsymbol{w})$ denotes the set of equilibrium labor allocations in industry $k$, i.e., for any $\boldsymbol{L}_k \in \mathscr{L}_k(\boldsymbol{w})$ , $\boldsymbol{L}_k$ satisfies complementary slackness conditions (4) for industry $k$.

### 4.1. Step 1: Equilibrium Labor Allocations

Before we proceed, let us introduce an additional assumption on the matrix of trade costs which we employ to prove our results in the case of $\alpha_k = 1$ for some $k$:

**Assumption 1.** *The matrix*

$$\begin{pmatrix} \tau_{11,k}^{-\varepsilon_k} & \cdots & \tau_{1N,k}^{-\varepsilon_k} \\ \vdots & & \vdots \\ \tau_{N1,k}^{-\varepsilon_k} & \cdots & \tau_{NN,k}^{-\varepsilon_k} \end{pmatrix}$$

*is non-singular.*

We discuss the role of this assumption at the end of this subsection. For now, note that this assumption is violated if trade is free (i.e., $\tau_{ni,k} = 1$ for all $n$ and $i$).

Given the previous definitions, we are now ready to state our first Proposition.

**Proposition 1.** *If either (a) $0 \leq \alpha_k < 1$, or (b) $\alpha_k = 1$ and Assumption 1 holds, then the set $\mathcal{L}_k(\boldsymbol{w})$ is a singleton; if $\alpha_k > 1$, then the set $\mathcal{L}_k(\boldsymbol{w})$ contains multiple allocations, including (but not necessarily limited to) one for each complete specialization allocation. Moreover, the unique allocation in $\mathcal{L}_k(\boldsymbol{w})$ is an interior allocation if $0 \leq \alpha_k < 1$, while it may be an interior or a corner allocation if $\alpha_k = 1$.*

The proof of this proposition for the standard CDK case with $\alpha_k = 0$ is trivial: given wages, labor allocations are explicitly obtained from the conditions $L_{i,k} G_{i,k}(\boldsymbol{w}, \boldsymbol{L}_k) = 0$. In the proof of the case with $\alpha_k \in (0,1]$ we follow a popular approach in the economics literature that consists of characterizing equilibria of general equilibrium models as solutions to optimization problems. The details of this proof are in Appendix C.1. In the case with $\alpha_k > 1$ it is easy to see that there are many solutions to (4). For example, choose some $i^*$ and set $L_{i^*,k} = \frac{1}{w_{i^*}} \sum_n \beta_{n,k} w_n \bar{L}_{n,k}$ and $L_{i,k} = 0$ for $i \neq i^*$. These labor allocations satisfy all conditions in (4), because $G_{i^*,k}(\boldsymbol{w}, \boldsymbol{L}_k) = 0$ and $G_{i,k}(\boldsymbol{w}, \boldsymbol{L}_k) = w_i \geq 0$ for all $i \neq i^*$. Of course, this is only one example and there are many other possible equilibrium allocations in the case with $\alpha_k > 1$.[13]

[13] Forces that lead to increasing unit costs as a sector expands would help in ensuring uniqueness. As an example, in Appendix B.3 we show that allowing for labor heterogeneity as in the Roy-Fréchet specification of Lagakos and Waugh (2013) — see also Galle, Rodríguez-Clare and Yi (2017) — results in a model that is isomorphic to the common framework presented in Section 3 but with a lower scale elasticity, leading to less stringent conditions for uniqueness.

Focusing on the Ricardian specification to provide some intuition for Proposition 1, we can think of a battle of two forces that are activated as the labor allocation to an industry vanishes. The first is intra-industry Fréchet heterogeneity — a force for cross-country industry diversification with strength $1/\theta_k$ — which implies that labor productivity increases to infinity as fewer and fewer goods are produced in an industry. The second is Marshallian externalities — a force for cross-country industry specialization with strength $\phi_k$ — which implies that productivity falls to zero as industry employment falls to zero. For $\alpha_k \in (0,1)$ we have that $1/\theta_k > \phi_k$ and the force for diversification dominates the one for specialization, implying a unique and strictly positive equilibrium labor allocation across countries.[14] For $\alpha_k > 1$ we have that $1/\theta_k < \phi_k$ and the converse holds, implying multiple equilibria since now complete specialization in any country is an equilibrium.[15] Finally, for $\alpha_k = 1$ the two forces are in balance so uniqueness requires an additional force of industry diversification in the form of strictly positive trade costs, as guaranteed by Assumption 1.

We finish this subsection by commenting on the role of Assumption 1 in Proposition 1.[16] While Assumption 1 is not required for uniqueness when $\alpha_k < 1$, we cannot rule out multiplicity of equilibria if it is violated when $\alpha_k = 1$. As mentioned above, this assumption is violated if trade is frictionless. In fact, it is violated if there is at least one pair of countries with frictionless trade (in both directions) between them. Suppose, for instance, that there were no trade costs between two countries $i$ and $j$. The triangular inequality implies the trade costs between $i$ and $j$ and all other countries are the same (i.e., $\tau_{ni,k} = \tau_{nj,k}$ and $\tau_{in,k} = \tau_{jn,k}$ for all $n \neq i, j$). It then follows that the $i$ and $j$ row in the matrix of Assumption 1 are the same, so the non-singularity requirement is violated. Notice, however, that the multiplicity that arises in this case is that at most the overall labor allocation $L_{i,k} + L_{j,k}$ is determined, but not $L_{i,k}$ or $L_{j,k}$. This type of non-uniqueness is similar to the one that can arise in Ricardian models under frictionless trade and is irrelevant for welfare: real wages are the same across any two equilibria

[14] Two assumptions in the EK framework are important here: first, that the Fréchet distribution is unbounded above — this is critical to have labor productivity grow to infinity as the industry shrinks; and second, that there is a continuum of goods within each industry is also of consequence — with a finite number of goods the maximum productivity would be finite, and there would be multiple equilibria.

[15] An important example of this obtains when $1/\theta_k \to 0$, which leads to a degenerate productivity distribution across goods within an industry, and our model collapses to the classic Ricardian model with Marshallian externalities.

[16] In Appendix C.1 we discuss sufficient conditions under which Assumption 1 holds.

in this set. Moreover, with any small trade costs between $i$ and $j$ the non-uniqueness disappears, rendering these cases non-generic.

### 4.2. Step 2: Equilibrium Wages

In the cases with $0 \le \alpha_k < 1$ or with $\alpha_k = 1$ and Assumption 1 satisfied, Proposition 1 implies that the solution of the system of complementary slackness conditions (4) determines a univalent function from wages to labor allocations, $\boldsymbol{L}_k(\boldsymbol{w})$ for $\boldsymbol{w} \in \mathbb{R}^N_{++}$. When we prove uniqueness of equilibrium below, we use this function to construct the labor excess demand system and to show that there is only one wage vector that clears all labor markets. Assumption 1 is a regularity assumption that helps us establish uniqueness in the case of $\alpha_k = 1$, but its violation does not affect existence of equilibrium. In fact, it is possible to show equilibrium existence for any set of non-negative $\alpha_k$ without relying on additional assumptions. The key result that can be invoked to establish existence is Theorem 8 from Debreu (1982). Applying this theorem requires checking a number of (standard) technical conditions about the labor excess demand system. Among these conditions the key one is upper-hemicontinuity. For industries with $\alpha_k \in [0,1]$ we prove upper-hemicontinuity by exploiting the equivalence between the system in (4) and a constrained optimization problem and invoking the Theorem of the Maximum from Stokey, Lucas and Prescott (1989). For each industry with $\alpha_k \ge 1$ we explicitly construct an equilibrium labor allocation with the required properties. Our existence result is summarized in the following proposition with all the technical details left to Appendix C:

**Proposition 2.** *If $\alpha_k \ge 0$ for all $k$ then an equilibrium exists.*

Our key results concerning uniqueness of equilibrium are stated in the following three propositions (proofs are in Appendix C).

**Proposition 3.** *If there is an industry $k$ with $\alpha_k > 1$, then there are multiple equilibria.*

**Proposition 4.** *Assume that $N = 2$ and that for all $k$ either (a) $0 \le \alpha_k < 1$, or (b) $\alpha_k = 1$ and Assumption 1 holds. Then there is a unique equilibrium.*

**Proposition 5.** *Assume that $0 \le \alpha_k < 1$ for all $k$ and trade is frictionless in all industries, i.e., that $\tau_{ni,k} = 1$ for all $n$, $i$, and $k$. Then there is a unique equilibrium.*

**Proposition 6.** *Assume that* $0 \leq \alpha_k \leq 1$ *for all* $k$*. Furthermore, assume that a particular country is a small open economy in the sense that changes in its labor allocations and wage do not impact labor allocations, price indices, and wages in other countries. Then the economy of this country has a unique equilibrium.*

Proposition 3 is proven by directly showing that for industries with $\alpha_k > 1$ there exist several complete specialization labor allocations with corresponding wage vectors that clear labor markets. Propositions 4-6 are proven by first showing that, under the assumptions invoked in those propositions, the excess demand for labor in country $i$ is a well-defined function of wages: $Z_i(\boldsymbol{w}) \equiv \sum_k L_{i,k}(\boldsymbol{w}) - \bar{L}_i$. We then show that the assumptions in Propositions 4 and 5 also imply that the labor excess demand system, $\boldsymbol{Z}(\boldsymbol{w}) \equiv (Z_1(\boldsymbol{w}), ..., Z_N(\boldsymbol{w}))$, satisfies the gross substitutes property, while the small open economy assumption of Proposition 6 implies that the excess demand for labor in the small economy is a decreasing function of wages. Uniqueness of equilibrium then follows from Proposition 17.F.3 from Mas-Colell, Whinston and Green (1995).[17]

The challenge in extending the uniqueness result beyond Propositions 4-6 is that, with positive trade costs and more than two countries, $\boldsymbol{Z}(\boldsymbol{w})$ does not satisfy the gross substitutes property. While scale economies act to reinforce the gross substitutes property when there are two countries, the same is not necessarily true for three or more countries. For instance, a rise in the wage in one country, say country 1, may reduce the demand for labor there, while at the same time raising the demand for labor in another country, say country 2, which is so far consistent with the gross substitutes property. The complexities arise because the increased labor demand in country 2 can generate productivity effects that can lead to increased exports to a third country, say country 3, which can, in turn, result in a fall in the demand for labor there. In other words, a rise in wages in country 1 can result in a fall in the demand for labor in country 3, thereby violating the gross substitutes property.

[17] In addition to the equilibrium uniqueness results of Propositions 4-6, Proposition 1 — uniqueness of labor allocations when wages are fixed — can also be considered as an equilibrium uniqueness result for an economy subject to standard assumptions that pin down wages. These assumptions are: (1) there is a freely traded "outside good" industry; (2) production in this industry exhibits constant returns to scale; (3) all countries produce a positive amount of the outside good. While the result of Proposition 1 is valid for any finite $N$ and admits non-zero trade costs, the requirement that all countries produce the outside good is critical for us to establish uniqueness. If some countries do not produce the outside good, then wages are not pinned down, and we do not have a proof of uniqueness for more than two countries even in this case.

A more powerful approach for proving that the equilibrium is unique is the Index Theorem, which — roughly speaking — says that if the determinant of the negative of the Jacobian of the normalized excess labor demand system (i.e., the Jacobian of $-\boldsymbol{Z}(\boldsymbol{w})$ after removing the last column and the last row) is positive at any equilibrium wage vector then there is at most one equilibrium (Kehoe, 1980). The challenge here is that the Jacobian of the aggregate labor demand is the sum of the Jacobians of the labor demand coming from each sector (i.e., $D\boldsymbol{Z}(\boldsymbol{w}) = \sum_k D\boldsymbol{L}_k(\boldsymbol{w})$), and establishing conditions on the determinant of a sum of matrices is extremely difficult. Indeed, we have a proof that the determinant of the negative of the normalized sector-level Jacobians (i.e., $-D\boldsymbol{L}_k(\boldsymbol{w})$ after removing the last column and the last row) is always positive, but we have not been able to extend this property to the sum of those modified Jacobians. It is natural to look for some property of matrices that implies a positive determinant and that survives under summation and check if this property is satisfied by the negative of the normalized sector-level Jacobians. Gross substitutes is one such property, but, as mentioned above, it does not hold for more than two countries.[18] Another such property is positive definiteness, but unfortunately our sector-level Jacobians are non-symmetric matrices and hence not positive definite.[19] In fact, even for a more general definition of positive definiteness, which applies to non-symmetric matrices, we have examples in which the negative of the normalized sector-level Jacobians do not satisfy this definition.

An entirely different approach is to transform the equilibrium system into a mapping whose fixed point is a solution of that system, and then show that this mapping is a contraction mapping. In principle, if there is a unique equilibrium, then such a contraction mapping exists, but finding the right transformation is of course very challenging. We view the techniques explored in Allen, Arkolakis and Li (2016) as a variant of this approach. In Appendix C.7 we show how the equilibrium system of our economy with $\alpha_k = \alpha$ for all $k$ can be transformed to the kind of system for which Allen, Arkolakis and Li (2016) establish sufficient conditions for uniqueness. Unfortunately, we find that

[18]Mathematically, gross substitutes implies that the negative of the normalized Jacobian is a diagonally dominant $Z$-matrix. This property survives under summation and implies a positive determinant.

[19]Another possibility is to apply Minkowski's Determinant Inequality (see, for example, Horn and Johnson (2013), Theorem 7.8.21) which establishes a lower bound for the determinant of a sum of matrices in terms of determinants of individual matrices. Unfortunately, again, the result only applies to (Hermitian) positive definite matrices.

those sufficient conditions are not satisfied in our economy whenever $\alpha_k \geq 0$.

Lacking a proof, we have systematically looked for counter-examples via numerical analysis, but we have found none. We report extensively on this work in Appendix C.8. In particular, in Appendix C.8.2 and in the online appendix we describe an algorithm based on interval analysis that allows us to rigorously find all solutions to our equilibrium system of equations for any fixed parametrization. That is, if there are multiple equilibria for a particular set of parameters, our algorithm is guaranteed to find them all. Conversely, if our algorithm finds only one equilibrium, it implies that we have a proof of uniqueness for a particular parametrization. This algorithm has a potential to be applied outside the scope of our framework.

## 5. Scale Economies and the Gains from Trade

In this section we explore the implications of scale economies for the gains from trade, defined as in ACR as the negative of the percentage change in real income as we move from the observed equilibrium to autarky. We first do so at a theoretical level and then use the World Input-Output Dataset (WIOD) to quantify the effects. In Section 6 we turn to the related but distinct question of how scale economies affect the gains from trade liberalization, defined as the welfare effects of a counterfactual decline in tariffs or trade costs starting at the observed equilibrium.

### 5.1. Theoretical Analysis

In principle, countries that specialize in industries with weak economies of scale could even lose from trade — the premise of Frank Graham's argument for protection. It turns out, however, that this cannot happen if $0 \leq \alpha_k \leq 1$ for all $k$. A simple argument establishes this result for an industry-country pair $(i,k)$ with $L_{i,k} > 0$. Using equations (6) and (7) and setting $w_i = 1$ by choice of numeraire, the industry-level price index is $P_{i,k} = \mu_k S_{i,k}^{-1/\varepsilon_k} L_{i,k}^{-\psi_k} \lambda_{ii,k}^{1/\varepsilon_k}$. Since autarky employment levels are $L_{i,k}^A = \beta_{i,k}\bar{L}_i$, autarky price indices are then $P_{i,k}^A = \mu_k S_{i,k}^{-1/\varepsilon_k} (\beta_{i,k}\bar{L}_i)^{-\psi_k}$. We can then write

$$P_{i,k} = P_{i,k}^A \left(\lambda_{ii,k}\right)^{1/\varepsilon_k} \left(\frac{L_{i,k}}{\beta_{i,k}\bar{L}_i}\right)^{-\psi_k}. \tag{8}$$

From (8) it is now evident that in the presence of scale economies there are two potentially countervailing forces acting on price indices: the standard gains from trade captured by $(\lambda_{ii,k})^{1/\varepsilon_k}$, and the potentially negative scale effect arising from a contraction in employment relative to autarky, which is captured by $\left(\frac{L_{i,k}}{\beta_{i,k}\bar{L}_i}\right)^{-\psi_k}$. Note, however, that in an equilibrium with $L_{i,k} > 0$ we must have $L_{i,k} > \lambda_{ii,k}\beta_{i,k}\bar{L}_i$, since the RHS is just total employment associated with domestic sales. This implies that $P_{i,k} < P^A_{i,k}(\lambda_{ii,k})^{1/\varepsilon_k - \psi_k}$, and hence $\psi_k \le 1/\varepsilon_k$ (or $\alpha_k \equiv \varepsilon_k \psi_k \le 1$) implies that $P_{i,k} < P^A_{i,k}$. In Appendix D.1 we show that this strict inequality also holds in the case in which $L_{i,k} = 0$, thus establishing the following result.

**Proposition 7.** *If* $0 \le \alpha_k \le 1$ *for all* $k$ *then all countries gain from trade.*

Intuitively, the strength of the standard gains from trade is regulated by the inverse of the trade elasticity, $1/\varepsilon_k$, while the strength of the potentially negative scale effects is regulated by $\psi_k$. For $\alpha_k \le 1$ the standard gains from trade always neutralize any potentially opposing scale effects that could lead to higher prices with trade than in autarky. On the contrary, if $\alpha_k > 1$, then one could have higher prices in some industries with trade than without, leading to the possibility of losses from trade.

Our Proposition 7 can be seen as a generalization of Proposition 1 in Venables (1987), which states that in a Krugman (1980) model with an “outside good" all countries gain from trade. Formally, the model in Venables (1987) is isomorphic to ours when we consider two countries and two industries, one having no trade costs, no scale economies, and an infinite trade elasticity (the “outside good"), and the other having trade costs, scale economies, and a finite trade elasticity, with $\alpha_k = 1$. Proposition 7 shows that this generalizes to a case without an outside good (i.e., with endogenous wages), with multiple sectors and microfoundations different than Krugman (e.g., Ricardian plus Marshallian externalities), and arbitrary scale economies as long as $\alpha_k \le 1$ for all $k$.

To further explore the implications of scale economies for the magnitude of the gains from trade, we assume that the equilibrium is interior so that all trade shares and labor allocations are strictly positive. This allows us to derive an expression for the gains from trade as a function of industry-level data and the *trade* and *scale* elasticities that extend the multi-sector expressions in ACR.[20]

[20] Given Proposition 1, the assumption that labor allocations are strictly positive is not restrictive for the case with $0 \le \alpha_k < 1$ for all $k$.

Equations (6) and (7) together with $P_n = \widetilde{\beta}_n \prod_{k=1}^{K} P_{n,k}^{\beta_{n,k}}$ imply that

$$W_n \equiv w_n / P_n \propto \mu_n \prod_k \left( S_{n,k} L_{n,k}^{\varepsilon_k \psi_k} \lambda_{nn,k}^{-1} \right)^{\beta_{n,k}/\varepsilon_k}.$$

Using hat notation, $\hat{x} = x'/x$, a foreign shock (i.e., a shock that does not affect the exogenous variables in country $n$) induces a change in welfare in country $n$ equal to

$$\hat{W}_n = \prod_k \hat{\lambda}_{nn,k}^{-\beta_{n,k}/\varepsilon_k} \cdot \prod_k \hat{L}_{n,k}^{\beta_{n,k}\psi_k}. \tag{9}$$

The first term on the RHS of this expression is the standard multi-industry formula for gains from trade (with upper-tier Cobb-Douglas preferences), while the second term is an adjustment for scale economies.

To better understand this expression, we can use the fact that the welfare effect of an infinitesimally small change in wages and prices is $d \ln W_n = d \ln w_n - \sum_k \beta_{n,k} d \ln P_{n,k}$. Totally log-differentiating $\lambda_{nn,k} = (L_{n,k}^{-\psi_k} w_n / P_{n,k})^{-\varepsilon_k}$ and substituting into the previous equation yields

$$d \ln W_n = -\sum_k \beta_{n,k} \frac{d \ln \lambda_{nn,k}}{\varepsilon_k} + \sum_k \beta_{n,k} \psi_k d \ln L_{n,k}. \tag{10}$$

The first term term on the right hand side captures the welfare effect of an infinitesimally small foreign shock taking home productivity as given, while the second term captures the welfare effect of that shock through home productivity changes caused by changing industry employment levels in the presence of scale effects. Integrating the first term over some discrete shock yields $\prod_k \hat{\lambda}_{nn,k}^{-\beta_{n,k}/\varepsilon_k}$, while integrating the second term yields $\prod_k \hat{L}_{n,k}^{\beta_{n,k}\psi_k}$.

Following ACR, we define the gains from trade as the negative of the percentage change in real income as we move from the observed equilibrium to autarky,

$$GT_n \equiv \frac{W_n - W_n^A}{W_n}.$$

We compute $GT_n$ by applying (9) and noting that for the move back to autarky we have $\hat{\lambda}_{nn,k} = 1/\lambda_{nn,k}$, and $\hat{L}_{n,k} = \beta_{n,k}/r_{n,k}$, where $r_{n,k} \equiv L_{n,k}/\bar{L}_n$ denotes the industry revenue (or employment) shares in the observed equilibrium. Using $e_{n,k} \equiv X_{n,k}/X_n$ for observed industry expenditure shares (of course, $e_{n,k} = \beta_{n,k}$ in the model), this leads to a formula for the gains from trade that depends only on the country's observables $\lambda_{nn,k}$, $e_{n,k}$ and $r_{n,k}$ as well as the trade and scale elasticities, $\varepsilon_k$ and $\psi_k$,

$$GT_n = 1 - \Delta_n \prod_k \lambda_{nn,k}^{e_{n,k}/\varepsilon_k}, \tag{11}$$

where

$$\Delta_n \equiv \prod_k \left(e_{n,k}/r_{n,k}\right)^{e_{n,k}\psi_k}.$$

The expression for the gains from trade in the standard perfectly competitive model with no scale economies obtains from (11) by setting $\psi_k = 0$ for all $k$, which leads to $\Delta_n = 1$. The effect of scale economies on the gains from trade then depends on whether $\Delta_n \gtrless 1$.

Consider first the case in which the scale elasticity is the same across industries ($\psi_k = \psi$ for all $k$) and note that

$$\Delta_n^{1/\psi} = \exp D_{KL}(\boldsymbol{e}_n \parallel \boldsymbol{r}_n),$$

where $\boldsymbol{r}_n \equiv (r_{n1}, ..., r_{nK})$, $\boldsymbol{e}_n \equiv (e_{n1}, ..., e_{nK})$, and

$$D_{KL}(\boldsymbol{e}_n \parallel \boldsymbol{r}_n) \equiv \sum_k e_{n,k} \ln(e_{n,k}/r_{n,k}) \tag{12}$$

is the Kullback-Leibler divergence of $\boldsymbol{r}_n$ from $\boldsymbol{e}_n$. We can think of $D_{KL}(\boldsymbol{e}_n \parallel \boldsymbol{r}_n)$ as a measure of industry specialization in country $n$ — in autarky we would have $\boldsymbol{r}_n = \boldsymbol{e}_n$ and $D_{KL}(\boldsymbol{e}_n \parallel \boldsymbol{r}_n) = 0$, while if $\boldsymbol{r}_n \neq \boldsymbol{e}_n$ then $D_{KL}(\boldsymbol{e}_n \parallel \boldsymbol{r}_n) > 0$. This implies that $\Delta_n > 1$ (except if $\boldsymbol{r}_n = \boldsymbol{e}_n$, in which case $\Delta_n = 1$) so that, *given trade shares*, scale economies actually *reduce* the gains from trade, with a larger decline for higher values of $\psi$ and for countries that exhibit higher levels of specialization.[21]

We can gain intuition about this result by going back to equation (10) and noting that if $\psi_k = \psi$ for all $k$ then the second term on the RHS of that equation can be written as $\psi \sum_k e_{n,k} \frac{dL_{n,k}}{L_{n,k}}$. A move back to autarky implies the expansion of industries with

[21] The opposite result would hold if instead of economies of scale we had diseconomies of scale. For example, in a setting with $\psi = 0$ and worker-level heterogeneity, Galle *et al.* (2017) show that $GT_n = 1 - \prod_k \lambda_{nn,k}^{e_{n,k}/\varepsilon_k} \left(e_{n,k}/r_{n,k}\right)^{-e_{n,k}/\kappa}$, where $\kappa$ is a parameter that determines the degree of heterogeneity. The argument above now implies that the gains from trade are *higher* than in the case with no scale economies, which obtains here in the limit as $\kappa \to \infty$, and corresponds to the case in which workers are homogeneous.

net imports and hence a high expenditure share or low employment. In either case, the expenditure-weighted productivity gain in expanding industries will be on average higher than the expenditure-weighted productivity loss in contracting industries, and hence $\psi \sum_k e_{n,k} \frac{dL_{n,k}}{L_{n,k}} > 0$. After integration, this leads to $\Delta_n > 1$.

Readers may be surprised by the result that the gains from trade are lower with economies of scale (i.e., $0 < \alpha < 1$) than without (i.e., $\alpha = 0$). As we mentioned above, this is because we are following the ACR approach of taking trade flows as data that is held fixed as we compare different models. The intuition that scale economies should lead to larger gains from trade through deeper industry-level specialization and larger trade flows is not operative here, although as we illustrate in Section 6 this intuition is relevant for the *gains from trade liberalization.*

Turning now to the more general case in which $\psi_k$ varies across $k$, $\Delta_n$ can be rewritten as

$$\Delta_n = \exp\left\{\bar{\psi}\,[DS - PS]\right\}, \tag{13}$$

where $DS \equiv D_{KL}(\boldsymbol{e}_n \parallel \boldsymbol{r}_n)$ and

$$PS \equiv \sum_k \frac{\psi_k - \bar{\psi}}{\bar{\psi}} \ln\left(\frac{r_{n,k}}{e_{n,k}}\right)^{e_{n,k}}, \tag{14}$$

with $\bar{\psi} \equiv (1/K)\sum_k \psi_k$. Expression (13) makes it clear that in general there are two (possibly competing) forces: the first measures the *degree of specialization* ($DS$) in country $n$, while the second measures the *pattern of specialization* ($PS$), i.e., the tendency of country $n$ to specialize in industries with either higher or lower than average scale economies. Since *DS* always pushes towards lower gains relative to the case with no scale economies, the overall effect of scale economies on gains from trade depends on the direction and magnitude of *PS*. Countries that tend to specialize in industries with lower than average scale economies — so that *PS* is negative — gain less from trade with scale economies than without. However, in countries that tend to specialize in industries with higher than average scale economies — so that *PS* is positive — the effect of scale economies on the gains from trade is ambiguous. If a country's *PS* is strong enough to overcome its *DS*, then such a country could have higher gains with than without scale economies.

## 5.2. Quantitative Analysis

We now explore the quantitative implications of the previous observations by computing the gains from trade using actual data. Using (11) as our reference point, we follow Costinot and Rodríguez-Clare (2014) and compute measures of $\lambda_{nn,k}$, $e_{n,k}$, and $r_{n,k}$ employing data on 31 sectors and 33 countries (plus the rest of the world aggregate) from the WIOD in 2008.[22] To focus on the role of scale economies, we assume throughout our quantitative analysis that there is a common trade elasticity of 5 for all industries (i.e., $\varepsilon_k = 5$ for all $k$).[23]

We start by assuming that there is a common scale elasticity across all industries and consider two possibilities: (i) no scale economies, $\psi_k = 0$ for all $k$; and (ii) scale economies, $\psi_k = 0.14$ for all $k$. Given a trade elasticity of 5, these two possibilities correspond to assuming $\alpha_k = 0$ for all $k$, and $\alpha_k = 0.7$ for all $k$. We choose $\alpha_k = 0.7$ in light of recent evidence in Bartelme *et al.* (2017, henceforth BCDR), Lashkaripour and Lugovskyy (2017), and Somale (2017).[24]

Columns 1 and 2 in Table 2 report the gains from trade for the two values of scale elasticities described in the previous paragraph, while column 4 reports the degree of industry specialization, $D_{KL}(\boldsymbol{e}_n \parallel \boldsymbol{r}_n)$. Consistent with Proposition 7, gains from trade decrease as we allow for stronger scale economies, and this decline is stronger for countries that have a higher degree of industry specialization. This is illustrated in Figure 1, which plots the gains from trade net of the standard ACR gains (i.e., the gains that would arise in the absence of scale economies) for $\psi_k = 0.14$, and $\psi_k = 0.2$. The latter was added for emphasis and corresponds to scale economies that would arise in a standard Krugman model.

Next, we study how heterogeneity in scale elasticities across industries affects the gains from trade. For the manufacturing industries we assign scale elasticities based on

[22] Equation (11) ignores trade deficits. As discussed in Costinot and Rodríguez-Clare (2014), this implies that our results in this subsection capture the change in real income rather than the change in real expenditure caused by shutting down trade and closing any trade deficits that exist in the data.

[23] We choose a value of 5 for the trade elasticity as this is a typical value used in the literature – see Head and Mayer (2014).

[24] Bartelme *et al.* (2017) use bilateral trade flows data between 61 countries and estimate that $\alpha = 0.68$ when data is pooled across 15 manufacturing industries. Lashkaripour and Lugovskyy (2017) use Colombian import transactions data and estimate that $\alpha = 0.6$ when pooling data across all manufacturing industries. Somale (2017) uses bilateral trade flows data and estimates that $\alpha \in [0.706, 0.811]$ for a sample of 29 mostly OECD countries and 18 industries.

| | Gains from Trade, % | | | | | | Gains from Trade, % | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\psi_k = 0$ | $\psi_k = 0.14$ | BCDR | $DS$ | $PS$ BCDR | | $\psi_k = 0$ | $\psi_k = 0.14$ | BCDR | $DS$ | $PS$ BCDR |
| Country | (1) | (2) | (3) | (4) | (5) | Country | (1) | (2) | (3) | (4) | (5) |
| AUS | 3.0 | 2.5 | 2.2 | 0.036 | −0.121 | IRL | 10.1 | 9.2 | 10.1 | 0.069 | 0.064 |
| AUT | 8.7 | 8.5 | 8.8 | 0.015 | 0.037 | ITA | 3.7 | 3.3 | 3.9 | 0.026 | 0.069 |
| BEL | 13.8 | 12.9 | 14.0 | 0.072 | 0.117 | JPN | 2.4 | 1.9 | 2.9 | 0.041 | 0.118 |
| BRA | 1.6 | 1.6 | 1.4 | 0.004 | −0.026 | KOR | 6.6 | 4.8 | 7.3 | 0.133 | 0.268 |
| CAN | 5.2 | 4.9 | 4.7 | 0.024 | −0.065 | MEX | 4.5 | 4.3 | 3.9 | 0.018 | −0.096 |
| CHN | 3.0 | 2.8 | 3.3 | 0.016 | 0.077 | NLD | 9.4 | 9.2 | 9.6 | 0.020 | 0.057 |
| CZE | 8.0 | 7.8 | 8.3 | 0.014 | 0.075 | POL | 6.0 | 5.9 | 5.8 | 0.009 | −0.025 |
| DEU | 6.0 | 5.6 | 6.7 | 0.033 | 0.157 | PRT | 6.3 | 5.9 | 5.9 | 0.032 | −0.042 |
| DNK | 9.5 | 8.3 | 8.8 | 0.090 | −0.032 | ROM | 6.2 | 5.9 | 5.3 | 0.020 | −0.134 |
| ESP | 4.1 | 3.7 | 3.9 | 0.023 | −0.003 | RUS | 3.2 | 2.2 | 1.4 | 0.076 | −0.245 |
| FIN | 5.5 | 5.0 | 6.2 | 0.036 | 0.169 | SVK | 11.5 | 11.1 | 11.9 | 0.038 | 0.102 |
| FRA | 4.0 | 3.8 | 4.1 | 0.020 | 0.033 | SVN | 13.3 | 11.9 | 12.0 | 0.112 | −0.143 |
| GBR | 4.6 | 4.5 | 4.4 | 0.008 | −0.034 | SWE | 6.7 | 6.5 | 7.2 | 0.019 | 0.114 |
| GRC | 5.7 | 4.2 | 4.1 | 0.110 | −0.182 | TUR | 4.0 | 3.8 | 4.0 | 0.018 | 0.017 |
| HUN | 12.3 | 11.9 | 12.1 | 0.027 | −0.007 | TWN | 9.6 | 8.1 | 10.2 | 0.112 | 0.225 |
| IDN | 3.5 | 3.4 | 3.2 | 0.013 | −0.052 | USA | 2.1 | 2.0 | 1.8 | 0.007 | −0.042 |
| IND | 3.1 | 2.9 | 3.1 | 0.013 | 0.001 | ROW | 6.9 | 6.3 | 5.9 | 0.050 | −0.131 |
| | | | | | | **Average** | 6.3 | 5.8 | 6.1 | 0.040 | 0.009 |

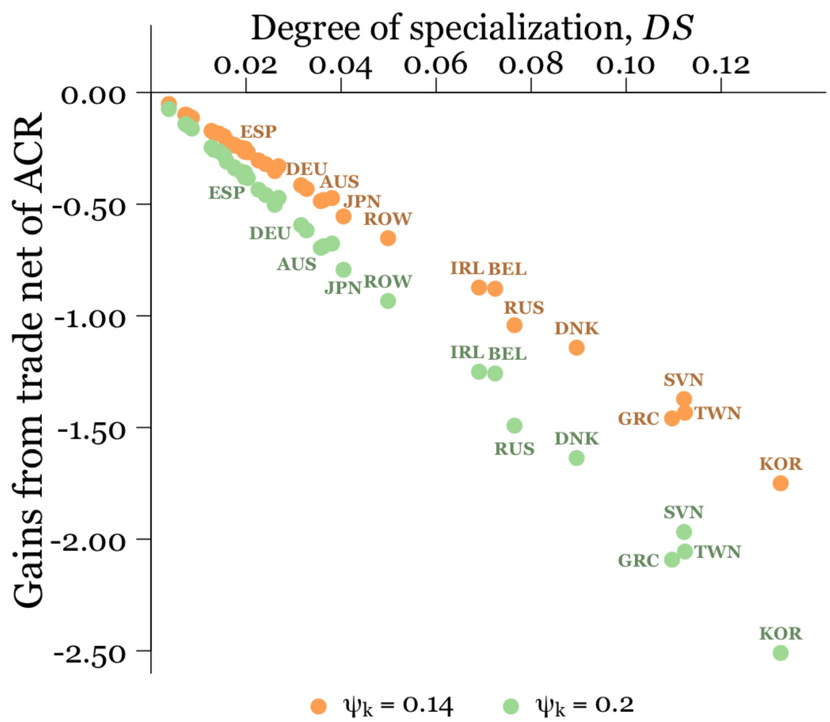
*Notes:* Gains from trade in columns 1-3 are calculated according to (11). For all columns 1-3, $\varepsilon_k = 5$ for all $k$. For columns 1 and 2, $\psi_k$ are set to 0 and 0.14 for all $k$, correspondingly. For column 3, $\psi_k$ for manufacturing industries are set based on estimates from Bartelme *et al.* (2017), and for non-manufacturing industries $\psi_k = 0$. $DS$ in column 4 is calculated according to expression (12) and does not depend on $\psi_k$ or $\varepsilon_k$. $PS$ in column 5 is calculated according to expression (14) and corresponds to the parametrization in column 3. Average values in the last row are calculated based on the full set of countries.

Table 2: Gains from Trade, Degree and Pattern of Specialization

the estimates obtained by BCDR. These elasticities are listed in Table 6 of Appendix D.5.[25] For the non-manufacturing industries we assume $\psi_k = 0$. The gains from trade and pattern of specialization ($PS$) corresponding to this case for a select set of countries are provided in columns 3 and 5 of Table 2, respectively. Note that since the degree of in-dustry specialization is calculated according to expression (12) and is independent of $\psi_k$, the results in column 4 also corresponds to this case.

The results in Table 2 illustrate the role of both the degree and pattern of specialization for the gains from trade, as highlighted in equation (13). When scale elasticities

[25] In the same appendix we explain how we use BCDR estimates to assign scale elasticities for the manufacturing industries in our exercises.

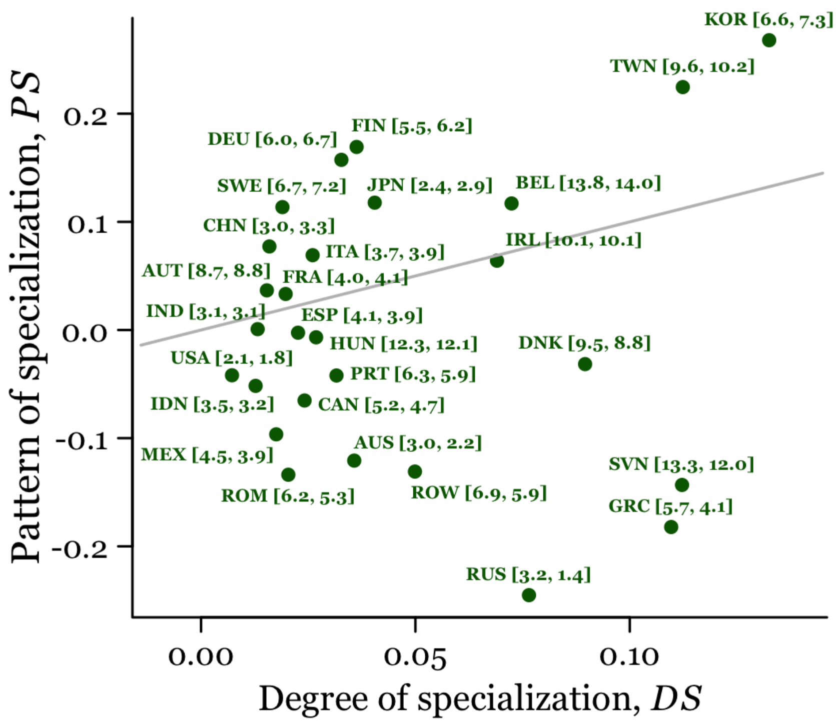



*Notes:* Gains from trade are calculated according to (11) with $\varepsilon_k = 5$ and $\psi_k = 0.14$ or 0.2 for all $k$. For ACR gains from trade $\psi_k = 0$. $DS$ is calculated according to (12) and does not depend on $\psi_k$.

Figure 1: Degree of Specialization and Gains from Trade with Common Scale Elasticities

vary across industries, it is possible that $PS$ is positive and high enough, so that the gains from trade are higher than those computed with the ACR formula. For example, comparing columns 1 and 3 we see that South Korea's and China's gains from trade increase from 6.6% to 7.3% and from 3% to 3.3%, respectively.

Results for the case corresponding to column 3 of Table 2 are also summarized in Figure 2, which plots $PS$ against the $DS$ for all countries in our sample. For each point we also report the name of the country, the standard ACR gains and the gains computed according to expression (11). Most countries lie below the 45 degree line, indicating that $PS < DS$, so that economies of scale lower the gains from trade. In contrast, $PS > DS$ for the countries above the 45 degree line. As we can see from the figure, in addition to China and South Korea, several other countries (for example, Germany, Japan and Taiwan) also have higher gains with economies of scale than without thanks to the pattern of specialization tilted in favor of sectors with relatively high scale elasticities.




*Notes:* $DS$ and $PS$ are calculated according to expressions (12) and (14), correspondingly. The numbers in parentheses are, respectively, the standard ACR gains, $GT^{ACR}$, and the gains from trade in the presence of economies of scale, $GT^{scale}$, both calculated according to (11). $PS$, $GT^{ACR}$ and $GT^{scale}$ are calculated by setting $\varepsilon_k = 5$ for all $k$. For $GT^{ACR}$, $\psi_k = 0$ for all $k$. For $PS$ and $GT^{scale}$, $\psi_k$) for manufacturing industries are set based on estimates from Bartelme *et al.* (2017), and $\psi_k = 0$ for non-manufacturing industries. The straight line is the 45 degree line. To avoid cluttering, only a subset of countries is included in the figure.

Figure 2: Degree of Specialization, Pattern of Specialization and Gains from Trade in the Presence of Economies of Scale

## 6. Scale Economies and the Gains from Trade Liberalization

Whereas the previous section was devoted to the gains from trade using autarky as the counterfactual, we now study the gains from a decline in trade barriers, focusing on how these gains are affected by scale economies. We again split the section in two parts: theory and quantitative analysis.

## 6.1. Theoretical Analysis

We consider two simple cases for which we can derive analytical results for the gains from a decline in trade costs. Both cases consist of two countries and two industries under conditions implying that relative wages are not affected by the trade shock.

### 6.1.1. Mirror-Image Countries

With two mirror-image countries the wage is equalized, and we can normalize both wages to one. For ease of exposition we index countries by $i = H, F$, where $H$ and $F$ represent Home and Foreign, respectively. Let $\bar{L} = 2$, $\beta_{i,k} = 1/2$ for all $(i,k)$, and let $S_{H,1} = S_{F,2} = 2$ and $S_{H,2} = S_{F,1} = 1$. Hence, Home has the comparative advantage in industry 1, and Foreign in industry 2. We assume that $\varepsilon_k = \varepsilon$ and $\psi_k = \psi$ for $k = 1,2$.

To establish a link with the results of Section 5, we first illustrate that the gains from trade are decreasing in $\psi$. We then show that the conclusion is reversed once we consider a trade liberalization exercise in which trade shares respond endogenously as we lower trade costs. There we find that the gains from trade liberalization are increasing in $\psi$.

Home's gains from trade are simply

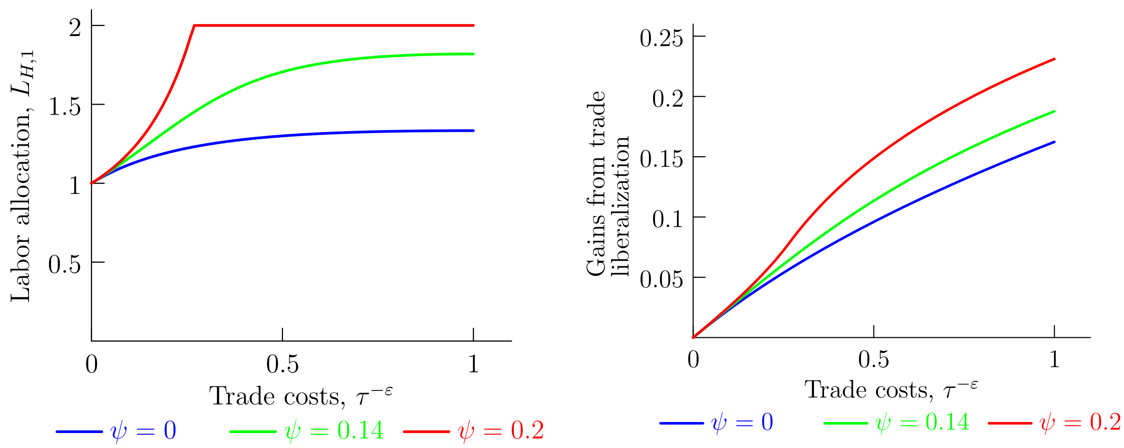
$$GT_H = 1 - \left(\frac{1/2}{r_{H,1}} \cdot \frac{1/2}{1 - r_{H,1}}\right)^{\psi/2} \left(\lambda_{HH,1}^{\frac{1}{2\varepsilon}} \cdot \lambda_{HH,2}^{\frac{1}{2\varepsilon}}\right).$$

The term $\left(\frac{1/2}{r_{H,1}} \cdot \frac{1/2}{1-r_{H,1}}\right)^{\psi/2}$ corresponds to $\Delta_H$ in Equation (11) and is higher than one as long as there is industry-level specialization.[26] Thus, given trade shares, gains are lower with scale effects ($\psi > 0$) than without ($\psi = 0$). It is also easy to see that these gains are decreasing in $\psi$.

Next we study the gains from trade liberalization, allowing for endogenous responses of trade shares to trade costs.[27] We use the autarky economy (i.e., the economy with $\tau = \infty$) as the baseline case and define gains from trade liberalization as welfare changes resulting from the move from autarky to an economy with a finite level of trade costs.

[26] The term $\frac{1/2}{r_{H,1}} \cdot \frac{1/2}{1-r_{H,1}}$ is minimized at $r_{H,1} = 1/2$, and specialization according to comparative advantage implies $r_{H,1} > 1/2$, hence we must have $\left(\frac{1/2}{r_{H,1}} \cdot \frac{1/2}{1-r_{H,1}}\right)^{\psi/2} > 1$.

[27] In the terminology of ACR, this corresponds to an "ex-ante analysis" whereas the results for the gains from trade above correspond to an "ex-post analysis".




(a) Industry specialization (b) Gains from trade liberalization

Figure 3: Effect of Economies of Scale

We set $\varepsilon = 5$ and, consistently with the exercises of Section 5, $\psi \in \{0, 0.14, 0.2\}$, which corresponds to $\alpha \in \{0, 0.7, 1\}$. In all these cases $L_{i,k} = 1$ for all $(i,k)$ under autarky. As $\tau$ falls from $\infty$, country $H$ specializes in industry 1 and country $F$ specializes in industry 2, but the extent of specialization will be stronger with $\psi = 0.2$ than $\psi = 0.14$, and with $\psi = 0.14$ than $\psi = 0$, as illustrated in Figure 3a. Figure 3b shows the implications for the gains from trade liberalization for each of these three cases. We see that the gains from trade liberalization increase with $\psi$. The intuition is simple: countries gain by specializing according to comparative advantage, and the concentration of production also allows for a greater exploitation of scale economies, which, in turn, generates additional efficiency gains.[28]

### 6.1.2. Outside Good

The result established in Proposition 7 that, in the region of uniqueness, countries always gain from trade does not necessarily imply that there are always gains from *further* trade liberalization. In fact, our model nests the one in Venables (1987), and

[28] To understand this further, note that the gains from trade liberalization can be seen as the increase in welfare $W_H = 1/(1 - GT_H) = \left(\frac{1/2}{r_{H,1}} \cdot \frac{1/2}{1-r_{H,1}}\right)^{-\psi/2} \left(\lambda_{HH,1}^{-1/2\varepsilon} \cdot \lambda_{HH,2}^{-1/2\varepsilon}\right)$ as $\tau$ falls. The decline in $\tau$ leads to deeper industry-level specialization, as captured by a higher $r_{H,1}$, and this decreases $\left(\frac{1/2}{r_{H,1}} \cdot \frac{1/2}{1-r_{H,1}}\right)^{-\psi/2}$ and decreases $W_H$. But there is also a change in trade shares, and this increases $\lambda_{HH,1}^{-1/2\varepsilon} \cdot \lambda_{HH,2}^{-1/2\varepsilon}$, which more than offsets the previous effect.

so we know that a unilateral decline in inward trade costs may decrease welfare. To see this more explicitly, consider a case with two countries and two industries, with $\varepsilon_1 = \infty > \varepsilon_2$, $\psi_1 = 0 < \psi_2 \leq 1/\varepsilon_2$ (so that $\alpha_2 \leq 1$), and with no trade costs in industry 1 (i.e., $\tau_{12,1} = \tau_{21,1} = 1$). If we start with an interior equilibrium (i.e., $L_{i,k} > 0$ for $i = 1,2$ and $k = 1,2$) then wages are pinned down by (exogenous) productivities in industry 1 (the *outside good*), and — suppressing the industry sub-index — the labor allocation in industry 2 is given by $(L_1, L_2)$ that solves

$$w_i L_i = \sum_n S_i L_i^{\alpha} \left(w_i \tau_{ni}\right)^{-\varepsilon} P_n^{\varepsilon} \beta_n w_n \bar{L}_n, \tag{15}$$

for $i = 1,2$, with $P_n^{-\varepsilon} = \sum_j S_j L_j^{\alpha} \left(w_j \tau_{nj}\right)^{-\varepsilon}$. The case considered by Venables (1987) entails $\alpha = 1$, in which case the previous system can be rewritten as a system in $(P_1, P_2)$,

$$w_i = \sum_n S_i \left(w_i \tau_{ni}\right)^{-\varepsilon} P_n^{\varepsilon} \beta_n w_n \bar{L}_n.$$

In this case it is easy to see that a decline in $\tau_{12}$ leads to an increase in $P_1$ and a decrease in $P_2$, exactly as in Venables (1987). Of course, if $\alpha = 0$ then $P_n^{-\varepsilon} = \sum_j S_j^{\alpha} \left(w_j \tau_{nj}\right)^{-\varepsilon}$, and so $P_1$ would decrease while there would be no change in $P_2$ implying further gains from unilateral trade liberalization. More generally, we can prove the following proposition (see Appendix D.2 for the details):

**Proposition 8.** *Assume there are two countries and two industries, with industry 1 playing the role of an outside good (i.e., $\varepsilon_1 = \infty$, $\psi_1 = 0$ and $\tau_{12,1} = \tau_{21,1} = 1$) and industry 2 having scale economies with $\alpha > 0$. Assume that the initial equilibrium is interior (i.e., $L_{i,k} > 0$ for $i = 1,2$ and $k = 1,2$). There exists a threshold $\bar{\alpha}^{n,\tau} \in (0,1)$ — depending on import and export shares — such that country $n$ loses from a small unilateral trade liberalization in industry 2 if and only if $\alpha \in (\bar{\alpha}^{n,\tau}, 1]$.*[29]

This Proposition generalizes the result of immiserizing inward trade liberalization in Venables (1987) in two ways. First, the result holds outside of the Krugman model – what is needed is that the scale economies vary across sectors, but the source of such economies (for example, love of variety or Marshallian externalities) is irrelevant. Second, the results are a manifestation of the more general idea that a shock that pushes

[29] In Appendix D.2 we also show that there exists a threshold $\bar{\alpha}^{n,S} \in (0,1)$ such that country $n$ loses from a small foreign productivity improvement in industry 2 if and only if $\alpha \in (\bar{\alpha}^{n,S}, 1]$.

a country to specialize in an industry with weak economies of scale (here the outside good) may lower the gains from trade. Using the notion of the pattern of specialization ($PS$) introduced in the previous section, we can restate this as implying that if a shock leads to a decline in a country's $PS$ then the gains from trade liberalization may be lower with economies of scale than without.

### 6.2. Quantitative Analysis

The results of the previous subsection imply that scale economies tend to increase the gains from trade liberalization except if they lead to a decline in the pattern of specialization. We illustrate the relevance of this finding by quantifying the welfare effects of removing all observed tariffs across the countries and sectors in the WIOD for the year 2008. We use data on *ad-valorem* tariffs for the year 2008 from the United Nations Statistical Division, Trade Analysis and Information System (UNCTAD-TRAINS), as described in detail in Appendix D.5. As in Section 5, we assume that $\varepsilon_k = 5$ for all $k$, and consider three scenarios for $\psi_k$: (i) no scale economies, (ii) $\psi_k = 0.14$ for all $k$, and (iii) $\psi_k$ set to the estimates in BCDR for manufacturing industries and set to zero for non-manufacturing industries.[30] To solve for counterfactual equilibria, we use the exact hat algebra approach popularized by Dekle, Eaton and Kortum (2008) and extended here to allow for complementary slackness conditions, as described in Appendix D.3.

Table 3 presents the results for the case in which all tariffs are simultaneously eliminated by all countries. Columns 1-3 report the gains for the cases (i)-(iii), respectively. Column 4 reports the percentage change in $PS$ for case (iii). Comparing columns 1 and 2 reveals that multilateral liberalization tends to generate larger gains with economies of scale than without: average gains increase from 0.16% to 0.19%, with almost all countries experiencing higher gains.

Comparing columns 1 and 3 we see that this tendency for larger gains with economies of scale persists, with average gains increasing from 0.16% to 0.22%. However, with heterogeneous scale elasticities across manufacturing we now see an additional effect: countries which are pushed to specialize more in industries with relatively weak scale

[30] In the absence of estimates of scale elasticities in non-manufacturing industries we assume scale elasticities are zero. A benefit of this assumption is that it allows us to highlight the role and importance of scale heterogeneity. We have also explored an alternative assumption of scale economies in non-manufacturing which are the average across those in manufacturing. The results for this case lie in between cases (i) and (ii), with the underlying message remaining the same.

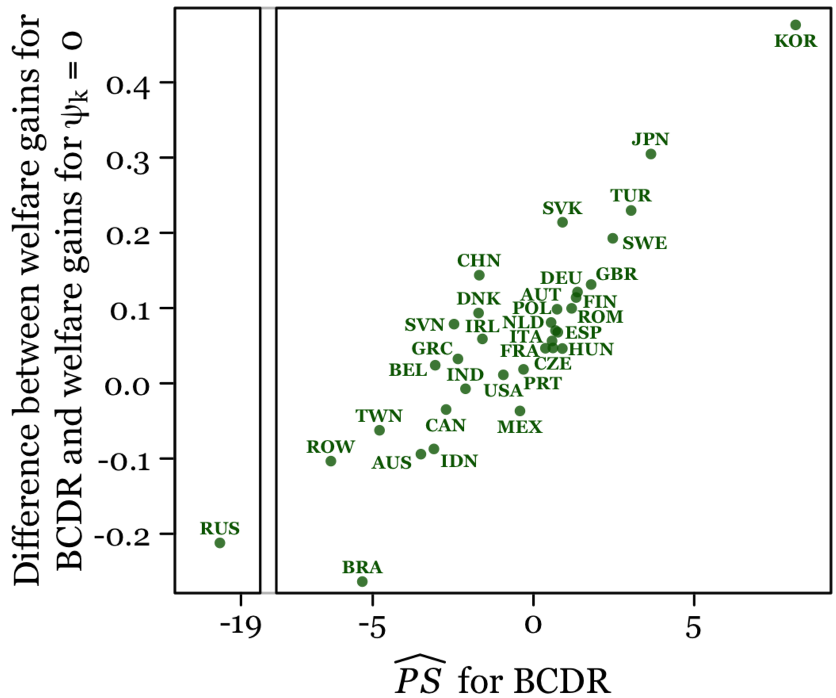
| | Gains from Elimination of Tariffs, % | | | $\widehat{PS}$, % | | Gains from Elimination of Tariffs, % | | | $\widehat{PS}$, % |
|---|---|---|---|---|---|---|---|---|---|
| | $\psi_k = 0$ | $\psi_k = 0.14$ | BCDR | BCDR | | $\psi_k = 0$ | $\psi_k = 0.14$ | BCDR | BCDR |
| Country | (1) | (2) | (3) | (4) | Country | (1) | (2) | (3) | (4) |
| AUS | 0.22 | 0.30 | 0.12 | −3.50 | IRL | 0.31 | 0.25 | 0.37 | −1.59 |
| AUT | 0.13 | 0.15 | 0.23 | 0.73 | ITA | 0.11 | 0.13 | 0.17 | 0.58 |
| BEL | 0.14 | 0.18 | 0.16 | −3.06 | JPN | 0.29 | 0.29 | 0.59 | 3.65 |
| BRA | 0.10 | 0.09 | −0.16 | −5.33 | KOR | 0.62 | 0.70 | 1.09 | 8.16 |
| CAN | 0.17 | 0.24 | 0.14 | −2.72 | MEX | −0.04 | −0.06 | −0.08 | −0.42 |
| CHN | 0.04 | 0.19 | 0.18 | −1.69 | NLD | 0.22 | 0.26 | 0.29 | 0.69 |
| CZE | 0.10 | 0.11 | 0.15 | 0.61 | POL | 0.11 | 0.15 | 0.19 | 0.55 |
| DEU | 0.21 | 0.23 | 0.33 | 1.37 | PRT | 0.09 | 0.12 | 0.11 | −0.31 |
| DNK | 0.22 | 0.32 | 0.32 | −1.71 | ROM | 0.04 | 0.09 | 0.14 | 1.19 |
| ESP | 0.05 | 0.08 | 0.12 | 0.76 | RUS | −0.13 | −0.16 | −0.35 | −19.66 |
| FIN | 0.24 | 0.23 | 0.35 | 1.33 | SVK | 0.14 | 0.19 | 0.35 | 0.90 |
| FRA | 0.13 | 0.16 | 0.18 | 0.37 | SVN | 0.19 | 0.29 | 0.26 | −2.47 |
| GBR | 0.09 | 0.11 | 0.22 | 1.79 | SWE | 0.14 | 0.16 | 0.34 | 2.47 |
| GRC | 0.07 | 0.14 | 0.10 | −2.35 | TUR | 0.10 | 0.11 | 0.33 | 3.04 |
| HUN | 0.09 | 0.13 | 0.14 | 0.90 | TWN | 1.00 | 0.65 | 0.94 | −4.79 |
| IDN | 0.21 | 0.28 | 0.12 | −3.10 | USA | 0.08 | 0.10 | 0.09 | −0.94 |
| IND | 0.00 | −0.01 | 0.00 | −2.11 | ROW | 0.11 | 0.11 | 0.01 | −6.30 |
| | | | | | **Average** | 0.16 | 0.19 | 0.22 | −0.97 |

*Notes:* Each of the columns 1-3 presents results on welfare gains of a separate counterfactual exercise that removes all tariffs observed in the data. For all exercises, $\varepsilon_k = 5$ for all $k$. For column 1, $\psi_k = 0$ for all $k$. For column 2, $\psi_k = 0.14$ for all $k$. For column 3, $\psi_k$ for manufacturing industries are set based on estimates from BCDR, and for non-manufacturing industries $\psi_k = 0$. $\widehat{PS}$ in column 4 is calculated for the exercise in column 3. Average values in the last row are calculated based on the full set of countries.

Table 3: Effects of Multilateral Elimination of Tariffs

economies, as captured by a decline in $PS$, could experience lower gains with economies of scale than without. For example, all five countries — Australia, Brazil, Canada, Indonesia and Taiwan — that experience a fall in gains also have a decline in $PS$. In the case of Brazil gains of 0.1% actually turn to losses of 0.16%. At the same time, countries that are pushed to specialize more in industries with relatively high scale economies (as captured by an increase in $PS$) experience higher gains from economies of scale. For instance, in the case of Japan, Spain, Great Britain and Sweden gains more than double, and for Romania and Turkey they more than triple.

Figure 4 sheds light on these heterogeneous effects by plotting the difference between welfare changes with BCDR elasticities and welfare changes for $\psi_k = 0$ for all $k$ against $\widehat{PS}$ for the full set of countries. The message that emerges is clear: economies




*Notes:* The figure compares results of two counterfactual exercises that have different parameterizations. In both exercises, all tariffs between all countries are eliminated, with all other parameters held unchanged. In both parametrizations, $\varepsilon_k = 5$ for all $k$. In the first parametrization, $\psi_k = 0$ for all $k$. In the second paramterization — "BCDR" — $\psi_k$ for manufacturing industries are based on estimates from Bartelme *et al.* (2017) and for non-manufacturing industries $\psi_k$ are set to 0.

Figure 4: Pattern of Specialization and Welfare Gains for Multilateral Elimination of Tariffs

of scale tend to increase the gains from tariff removal unless a country is pushed to specialize in industries with low economies of scale.

## 7. Scale Economies and Trade Flows

In this section we quantify the role of scale economies in determining industry-level specialization and trade flows. In particular, we ask how these variables would change if we shut down scale economies but leave all other exogenous variables unchanged.

We rely on the fact that if $\boldsymbol{L}$ is an equilibrium of the actual economy with scale economies then it is also an equilibrium of the economy with no scale economies given by

$$w_i L_{i,k} = \sum_{n=1}^{N} \frac{T_{i,k}\left(w_i \tau_{ni,k}\right)^{-\varepsilon_k}}{\sum_{l=1}^{N} T_{l,k}\left(w_l \tau_{nl,k}\right)^{-\varepsilon_k}} \beta_{n,k}(w_n \bar{L}_n + D_n) \tag{16}$$

and

$$\sum_{k=1}^{K} L_{i,k} = \bar{L}_i, \tag{17}$$

where $T_{i,k} \equiv S_{i,k} L_{i,k}^{\alpha_k}$ and where $D_n$ are trade deficits satisfying $\sum_n D_n = 0$. Thus, if we want to know the counterfactual allocation for the economy with $\alpha_k = 0$ for all $k$ but everything else equal, we can use the exact hat algebra approach in the economy with no scale effects subjected to a shock to productivities $T_{i,k}$ given by $\hat{T}_{i,k} = S'_{i,k} / \left(S_{i,k} L_{i,k}^{\alpha_k}\right) = \hat{S}_{i,k} L_{i,k}^{-\alpha_k}$.[31] To focus on the interaction between specialization and scale economies, we assume that $\hat{S}_{i,k}$ is such that if country $i$ was in autarky then the shock would have no effect on productivity. Since in autarky $L_{i,k} = \beta_{i,k} \bar{L}_i$, this requires $\hat{S}_{i,k} = (\beta_{i,k} \bar{L}_i)^{\alpha_k}$. Using $L_{i,k} = r_{i,k} \bar{L}_i$ and measuring $\beta_{i,k}$ by $e_{i,k}$, this implies that $\hat{T}_{i,k} = (e_{i,k}/r_{i,k})^{\alpha_k}$. Using (16) and (17) we can derive a system in wage changes. The solution for $\hat{w}_i$ can then be used to get the implied changes in labor allocations $\hat{L}_{i,k}$, and by extension the implied changes in trade flows $\hat{X}_{ni,k}$. The details of this derivation are in Appendix D.4.[32]

Table 4 presents results of this exercise. Column 1 reports the degree of specialization ($DS$) in the data as defined in (12). Columns 2 and 3 report, respectively, the implied change in $DS$ and in total exports for each country when $\psi_k = 0.14$ and $\varepsilon_k = 5$ for all $k$. Columns 4 and 5 do the same for the case when $\varepsilon_k = 5$ for all $k$ and scale elasticities in manufacturing industries are set based on the estimates from BCDR, while for non-manufacturing industries they are set to zero.

As expected, the removal of economies of scale implies a decline in the degree of specialization and total trade, but the effects are small. Without economies of scale amplifying Ricardian comparative advantage, industry-level specialization would be 1-3% less than what we observe in the data, implying that scale economies are much less important than Ricardian comparative advantage in driving industry-level specialization. In addition, without economies of scale world trade would be between 2.5% and

[31] We can ignore corner solutions because the data has no zeros at the industry level (i.e., $r_{i,k} > 0$ for all $i, k$) and the shock that we consider moves us away from corners.

[32] In order to solve for $\hat{w}_i$, we only need to work with the model for $\alpha_k = 0$ for all $k$ (the standard multi-sector Eaton and Kortum, 2002), and so for this case the algorithm that solves for $\hat{w}_i$ is just a straightforward extension of the Alvarez and Lucas (2007) algorithm to multiple sectors.

| | | Effects of Elimination of Economies of Scale, % | | | | | | Effects of Elimination of Economies of Scale, % | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $\psi_k = 0.14$ | | BCDR | | | | $\psi_k = 0.14$ | | BCDR | |
| | $DS$ | $\widehat{DS}$ | $\widehat{EX}$ | $\widehat{DS}$ | $\widehat{EX}$ | | $DS$ | $\widehat{DS}$ | $\widehat{EX}$ | $\widehat{DS}$ | $\widehat{EX}$ |
| Country | (1) | (2) | (3) | (4) | (5) | Country | (1) | (2) | (3) | (4) | (5) |
| AUS | 0.04 | −2.8 | −12.7 | −1.5 | −3.7 | IRL | 0.07 | −5.4 | −5.2 | −4.1 | −3.7 |
| AUT | 0.02 | −1.2 | −1.1 | −0.4 | −0.2 | ITA | 0.03 | −2.2 | −5.1 | −0.5 | −2.6 |
| BEL | 0.07 | −6.1 | −0.9 | −0.8 | −1.0 | JPN | 0.04 | −3.2 | −11.4 | −0.6 | −5.5 |
| BRA | 0.00 | −0.1 | −5.5 | −0.1 | −2.6 | KOR | 0.13 | −10.8 | −11.1 | −1.5 | −5.9 |
| CAN | 0.02 | −1.9 | −2.3 | −0.8 | 0.9 | MEX | 0.02 | −1.1 | −4.7 | −0.8 | −2.1 |
| CHN | 0.02 | −1.0 | −7.7 | −0.4 | −3.8 | NLD | 0.02 | −1.4 | −3.8 | −0.9 | −2.0 |
| CZE | 0.01 | −1.1 | −3.0 | −0.7 | −2.2 | POL | 0.01 | −0.7 | −2.8 | −0.4 | −1.1 |
| DEU | 0.03 | −2.4 | −4.8 | −0.8 | −3.0 | PRT | 0.03 | −2.5 | −2.2 | −0.7 | 1.5 |
| DNK | 0.09 | −7.8 | −9.2 | −1.9 | 0.1 | ROM | 0.02 | −1.3 | −2.4 | −0.7 | 1.4 |
| ESP | 0.02 | −1.8 | −2.9 | −0.2 | −0.5 | RUS | 0.08 | −4.9 | −18.7 | −3.5 | −9.1 |
| FIN | 0.04 | −2.6 | −7.0 | −1.3 | −5.1 | SVK | 0.04 | −3.2 | −3.5 | −1.0 | −2.1 |
| FRA | 0.02 | −1.6 | −3.9 | −0.2 | −1.7 | SVN | 0.11 | −9.0 | −1.6 | −8.3 | −0.7 |
| GBR | 0.01 | −0.5 | −2.9 | −0.3 | −0.1 | SWE | 0.02 | −1.4 | −3.7 | −0.7 | −2.1 |
| GRC | 0.11 | −9.6 | −35.9 | −3.0 | −3.0 | TUR | 0.02 | −1.2 | −10.0 | −0.7 | −6.6 |
| HUN | 0.03 | −2.0 | −3.2 | −1.1 | −2.1 | TWN | 0.11 | −9.0 | −11.8 | −2.7 | −8.3 |
| IDN | 0.01 | −0.7 | −4.7 | −0.5 | −1.5 | USA | 0.01 | −0.4 | −6.5 | −0.2 | −1.4 |
| IND | 0.01 | −0.9 | −3.9 | −0.1 | −1.9 | ROW | 0.05 | −3.4 | −11.0 | −1.1 | −1.8 |
| | | | | | | **Average** | | −3.1 | −6.7 | −1.2 | −2.5 |

*Notes:* $DS$ in column 1 is calculated from data according to expression (12) and does not depend on $\varepsilon_k$ and $\psi_k$. The two pairs of columns 2-3 and 4-5 correspond to two exercises that differ in the assumption about what model generated the data: for columns 2-3 $\psi_k = 0.14$ for all $k$, while for columns 4-5 $\psi_k$ for manufacturing industries are set based on estimates from Bartelme *et al.* (2017) and for non-manufacturing industries $\psi_k = 0$. In both exercises $\varepsilon_k = 5$ for all $k$. In the counterfactual equilibria, $\psi_k$ is set to zero for all $k$, and all productivities are adjusted so that in autarky the changes in $\psi_k$ would have no effect. All other parameters are held unchanged. $\widehat{DS}$ and $\widehat{ES}$ are the implied changes in the degree of specialization and total exports in terms of the world GDP. Average values in the last row are calculated based on the full set of countries.

Table 4: Scale and Trade Flows

7% lower than what we see in the data.

## 8. Concluding Remarks

For over a century since Alfred Marshall's initial exposition, economists have been intrigued with the implications of industry-level external economies of scale for trading economies. Despite such interest, however, the discomfort with the plethora of equilibria and counter-intuitive implications in early work implied that Marshallian externalities were mostly ignored in the recent trade literature. In this paper we show how one can add Marshallian externalities to the Eaton and Kortum (2002) framework while retaining the property that the equilibrium is unique. The resulting model has exactly the same mathematical structure as generalized versions of the multi-industry Krugman and Melitz-Pareto models and so our uniqueness result applies to these well-known models as well. The key condition for uniqueness is simple and intuitive: the scale elasticity must be weakly lower than the inverse of the trade elasticity for all industries.

The model is rich in welfare implications. Most importantly, if parameters are in the region of uniqueness then all countries gain from trade. Economies of scale tend to make gains from trade lower and gains from trade liberalization higher relative to models without scale economies, with results more positive for countries that specialize in industries with stronger than average scale economies.

Bartelme *et al.* (2017) have proposed a way to estimate industry-level external economies of scale elasticities and to quantify the consequences for industrial policy in a multi-sector gravity model with scale economies like that developed in this paper. Their findings imply that scale economies are positive but lower than the inverse of the trade elasticity, and that scale elasticities substantially vary across manufacturing industries. We have explored the implications of these estimated elasticities — assuming zero scale economies outside of manufacturing industries — for the gains from removal of all tariffs observed in the data for the countries in the WIOD. We find that the presence of economies of scale implies on average larger gains from the removal of all tariffs, with gains more than doubling for some countries — for example, Japan, Spain, Great Britain and Sweden. Still, for countries that are pushed to specialize in industries with low economies of scale gains could fall and even become negative, as, for example, is the case of Brazil.

We also found that while economies of scale do play a role in explaining trade flows and industry-level specialization in the data, they pale in comparison to the importance of Ricardian comparative advantage: shutting down economies of scale would imply a decline in industry-level specialization of less than 3%, with an implied fall in trade of between 2.5% and 7%.

## References

ABDEL-RAHMAN, H. and FUJITA, M. (1990). Product variety, marshallian externalities, and city sizes. *Journal of Regional Science*, **30** (2), 165–183.

ALLEN, T., ARKOLAKIS, C. and LI, X. (2016). On the existence and uniqueness of trade equilibria. *mimeo, Northwestern and Yale Universities*.

—, — and TAKAHASHI, Y. (2014). *Universal Gravity*. NBER Working Papers w20787, National Bureau of Economic Research, Inc.

ALVAREZ, F. and LUCAS, R. (2007). General Equilibrium Analysis of the Eaton-Kortum Model of International Trade. *Journal of Monetary Economics*, **54** (6), 1726–1768.

ARKOLAKIS, C., COSTINOT, A. and RODRÍGUEZ-CLARE, A. (2012). New trade models, same old gains? *American Economic Review*, **102** (1), 94–130.

—, DEMIDOVA, S., KLENOW, P. J. and RODRÍGUEZ-CLARE, A. (2008). Endogenous Variety and the Gains from Trade. *American Economic Review: Papers and Proceedings*, **98** (2), 444–50.

BARTELME, D., COSTINOT, A., DONALDSON, D. and RODRIGUEZ-CLARE, A. (2017). External economies of scale and industrial policy: A view from trade. *mimeo*.

BEHRENS, K., LAMORGESE, A. R., OTTAVIANO, G. I. and TABUCHI, T. (2004). Testing the home market effect in a multi-country world: The theory. *CEPR Discussion Paper #4468*.

—, —, — and — (2009). Beyond the home market effect: Market size and specialization in a multi-country world. *Journal of International Economics*, **79** (2), 259 – 265.

CALIENDO, L. and PARRO, F. (2015). Estimates of the Trade and Welfare Effects of NAFTA. *The Review of Economic Studies*, **82** (1), 1–44.

CHANEY, T. (2008). Distorted gravity: The intensive and extensive margins of international trade. *American Economic Review*, **98** (4), 1707–21.

CHIPMAN, J. S. (1965). A survey of the theory of international trade: Part 2, the neoclassical theory. *Econometrica*, **33** (4), 685–760.

COSTINOT, A., DONALDSON, D. and KOMUNJER, I. (2012). What Goods Do Countries Trade? A Quantitative Exploration of Ricardo's Ideas. *Review of Economic Studies*, **79** (2), 581–608.

— and RODRÍGUEZ-CLARE, A. (2014). Trade theory with numbers: Quantifying the consequences of globalization. In G. Gopinath, E. Helpman and K. Rogoff (eds.), *Handbook of International Economics*, vol. 4, *4*, Elsevier, pp. 197–261.

DEBREU, G. (1982). Chapter 15: Existence of Competitive Equilibrium. *Handbook of Mathematical Economics*, vol. 2, Elsevier, pp. 697 – 743.

DEKLE, R., EATON, J. and KORTUM, S. (2008). Global Rebalancing with Gravity: Measuring the Burden of Adjustment. *IMF Staff Papers*, **55** (3), 511–540.

DORNBUSCH, R., FISCHER, S. and SAMUELSON, P. A. (1977). Comparative advantage, trade, and payments in a ricardian model with a continuum of goods. *American Economic Review*, **67** (5), 823–839.

EATON, J. and KORTUM, S. (2001). Technology, Trade, and Growth: A Unified Framework. *European Economic Review*, **45** (4-6), 742–755.

— and — (2002). Technology, geography, and trade. *Econometrica*, **70** (5), 1741–1779.

ETHIER, W. J. (1982). Decreasing Costs in International Trade and Frank Graham's Argument for Protection. *Econometrica*, **50** (5), 1243–1268.

FEENSTRA, R. C., LUCK, P. A., OBSTFELD, M. and RUSS, K. N. (2014). *In Search of the Armington Elasticity*. NBER Working Paper 20063, National Bureau of Economic Research, Inc.

FERRIS, M. C. and MUNSON, T. S. (1999). Interfaces to PATH 3.0: Design, Implementation and Usage. *Computational Optimization and Applications*, **12** (1), 207–227.

GALLE, S., RODRÍGUEZ-CLARE, A. and YI, M. (2017). Slicing the Pie: Quantifying the Aggregate and Distributional Effects of Trade. Mimeo, University of California, Berkeley.

GRAHAM, F. D. (1923). Some aspects of protection further considered. *The Quarterly Journal of Economics*, **37** (2), 199–227.

GROSSMAN, G. M. and ROSSI-HANSBERG, E. (2010). External Economies and International Trade Redux. *The Quarterly Journal of Economics*, **125** (2), 829–858.

HANSEN, E. and WALSTER, G. (2003). *Global Optimization Using Interval Analysis: Revised And Expanded*. Monographs and textbooks in pure and applied mathematics, CRC Press.

HANSON, G. H. and XIANG, C. (2004). The home-market effect and bilateral trade patterns. *American Economic Review*, **94** (4), 1108–1129.

HEAD, K. and MAYER, T. (2014). Gravity equations: Workhorse, toolkit, and cookbook. In G. Gopinath, E. Helpman and K. Rogoff (eds.), *Handbook of International Economics*, vol. 4, *3*, Elsevier, pp. 131 – 195.

HELPMAN, E. and KRUGMAN, P. (1985). *Market Structure and Foreign Trade: Increasing Returns, Imperfect Competition, and the International Economy*. MIT press.

HORN, R. A. and JOHNSON, C. R. (2013). *Matrix Analysis*. New York: Cambridge University Press, 2nd edn.

KEHOE, T. J. (1980). An Index Theorem for General Equilibrium Models with Production. *Econometrica*, **48** (5), 1211–32.

—, LEVINE, D. K. and ROMER, P. M. (1992). On Characterizing Equilibria of Economies with Externalities and Taxes as Solutions to Optimization Problems. *Economic Theory*, **2** (1), 43–68.

KRUGMAN, P. (1980). Scale Economies, Product Differentiation, and the Pattern of Trade. *American Economic Review*, **70** (5), 950–59.

— (1995). Increasing Returns, Imperfect Competition and the Positive Theory of International trade. In G. M. Grossman and K. Rogoff (eds.), *Handbook of International Economics*, vol. 3, *24*, Elsevier, pp. 1243–1277.

KUBLER, F., RENNER, P. and SCHMEDDERS, K. (2014). Chapter 11 — Computing All Solutions to Polynomial Equations in Economics. In K. Schmedders and K. L. Judd (eds.), *Handbook of Computational Economics Vol. 3*, *Handbook of Computational Economics*, vol. 3, Elsevier, pp. 599 – 652.

LAGAKOS, D. and WAUGH, M. E. (2013). Selection, agriculture, and cross-country productivity differences. *The American Economic Review*, **103** (2), 948–980.

LASHKARIPOUR, A. and LUGOVSKYY, V. (2017). National Differentiation and Industry-Wide Scale Effects. Mimeo, Indiana University.

LYN, G. and RODRÍGUEZ-CLARE, A. (2013a). External Economies and International Trade Redux: Comment. *The Quarterly Journal of Economics*, **128** (4), 1895–1905.

— and — (2013b). Marshallian Externalities, Comparative Advantage and International Trade. Mimeo, Department of Economics, University of California, Berkeley.

MARSHALL, A. (1890). *Principles of Economics*, vol. 1. London: Macmillan and Co., 1st edn.

MAS-COLELL, A., WHINSTON, M. D. and GREEN, J. R. (1995). *Microeconomic Theory*, vol. 1. Oxford university press New York.

MELITZ, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, **71** (6), 1695–1725.

MOORE, R., KEARFOTT, R. and CLOUD, M. (2009). *Introduction to Interval Analysis*. SIAM e-books, Society for Industrial and Applied Mathematics (SIAM, 3600 Market Street, Floor 6, Philadelphia, PA 19104).

NEGISHI, T. (1960). Welfare Economics and Existence of Equilibrium for a Competitive Economy. *Metroeconomica*, **12** (2-3), 92–97.

ORTEGA, J. M. and RHEINBOLDT, W. C. (2000). *Iterative Solution of Nonlinear Equations in Several Variables*. Classics in Applied Mathematics, Society for Industrial and Applied Mathematics.

REDDING, S. J. (2016). Goods trade, factor mobility and welfare. *mimeo, Princeton University*.

RUDIN, W. (1976). *Principles of Mathematical Analysis*. International series in pure and applied mathematics, McGraw-Hill, Inc., 3rd edn.

SCHOENBERG, I. J. (1938). Metric spaces and completely monotone functions. *Annals of Mathematics*, **39** (4), 811–841.

SOMALE, M. (2017). Comparative Advantage in Innovation and Production. Mimeo, Federal Reserve Board.

STOKEY, N. L., LUCAS, R. E., JR. and PRESCOTT, E. C. (1989). *Recursive Methods in Economic Dynamics*. Harvard University Press, Cambridge, MA.

TIMMER, M. P., DIETZENBACHER, E., LOS, B., STEHRER, R. and DE VRIES, G. J. (2015). An Illustrated User Guide to the World InputâĂŞOutput Database: The Case of Global Automotive Production. *Review of International Economics*, **23** (3), 575–605.

VENABLES, A. J. (1987). Trade and trade policy with differentiated products: A chamberlinian-ricardian model. *The Economic Journal*, **97** (387), 700–715.

YATSYNOVICH, Y. (2014). Technological Spillovers and Dynamics of Comparative Advantage. Mimeo, University of California, Berkeley.

# Appendices

## A. Definitions

It is clear that both functions $G_{i,k}(\boldsymbol{w},\boldsymbol{L}_k)$ and $L_{i,k}G_{i,k}(\boldsymbol{w},\boldsymbol{L}_k)$ which appear in the non-linear complementarity problem (4) are well-defined for all positive wages and positive labor allocations, i.e., for all $\boldsymbol{w}\in\mathbb{R}^N_{++}$ and $\boldsymbol{L}_k\in\mathbb{R}^N_{++}$. We are interested in extending the definitions of $G_{i,k}(\boldsymbol{w},\boldsymbol{L}_k)$ and $L_{i,k}G_{i,k}(\boldsymbol{w},\boldsymbol{L}_k)$ to the set of all non-negative labor allocations excluding the point with $L_{i,k}=0$ for all $i$, i.e., to the set $\mathbb{R}^N_+\setminus\{\boldsymbol{0}\}$. To this end, we allow for function $G_{i,k}(\boldsymbol{w},\boldsymbol{L}_k)$ to take infinite values. Formally, we consider the function $G_{i,k}:\mathbb{R}^N_{++}\times\mathbb{R}^N_+\setminus\{\boldsymbol{0}\}\to\mathbb{R}\cup\{-\infty,+\infty\}$,[33] and for each given vector of wages $\boldsymbol{w}\in\mathbb{R}^N_{++}$ and vector of labor allocations $\boldsymbol{L}_k\in\mathbb{R}^N_+\setminus\{\boldsymbol{0}\}$ we formally define $G_{i,k}(\boldsymbol{w},\boldsymbol{L}_k)$ and $L_{i,k}G_{i,k}(\boldsymbol{w},\boldsymbol{L}_k)$ by the limits

$$G_{i,k}(\boldsymbol{w},\boldsymbol{L}_k)\equiv\lim_{\boldsymbol{x}^t\to\boldsymbol{L}_k}\left[w_i-\frac{1}{x_i^t}\sum_n\lambda_{ni,k}(\boldsymbol{w},\boldsymbol{x}^t)\beta_{n,k}w_n\bar{L}_n\right]$$

and

$$L_{i,k}G_{i,k}(\boldsymbol{w},\boldsymbol{L}_k)\equiv\lim_{\boldsymbol{x}^t\to\boldsymbol{L}_k}x_i^t\left[w_i-\frac{1}{x_i^t}\sum_n\lambda_{ni,k}(\boldsymbol{w},\boldsymbol{x}^t)\beta_{n,k}w_n\bar{L}_n\right],$$

where $\{\boldsymbol{x}^t\}_{t=1}^\infty$ is any sequence converging to $\boldsymbol{L}_k$ such that $\boldsymbol{x}^t\in\mathbb{R}^N_{++}$ for $t=1,2,\ldots$.

Let us verify that functions $G_{i,k}(\boldsymbol{w},\boldsymbol{L}_k)$ and $L_{i,k}G_{i,k}(\boldsymbol{w},\boldsymbol{L}_k)$ are well-defined. Since for all $\boldsymbol{L}_k\in\mathbb{R}^N_{++}$ functions $G_{i,k}(\boldsymbol{w},\boldsymbol{L}_k)$ and $L_{i,k}G_{i,k}(\boldsymbol{w},\boldsymbol{L}_k)$ are well-defined and continuous, the above limits coincide with the values of these functions in the corresponding points.

Next, consider any sequence $\{\boldsymbol{x}^t\}_{t=1}^\infty$ with $\boldsymbol{x}^t\in\mathbb{R}^N_{++}$ for $t=1,2,\ldots$ and converging to $\boldsymbol{L}_k$. We have

$$\lim_{\boldsymbol{x}^t\to\boldsymbol{L}_k}\frac{1}{x_i^t}\sum_n\lambda_{ni,k}(\boldsymbol{w},\boldsymbol{x}^t)=\lim_{\boldsymbol{x}^t\to\boldsymbol{L}_k}\left[x_i^t\right]^{\alpha_k-1}\sum_n\frac{S_{i,k}\left(w_i\tau_{ni,k}\right)^{-\varepsilon_k}}{\sum_l S_{l,k}\left[x_l^t\right]^{\alpha_k}\left(w_l\tau_{nl,k}\right)^{-\varepsilon_k}}\beta_{n,k}w_n\bar{L}_n.$$

[33] Here the set $\mathbb{R}\cup\{-\infty,+\infty\}$ is the extended real number system with symbols $-\infty$ and $+\infty$ following the standard conventions (see, for example, p. 11-12 in Rudin, 1976). In particular, for any $x\in\mathbb{R}$, $-\infty<x<+\infty$.

Then, since, $\boldsymbol{L}_k \neq \boldsymbol{0}$,

$$\lim_{\boldsymbol{x}^t \to \boldsymbol{L}_k} \sum_l S_{l,k} \left[x_l^t\right]^{\alpha_k} \left(w_l \tau_{nl,k}\right)^{-\varepsilon_k} = \sum_l S_{l,k} L_{l,k}^{\alpha_k} \left(w_l \tau_{nl,k}\right)^{-\varepsilon_k} > 0 \text{ for all } n.$$

Hence, $\lim_{\boldsymbol{x}^t \to \boldsymbol{L}_k} \frac{1}{x_i^t} \sum_n \lambda_{ni,k}(\boldsymbol{w}, \boldsymbol{x}^t) = \infty$ if $L_{i,k} = 0$ and $0 \leq \alpha_k < 1$, and $\lim_{\boldsymbol{x}^t \to \boldsymbol{L}_k} \frac{1}{x_i^t} \sum_n \lambda_{ni,k}(\boldsymbol{w}, \boldsymbol{x}^t)$ is a positive number if $L_{i,k} > 0$ or if $\alpha_k \geq 1$. This, in turn, implies that

$$\begin{aligned} G_{i,k}(\boldsymbol{w}, \boldsymbol{L}_k) &= \lim_{\boldsymbol{x}^t \to \boldsymbol{L}_k} \left[ w_i - \frac{1}{x_i^t} \sum_n \lambda_{ni,k}(\boldsymbol{w}, \boldsymbol{x}^t) \beta_{n,k} w_n \bar{L}_n \right] \\ &= \begin{cases} -\infty, & \text{if } L_{i,k} = 0 \text{ and } 0 \leq \alpha_k < 1, \\ \text{finite number}, & \text{if } L_{i,k} > 0 \text{ or } \alpha_k \geq 1. \end{cases} \end{aligned}$$

So, the limit always exists and is either $-\infty$ or a finite number. Hence, function $G_{i,k}$ is well-defined with its codomain given by the extended real line $\mathbb{R} \cup \{-\infty, +\infty\}$.

Similarly, it is easy to verify that $\lim_{\boldsymbol{x}^t \to \boldsymbol{L}_k} x_i^t \left[ w_i - \frac{1}{x_i^t} \sum_n \lambda_{ni,k}(\boldsymbol{w}, \boldsymbol{x}^t) \beta_{n,k} w_n \bar{L}_n \right]$ always exists. Moreover, this limit is always a finite number. Hence, function $L_{i,k} G_{i,k}$ is also well-defined.

## B. A Common Framework

In this appendix we derive expressions for price indices and trade shares for the multi-industry versions of Krugman (1980) and Melitz (2003) models. The derivations show that both Krugman and Melitz-Pareto models lead to equilibrium conditions that are almost identical to the ones for the Ricardian model.

### B.1. A Krugman Model with Two-Tier CES preferences

There is a continuum of differentiated varieties within each industry. Preferences are multi-tiered: Cobb-Douglas across industries with weights $\beta_{i,k}$, CES across country bundles within an industry with elasticity $\eta_k$, and CES across varieties within a country bundle with elasticity of substitution $\sigma_k > 1$.

Let $A_{i,k}$ be the exogenous productivity in $(i,k)$ which is common across firms in that industry, let $F_{i,k}$ denote the fixed cost (in terms of labor) associated with the production of any variety in $(i,k)$, and let $M_{i,k}$ the measure of goods produced in $(i,k)$. There is monopolistic competition and trade shares are $\lambda_{ni,k} = \left(P_{ni,k}/P_{n,k}\right)^{1-\eta_k}$, where $P_{ni,k} = M_{i,k}^{1/(1-\sigma_k)}\left(\bar{\sigma}_k w_i \tau_{ni}/A_{i,k}\right)$ is the price index in country $n$ of country $i$ varieties of industry $k$, $\bar{\sigma}_k \equiv \sigma_k/(\sigma_k - 1)$ is the mark-up, and $P_{n,k} = \left(\sum_i P_{ni,k}^{1-\eta_k}\right)^{1/(1-\eta_k)}$.

We now solve for equilibrium variety $M_{i,k}$ as a function of industry employment $L_{i,k}$ and then use the result to derive an expression for trade shares for this model. Variable profits in $(i,k)$ are simply total industry revenues divided by $\sigma_k$. Letting $\Pi_{i,k}$ be total profits net of fixed costs in industry $(i,k)$, we then have $\Pi_{i,k} = \sum_n \lambda_{ni,k} X_{n,k}/\sigma_k - w_i M_{i,k} F_{i,k}$. If $L_{i,k} > 0$, then free entry implies zero profits, so total revenues must equal total wage payments in industry $(i,k)$, $\sum_n \lambda_{ni,k} X_{n,k} = w_i L_{i,k}$. Combined with $\Pi_{i,k} = 0$ we then have $M_{i,k} = L_{i,k}/\left(\sigma_k F_{i,k}\right)$. Trade shares are then

$$\lambda_{ni,k} = \frac{A_{i,k}^{\eta_k-1} F_{i,k}^{-\frac{\eta_k-1}{\sigma_k-1}} L_{i,k}^{\frac{\eta_k-1}{\sigma_k-1}} \left(w_i \tau_{ni,k}\right)^{-(\eta_k-1)}}{\sum_l A_{l,k}^{\eta_k-1} F_{l,k}^{-\frac{\eta_k-1}{\sigma_k-1}} L_{l,k}^{\frac{\eta_k-1}{\sigma_k-1}} \left(w_l \tau_{nl,k}\right)^{-(\eta_k-1)}}$$

with price indices given by

$$P_{n,k} = \mu_k^{Krug} \left(\sum_l A_{l,k}^{\eta_k-1} F_{l,k}^{-\frac{\eta_k-1}{\sigma_k-1}} L_{l,k}^{\frac{\eta_k-1}{\sigma_k-1}} \left(w_l \tau_{nl,k}\right)^{-(\eta_k-1)}\right)^{-1/(\eta_k-1)},$$

where $\mu_k^{Krug} = \sigma_k^{\frac{1}{\sigma_k-1}} \bar{\sigma}_k$.

It is then immediately evident that the expressions for trade shares and industry price indexes in this Krugman model collapses to those in equations (6) and (7) by setting $\mu_{n,k} = \mu_k^{Krug}$, $S_{i,k} = A_{i,k}^{\eta_k-1} F_{i,k}^{-\frac{\eta_k-1}{\sigma_k-1}}$, $\psi_k = (\sigma_k - 1)^{-1}$, and $\varepsilon_k = (\eta_k - 1)$. Note also that if we set $\sigma_k = \eta_k$ for all $k$, then this is just the standard multi-industry Krugman model, while if $\sigma_k \to \infty$, then $(\eta_k - 1)/(\sigma_k - 1) \to 0$ and we obtain the multi-industry Armington model. See rows 2 and 4 of Table 1.

Given the expressions for the price indices and trade shares derived above, the equilibrium conditions are given by the same expressions as in the Ricardian model: goods market clearing conditions (3)-(4) and labor market clearing conditions (5).

### B.2. A Melitz-Pareto Model with Two-Tier Preferences

We now present a model à la Melitz (2003) with Pareto-distributed productivity.

After paying a fixed "entry" cost $F_{i,k}$ in units of labor in country $i$, firms are able to produce a variety in industry $(i,k)$ with labor productivity drawn from a Pareto distribution with shape parameter $\theta_k > \sigma_k - 1$ and location parameter $b_{i,k}$. Firms from $i$ can then pay a fixed "marketing" cost $f_{n,k}$ in units of labor of $n$ to serve that market.

Let us ignore the industry subscript for a moment. Let $\Omega_{ni}$ denote the set of varieties that $i$ sells to $n$. The price index of these goods is $P_{ni} \equiv \left(\int_{\omega\in\Omega_{ni}} p_{ni}(\omega)^{1-\sigma} d\omega\right)^{\frac{1}{1-\sigma}}$. Let $M_i$ denote total entry in country $i$ and $\varphi^*_{ni}$ denote the cutoff productivity such that $i$ exports to $n$ all goods with productivity higher than $\varphi^*_{ni}$. We have

$$\begin{aligned} P_{ni}^{1-\sigma} &= M_i \int_{\varphi^*_{ni}}^{\infty} [p_{ni}(\varphi)]^{1-\sigma} \, dG_i(\varphi) \\ &= \theta b_i^{\theta} M_i \left[\frac{\sigma}{\sigma-1} w_i \tau_{ni}\right]^{1-\sigma} \int_{\varphi^*_{ni}}^{\infty} \varphi^{\sigma-\theta-2} d\varphi \\ &= M_i \left[\frac{\sigma}{\sigma-1} w_i \tau_{ni}\right]^{1-\sigma} \frac{\theta b_i^{\theta} \left(\varphi^*_{ni}\right)^{\sigma-\theta-1}}{\theta-(\sigma-1)}. \end{aligned}$$

The condition that determines the cutoff $\varphi^*_{ni}$ is

$$\frac{1}{\sigma}\left(\frac{\sigma}{\sigma-1}\cdot\frac{w_i\tau_{ni}}{\varphi^*_{ni}}\right)^{1-\sigma} P_{ni}^{\sigma-1}\left(\frac{P_{ni}}{P_n}\right)^{1-\eta} X_n = w_n f_n.$$

This implies that

$$\varphi^*_{ni} = \frac{w_i\tau_{ni}}{P_{ni}}\left(\frac{\sigma}{\tilde{\sigma}}\cdot\frac{w_n f_n}{X_n}\right)^{\frac{1}{\sigma-1}}\left(\frac{P_{ni}}{P_n}\right)^{\frac{1-\eta}{1-\sigma}},$$

where $\tilde{\sigma} \equiv \left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma}$.[34] Plugging this expression into the expression for the price index yields

$$P_{ni} = w_i\tau_{ni}\left(\frac{w_i\tau_{ni}}{P_{ni}}\left(\frac{X_n}{w_n f_n}\right)^{\frac{1}{1-\sigma}}\left(\frac{P_{ni}}{P_n}\right)^{\frac{1-\eta}{1-\sigma}}\right)^{\frac{\theta}{\sigma-1}-1} M_i^{\frac{1}{1-\sigma}}\bar{\kappa}^o_{i,k},$$

where

$$\bar{\kappa}^o_{i,k} \equiv \frac{\sigma}{\sigma-1}\left(\frac{\theta b_i^{\theta}\left[\frac{\sigma}{\tilde{\sigma}}\right]^{1-\frac{\theta}{\sigma-1}}}{\theta-(\sigma-1)}\right)^{\frac{1}{1-\sigma}}.$$

[34] This expression is valid as long as $\varphi^*_{ni} \geq b_i$. We assume that this inequality holds for all $i,n$.

Bringing back the $k$ subindex and using the well known result that in this model equilibrium entry in each industry must satisfy $M_{i,k} = \frac{\sigma_k - 1}{\sigma_k \theta_k} \frac{L_{i,k}}{F_{i,k}}$, we then have

$$P_{ni,k} = w_i \tau_{ni,k} \left( \frac{w_i \tau_{ni,k}}{P_{ni,k}} \left( \frac{X_{n,k}}{w_n f_n} \right)^{\frac{1}{1-\sigma_k}} \left( \frac{P_{ni,k}}{P_{n,k}} \right)^{\frac{1-\eta_k}{1-\sigma_k}} \right)^{\frac{\theta_k}{\sigma_k - 1} - 1} L_{i,k}^{\frac{1}{1-\sigma_k}} \kappa_{i,k}^{o}.$$

where $\kappa_{i,k}^{o} \equiv \left( \frac{\sigma_k - 1}{\sigma_k \theta_k} F_{i,k}^{-1} \right)^{\frac{1}{1-\sigma_k}} \bar{\kappa}_{i,k}^{o}$. Using $X_{n,k} = \beta_{n,k} w_n \bar{L}_n$ we get

$$P_{ni,k} = w_i \tau_{ni,k} \left( \frac{w_i \tau_{ni,k}}{P_{ni,k}} \left( \frac{P_{ni,k}}{P_{n,k}} \right)^{\frac{1-\eta_k}{1-\sigma_k}} \right)^{\frac{\theta_k}{\sigma_k - 1} - 1} L_{i,k}^{\frac{1}{1-\sigma_k}} \kappa_{n,k}^{d} \kappa_{i,k}^{o},$$

where

$$\kappa_{n,k}^{d} \equiv \left[ \frac{\beta_{n,k} \bar{L}_n}{f_{n,k}} \right]^{\frac{1}{1-\sigma_k} \left( \frac{\theta_k}{\sigma_k - 1} - 1 \right)}.$$

Solving for $P_{ni,k}$ we get

$$P_{ni,k}^{1-\eta_k} = \left[ \left( w_i \tau_{ni,k} \right)^{-\theta_k} P_{n,k}^{\frac{1-\eta_k}{1-\sigma_k} (\theta_k - \sigma_k + 1)} L_{i,k} (\kappa_{n,k}^{d} \kappa_{i,k}^{o})^{1-\sigma_k} \right]^{\xi_k}.$$

where $\xi_k \equiv \frac{1}{1 + \theta_k \left( \frac{1}{\eta_k - 1} - \frac{1}{\sigma_k - 1} \right)}$. The expression for trade shares can then be derived by using $P_{n,k}^{1-\eta_k} = \sum_i P_{ni,k}^{1-\eta_k}$ together with $\lambda_{ni,k} = \left( \frac{P_{ni,k}}{P_{n,k}} \right)^{1-\eta_k}$, that is,

$$\lambda_{ni,k} = \frac{b_{i,k}^{\theta_k \xi_k} F_{i,k}^{-\xi_k} L_{i,k}^{\xi_k} \left( w_i \tau_{ni,k} \right)^{-\theta_k \xi_k}}{\sum_l b_{l,k}^{\theta_k \xi_k} F_{l,k}^{-\xi_k} L_{l,k}^{\xi_k} \left( w_l \tau_{nl,k} \right)^{-\theta_k \xi_k}}.$$

Finally, combining $P_{n,k}^{1-\eta_k} = \sum_i P_{ni,k}^{1-\eta_k}$ with the result above for $P_{ni,k}$ yields price indices

$$P_{n,k} = \mu_{n,k}^{Mel} \left( \sum_l b_{l,k}^{\theta_k \xi_k} F_{l,k}^{-\xi_k} L_{l,k}^{\xi_k} \left( w_l \tau_{nl,k} \right)^{-\theta_k \xi_k} \right)^{-1/\theta_k \xi_k}$$

where $\xi_k \equiv \frac{1}{1 + \theta_k \left( \frac{1}{\eta_k - 1} - \frac{1}{\sigma_k - 1} \right)}$, $\mu_{n,k}^{Mel} \equiv \bar{\mu}_k^{Mel} \left( \frac{f_{n,k}}{\beta_{n,k} L_n} \right)^{\left( \frac{1}{\sigma_k - 1} - \frac{1}{\theta_k} \right)}$ and $\bar{\mu}_k^{Mel} \equiv \left( \frac{\sigma_k - 1}{\sigma_k} \right)^{\frac{\sigma_k}{\theta_k}} \left( \frac{\frac{1}{\theta_k} \left( \frac{\sigma_k}{\bar{\sigma}_k} \right)^{1 - \frac{\theta_k}{\sigma_k - 1}}}{\theta_k - (\sigma_k - 1)} \right)^{\frac{1}{\theta_k}}$. Note that if we set $\sigma_k = \eta_k$ for all $k$ then $\xi_k = 1$ and this model is just a multi-industry version of the Melitz-Pareto model in Arkolakis *et al.*

(2008).

We get expressions (6) and (7) for trade shares and price indices by setting $\mu_{n,k} = \bar{\mu}_k^{Mel}\left(\frac{f_{n,k}}{\beta_{n,k}L_n}\right)^{\left(\frac{1}{\sigma_k-1}-\frac{1}{\theta_k}\right)}$, $S_{i,k} = b_{i,k}^{\theta_k\xi_k}F_{i,k}^{-\xi_k}$, $\psi_k = 1/\theta_k$, and $\varepsilon_k = \theta_k\xi_k$. The goods market clearing conditions are then given by (3)-(4). The labor market clearing conditions need to take into account that labor is used not only for production, but also for paying fixed entry costs:

$$\sum_k w_i L_{i,k} = \left[1 - \sum_k \beta_{i,k}\left(\frac{\theta_k - (\sigma_k - 1)}{\sigma_k\theta_k}\right)\right] w_i \bar{L}_i.$$

This expression differs from the labor market clearing condition (5) in the Ricardian model by a constant term. This difference is inconsequential for equilibrium analysis and comparative statics, therefore, in the analysis pertaining to the general model we ignore this term.

The above derivations for the Melitz model rely on the balanced trade assumption. Without this assumption, the multiplicative term for the price index and the term on the right-hand side of the labor market clearing condition would depend on trade deficits. This has implications for the comparative statics exercises that involve changing trade deficits. At the same time, the derivations for the Ricardian model with Marshallian externalities and for the Krugman model do not depend on the balanced trade assumption.

### B.3. A Ricardian Model with Heterogeneous Labor Supply

In this section we present a version of a Ricardian model with external economies of scale and heterogeneous labor supply as in Galle *et al.* (2017). Similarly to our baseline model of Section 2, each of $k = 1,\ldots,K$ industries in any country $i = 1,\ldots,N$ is composed of a continuum of goods $\omega \in [0,1]$. Labor productivity for good $\omega$ in industry $(i,k)$ is $z_{i,k}(\omega)E_{i,k}^{\phi_k}$, where $E_{i,k}$ is the amount of efficiency units of labor employed in industry $(i,k)$, and $z_{i,k}$ is drawn from a Fréchet distribution with shape parameter $\theta_k$ and scale parameter $T_{i,k}$. Parameter $\phi_k$ plays the role of the strength of economies of scale in terms of efficiency units of labor. A worker in country $i$ can supply $a_{i,k}$ efficiency units to industry $k$, with $a_{i,k}$ drawn from a Fréchet distribution with shape parameter $\kappa > 1$ and scale parameter $A_{i,k}$. The total number of workers in country $i$ is $\bar{L}_i$.

Since workers are heterogeneous in their sector productivities, wages can differ across sectors. Denote by $w_{i,k}$ wage per efficiency unit in industry $(i,k)$. The industry-level bilateral trade shares are given by

$$\lambda_{ni,k}(\boldsymbol{w}_k,\boldsymbol{E}_k)=\frac{T_{i,k}\left(\tau_{ni,k}w_{i,k}/E_{i,k}^{\phi_k}\right)^{-\theta_k}}{\sum_l T_{l,k}\left(\tau_{nl,k}w_{l,k}/E_{l,k}^{\phi_k}\right)^{-\theta_k}}, \tag{18}$$

where $\boldsymbol{w}_k\equiv\left(w_{1,k},\ldots,w_{N,k}\right)$ and $\boldsymbol{E}_k\equiv\left(E_{1,k},\ldots,E_{N,k}\right)$. Equilibrium allocations of efficiency units in industry $(i,k)$ satisfy the following complementary slackness condition:

$$\frac{w_{i,k}E_{i,k}}{\Phi_i}\geq 0,\quad G_{i,k}(\boldsymbol{w}_k,\boldsymbol{E}_k)\geq 0,\quad \frac{w_{i,k}E_{i,k}}{\Phi_i}G_{i,k}(\boldsymbol{w}_k,\boldsymbol{E}_k)=0, \tag{19}$$

where

$$G_{i,k}(\boldsymbol{w}_k,\boldsymbol{E}_k)=\Phi_i-\frac{\Phi_i}{w_{i,k}E_{i,k}}\sum_n\lambda_{ni,k}(\boldsymbol{w}_k,\boldsymbol{E}_k)\beta_{n,k}\eta\Phi_n\bar{L}_n, \tag{20}$$

$$\Phi_i\equiv\left(\sum_s A_{i,s}w_{i,s}^{\kappa}\right)^{\frac{1}{\kappa}}, \tag{21}$$

and $\eta>0$ is a constant. Since $w_{i,k}E_{i,k}/\Phi_i$ is proportional to employment in sector $(i,k)$ and $\Phi_i$ is proportional to the wage per person in country $i$, this definition of function $G_{i,k}$ is a straightforward extension of the one used for the baseline model with homogeneous labor.[35] Supply of efficiency units in industry $(i,k)$ is given by

$$E_{i,k}=\eta\frac{A_{i,k}w_{i,k}^{\kappa-1}}{\Phi_i^{\kappa-1}}\bar{L}_i, \tag{22}$$

and the total income in country $i$ can be calculated as

$$Y_i=\sum_k w_{i,k}E_{i,k}=\eta\Phi_i\bar{L}_i. \tag{23}$$

[35]Specifically, as we show below, the number of workers employed in sector $(i,k)$ is proportional to $w_{i,k}E_{i,k}/\Phi_i$, while $\Phi_i$ is proportional to the wage in country $i$. It is easy to check that for $\tilde{\alpha}_k\in(1-1/\kappa,1]$ the alternative definition $G_{i,k}=w_{i,k}-\frac{1}{E_{i,k}}\sum_n\lambda_{ni,k}\beta_{n,k}Y_n$ with complementary slackness condition (19) changed accordingly, would generate corner labor allocations that satisfy all conditions (18)-(23) and, nevertheless, are not equilibrium allocations.

The equilibrium of the model is given by sector wages $\boldsymbol{w}_k$ and efficiency labor allocations $\boldsymbol{E}_k$, for $k = 1, \dots, K$, that satisfy the equilibrium system given by (18)-(23). In order to show equivalence of this equilibrium system to the one in the common framework of Section 3, use (22) to express $w_{i,k} = \left(\eta \bar{L}_i A_{i,k}\right)^{\frac{1}{1-\kappa}} E_{i,k}^{\frac{1}{\kappa-1}} \Phi_i$, relabel $\Phi_i$ as $w_i$, and let

$$L_{i,k} \equiv \eta^{-1} \frac{w_{i,k} E_{i,k}}{w_i} = \eta^{\frac{\kappa}{1-\kappa}} \left(\bar{L}_i A_{i,k}\right)^{\frac{1}{1-\kappa}} E_{i,k}^{\frac{\kappa}{\kappa-1}}.$$

In this new notation $E_{i,k} = \eta \left(\bar{L}_i A_{i,k}\right)^{\frac{1}{\kappa}} L_{i,k}^{\frac{\kappa-1}{\kappa}}$ and trade shares are given by

$$\tilde{\lambda}_{ni,k}\left(\boldsymbol{w}, \boldsymbol{L}_k\right) = \frac{\tilde{T}_{i,k} \left(\tau_{ni,k} w_i\right)^{-\theta_k} L_{i,k}^{\tilde{\alpha}_k}}{\sum_l \tilde{T}_{l,k} \left(\tau_{nl,k} w_l\right)^{-\theta_k} L_{l,k}^{\tilde{\alpha}_k}},$$

where $\tilde{T}_{i,k} \equiv T_{i,k} \eta^{\theta_k \phi_k} \left(\bar{L}_i A_{i,k}\right)^{\frac{\theta_k (1+\phi_k)}{\kappa}}$ and $\tilde{\alpha}_k \equiv \theta_k \left(\phi_k - \frac{1}{\kappa-1}\right) \frac{\kappa-1}{\kappa}$. Goods market clearing conditions (19)-(20) can be rewritten as

$$L_{i,k} \geq 0, \quad \tilde{G}_{i,k}\left(\boldsymbol{w}, \boldsymbol{L}_k\right) \geq 0, \quad L_{i,k} \tilde{G}_{i,k}\left(\boldsymbol{w}, \boldsymbol{L}_k\right) = 0,$$

where

$$\tilde{G}_{i,k}\left(\boldsymbol{w}, \boldsymbol{L}_k\right) = w_i - \frac{1}{L_{i,k}} \sum_n \tilde{\lambda}_{ni,k}\left(\boldsymbol{w}, \boldsymbol{L}_k\right) \beta_{n,k} w_n \bar{L}_n.$$

Finally, expression (21) can be manipulated to yield the labor market clearing condition:

$$\begin{aligned} w_i &= \left(\sum_s A_{i,s} w_{i,s}^{\kappa}\right)^{\frac{1}{\kappa}} = \left(\sum_s A_{i,s} \left(\eta \bar{L}_i A_{i,s}\right)^{\frac{\kappa}{1-\kappa}} E_{i,s}^{\frac{\kappa}{\kappa-1}} w_i^{\kappa}\right)^{\frac{1}{\kappa}} \\ &= \left(\sum_s A_{i,s} \left(\bar{L}_i A_{i,s}\right)^{-1} L_{i,s}\right)^{\frac{1}{\kappa}} w_i, \end{aligned}$$

which is equivalent to

$$\sum_s L_{i,s} = \bar{L}_i.$$

Hence, indeed, we get the equilibrium system that is equivalent to the one for the common framework of Section 3. Condition for uniqueness in the common framework, $\tilde{\alpha}_k \le 1$, is equivalent to $\left(\phi_k - \frac{1}{\kappa-1}\right)\frac{\kappa-1}{\kappa} \le \frac{1}{\theta_k}$. As $\kappa \to \infty$, we get the same condition in terms of parameters $\theta_k$ and $\phi_k$ as in our baseline Ricardian model of Section 2, but in general we can allow for stronger economies of scale in the model with heterogenous labor — given by $\phi_k$ — while still having condition $\tilde{\alpha}_k \le 1$ satisfied.

# C. Existence and Uniqueness

## C.1. Proof of Proposition 1

As it is mentioned in the main text, the case with $\alpha_k = 0$ is trivial: given wages, labor allocations are explicitly obtained from the conditions $L_{i,k}G_{i,k}(\boldsymbol{w}, \boldsymbol{L}_k) = 0$. Below we focus on the case with $\alpha_k \in (0,1]$.

For brevity of notation suppress the sub-index $k$ and let $a_{ni} \equiv S_i \left(w_i \tau_{ni}\right)^{-\theta} w_i^{-\alpha}$ and $b_n \equiv \beta_n w_n \bar{L}_n$. Combining equations (1) and (3) we then have

$$\frac{G_i\left(\boldsymbol{w}, \boldsymbol{L}\right)}{w_i} = 1 - \frac{1}{w_i L_i} \sum_n \frac{a_{ni}(w_i L_i)^{\alpha}}{\sum_l a_{nl}(w_l L_l)^{\alpha}} b_n.$$

Transforming variables with $x_i \equiv w_i L_i$, letting $\boldsymbol{x} \equiv (x_1, ..., x_N)$, and with a slight abuse of notation, we can write

$$G_i\left(\boldsymbol{x}\right) = 1 - \sum_n \frac{a_{ni} x_i^{\alpha-1}}{\sum_l a_{nl} x_l^{\alpha}} b_n.^{36}$$

The system in (4) can now be written as a non-linear complementarity problem (NCP) in $\boldsymbol{x}$:

$$x_i \ge 0, \quad G_i(\boldsymbol{x}) \ge 0, \quad x_i G_i(\boldsymbol{x}) = 0, \quad i = 1, \ldots, N. \tag{24}$$

Note that if $\boldsymbol{x}$ solves (24) then $\sum_i x_i G_i(\boldsymbol{x}) = 0$ and hence $\sum_i x_i = \sum_i b_i$. This implies that the solution to (24) satisfies $\boldsymbol{x} \in \Gamma \equiv \left\{\boldsymbol{x} \in \mathbb{R}^N \,|\, x_i \ge 0,\ i = 1, \ldots, N;\ \sum_i x_i = \sum_i b_i\right\}$.

To prove Proposition 1 we follow a popular approach in the economics literature that consists of characterizing equilibria of general equilibrium models as solutions to optimization problems.[37] Doing this is possible if, for example, the function $\boldsymbol{G}(\boldsymbol{x}) \equiv$

[36] Analogously to our treatment of the original functions $G_{i,k}(\boldsymbol{w}, \boldsymbol{L}_k)$ and $L_{i,k}G_i(\boldsymbol{w}\boldsymbol{L}_k)$ at $L_{i,k} = 0$ (see Appendix A), we define values of $G_i(\boldsymbol{x})$ and $x_i G_i(\boldsymbol{x})$ at $x_i = 0$ by their limits.

[37] Negishi (1960) is probably the most well-known example of this approach in which market equilibria are characterized as solutions to a social planner's problem. Kehoe, Levine and Romer (1992) describe a more general framework in which the optimization problem does not necessarily have an economic

$(G_1(\boldsymbol{x}),\ldots,G_N(\boldsymbol{x}))$ has a Jacobian that is symmetric at all points in its domain, since in this case the function $\boldsymbol{G}$ is the gradient of some other function $F$ that we can use in the optimization problem.[38] Fortunately, our function $\boldsymbol{G}$ satisfies this symmetry condition. In fact, it is easy to see that $\boldsymbol{G}$ is the gradient of function $F : \mathbb{R}^N_+ \setminus \{\mathbf{0}\} \to \mathbb{R}$ defined by

$$F(\boldsymbol{x}) \equiv \alpha \sum_n x_n - \sum_n b_n \ln\left(\sum_i a_{ni} x_i^\alpha\right). \tag{25}$$

As we establish formally below, this makes it possible to solve the NCP in (24) by way of solving $\arg\min_{\boldsymbol{x}\in\Gamma} F(\boldsymbol{x})$, where $\Gamma$ is the compact set defined above.[39]

In Lemmas 1 and 2 below we characterize the optimization problem $\arg\min_{\boldsymbol{x}\in\Gamma} F(\boldsymbol{x})$. After that, in Lemma 3, we establish the connection between this problem and the NCP (24).

**Lemma 1.** *If either (a)* $0 < \alpha < 1$*, or (b)* $\alpha = 1$ *and Assumption 1 holds, then* $F(\cdot)$ *defined in* (25) *has a unique global minimum on* $\Gamma$.

*Proof.* We prove this lemma by showing that function $F(\cdot)$ is strictly convex on $\Gamma$. The uniqueness of solution of the minimization problem $\min_{\boldsymbol{x}\in\Gamma} F(\boldsymbol{x})$ then follows from the fact that $\Gamma$ is a convex set, and a strictly convex function can have at most one global minimum on a convex set.

If $0 < \alpha < 1$, then for any $n$ function $\sum_i a_{ni} x_i^\alpha$ is a strictly concave function. And since the logarithm is a strictly concave function, $F(\cdot)$ is strictly convex.

If $\alpha = 1$, then we need to make sure that for any two vectors $\boldsymbol{x} \neq \boldsymbol{y}$ we cannot have that $\sum_i a_{ni} x_i = \sum_i a_{ni} y_i$ for all $n$. Otherwise, we would have $F(\gamma\boldsymbol{x} + (1-\gamma)\boldsymbol{y}) = \gamma F(\boldsymbol{x}) + (1-\gamma)F(\boldsymbol{y})$ for any $\gamma \in [0,1]$ and strict convexity would be violated. Assumption 1 guarantees that matrix $A = (a_{ni})$ is non-singular. Hence, for any $\boldsymbol{x} \neq \boldsymbol{y}$ we have $\sum_i a_{ni} x_i \neq \sum_i a_{ni} y_i$ for at least one $n$. Hence, in case of $\alpha = 1$, $F(\cdot)$ is also strictly convex under Assumption 1. Note that without Assumption 1 function $F(\cdot)$ is just convex, but not necessarily strictly convex. □

interpretation. Our case fits into their general framework.

[38]A classical result in mathematics states that a vector function is a gradient map if and only if its Jacobian is symmetric in the domain of the function (see, for example, Theorem 4.1.16 on page 95 in Ortega and Rheinboldt, 2000).

[39]We thank Anca Ciurte and Ioan Rasa for pointing us in this direction.

**Lemma 2.** *Let $\boldsymbol{x}^*$ be the unique global minimum of $F(\cdot)$ on $\Gamma$. If $0<\alpha<1$, then $x_i^*>0$ for all $i=1,\dots,N$.*

*Proof.* Suppose, without loss of generality, that $x_1^*=0$. Since $\sum_n x_n^* = \sum_n b_n > 0$, we can also suppose without loss of generality that $x_2^* \neq 0$. Consider the vector $\boldsymbol{x}(\varepsilon) = \left(\varepsilon, x_2^*-\varepsilon, x_3^*,\dots,x_N^*\right)$, where $\varepsilon \in \left[0, x_2^*\right]$. Clearly, $\boldsymbol{x}(\varepsilon)\in\Gamma$. Define

$$\tilde{F}(\varepsilon) \quad\equiv\quad F(\boldsymbol{x}(\varepsilon)) = \alpha\sum_n x_n^* - \sum_n b_n \ln\left(a_{n1}\varepsilon^\alpha + a_{n2}(x_2^*-\varepsilon)^\alpha + \sum_{i=2}^N a_{ni}\left(x_i^*\right)^\alpha\right).$$

We now show that $\tilde{F}(\varepsilon)<\tilde{F}(0)$ for small enough $\varepsilon>0$. We have

$$\frac{\partial\left(a_{j1}\varepsilon^\alpha + a_{j2}(x_2^*-\varepsilon)^\alpha\right)}{\partial\varepsilon} = \alpha a_{j1}\varepsilon^{\alpha-1} - \alpha a_{j2}(x_2^*-\varepsilon)^{\alpha-1},$$

which, given $\alpha\in(0,1)$, is positive for small enough $\varepsilon$. This implies that $a_{j1}\varepsilon^\alpha + a_{j2}(x_2^*-\varepsilon)^\alpha + \sum_{i=2}^N a_{ji}\left[x_i^*\right]^\alpha > a_{j2}\left[x_2^*\right]^\alpha + \sum_{i=2}^N a_i\left[x_i^*\right]^\alpha$ for small enough $\varepsilon$. Since $\ln(\cdot)$ is a strictly increasing function, we then get that $\tilde{F}(\varepsilon)<\tilde{F}(0)$ for small enough $\varepsilon$, a contradiction. This implies that $\boldsymbol{x}^*$ cannot be a global minimum of $F(\cdot)$ on $\Gamma$. Hence, in the case of $\alpha\in(0,1)$ we must have $x_i^*>0$ for all $i$. □

**Lemma 3.** *If $0<\alpha\le 1$, then $\boldsymbol{x}^*$ is a global minimum of $F(\cdot)$ on $\Gamma$ if and only if $\boldsymbol{x}^*$ is a solution to* (24).

*Proof.* The proof of this lemma is almost trivial, because the conditions in (24) are just the first-order conditions for the minimization of $F(\cdot)$ on $\Gamma$. The only complication is that, to invoke the first-order conditions, we need to have differentiability of $F(\cdot)$ on $\Gamma$, which is understood as differentiability of $F(\cdot)$ on some open set containing $\Gamma$. In case of $\alpha<1$ any such open set necessarily includes points $\boldsymbol{x}$ with $x_i\le 0$, at which $F(\cdot)$ is not differentiable. The proof below formally deals with this complication.

Let us start with the simpler case of $\alpha=1$. We can take the set $D\equiv\left\{\boldsymbol{x}\in\mathbb{R}^N \mid \sum_i a_{ni}x_i>0 \text{ for all } n\right\}$ as the domain of $F(\cdot)$ and use $\Gamma$ as the constraint set. Clearly, $\Gamma\subset D\cap\left\{\boldsymbol{x}\in\mathbb{R}^N \mid \sum_i x_i=\sum_i b_i\right\}$ and $F(\cdot)$ is differentiable on $D$ for $\alpha=1$. Consider the minimization problem

$$\min_{\boldsymbol{x}\in D} F(\boldsymbol{x})$$
$$\text{s.t.}$$
$$x_i \geq 0, \quad i = 1,\dots,N;$$
$$\sum_i x_i = \sum_i b_i. \tag{26}$$

Its first-order conditions, after some manipulations, can be written as

$$x_i \geq 0, \quad \frac{\partial F(\boldsymbol{x})}{\partial x_i} \geq 0, \quad x_i \frac{\partial F(\boldsymbol{x})}{\partial x_i} = 0, \quad i = 1,\dots,N,$$
$$\sum_i x_i = \sum_i b_i.$$

These conditions are also sufficient, because $F(\cdot)$ is convex and the set that satisfies the constraints of the minimization problem is a convex set. Then, since $G_i(\boldsymbol{x}) = \frac{\partial F(\boldsymbol{x})}{\partial x_i}$, we get that any solution of the minimization problem (26) is also a solution of the NCP in (24) and vice versa.

Let us now turn to the case with $0 < \alpha < 1$. Let $\boldsymbol{x}^*$ be a minimum of $F(\cdot)$ on $\Gamma$. By Lemma 2, $x_i^* > 0$ for all $i$. Let $\delta > 0$ be some number such that $\delta \leq x_i^*$ for all $i = 1,\dots,N$. Define the domain of $F$ by $\tilde{D} \equiv \{\boldsymbol{x} \in \mathbb{R}^N \mid x_i > \delta/2,\ i = 1,\dots,N\}$. Consider the minimization problem:

$$\min_{\boldsymbol{x}\in \tilde{D}} F(\boldsymbol{x})$$
$$\text{s.t.} \tag{27}$$
$$\sum_i x_i = \sum_i b_i.$$

Since $\tilde{D} \cap \{\boldsymbol{x} \in \mathbb{R}^N \mid \sum_i x_i = \sum_i b_i\} \subset \Gamma$, if $\boldsymbol{x}^*$ minimizes $F$ on $\Gamma$, it also solves the minimization problem (27). Since $x_i^* > \delta/2$ for all $i$, the first-order conditions for (27) are given by

$$\frac{\partial F(\boldsymbol{x}^*)}{\partial x_i} = 0,\ i = 1,\dots,N, \quad \text{and} \quad \sum_i x_i^* = \sum_i b_i. \tag{28}$$

Hence, $\boldsymbol{x}^*$ solves NCP in (24). Conversely, if $\boldsymbol{x}^*$ solves NCP, then $x_i^* > 0$ for all $i$ because the condition $G_i(\boldsymbol{x}^*) \geq 0$ cannot be satisfied for $x_i^* = 0$ if $0 < \alpha < 1$. Hence, $\boldsymbol{x}^*$ satisfies conditions (28), which are the first order conditions for an interior solution of (27) with an appropriately chosen $\delta > 0$. Since these first-order conditions are also sufficient,

$\boldsymbol{x}^*$ solves (27). Now, suppose by contradiction that the minimum of $F(\cdot)$ on $\Gamma$ is some $\boldsymbol{x}^{**} \neq \boldsymbol{x}^*$. Then, by Lemma 2, $x_i^{**} > 0$ for all $i$. Therefore we can extend the open set on which $F(\cdot)$ is differentiable to include both $\boldsymbol{x}^*$ and $\boldsymbol{x}^{**}$. Then both $\boldsymbol{x}^*$ and $\boldsymbol{x}^{**}$ satisfy the first-order conditions (28), which gives a contradiction given that $F(\cdot)$ is strictly convex and the constraint set is convex. Hence, $\boldsymbol{x}^*$ is the minimum of $F(\cdot)$ on $\Gamma$. □

Combing the results from Lemmas 1-3, we get a proof of the part of Proposition 1 concerning the case of $0 < \alpha \leq 1$.

#### C.1.1. Continuity of solution with respect to $\alpha$

One might wonder if the equilibrium labor allocation is continuous in $\alpha$ as we approach $\alpha = 1$ from below. Economically speaking, one would expect this to be the case, so that if at $\alpha = 1$ we have a corner allocation with $x_i = 0$ for some country $i$ then $x_i(\alpha) > 0$ for all $\alpha < 1$ but $\lim_{\alpha \uparrow 1} x_i(\alpha) \to 0$. Mathematically, however, this result is not trivial because the function $\boldsymbol{G}$ is not jointly continuous in $\boldsymbol{x}$ and $\alpha$ for $\alpha = 1$ and points $\boldsymbol{x}$ with $x_i = 0$ for some $i$. Still, thanks to the optimization approach followed in the previous lemmas, we can establish the left continuity of $\boldsymbol{x}(\alpha)$.

**Lemma 4.** *If Assumption 1 holds, then $\boldsymbol{x}(\alpha)$ is continuous as a function of $\alpha$ for all $\alpha \in (0,1]$. In particular,* $\lim_{\alpha \uparrow 1} \boldsymbol{x}(\alpha) = \boldsymbol{x}(1)$.

*Proof.* Let us formally bring argument $\alpha$ into the notation of function $F$ defined in (25), i.e., consider the function $F(\boldsymbol{x};\alpha)$. Lemma 1 establishes that under Assumption 1 the solution to the optimization problem $\min_{x \in \Gamma} F(\boldsymbol{x};\alpha)$ defines a function $\boldsymbol{x} : (0,1] \to \mathbb{R}_+^N \setminus \{\boldsymbol{0}\}$. Clearly, $F(\boldsymbol{x};\alpha)$ is continuous for all $\boldsymbol{x} \in \mathbb{R}_+^N \setminus \{\boldsymbol{0}\}$ and $\alpha \in (0,1]$. $\Gamma$ is a compact set which is the same for all $\alpha \in (0,1]$. Thus, all conditions for Theorem of the Maximum (Theorem 3.6) from Stokey, Lucas and Prescott (1989) are statisfied, and $\boldsymbol{x}(\alpha)$ is continuous for all $\alpha \in (0,1]$. □

#### C.1.2. Sufficient conditions for Assumption 1

While it is easy to check if Assumption 1 is satisfied for a particular parametrization, we can say a little bit more about the conditions which guarantee that this assumption holds. Behrens, Lamorgese, Ottaviano and Tabuchi (2004) invoke classical results by Schoenberg (1938) to show that, if trade costs $\tau_{ni}$ correspond to the Euclidean distance between countries $n$ and $i$, then the matrix in Assumption 1 is positive definite (and, hence, non-singular) as long as all countries are at distinct locations. In fact, any three distinct numbers that satisfy the triangle inequality can be mapped to lengths of sides of a triangle in $\mathbb{R}^2$, which means that any such numbers correspond to Euclidean distances between vertices of a triangle in $\mathbb{R}^2$. Together with the results from Schoenberg (1938), this observation implies that for $N = 3$ the matrix in Assumption 1 is positive definite if (i) the iceberg trade costs are symmetric, (ii) greater than 1 for different countries, and (iii) satisfy the triangle inequality. For $N > 3$, conditions (i)-(iii) do not generally imply that the iceberg trade costs correspond to distances in an Euclidean space. Still, extensive simulations for trade-freeness matrices for $N = 4,5,6$ lead us to conjecture that conditions (i)-(iii) guarantee that the matrix in Assumption 1 is positive definite. Moreover, we conjecture that we can even dispense with the symmetry condition (i) — in this case it is the sum of the matrix in Assumption 1 with its transpose that is positive definite.

### C.2. Proof of Proposition 2

We start the proof of Proposition 2 with two lemmas:

**Lemma 5.** *If either (a) $0 \le \alpha_k < 1$, or (b) $\alpha_k = 1$ and Assumption 1 holds, then the function $\boldsymbol{L}_k(\boldsymbol{w})$ is continuous for all $\boldsymbol{w} \in \mathbb{R}^N_{++}$.*

**Lemma 6.** *If $\alpha_k = 1$, then the solution to (4) determines a non-empty convex-valued upper hemi-continuous correspondence $\mathscr{L}_k(\boldsymbol{w})$ for all $\boldsymbol{w} \in \mathbb{R}^N_{++}$.*

**Proof of Lemmas 5 and 6.** We prove Lemmas 5 and 6 simultaneously. The case with $\alpha_k = 0$ is trivial because labor allocations are explicitly obtained from the goods market clearing conditions $L_{i,k}G_{i,k}(\boldsymbol{w},\boldsymbol{L}_k) = 0$, and the resulting expressions for $L_{i,k}(\boldsymbol{w})$ are obviously continuous. Below we focus on the case with $\alpha_k \in (0,1]$.

Define a multi-valued correspondence $\Gamma_k : \mathbb{R}^N_{++} \to \mathbb{R}^N_+ \setminus \{\boldsymbol{0}\}$ by

$$\Gamma_k(\boldsymbol{w}) = \left\{ \boldsymbol{L}_k \in \mathbb{R}^N \,|\, L_{i,k} \ge 0, \sum_i w_i L_{i,k} = \sum_i \beta_{i,k} w_i \bar{L}_i \right\}.$$

Define function $F_k : \left(\mathbb{R}_+^N \setminus \{\mathbf{0}\}\right) \times \mathbb{R}_{++}^N \to \mathbb{R}$ by

$$F_k(\boldsymbol{L}_k; \boldsymbol{w}) = \alpha_k \sum_n w_n L_{n,k} - \sum_n \beta_{n,k} w_n \bar{L}_n \ln\left(\sum_i S_{i,k} L_{i,k}^{\alpha_k} \left(w_i \tau_{ni,k}\right)^{-\varepsilon_k}\right).$$

Denote the set of labor allocations at which $F_k(\boldsymbol{L}_k; \boldsymbol{w})$ achieves its minimum on $\Gamma_k(\boldsymbol{w})$ by $\mathscr{L}_k(\boldsymbol{w}) \equiv \arg\min_{\boldsymbol{L}_k \in \Gamma_k(\boldsymbol{w})} F_k(\boldsymbol{L}_k; \boldsymbol{w})$. It is straightforward to show that $\Gamma_k(\boldsymbol{w})$ is both lower hemi-continuous and upper hemi-continuous for all $\boldsymbol{w} \in \mathbb{R}_{++}^N$ (see the corresponding definitions in Stokey, Lucas and Prescott, 1989). Hence, $\Gamma_k(\boldsymbol{w})$ is continuous for all $\boldsymbol{w} \in \mathbb{R}_{++}^N$. Clearly, $\Gamma_k(\boldsymbol{w})$ is also compact-valued for all $\boldsymbol{w} \in \mathbb{R}_{++}^N$. Function $F_k(\boldsymbol{L}_k; \boldsymbol{w})$ is continuous for all $\boldsymbol{L}_k \in \mathbb{R}_+^N \setminus \{\mathbf{0}\}$ and $\boldsymbol{w} \in \mathbb{R}_{++}^N$. Thus, all conditions for Theorem 3.6 (Theorem of the Maximum) from Stokey, Lucas and Prescott (1989) are satisfied, and the correspondence $\mathscr{L}_k : \mathbb{R}_{++}^N \to \mathbb{R}_+^N \setminus \{\mathbf{0}\}$ is nonempty and upper hemi-continuous.

Lemma 1 establishes that under conditions (a) or (b) $F_k(\boldsymbol{L}_k; \boldsymbol{w})$ is strictly convex in $\boldsymbol{L}_k$ and, hence, $\mathscr{L}_k(\boldsymbol{w})$ is a singleton. In this case upper hemi-continuity of $\mathscr{L}_k(\boldsymbol{w})$ simply means continuity. Lemma 3 implies that all global minima of $F_k(\cdot; \boldsymbol{w})$ on $\Gamma_k(\boldsymbol{w})$ are solutions to problem (4). Therefore, under conditions (a) or (b) the solution to (4) defines a continuous function $\boldsymbol{L}_k(\boldsymbol{w})$ from wages to labor allocations.

If $\alpha_k = 1$ and Assumption 1 does not hold, function $F_k(\boldsymbol{L}_k; \boldsymbol{w})$ is convex in $\boldsymbol{L}_k$, but not necessarily strictly convex. Then, since $\Gamma_k(\boldsymbol{w})$ is a convex set, $\mathscr{L}_k(\boldsymbol{w})$ is also convex. Again, Lemma 3 implies that $\mathscr{L}_k(\boldsymbol{w})$ consists of all solutions to problem (4). So, in this case, solution to (4) determines a correspondence $\mathscr{L}_k(\boldsymbol{w})$ between wages and equilibrium labor allocations which is non-empty, convex-valued, and upper hemi-continuous. □

**Proof of Proposition 2.** Without loss of generality assume that $\alpha_k > 1$ for $k = 1, \ldots, K^*$ and $0 \le \alpha_k \le 1$ for $k = K^* + 1, \ldots, K$ and consider the following three cases: (a) $0 \le K^* < K$; (b) $K^* = K$ and $K < N$; (c) $K^* = K$ and $K \ge N$. In what follows, refer to Figure 5 for an illustration of the patterns of labor allocations that we choose for the cases (a)-(c). In this figure, rows of matrices correspond to countries and columns correspond to industries. In the next paragraph we formally define these patterns.

If we are in case (a), then for $i = 1, \ldots, N-1$ and $k = 1, \ldots, K^*$ set $L_{i,k} = 0$. In this case industries $k = 1, \ldots, K^*$ are arbitrary chosen to be supplied by country $N$ only. Next, if

$$\begin{pmatrix} 0 & \dots & 0 & L_{1,K^*+1} & \dots & L_{1,K} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & L_{N-1,K^*+1} & \dots & L_{N-1,K} \\ L_{N,1} & \dots & L_{N,K^*} & L_{N,K^*+1} & \dots & L_{N,K} \end{pmatrix}$$

Case (a)

$$\begin{pmatrix} \bar{L}_1 & & & \\ & \ddots & 0 & \\ & 0 & \ddots & \\ & & & \bar{L}_{K^*} \\ 0 & \dots & 0 & \bar{L}_{K^*+1} \\ \vdots & & \vdots & \vdots \\ 0 & \dots & 0 & \bar{L}_N \end{pmatrix}$$

Case (b)

$$\begin{pmatrix} \bar{L}_1 & & 0 & 0 & \dots & 0 \\ & \ddots & & \vdots & & \vdots \\ 0 & & \bar{L}_{N-1} & 0 & \dots & 0 \\ 0 & \dots & 0 & L_{N,N} & \dots & L_{N,K^*} \end{pmatrix}$$

Case (c)

Figure 5: Labor Allocation Patterns in Proposition 2

we are in case (b), then for $i = 1,\dots,K^*$ and $k = 1,\dots,K^*$ set $L_{i,k} = 0$ if $i \neq k$; and for $i = K^*+1,\dots,N$ and $k = 1,\dots,K^*-1$ set $L_{i,k} = 0$. In this case we arbitrary assign each country with index $i = 1,\dots,K^*-1$ to be the only supplier of the industry with the corresponding index $k = 1,\dots,K^*-1$, while the remaining countries allocate all their labor to industry $K^*$. Finally, if we are in case (c), then for $i = 1,\dots,N-1$ and $k = 1,\dots,K$ set $L_{i,k} = 0$ if $i \neq k$; and for $k = 1,\dots,N-1$ set $L_{N,k} = 0$. In this case, similarly to case (b), each country with index $i = 1,\dots,N-1$ is the only supplier of the industry with the corresponding index $k = 1,\dots,N-1$, while the remaining industries are all supplied by country $N$ only.

In cases (b) and (c), when some country $i$ allocates all its labor to only one industry, we label the corresponding entries of the labor allocation matrices in Figure 5 by $\bar{L}_i$. At the same time, in the formal definitions of the labor allocation patterns above we do not explicitly set labor allocations in the corresponding cases to the full labor endowments. The reason is that we are going to use the two-step definition of equilibrium from the main text to prove existence. In the first step we fix wages and derive equilibrium labor allocations, and in the second step we find wages that clear labor markets. So, labeling of non-zero entries in Figure 5 shall be understood as equilibrium outcomes rather than predetermined allocations.

It is easy to verify that for all country-industry pairs $(i,k)$ for which we assigned $L_{i,k} = 0$ the corresponding complementary slackness conditions (4) are satisfied for any positive vector of wages. This is because in all these cases we have $\alpha_k > 1$. For all other cases we can either explicitly (for $\alpha_k = 0$ or $\alpha_k > 1$) or implicitly (for $0 < \alpha_k \le 1$) solve (4) to find (first-step) equilibrium labor allocations.

Importantly for what follows, the allocations described in cases (a)-(c) imply that for any country $i$ we have the following three mutually exclusive possibilities:

1. There is some industry $k$ for which country $i$ is the only supplier. In this case country $i$'s equilibrium labor allocation in industry $k$ is given by

$$L_{i,k}(\boldsymbol{w}) = \frac{1}{w_i}\sum_n \beta_{n,k} w_n \bar{L}_n. \tag{29}$$

2. Country $i$ allocates all its labor to some industry $k$ that is supplied by multiple countries each of which allocates all its labor to this industry — this happens in case (b) above if $i \ge K^*$. The equilibrium labor allocation in industry $k$ is given by

$$L_{i,k}(\boldsymbol{w}) = \sum_n \frac{S_{i,K^*}\left[\tau_{ni,K^*}\right]^{-\varepsilon_{K^*}} \bar{L}_i^{\alpha_{K^*}} w_i^{-\varepsilon_{K^*}-1}}{\sum_{l\ge K^*} S_{l,K^*}\left[\tau_{nl,K^*}\right]^{-\varepsilon_{K^*}} \bar{L}_l^{\alpha_{K^*}} w_l^{-\varepsilon_{K^*}}} \beta_{n,K^*} w_n \bar{L}_n, \tag{30}$$

while $L_{i,k'}(\boldsymbol{w}) = 0$ for all $k' \ne k$.

3. Country $i$ allocates all its labor to industries with $\alpha_k \le 1$. In this case country $i$'s equilibrium labor allocations satisfy (4), which defines a function $L_{i,k}(\boldsymbol{w})$ if $\alpha_k < 1$ and a correspondence $\mathscr{L}_{i,k}(\boldsymbol{w})$ if $\alpha_k = 1$.

Let

$$\mathcal{Z}_i(\boldsymbol{w}) \equiv \left\{ \sum_k L_{i,k} - \bar{L}_i \,\middle|\, L_{i,k} = L_{i,k}(\boldsymbol{w}) \text{ if } \alpha_k \ne 1 \text{ and } L_{i,k} \in \mathscr{L}_{i,k}(\boldsymbol{w}) \text{ if } \alpha_k = 1 \right\}$$

be the excess labor demand correspondence in country $i$, and let

$$\mathcal{Z}(\boldsymbol{w}) \equiv (\mathcal{Z}_1(\boldsymbol{w}), \ldots, \mathcal{Z}_N(\boldsymbol{w})).$$

We are going to use Theorem 8 from Debreu (1982) to show that there exists a positive vector of wages $\boldsymbol{w}$ such that $\mathbf{0} \in \mathcal{Z}(\boldsymbol{w})$. For that we need to verify that $\mathcal{Z}$ satisfies the following properties: (i) $\mathcal{Z}$ is homogeneous of degree zero;[40] (ii) $\mathcal{Z}$ is convex-valued; (iii) $\mathcal{Z}$ is bounded below; (iv) $\mathcal{Z}$ is upper hemi-continuous; (v) (Walras' Law) $\sum w_i Z_i = 0$ for any $\boldsymbol{w} \in \mathbb{R}^N_{++}$ and any $(Z_1,\dots,Z_N) \in \mathcal{Z}(\boldsymbol{w})$; (vi) (Boundary Condition) if $\{\boldsymbol{w}^t\}_{t=1}^{\infty}$ is a wage sequence such that $\boldsymbol{w}^t \to \boldsymbol{w}$ as $t \to \infty$, where $\boldsymbol{w} \neq 0$ is a finite vector of wages and $w_i = 0$ for some $i$, then for any sequence $\left(Z_1^t,\dots,Z_N^t\right) \in \mathcal{Z}(\boldsymbol{w}^t)$ for $t = 1,2,\dots$, we have $\max\{Z_1^t,\dots,Z_N^t\} \to \infty$ as $t \to \infty$.

It is immediate to see that $\mathcal{Z}(\boldsymbol{w})$ is homogeneous of degree zero, that Walras' Law is satisfied for any positive $\boldsymbol{w}$, and that $Z_i > -\bar{L}_i$ for any $(Z_1,\dots,Z_N) \in \mathcal{Z}(\boldsymbol{w})$ and all positive $\boldsymbol{w}$. The property that $\mathcal{Z}(\boldsymbol{w})$ is convex-valued follows from the fact that $\mathcal{Z}(\boldsymbol{w})$ consists of the sum of functions $L_{i,k}(\boldsymbol{w})$ and correspondences $\mathcal{L}_{i,k}(\boldsymbol{w})$ which are convex-valued by virtue of Lemma 6. Upper hemi-continuity of $\mathcal{Z}(\boldsymbol{w})$ follows from upper hemi-continuity of $\mathcal{L}_{i,k}(\boldsymbol{w})$ established in Lemma 6 and from the fact $L_{i,k}(\boldsymbol{w})$ are given by (29), or by (30), or by the solution to (4), which is continuous by Lemma 5.

The only non-trivial condition to check is the boundary condition (vi). Consider any wage sequence from this condition. Let index $j$ be such that wage $w_j^t$ converges to 0 weakly "faster" than other wages. Formally, index $j$ is such that the limit $\lim_{t\to\infty} w_j^t / w_l^t$ is finite for all $l$. Such index always exists because there is a finite number of indices.

Consider the three possibilities above. Under the first possibility, country $j$ is the only supplier of some industry $k$, therefore by expression (29) we have $L_{j,k}\left[\boldsymbol{w}^t\right] = \sum_n \beta_{n,k}\left[w_n^t / w_j^t\right]\bar{L}_n$. This converges to $\infty$ as $t \to \infty$, and so $\max\{Z_1^t,\dots,Z_N^t\} \to \infty$ for any sequence $\left(Z_1^t,\dots,Z_N^t\right) \in \mathcal{Z}(\boldsymbol{w}^t)$ for $t = 1,2,\dots$.

Under the second possibility, country $j$'s excess labor demand function is given by

$$Z_j(\boldsymbol{w}) = \sum_n \frac{S_{j,K^*}\left[\tau_{nj,K^*}\right]^{-\varepsilon_{K^*}} \bar{L}_j^{\alpha_{K^*}} (w_j^t)^{-1}}{\sum\limits_{l \geq K^*} S_{l,K^*}\left[\tau_{nl,K^*}\right]^{-\varepsilon_{K^*}} \bar{L}_l^{\alpha_{K^*}} \left[w_j^t / w_l^t\right]^{\varepsilon_{K^*}}} \beta_{n,K^*} w_n^t \bar{L}_n - \bar{L}_i.$$

The denominator of any term in the above sum converges to a finite positive number as $t \to \infty$. The numerator converges either to a finite positive number or to infinity. Moreover, since for at least one index $n$ wage $w_n^t$ converges to a positive number and

[40]Homogeneity of degree zero is not explicitly mentioned in Theorem 8 in Debreu (1982). Instead, the excess demand correspondence is assumed to be defined on a simplex of prices. For our purposes this is the same as assuming homogeneity of degree zero.

$w_j^t$ converges to 0, we have that for at least one index $n$ the numerator of the corresponding term in the above sum converges to $\infty$. Hence, the whole sum converges to $\infty$. Therefore, again, the boundary condition is satisfied.

Finally, under the third possibility country $j$ supplies all its labor to industries with $\alpha_k \le 1$. Pick any such industry $k$. Equilibrium labor allocations to industry $k$ in all countries satisfy (4) (see case (a) in Figure 5). Let us use the general notation $\mathscr{L}_{i,k}(\boldsymbol{w})$ for all such labor allocations (if $\alpha_k = 1$, then $\mathscr{L}_{i,k}(\boldsymbol{w})$ is a singleton). If there is some country $i$ with the corresponding sequence of sets $\mathscr{L}_{i,k}(\boldsymbol{w}^t)$ such that any sequence $L_{i,k}^t \in \mathscr{L}_{i,k}(\boldsymbol{w}^t)$ converges to $\infty$ as $t \to \infty$, then the boundary condition is satisfied. Let us show that there always exists such a country by supposing the contrary. That is, suppose that for any country $i$ there exists a sequence $L_{i,k}^t \in \mathscr{L}_{i,k}(\boldsymbol{w}^t)$ converging to a finite number as $t \to \infty$. That means that there exists a sequence $\left(L_{1,k}^t, \ldots, L_{N,k}^t\right) \in \mathscr{L}_k(\boldsymbol{w}^t)$ converging to a finite vector as $t \to \infty$. Consider this sequence. Let us focus again on the country $j$ for which the wage converges to 0 weakly "faster" than for any other country. For any $t$, $L_{j,k}^t$ satisfies (4) and, in particular, $L_{j,k}^t$ satisfies the inequality

$$w_j^t \ge \sum_n \frac{S_{j,k}\left[L_{j,k}^t\right]^{\alpha_k-1}\left[w_j^t \tau_{nj,k}\right]^{-\varepsilon_k}}{\sum_l S_{l,k}\left[L_{l,k}^t\right]^{\alpha_k}\left[w_l^t \tau_{nl,k}\right]^{-\varepsilon_k}} \beta_{n,k} w_n^t \bar{L}_n.$$

This inequality can be equivalently rewritten as

$$w_j^t \ge \sum_{n=1}^{N} \frac{S_{j,k}\left[L_{j,k}^t\right]^{\alpha_k-1} \tau_{nj,k}^{-\varepsilon_k}}{\sum_l S_{l,k}\left[L_{l,k}^t\right]^{\alpha_k} \tau_{nl,k}^{-\varepsilon_k}\left[w_j^t / w_l^t\right]^{\varepsilon_k}} w_n^t \bar{L}_n. \tag{31}$$

The denominator of any term in the above summation (31) converges to a finite number (which can be either positive or zero). The numerator of any term in the summation (31) converges to either a finite positive number or to infinity. Also, there exists at least one index $n$ such that $\lim_{t\to\infty} w_n^t > 0$. Then, for this index $n$ the corresponding term in the summation (31) converges to either a finite positive number or to $\infty$. This, in turn, implies that the whole sum in (31) converges to either a finite positive number or to $\infty$. At the same time, the left-hand side of inequality (31) converges to 0. A contradiction. □

### C.3. Proof of Proposition 3

If $\alpha_k > 1$ for some $k$ then in the proof of Proposition 2 $K^* > 0$. This implies that there are different allocations that we can assign (i.e., one for each country), and since there is an equilibrium for each one, this immediately establishes that there are multiple equilibria. □

### C.4. Proof of Proposition 4

This proof proceeds by showing that $\boldsymbol{Z}(\boldsymbol{w})$ satisfies the gross substitutes property (GSP). Uniqueness of wages then follows from Proposition 17.F.3 from Mas-Colell, Whinston and Green (1995).

Consider any particular industry $k$. Let us separately analyze the two possibilities $0 \le \alpha_k < 1$ and $\alpha_k = 1$.

If $0 \le \alpha_k < 1$, then for any $i$ we have that $L_{i,k}(\boldsymbol{w}) > 0$ for any wage vector $\boldsymbol{w} \in \mathbb{R}^N_{++}$, and $L_{i,k}(\boldsymbol{w})$ solves:

$$w_i L_{i,k}(\boldsymbol{w}) = \sum_n \lambda_{ni,k}(\boldsymbol{w}, \boldsymbol{L}_k(\boldsymbol{w}))\beta_{n,k} w_n \bar{L}_n.$$

By differentiating both sides of this expression w.r.t. wages, we can get a linear system of equations which determines the effect of wages on labor allocations. Let us introduce additional notation to write in matrix form this effect. Denote $x_{ij,k} \equiv \dfrac{d \ln L_{i,k}(\boldsymbol{w})}{d \ln w_j}$, $q_{i,k} \equiv w_i L_{i,k}(\boldsymbol{w})$, and $b_{i,k} \equiv \beta_{i,k} w_i \bar{L}_i$. Let $B_k$ denote the diagonal matrix with elements $b_{i,k}$ along the diagonal, $Q_k$ the diagonal matrix with elements $q_{i,k}$ along the diagonal, $\Lambda_k$ the matrix of sector level expenditure shares $\lambda_{ij,k}$, and $X_k$ the matrix of partials $x_{ij,k}$. Finally, let $U_k \equiv \left((1-\alpha_k) Q_k + \alpha_k \Lambda_k^T B_k \Lambda_k\right)$ and $V_k \equiv \left(\Lambda_k^T B_k + \varepsilon_k \Lambda_k^T B_k \Lambda_k - (1+\varepsilon_k) Q_k\right)$. In this notation the effect of wages on labor allocations is obtained from the system:

$$U_k X_k = V_k.$$

It straightforward to check that matrix $U_k$ is a positive definite matrix with all positive elements, and matrix $V_k$ has negative diagonal and positive off-diagonal elements. Since $U_k$ is positive definite, the inverse exists and its determinant is positive. Moreover, $U_k^{-1} = \frac{1}{\det(U_k)} C_k^T$, where $C_k^T$ is the transpose of the matrix of cofactors $C_k$ of $U_k$. Since all the elements of $U_k$ are positive, then for $N=2$, $C_k$ is a $2\times 2$ matrix consisting of positive diagonal elements and negative off-diagonal elements.[41] Therefore, $U_k^{-1}$ has this property as well. One can then readily verify that $U_k^{-1}V_k$ is a matrix with the same properties as $V_k$ — it has negative diagonal and positive off diagonal elements. Thus the Jacobian matrix of wages effects on labor allocations in industry $k$ with $0\le\alpha_k<1$ satisfies the GSP.

If $\alpha_k=1$, then $L_{i,k}(\boldsymbol{w})$ can be equal to 0 for some $i$, and we cannot establish differentiability of labor allocations in that region. We are going to check directly what happens to labor allocations as wages change. To that end, assume without loss of generality that $\boldsymbol{w}'$ and $\boldsymbol{w}''$ are such that $w_1''>w_1'$ and $w_2''=w_2'=1$. Let us show that $L_{2,k}(\boldsymbol{w}'')\ge L_{2,k}(\boldsymbol{w}')$ for all $k$ and there is some industry $\tilde{k}$ such that $L_{2,\tilde{k}}(\boldsymbol{w}'')>L_{2,\tilde{k}}(\boldsymbol{w}')$.

In general, given wage $\boldsymbol{w}'$ there are three cases: (a) $L_{1,k}(\boldsymbol{w}')=0$ and $L_{2,k}(\boldsymbol{w}')=\beta_{1,k}w_1'\bar{L}_1+\beta_{2,k}\bar{L}_2$; (b) $L_{i,k}(\boldsymbol{w}')>0$ for $i=1,2$; (c) $L_{2,k}(\boldsymbol{w}')=0$ and
$L_{1,k}(\boldsymbol{w}')=\frac{1}{w_1'}\left(\beta_{1,k}w_1'\bar{L}_1+\beta_{2,k}\bar{L}_2\right)$.

Let us consider these different cases.

**Case (a).** In this case we have $G_{1,k}(\boldsymbol{w}')\ge 0$ and $G_{1,k}(\boldsymbol{w}')$ simplifies to:

$$G_{1,k}(\boldsymbol{w}')=w_1-\frac{S_{1,k}\left[w_1'\right]^{-\varepsilon_k}}{S_{2,k}L_{2,k}(\boldsymbol{w}')\tau_{12,k}^{-\varepsilon_k}}\beta_{1,k}w_1'\bar{L}_1-\frac{S_{1,k}\left(w_1'\tau_{21,k}\right)^{-\varepsilon_k}}{S_{2,k}L_{2,k}(\boldsymbol{w}')}\beta_{2,k}\bar{L}_2.$$

After substituting $L_{2,k}(\boldsymbol{w}')=\beta_{1,k}w_1'\bar{L}_1+\beta_{2,k}\bar{L}_2$ into the above expression for $G_{1,k}(\boldsymbol{w}')$, and dividing both sides of this expression by $w_1'$, we get:

$$\frac{G_{1,k}(\boldsymbol{w}')}{w_1'}=1-\frac{S_{1,k}}{S_{2,k}\tau_{12,k}^{-\varepsilon_k}}\cdot\frac{\left[w_1'\right]^{-\varepsilon_k}\beta_{1,k}\bar{L}_1}{\beta_{1,k}w_1'\bar{L}_1+\beta_{2,k}\bar{L}_2}-\frac{S_{1,k}\tau_{21,k}^{-\varepsilon_k}}{S_{2,k}}\cdot\frac{\left[w_1'\right]^{-1-\varepsilon_k}\beta_{2,k}\bar{L}_2}{\beta_{1,k}w_1'\bar{L}_1+\beta_{2,k}\bar{L}_2}.$$

Clearly, the right-hand side of this expression is increasing in $w_1'$. Hence, $G_{1,k}(\boldsymbol{w}'')/w_1''>G_{1,k}(\boldsymbol{w}')/w_1'\ge 0$, which in turn implies that $G_{1,k}(\boldsymbol{w}'')>0$. Therefore, $L_{1,k}(\boldsymbol{w}'')=0$ and $L_{2,k}(\boldsymbol{w}'')=\beta_{1,k}w_1''\bar{L}_1+\beta_{2,k}\bar{L}_2$ solve the complementary slackness problem (4). In other words, we still remain in case (a) after we increase the wage of the country 1 from $w_1'$ to $w_1''$. Clearly, in this case $L_{2,k}(\cdot)$ is a strictly increasing function of the wage of the first country, $L_{2,k}(\boldsymbol{w}'')>L_{2,k}(\boldsymbol{w}')$.

**Case (b).** We know that, as long as we are in case (b), $L_{1,k}(\cdot)$ is a decreasing function

[41] This is no longer true with $N>2$.

and $L_{2,k}(\cdot)$ is an increasing function of $w_1$. Therefore, starting in case (b) with $\boldsymbol{w}'$ and gradually increasing $w_1$ from $w_1'$ to $w_1''$, we either remain in case (b) or switch to case (a) at some point. The above argument for case (a) implies that, once we switch to case (a), we will remain in case (a) as we keep increasing $w_1$. Thus, for $\boldsymbol{w}''$ we can either be in case (a) or in case (b), but not in case (c), and since in both cases (a) and (b) $L_{2,k}(\cdot)$ is a strictly increasing function of $w_1$, we must have $L_{2,k}(\boldsymbol{w}'') > L_{2,k}(\boldsymbol{w}')$.

**Case (c).** In this case, we can be in any of the cases (a)-(c) for $\boldsymbol{w}''$. If we are in cases (a) or (b) for $\boldsymbol{w}''$, then $L_{2,k}(\boldsymbol{w}'') > L_{2,k}(\boldsymbol{w}') = 0$. If we are in case (c) for $\boldsymbol{w}''$, then $L_{2,k}(\boldsymbol{w}'') = L_{2,k}(\boldsymbol{w}') = 0$, but there must exist some industry $\tilde{k}$, for which we are in case (a) or (b) for $\boldsymbol{w}'$ (regardless of the value of $\alpha_{\tilde{k}}$ in this industry). Applying the arguments above, for any such industry we have $L_{2,\tilde{k}}(\boldsymbol{w}'') > L_{2,\tilde{k}}(\boldsymbol{w}')$.

Since the effect of changes in wages on $\boldsymbol{Z}(\boldsymbol{w})$ consists of the sum (across industries) of effects on industry-level labor allocations, we conclude that $\boldsymbol{Z}(\boldsymbol{w})$ satisfies the GSP. □

## C.5. Proof of Proposition 5

In this proof we will use matrices $B_k$, $Q_k$, $\Lambda_k$, $U_k$, and $V_k$ defined in the proof of Proposition 4 in Appendix C.4. In addition to that, let $L_k$ be a diagonal matrix with elements $L_{i,k}$ along the diagonal; $W$ be a diagonal matrix with elements $w_i$ along the diagonal; $D\boldsymbol{Z}(\boldsymbol{w})$ be the Jacobian matrix of the excess demand system, $Z(w)$, with elements $\partial Z_i(\boldsymbol{w})/\partial w_j$; and $D\boldsymbol{L}_k(\boldsymbol{w})$ be the Jacobian matrix of industry-level labor allocations with elements $\partial L_{i,k}(\boldsymbol{w})/\partial w_j$.

We have

$$D\boldsymbol{Z}(\boldsymbol{w}) = \sum_k D\boldsymbol{L}_k(\boldsymbol{w}) = \sum_k L_k U_k^{-1} V_k W^{-1}.$$

Matrix $V_k$ has the following properties: (i) entries in each row add up to 0; (ii) diagonal entries are negative; (iii) off-diagonal entries are positive. For all industries $k$ with $\alpha_k = 0$ matrix $U_k$ reduces to diagonal matrix $Q_k$ with positive diagonal elements. Therefore, we can immediately conclude that for all such industries $D\boldsymbol{L}_k(\boldsymbol{w})$ has properties (i)-(iii) as well. The rest of this appendix section is devoted to proving that for all industries $k$ with $0 < \alpha_k < 1$ matrix $D\boldsymbol{L}_k(\boldsymbol{w})$ also has properties (i)-(iii) under free trade. Since summation of matrices with properties (i)-(iii) again gives a matrix with these properties, the whole

Jacobian of the excess demand system, $D\boldsymbol{Z}(\boldsymbol{w})$, has properties (i)-(iii) under free trade. This means that the excess demand system $\boldsymbol{Z}(\boldsymbol{w})$ has the gross substitutes property. Hence, there is at most one normalized vector of wages such that $\boldsymbol{Z}(\boldsymbol{w}) = \mathbf{0}$.

Consider any industry $k$ with $0 < \alpha_k < 1$. For brevity of notation we drop the industry index $k$ in the rest of this proof. According to Proposition 1 all industry-level labor allocations are interior, and so, $L_i > 0$, $\lambda_{ii} > 0$, $q_i > 0$ for all $i$. We start with three lemmas which apply to the general case of costly trade.

**Lemma 7.** *Let $\mu_1, \ldots, \mu_N$ be eigenvalues of matrix $Q^{-1}\Lambda^T B\Lambda$. Then $\mu_i$ is real and $0 \le \mu_i \le 1$ for each $i$.*

*Proof.* Consider matrix $Q^{-1/2}\Lambda^T B\Lambda Q^{-1/2}$, and let $\mu$ be any eigenvalue of this matrix with the corresponding eigenvector $v$. By definition of an eigenvalue, $Q^{-1/2}\Lambda^T B\Lambda Q^{-1/2} v = \mu v$. This is equivalent to $Q^{-1}\Lambda^T B\Lambda\left(Q^{-1/2}v\right) = \mu\left(Q^{-1/2}v\right)$. Hence, $\mu$ is an eigenvalue of $Q^{-1}\Lambda^T B\Lambda$ with the corresponding eigenvector $Q^{-1/2}v$. Therefore, matrices $Q^{-1}\Lambda^T B\Lambda$ and $Q^{-1/2}\Lambda^T B\Lambda Q^{-1/2}$ have the same eigenvalues, and so $\mu_1, \ldots, \mu_N$ are eigenvalues of $Q^{-1/2}\Lambda^T B\Lambda Q^{-1/2}$.

Clearly, matrix $Q^{-1/2}\Lambda^T B\Lambda Q^{-1/2}$ is positive semi-definite. Hence, all its eigenvalues are real and nonnegative, i.e., $\mu_i$ is real and $\mu_i \ge 0$ for each $i$. Next, matrix $Q^{-1}\Lambda^T B\Lambda$ is a positive stochastic matrix (its entries in each row add up to 1). Therefore, the Perron-Frobenius theorem implies that 1 is its eigenvalue with algebraic multiplicity one and $|\mu_i| < 1$ for any $|\mu_i| \ne 1$. Since $\mu_i \ge 0$ for all $i$, we have the statement of the lemma. □

**Lemma 8.** $\lim_{t\to\infty} \alpha^t \left(I_N - Q^{-1}\Lambda^T B\Lambda\right)^t = 0$.

*Proof.* Eigenvalues of matrix $I_N - Q^{-1}\Lambda^T B\Lambda$ are $1-\mu_1, \ldots, 1-\mu_N$, where $\mu_1, \ldots, \mu_N$ are eigenvalues of $Q^{-1}\Lambda^T B\Lambda$. Lemma 7 implies that $0 \le 1-\mu_i \le 1$ for all $i$. Then, since eigenvalues of matrix $\alpha\left(I_N - Q^{-1}\Lambda^T B\Lambda\right)$ are $\alpha\left(1-\mu_1\right), \ldots, \alpha\left(1-\mu_1\right)$, we have that $\rho\left(\alpha\left[I_N - Q^{-1}\Lambda^T B\Lambda\right]\right) < 1$, where $\rho(\cdot)$ is the spectral radius of a matrix. Therefore, $\alpha^t\left(I_N - Q^{-1}\Lambda^T B\Lambda\right)^t \to 0$ as $t \to \infty$ (see, for example, Theorem 5.6.12 in Horn and Johnson, 2013). □

**Lemma 9.** $U^{-1}V = \varepsilon\alpha^{-1} I_N - \sum_{t=0}^{\infty} \alpha^t \left(I_N - Q^{-1}\Lambda^T B\Lambda\right)^t \left[\left(1+\varepsilon\alpha^{-1}\right) I_N - Q^{-1}\Lambda^T B\right]$.

*Proof.* Consider $U^{-1}$:

$$U^{-1} = \left[(1-\alpha)\,Q + \alpha\Lambda^T B\Lambda\right]^{-1} = \left[(1-\alpha)\,I_N + \alpha Q^{-1}\Lambda^T B\Lambda\right]^{-1} Q^{-1}$$
$$= \left[I_N - \alpha\left(I_N - Q^{-1}\Lambda^T B\Lambda\right)\right]^{-1} Q^{-1}.$$

Lemma 8 implies that we can write

$$\left[I_N - \alpha\left(I_N - Q^{-1}\Lambda^T B\Lambda\right)\right]^{-1} = \sum_{t=0}^{\infty} \alpha^t \left(I_N - Q^{-1}\Lambda^T B\Lambda\right)^t$$

(see, for example, Corollary 5.6.15 in Horn and Johnson, 2013). Then

$$U^{-1}V = -\sum_{t=0}^{\infty} \alpha^t \left(I_N - Q^{-1}\Lambda^T B\Lambda\right)^t \left[(1+\varepsilon)\,I_N - Q^{-1}\Lambda^T B - \varepsilon Q^{-1}\Lambda^T B\Lambda\right]$$
$$= -\varepsilon \sum_{t=0}^{\infty} \alpha^t \left(I_N - Q^{-1}\Lambda^T B\Lambda\right)^{t+1} - \sum_{t=0}^{\infty} \alpha^t \left(I_N - Q^{-1}\Lambda^T B\Lambda\right)^t \left[I_N - Q^{-1}\Lambda^T B\right]$$
$$= \varepsilon\alpha^{-1} I_N - \sum_{t=0}^{\infty} \alpha^t \left(I_N - Q^{-1}\Lambda^T B\Lambda\right)^t \left[\left(1+\varepsilon\alpha^{-1}\right) I_N - Q^{-1}\Lambda^T B\right].$$

□

Let us now consider the case of frictionless trade. In this case the matrix of trade shares, $\Lambda$, has the same entries in each column:

$$\Lambda \equiv \begin{pmatrix} \lambda_{11} & \dots & \lambda_{NN} \\ \vdots & & \vdots \\ \lambda_{11} & \dots & \lambda_{NN} \end{pmatrix}.$$

So, it can be represented (with a slight abuse of notation) as $O\Lambda$ where $O$ is an $N \times N$ matrix of ones (i.e., $O = \iota \cdot \iota^T$ with $\iota^T \equiv (1, \dots, 1)$) and

$$\Lambda \equiv \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_N \end{pmatrix}.$$

Then, in this notation

$$U = (1-\alpha)\,Q + \alpha\Lambda OBO\Lambda \quad \text{and} \quad V = \Lambda OB + \varepsilon\Lambda OBO\Lambda - (1+\varepsilon)\,Q.$$

Denote $b \equiv \sum_n b_n$ and observe that $OBO = bO$. Also, since $L_i$ satisfies the goods market clearing condition, $w_i L_i = \sum_n \lambda_i b_n = \lambda_i \sum_n b_n = b\lambda_i$. Then, since in our notation $q_{ii} = w_i L_i$, we have that $Q = WL = b\Lambda$. These equalities together with Lemma 9 allow us to write:

$$U^{-1}V = \varepsilon\alpha^{-1} I_N - \sum_{t=0}^{\infty} \alpha^t \left(I_N - Q^{-1}\Lambda O B O \Lambda\right)^t \left[\left(1+\varepsilon\alpha^{-1}\right) I_N - Q^{-1}\Lambda O B\right]$$
$$= \varepsilon\alpha^{-1} I_N - \sum_{t=0}^{\infty} \alpha^t \left(I_N - O\Lambda\right)^t \left[\left(1+\varepsilon\alpha^{-1}\right) I_N - b^{-1} O B\right].$$

Using the fact that $\sum_i \lambda_i = 1$ and, hence, $O\Lambda O = O$, we get:

$$\left(I_N - O\Lambda\right)\left[\left(1+\varepsilon\alpha_k^{-1}\right) I_N - b^{-1} O B\right] = \left(1+\varepsilon\alpha^{-1}\right) I_N - b^{-1} O B$$
$$- \left(1+\varepsilon\alpha^{-1}\right) O\Lambda + b^{-1} O\Lambda O B$$
$$= \left(1+\varepsilon\alpha^{-1}\right)\left(I_N - O\Lambda\right),$$

and

$$\left(I_N - O\Lambda\right)\left(I_N - O\Lambda\right) = I_N - O\Lambda - O\Lambda + O\Lambda O\Lambda = I_N - O\Lambda.$$

Therefore,

$$U^{-1}V = \varepsilon\alpha^{-1} I_N - \left[\left(1+\varepsilon\alpha^{-1}\right) I_N - b^{-1} O B\right] - \left(1+\varepsilon\alpha^{-1}\right) \sum_{t=1}^{\infty} \alpha^t \left(I_N - O\Lambda\right)^t$$
$$= -\left(I_N - b^{-1} O B\right) - \left(1+\varepsilon\alpha^{-1}\right)\left(I_N - O\Lambda\right) \sum_{t=1}^{\infty} \alpha^t$$
$$= \left(b^{-1} O B - I_N\right) + \frac{\alpha+\varepsilon}{1-\alpha}\left(O\Lambda - I_N\right).$$

Observe that both matrices $\left(b^{-1}OB - I_N\right)$ and $\left(O\Lambda - I_N\right)$ have properties (i)-(iii) listed at the beginning of this appendix section. Hence, matrix $U^{-1}V$ has properties (i)-(iii) as well. This, in turn, implies that matrix $\Lambda U^{-1} V W^{-1}$ also has properties (i)-(iii). This concludes our proof. □

## C.6. Proof of Proposition 6

Fix country $i$ and let it be a small open economy in the sense that changes in its labor allocations and wage do not impact labor allocations, price indices, and wages in other countries. The equilibrium system for country $i$'s economy consists of $K+1$ conditions that are a subset of equilibrium conditions (3)-(5) for the world economy. Specifically, country $i$'s equilibrium conditions are $K$ goods market clearing conditions

$$L_{i,k} \geq 0, \quad G_{i,k}\left(w_i, L_{i,k}\right) \geq 0, \quad L_{i,k} G_{i,k}\left(w_i, L_{i,k}\right) = 0, \quad \text{for } k = 1, \ldots, K, \tag{32}$$

and one labor market clearing condition

$$\sum_k L_{i,k} - \bar{L}_i = 0, \tag{33}$$

where

$$G_{i,k}\left(w_i, L_{i,k}\right) \equiv w_i - \frac{S_{i,k}\beta_{i,k}\bar{L}_i}{S_{i,k}L_{i,k}^{\alpha_k} w_i^{-\varepsilon_k} + A_{i,k}} L_{i,k}^{\alpha_k - 1} w_i^{-\varepsilon_k + 1} - B_{i,k} S_{i,k} w_i^{-\varepsilon_k} L_{i,k}^{\alpha_k - 1}, \tag{34}$$

with

$$A_{i,k} \equiv \sum_{l \neq i} S_{l,k} L_{l,k}^{\alpha_k} \left(w_l \tau_{il,k}\right)^{-\varepsilon_k},$$

$$B_{i,k} \equiv \mu_k^{-\varepsilon_k} \sum_{n \neq i} \tau_{ni,k}^{-\varepsilon_k} P_{n,k}^{\varepsilon_k} \beta_{n,k} w_n \bar{L}_n.$$

The equilibrium of country $i$'s economy is given by wage $w_i$ and labor allocations $L_{i,k}$ for $k = 1, \ldots, K$ that solve (32)-(34). Here we slightly abuse the notation used in conditions (3)-(5) by dropping the dependence of $G_{i,k}$ on wages and labor allocations of countries different from $i$. This is done to emphasize the small open economy assumption.[42]

In what follows, drop country $i$'s index from notation for brevity. Fix wage $w$ and focus on the complementary slackness conditions (32). Consider first any industry $k$ with $\alpha_k < 1$. Condition $G_k(w, L_k) \geq 0$ can be written as:

$$1 \geq \frac{S_k \beta_k \bar{L}}{S_k L_k^{\alpha_k} w^{-\varepsilon_k} + A_k} L_k^{\alpha_k - 1} w^{-\varepsilon_k} + B_k S_k w^{-\varepsilon_k - 1} L_k^{\alpha_k - 1}. \tag{35}$$

Since $\alpha_k < 1$, the right-hand side of (35) goes to $\infty$ as $L_k \to 0$, while the left-hand side

[42]Observe that we assume that country $i$ can impact its own industry-level price indices given by $\mu_k\left(S_{i,k}L_{i,k}^{\alpha_k} w_i^{-\varepsilon_k} + A_{i,k}\right)^{-\frac{1}{\varepsilon_k}}$. This guarantees that, as we increase trade costs to infinity, country $i$'s production pattern converges to the production pattern in autarky.

does not depend on $L_k$. Hence, for a fixed wage $w$, only positive labor allocations can satisfy (35). The complementarity slackness condition (32) then implies that condition $G_k(w, L_k) \geq 0$ holds with equality. Next, the right-hand side of (35) is a decreasing function of $L_k$, and it falls from $\infty$ to 0 as $L_k$ increases from 0 and $\infty$. Hence, for each fixed wage $w$ there is a unique solution to $G_k(w, L_k) = 0$. In other words, equation $G_k(w, L_k) = 0$ defines an implicit function from wages to labor allocations, $L_k(w)$. Since the right-hand side of (35) is also a decreasing function of wages, the implicit function theorem implies that $L_k(\cdot)$ is a decreasing function. We can easily show that this function ranges from 0 to $\infty$ for $w \in (0, \infty)$. For that, rewrite $G_k(w, L_k) = 0$ as

$$L_k^{1-\alpha_k} = \frac{S_k \beta_k \bar{L}}{S_k L_k^{\alpha_k} w^{-\varepsilon_k} + A_k} w^{-\varepsilon_k} + B_k S_k w^{-\varepsilon_k - 1}.$$

As $w$ goes to $\infty$, the right-hand side of the above expression converges to 0, and, hence, $L_k(w)$ converges to 0. Similarly, as $w$ goes to 0, the right-hand side of the above expression converges to $\infty$, and, hence, $L_k(w)$ converges to $\infty$. By the implicit function theorem $L_k(\cdot)$ is a continuous function, and, therefore, it takes the whole range of values from 0 to $\infty$ as $w$ ranges from 0 to $\infty$.

Next, consider an industry $k$ with $\alpha_k = 1$. Write condition $G_k(w, L_k) \geq 0$ as

$$1 \geq \frac{S_k \beta_k \bar{L}}{S_k L_k w^{-\varepsilon_k} + A_k} w^{-\varepsilon_k} + S_k B_k w^{-\varepsilon_k - 1}. \tag{36}$$

If $S_k B_k w^{-\varepsilon_k - 1} \geq 1$ or, equivalently, if $w \leq (S_k B_k)^{\frac{1}{1+\varepsilon_k}}$, then (36) cannot be satisfied for any finite $L_k \geq 0$. So, in any equilibrium we must have $w > \underline{w}_k$ with $\underline{w}_k \equiv (S_k B_k)^{\frac{1}{1+\varepsilon_k}}$. Now consider equation $G_k(w, 0) = 0$:

$$1 = \frac{S_k \beta_k \bar{L}}{A_k} w^{-\varepsilon_k} + S_k B_k w^{-\varepsilon_k - 1}. \tag{37}$$

The left-hand side of (37) does not depend on $w$, while the right-hand side is a decreasing function of $w$, which falls from $\infty$ to 0 as $w$ increases from 0 to $\infty$. Hence, there is a unique $w$ — denote it by $\overline{w}_k$ — that solves (37), and $\overline{w}_k > \underline{w}_k$.

Next, consider any $w \in \left(\underline{w}_k, \overline{w}_k\right)$. The definition of $\overline{w}_k$ implies that, if $L_k = 0$, then the right-hand side of (36) is higher than 1. Therefore, we must have $L_k > 0$, and, hence, the complementarity slackness condition (32) has to hold with equality. Since the right-hand side of (36) is a decreasing function of $L_k$ and since $w > \underline{w}_k$, there exists a unique $L_k$ that solves $G_k(w, L_k) = 0$ for a given $w$. In other words, for $w \in (\underline{w}_k, \overline{w}_k)$, condition $G_k(w, L_k) = 0$ defines an implicit function $L_k(w)$ — the same as in the case with $\alpha_k < 1$. Moreover, as in the case with $\alpha_k < 1$, $L_k(\cdot)$ is a decreasing function. Importantly, $L_k(w) \to \infty$ as $w \to \underline{w}_k$.

Now consider $w \geq \overline{w}_k$. For such $w$ the right-hand side of (37) is weakly smaller than 1. Therefore, any positive $L_k$ will make the right-hand side of (36) strictly smaller than 1, while the complementary slackness condition (32) requires that for $L_k > 0$ condition (36) holds with equality. Hence, the only possibility to satisfy (32) for $w \geq \overline{w}_k$ is to have $L_k = 0$, which is the unique solution of (32) in this case. Furthermore, $L_k(w) \to 0$ as $w$ converges to $\overline{w}_k$ from the left.

The arguments in the above two paragraphs imply that for industries $k$ with $\alpha_k = 1$ condition (32) defines a function $L_k(w)$ for $w > \underline{w}_k$. This function is decreasing for $w \in (\underline{w}_k, \overline{w}_k)$, is zero for all $w \geq \overline{w}_k$, and it takes the full range of values from 0 to $\infty$ as $w$ varies from $\underline{w}_k$ to $\infty$.

Let us now turn to the labor market clearing condition (33). We can write the excess demand for labor as a function of the wage:

$$Z(w) = \sum_k L_k(w) - \bar{L} \quad \text{for } w > \underline{w},$$

where $\underline{w} \equiv \max\{\underline{w}_k \mid \text{for } k \text{ such that } \alpha_k = 1\}$ if there are industries $k$ with $\alpha_k = 1$ and $\underline{w} \equiv 0$ if there are no such industries. $Z(w)$ is a decreasing function $w$, and it falls from $\infty$ to $-\bar{L}$ as $w$ increases from $\underline{w}$ to $\infty$. Hence, there is a unique wage that solves $Z(w) = 0$.

□

## C.7. Applying Uniqueness Results in Allen, Arkolakis and Li (2016)

We can map the equilibrium system of our common framework into the system in Equation (3) in Allen, Arkolakis and Li (2016, henceforth AAL) and explore if their Theorem 2 can be invoked to establish uniqueness. Ignoring the inequality part of the complementary slackness conditions, the equilibrium conditions in the common framework can be written as

$$w_i^{1+\varepsilon_k} L_{i,k}^{1-\alpha_k} = \sum_n S_{i,k} \tau_{ni,k}^{-\varepsilon_k} P_{n,k}^{\varepsilon_k} \beta_{n,k} w_n \bar{L}_n,$$
$$P_{n,k}^{-\varepsilon_k} = \sum_l S_{l,k} \tau_{nl,k}^{-\varepsilon_k} w_l^{-\varepsilon_k} L_{l,k}^{\alpha_k},$$
$$w_i = \sum_s \frac{L_{i,s}}{\bar{L}_i} w_i.$$

To turn this into the AAL structure, assume that $\varepsilon_k = \varepsilon$ and $\alpha_k = \alpha$ for all $k$, and let $x_{ik}^1 \equiv w_i$, $x_{ik}^2 \equiv L_{i,k}$, and $x_{ik}^3 \equiv P_{i,k}$. Then the above equilibrium conditions can be written as

$$\left(x_{ik}^1\right)^{1+\varepsilon} \left(x_{ik}^2\right)^{1-\alpha} = \sum_{ns} K_{ik,ns}^1 x_{ns}^1 \left(x_{ns}^3\right)^{\varepsilon}, \tag{38}$$

$$\left(x_{ik}^3\right)^{-\varepsilon} = \sum_{ns} K_{ik,ns}^2 \left(x_{ns}^1\right)^{-\varepsilon} \left(x_{ns}^2\right)^{\alpha}, \tag{39}$$

$$x_{ik}^1 = \sum_{ns} K_{ik,ns}^3 x_{ns}^1 x_{ns}^2, \tag{40}$$

where $K_{ik,ns}^1$, $K_{ik,ns}^2$, and $K_{ik,ns}^3$ are appropriate (nonnegative) constants. This system maps into the system of equations (3) in AAL with each "location" being an $(i,s)$ pair. Following AAL's notation, we have

$$\Gamma = \begin{pmatrix} 1+\varepsilon & 1-\alpha & 0 \\ 0 & 0 & -\varepsilon \\ 1 & 0 & 0 \end{pmatrix}$$

and

$$B = \begin{pmatrix} 1 & 0 & \varepsilon \\ -\varepsilon & \alpha & 0 \\ 1 & 1 & 0 \end{pmatrix}.$$

Assuming that $\alpha \in [0,1)$, we have

$$\left|B\Gamma^{-1}\right| = \begin{pmatrix} 0 & 1 & 1 \\ \frac{\alpha}{1-\alpha} & 0 & \frac{\alpha+\varepsilon}{1-\alpha} \\ \frac{1}{1-\alpha} & 0 & \frac{\alpha+\varepsilon}{1-\alpha} \end{pmatrix}.$$

Each element of this matrix is a non-decreasing function of $\alpha$ and $\varepsilon$. Hence, the spectral radius of this matrix for arbitrary $\alpha \in [0,1)$ and $\varepsilon \geq 0$ is at least as large as the spectral radius of the same matrix with $\alpha = 0$ and $\varepsilon = 0$, i.e., of the matrix:

$$\begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

(see Corollary 8.1.19 in Horn and Johnson, 2013), which has a spectral radius of 1. In simulations we see that, in order to have the spectral radius of $\left|B\Gamma^{-1}\right|$ to be not larger than 1, we need to have a negative $\alpha$. So, we cannot invoke AAL's Theorem 2 to establish uniqueness.

It is interesting to explore how the AAL approach can be used to establish uniqueness for labor allocations given wages. That would correspond to the case in which we take $x^1_{ik}$ as given and ignore equation (40) in the system (38)-(40). Relabeling $x^2_{ik}$ as $y^1_{ik}$ and $x^3_{ik}$ as $y^2_{ik}$, the relevant system can be written as

$$\left(x^1_{ik}\right)^{1-\alpha} = \sum_{ns} \tilde{K}^1_{ik,ns}\left(x^2_{ns}\right)^{\varepsilon},$$
$$\left(x^2_{ik}\right)^{-\varepsilon} = \sum_{ns} \tilde{K}^2_{ik,ns}\left(x^1_{ns}\right)^{\alpha}.$$

This entails

$$\Gamma = \begin{pmatrix} 1-\alpha & 0 \\ 0 & -\varepsilon \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 0 & \varepsilon \\ \alpha & 0 \end{pmatrix}.$$

Then for $\alpha \in [0,1)$ we have

$$\left|B\Gamma^{-1}\right| = \begin{pmatrix} 0 & 1 \\ \dfrac{\alpha}{1-\alpha} & 0 \end{pmatrix}.$$

The spectral radius of this matrix is $\sqrt{\frac{\alpha}{1-\alpha}}$, which is lower than 1 only if $\alpha < 1/2$. This condition is more stringent than the one in Proposition 1.

## C.8. Computation of Equilibrium

### C.8.1. A Practical Algorithm for Finding an Equilibrium

The analysis of Section 4 suggests two alternative approaches to numerically compute an equilibrium. First, one can use an algorithm that properly deals with the complementary slackness conditions in the system of equations (4) and (5) for $(\boldsymbol{w}, \boldsymbol{L})$. This requires an algorithm for non-linear complementarity problems, such as the PATH solver (Ferris and Munson, 1999). Second, one can follow the approach used in Section 4 to prove existence and uniqueness of equilibrium and break the problem in two steps: first, for each wage vector $\boldsymbol{w}$ find $\boldsymbol{L}_k(\boldsymbol{w})$ for each $k$ by solving the optimization problem associated with (25), and second, find the wage vector such that the excess labor demand $\boldsymbol{Z}(\boldsymbol{w}) \equiv \sum_k \boldsymbol{L}_k(\boldsymbol{w}) - \bar{\boldsymbol{L}}$ is zero using the *tatonnement* iterative procedure proposed by Alvarez and Lucas (2007).

It turns out, however, that a third approach does best. Consider the function $\boldsymbol{w}(\boldsymbol{T})$ that one would get simply by solving for wages in the standard multi-sector model with no scale economies and technology parameters $\boldsymbol{T} = \{T_{i,k}\}$, and let $L^d_{i,k}(\boldsymbol{T}, \boldsymbol{w})$ be labor demand as a function of technology parameters and wages also in that model. Let $\boldsymbol{T}(\boldsymbol{L})$ be defined by $T_{i,k}(\boldsymbol{L}) = S_{i,k} L^{\alpha_k}_{i,k}$ and let $\boldsymbol{H}(\boldsymbol{L}) \equiv \boldsymbol{L}^d(\boldsymbol{T}(\boldsymbol{L}), \boldsymbol{w}(\boldsymbol{T}(\boldsymbol{L})))$. By definition of $\boldsymbol{w}(\boldsymbol{T})$ we must have $\sum_k L^d_{i,k}(\boldsymbol{T}(\boldsymbol{L}), \boldsymbol{w}(\boldsymbol{T}(\boldsymbol{L}))) = \bar{L}_i$ for all $i$ and $\boldsymbol{L}$. Thus, if $\boldsymbol{L}^*$ is an interior fixed point of the mapping $\boldsymbol{H}(\boldsymbol{L})$ then $(\boldsymbol{w}^*, \boldsymbol{L}^*) = (\boldsymbol{w}(\boldsymbol{T}(\boldsymbol{L}^*), \boldsymbol{L}^*)$ is an equilibrium of our economy with economies of scale. Since $\boldsymbol{H}(\boldsymbol{L})$ is a continuous mapping from the compact set $\Lambda \equiv \{\boldsymbol{L} | \sum_k L_{i,k} = \bar{L}_i\}$ to itself, then we can use the iterative procedure given by $\boldsymbol{L}_{t+1} = \boldsymbol{H}(\boldsymbol{L}_t)$ to compute the equilibrium points.

We have used this algorithm for counterfactual analysis with many countries and sectors (see Subsection 6.2 and Section 7; also see Appendix D.3 for a more detailed description this algorithm) and found that it can easily handle corners and that it is very robust. We have also used this algorithm on economies with three or four countries, two sectors, $\alpha = 0.9$ and randomly chosen values for all other parameters. Compared to a standard Newton method, it is slower but way more robust. We randomly generated more than a million economies with three countries and two sectors, and more than half a million economies with four countries and two sectors. In all cases the algorithm using the iterative procedure with $\boldsymbol{L}_{t+1} = \boldsymbol{H}(\boldsymbol{L}_t)$ found a solution, whereas the Newton method found a solution only for some initial conditions.

Because the Newton method is faster, we used it in combination with our iterative procedure in an effort to find examples with multiple equilibria. For each of the random economies mentioned above, we computed the equilibrium with the iterative procedure, and also with the Newton method with 400 different starting points. If there were multiple equilibria, we would likely have one of the solutions of the Newton method be different than the one found by the iterative procedure, but this never happened.[43] For the case with $\alpha = 0.9$ we also computed the sign of the determinant of the (negative of the) normalized excess labor demand evaluated at the equilibrium we found. By the Index Theorem, a negative value would imply multiplicity. We always found this sign to be positive.

### C.8.2. A Rigorous Algorithm for Finding All Equilbria

The fact that we do not find different equilibria when starting from different initial conditions does not prove that there is no multiplicity. To prove uniqueness of equilibrium for a particular parameterization of our economy, we need a procedure that guarantees to find all solutions to our system of complementary slackness conditions. One approach for doing so is to transform our equilibrium system into a system of polynomial equations, and then apply methods from algebraic geometry to find all solutions of that transformed system. This is feasible in our case as long as all trade and scale elasticities are rational numbers.[44] An alternative approach is to exploit interval analysis, which provides methods for direct computation of outer bounds for the range of values of a function evaluated in some interval.[45] This approach has the benefit that it can deal with the complementary slackness conditions that become relevant when some $\alpha \geq 1$. Moreover, interval-arithmetic based algorithms for finding all solutions of a system of non-linear equations are natural extensions of the bisection and Newton methods that are well-known in economics.

The key idea of interval arithmetic is that instead of working with functions defined over variables, we work with extensions of those functions that are defined over in-

[43]To check that this procedure delivers multiplicity when we know they exist, we used the same code for $\alpha = 2$. We find that this leads to multiple equilibria for randomly generated parameters with three economies and two sectors.

[44]For a review of the use of algebraic geometry methods to find all solutions to systems of polynomial equations, and the use of these techniques in economics, see Kubler, Renner and Schmedders (2014).

[45]See Moore, Kearfott and Cloud (2009) for an excellent introduction to interval analysis. For a thorough coverage of interval analysis methods see Hansen and Walster (2003).

tervals. Consider a function $F : A \to \mathbb{R}^N$, with $A \subseteq \mathbb{R}^N$. Given that $F$ satisfies certain conditions (which are most likely to be satisfied for functions typically encountered in economics), interval arithmetic can be used to construct a function $\mathbb{F}$ operating directly on intervals $\mathbb{X} \subseteq A$ with the following properties: (i) for any interval $\mathbb{X} \subseteq A$ and any vector $x \in \mathbb{X}$, if $y = F(x)$, then $y \in \mathbb{F}(\mathbb{X})$;[46] (ii) if $[x, x]$ is the degenerate interval with unique element $x$, then $\mathbb{F}([x, x]) = F(x)$; and (iii) the "smaller" is an interval $\mathbb{X}$, the "tighter" is the enclosure of the range of values of $F$ over $\mathbb{X}$ by interval $\mathbb{F}(\mathbb{X})$.

Such interval extensions of regular functions are useful for characterizing solutions of $F(x) = 0$. Indeed, property (i) guarantees that if $0 \notin \mathbb{F}(\mathbb{X})$ then there is no solution to $F(x) = 0$ in $\mathbb{X}$. Imagine then that $0 \in \mathbb{F}(\mathbb{X}_0)$ for $\mathbb{X}_0 \subseteq A$. We can partition $\mathbb{X}_0$ into $I$ subintervals $\mathbb{X}_i$ for $i = 1, \ldots, I$ and check whether $0 \in \mathbb{F}(\mathbb{X}_i)$. If $0 \in \mathbb{F}(\mathbb{X}_i)$ for some $i$, we can further partition the interval $\mathbb{X}_i$ and iterate, at each point discarding intervals $\mathbb{X}$ for which $0 \notin \mathbb{F}(\mathbb{X})$. Up to a level of precision supplied in the algorithm, we can then find all "small" intervals $\mathbb{X}$ such that $0 \in \mathbb{F}(\mathbb{X})$. Property (iii) from above guarantees that, if such intervals $\mathbb{X}$ are "small" enough, then they contain solutions to $F(x) = 0$. This procedure is a multi-dimensional bisection algorithm and is essentially a rigorous way of doing what economists sometimes call "grid search". And, just as Newton-based methods are preferable to bisection/"grid search" for finding solutions of systems of equations whenever the relevant functions are differentiable, here too it is possible and preferable to use interval analogs of Newton-based algorithms if $F$ is differentiable.

We have built a Newton-based algorithm using interval arithmetic to find all solutions of our system of complementary-slackness conditions (see Online Appendix) and used this algorithm to find all equilibria for our economy with three countries and two sectors for $\alpha = 0.9$ and randomly chosen values for all other parameters. The typical time it takes to run the algorithm for economies with three countries is about 1 hour. At the time of this writing, we have been able to run the algorithm for about 3 thousand parameterizations of our economy with three countries, and we always found a unique

[46]The converse is not required: we may have $y \in \mathbb{F}(\mathbb{X})$ with $y \neq F(x)$ for all $x \in \mathbb{X}$. To guarantee this converse property, we essentially need to solve global minimization and maximization problems, $\underline{F} \equiv \min_{x \in \mathbb{X}} F(x)$ and $\overline{F} = \max_{x \in \mathbb{X}} F(x)$, and assign $\mathbb{F}(\mathbb{X}) = \left[\underline{F}, \overline{F}\right]$. Solving such optimization problems for each interval $\mathbb{X}$ is, of course, prohibitively expensive. One of the purposes of interval arithmetic is construction of estimated bounds on the range of values of function $F$ over an interval $\mathbb{X}$ with acceptable tolerance and with as little effort as possible.

equilibrium.[47]

## D. Scale Economies, Welfare and Trade Flows

### D.1. Proof of Proposition 7

The argument just above Proposition 7 has already established that, setting $w_i = 1$ by choice of numeraire, if $L_{i,k} > 0$ then $P_{i,k} < P^A_{i,k}$ and so there are strictly positive gains from trade in industry $k$ for country $i$. Let's now consider the case with $L_{i,k} = 0$. In this case we know that in equilibrium we must have $G_{i,k} \geq 0$, and using again $w_i = 1$ this can be rewritten as

$$P_{i,k} \leq \mu_k \left( \frac{1 - \Delta_{i,k}}{S_{i,k}\beta_{i,k}\bar{L}_i} \right)^{1/\varepsilon_k},$$

where $\Delta_{i,k} \equiv \sum_{n \neq i} S_i \tau_{ni}^{-\varepsilon} \mu_k^{-\varepsilon_k} P_{n,k}^{\varepsilon_k} \beta_{n,k} w_n \bar{L}_n$. But in autarky we have $P^A_{i,k} = \mu_k \left(S_{i,k}\beta_{i,k}\bar{L}_i\right)^{-1/\varepsilon_k}$, and hence

$$P_{i,k}/P^A_{i,k} \leq \left(1 - \Delta_{i,k}\right)^{1/\varepsilon_k}.$$

With trade we have $0 < \Delta_{i,k} < 1$, and hence $P_{i,k} < P^A_{i,k}$. □

### D.2. Proof of Proposition 8

Letting $x_i \equiv w_i L_i$, $y_i \equiv P^{\varepsilon}$, $a_{ni} \equiv S_i w_i^{-\alpha-\varepsilon} \tau_{ni}^{-\varepsilon}$, $b_n \equiv \beta_n w_n \bar{L}_n$ and log-differentiating the system in (15) around an equilibrium point for some change in $a_{ni}$ we get

$$d\ln x_i = \frac{1}{1-\alpha} \sum_n \chi_{ni} \left(d\ln a_{ni} + d\ln y_n\right) \quad \text{for} \quad i = 1, ..., N,$$

$$d\ln y_i = -\sum_j \lambda_{ij} \left(d\ln a_{ij} + \alpha d\ln x_j\right) \quad \text{for} \quad n = 1, ..., N,$$

where $\lambda_{ni} \equiv a_{ni} x_i^{\alpha} y_n$ are import shares and $\chi_{ij} \equiv \dfrac{a_{ij} x_j^{\alpha} y_i b_i}{\sum_n a_{nj} x_j^{\alpha} y_n b_n}$ are export shares (i.e. $\chi_{ij}$ is the share of total country $j$ exports directed to country $i$).

[47]The algorithm can also be used to find all equilibria in the case when $\alpha = 1$ or $\alpha > 1$ for some industries. We have tried our algorithm for economies with two countries and two industries with $\alpha > 1$ in one industry and $\alpha < 1$ in the other industry. In this case the algorithm was able to find multiple equilibria.

Let $\mathscr{X}$ be the matrix of export shares with elements $\chi_{ni}$, $\Lambda$ be the matrix of import shares with elements $\lambda_{ni}$, let $X$ and $Y$ be column vectors with elements $d\ln x_i$ and $d\ln y_i$, let $A$ be the matrix with typical element $d\ln a_{ni}$, and let matrix $\mathbf{1}$ be a column vector whose entries are all ones. We can rewrite the system in matrix form as

$$X = \frac{1}{1-\alpha}\left(\left[\mathscr{X}^T \circ A^T\right]\mathbf{1} + \mathscr{X}^T Y\right),$$

$$Y = -\left[\Lambda \circ A\right]\mathbf{1} - \alpha\Lambda X,$$

where the symbol "$\circ$" denotes the Hadamard product. Substituting the first equation into the second and rearranging we get

$$\left(\gamma I + \Lambda\mathscr{X}^T\right)Y = -\left(\gamma\left[\Lambda \circ A\right]\mathbf{1} + \Lambda\left[\mathscr{X}^T \circ A^T\right]\mathbf{1}\right), \tag{41}$$

where $\gamma = \dfrac{1-\alpha}{\alpha}$.

Since $\sum_n \lambda_{ni} b_n = x_i$ implies $\chi_{ij} = \dfrac{\lambda_{ij} b_i}{x_j}$, we can write $\mathscr{X} = B\Lambda L^{-1}$ and by extension $\Lambda\mathscr{X}^T = \Lambda L^{-1}\Lambda^T B$, where $L$ is a diagonal matrix with elements $x_i$ on the diagonal and $B$ is a diagonal matrix with elements $b_i$ on the diagonal. Observe that matrices $\Lambda L^{-1}\Lambda^T B$ and $\left(B^{\frac{1}{2}}\Lambda\right)L^{-1}\left(B^{\frac{1}{2}}\Lambda\right)^T$ have the same eigenvalues, and that matrix $\left(B^{\frac{1}{2}}\Lambda\right)L^{-1}\left(B^{\frac{1}{2}}\Lambda\right)^T$ is positive semidefinite. It then follows that all eigenvalues of $\Lambda L^{-1}\Lambda^T B$ are real and non-negative, which, in turn, implies that eigenvalues of $\gamma I + \Lambda\mathscr{X}^T$ are real and positive for any $\gamma > 0$, and so $\det\left(\gamma I + \Lambda\mathscr{X}^T\right) > 0$ for $\gamma > 0$. Since we are interested only in the signs of entries of $Y$ in expression (41), we can then focus on

$$-\det\left(\gamma I + \Lambda\mathscr{X}^T\right)Y = adj\left(\gamma I + \Lambda\mathscr{X}^T\right)\left(\gamma\left[\Lambda \circ A\right]\mathbf{1} + \Lambda\left[\mathscr{X}^T \circ A^T\right]\mathbf{1}\right), \tag{42}$$

where $adj(\cdot)$ is the adjugate of a matrix.

Consider now the case $N = 2$ and without loss of generality consider a unilateral trade liberalization for country 1. We are then interested in the sign of $\partial\ln y_1/\partial\ln a_{12}$, and so for this case we have $d\ln a_{11} = d\ln a_{22} = d\ln a_{21} = 0$ and $d\ln a_{12} \neq 0$. Using the facts $adj\left(\gamma I + \Lambda\mathscr{X}^T\right) = \gamma I + adj\left(\Lambda\mathscr{X}^T\right)$ (this is true only in the case of $2\times 2$ matrices), $adj\left(\Lambda\mathscr{X}^T\right) = adj\left(\mathscr{X}^T\right)adj(\Lambda)$ and $adj(\Lambda)\Lambda = \det(\Lambda)$, and applying the result in (42) together with some manipulation we have

$$-\frac{\det(\gamma I+\Lambda\mathscr{X})}{\lambda_{12}d\ln a_{12}}Y=\gamma^2\begin{pmatrix}1\\0\end{pmatrix}+\gamma\begin{pmatrix}\lambda_{21}\chi_{21}+\lambda_{22}\chi_{22}+\chi_{12}\\-\lambda_{21}\chi_{11}-\lambda_{22}\chi_{12}+\frac{\lambda_{22}}{\lambda_{12}}\chi_{12}\end{pmatrix}+\frac{\chi_{12}}{\lambda_{12}}\det(\Lambda)\begin{pmatrix}-\chi_{21}\\\chi_{11}\end{pmatrix}.$$

Using the expression above together with some algebra one can then show that there exists $\bar{\gamma}^{1,\tau}>0$ such that for any $\gamma\in(0,\infty)$ we have that $\partial\ln y_1/\partial\ln a_{12}$ is negative if and only if $\gamma>\bar{\gamma}^{1,\tau}$, with $\bar{\gamma}^{1,\tau}$ given by

$$\bar{\gamma}^{1,\tau}=\frac{\sqrt{D^{\tau}}-(\lambda_{21}\chi_{21}+\lambda_{22}\chi_{22}+\chi_{12})}{2}>0,$$

where

$$D^{1,\tau}\equiv(\lambda_{21}\chi_{21}+\lambda_{22}\chi_{22}+\chi_{12})^2+4(\chi_{11}\chi_{22}-\chi_{12}\chi_{21})\lambda_{21}$$

is always positive.[48] Since $\gamma>\bar{\gamma}^{1,\tau}\Leftrightarrow\alpha<\bar{\alpha}^{1,\tau}=1/(1+\bar{\gamma}^{1,\tau})$ and since

$$-\partial\ln P_1/\partial\ln\tau_{12}=\partial\ln P_1^{\varepsilon}/\partial\ln\tau_{12}^{-\varepsilon}=\partial\ln y_1/\partial\ln a_{12},$$

the result in the text immediately follows.

Consider now a productivity increase in country 2. Here we are interested in the sign of $\partial\ln y_1/\partial\ln a_{22}$, and so for this case we have $d\ln a_{11}=d\ln a_{21}=0$ and $d\ln a_{22}\neq 0$. Note also that we have $d\ln a_{12}=d\ln\tau_{12}^{-\varepsilon}+d\ln a_{22}=d\ln a_{22}$ since $d\ln\tau_{12}^{-\varepsilon}=0$. Analogous to the trade liberalization exercise above one can readily show that $\partial\ln P_1/\partial\ln S_2=\partial\ln P_1/\partial\ln a_{22}<0$ if and only if $\bar{\gamma}_n^S\Leftrightarrow\alpha<\bar{\alpha}_n^S=1/(1+\bar{\gamma}^{1,S})$, with $\bar{\gamma}^{1,S}$ given by

$$\bar{\gamma}^{1,S}\equiv\frac{\sqrt{D^{1,S}}-\left(\lambda_{21}\chi_{21}+\chi_{12}-\lambda_{11}\lambda_{22}\lambda_{12}^{-1}\chi_{21}\right)}{2}>0$$

where

$$\begin{aligned}D^{1,S}\equiv&\left(\lambda_{21}\chi_{21}+\chi_{12}-\lambda_{11}\lambda_{22}\lambda_{12}^{-1}\chi_{21}\right)^2+4(\chi_{11}\chi_{22}-\chi_{12}\chi_{21})\lambda_{21}\\&+4\left(\lambda_{11}\lambda_{22}\lambda_{12}^{-1}\chi_{21}+\lambda_{22}\chi_{22}\right)\chi_{22}.\end{aligned}$$

The result in the text then immediately follows.

[48]In particular, it is straight forward to verify that $\tau_{12}\tau_{21}\geq 1$ implies $\chi_{11}\chi_{22}-\chi_{12}\chi_{21}\geq 0$.

### D.3. System in Changes and Algorithm for Section 6.2

In this appendix we derive the system in changes and describe the algorithm used to perform counterfactual exercises in Section 6.2.

#### D.3.1. Derivation of System in Changes

In the presence of tariffs, total (*ad-valorem*) trade costs are given by $\tau_{ni,k}(1+t_{ni,k})$, where $t_{ni,k}$ is a tariff that importer $n$ imposes on goods from exporter $i$'s industry $k$, and $\tau_{ni,k}$ captures all other (iceberg) costs of trade. Since data on trade flows features trade imbalances, we assume that country $n$'s value of net imports is given by $D_n$, which can be different from zero and satisfy $\sum_n D_n = 0$. The exact hat algebra approach by Dekle, Eaton and Kortum (2008) works as long as we start from an equilibrium that does not have corners, which is the case in our data as there are no $(i,k)$ pairs with $L_{i,k} = 0$.

Let us first derive the equilibrium system of equations for the version of our common framework that features tariffs and trade imbalances. Denote by $E_n$ the total expenditure in country $n$. Then country $n$'s expenditure on goods from industry $(i,k)$ is given by $X_{ni,k} = \lambda_{ni,k}\beta_{n,k}E_n$, with trade shares given by

$$\lambda_{ni,k} = \frac{S_{i,k}L_{i,k}^{\alpha_k}\left(w_i\tau_{ni,k}\left(1+t_{ni,k}\right)\right)^{-\varepsilon_k}}{\sum_l S_{l,k}L_{l,k}^{\alpha_k}\left(w_l\tau_{nl,k}\left(1+t_{nl,k}\right)\right)^{-\varepsilon_k}}.$$

Budget balance requires that $E_n = w_n\bar{L}_n + D_n + R_n$, where $D_n$ are trade imbalances in the data with $\sum_n D_n = 0$, and

$$R_n \equiv \sum_k\sum_i \frac{t_{ni,k}}{1+t_{ni,k}}X_{ni,k} = \sum_k\sum_i \frac{t_{ni,k}\lambda_{ni,k}}{1+t_{ni,k}}\beta_{n,k}E_n$$

denotes total tariff revenues in country $n$. Substituting $E_n$ into the definition of $R_n$ yields:

$$R_n = \frac{\pi_n}{1-\pi_n}\left(w_n\bar{L}_n + D_n\right),$$

where

$$\pi_n \equiv \sum_k\sum_i \frac{t_{ni,k}\lambda_{ni,k}}{1+t_{ni,k}}\beta_{n,k}.$$

Demand for goods from industry $(i,k)$ is given by

$$\begin{aligned}\sum_n \frac{1}{1+t_{ni,k}} X_{ni,k} &= \sum_n \frac{\lambda_{ni,k}}{1+t_{ni,k}} \beta_{n,k} E_n \\ &= \sum_n \frac{\lambda_{ni,k}}{1+t_{ni,k}} \left(\frac{\beta_{n,k}}{1-\pi_n}\right)\left(w_n \bar{L}_n + D_n\right),\end{aligned}$$

and so the goods market clearing condition is

$$L_{i,k} \geq 0, \quad G_{i,k}\left(\boldsymbol{w}, \boldsymbol{L}_k\right) \geq 0, \quad L_{i,k} G_{i,k}\left(\boldsymbol{w}, \boldsymbol{L}_k\right) = 0,$$

with

$$\begin{aligned}G_{i,k}\left(\boldsymbol{w}, \boldsymbol{L}_k\right) &\equiv w_i - \frac{1}{L_{i,k}} \sum_n \frac{\lambda_{ni,k}}{1+t_{ni,k}} \beta_{n,k} E_n \\ &= w_i - \frac{1}{L_{i,k}} \sum_n \frac{\lambda_{ni,k}}{1+t_{ni,k}} \left(\frac{\beta_{n,k}}{1-\pi_n}\right)\left(w_n \bar{L}_n + D_n\right).\end{aligned}$$

Finally, the labor market clearing condition is the same as in the case without tariffs and trade imbalances:

$$\sum_k L_{i,k} = \bar{L}_i.$$

Now we can formulate the system in changes. For any observed variable $x$, denote its value in a counterfactual equilibrium by $x'$ and the relative change in $x$ by $\hat{x} \equiv x'/x$. Assuming that the observed equilibrium does not have corner labor allocations, we can use the hat notation to write the system of equations for a counterfactual equilibrium with new values for trade costs, tariffs, and productivities:

$$\begin{aligned}&\hat{L}_{i,k} L_{i,k} \geq 0, \quad \tilde{G}_{i,k}\left(\hat{\boldsymbol{w}}, \hat{\boldsymbol{L}}_k\right) \geq 0, \quad \hat{L}_{i,k} Y_{i,k} \tilde{G}_{i,k}\left(\hat{\boldsymbol{w}}, \hat{\boldsymbol{L}}_k\right) = 0, \\ &\sum_k \hat{L}_{i,k} Y_{i,k} = Y_i,\end{aligned}$$

with

$$\tilde{G}_{i,k}\left(\hat{\boldsymbol{w}}, \hat{\boldsymbol{L}}_k\right) \equiv \hat{w}_i - \frac{1}{\hat{L}_{i,k} Y_{i,k}} \sum_n \frac{\lambda'_{ni,k}\left(\hat{\boldsymbol{w}}, \hat{\boldsymbol{L}}_k\right)}{1+t'_{ni,k}} \cdot \frac{\beta_{n,k}\left(\hat{w}_n Y_n + D_n\right)}{1-\pi'_n\left(\hat{\boldsymbol{w}}, \hat{\boldsymbol{L}}\right)},$$

and

$$\lambda'_{ni,k}\left(\hat{\boldsymbol{w}},\hat{\boldsymbol{L}}_k\right)=\frac{\hat{L}_{i,k}^{\alpha_k}\left(\hat{w}_i\left(1+t'_{ni,k}\right)\hat{\tau}_{ni,k}\right)^{-\varepsilon_k}\left(1+t_{ni,k}\right)^{\varepsilon_k}\lambda_{ni,k}}{\sum_l \hat{L}_{l,k}^{\alpha_k}\left(\hat{w}_l\left(1+t'_{nl,k}\right)\hat{\tau}_{nl,k}\right)^{-\varepsilon_k}\left(1+t_{nl,k}\right)^{\varepsilon_k}\lambda_{nl,k}},$$

$$\pi'_n\left(\hat{\boldsymbol{w}},\hat{\boldsymbol{L}}\right)=\sum_k\sum_i\frac{t'_{ni,k}\lambda'_{ni,k}\left(\hat{\boldsymbol{w}},\hat{\boldsymbol{L}}_k\right)}{1+t'_{ni,k}}\beta_{n,k},$$

where $Y_{i,k}\equiv w_i L_{i,k}$ and $Y_n\equiv w_n\bar{L}_n$.

The above system in changes still allows the counterfactual equilibrium to exhibit corner allocations. Therefore, we need to calculate changes in welfare explicitly as $\hat{I}_n/\hat{P}_n$, where $\hat{I}_n$ is the change in income given by

$$\hat{I}_n=\frac{w'_n\bar{L}_n+R'_n}{w'_n\bar{L}_n+R_n}=\frac{w'_n\bar{L}_n+\pi'_nD_n}{w'_n\bar{L}_n+\pi_nD_n}\cdot\frac{1-\pi_n}{1-\pi'_n},$$

and

$$\hat{P}_n=\prod_k\left(\sum_l \hat{L}_{l,k}^{\alpha_k}\left(\hat{w}_l(1+t'_{nl,k})\hat{\tau}_{nl,k}\right)^{-\varepsilon_k}(1+t_{nl,k})^{\varepsilon_k}\lambda_{nl,k}\right)^{-\beta_{n,k}/\varepsilon_k}.$$

### D.3.2. Algorithm for Counterfactuals in Section 6.2

The algorithm consists of two logical parts: an inner loop and an outer loop. The inner loop keeps $\hat{\boldsymbol{L}}$ fixed and finds wages $\hat{\boldsymbol{w}}$ that clear labor markets. The outer loop finds labor allocations $\hat{\boldsymbol{L}}$ that clear goods markets.

The inner loop exploits the tatonnement process proposed by Alvarez and Lucas (2007). For any variable $x$ calculated in the inner loop, let us denote the value of $x$ on the $t$-th iteration of the inner loop by $x^{(t)}$ with $x^{(0)}$ denoting the value in the baseline equilibrium (corresponding by assumption to the data). Let us also use the hat notation for the change in $x^{(t)}$: $\hat{x}^{(t)}\equiv x^{(t)}/x^{(0)}$. The $(t+1)$-th inner loop iteration for wages can be written as

$$\hat{w}_i^{(t+1)}=\hat{w}_i^{(t)}+\nu\frac{\sum_k X_{i,k}^{(t)}-\hat{w}_i^{(t)}Y_i^{(0)}}{Y_i^{(0)}},$$

where $v$ is some small positive number that is a parameter of the algorithm,

$$X_{i,k}^{(t)} = \sum_n \frac{\lambda_{ni,k}^{(t)}}{1+t'_{ni,k}} \cdot \frac{\beta_{n,k}\left(\hat{w}_n^{(t)} Y_n^{(0)} + D_n^{(0)}\right)}{1 - \sum_k \sum_l \frac{t'_{nl,k}\lambda_{nl,k}^{(t)}}{1+t'_{nl,k}} \beta_{n,k}},$$

and

$$\lambda_{ni,k}^{(t)} = \frac{\hat{L}_{i,k}^{\alpha_k}\left(\hat{w}_i^{(t)}\right)^{-\varepsilon_k} \hat{\tau}_{ni,k}^{-\varepsilon_k}\left(\frac{1+t'_{ni,k}}{1+t_{ni,k}}\right)^{-\varepsilon_k} \lambda_{ni,k}^{(0)}}{\sum_l \hat{L}_{l,k}^{\alpha_k}\left(\hat{w}_l^{(t)}\right)^{-\varepsilon_k} \hat{\tau}_{nl,k}^{-\varepsilon_k}\left(\frac{1+t'_{nl,k}}{1+t_{nl,k}}\right)^{-\varepsilon_k} \lambda_{nl,k}^{(0)}}.$$

The inner loop iterates until there is no significant change between $\hat{\boldsymbol{w}}^{(t)}$ and $\hat{\boldsymbol{w}}^{(t+1)}$.

For a given $\hat{\boldsymbol{L}}$, the inner loop gives the set of wages $\hat{w}_i(\hat{\boldsymbol{L}})$ that clear labor markets. The outer loop iterates on $\hat{\boldsymbol{L}}$ using labor demand (in value) for sector $(i,k)$. Denoting by $\hat{L}_{i,k}^{(l)}$ labor allocations on the $l$-th iteration of the outer loop, the $(l+1)$-th iteration of the outer loop can be written as:

$$\hat{L}_{i,k}^{(l+1)} = \frac{1}{\hat{w}_i(\hat{\boldsymbol{L}}^{(l)})Y_{ik}^{(0)}} \sum_n \lambda'_{ni,k}\left(\hat{\boldsymbol{w}}(\hat{\boldsymbol{L}}^{(l)}), \hat{\boldsymbol{L}}^{(l)}\right)\beta_{n,k}\hat{w}_n(\hat{\boldsymbol{L}}^{(l)})Y_n^{(0)}.$$

The outer loop iterates until there is no significant change between $\hat{\boldsymbol{L}}^{(l)}$ and $\hat{\boldsymbol{L}}^{(l+1)}$.

## D.4. Derivation of Algebra for Section 7

In this appendix we derive the system in changes for the counterfactuals in Section 7. Combining (16) and (17), and using $Y_i \equiv w_i \bar{L}_i$ together with shock $\hat{T}_{i,k} = (e_{i,k}/r_{i,k})^{\alpha_k}$, we get a system in wage changes given by

$$\hat{w}_i Y_i = \sum_{k=1}^K \sum_{n=1}^N \frac{(e_{i,k}/r_{i,k})^{\alpha_k}(\hat{w}_i)^{-\varepsilon_k}\lambda_{ni,k}}{\sum_{l=1}^N (e_{l,k}/r_{l,k})^{\alpha_k}(\hat{w}_l)^{-\varepsilon_k}\lambda_{nl,k}} e_{n,k}(\hat{w}_n Y_n + D_n).$$

The solution for $\hat{w}_i$ can then be used to get the implied hat change in the labor allocation from

$$\hat{L}_{i,k} = \frac{1}{\hat{w}_i Y_{i,k}} \sum_{n=1}^N \frac{(e_{i,k}/r_{i,k})^{\alpha_k}(\hat{w}_i)^{-\varepsilon_k}\lambda_{ni,k}}{\sum_{l=1}^N (e_{l,k}/r_{l,k})^{\alpha_k}(\hat{w}_l)^{-\varepsilon_k}\lambda_{nl,k}} e_{n,k}(\hat{w}_n Y_n + D_n),$$

where $Y_{i,k} \equiv w_i L_{i,k}$. Finally, we can then get the implied change in trade flows from

$$\hat{X}_{ni,k} = \frac{1}{X_{ni,k}} \cdot \frac{(e_{i,k}/r_{i,k})^{\alpha_k} (\hat{w}_i)^{-\varepsilon_k} \lambda_{ni,k}}{\sum_{l=1}^{N} (e_{l,k}/r_{l,k})^{\alpha_k} (\hat{w}_l)^{-\varepsilon_k} \lambda_{nl,k}} e_{n,k} (\hat{w}_n Y_n + D_n).$$

## D.5. Data on Tariffs and Estimates of Elasticities of Scale

### D.5.1. Data on Tariffs

We use data on *ad-valorem* tariffs for the year 2008 from the United Nations Statistical Division, Trade Analysis and Information System (UNCTAD-TRAINS). For some countries 2008 tariff data are missing. In these cases we use tariff data for the nearest available year. Following Caliendo and Parro (2015), we focus on effectively applied tariffs and use simple tariff line average as the measure of tariff for each importer-exporter-sector combination.

The original UNCTAD-TRAINS data for 2008 do not cover individual European Union (EU) member countries, but they include the EU as a reporter. We infer tariffs for the EU member countries by using the fact that the EU members enjoy free circulation of goods within the EU and exert a common external tariff on all goods entering the EU market.

The UNCTAD-TRAINS data cover 31 two-digit (in ISIC rev.3) sectors, 181 reporters, and 245 partners. We aggregate these data into 16 sectors and 34 countries that we use in the current paper: we first aggregate tariffs across reporters by using reporters' total imports as weights; then we aggregate tariffs across partners by using partners' total exports as weights; and, finally, we aggregate tariffs across sectors by using sectors' total imports as weights. When calculating the weights, we use imports and exports only in the sectors for which UNCTAD-TRAINS has data on tariffs.

Tables 5 and 6 present average tariffs for importers and sectors. Averaging is done only for the sectors that have data on tariffs in UNCAD-TRAINS. In the exercises with tariffs in Section 6.2, we assume zero tariff for all sectors that have empty values in the "Average Tariff, %" column of Table 6.

| Country | Average Tariff, % | Country | Average Tariff, % | Country | Average Tariff, % | Country | Average Tariff, % |
|---|---|---|---|---|---|---|---|
| AUS | 2.71 | ESP | 3.12 | ITA | 3.16 | SVK | 3.05 |
| AUT | 3.06 | FIN | 3.06 | JPN | 6.68 | SVN | 3.05 |
| BEL | 3.07 | FRA | 3.15 | KOR | 12.64 | SWE | 3.06 |
| BRA | 9.66 | GBR | 3.12 | MEX | 5.35 | TUR | 6.18 |
| CAN | 3.42 | GRC | 3.06 | NLD | 3.08 | TWN | 5.97 |
| CHN | 8.92 | HUN | 3.06 | POL | 3.08 | USA | 2.81 |
| CZE | 3.06 | IDN | 7.16 | PRT | 3.06 | ROW | 6.50 |
| DEU | 3.21 | IND | 13.83 | ROM | 3.06 | | |
| DNK | 3.06 | IRL | 3.06 | RUS | 8.20 | | |

*Notes:* Original data on tariffs are from UNCTAD-TRAINS for year 2008. Average tariffs are caluclated across exporters and sectors with exporters weighted by their total exports and sectors weighted by total sector imports.

Table 5: Average Tariffs by Importers

### D.5.2. Estimates of Elasticities of Scale

We use sector-level estimates of elasticities of scale obtained by Bartelme *et al.* (2017, BCDR). These values are provided in Table 6. Several comments are in order.

First, in their paper, BCDR work with a list of sectors that is more disaggregated than the one in the current paper. In particular, three sectors from our list: "Basic metals and fabricated metal", "Electrical and optical equipment", and "Transport equipment" — are broken into subsectors in BCDR, with each subsector having its own estimate of scale elasticity. To assign values of scale elasticities to sectors from our list, we compute simple averages over corresponding subsectors in BCDR.

Second, in the current paper we use BCDR estimates only for the manufacturing sectors. BCDR have an estimate of scale elasticity for "Coke, refined petroleum and nuclear fuel", which we classify as non-manufacturing, and, thus, do not use BCDR estimate for this sector.

Third, BCDR do not have an estimate for one sector that we classify as manufacturing: "Manufacturing, N.E.C.; recycling". In our counterfactual exercises we assign the same scale elasticity to this sector as for the non-manufacturing sectors.

| | Sector | $\psi_k$ | Average Tariff, % |
|---|---|---|---|
| | Agriculture, hunting, forestry and fishing | | 12.05 |
| | Mining and quarrying | | 0.94 |
| (M) | Food, beverages and tobacco | 0.02 | 16.47 |
| (M) | Textiles and textile products. Leather, leather products and footwear | 0.16 | 10.27 |
| (M) | Wood and products of wood and cork | 0.13 | 4.15 |
| (M) | Pulp, paper, printing and publishing | 0.17 | 3.10 |
| | Coke, refined petroleum and nuclear fuel | | 2.87 |
| (M) | Chemicals and chemical products | 0.14 | 3.76 |
| (M) | Rubber and plastics | 0.20 | 5.59 |
| (M) | Other non-metallic mineral | 0.20 | 6.02 |
| (M) | Basic metals and fabricated metal | 0.14 | 3.55 |
| (M) | Machinery, not elsewhere classified | 0.15 | 3.38 |
| (M) | Electrical and optical equipment | 0.15 | 5.82 |
| (M) | Transport equipment | 0.18 | 6.31 |
| (M) | Manufacturing, not elsewhere classified; recycling | | 5.82 |
| | Electricity, gas and water supply | | 0.60 |
| | Construction | | |
| | Sale and repair of motor vehicles and motorcycles; retail sale of fuel. Wholesale trade, except of motor vehicles and motorcycles | | |
| | Retail trade and repair, except of motor vehicles and motorcycles | | |
| | Hotels and restaurants | | |
| | Inland transport | | |
| | Water transport | | |
| | Air transport | | |
| | Other supporting transport activities | | |
| | Post and telecommunications | | |
| | Financial intermediation | | |
| | Real estate activities | | |
| | Renting of machinery and equipment and other business activities | | |
| | Education | | |
| | Health and social work | | |
| | Other community, social and personal services. Private households with employed persons. Public administration and defence; compulsory social security | | |

*Notes:* Manufacturing sectors are marked by "(M)". Values for scale elasticities, $\psi_k$, for manufacturing sectors are based on estimates from Bartelme *et al.* (2017) (see description in the text of how these estimates are used to get the values for $\psi_k$). Original data on tariffs are from UNCTAD-TRAINS for year 2008. Average tariffs are caluclated across importers and exporters with importers weighted by their total imports and exporters weighted by their total exports. Empty entries in the column for $\psi_k$ mean that either Bartelme *et al.* (2017) do not have estimates for the corresponding sectors or their estimates are not used in the current paper. Empty values in the column for average tariffs mean that UNCTAD-TRAINS does not have data for the corresponding sectors.

Table 6: List of Sectors with Scale Elasticities for Exercises in Sections 5.2 and 6.2. Average Tariffs by Sectors