Trade, Labor Reallocation Across Firms and Wage Inequality

Mariano Somale*
Federal Reserve Board

March 2024

Abstract

This paper develops a framework for studying the distributional effects of trade in which strong skill-productivity complementarities in production imply that inequality rises as workers reallocate toward more-productive (skill-intensive) firms in the same industry. The model features a large number of skill groups and can accommodate empirically relevant restrictions on firm selection into exporting. An autarkic economy that opens to trade always experiences a pervasive rise in wage inequality under no firm entry, with wage polarization being another possibility under free entry. Theoretically, more outcomes are possible following a trade liberalization in a trading economy. In a calibrated version of the framework, any increase in trade openness always leads to pervasively higher wage inequality. The analysis highlights the importance of properly accounting for the role of new exporters (extensive margin) in shaping the aggregate relative demand for skills.

Keywords: Trade, firms, workers, supermodularity, wage inequality.
JEL codes: F10, F12, F16.

---

*Federal Reserve Board (email: mariano.a.somale@frb.gov, website: www.marianosomale.com). All the views expressed in this paper are mine and do not necessarily represent those of the Federal Reserve System. I am grateful to Simon Fuchs, Maximiliano Dvorkin, Sebastian Heise, David Kohn, Jing Zhang, Logan Lewis, Lorenzo Caliendo, and Julian Neira, as well as to many conference participants for helpful comments and suggestions. This paper builds on the second chapter of my dissertation, Somale (2015), which can found in my personal website.

1 Introduction

Wage inequality has risen significantly in many countries since the late 70s, a period that also saw a rapid expansion of international trade. Three broad lessons follow from the empirical research exploring the connection between both trends. First, as discussed in Goldberg and Pavcnik (2007) and Helpman (2016), the rise in inequality is largely accounted for by within-industry effects, with the evidence providing little support for the between-industry channels emphasized by the traditional factor-proportions trade theory.[1] Second, firms may be an important part of the story behind the changes in the wage distribution. For example, Krishna, Poole, and Senses (2014) find substantial within-industry labor reallocation across firms following a trade liberalization that cannot be explained by a random assignment of workers to firms.[2] Third, divergent trends in inequality in different parts of the wage distribution (Autor, Katz, and Kearney 2008) and a rise in within-group (residual) wage inequality (Acemoglu 2002; Attanasio, Goldberg, and Pavcnik 2004) indicate that grouping workers into a few large skill-groups (as typically done in the literature) does not provide enough detail to understand the full distributional consequences of international trade.

In light of these lessons, this paper develops a general equilibrium trade model with a large number of skill groups that emphasizes within-industry labor reallocation across heterogeneous firms as the mechanism through which trade affects the wage distribution. In particular, strong skill-productivity complementarities in production imply that an increase in trade openness raises wage inequality when it induces a reallocation of workers toward more-productive (skill-intensive) firms in the same industry. I use the model to study the channels through which a trade-induced labor reallocation affects the wage distribution, including the entry and exit of firms into and out of the market, the increased demand of incumbent exporters, and the demand of new exporters.

The framework builds on standard heterogenous-firm trade models. As in Melitz (2003), labor is the only factor of production, the labor market is perfectly competitive, and final goods are produced by monopolistically competitive firms that differ in their

[1]This evidence includes a rise in the skill-premium in developed and developing countries (Goldberg and Pavcnik 2007), and little inter-industry labor reallocation following trade liberalizations.

[2]In addition, as discussed in Card et al. (2016), numerous studies find similar trends in the aggregate dispersion of wages and firms' productivity.

productivity. In addition, the presence of fixed production and export costs leads to selection into activity and into exporting—i.e., only some firms find it optimal to produce, and only a subset of them export. Departing from Melitz (2003), the labor force comprises heterogeneous workers of a continuum of skill types, so firms must choose not only the total number of *production* workers to hire but also the mix of skill-types to employ. Strong production complementarities between worker skill and firm productivity imply that more-productive firms have workforces of higher average ability in equilibrium.

The core of the framework lies in the production and export technology of firms. The output of a firm depends linearly on the number of *production* workers of each skill type that it employs. The productivity of a production worker at a given firm is a strictly log supermodular function of the worker's skill and the firm's productivity, giving more able workers a comparative advantage in production at more-productive firms. As in Costinot and Vogel (2010), these assumptions permit the analysis of market equilibrium to transform into the analysis of a matching problem. In particular, the equilibrium allocation of *production* workers among *active* firms is characterized by a strictly increasing and continuous matching function that maps the skill types of the former to the productivity types of the latter. Moreover, this matching function is a sufficient statistic for the dispersion wages in this setting, facilitating the analysis of comparative static predictions about wage inequality.

Fixed export costs also play an important role, as they determine firm selection into exporting, shaping the set exporters and their collective demand for skills. Therefore, I consider a flexible specification of fixed export costs that can accommodate weaker and more empirically relevant restrictions on firm selection into exporting than standard heterogeneous-firms trade models.[3] Specifically, I posit that fixed export costs vary across firms, and model their firm-specific sizes as independent realizations of a nonnegative random variable with an absolutely continuous and increasing cumulative distribution fuction (CDF). As a result, exporters are, on average, more productive than nonexporters in equilibrium, but high-productivity nonexporters coexists with low-productivity exporters. Finally, all fixed costs are paid in terms of a "skill bundle" that comprises *nonproduction* workers of all skill levels, an assumption that allows me to isolate the impact on the wage

[3] Assuming common fixed export costs across firms has been standard since Melitz (2003). This unrealistic assumption is not inocuous in this setting as it affects the distributional effects of trade.

distribution of the endogenous assignment of production workers to firms.

The cross section of the model captures several features of the data identified by the trade and labor literatures. The dispersion of wages in the model reflects between-firms wage differences (rather than within-firm differences), a channel that represents around 60% of the wage dispersion in the United States (Davis and Haltiwanger 1991). In addition, more-productive firms tend to be larger (in terms of output), have workforces of higher average ability, and pay higher average wages (Card et al. 2016). Per the stochastic representation of fixed export costs, the model features an imperfect positive correlation between size, firm wages and export status (Bernard and Jensen 1995) as well as between the latter and firm productivity, leading to overlapping productivity distributions for exporters and nonexports (Bernard, Eaton, Jensen, and Kortum 2003). Finally, if workers are classified in large skill groups, possibly reflecting imperfect observability of worker ability, then the model features wage heterogeneity within each of these skill groups (Acemoglu 2002; Attanasio et al. 2004).

I carry out the analysis of the effects of trade on the wage distribution under two widely used assumptions about firm entry into the industry: no free entry a-lá Chaney (2008) and free entry a-lá Melitz (2003). These alternative entry assumptions lead to the no-free-entry and free-entry models analyzed in the paper, whose predictions can be interpreted, respectively, as the short- and long-term effects of trade.[4] These models differ only in the equilibrium condition that pins down the *activity cutoff*, the productivity value below which firms do not find it profitable to produce. Conditional on the activity cutoff, the two models are identical, so they share the cross-sectional features discussed above.

To study the impact of higher trade openness on the wage distribution, I decompose the associated labor reallocation across firms into three channels. The first channel, the *selection-into-activity* channel, captures the reallocation of resources driven by changes in the set of active firms—i.e., by changes in the activity cutoff. The second channel, the *intensive margin* of trade, reflects the changes in the production and employment decisions of incumbent exporters that continue serving the foreign market after the decline in trade frictions. Finally, the third channel, the *extensive margin* of trade, captures the

[4]Exploring the implications of these two alternative entry assumptions also serves a pedagogical purpose. By delivering sharper results, the no-free-entry model facilitates the analysis of the main forces at play, which in turn simplifies the discussion of the more nuanced implications of the free-entry model.

reallocation of employment associated with changes in the set of exporters. These last two channels are largely determined by firm selection into exporting, highlighting the importance of not arbitrarily restricting this margin of adjustment in the model. This decomposition not only highlights the key elements driving the results in the current setting, but also facilitates the comparison with the implications of other frameworks in the literature exploring the connection between international trade, firms, and wages.

I analyze two instances of increased trade openness, opening to international trade and a trade liberalization, where the latter is defined as a decline in the variable trade costs faced by an economy that already participates in international trade. In the *no-free-entry model*, an initially autarkic economy that opens to trade always experiences an increase in the activity cutoff and a pervasive rise in wage inequality, in the sense that for any pair of workers, the relative wage of the more-skilled one rises. In terms of the three channels discussed above, the selection-into-activity channel induces a pervasive rise in wage inequality, as the exit of the least productive (low-skill-intensive) firms leads to a decline in the relative demand of less-skilled workers. With no exporters in the initial autarkic equilibrium, the intensive margin channel is not operational in this counterfactual. Finally, the extensive margin channel also leads to a pervasive rise in wage inequality; the (new) exporters in the open economy are, on average, more productive than nonexporters, so their collective labor demand is biased toward more-skilled workers. The importance of this channel, which depends on how fast the fraction of exporting firms increases with productivity, is determined by the CDF of fixed export costs.

A trade liberalization can lead to additional outcomes. Although a decline in variable trade costs in the no-free-entry model always leads to an increase in the activity cutoff and a rise in wage inequality at the lower end of the wage distribution, little can be said about its impact elsewhere in the distribution. As the activity cutoff rises, the selection-into-activity channel leads to a pervasive rise in wage inequality. The intensive-margin channel also leads to a pervasive rise in wage inequality, reflecting a rise in the more-skill-intensive labor demand of incumbent exporters as they expand their production to satisfy a higher foreign demand. In contrast, the impact of the extensive-margin channel on the wage distribution is theoretically ambiguous. Without additional restrictions on the CDF of fixed export costs, new exporters can be (on average) more or less productive than incumbent firms, so their collective demand may be biased toward more- or less-skilled workers. Moreover, the ambiguity about the effects of this third channel extends to the overall impact of a trade liberalization on the wage distribution. This result highlights the importance of paying close attention to the modeling of the extensive-margin channel in any study emphasizing the role of heterogenous firms in the distributional consequences of higher trade openness. I present sufficient conditions on the CDF of export costs under which wage inequality rises pervasively after a trade liberalization.

Assuming *free entry* brings an additional source of ambiguity relative to the previous results, as the effects of increased trade openness on the activity cutoff cannot be determined without imposing additional restrictions on primitives. If the *activity cutoff rises* after the economy opens to trade or after a liberalization, then the distributional effects predicted by the free-entry model are qualitatively the same as those described earlier for the no-free-entry model. If the *activity cutoff declines* after the economy opens to trade, then wages polarize—i.e., wage inequality decreases among the least-skilled workers but increases among the most-skilled ones. In this case, the selection-into-activity and extensive-margin channels lead to a pervasive decline and a pervasive rise in wage inequality, respectively, with the former channel dominating at the lower end of the wage distribution and the latter at the upper end. Finally, if a trade liberalization leads to a decline in the activity cutoff, then wage inequality necessarily decreases at the lower end of the distribution and increases somewhere else, but additional outcomes beyond wage polarization are possible. Of note, regardless of entry assumptions, an increase in trade openness never leads to a pervasive decline in wage inequality in this framework.

I also explore the effects of higher trade openness on the level of real wages. For both entry assumptions, an increase in trade openness (opening to trade or liberalization) always raises average real wages, but the least-skilled workers in the economy could see their real wage decline. In the free-entry-model, the fate of the real wages of these workers is completely determined by the response of activity cuttoff, leading to interesting connections between the effects of higher trade openness on the level and distribution of wages. For example, opening to international trade raises the real wage of the poorest workers in the economy only if it also induces a pervasive rise in wage inequality.

To assess the empirical relevance of the theoretical possibilities described above, I calibrate the model based on estimates from the literature and some broad features of firm data from Portugal. Given its informational content about the extensive-margin channel in the model, which drives much of the ambiguity in the theoretical results, a crucial target of the calibration is the fraction of firms that export in each decile of the empirical distribution of firms by value added per worker. For both entry assumptions, the calibrated model predicts pervasively higher wage inequality and higher real wages for all workers following any increase in trade openness. In the case of a trade liberalization, wage inequality always increases through the selection-into-activity and intensive-margin channels, while it decreases slightly through the extensive-margin channel. These results suggest that a decline in trade costs is likely to lead to pervasively higher wage inequality, in both the short and long run, through the labor-reallocation mechanisms emphasized in this paper. The analysis also highlights the importance of accurately quantifying the extensive-margin channel in the model. Indeed, assuming common fixed export costs across firms, as has been standard since Melitz (2003), results in much larger distributional effects through this channel, leading in some cases to declines in inequality in some parts of the wage distribution following a liberalization.

This paper is related to a growing number of studies using assignment models to study the distributional consequences of international trade and offshoring. Studies based on two-region competitive models, such as Grossman and Maggi (2000), Ohnsorge and Trefler (2007), Antràs, Garicano, and Rossi-Hansberg (2006), and Costinot and Vogel (2010), emphasize differences in higher moments of the skill distribution across regions as the drivers of trade and its distributional effects, with these effects generally differing qualitatively across regions as a result.[5] In contrast, different countries can experience similar distributional effects from trade through the mechanisms emphasized in this paper, as they do not rely on differences across countries. As such, the framework in this paper is better suited to think about the expansion international trade as a common factor contributing to the rise in wage inequality observed in many economies since the late 70s.

Methodologically, this paper is closer to a branch of this literature that, building on Costinot (2009), develops two-sided heterogeneity models by embedding in different general equilibrium frameworks a production technology similar to the one considered in this paper, giving rise to similar assignment problems. In models with neoclassical roots, Costinot and Vogel (2010) study the assignment of workers to tasks while Grossman, Helpman, and Kircher (2017) study the matching of managers and workers and

[5] Of note, trade among identical countries has no distributional effects.

their sorting into different industries.[6] In monopolistically competitive settings, Sampson (2014) and Somale (2015) analyze the assignment of workers to firms in models that extend Yeaple (2005) and Chaney (2008), respectively. However, a general equilibrium analysis of a similar extension of Melitz (2003), the canonical heterogeneous-firm trade model, has proved technically challenging.[7] I contribute to this literature by presenting said analysis under weaker assumptions about selection into exporting and by showing how this type of models can be taken to the data.

This paper contributes methodologically to this branch of the assignment trade literature by deriving a set of lemmas and propositions that facilitate the general equilibrium analysis of models featuring similar assignment problems. Among other results, I establish the existence and uniqueness of the equilibrium, a prerequisite for a theoretical analysis of comparative statics. Conditional on the activity cutoff, the market equilibrium is characterized by a system of nonlinear differential equations and a set of boundary conditions that together define a nonlinear two-point boundary value problem (BVP). In contrast to the cases of initial value problems (IVP) and linear BVPs, establishing existence and uniqueness of solutions is not trivial in the case of nonlinear BVPs, with off-the-shelf mathematical results typically covering particular cases of the problem. Despite these difficulties, several studies in the trade literature that use assignment models leading to similar BVPs simply assume or state without proof the existence and uniqueness of the solution. In this paper, I fill this gap in the trade literature for the case of a nonlinear two-point boundary BVP that encompasses those in this paper and others in the literature.[8]

This paper also relates to a literature proposing heterogeneous-firms models in which international trade can affect wage inequality through within-industry mechanisms. Motivated by developments in within-group wage inequality, one line of research develops models with labor market frictions in which ex-ante identical workers earn different wages at different firms, reflecting differences in efficiency wages (Davis and Harrigan 2011) or

[6] The distributional effects of trade in these tudies also relies on differences across countries.

[7] Aducing intractability, Sampson (2014) presents only some partial equilibrium results in this setting. In Somale (2015), I only considered the effects of opening to trade under no free entry and common fixed export costs accross firms.

[8] The general BVP considered in this paper encompasses those in Costinot and Vogel (2010), Sampson (2014), Somale (2015), Grossman, Helpman, and Kircher (2017).

fair wages (Egger and Kreickemeier 2009, 2012; Amiti and Davis 2012) required to induce worker effort, as well as differences in average ex-post worker ability amid search-and-matching frictions, unobservable worker ability and costly screening (Helpman, Itskhoki, and Redding 2010; Helpman et al. 2016). Given their focus on ex-ante indentical workers, these models cannot speak to the effects of trade on the relative reward to observable worker characteristics, such as the effects on the skill premium. By contrast, the framework in this paper can speak to these issues as well as to within-group inequality if workers are classified in large skill groups.

Another strand of this literature focuses on the effects of trade on the relative earnings of ex-ante heterogeneous workers (from the perspective of firms) through firms' technological choices (Yeaple 2005; Bustos 2011; Sampson 2014), workers' occupational choices (Monte 2011) or changes in the distribution of labor demand across firms differing in skill intensity (Somale 2015, Burstein and Vogel 2017). While these studies typically contemplate only a few large skill groups or place strong restrictions on selection into exporting, I consider a continuum of skill groups and a flexible specification of the latter.[9] This allows me to study the effects of trade on the entire wage distribution under empirically relevant restrictions on selection into exporting, showing that restrictions typically imposed on this margin can lead to significantly different distributional effects.

The rest of the paper is organized as follows. Section 2 describes the basic setup of the framework. Sections 3 and 4 characterize the equilibrium in the no-free-entry model and present existence and uniqueness results. Section 5 studies the effects of higher trade openness on wage inequality in the no-free-entry model, while section 6 extends the analysis to the free-entry model. After describing the calibration approach, section 7 discusses the implications of a calibrated version of the model. Section 8 concludes.

2 Basic Setup

This section develops a framework for studying the effects of higher trade openness on the wage distribution in which strong skill-productivity complementarities in production

[9]These strong restrictions are introduced by assuming common fixed export costs across firms in models based on Melitz (2003) and by imposing strong functional form assumptions on productivity distributions in models based on Eaton and Kortum (2002).

imply that inequality rises as workers reallocate towards more productive firms in the same industry. The model features a large number of skill groups and a flexible specification of fixed export costs that can accommodate weaker and more empirically relevant restrictions on firm selection into exporting than standard heterogeneous-firms trade models.

2.1 Demand

The preferences of the representative consumer are given by a C.E.S utility function over a continuum of goods indexed by ω :

$$U=\left[\int_{\omega\in\Omega} u(\omega)^{\frac{\sigma-1}{\sigma}}\,d\omega\right]^{\frac{\sigma}{\sigma-1}},$$

where $u(\omega)$ is the quantity consumed of good ω, the measure of the set Ω represents the mass of available goods and $\sigma>1$ is the elasticity of substitution between goods. The demand and expenditure for individual varieties generated by this utility function are

$$u(\omega)=EP^{\sigma-1}p(\omega)^{-\sigma}, \qquad E(\omega)=EP^{\sigma-1}p(\omega)^{1-\sigma}, \tag{1}$$

where P is the aggregate price level and E is aggregate expenditure,

$$P=\left[\int_{\omega\in\Omega} p(\omega)^{1-\sigma}\,d\omega\right]^{\frac{1}{1-\sigma}}, \qquad E=\int_{\omega\in\Omega} E(\omega)\,d\omega. \tag{2}$$

2.2 Production

There is a continuum of *active*, monopolistically competitive firms in the market, each producing a different variety ω.[10] As in Melitz (2003), firms differ in their productivity level ϕ, which they obtain as an independent draw from a distribution $G(\phi)$ with density function $g(\phi)$. I assume that the support of G, $\Phi\equiv\{\phi: g(\phi)>0\}$, is equal to some bounded interval of nonnegative real numbers, $[\underline{\phi},\overline{\phi}]\subseteq\mathbb{R}_+$. In contrast to Melitz (2003), the labor force is heterogenous, consisting of a continuum of workers of mass L that differ in their skill level s. The distribution of worker's skills is represented by a nonnegative density $V(s)$, so $LV(s)\geq 0$ represents the inelastic supply of workers with skill s. I only

[10]A firm is active in the market if it produces positive output.

consider skill distributions such that the support of V, denoted by S, is equal to some bounded interval of nonnegative real numbers—i.e., $S \equiv \{s : V(s) > 0\} = [\underline{s}, \overline{s}] \subseteq \mathbb{R}_+$.

The production technology of firms is represented by a cost function that exhibits constant marginal cost and fixed overhead costs. After paying the fixed costs described below, a firm must decide the mix of workers to use in production. The total output of a firm with productivity ϕ, $q(\phi)$, is given by

$$q(\phi) = \int_{s \in S} A(s,\phi)\, l(s,\phi)\, ds, \tag{3}$$

where $A(s,\phi)$ is the marginal productivity of a worker of skill s, and $l(s,\phi)$ is the total number of *production* workers of that skill level employed by the firm.[11] More-skilled workers are more productive than less skilled workers, regardless of the productivity of the firm that employs them. Also, more-productive firms have lower labor input requirements than less-productive firms no matter the type of worker considered. In terms of the production function (3), I formally assume that the productivity function $A(s,\phi)$ is strictly positive, strictly increasing and continuously differentiable—i.e., $A(s,\phi) > 0$, $A_s(s,\phi) > 0$ and$A_\phi(s,\phi) > 0$.[12]

In addition to the absolute productivity advantage described above, more-skilled workers have a comparative advantage in production at more-productive firms. Specifically, I follow Costinot and Vogel (2010) and assume that the function $A(.,.)$ is strictly log-supermodular, $A(s',\phi')\, A(s,\phi) > A(s',\phi)\, A(s,\phi')$ for all $s' > s$ and $\phi' > \phi$. Since $A(s,\phi) > 0$, the previous inequality can be rearranged as $\frac{A(s',\phi')}{A(s',\phi)} > \frac{A(s,\phi')}{A(s,\phi)}$, showing that the productivity gains from switching to a more-productive firm are higher for more-skilled workers. Alternatively, the gains from hiring a more-skilled worker are higher for more-productive firms.

Following a standard practice in the international trade literature, I assume that fixed costs are paid in terms of *nonproduction* workers. Specifically, I assume that firms pay a fixed cost of $fV(s)$ units of each skill $s \in S$, implying that the total fixed cost of a firm is $f\int_{\underline{s}}^{\overline{s}} w(s)\, V(s)\, ds = f\overline{w}$, where $w(s)$ is the wage of a worker with skill level s, and $\overline{w}$ is the average wage in the economy and the numeraire, $\overline{w} = 1$. This specification of fixed

[11] Firms also employ nonproduction workers as part of their fixed costs requirements.

[12] For any function $F(x_1, ..., x_n)$, F_{x_i} denotes the partial derivative of F with respect to variable x_i.

costs guarantees that the distribution of skills in the economy is still given by $V(s)$ after all fixed costs have been paid, implying that the demand of labor induced by fixed-costs requirements has no effect on the wage schedule $\{w(s)\}$. As such, the wage schedule is completely determined by the interactions between the exogenous relative supply of skills, captured by the distribution $V(s)$, and the endogenous relative demand of skills derived from the firm's demand of *production* workers.

Of note, I view the category of production workers in the model as somewhat broader than that used in most empirical studies. For example, larger firms typically have larger organizational structures and hence more middle managers. As such, middle managers are production workers through the lenses of the model, although they are typically classified as nonproduction workers in empirical studies. This observation implies that the model's implications about relative earnings of production and nonproduction workers are not comparable to available estimates from the the empirical literature, so I do not focus on these implications in the analysis.

2.3 Variable Costs and Prices

Per the linear production technology (3), workers are perfect substitutes in production. Accordingly, firms employ only those worker types that entail the lowest cost per unit of output, implying that the marginal cost of a firm with productivity ϕ, $c(\phi)$, is given by

$$c(\phi) = \min_{s \in S} \left\{ \frac{w(s)}{A(s,\phi)} \right\}. \tag{4}$$

For any wage schedule, the marginal cost $c(\phi)$ is strictly decreasing in the productivity level ϕ, as a firm can always hire the same type of workers employed by a less-productive competitor and obtain a strictly lower marginal cost due its absolute productivity advantage, $\phi' > \phi \Leftrightarrow c(\phi') < c(\phi)$.

Faced with the iso-elastic demands in (1), firms optimally set their price equal to a constant markup over their marginal costs, $p(\phi) = \frac{\sigma}{\sigma-1} c(\phi)$. This pricing rule and the cost minimization condition (4) imply

$$p(\phi) \leq \frac{\sigma}{\sigma-1} \frac{w(s)}{A(s,\phi)} \text{ for all } s \in S; \qquad p(\phi) = \frac{\sigma}{\sigma-1} \frac{w(s)}{A(s,\phi)} \text{ if } l(s,\phi) > 0. \tag{5}$$

2.4 Entry

I carry out the analysis under two widely-used assumptions regarding entry; no free entry a-lá Chaney (2008) and free entry a-lá Melitz (2003). In the first case, there is a fixed mass of firms in the industry. In the second case, there is unbounded pool of prospective firms that must pay a fixed entry cost to develop a new product variety and enter the industry. The results obtained under the no-free-entry assumption can be interpreted as the short-term consequences of trade, before investment in the development of new varieties leads new firms to enter the industry. In contrast, the results obtained under the free-entry assumption can be viewed as the long-term effects of trade.

3 No-Free-Entry Model: the Closed Economy

As in Chaney (2008), there is a fixed mass $\overline{M}$ of firms in the industry. A firm is *active* in the market if and only if it finds it profitable to produce. The pricing rule (5), the consumer's demand and expenditure functions in (1), and the goods-market clearing condition $(u(\omega) = q(\omega))$, imply that a firm's output, revenue and profit from serving the *domestic* market are given by

$$q^d(\phi) = EP^{\sigma-1}\left[\frac{\sigma}{\sigma-1}c(\phi)\right]^{-\sigma}; \; r^d(\phi) = EP^{\sigma-1}\left[\frac{\sigma}{\sigma-1}c(\phi)\right]^{1-\sigma}; \; \pi^d(\phi) = \frac{r^d(\phi)}{\sigma} - f, \tag{6}$$

where aggregate expenditure, E, equals aggregate income. The last expression, together with a decreasing marginal cost function $c(\phi)$, implies that a firm's profit is an increasing function of the firm's productivity.

There are combinations of parameters such that all firms are active in equilibrium, $\pi^d(\underline{\phi}) \geq 0$. However, since this case is not theoretically interesting nor empirically relevant, I focus on equilibria featuring selection into activity—i.e., the least-productive firms find it unprofitable to produce and remain inactive, $\pi^d(\underline{\phi}) < 0$.[13] In such an equilibrium, there is a cutoff productivity value $\phi^* \in (\underline{\phi}, \overline{\phi})$ such that only firms with productivity above this value are active in the market. The value of this *activity cutoff* corresponds

[13]Proposition 1 presents conditions on primitives that rule out this possibility.

to the level of productivity at which firms make zero profits,[14]

$$\pi^d(\phi^*) = 0. \tag{7}$$

In turn, the activity cutoff ϕ^* determines the total mass of active firms in the industry,

$$M = [1 - G(\phi^*)]\,\overline{M}. \tag{8}$$

Finally, the labor market of each type of worker must clear,

$$LV(s) = \int_{\phi^*}^{\overline{\phi}} l^d(s,\phi)\,\frac{g(\phi)}{[1 - G(\phi^*)]}\,d\phi M + MfV(s) \;\text{ for all } s \in S. \tag{9}$$

The left- and right-hand sides of the last expression capture, respectively, the total supply and demand of workers of skill s, with the total demand comprising the demand of *production* workers (first term), and the demand of *nonproduction* workers derived form the presence of fixed costs of production (second term). Having described all the components of the economy, I state the formal definition of the equilibrium.

Definition 1 *A no-free-entry equilibrium of the closed economy is a mass of active firms $M > 0$, a productivity activity-cutoff , $\phi^* \in (\underline{\phi}, \overline{\phi})$, an output function $q^d : [\phi^*, \overline{\phi}] \to \mathbb{R}_+$, a labor allocation function $l^d : S \times [\phi^*, \overline{\phi}] \to \mathbb{R}_+$, a price function $p : [\phi^*, \overline{\phi}] \to \mathbb{R}_+$ and a wage schedule $w : S \to \mathbb{R}_+$ such that the following conditions hold,*[15]
(i) consumers behave optimally, equations (1) and (2);
(ii) firms behave optimally given their technology, equations (3), (5), (7) and (8);
(iii) goods and labor markets clear, equations (6) and (9), respectively;
(iv) the numeraire assumption holds, $\overline{w} = 1$.

3.1 Characterization of the Equilibrium

The log-supermodularity of the productivity function, A, implies that the equilibrium labor allocation is characterized by positive assortative matching—i.e., more-productive

[14] If all firms are active in the market, then $\phi^* = \underline{\phi}$, and condition (7) may not hold.
[15] Technically, this definition corresponds to an equilibrium featuring selection into activity.

firms employ *production* workers of higher ability. Specifically, there exists a continuous and strictly increasing matching function $N : S \to [\phi^*, \overline{\phi}]$ such that, all firms of productivity $N(s)$ employ production workers of skill s, and all production workers of skill s are employed at firms with the productivity $N(s)$. Behind this result, formally stated in lemma 1, lies a simple intuition. The cost-minimization condition (4) implies that a firm of productivity ϕ' employing a worker of skill s' cannot reduce its marginal cost of production by employing a worker of a different skill, that is, $w(s')/A(s', \phi') \leq w(s)/A(s, \phi')$ for all $s \in S$. This observation and the strict log-supermodularity of A imply that, for any skill level $s > s'$ and any productivity level $\phi < \phi'$, the following inequalities hold, $\frac{A(s,\phi)}{A(s',\phi)} < \frac{A(s,\phi')}{A(s',\phi')} \leq \frac{w(s)}{w(s')}$. Accordingly, a firm with productivity $\phi < \phi'$ does not employ workers of skill $s > s'$, as it can obtain a strictly lower marginal cost by hiring a worker of skill s'. Although this argument only proves that the matching function is weakly increasing, it highlights the connection between the log-supermodularity of A and positive assortative matching in equilibrium.

Armed with the previous result, the equilibrium can be characterized in terms of the matching function N, revealing a tight connection between the latter and wage inequality in the current framework. A worker of skill s is matched to a firm with productivity $N(s)$ in equilibrium if and only if the skill level s solves the cost minimization problem (4) for any firm with productivity $\phi = N(s)$. The first order condition for an interior solution of this problem yields the following equilibrium condition,[16]

$$\frac{d \ln w(s)}{ds} = \frac{\partial \ln A(s, N(s))}{\partial s}. \tag{10}$$

The last expression is central in the analysis of wage inequality. It implies that the matching function N is a sufficient statistic for the dispersion of wages in the economy, as it is the only endogenous variable affecting the slope of the wage schedule. The connection between N and wage inequality can be seen more clearly by integrating (10) between s' and $s'' > s'$ to get $w(s'')/w(s') = \exp\{\int_{s'}^{s''} \frac{\partial \ln A(t,N(t))}{\partial s} dt\}$. The last expression, together with the strict log-supermodularity of A, implies that the ratio $w(s'')/w(s')$ is increasing in the values that the matching function takes on the interval $[s', s'']$. Then, any change in the environment leading to an upward shift of the matching function on a given interval

[16] As stated in lemma 1, all the endgogenous functions considered in this section are differentiable.

also leads to higher relative wages for more-skilled workers in that interval. Moreover, the new distribution of wages in the interval is second-order stochastically dominated by the old one, so inequality is pervasively higher after the change.[17]

Letting $H : [\phi^*, \overline{\phi}] \rightarrow S$ denote the inverse function of the matching function N, the optimal pricing rule (5) and the expression for revenues in (6) can be used to express firm's prices and revenues as functions of the productivity level ϕ and the value of the function H at that productivity level. Totally differentiating these functions with respect to ϕ and using equation (10) in the resulting expressions yields

$$p_{\phi}(\phi) = -p(\phi) \frac{\partial \ln A(H(\phi), \phi)}{\partial \phi}, \tag{11}$$

$$r^d_{\phi}(\phi) = (\sigma - 1) r^d(\phi) \frac{\partial \ln A(H(\phi), \phi)}{\partial \phi}. \tag{12}$$

The last two equations imply that the equilibrium matching of workers and firms is also a sufficient statistic for the dispersion of firms' prices and revenues. In particular, integrating equation (12) reveals that for $\phi'' > \phi'$, the ratio of revenues $r^d(\phi'')/r^d(\phi')$ is increasing in the values that the *inverse* of the matching function takes on $[\phi', \phi'']$, so a shift in the matching function will have opposite effects on the dispersion of wages and revenues in the closed economy.

The equilibrium labor allocation must be consistent with market clearing in the labor and goods markets—i.e., N (or H) must be consistent with conditions (1), (3), (6) and (9). This consistency requirement yields the following equilibrium condition,

$$H_{\phi}(\phi) = \frac{r^d(\phi)\, g(\phi)\, \overline{M}}{A(H(\phi), \phi) \left[L - f\left[1 - G(\phi^*)\right] \overline{M}\right] V(H(\phi))\, p(\phi)}, \tag{13}$$

which, after some re-arrangement, states that consumers' expenditure accruing to firms with productivity ϕ, $r^d(\phi)\, g(\phi)\, \overline{M}$, must equal the total value of the output that those firms can produce with the workers they employ.

Given the equilibrium activity cutoff, ϕ^*, equations (11)-(13) form a system of non-linear differential equations that the price function, p, the revenue function, r^d, and the inverse of the matching function, H, must satisfy in equilibrium. As is well-known, there is

[17] In appendix B.1.2, I show that the new distribution is Lorenz dominated by the previous one.

an uncountable family of functions that satisfy a system like (11)-(13), so a set of *boundary conditions* is needed to pin down a particular solution. Two of these boundary conditions are provided by the labor market clearing condition, as all workers must be assigned to some firm in equilibrium, $H(\phi^*) = \underline{s}$, $H(\overline{\phi}) = \overline{s}$. A third boundary condition is provided by the zero-profit condition for firms with productivity ϕ^*, $r^d(\phi^*) = \sigma f$. Finally, the activity cutoff ϕ^* can be determined from the the following equilibrium condition,

$$\tfrac{\sigma-1}{\sigma}\int_{\phi^*}^{\overline{\phi}} r^d(\phi) g(\phi)\, d\phi \overline{M} + f\left[1 - G(\phi^*)\right]\overline{M} = L, \tag{14}$$

which states that the total wages paid by firms to production and nonproduction workers (left) equals total labor income in the economy, where the expression for the latter uses the numeraire assumption. I summarize the results in this section in the following lemma.

Lemma 1 *In a no-free-entry equilibrium of the closed economy there exists a continuous and strictly increasing matching function $N : S \to [\phi^*, \overline{\phi}]$ (with inverse function H) such that (a) $l^d(s, \phi) > 0$ if and only if $N(s) = \phi$, (b) $N(\underline{s}) = \phi^*$, and $N(\overline{s}) = \overline{\phi}$. In addition, the following conditions hold*
(i) The wage schedule w is continuously differentiable and satisfies (10).
(ii) The price, revenue and matching functions, $\left\{p, r^d, N(\text{and } H)\right\}$, are continuously differentiable. Given ϕ^*, the triplet $\left\{p, r^d, H\right\}$ solves the boundary value problem (BVP) comprising the system of differential equations (11)-(13) and the boundary conditions $r^d(\phi^*) = \sigma f$, $H(\phi^*) = \underline{s}$, $H(\overline{\phi}) = \overline{s}$.
(iii) The activity cutoff ϕ^* and the revenue function r^d satisfy (14).
Moreover, if a number $\phi^* \in (\underline{\phi}, \overline{\phi})$, and functions $p, r^d : [\phi^*, \overline{\phi}] \to \mathbb{R}_+$ and $H : [\phi^*, \overline{\phi}] \to S$ satisfy conditions (ii)-(iii), then they are, respectively, the productivity activity-cutoff, the price function, the revenue function, and the inverse of the matching function of a no-free-entry equilibrium of the closed economy.

4 No-Free-Entry Model: the Open Economy

Balanced trade takes place between $n + 1$ symmetric (identical) economies of the type described above, so the description presented in section 2, including equations (1)-(5), holds for each of these economies. Given that the symmetry assumption ensures that all countries share the same equilibrium variables, I restrict the analysis to the home country. Firms face fixed and variable trade costs. Per-unit trade costs are common to all firms and are modeled in the standard iceberg formulation, whereby $\tau > 1$ units of a good must be shipped in order for 1 unit to arrive in a foreign destination. In contrast, fixed export costs vary across firms. A firm that wishes to export to country i must incur an idiosyncratic fixed cost of y units of a "bundle of skills" comprising $f^x V(s)$ workers of each skill $s \in S$. With the average wage as the numeraire, the total fixed export cost of the firm is $f^x y$ per foreign market. I model the firm-specific *size* of fixed export costs, y, as the realization of a nonnegative random variable Y with CDF F, which I assume is independent of the productivity distribution, absolutely continuous, and satisfies $F(y) = 0$ for $y \leq \underline{y}$, $dF(y) > 0$ for $y \geq \underline{y}$, where $\underline{y}$ is the lower bound of the support of Y. In addition, I assume that $f_x \underline{y} \tau^{\sigma-1} > f$, which guarantees that a firm's profit in the domestic market is always higher than in any individual foreign market.[18]

These assumptions about fixed export costs have three important implications. First, as in the case of fixed production costs, formulating fixed export costs in terms of said bundle of skills guarantees that the demand of labor induced by fixed-export-costs requirements does not affect the wage schedule. Second, in the presence of heterogeneous fixed export costs, a highly productive firm may not find it profitable to export if it faces high fixed export costs, while a less productive competitor may choose to serve the foreign market if its fixed export costs are sufficiently low. As a result, the productivity distributions of exporters and nonexporters overlap in equilibrium, consistent with the evidence in Bernard, Eaton, Jensen, and Kortum (2003). Third, an implication of the restriction $f_x \underline{y} \tau^{\sigma-1} > f$ is that, as in Melitz (2003), the activity status of a firm in the open economy continues to be determined by its domestic profit. Although not essential for the qualitative results in the paper, this implication simplifies the exposition.[19]

The determination of the set of active firms and their operations in the domestic market are little changed relative to the closed economy. There is a fixed mass $\overline{M}$ of *potential* firms in the industry. A firm is *active* if and only if it makes nonnegative profits in the

[18] A similar relationship between domestic and foreing profits is featured in Melitz (2003).

[19] Alternatively, I could have just assumed that a firm is active if and only if it makes positive profits in the domestic market, regardless of its potential export profits.

domestic market. The pricing rule (5) and the expenditure functions in (1) imply that the *potential* domestic output, q^d, revenue, r^d, and profit, π^d, of a firm with productivity ϕ are still given by (6). As before, domestic profits are strictly increasing in ϕ, so the equilibrium is characterized by a cutoff productivity level, $\phi^* \in (\underline{\phi}, \overline{\phi})$, such that a firm is active in the market if and only if its productivity is above this level.[20] Firms with productivity ϕ^* make zero domestic profit, condition (7), while the mass of active firms, M, is given by (8).

The equilibrium in the open economy features selection into trade—i.e., only a subset of active firms export. An active firm serves a foreign market if and only if it can make nonnegative profits there. In the presence of variable trade costs, consumers in each country face higher prices for imported goods, $p^x(\phi) = \tau p(\phi)$, so conditions (5) and (1) and the symmetry assumption imply that the *potential* export output, revenue and profit of a firm with productivity ϕ and fixed export costs $f^x y$ are given by

$$q^x(\phi) = \tau^{1-\sigma} q^d(\phi), \; r^x(\phi) = \tau^{1-\sigma} r^d(\phi), \; \pi^x(\phi) = \frac{\tau^{1-\sigma} r^d(\phi)}{\sigma} - f^x y. \tag{15}$$

Then, such a firm exports if and only if $y \leq \tau^{1-\sigma} r_d(\phi) / \sigma f^x$, which, together with the assumptions about y, implies that only a fraction $F(\tau^{1-\sigma} r_d(\phi) / \sigma f_x)$ of firms with productivity $\phi \geq \phi^*$ export. Note that this fraction is a continuous and increasing function of the productivity level ϕ, so exporters are, on average, more productive than nonexporters. These observations imply that the mass of exporters with productivity ϕ is

$$M^x(\phi) = g(\phi) F\left(\tau^{1-\sigma} r^d(\phi) / \sigma f^x\right) \overline{M}. \tag{16}$$

Finally, the labor market of each type of worker must clear,

$$\begin{aligned} LV(s) = {} & \textstyle\int_{\phi^*}^{\overline{\phi}} [l^d(s,\phi) g(\phi) \overline{M} + l^x(s,\phi) M^x(\phi)] d\phi + \cdots \\ & \cdots f M V(s) + \textstyle\int_{\phi^*}^{\overline{\phi}} n f^x \int_0^{\frac{\tau^{1-\sigma} r^d(\phi)}{\sigma f_x}} y dF(y) g(\phi) \overline{M} d\phi V(s).^{21} \end{aligned} \tag{17}$$

The left- and right-hand sides of the last expression capture, respectively, the total supply and demand for workers of skill s. Total demand comprises the demand of *production*

[20] As before, I focus on equilibria featuring selection into activity, i.e. $\pi^d(\underline{\phi}) < 0$.
[21] $l^x(s,\phi)$ represents exports-induced demand for production workers.

workers to supply the domestic and foreign markets, first term, and the demand of *non-production* workers derived form the presence of fixed costs of production and fixed export costs, the second and third terms. Conditions (1)-(3), (5)-(8), (15)-(17) and the numeraire assumption completely describe the equilibrium, prompting the formal definition of equilibrium in the appendix, analogous to that for the closed economy.

4.1 Characterization of the Equilibrium

The equilibrium of the open economy shares several features with its closed-economy counterpart. Cost minimization by firms and the strict log-supermodularity of A imply that the equilibrium labor allocation in the open economy is characterized by a strictly increasing matching function, N, that maps the set of skills, S, to the set of productivity levels of active firms, $[\phi^*, \overline{\phi}]$. In addition, equation (10), connecting the wage schedule to the matching function, and equations (11) and (12), connecting the price and domestic-revenue functions to the *inverse* of the matching function, H, continue to hold. As before, these equilibrium conditions imply that the matching function N (and its inverse H) is a sufficient statistic for the dispersion of wages, prices and domestic revenues.

The equilibrium labor allocation must be consistent with labor and goods markets clearing—i.e., N (or H) must be consistent with conditions (3), (6), (15) and (17). This observation and the expression for the mass of exporters, equation (16), yield the following equilibrium condition,

$$H_{\phi}\left(\phi\right) = \frac{r^d(\phi)\left[1+F\left(\frac{r^d(\phi)\tau^{1-\sigma}}{\sigma f^x}\right)n\tau^{1-\sigma}\right]g(\phi)\overline{M}}{A(H(\phi),\phi)V(H(\phi))p(\phi)\left[L-fM-\int_{\phi^*}^{\overline{\phi}} nf^x \int_0^{\frac{r^d(\phi')\tau^{1-\sigma}}{\sigma f^x}} ydF(y)g(\phi')\overline{M}d\phi'\right]}. \tag{18}$$

After some re-arrangement, the last expression states that the total revenue that firms with productivity ϕ make from their sales in the domestic and foreign markets, the numerator on the right-hand side of (18), must equal the total value of the output that those firms can produce with the workers they employ.

Given the equilibrium activity cutoff, ϕ^*, equations (11), (12) and (18) form a system of nonlinear differential equations that the price function, p, the domestic revenue function, r^d, and the inverse of the matching function, H, must satisfy in equilibrium. Two boundary conditions for this system are provided by the labor market clearing condition, as all workers must be assigned to some firm in equilibrium, $H(\phi^*) = \underline{s}$, $H(\overline{\phi}) = \overline{s}$. A third boundary condition is provided by the zero-domestic-profit condition for firms with productivity ϕ^*, $r^d(\phi^*) = \sigma f$. Finally, the open-economy counterpart of equation (14) can be used to determine the activity cutoff ϕ^*,

$$\begin{array}{c} \frac{\sigma-1}{\sigma}\int_{\phi^*}^{\overline{\phi}} r^d(\phi)[1+F\left(\frac{r^d(\phi)\tau^{1-\sigma}}{\sigma f^x}\right)n\tau^{1-\sigma}]g(\phi)\,d\phi\overline{M}+\cdots \\ \cdots fM+\int_{\phi^*}^{\overline{\phi}} nf^x\int_0^{\frac{r^d(\phi')\tau^{1-\sigma}}{\sigma f^x}} ydF(y)\,g(\phi')\,\overline{M}d\phi' \end{array} = L, \tag{19}$$

which states that the total value of wages paid by firms to production and nonproduction workers (left) equals total labor income in the economy, where the expression for the latter uses the numeraire assumption. As in the closed economy case, the conditions derived in this section are not only necessary, but also sufficient for an equilibrium. This characterization of the equilibrium is summarized in lemma 3 in the appendix, which can be easily proved adapting the arguments in the proof of lemma 1.

I conclude this section with a summary of the qualitative properties of the equilibrium in the open economy. In equilibrium, more-productive firms employ *production* workers of higher ability and pay them higher wages. The stochastic specification of fixed export costs yields an imperfect positive correlation between firms' productivity, average workforce ability, size and export status, which is consistent with the empirical evidence documented in Bernard and Jensen (1995) and Bernard, Eaton, Jensen, and Kortum (2003).

4.2 Existence and Uniqueness of the Equilibrium

I start this section by studying the existence and uniqueness of solutions to the nonlinear, two-point BVPs characterizing the equilibrium in the closed and open economies. In contrast to the cases of initial value problems (IVPs) and linear BVPs, for which there is a standard theory that provides fairly general results under relatively mild restrictions on the *data* of the problem, such a study is not trivial in the case of nonlinear BVPs for two reasons.[22] First, there is no unified theory that can be applied to study these issues for an arbitrary problem. Because of the complexity of the subject, the mathematical literature has typically focused on particular cases of the problem, leading to a multitude of theoretical approaches tailored to these cases.[23] Second, most results in the literature are based on restrictive and not-easily-verifiable assumptions, while those results based on less restrictive assumptions, resembling those used in the standard theory of IVPs, have a local flavor.[24] Despite these difficulties, several studies in the trade literature that use assignment models and arrive to characterizations of the equilibrium involving a BVP similar to those above, simply assume or state without proof the existence and uniqueness of the solution. In this section, I fill this gap in the trade literature by presenting existence and uniqueness results for a nonlinear BVP that encompasses the two BVPs considered above and others in the literature.[25]

For any $\phi_0, \phi_1 \in [\underline{\phi}, \overline{\phi}]$ and $s_0, s_1 \in [\underline{s}, \overline{s}]$, with $\phi_0 < \phi_1$ and $s_0 < s_1$, I consider the nonlinear, two-point BVP (20), comprising the system of differential equations (20a)-(20c) and the boundary conditions (20d),

$$z_\phi(\phi) = -z(\phi)\frac{\partial \ln A(\Gamma(\phi), \phi)}{\partial \phi}, \tag{20a}$$

$$x_\phi(\phi) = (\sigma - 1)\, x(\phi)\frac{\partial \ln A(\Gamma(\phi), \phi)}{\partial \phi}, \tag{20b}$$

$$\Gamma_\phi(\phi) = \frac{x(\phi)\left[1 + F(K_0 x(\phi))\, K_1\right]\alpha(\phi) g(\phi)}{A(\Gamma(\phi), \phi)\, V(\Gamma(\phi))\, z(\phi)}, \tag{20c}$$

$$x(\phi) = 1,\ \Gamma(\phi_0) = s_0,\ \Gamma(\phi_1) = s_1, \tag{20d}$$

where $\alpha(\phi)$ is a strictly positive continuous function, $\alpha : [\underline{\phi}, \overline{\phi}] \to \mathbb{R}_{++}$, K_0 and K_1 are nonnegative constants and $\{A, g, V, F\}$ are the functions defined earlier.

[22] For a discussion of standard existence and uniqueness theory for IVPs see Agarwal and O'Regan (2008a), which also covers basic results for linear BVPs. For a more comprehensive treatment of linear BVPs see Stakgold (1998) and Agarwal and O'Regan (2008b).

[23] Bernfeld and Lakshmikantham (1974) survey the most common problems and theoretical approaches considered in the literature. See Kiguradze (1988) for some results for the general two-point BVP.

[24] Bailey, Shampine, and Waltman (1968) present several existence and uniqueness results for nonlinear BVPs using Piccard's Iteration method when the functions involved satisfy certain Lipschitzian conditions. In all cases, the interval over which the solution is defined has to be sufficiently small.

[25] The general BVP considered in this section encompasses those in Costinot and Vogel (2010), Sampson (2014), Somale (2015), Grossman, Helpman, and Kircher (2017).

The general BVP defined above nests the BVPs corresponding to the closed and open economies, as the latter can be obtained as particular parametrizations of the former. If we set $K_0 = (f/f_x)\,\tau^{1-\sigma}$, $K_1 = n\tau^{1-\sigma}$, $\phi_0 = \phi^*$, $\phi_1 = \overline{\phi}$ and $\alpha(\phi) = 1$ for all $\phi \in [\underline{\phi}, \overline{\phi}]$, the resulting BVP is equivalent to the BVP of the open economy, in the sense that any solution to one of these two BVPs can be used to construct a solution to the other. To see this, let $\{z, x, \Gamma\}$ be a solution to the BVP (20) parametrized as above. If we define $r^d(\phi) \equiv \sigma f x(\phi)$, $p(\phi) \equiv z(\phi)\,\sigma f \overline{M}/[L - fM - \int_{\phi^*}^{\overline{\phi}} n f_x \int_0^{f x(\phi')\tau^{1-\sigma}/f_x} y dF(y)\, g(\phi')\, \overline{M} d\phi']$ and $H = \Gamma$, then $\{p, r^d, H\}$ is a solution to the BVP of the open economy. A similar argument shows that any solution to the BVP of the open economy can be used to construct a solution to this particular parametrization of BVP (20). Finally, if we set $K_1 = 0$ in the parametrization above, the resulting BVP is equivalent to the BVP of the closed economy defined in lemma 1.ii.

Lemma 2 states some important results about the general BVP (20).

Lemma 2 *If the right-hand side of equations (20a)-(20c) are locally Lipschitz continuous with respect to $\{z, x, \Gamma\}$, then there is a unique continuously differentiable solution to the BVP (20) for any $\phi_0, \phi_1 \in [\underline{\phi}, \overline{\phi}]$ and $s_0, s_1 \in [\underline{s}, \overline{s}]$, with $\phi_0 < \phi_1$ and $s_0 < s_1$. As a function of (ϕ_0, s_0), the solution to the BVP, $\{z(.;\phi_0, s_0), x(.;\phi_0, s_0), \Gamma(.;\phi_0, s_0)\}$, satisfies the following conditions,*
(i) (no crossing) If $K_1 = 0$ and Γ^{-1} is the inverse of Γ, then $s_0^a < s_0^b$ implies $\Gamma(\phi;\phi_0, s_0^a) < \Gamma(\phi;\phi_0, s_0^b)$ on $[\phi_0, \phi_1)$, while $\phi_0^a > \phi_0^b$ implies $\Gamma^{-1}(s;\phi_0^a, s_0) > \Gamma^{-1}(s;\phi_0^b, s_0)$ on $[s_0, s_1)$.
(ii) $\phi_0^a > \phi_0^b$ implies $x(\phi;\phi_0^a, s_0) < x(\phi;\phi_0^b, s_0)$ on $[\phi_0^a, \phi_1]$.

I present a brief outline of the proof of the last lemma below, relegating the details to the appendix. To prove existence, I follow O'Regan (2013) and recast the BVP as a fixed point problem. In particular, I show that a triplet $\{z, x, \Gamma\}$ solves BVP (20) if and only if Γ is a fixed point of some compact functional, Ψ, defined over a convex and closed set K, $\Psi(\Gamma) = \Gamma$. Then, a direct application of Schauder fixed point theorem yields the existence result. The uniqueness of the solution is established as a consequence of the particular structure of the problem and the strict log-supermodularity of A. Lemma 1.i is obtained as a corollary of the uniqueness result. For $K_1 = 0$ (closed economy), lemma 1.ii immediately follows from the previous no-crossing result, (20b) and the log-supermodularity of A. However, this argument cannot be extended to the case $K_1 > 0$

(open economy), as the no-crossing property no longer holds. In the appendix, I present a slightly longer argument that is valid for the general case $K_1 \geq 0$, which also establishes the result as a consequence of the strict log-supermodularity of A.

An important corollary of the discussion so far is that, for a given activity cutoff ϕ^*, the functions r^d and H that solve the BVPs of the closed and open economies do not depend on the mass of firms, $\overline{M}$, nor the mass of production workers.[26] This feature of the solution follows from the uniqueness result in lemma 2, equation (20c) and the correspondence between said BVPs and BVP (20) described above. In fact, the mass of firms and the mass of production workers affect only the level of the solution function p. This result will prove useful in the analysis of the free-entry model in section 6.

As the BVP of the open economy has a unique solution conditional on the activity cutoff ϕ^*, then there exists a unique equilibrium of the open economy if and only if there is a unique value of ϕ^* that solves equation (19). Given the correspondence between the open-economy BVP and the general BVP (20), lemma 2.ii implies that $r^d(\phi)$ is strictly decreasing in the activity cutoff ϕ^*, making the left-hand side of (19) strictly decreasing in the value of ϕ^*. As the right-hand side of (19) does not depend on ϕ^*, there is a unique solution to (19) if the size of the market, as captured by L, is not too large.[27] A similar argument shows that there is a unique equilibrium in the closed economy. I summarize this discussion in the next proposition, which also establishes the (constrained) efficiency of the equilibrium.

Proposition 1 *Let $\left\{\underline{p}, \underline{r}^d, \underline{H}\right\}$ and $\left\{\underline{p}^a, \underline{r}^{d,a}, \underline{H}^a\right\}$ be, respectively, the solution to the BVPs characterizing the open- and closed-economy equlbria with $\phi^* = \underline{\phi}$. In addition, let $\beta(r^d, \phi^*)$ and $\beta^a(r^d, \phi^*)$ denote the functions defined by the left-hand sides of equations (19) and (14), respectively, in terms of ϕ^* and r^d.*
(i) For $\beta(\underline{r}^d, \underline{\phi}) > L$, there is a unique no-free-entry equilibrium of the open economy.
(ii) For $\beta^a(\underline{r}^d, \underline{\phi}) > L$, there is a unique no-free-entry equilibrium of the closed economy. In addition, the equilibrium of the closed economy is efficient, while that of the open economy is efficient when $f \leq f_x \tau^{1-\sigma}$, and constrained efficient when $f > f_x \tau^{1-\sigma}$.

[26] The mass of production workers in the closed and open economies are given by the term in brackets in the denominator of the right-hand side of equations (13) and (18), respectively.

[27] If L is too large relative to the mass of firms, $\overline{M}$, then there is no equilibrium featuring selection into activity as all firms make postive profits.

5 No Free Entry, Trade and Wage Inequality

In this section, I study the effects of higher trade openness on wage inequality in the no-free-entry model described above. In the model, a decline in trade frictions induces a reallocation of production and employment across firms with heterogenous skill demand, affecting the aggregate relative demand for skills and the relative wages in the economy. In the analysis, I decompose these effects into the contributions of each of the three channels defined in the introduction—the selection-into-activity, intensive-margin and extensive-margin channels.

Being a sufficient statistic for the dispersion of wages in the model, the matching function takes center stage in the subsequent analysis, as any result about wage inequality in this framework is a statement about the impact on the matching function of the shock under consideration. Lemma 4 in the appendix collects several results related to the general BVP in (20) that are instrumental to the analysis. In particular, this lemma characterizes the dependence of the solution function Γ (and some functionals of Γ) on the parameters of the problem.

5.1 Autarky vs. Trade

The first instance of higher trade openness that I consider is the case of an initially autarkic economy that opens up to trade. I start this section with one of the main results of the paper, Proposition 2, which states that opening to trade leads to a pervasive increase in wage inequality.

Proposition 2 *Let $\{\phi_a^*, N^a\}$ and $\{\phi_\tau^*, N^\tau\}$ be the activity cutoffs and matching functions corresponding to the no-free-entry equilibrium of the closed and open economies, respectively. Then the following conditions hold:*
(i) $\phi_\tau^* > \phi_a^*$ and $N^\tau(s) > N^a(s)$ for all $s \in [\underline{s}, \overline{s})$, so inequality is pervasively higher in the open economy.
(ii) The selection-into-activity and extensive-margin channels lead to pervasively higher inequality (intensive-margin channel not operational).

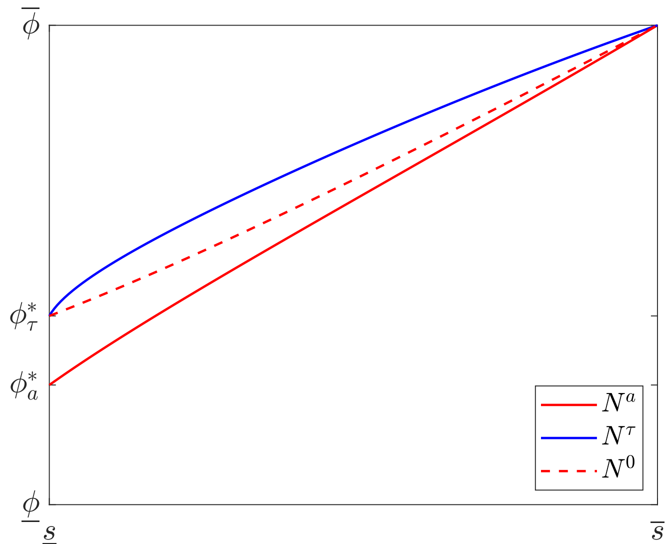
The first result in the last proposition, $\phi_\tau^* > \phi_a^*$, states that the selection-into-activity effects of trade highlighted in Melitz (2003) always hold in the no-free-entry model of this paper—i.e., trade induces the least productive firms to exit the market. Although somewhat trivial in homogenous-workers models a-lá Melitz/Channey, this result is not immediate in the current framework. For example, in an homogenous-workers version of the no-free-entry model above, assuming that firms with productivity ϕ_a^* are still active after the economy starts trading results in unchanged domestic revenues and labor costs. With aggregate labor costs pinned down by an equilibrium condition, this observation, together with positive export labor costs, implies that a higher activity cutoff is required in the open economy. In contrast, making the same assumption in the heterogeneous-worker framework above leads to lower domestic revenues and labor costs, so establishing the result requires proving that the decline in the latter is more than offset by the new labor costs of exporting (variable and fixed). I do so in the appendix by showing that total wages paid to *production* workers necessarily increase if the activity cutoff remains unchanged, which together with the presence of fixed export labor costs, leads to a rise in the the total wages paid by firms. With total wages pinned down by the numeraire assumption, condition (19), a higher activity cutoff is required in the open economy.[28]

To gain more insight into the effects of opening to trade on wage inequality, I decompose the overall effect into the three channels defined earlier. First of all, note that the intensive-margin channel is not operational in this case, as there were no exporters before the economy started to trade. The selection-into-activity channel captures the impact on wage inequality of the trade-induced increase in the activity cutoff, excluding the impact of changes in the set of exporters. To isolate the effect of this channel, I contrast the matching function of the closed economy with that of an ancillary autarkic economy that differs from the former only in that its activity cutoff is given by that of the open economy. That is, the equilibria of the closed and ancillary economies are characterized by the BVP in lemma 1.ii with $\phi^* = \phi_a^*$ and $\phi^* = \phi_\tau^*$, respectively. The typical situation is depicted in figure 1, where the solid and dashed red lines are, respectively, the matching functions of the closed (N^a) and ancillary (N^0) economies. The no-crossing result in lemma 2.i. implies that the latter lies strictly above the former on $[\underline{s}, \overline{s})$ as shown in the figure. Intuitively, as the firms with productivity in the range $[\phi_a^*, \phi_\tau^*)$ become inactive, the aggregate demand for workers with skills in the range $[\underline{s}, N^a(\phi_\tau^*))$ drops to zero barring any change in the wage schedule. Per the labor market clearing condition, these

[28] As explained earlier, the left-hand side of (19) is strictly decreasing in the activity cutoff.

workers must be reallocated among the firms that remain active, requiring a decline in their relative wages.

Figure 1: Opening to Trade and the Matching Function



Note: The solid red and blue lines represent, respectively, the matching functions of the closed (N^a) and open (N^τ) economies. The dashed red line depicts the matching function of the ancillary autarkic economy (N^0) described in the text. The differences between N^a and N^0 and between N^0 and N^τ capture the impact of the selection-into-activity and extensive-margin channels, respectively.

The extensive-margin channel reflects the impact on wage inequality of the increased labor demand by new exporters as they expand their production to serve the foreign market, excluding the effects of changes in the activity cutoff. Put another way, this channel captures the effects of replacing $[1 + F\left(r^d\left(\phi\right)\tau^{1-\sigma}/\sigma f^x\right) n\tau^{1-\sigma}]$ with 1 in the BVP of the open economy, precisely what the difference between the matching functions of the ancillary (N^0) and open (N^τ) economies in figure 1 captures, with the latter shown in blue. To see why N^τ necessarily lies above N^0 as depicted in the figure, suppose for a moment that the wages of the ancillary economy also prevail in the open economy. In this case, firms of a given productivity level demand the same skill type of workers in both economies, with exporters in the open economy demanding more labor than nonexporters due to the foreign demand they face. If the fraction of exporters was constant across productivity levels, this additional export-driven labor demand would affect all skill levels proportionally, leaving unchanged the overall relative demand for skills in the economy.

However, as the fraction of exporters in the model increases with firms' productivity, this additional export-driven labor demand is tilted towards more-able workers, resulting in an excess demand for this type of labor. As such, market clearing requires higher relative wages for more-skilled workers in the open economy.[29]

I conclude this section with a discussion of the impact of trade on the *level* of real wages. Although trade always raises the average real wage, the least-skilled workers in the economy may see their real wage decline. The pricing rule (5) and the zero profit condition (7) imply that the aggregate price indices of the closed (P^a) and open (P^τ) economies satisfy

$$\left(P^i\right)^\sigma = \frac{\sigma f}{U^i}\left[\frac{\sigma}{(\sigma-1)}\frac{w^i\left(\underline{s}\right)}{A\left(\underline{s},\phi_i^*\right)}\right]^{\sigma-1} \text{ for } i=a,\tau, \tag{21}$$

where U^i is the aggregate real expenditure/income in the economy. Per the efficiency result in proposition 1, real income is higher in the open economy, $U^\tau > U^a$.[30] In addition, proposition 2.i, together with the numeraire assumption ($\overline{w}^i = 1$), implies that the open economy exhibits a higher activity cutoff, $\phi_\tau^* > \phi_a^*$, and a lower wage for the least-able workers, $w^\tau\left(\underline{s}\right) < w^a\left(\underline{s}\right)$. Accordingly, $P^\tau < P^a$, so the average real wage, $\overline{w}/P$, is higher in the open economy.

Finally, recalling that $U^i = E^i/P^i$, equation (21) can be rearranged to get the an expression for the real wage of the least-able workers, $\frac{w^i(\underline{s})}{P^i} = \frac{(\sigma-1)}{\sigma}A\left(\underline{s},\phi_i^*\right)\left[E^i/\sigma f\right]^{\frac{1}{\sigma-1}}$. This expression implies that opening to trade necessarily improves the real wage of these workers when it induces a rise in aggregate expenditure/income. However, in some parameterizations of the model, opening to trade can induce a decline in the real wage of the poorest workers, as the drop in aggregate income more than offsets the boost from working at a more productive employer (higher activity cutoff).

[29] Formally, in the appendix I show that the BVPs of the ancillary and open economies can be conceived as particular parameterizations of the general BVP (20) with $K_1 = 0$ that differ only in the parameter function $\alpha(\phi)$, which is constant in the former and increasing in the latter. The result then follows from a direct application of lemma 4.i in the appendix.

[30] Note that the closed economy allocation is available to the planner of the open economy, so a simple revealed-preference argument yields $U^\tau > U^a$.

5.2 Trade Liberalization

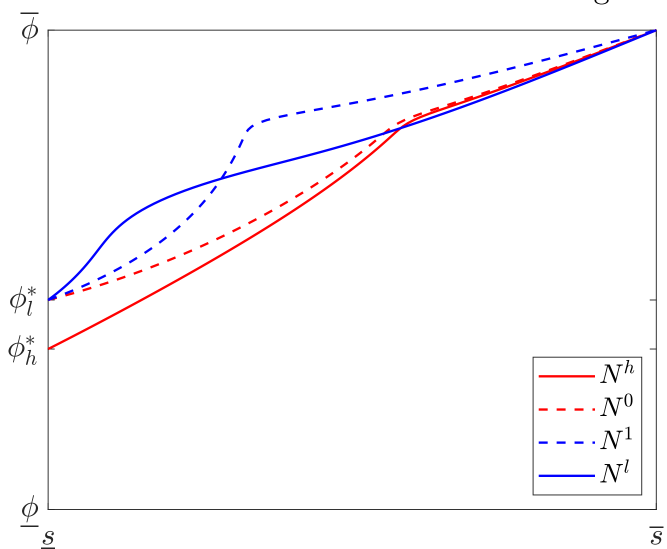
Although the preceding analysis sheds light into the effects of higher trade openness on wage inequality, very few, if any, of the countries in the world operate in autarky. For this reason, in this section I study the effects on wage inequality of a trade liberalization, defined as a decline in the variable trade costs faced by an economy that already participates in international trade. As described in proposition 3, I find that these effects may differ from those described in the previous section. In particular, although a trade liberalization necessarily raises wage inequality among the least-skilled workers in the economy, wage inequality may decline elsewhere in the wage the distribution.

Proposition 3 *Consider a trade liberalization that reduces variable trade costs from τ_h to τ_l, and let $\left\{\phi_h^*, N^h\right\}$ and $\left\{\phi_l^*, N^l\right\}$ represent, respectively, the pre- and post-liberalization activity cutoffs and matching functions. Then, the following conditions hold:*
(i) $\phi_l^* > \phi_h^*$, so a trade liberalization raises wage inequality among the least-skilled workers.
(ii) The selection-into-activity and intensive-margin channels lead to pervasively higher inequality, while the effect of the extensive-margin channel is ambiguous.
(iii) Let $\eta_0^F(t,\lambda) \equiv \frac{F_y(t\lambda)\lambda}{[1+F(t\lambda)k]}$, $\eta_1^F(t,\lambda) \equiv \frac{F_y(t\lambda)\lambda^2}{[1+F(t\lambda)k]}$, and $\bar{t} \equiv \frac{r^{d,a}(\bar{\phi})}{\sigma f^x}$, where $r^{d,a}(\overline{\phi})$ is the autarky revenue function. If the functions η_0^F and η_1^F are, respectively, strictly decreasing and strictly increasing in λ for $\lambda \geq 1$, $k \in (0,n)$ and $t \in (\underline{y}, \bar{t})$, then a trade liberalization raises wage inequality pervasively.

The first result of the proposition states that, as in the Melitz/Channey models, a trade liberalization always leads to the exit of the least productive of firms from the market, $\phi_l^* > \phi_h^*$. The general line of argument used in the proof of proposition 2.i. can be applied here as well. If the activity cutoff remains unchanged after the decline in trade costs, then total wages paid to production and nonproduction workers necessarily increase. With total wages pinned down by condition (19), the activity cutoff must be higher after the liberalization. This result and the continuity of the matching functions imply that $N^l(s) > N^h(s)$ on some interval of the form $[\underline{s}, s')$, which is equivalent to the second part of the claim in proposition 3.i.[31]

[31] Of note, establishing the consequences of an unchanged activity cutoff is more complicated in the case of a trade liberalization, as multiple crossings of relevant matching functions cannot be ruled out. In this case, the formal argument is based on the results in lemma 4.iv-v.

Figure 2: Trade Liberalization and the Matching Function



Note: The solid red and blue lines represent, respectively, the pre- (N^h) and post-liberalization (N^l) matching functions described in Proposition 3. The dashed red (N^0) and dashed blue lines (N^1) depict the matching functions of the ancillary economies described in the text. The effects of the selection-into-activity, intensive-margin, and extensive-margin channels on the matching function are captured, respectively, by the differences between the pairs $\{N^h, N^0\}$, $\{N^0, N^1\}$, and $\{N^1, N^l\}$.

As before, the overall impact of a trade liberalization on wage inequality can be decomposed into the three channels defined earlier. The selection-into-activity channel captures the changes in wage dispersion associated with the rise in the activity cutoff, excluding the impact of changes in the labor demand of incumbent exporters and of changes in the set of exporters. To isolate the effect of this channel, I contrast the matching function of the open economy before the liberalization, N^h, with that of an ancillary open economy, N^0, that differs from the former only in that its activity cutoff is given by that prevailing after the liberalization, ϕ_l^*. That is, as I explain in more detail in the appendix, the BVPs associated with N^h and N^0 can be conceived as parameterization of the general BVP (20), with $K_1 = 0$ and $\alpha^h(\phi) \equiv \left[1 + F\left(r^{d,h}(\phi)\,\tau_h^{1-\sigma}/\sigma f^x\right) n\tau_h^{1-\sigma}\right]$, that differ only in their boundary conditions.[32] Accordingly, the no crossing result in lemma 2.i. implies that N^0 lies strictly above N^h on $[\underline{s}, \overline{s})$ as depicted by the dashed and solid red lines in figure 2. The intuition for the effects of this channel are the same as before—i.e., the exit

[32] $r^{d,h}$ is the domestic revenue function of the open economy with variable trade costs τ_h.

of the least-productive firms from the market reduces the relative demand for less-skilled workers, pushing down their relative wages.

The intensive-margin channel captures the impact on wage inequality of the liberalization-induced rise in the labor demand of incumbent exporters. I isolate this channel by contrasting the matching function N^0 with that of a second ancillary open economy, N^1, with the same set of exporters and active firms, but with variable trade costs given by τ_l. That is, N^1 is obtained by replacing the parameter function $\alpha^h(\phi)$ with $\alpha^1(\phi) \equiv [1 + F\left(r^{d,h}(\phi)\,\tau_h^{1-\sigma}/\sigma f^x\right) n\tau_l^{1-\sigma}]$ in the BVP associated with N^0. As shown by the dashed blue and red lines in figure 2, N^1 necessarily lies above N^0 on $(\underline{s}, \overline{s})$ for the same reasons laid out in the discussion of the extensive-margin channel in proposition 2. If these ancillary economies shared the same wage schedule, then firms of a given productivity level would demand the same worker type in both economies, with the N^1-economy exhibiting a larger labor demand from exporters (lower trade costs). As the (common) fraction of exporters in these economies is increasing in firms's productivity, this additional export-driven labor demand in the N^1-economy results in a higher relative demand for more-skilled workers, which is inconsistent with labor market clearing. Accordingly, the wages of these workers must be higher in the N^1-economy.[33]

The extensive-margin channel captures the impact on relative wages of allowing the fraction of exporters to adjust—i.e., the effects on wages of replacing $\alpha^1(\phi)$ with $[1 + F\left(r^{d,l}(\phi)\,\tau_l^{1-\sigma}/\sigma f^x\right) n\tau_l^{1-\sigma}]$ in the BVP associated with N^1. Little can be said about these effects without making additional assumptions about the primitives of the model. In figure 2, which illustrates only one of the many possibilities, the impact of this channel is given by the difference between N^1 and N^l, the dashed and solid blue lines, respectively. In this example, the weight of some middle-productivity firms among exporters in the post-liberalization economy is larger than in the ancillary N^1-economy. Then, the change in the set of exporters drives up the relative demand for some middle-skill workers, pushing up their wages relative to those of workers with lower and higher skill levels. That said, the impact of this channel could take other forms depending on the CDF of fixed export costs, F, including a pervasive rise and a pervasive decline in wage inequality. Moreover, the effects of this channel can be strong enough to offset the impact of the other two

[33] The result follows from a direct application of lemma 4.i in the appendix, with α^1 taking the role of α^a in the lemma.

channels in some parts of the wage distribution, as shown by the crossing of N^h and N^l in figure 2.

Proposition 3.iii presents a set of sufficient conditions on the CDF of fixed exports costs, F, that guarantee that a trade liberalization always leads to a pervasive rise in wage inequality. When the condition on the function η_1^F is satisfied, reducing variable trade costs while keeping the activity cutoff unchanged in the BVP of the open economy (that allows the set of exporters to change) always leads to pervasively higher wage inequality. In addition, when the condition on η_0^F is satisfied, increasing the activity cutoff while keeping variable trade costs constant in said BVP also leads to a pervasive rise in wage dispersion. Accordingly, when both conditions are met, wage inequality increases pervasively following a liberalization, as the effect on relative wages of changes in the set of exporters (extensive-margin channel) never offsets the combined impact of the selection-into-activity and intensive-margin channels. Although these restrictions on F may appear very restrictive to some readers, one should bear in mind that they are sufficient conditions under all parameterizations of the model.[34]

Regarding the impact of a trade liberalization on the *level of wages*, the analysis and conclusions of the previous section also apply to this case. A liberalization increases real income and average real wages, but the least productive workers in the economy could see their real wage decline in some parameterizations of the model.

5.3 Trade and Wage Dispersion in Other Frameworks

The three-channel decomposition of the effects of higher trade openness on wage inequality described above can be a useful tool to analyze differences in the implications of alternative frameworks in the literature. For illustration purposes, I compare the effects of opening to trade on wage inequality in the no-free-entry model in this paper with those in Helpman, Itskhoki, and Redding (2010), henceforth HIR. In the HIR model, firms screen workers to improve the composition of their labor forces as worker ability is not directly observable. As larger firms have higher returns from screening, they do so more intensively and have workforces of higher average ability than smaller firms. This mechanism generates a wage-

[34] For a Pareto distribution, the condition on η_0^F is always satisfied, while that on η_1^F is satisfied when the shape parameter is small enough. Moroever, a sufficiently small shape parameter typically precludes the crossing of the matching function even when the condition on η_1^F is not satisfied.

size premium, implying that both productivity and exporting positively affect the average wages paid by a firm.

In the HIR model, wage inequality increases after an economy opens to trade only when there is selection into exporting (only some firms export), but is unchanged when all firms become exporters. In terms of the three channels defined earlier, the selection-into-activity channel is not operational in the HIR model, as changes in the activity cutoff do not modify the relative size of firms. In addition, the extensive-margin channel affects wage inequality only when it changes the relative size of firms in the economy— i.e., only when some but not all firms export. In contrast, trade always leads to higher wage inequality in the no-free-entry model of this paper. Although trade may not affect wage inequality through the extensive-margin channel if all firms export (as in HIR), it always drives up wage dispersion through the selection–into-activity channel.

6 The Free-Entry Model

In the model outlined above, the mass of firms in the industry is fixed at an exogenous level. Although this assumption may be a good approximation to the firm-entry dynamics in the short-run, it does not capture the change in the number of firms through endogenous entry and exit over time. In this section, I relax this assumption by allowing firms to enter the industry for a cost, making the mass of firms in the industry, $\overline{M}$, an additional endogenous variable. Specifically, I assume that there is an unbounded pool of prospective firms that can enter the industry by incurring a fixed entry cost of $f^e V(s)$ units of each skill $s \in S$. Accordingly, the aggregate expenditure on entry costs is $\overline{M} f^e$ when a mass $\overline{M}$ of firms enters the industry. Upon entry, firms obtain their productivity as independent draws from the distribution G, as explained in section 2.2. All the other primitives of the model remain unchanged.

The new assumptions above do not affect the basic structure of the model described in section 2, so equations (1)-(5) continue to hold. Conditional on the mass of firms, $\overline{M}$, the equilibrium analysis in section 4 applies almost unchanged to the free-entry model, with the caveat that equilibrium conditions now reflect the labor demand derived from the presence of fixed entry costs—i.e., L must be replaced with $L - f^e\overline{M}$ throughout the analysis. The new free-entry assumption implies that, in equilibrium, prospective entrants must be indifferent between entering and not entering the industry. Accordingly, expected profits from entering the industry must equal the cost of entry, $[1 - G(\phi^*)]\left[\overline{\pi}^d + \overline{\pi}^x\right] = f^e$, where $\overline{\pi}^d$ and $\overline{\pi}^x$ are, respectively, the average domestic and export profits among active firms.[35] Per the optimal pricing rule, this *free-entry* condition can be written as

$$\int_{\phi^*}^{\overline{\phi}} \left[\frac{r^d(\phi)}{\sigma} - f\right] g(\phi)\, d\phi + \int_{\phi^*}^{\overline{\phi}} \int_0^{\frac{r^d(\phi)\tau^{1-\sigma}}{\sigma f^x}} n\left[\frac{r^d(\phi)\,\tau^{1-\sigma}}{\sigma} - f^x y\right] dF(y)\, g(\phi)\, d\phi = f^e. \tag{22}$$

The last equation completes the description of the open-economy equilibrium in the free-entry model, prompting a definition analogous to that in definition 1.

The free-entry equilibrium of the open economy is subject to a characterization analogous to that given in section 4.1 for the no-free-entry model. In particular, given the activity cutoff, ϕ^*, the price, domestic-revenue and inverse-matching functions, $\left\{p, r^d, H\right\}$, solve a BVP that differs from that of the no-free-entry model in lemma 3.iii. only in that L is replaced by $L - f^e\overline{M}$ in the equation defining the slope of the inverse-matching function. Moreover, the discussion in section 4.2 implies that conditional on ϕ^*, the BVPs of the no-free-entry and free-entry models have the same parameterization in terms of the general BVP (20), so they share the same solution functions r^d and H. The equilibrium value for ϕ^* is pinned down by the free entry condition (22).[36]

The observations above have important implications. First, all the conclusions reached in section 4.2 about the dependence of $\left\{r^d, H\right\}$ on the activity cutoff ϕ^* continue to hold in the free-entry model. Accordingly, many results, such as the existence and uniqueness of the equilibrium in the free-entry model, can be derived in a similar way.[37] Second, the only relevant difference between the no-free-entry and free-entry models regarding the determination of the equilibrium matching function is given by the equations that pins down the activity cutoff in these models, equations (19) and (22), respectively. In the remainder of this section I explore how this difference affects the impact of increased trade

[35] Note that $\overline{\pi}^x$ is not the average export profits among exporters, but among all active firms.

[36] This is the case because ϕ^* and r^d are the only endogenous variables appearing in equation (22). Note that using the analog of equation (19) for the free-entry model to determine the activity cutoff ϕ^* would only give us ϕ^* as a function of the endogenous mass of firms $\overline{M}$.

[37] As $r^d(\phi)$ depends negatively on the activity cutoff, the left-hand side of equation (22) is strictly decreasing in ϕ^*, implying that there is unique free-entry equilibrium if entry costs are not too high.

openness on wage inequality.

6.1 Autarky vs. Trade in the Free-entry Model

Unlike the case of the no-free-entry model, an increase in trade openes may lead to a rise or fall in the activity cutoff in the free-entry model, with ambiguous effects on the wage distribution through the selection-into-activity channel. As formally stated in proposition 4, this additional source of ambiguity in the free-entry model implies that opening to trade can lead a pervasive rise in wage inequality or a wage polarization.

Proposition 4 *Let $\{\phi_a^*, N^a\}$ and $\{\phi_\tau^*, N^\tau\}$ be the activity cutoffs and matching functions corresponding to the free-entry equilibrium of the closed and open economies, respectively. Then ϕ_τ^* could be lower or higher than ϕ_a^* depending on the model's parameters.*
(i) If $\phi_\tau^* \geq \phi_a^*$, then $N^\tau(s) > N^a(s)$ on $s \in (\underline{s}, \overline{s})$, so opening to trade leads to pervasively higher wage inequality. The selection-into-activity channel leads to a pervasive rise (no change) in wage inequality if $\phi_\tau^* > (=) \phi_a^*$. The extensive-margin channel always leads to a pervasive rise in wage inequality.
(ii) If $\phi_\tau^* < \phi_a^*$, then $N^\tau(s)$ and $N^a(s)$ intersect exactly once on $(\underline{s}, \overline{s})$, so opening to trade leads to wage polarization. The selection-into-activity and extensive-margin channels lead, respectively, to pervasively lower and pervasively higher wage inequality.

I start the discussion of propostition 4 by analyzing why opening to trade may lead to a decline in the activity cuoff in the free-entry model. As this theoretical possibility is not present in the no-free-entry model in this paper nor in standard free-entry models with homogeneous workers, such as Melitz (2003), I discuss the differences between these two frameworks and the free-entry model in this paper that allow for this additional possibility in the latter.

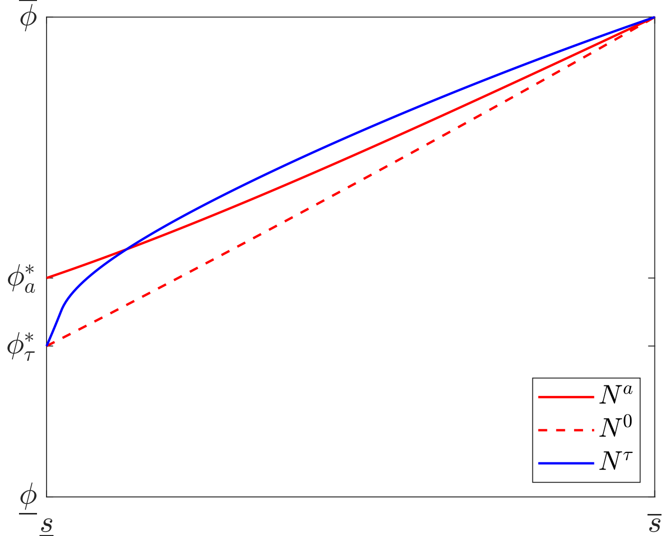
Opening to trade may have different qualitative effects on the activity cutoff in the no-free-entry and free-entry models of this paper, reflecting the different equilibrium conditions that determine this cutoff in these models. These differences are better understood by comparing the impact that trade has on these equilibrium conditions when the set of active firms and the revenue of the least-productive ones are assumed to remain unchanged, $r^d(\phi_a^*) = \sigma f$. As discussed in section 5, in this scenario, trade leads to a rise in the implied total wages paid to production and nonproduction workers, as total firms' revenue and fixed export costs increase. Accordingly, equation (19) implies that a higher activity cutoff is required in the open economy of the no-free-entry model. In contrast, in the free-entry model, total firms' revenue and fixed export costs enter with opposite signs on the left-hand side of the free-entry condition (22), with an ambiguous net effect, so a lower activity cutoff may be required in the open economy.

Relative to standard free-entry models with homogeneous workers, a trade-induced decline in the activity cutoff is possible in the free-entry model because of the endogenous changes in the matching of heterogeneous workers to firms.[38] As before, it is instructive to compare the impact that trade has on the free-entry condition in these models when the set of active firms and the revenue of the least-productive ones are assumed to remain unchanged. In such a scenario, trade increases export profits from zero (in autarky) to some strictly positive number in both models. With domestic profits remaining unchanged in the homogeneous-workers model (before adjusting the activity cutoff), average/expected profits necessarily increase, so the free-entry condition requires a higher activity cutoff in the open economy. In contrast, in the free-entry model of this paper, trade may lead to a decline in aggregate profits due to changes in the matching function. Specifically, as the matching function N shifts up (H shifts down) in the scenario considered, domestic revenues and profits decline for firms with productivity above ϕ_a^*. For some parameter values, the decline in aggregate domestic profits more than offsets the rise in export profits, so the free-entry condition (22) requires a lower activity cutoff in the open economy.

Per proposition 4, conditional on its impact on the activity cutoff, trade has a unique qualitative effect on the dispersion of wages, with an unambiguous effect through the selection-into-activity and extensive-margin channels. The case in proposition 4.i, $\phi_\tau^* \geq \phi_a^*$, is essentially the same situation considered in section 5.1 for the no-free-entry model. If $\phi_\tau^* > \phi_a^*$, then the situation is identical to that depicted in figure 1, so the corresponding analysis applies here as well. When $\phi_\tau^* = \phi_a^*$, the only difference is that the selection-into-activity channel has no effect on wage dispersion.

[38] The stochastic modeling of fixed costs is another difference between the free-entry model in this paper and standard Melitz-type models. However, said difference alone cannot produce a trade-induced declined in the activity cutoff.

Figure 3: Opening toTrade and the Matching Function in Free-entry Model



Note: The solid red and blue lines represent, respectively, the matching functions of the closed (N^a) and open (N^τ) economies. The dashed red line depicts the matching function of the ancillary autarkic economy (N^0) described in the text. The differences between N^a and N^0 and between N^0 and N^τ capture the impact of the selection-into-activity and extensive-margin channels, respectively. The figure depicts the case in which trade induces a decline in the activity cutoff.

The case in proposition 4.ii, $\phi_\tau^* < \phi_a^*$, requires some additional explanation. As I discuss in the appendix, the matching function of the open economy, N^τ, cannot remain completely below that of the closed economy, N^a, on $[\underline{s}, \overline{s})$. Otherwise, per lemma 2.ii, expected domestic profits in the open economy would be strictly higher than in autarky, implying a violation of the free-entry condition (22). Then, N^τ and N^a must intersect at least once on $(\underline{s}, \overline{s})$. Moreover, adapting the analysis of the extensive-margin channel in section 5.1 to assess the relative position of N^a and N^τ to the right of the first intersection, it can be shown that N^a must remain below N^τ there, so the matching functions must intersect exactly once on $(\underline{s}, \overline{s})$.[39] The situation is depicted in figure 3, where the solid red and blue lines represent N^a and N^τ, respectively. As before, the dashed red line is the matching function of an ancillary autarkic economy, N^0, that is obtained by changing

[39] Formally, to the right of the first intersection point, the matching functions of the closed and open economies can be conceived as solutions to particular parameterizations of the general BVP (20) with $K_1 = 0$ that differ only in the parameter function $\alpha(\phi)$, which is constant in the former and increasing in the latter. The result then follows from a direct application of lemma 4.i in the appendix.

the activity cutoff in the BVP corresponding to N^a from ϕ_a^* to ϕ_τ^*. As discussed in section 5.1, the effects of trade on wage inequality through the selection-into-activity and extensive-margin channels are captured, respectively, by the difference between the pairs $\{N^a, N^0\}$ and $\{N^0, N^\tau\}$. While the selection-into-activity channel pervasively reduces wage inequality, the extensive-margin channel pervasively increases it, with the former channel dominating to the left of the interior intersection point of N^a and N^τ, and the latter dominating to the right. As a result, workers with skill level corresponding to this (interior) intersection point see their wages decline relative to those of all other workers—i.e., opening to trade leads to wage polarization.

Turning to the effects of trade on the *level of real wages*, the results obtained for the no-free-entry model generally go through. First, the average real wage is always higher in the open economy. As before, the result follows from the (constrained) efficiency of the equilibrium. Second, opening to trade may induce a decline in the real wage of the least-skilled workers in the economy, although in the free-entry model this possibility is fully determined by the impact of trade on the activity cutoff. As the free-entry condition implies that the economy's total income and expenditure is given by total labor income, $E = \overline{w}L$, rearranging equation (21) yields $w^i(\underline{s})/P^i = \frac{(\sigma-1)}{\sigma} A(\underline{s}, \phi_i^*) [L/\sigma f]^{\frac{1}{\sigma-1}}$ for $i = a, \tau$, so trade rises the real wage of even the least-skilled workers in the economy if and only if it rises the activity cutoff. Note that this observation, together with proposition 4, implies that opening to international trade raises the real wage of the poorest workers in the economy only if it also induces a pervasive rise in wage inequality.

6.2 Trade Liberalization in the Free-Entry Model

The effects of a trade liberalization on the wage distribution in the free-entry-model can be derived by resorting to the results in propositions 2 to 4, as they largely cover the range of possible outcomes in this case. For the same reasons behind the corresponding result in proposition 4, a trade liberalization could lead to a rise or a fall in the activity cutoff. If the activity cutoff increases, then the situation is identical to that considered in proposition 3. If the activity cutoff declines, then the pre- and post-liberalization matching functions must intersect at least once on $(\underline{s}, \overline{s})$ to avoid a violation of the free entry condition as discussed in the case of proposition 4.ii. However, in the case of a trade liberalization, more than one crossing on $(\underline{s}, \overline{s})$ cannot be ruled out even when the conditions on the functions $\eta_0^F(t, \lambda)$ and $\eta_1^F(t, \lambda)$ in proposition 3 are satisfied.

7 Empirically Relevant Distributional Effects of Trade

The analysis of the previous sections shows that an increase in trade openness generally has ambiguous theoretical implications for the wage distribution. The goal of this section is to explore which of the theoretical possibilities described in that analysis are the most empirically relevant. To that end, I calibrate the primitives of the framework based on estimates from the literature and some broad features of firm data from Portugal for 2006. As much of the theoretical ambiguity is driven by the extensive-margin channel, a crucial target of the calibration is the fraction of firms that export in each decile of the empirical distribution of firms by value added per worker.[40]

7.1 Data and Calibration

I compute all empirical moments targeted in the calibration from a summary of the dataset constructed in **?**), which in turn draws from annual information on Portuguese firms reported under the *Informação Empresarial Simplificada*. For each decile of manufacturing firms in terms of value added per worker, this summary includes information on total employment, total labor costs, average wages, the share of firms that are exporters, and average value added per worker across firms.

Following Melitz and Redding (2015), I set the elasticity of substitution between final goods to four, choose variable trade costs to match the average exports-to-sales ratio among Portuguese firms, and make assumptions that guarantee that firms' revenue in the model, r^d, is distributed Pareto with shape parameter equal to one.[41] In particular, I assume that firm productivity has a truncated Pareto distribution and that the productivity function takes the form $A(s, \phi) = B_0^A[\alpha_s s^\rho + \alpha_\phi \phi^\rho]^{B_1^A/\rho}$, where $A(s, \phi)$ is homogeneous of degree $B_1^A > 0$, $\rho < 0$ is the constant elasticity of substitution between worker skill and

[40] This section presents a brief summary of my calibration approach and main results. For more details, see the corresponding section of the extended version of the paper in my personal website.

[41] Setting $n\tau^{1-\sigma}/(1 + n\tau^{1-\sigma}) = 0.31$ yields a value for $n\tau^{1-\sigma}$. All relevant calibrated variables and the counterfactual exercises discussed later depend on $\{n, \tau\}$ only through $n\tau^{1-\sigma}$ (see extended paper).

firm productivity, and the positive parameters $\{B_0^A, \alpha_s, \alpha_\phi\}$ are overall and input-specific productivity shifters.[42] In addition, I make assumptions such that the endogenous fraction of exporters, $FX(\phi) \equiv F(r^d(\phi)\,\tau^{1-\sigma}/\sigma f^x)$, is a power function of firm productivity in the calibrated equilibrium. These restrictions on $\{G(\phi),\, r^d,\, A(s,\phi),\, FX(\phi)\}$, together with the model's equilibrium conditions, allow me to back out the implied functional forms of all remaining endogenous and exogenous elements of the model, including the CDF of fixed export costs. As a result, I can compute the model's implications for several moments of the data described above.

In the calibration of the model, I target the distribution of (i) total employment and (ii) the total wage bill across deciles of value added per worker, as well as (iii) the fraction of firms that export and (iv) the average value added per worker in each decile. It can be shown that the model-implied values for these moments depend only on a subset of parameters of $\{G(\phi),\, A(s,\phi),\, FX(\phi)\}$, which also completely determines the wage distribution in the calibrated equilibrium. Accordingly, after all remaining parameters are normalized or chosen to satisfy equilibrium conditions of the model, I estimate this subset of parameters via simulated methods of moments, targeting moments (i)-(iv) in the Portuguese data.[43]

Despite being highly stylized, the model does a good job at fitting the targeted moments (i)-(iv) in the Portuguese data as shown in figure 4. In particular, the calibrated model fits particularly well the fraction of firms that export in each decile of the distribution of firms' value added per worker (panel c), a crucial target of the calibration. This moment plays a major role in pinning down the CDF of the firm-specific component of fixed export costs, the primitive of the model controlling the extensive-margin channel. As this margin drives much of the theoretical ambiguity regarding the distributional effects of higher trade openness, it is especially important that the calibrated model fits well this moment of the data. Figure 6 and table 1 in the appendix show that the calibrated model also fits relatively well untargeted moments in the Portuguese data.

Given my calibration approach, the elasticity of substitution between worker skill and firm productivity in the productivity function, ρ, does not affect wage inequality in the calibrated equilibrium. However, ρ does affect the distributional effects of increased trade

[42] The assumption $\rho < 0$ guarantees that $A(s,\phi)$ is strictly log-supermodular.

[43] For more details on the calibration, see the extended appendix in my personal website.

Figure 4: Model vs. Data

Note: For the moments targeted in the calibration of the model, the figure shows the model's prediction (line) and the target values computed from the Portuguese data for 2006 described in the text (bars).

openness in the model. In particular, when s and ϕ are hard to substitute (lower values of ρ), a given change in trade costs is associated with less labor reallocation across firms and larger changes in relative wages. As such, I explore the implications of the model for different values of ρ, with the baseline results discussed in this section corresponding to $\rho = -10$. For this value of ρ, the largest liberalization I consider—which is significantly larger than liberalizations typically featured in the literature—induces a change in the Gini index of about ten points, which is somewhat higher than the six-point range of variation of Portugal's Gini index over the last two decades.[44] The main messages go

[44] World Bank estimates of Portugal's Gini index, which start in 2003, show a maximum value of 38.9

through for $\rho = -5, -15$.[45]

7.2 Revisiting the Distributional Effects of Trade

Armed with calibrated parameter values, I revisit the distributional effects of higher trade openness implied by the framework. Under the weak assumptions of section 5, *a trade liberalization in the no-free-entry model* has ambiguous distributional effects, with the ambiguity driven by extensive margin channel. As discussed in section 6, assuming *free entry* brings an additional source of theoretical ambiguity through the intensive-margin channel, as the activity cutoff may rise or fall following an increase in trade openness. In contrast, a decline in trade costs always raises wage inequality pervasively in the calibrated framework under both entry assumptions. Indeed, the calibrated CDF of fixed exports costs, $F(y)$, satisfies the sufficient condition in proposition 3.iii and the activity cutoff always rises in the calibrated free-entry model.[46] As such, an increase in trade openness always raises wage inequality pervasively, regardless of its magnitude, initial level of trade costs, or the calibrated values of other primitives (including ρ).

To gain further insight on the implications of the calibrated model, I quantify the effects of trade-costs declines on overall wage inequality in the no-free-entry model through each of the three channels defined in section 5—selection-into-activity, intensive-margin, and extensive-margin channels. Panel (a) of figure 5 shows the *incremental* change in the Gini index (black dots) and the contribution of each of these channels (stacked bars) as variable trade costs are incrementally reduced by the same proportion $\widehat{\tau}_{step} \equiv \frac{\tau_{post}}{\tau_{pre}}$, where τ_{pre} and τ_{post} are, respectively, the level of trade costs prevailing before and after the liberalization.[47] For example, the height of the first black dot in the chart captures the change in the Gini index induced by a decline in trade costs from their value in the

in 2004 at 38.9 and a minimum value of 32.8 in 2019.

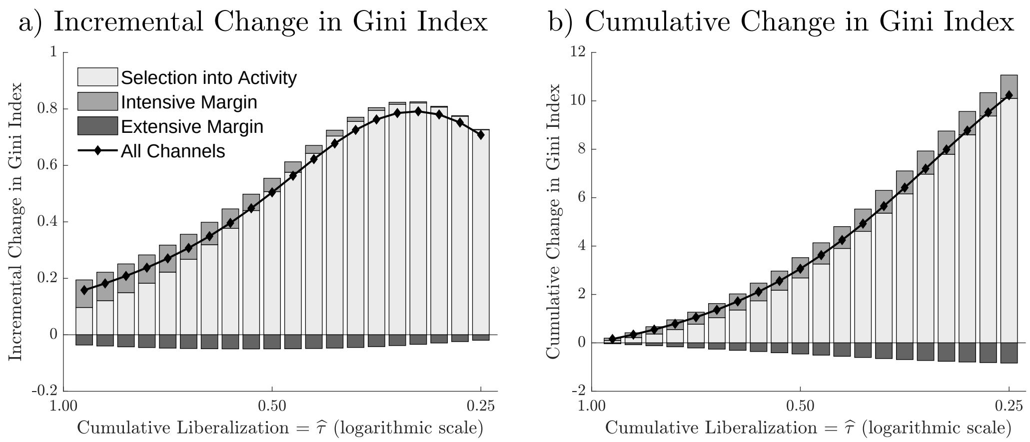
[45] Although the effects of higher trade openness on wage inequality through each of the channels defined earlier are magnified for lower values of ρ, the *relative* quantitative importance of each of the channels is largely unchanged. As such, the conclusions about the most likely qualitative effects of trade on wage inequality are also unchanged. See appendix D of extended paper.

[46] Calibrated fixed export costs have a truncated Pareto distribution with shape parameter $\nu \approx 0.24$ and support $S_F \approx [0.25, 3043]$.

[47] Specifically, trade costs decline by about 7 perent in each liberalization, $\widehat{\tau}_{step} \approx 0.93$, which follows from dividing the maximum cummulative decline in trade costs considred in the chart (75 percent) into 20 liberalization steps of the same size, $0.25 = (\widehat{\tau}_{step})^{20}$.

calibrated equilibrium, τ_0, to $\tau_1 = \widehat{\tau}_{step}\tau_0$. Similarly, the height of the second one indicates the additional change in the Gini index as trade costs further decline to $\tau_2 = \widehat{\tau}_{step}\tau_1$. The horizontal axis of the chart indicates the cumulative decline in trade costs after k sequential liberalizations, $\widehat{\tau} = [\widehat{\tau}_{step}]^k$. Panel (b) shows the cumulative change in the Gini index—i.e., the values in panel (b) are the cumulative sum of those in panel (a).

Figure 5: Trade Liberalization in the No-Free-Entry Model



Note: Panel (a) shows the incremental change in the Gini index (black dots) as variable trade costs are incrementally reduced by the same proportion $\widehat{\tau}_{step} \approx 0.93$. These changes are decomposed into the contributions of the selection-into-activity, intensive-margin and extensive-margin channels (stacked bars). The horizontal axis indicates the cumulative decline in trade costs after k sequential liberalizations. Panel (b) shows the corresponding cumulative changes in the index and cumulative contributions of each channel.

A few lessons follow from figure 5 relating to the *relative* quantitative importance of each of the three channels in affecting inequality. Notably, the contribution of the extensive-margin channel (always negative in the figure) is dwarfed by the combined (positive) contributions of the of the other two channels. In addition, the selection-into-activity channel increasingly dominates as trade costs decline. Specifically, while the contribution of the intensive margin gradually declines until vanishing, that of the selection-into-activity channel increases for the most part, remaining significant in the range of cumulative trade costs declined considered in the figure. The general picture painted by figure 5 for the case of the Gini index also holds for other measures of wage inequality, as indicated by figure 7 of the appendix for the cases of the 90/10, 90/50 and

50/10 ratios. The free-entry model yields similar results.

The results of this section suggest that a decline in trade costs is likely to lead to pervasively higher wage inequality, in both the short and long run, through the labor-reallocation mechanisms emphasized in this paper. That said, the calibration exercise also shows that *a decline in trade costs always raises the real wage of all workers.*

I conclude this section by discussing the effects of changing the specification of fixed export costs in the calibration above to one in which all firms face the same cost, a standard assumption since Melitz (2003). For this alternative specification, figure 8 in the appendix repeats the decomposition analysis of figures 5 and 7 for the baseline model. These figures reveal major differences in the contribution of the extensive-margin channel across these two specifications. Notably, with common fixed export costs across firms, this channel exerts a much stronger downward pressure on wage inequality for initial liberalizations (before all firms export). Indeed, in some cases, this channel is strong enough to induce a decline in inequality among more-skilled workers (crossing of matching functions), leading to slight declines in the 90/50 ratio. In contrast, this channel has a much less significant role in the baseline specification of the model, so a decline in trade costs always leads to a pervasive rise in wage inequality. This result further stresses the importance of carefully quantifying the extensive-margin channel.

8 Conclusion

This paper develops a framework for studying the effects of higher trade openness on the wage distribution in which strong skill-productivity complementarities in production imply that inequality rises as workers reallocate towards more productive (skill-intensive) firms in the same industry. The model features a large number of skill groups and can accommodate weaker and more empirically relevant restrictions on firm selection into exporting than standard heterogenous-firms models. The cross-sectional structure of the model captures several features of the data identified by the trade and labor literatures.

I shed light into the distributional effects of increased trade openness by decomposing them into those associated with the selection-into-activity, intensive-margin, and extensive-margin channels of trade. Although an autarkic economy that opens to trade always experiences a pervasive rise in wage inequality under no firm entry, more theoretical outcomes are possible following a trade liberalization, with the ambiguity driven by the extensive-margin channel. Assuming *free entry* brings an additional source of theoretical ambiguity through the intensive-margin channel, as the activity cutoff may rise or fall following an increase in trade openness. In a calibrated version of the framework, an increase in trade openness always leads to pervasively higher wage inequality under both entry assumptions, as the extensive-margin channel has a small quantitative role and the activity cutoff always rises. The analysis highlights the importance of properly accounting for the role of new exporters (extensive margin) in shaping the aggregate relative demand for skills, which in the framework is controlled by the specification of fixed export costs.

Finally, this paper contribute methodologically to the analysis of assignment problems. In addition to presenting existence and uniqueness results for a general BVP that encompasses those in this paper and others in the literature, I derive results about the dependence of the solution on the parameters of the problem. These results can be used to analyze comparative statics exercises beyond those considered in this paper.

References

Acemoglu, D. (2002). Technical Change, Inequality, and the Labor Market. *Journal of Economic Literature 40*(1), 7–72.

Agarwal, R. and D. O'Regan (2008a). *An Introduction to Ordinary Differential Equations*. Universitext. Springer New York.

Agarwal, R. and D. O'Regan (2008b). *Ordinary and Partial Differential Equations: With Special Functions, Fourier Series, and Boundary Value Problems*. Universitext. Springer New York.

Amiti, M. and D. R. Davis (2012). Trade, Firms, and Wages: Theory and Evidence. *The Review of Economic Studies 79*(1), 1–36.

Antràs, P., L. Garicano, and E. Rossi-Hansberg (2006). Offshoring in a Knowledge Economy. *The Quarterly Journal of Economics 121*(1), 31–77.

Atkinson, A. B. (1970). On the Measurement of Inequality. *Journal of Economic Theory 2*(3), 244–263.

Attanasio, O., P. K. Goldberg, and N. Pavcnik (2004). Trade Reforms and Wage Inequality in Colombia. *Journal of Development Economics 74*(2), 331–366.

Autor, D. H., L. F. Katz, and M. S. Kearney (2008). Trends in U.S. Wage Inequality: Revising the Revisionists. *The Review of Economics and Statistics 90*(2), 300–323.

Bailey, P., L. Shampine, and P. Waltman (1968). *Nonlinear Two Point Boundary Value Problems*. Mathematics in science and engineering. Academic Press.

Bernard, A., J. Eaton, J. Jensen, and S. Kortum (2003). Plants and Productivity in International Trade. *American Economic Review 93*(4), 1268–1290.

Bernard, A. and J. Jensen (1995). Exporters, Jobs, and Wages in U.S. Manufacturing: 1976-1987. *Brookings Papers on Economic Activity 26*, 67–119.

Bernfeld, S. and V. Lakshmikantham (1974). *An Introduction to Nonlinear Boundary Value Problems*. Mathematics in science and engineering : a series of monographs and textbooks. Academic Press.

Burstein, A. and J. Vogel (2017). International trade, technology, and the skill premium. *Journal of Political Economy 125*(5), 1356–1412.

Bustos, P. (2011). Trade Liberalization, Exports, and Technology Upgrading: Evidence on the Impact of MERCOSUR on Argentinean Firms. *American Economic Reveiw 101*(1), 304–340.

Card, D., A. R. Cardoso, J. Heining, and P. Kline (2016). Firms and Labor Market Inequality: Evidence and Some Theory. Working Paper 22850, National Bureau of Economic Research.

Chaney, T. (2008). Distorted Gravity: The Intensive and Extensive Margins of International Trade. *American Economic Review 98*(4), 1707–21.

Costinot, A. (2009). An Elementary Theory of Comparative Advantage. *Econometrica 77*(4), 1165–1192.

Costinot, A. and J. Vogel (2010). Matching and Inequality in the World Economy. *Journal of Political Economy 118*(4), 747–786.

Davis, D. R. and J. Harrigan (2011). Good Jobs, Bad Jobs, and Trade Liberalization. *Journal of International Economics 84*(1), 26–36.

Davis, S. J. and J. Haltiwanger (1991). Wage Dispersion between and within U.S. Manufacturing Plants, 1963-86. *Brookings Papers on Economic Activity 22*(1991 Micr), 115–200.

Eaton, J. and S. Kortum (2002). Technology, geography, and trade. *Econometrica 70*(5), 1741–779.

Egger, H. and U. Kreickemeier (2009). Firm Heterogeneity and the Labor Market Effects of Trade Liberalization. *International Economic Review 50*(1), 187–216.

Egger, H. and U. Kreickemeier (2012). Fairness, Trade, and Inequality. *Journal of International Economics 86*(2), 184–196.

Goldberg, P. K. and N. Pavcnik (2007). Distributional Effects of Globalization in Developing Countries. *Journal of Economic Literature 45*(1), 39–82.

Grossman, G. M., E. Helpman, and P. Kircher (2017). Matching, Sorting, and the Distributional Effects of International Trade. *Journal of Political Economy 125*(1), 224–264.

Grossman, G. M. and G. Maggi (2000). Diversity and Trade. *American Economic Review 90*(5), 1255–1275.

Helpman, E. (2016). Globalization and Wage Inequality. Working Paper 22944, National Bureau of Economic Research.

Helpman, E., O. Itskhoki, M.-A. Muendler, and S. J. Redding (2016). Trade and Inequality: From Theory to Estimation. *The Review of Economic Studies 84*(1), 357–405.

Helpman, E., O. Itskhoki, and S. Redding (2010). Inequality and Unemployment in a Global Economy. *Econometrica 78*(4), 1239–1283.

Kiguradze, I. (1988). Boundary-value Problems for Systems of Ordinary Differential Equations. *Journal of Soviet Mathematics 43*, 2259–2339.

Krishna, P., J. P. Poole, and M. Z. Senses (2014). Wage Effects of Trade Reform with Endogenous Worker Mobility. *Journal of International Economics 93*(2), 239–252.

Melitz, M. (2003). The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. *Econometrica 71*(6), 1695–1725.

Melitz, M. and S. J. Redding (2015). New trade models, new welfare implications. *American Economic Review 105*(3), 1105–46.

Monte, F. (2011). Skill Bias, Trade, and Wage Dispersion. *Journal of International Economics 83*(2), 202–218.

Ohnsorge, F. and D. Trefler (2007). Sorting It Out: International Trade with Heterogeneous Workers. *Journal of Political Economy 115*(5), 868–892.

O'Regan, D. (2013). *Existence Theory for Nonlinear Ordinary Differential Equations.* Mathematics and Its Applications. Springer Netherlands.

Sampson, T. (2014). Selection into Trade and Wage Inequality. *American Economic Journal: Microeconomics 6*(3), 157–202.

Somale, M. (2015). *Essays in Technology and International Trade*. Ph. D. thesis, Princeton University.

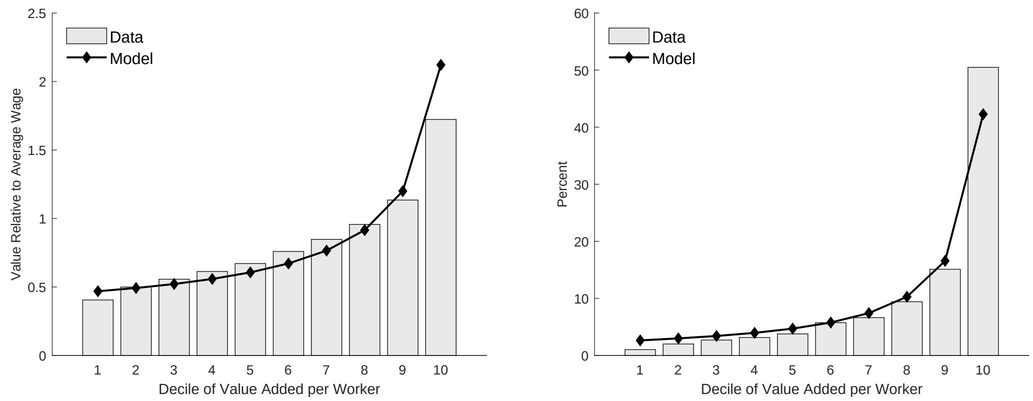
Stakgold, I. (1998). *Green's Functions and Boundary Value Problems.* Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts. Wiley.

Yeaple, S. R. (2005). A Simple Model of Firm Heterogeneity, International Trade, and Wages. *Journal of International Economics 65*(1), 1–20.

A Additional Figures and Tables

Figure 6: Model vs. Data, Nontargeted Moments

a) Average Wage   b) Distribution of Total Value Added



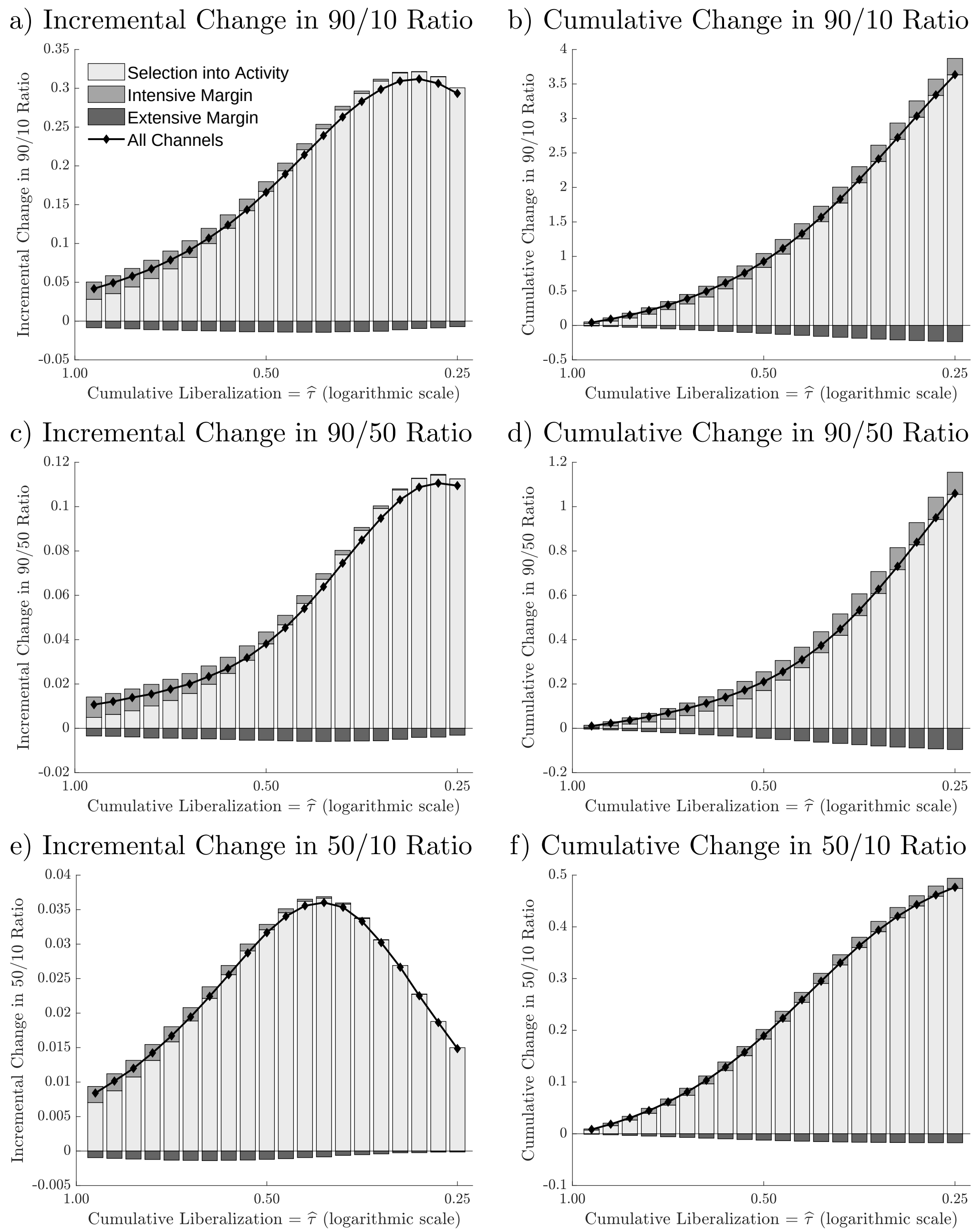
Note: For two untargeted moments in the calibration, the figure compares the model's predictions (line) against their values in the Portuguese data for 2006 (bars) described in the text.

Table 1: Measures of Wage Inequality: Model vs. Data

	Model	Data	
		2005	2007
Gini Index	34	36	35
90/10 Ratio	4.75	4.06	3.97
90/50 Ratio	2.80	2.64	2.59
50/10 Ratio	1.69	1.53	1.53

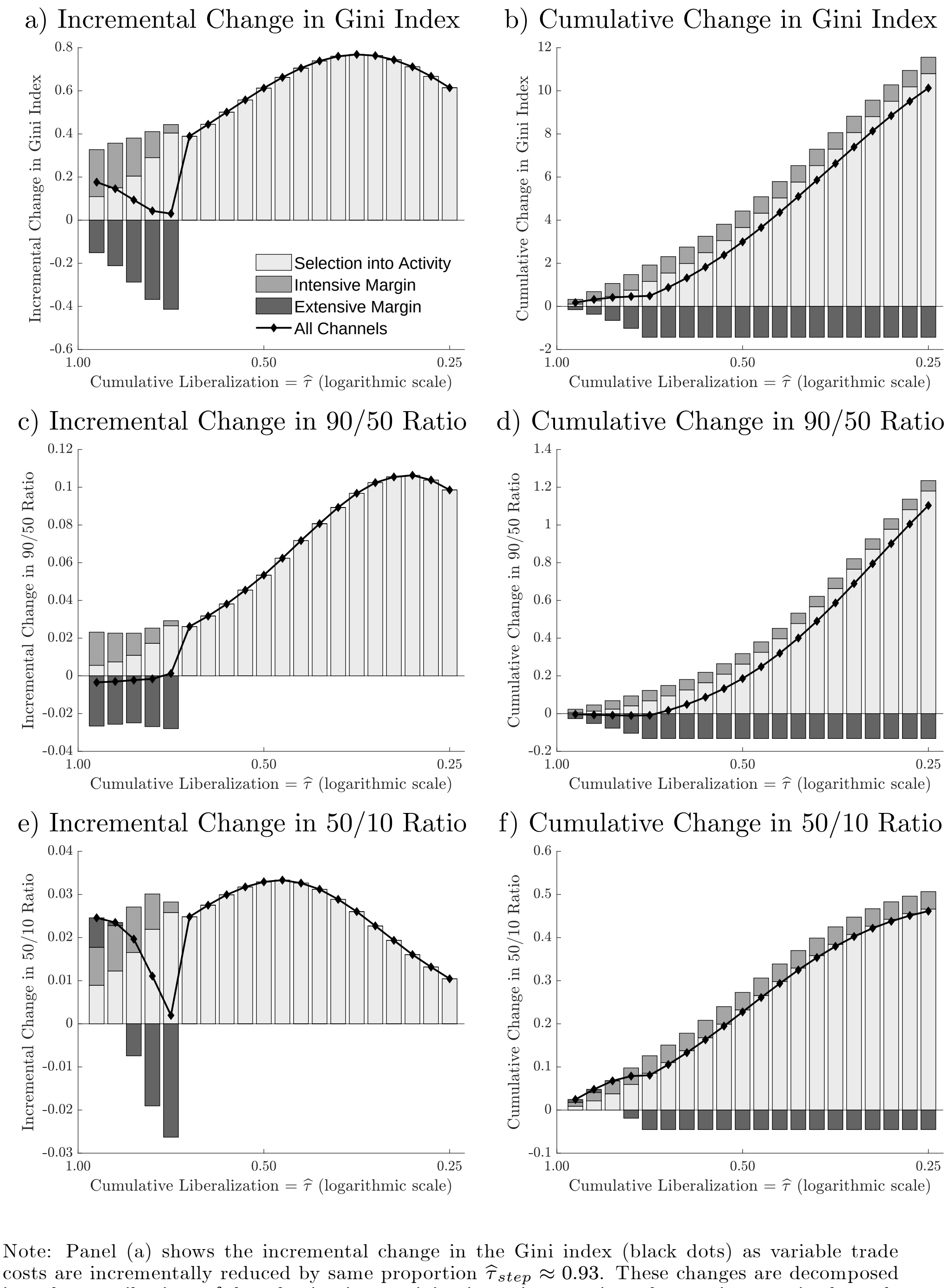
Note: The values in the first column correspond to the calibration of the model discussed in the text. Those in the second and third columns are taken from table 1 in Pereira (2021), and are based on wage data from the Portuguese dataset "Quadros de Pessoal" for the years 2005 and 2007. Values for 2006 are not reported in Pereira (2021). None of these inequality measures was targeted in the calibration.

Figure 7: Trade Liberalization in the No-Free-Entry Model, Countinued



Note: Panel (a) shows the incremental change in the 90/10 ratio (black dots) as variable trade costs are incrementally reduced by the same proportion $\hat{\tau}_{step} \approx 0.93$. These changes are decomposed into the contributions of the selection-into-activity, intensive-margin and extensive-margin channels (stacked bars). Panel (b) shows the corresponding cumulative changes in the ratio and cumulative contributions of each channel. The rest of the panels show similar calculations for the 90/50 and 50/10 ratios.

Figure 8: Trade Liberalization, Common Fixed Export Costs Across Firms



Note: Panel (a) shows the incremental change in the Gini index (black dots) as variable trade costs are incrementally reduced by same proportion $\hat{\tau}_{step} \approx 0.93$. These changes are decomposed into the contributions of the selection-into-activity, intensive-margin and extensive-margin channels (stacked bars). Panel (b) shows the corresponding cumulative changes in the ratio and cumulative contributions of each channel. The rest of the panels show similar calculations for the 90/50 and the 50/10 ratios.

B Theoretical Appendix

B.1 Section 3

B.1.1 Proof of Lemma 1

Existence of an strictly increasing matching function N. This result was initially proved in Costinot and Vogel (2010) and later extended to the case of monopolistic competition in final goods in Somale (2015) and Sampson (2014). For a step-by-step proof of the result see Somale (2015) or the extended appendix in my personal website, with the latter also including step-by-step versions of the proofs in this appendix.[48]

Conditions i-iii. Consider a no-free-entry equilibrium of the closed economy with activity cutoff ϕ^*, wage schedule $w(s)$, price function $p(\phi)$, domestic revenue function $r^d(\phi)$ and matching function $N(s)$. The cost minimization condition (4) and the existence of the matching function N imply that $s = \arg\min_z w(z)/A(z, N(s))$, so $\frac{w(s)}{A(s,N(s))} \leq \frac{w(s+ds)}{A(s+ds,N(s))}$ and $\frac{w(s+ds)}{A(s+ds,N(s+ds))} \leq \frac{w(s)}{A(s,N(s+ds))}$. Combining these inequalities yields

$$\frac{A(s+ds, N(s))}{A(s, N(s))} \leq \frac{w(s+ds)}{w(s)} \leq \frac{A(s+ds, N(s+ds))}{A(s, N(s+ds))},$$

from which we can obtain the differentiability of $w(s)$ and equation (10), after taking logs, dividing by ds and taking limits as $ds \to 0$.[49] This proves **condition i**.

The pricing rule (5) and the existence of H imply $\phi = \arg\max_\gamma p(\gamma) A(H(\phi), \gamma)$, so similar arguments to those used in the last paragraph yield the differentiability of $p(\phi)$ and condition (11). The differentiability of $r^d(\phi)$ and condition (12) follow from the definition of $r^d(\phi)$ in (6) and the differentiability of $p(\phi)$.

The pricing rule (5) implies that the variable production cost of a firm equals a fraction $(\sigma-1)/\sigma$ of its revenue. Then, the total wages paid to production workers employed at firms with productivity weakly lower than ϕ must be equal to a fraction $(\sigma-1)/\sigma$ of the total revenue generated by those firms,

$$\int_{\underline{s}}^{H(\phi)} w(s) V(s) [L - fM] ds = \tfrac{(\sigma-1)}{\sigma} \int_{\underline{\phi}}^{\phi} r^d(\phi') g(\phi') d\phi' \overline{M} \text{ for all } \phi \in \left[\phi^*, \overline{\phi}\right]. \quad (23)$$

[48] https://www.marianosomale.com/

[49] The limits are well defined since all the functions involved are continuous.

As $r^d(\phi')\,g(\phi')$ is continuous, the right hand side of (23) is a differentiable function of the limit of integration ϕ. Accordingly, the left-hand side is also a differentiable function of ϕ, implying that $H(\phi)$ is differentiable. Differentiating (23) with respect to ϕ and using the pricing rule (5) to substitute for the wage $w(s)$ yields condition (13). Concluding the proof of **condition ii**, the boundary conditions on H and r^d are obtained as indicated in the main text. Finally, **condition iii** follows from equation (23), evaluated at $\phi = \overline{\phi}$, and the numeraire assumption, $\int_{\underline{s}}^{\overline{s}} w(s)\,V(s)\,ds = 1$.

Turning to the **sufficiency result in the lemma**, suppose that $\left\{\phi^*, p, r^d, H\right\}$ satisfy conditions (ii)-(iii). From these variables, define $N \equiv H^{-1}$, $M \equiv [1 - G(\phi^*)]\overline{M}$, $w(s) \equiv \frac{\sigma-1}{\sigma} A(s, N(s))\,p(N(s))$, $q(\phi) \equiv \frac{r(\phi)}{p(\phi)}$, and $l(s,\phi) \equiv V(s)[L - fM]\delta(\phi - N(s))$, where $\delta(x)$ is the Dirac-delta function. It can be easily checked that the variables $\{\phi^*$, M, $w(s)\,, p(\phi)\,, q(\phi)\,, l(s,\phi)\}$ satisfy all the conditions in the definition of equilibrium. This concludes the proof of the lemma.

B.1.2 Matching function and Lorenz dominance

The the poorest ρ fraction of workers in the interval $[s_0, s_1]$ is associated with the skill level $s(\rho)$ that satisfies $\rho = \int_{s_0}^{s(\rho)} V(s)\,ds \Big/ \int_{s_0}^{s_1} V(s)\,ds$. Accordingly, the Lorenz Curve of wage income is given by

$$L(\rho) \equiv \frac{\int_{s_0}^{s(\rho)} w(s)\,V(s)\,ds}{\int_{s_0}^{s_1} w(s)\,V(s)\,ds} = \frac{\int_{s_0}^{s(\rho)} \frac{w(s)}{w(s(\rho))} V(s)\,ds}{\int_{s_0}^{s(\rho)} \frac{w(s)}{w(s(\rho))} V(s)\,ds + \int_{s(\rho)}^{s_1} \frac{w(s)}{w(s(\rho))} V(s)\,ds}$$

Consider two economies A and B with matching functions N^A and N^B such that $N^B(s) > N^A(s)$ for all $s \in [s_0, s_1] \subseteq [\underline{s}, \overline{s}]$. The strict log-supermodularity of the productivity function implies $w^A(s')/w^A(s) < w^B(s')/w^B(s)$, for all $s' > s$ in $[s_0, s_1]$. Using this observation in the last expression implies $L^A(\rho) > L^B(\rho)$ for all $\rho \in (0,1)$. Atkinson (1970) showed that Lorenz dominance is equivalent to *Normalized* Second-Order Stochastic Dominance.

B.2 Section 4

B.2.1 Definition of Equilibrium, Open Economy

Definition 2 *A no-free-entry equilibrium of the open economy is an activity cutoff* ϕ^**, a mass of active firms* $M > 0$*, a mass of exporters* $M^x(\phi) > 0$ *for each productivity level* $\phi \geq \phi^*$*, output functions* $q^d, q^x : [\phi^*, \overline{\phi}] \to \mathbb{R}_+$*, labor allocations functions* $l^d, l^x : S \times [\phi^*, \overline{\phi}] \to \mathbb{R}_+$*, a price function* $p : [\phi^*, \overline{\phi}] \to \mathbb{R}_+$ *and a wage schedule* $w : S \to \mathbb{R}_+$ *such that the following conditions hold,*
(i) consumers behave optimally, equations (1) and (2);
(ii) firms behave optimally given their technology, equations (3), (5), (7), (8) and (16);
(iii) goods and labor markets clear, equations (6), (15) and (17);
(iv) the numeraire assumption holds, $\overline{w} = 1$*.*

B.2.2 Characterization of Equilibrium, Open Economy

Lemma 3 *In a no-free-entry equilibrium of the open economy with activity cutoff* $\phi^* \in (\underline{\phi}, \overline{\phi})$ *the following conditions hold.*
(i) There exists a continuous and strictly increasing matching function $N : S \to [\phi^*, \overline{\phi}]$*, (with inverse function* H*) such that (i)* $l^d(s,\phi) + l^x(s,\phi) > 0$ *if and only if* $N(s) = \phi$*, (ii)* $N(\underline{s}) = \phi^*$*, and* $N(\overline{s}) = \overline{\phi}$*.*
(ii) The wage schedule w *is continuously differentiable and satisfies (10)*
(iii) The price, domestic revenue and matching functions, $\{p, r^d, N\}$*, are continuously differentiable. Given* ϕ^**, the triplet* $\{p, r^d, H\}$ *solves the BVP comprising the system of differential equations {(11), (12), (18)} and the boundary conditions* $r^d(\phi^*) = \sigma f$*,* $H(\phi^*) = \underline{s}$*,* $H(\overline{\phi}) = \overline{s}$*.*
(iv) The activity cutoff ϕ^* *and the revenue function* r^d *satisfy (19).*
Moreover, if a number $\phi^* \in (\underline{\phi}, \overline{\phi})$*, and functions* $p, r^d : [\phi^*, \overline{\phi}] \to \mathbb{R}_+$ *and* $H : [\phi^*, \overline{\phi}] \to S$ *satisfy the conditions (iii)-(iv), then they are, respectively, the activity cutoff, the price function, the domestic revenue function, and the inverse of the matching function of a no-free-entry equilibrium of the open economy.*

Proof. Adapt arguments in the proof of lemma 1. ■

B.2.3 Proof of Lemma 2

Existence. Define the functional Ψ, mapping the space of continuous functions into itself, as follows

$$\Psi(y)(\phi) \equiv s_0 + [s_1 - s_0] \frac{\int_{\phi_0}^{\phi} h(t, y(t))\, e^{\sigma \int_{\phi_0}^{t} \frac{\partial \ln A(y(u),u)}{\partial \phi} du} \left[1 + F\left(K_0 e^{(\sigma-1)\int_{\phi_0}^{t} \frac{\partial \ln A(y(u),u)}{\partial \phi} dt}\right) K_1\right] dt}{\int_{\phi_0}^{\phi_1} h(t, y(t))\, e^{\sigma \int_{\phi_0}^{t} \frac{\partial \ln A(y(u),u)}{\partial \phi} du} \left[1 + F\left(K_0 e^{(\sigma-1)\int_{\phi_0}^{t} \frac{\partial \ln A(y(u),u)}{\partial \phi} dt}\right) K_1\right] dt}, \tag{24}$$

where

$$h(t, y(t)) \equiv \frac{A(s_0, \phi_0)}{A(y(t), t)} \frac{V(s_0)}{V(y(t))} \frac{g(t)}{g(\phi_0)} \frac{\alpha(t)}{\alpha(\phi_0)}. \tag{25}$$

Two observations about the functional Ψ are in order. First, the problem of finding a solution to BVP (20) is equivalent to the problem of finding a fixed point of Ψ. To see this, let $\{z, x, \Gamma\}$ be a solution to the BVP (20). Then, equations (20a)-(20c) and condition $x(\phi_0) = 1$ can be combined to find an expression for $\Gamma_\phi(t)/\Gamma_\phi(\phi_0)$ for any $t \in (\phi_0, \phi_1]$. Integrating this expression with respect to dt between ϕ_0 and ϕ yields

$$\Gamma(\phi) = \Gamma(\phi_0) + \Gamma_\phi(\phi_0) \int_{\phi_0}^{\phi} \frac{h(t, \Gamma(t))}{[1 + F(K_0) K_1]} e^{\sigma \int_{\phi_0}^{t} \frac{\partial \ln A(\Gamma(u),u)}{\partial \phi} du} [1 + F(K_0 x(t)) K_1]\, dt. \tag{26}$$

Evaluating the last expression at $\phi = \phi_1$, and using the boundary conditions on Γ, we can solve for $\Gamma_\phi(\phi_0)$. Using this expression for $\Gamma_\phi(\phi_0)$ back in (26), $x(t) = e^{(\sigma-1)\int_{\phi_0}^{t} \frac{\partial \ln A(H(t),t)}{\partial \phi} dt}$, and the definition of Ψ in (24) yields $\Gamma = \Psi(\Gamma)$—i.e., Γ is a fixed point of Ψ.

On the other direction, let Γ be a fixed point of Ψ. If we define $x(\phi) = e^{(\sigma-1)\int_{\phi_0}^{t} \frac{\partial \ln A(\Gamma(u),u)}{\partial \phi} du}$ and $z(\phi) = \frac{[1+F(K_0)K_1]\alpha(\phi_0) g(\phi_0)}{A(s_0,\phi_0) V(s_0) \Gamma_\phi(\phi_0)} e^{-\int_{\phi_0}^{t} \frac{\partial \ln A(\Gamma(u),u)}{\partial \phi} du}$, then it is easy to verify that $\{z, x, \Gamma\}$ is a solution to BVP (20).

The second observation is that Ψ is a compact self-map on the convex and closed set

$$K \equiv \{y \in \mathbf{C}[\phi_0, \phi_1] : s_0 \le y(\phi) \le s_1 \text{ for all } \phi \in [\phi_0, \phi_1]\}. \tag{27}$$

Per the Arzela-Ascoli theorem, Ψ is compact if the the set $\Psi(K)$ is bounded and equicontinuous. Indeed, letting $\{\underline{h}, \overline{h}\}$ be the minimum and maximum values of $h(\phi, y)$ on $[\phi_0, \phi_1] \times [s_0, s_1]$ and $\{\underline{r}, \overline{r}\}$ be the corresponding values associated to the function $\frac{\partial \ln A(y,\phi)}{\partial \phi}$, it can be shown that $\|\Psi(y)\|_\infty \leq s_0 + [s_1 - s_0] \frac{\overline{h}}{\underline{h}} e^{\sigma\overline{r}(\phi_1-\phi_0)} (1 + K_1)$ for any $y \in K$, which shows that $\Psi(K)$ is bounded. In addition, it can be shown that $|\Psi(y)(\phi') - \Psi(y)(\phi)| \leq \frac{[s_1-s_0]}{(\phi_1-\phi_0)} \frac{\overline{h}}{\underline{h}} e^{\sigma\overline{r}(\phi_1-\phi_0)} (1 + K_1) |\phi' - \phi|$ for any $y \in K$ and $\phi' > \phi$, which implies that $\Psi(K)$ is equicontinuous on $[\phi_0, \phi_1]$.

These observations, together with the Schauder fixed point theorem, imply that Ψ has a fixed point in K, and that this fixed point is a solution to BVP (20). Finally, the continuity of $\{A, V, g, \alpha, F\}$ and (20c) implies that Γ is continuously differentiable.

Uniqueness. As a first step, note that the initial value problem (IVP) given by the differential equations (20a)-(20c) and initial conditions $\{x(\phi_0) = 1, \Gamma(\phi_0) = s_0, z(\phi_0) = z_0\}$ has at most one solution, as the right-hand side of equations (20a)-(20c) are locally Lipschitz continuous with respect to $\{z, x, \Gamma\}$.

To prove that the solution to BVP (20) is unique, I proceed by contradiction. Suppose that there are two different solutions $\{z', x', \Gamma'\}$ and $\{z, x, \Gamma\}$ to BVP (20). Then, the uniqueness result in the previous paragraph implies that $z'(\phi_0) \neq z(\phi_0)$, which, together with equation (20c), implies $\Gamma'_\phi(\phi_0) \neq \Gamma_\phi(\phi_0)$. Without loss of generality suppose $\Gamma'_\phi(\phi_0) < \Gamma_\phi(\phi_0)$, which, in turn, yields $\Gamma(\phi) > \Gamma'(\phi)$ in some neighborhood (ϕ_0, c), with $c > \phi_0$. By assumption, we know that the functions Γ' and Γ must intersect at least once again on $(\phi_0, \phi_1]$ since $\Gamma(\phi_1) = \Gamma'(\phi_1)$. Let ϕ^+ be the first value to the right of ϕ_0 at which the functions Γ' and Γ intersect—$\phi^+ \equiv \inf\{\phi \in (\phi_0, \phi_1] : \Gamma'(\phi) = \Gamma(\phi)\}$—and note that ϕ^+ is well-defined since Γ' and Γ are continuous. These observations—$\Gamma(\phi) > \Gamma'(\phi)$ on (ϕ_0, ϕ^+) and $\Gamma(\phi^+) = \Gamma'(\phi^+)$—imply $\Gamma'_\phi(\phi^+) \geq \Gamma_\phi(\phi^+)$, which together with $\Gamma'_\phi(\phi_0) < \Gamma_\phi(\phi_0)$ yields

$$\frac{\Gamma'_\phi(\phi^+)/\Gamma'_\phi(\phi_0)}{\Gamma_\phi(\phi^+)/\Gamma_\phi(\phi_0)} > 1. \tag{28}$$

As discussed above, Γ' and Γ are fixed points of the functional Ψ defined in (24), so $\Gamma_\phi(\phi)$ and $\Gamma'_\phi(\phi)$ can be obtained differentiating the right-hand side of (24). Combining the resulting expressions yields

$$\frac{\Gamma'_\phi\left(\phi^+\right)/\Gamma'_\phi\left(\phi_0\right)}{\Gamma_\phi\left(\phi^+\right)/\Gamma_\phi\left(\phi_0\right)} = \exp\left\{\sigma \int_{\phi_0}^{\phi^+}\left[\frac{\partial \ln A(\Gamma'(u),u)}{\partial\phi} - \frac{\partial \ln A(\Gamma(u),u)}{\partial\phi}\right]du\right\} \times \cdots$$
$$\cdots \frac{\left[1 + F\left(K_0 \exp\left\{(\sigma-1)\int_{\phi_0}^{\phi^+}\frac{\partial \ln A(\Gamma'(u),u)}{\partial\phi}du\right\}\right)K_1\right]}{\left[1 + F\left(K_0 \exp\left\{(\sigma-1)\int_{\phi_0}^{\phi^+}\frac{\partial \ln A(\Gamma(u),u)}{\partial\phi}du\right\}\right)K_1\right]} < 1, \tag{29}$$

where in the last expression I used the fact that $\Gamma'\left(\phi^+\right) = \Gamma\left(\phi^+\right)$, so $h\left(\phi^+, \Gamma'\left(\phi^+\right)\right) = h\left(\phi^+, \Gamma\left(\phi^+\right)\right)$. The log-supermodularity of A, $\Gamma(\phi) > \Gamma'(\phi)$ on $\left(\phi_0, \phi^+\right)$, and the fact that F strictly increasing imply that each of the terms on the right-hand side of the last expression is strictly less than 1. As the inequality in (29) contradicts that in (28), it must be the case that BVP (20) has a unique solution.

Condition i. Let $\{z^i, x^i, \Gamma^i\}$ be the unique solution to BVP (20) with $K_1 = 0$ and $s_0 = s_0^i$, for $i = a, b$ and $s_0^a > s_0^b$. To derive a contradiction, suppose that there is a $\phi^+ \in (\phi_0, \phi_1)$ with $\Gamma^a\left(\phi^+\right) = \Gamma^b\left(\phi^+\right) \equiv s^+$. If we define the functions $y^i, w^i : [\phi_0, \phi_1] \rightarrow \mathbb{R}_+$ as $y^i(\phi) = z^i(\phi)/x^i\left(\phi^+\right)$, $w^i = x^i(\phi)/x^i\left(\phi^+\right)$, it is readily seen that on $\left[\phi^+, \phi_1\right]$ and for $i = a, b$, $\{y^i, w^i, \Gamma^i\}$ is a solution to the BVP given by the system of differential equations (20a)-(20c) and boundary conditions $w(\phi^+) = 1$, $\Gamma(\phi^+) = s^+$, $\Gamma(\phi_1) = s_1$. As this BVP is just a particular case of BVP (20), it has a unique solution, implying that $\{y^a, w^a, \Gamma^a\} = \left\{y^b, w^b, \Gamma^b\right\}$ on $\left[\phi^+, \phi_1\right]$. In turn, this result implies that $\{w^a, y^a, \Gamma^a\}$ and $\left\{w^b, y^b, \Gamma^b\right\}$ solve the same IVP on $[\phi_0, \phi_1]$ given by the system (20a)-(20c) and the same initial conditions at any $\phi \in \left(\phi_+, \phi_1\right)$. As such, our earlier uniqueness result for IVPs implies that $\Gamma^a(\phi) = \Gamma^b(\phi)$ on $[\phi_0, \phi_1]$, which contradicts $s_0^a > s_0^b$. The no-crossing result related to the inverses of Γ^i can be establish in a similar way.

Condition ii. Let $\phi_0^a > \phi_0^b$ and suppose that $x^a(\phi) \equiv x(\phi; \phi_0^a) \geq x\left(\phi; \phi_0^b\right) \equiv x^b(\phi)$ for some ϕ on $[\phi_0^a, \phi_1]$. Note that $x^a\left(\phi_0^a\right) < x^b(\phi_0^a)$, so there is a productivity level ϕ' such that $x^a(\phi) = x^b(\phi)$ for the first time. By definition of ϕ', we have $x^a(\phi') = x^b(\phi')$ and $x^a(\phi) < x^b(\phi)$ for $\phi < \phi'$, so $x^a(\phi)$ grows faster than $x^b(\phi)$ in some neighborhood to the left of ϕ'. This observation, equation (20b), and the log-supermodularity of A imply that there is a $\phi'' < \phi'$, such that $\Gamma^a(\phi) > \Gamma^b(\phi)$ on (ϕ'', ϕ'). As $\Gamma^a\left(\phi_0^a\right) < \Gamma^b\left(\phi_0^a\right)$ and $\Gamma^a(\phi_1) = \Gamma^b(\phi_1)$, then Γ^a and Γ^b must intersect at least once to left of ϕ'' and to the right of ϕ'. I use ϕ_- and ϕ_+ to denote, respectively, the productivity levels corresponding to the first intersections of Γ^a and Γ^b that are weakly to the left of ϕ'' and weakly to the right of ϕ'. Note that $\phi_- \leq \phi'' < \phi' < \phi_+$.

The previous discussion implies $\Gamma^a\left(\phi_-\right) = \Gamma^b\left(\phi_-\right)$, $\Gamma^a\left(\phi\right) > \Gamma^b\left(\phi\right)$ on $\left(\phi_-, \phi_+\right)$ and $\Gamma^a\left(\phi_+\right) = \Gamma^b\left(\phi_+\right)$. Then $\Gamma^a_\phi\left(\phi_-\right) \geq \Gamma^b_\phi\left(\phi_-\right)$ and $\Gamma^a_\phi\left(\phi_+\right) \leq \Gamma^b_\phi\left(\phi_+\right)$, so $\frac{\Gamma^a_\phi(\phi_+)/\Gamma^a_\phi(\phi_-)}{\Gamma^b_\phi(\phi_+)/\Gamma^b_\phi(\phi_-)} \leq 1$. In addition, differentiating the right-hand side of (24) to get an expression for Γ^i_ϕ yields

$$\frac{\Gamma^a_\phi(\phi_+)/\Gamma^a_\phi(\phi_-)}{\Gamma^b_\phi(\phi_+)/\Gamma^b_\phi(\phi_-)} > \frac{x^a\left(\phi_+\right)/x^a\left(\phi_-\right)}{x^b\left(\phi_+\right)/x^b\left(\phi_-\right)} \frac{\left[1+F\left(K_0 x^a\left(\phi_+\right)\right)K_1\right]/\left[1+F\left(K_0 x^a\left(\phi_-\right)\right)K_1\right]}{\left[1+F\left(K_0 x^b\left(\phi_+\right)\right)K_1\right]/\left[1+F\left(K_0 x^b\left(\phi_-\right)\right)K_1\right]}.$$

By definition, $x^a\left(\phi'\right) = x^b(\phi')$, $x^a\left(\phi_+\right) \geq x^b(\phi_+)$ (as $\Gamma^a\left(\phi\right) \geq \Gamma^b\left(\phi\right)$ on $\left[\phi', \phi_+\right]$), and $x^a\left(\phi_-\right) < x^b(\phi_-)$, so the last expression implies $\frac{\Gamma^a_\phi(\phi_+)/\Gamma^a_\phi(\phi_-)}{\Gamma^b_\phi(\phi_+)/\Gamma^b_\phi(\phi_-)} > 1$, contradicting our earlier result. Then, it must be the case that $x^a\left(\phi\right) < x^b\left(\phi\right)$ for all $\phi \in [\phi^a_0, \phi_1]$.

B.2.4 Proof of Proposition 1

The proof of the existence and uniqueness of the equilibrium in the closed and open economies was laid out in the text. Here, I prove the (constrained) efficiency of the equilibrium by showing the equivalence between equilibrium allocations and solutions to the relevant planner's problems.

Consider the following **closed-economy** planner's problem

$$\max_{\phi^*, \widetilde{q}(\phi), \widetilde{H}(\phi)} \int_{\phi^*}^{\overline{\phi}} \widetilde{q}\left(\phi\right)^{\frac{\sigma-1}{\sigma}} g\left(\phi\right)\overline{M} d\phi \quad \text{subject to}$$

$$\begin{gathered}\widetilde{H}_\phi\left(\phi\right) = \frac{\widetilde{q}(\phi) g(\phi)\overline{M}}{A(\widetilde{H}(\phi),\phi)V(\widetilde{H}(\phi))\left[L - f[1-G(\phi^*)]\overline{M}\right]} \equiv h^H(\phi^*, \widetilde{q}\left(\phi\right), \widetilde{H}\left(\phi\right), \phi) \text{ for all } \phi \in \left[\phi^*, \overline{\phi}\right], \\ \widetilde{H}(\phi^*) = \underline{s};\ \widetilde{H}(\overline{\phi}) = \overline{s}.\end{gathered} \tag{30}$$

The Lagrangian can be expressed as

$$\begin{aligned} L(\phi^*, \widetilde{q}, \widetilde{H}) = \int_{\phi^*}^{\overline{\phi}} \widetilde{q}\left(\phi\right)^{\frac{\sigma-1}{\sigma}} g\left(\phi\right) d\phi \overline{M} + \int_{\phi^*}^{\overline{\phi}} \widetilde{H}\left(\phi\right) \lambda^H_\phi\left(\phi\right) d\phi + \lambda^H(\phi^*)\underline{s} - \widetilde{H}\left(\overline{\phi}\right)\lambda^H(\overline{\phi}) + \cdots \\ \cdots \int_{\phi^*}^{\overline{\phi}} h^H(\phi^*, \widetilde{q}\left(\phi\right), \widetilde{H}\left(\phi\right), \phi)\lambda^H(\phi) d\phi + \mu^H\left[H\left(\overline{\phi}\right) - \overline{s}\right] \end{aligned}$$

The stationarity condition, together with the constraints of the problem, yields the following first order necessary conditions for an optimum

$$\widetilde{H}_{\phi}(\phi) = h^{H}(\phi^{*}, \widetilde{q}(\phi), \widetilde{H}(\phi), \phi)$$
$$h^{H}_{H}(\phi^{*}, \widetilde{q}(\phi), \widetilde{H}(\phi), \phi)\lambda^{H}(\phi) + \lambda^{H}_{\phi}(\phi) = 0$$
$$\tfrac{\sigma-1}{\sigma}\widetilde{q}(\phi)^{-\frac{1}{\sigma}} g(\phi)\overline{M} + h^{H}_{q}(\phi^{*}, \widetilde{q}(\phi), \widetilde{H}(\phi), \phi)\lambda^{H}(\phi) = 0$$
$$\left[\mu^{H} - \lambda^{H}\left(\overline{\phi}\right)\right] = 0$$
$$\widetilde{H}(\phi^{*}) = \underline{s}, \quad \widetilde{H}(\overline{\phi}) = \overline{s}$$
$$\textstyle\int_{\phi^{*}}^{\overline{\phi}} h^{H}_{\phi^{*}}(\phi^{*}, \widetilde{q}(\phi), \widetilde{H}(\phi), \phi)\lambda^{H}(\phi)d\phi = \widetilde{q}(\phi^{*})^{\frac{\sigma-1}{\sigma}} g(\phi^{*})\overline{M} + h^{H}(\phi^{*}, \widetilde{q}(\phi), \widetilde{H}(\phi), \phi)\lambda^{H}(\phi^{*}). \tag{31}$$

It can be shown that if $\{\phi^{*}, \widetilde{q}, \widetilde{H}, \lambda^{H}\}$ satisfies (31), then we can define functions $\{\widetilde{p}(\phi), \widetilde{r}(\phi)\}$ such that $\{\phi^{*}, \widetilde{p}(\phi), \widetilde{r}(\phi), \widetilde{H}\}$ satisfy the conditions of lemma 1, proving that a solution to the planner's problem is an equilibrium of the closed economy. On the other direction, it can be shown that if $\left\{\phi^{*}_{a}, p, r^{d}, H\right\}$ are the activity cutoff, price, revenue and inverse matching functions of the closed economy equilibrium, with output function $q^{d}(\phi) = r^{d}(\phi)/p(\phi)$, then $\left\{q^{d}, H, \lambda, \phi^{*}_{a}\right\}$ solves the planner's problem (30). For a step-by-step proof, see the extended appendix to this paper in my personal website.

In a similar way, one can show that an allocation is an equilibrium of the **open economy** if and only if it is a solution to the open-economy analog of problem (30) when $f\tau^{1-\sigma} \leq f_{x}$. When this restriction on parameters is not satisfied, the equivalence between equilibria of the open economy and solutions to said problem no longer holds. Intuitively, if $f\tau^{1-\sigma} > f_{x}$, then the planner is willing to accept some "negative domestic profits", $\widetilde{r}^{d}(\phi^{*}) < \sigma f$, because they are more than offset by positive export profits. However, by changing slightly the arguments, it can be shown that when $f\tau^{1-\sigma} > f_{x}$, the equilibria of the open economy are equivalent to solutions to constrained planner's problems that feature the following additional constraint

$\sigma f \int_{\phi^{*}}^{\overline{\phi}} \left[\frac{\widetilde{q}^{d}(\phi)}{\widetilde{q}^{d}(\phi^{*})}\right]^{\frac{\sigma-1}{\sigma}} \left[1 + F(\widetilde{y}(\phi))\tau^{1-\sigma}n\right] g(\phi)\overline{M}d\phi = \frac{\sigma}{\sigma-1}L^{pw}(\phi^{*}, \widetilde{y}(\phi))$,

where L^{pw} is the mass of production workers. Accordingly, the equilibrium is constrained efficient in this case.

B.3 Additional Results related to BVP (20)

In this section, I present some results related to BVP (20) that are used in the text and in the proof of other results.

Lemma 4 *For $i = a, b$, let $\{z^i, x^i, \Gamma^i\}$ be the unique solution to the BVP (20) with parameters $\{\alpha^i(\phi), K_0^i, K_1^i\}$ and boundary conditions $\{x^i(\phi_0) = 1,\ \Gamma^i(\phi_0) = s_0,\ \Gamma^i(\phi_1) = s_1\}$.*

(i) Suppose that $K_1^i = 0$, $\frac{\alpha^a(\phi')}{\alpha^a(\phi)} \geq \frac{\alpha^b(\phi')}{\alpha^b(\phi)}$ for all $\phi' > \phi \in [\phi_0, \phi_1]$, and $\frac{\alpha^a(\phi')}{\alpha^a(\phi)} > \frac{\alpha^b(\phi')}{\alpha^b(\phi)}$ for all $\phi' > \phi$ on some subinterval $[\phi_l, \phi_h] \subseteq [\phi_0, \phi_1]$. Then $\Gamma^a(\phi) < \Gamma^b(\phi)$ for all $\phi \in (\phi_0, \phi_1)$ and $\Gamma^a_\phi(\phi_0) < \Gamma^b_\phi(\phi_0)$ and $\Gamma^a_\phi(\phi_1) > \Gamma^b_\phi(\phi_1)$.

(ii) Suppose that $K_0^i = K_0$, $\alpha^i(\phi) = \alpha(\phi)$ and $K_1^b < K_1^a$. Then $\Gamma^a_\phi(\phi_0) < \Gamma^b_\phi(\phi_0)$, so there is a $\phi^+ \in (\phi_0, \phi_1]$ such that $\Gamma^a(\phi^+) = \Gamma^b(\phi^+)$ and $\Gamma^a(\phi) < \Gamma^b(\phi)$ for all $\phi \in (\phi_0, \phi^+)$.

(iii) Let $\Phi^i \equiv \int_{\phi_0}^{\phi_1} x^i(\phi) \frac{\left[1+F\left(K_0^i x^i(\phi)\right)K_1^i\right]}{\left[1+F\left(K_0^i\right)K_1^i\right]} \frac{\alpha^i(\phi)}{\alpha^i(\phi_0)} g(\phi)\, d\phi$. If $\Gamma^a(\phi) < \Gamma^b(\phi)$ for $\phi \in (\phi_0, \phi_1)$, then $\Phi^a > \Phi^b$.

(iv) If $\alpha^i(\phi) = \alpha(\phi)$, $K_0^b = \lambda K_0^a$ and $K_1^b = \lambda K_1^a$ for $\lambda > 1$, then $x^b(\phi)\lambda > x^a(\phi)$ for all for all $\phi \in [\phi_0, \phi_1]$.

(v) Let $\delta^i(\phi) \equiv [1 + F(K_0^i x^i(\phi)) K_1^i]\, \alpha^i(\phi)$. If $\Gamma^a \neq \Gamma^b$ and, $\delta^a(\phi) < \delta^b(\phi)$ for all $\phi \in [\phi_0, \phi_1]$, then

$$\int_{\phi_0}^{\phi_1} x^a(\phi)\, \delta^a(\phi)\, g(\phi)\, d\phi < \int_{\phi_0}^{\phi_1} x^b(\phi)\, \delta^b(\phi)\, g(\phi)\, d\phi. \tag{32}$$

(vi) Suppose that $\{\alpha^i(\phi), K_1^i\} = \{\alpha(\phi), K_1\}$, $K_0^i, K_1 \in \mathbb{R}_{++}$ and $K_0^a > K_0^b$. If the function $\eta_0(t, \lambda) \equiv \frac{F_y(t\lambda)\lambda K_1}{[1+F(t\lambda)K_1]}$ is strictly decreasing (increasing) in λ on $[1, \infty)$ for $t \in [K_0^b, K_0^b x^b(\phi_1)]$, then $\Gamma^a(\phi) > (<)\Gamma^b(\phi)$ on (ϕ_0, ϕ_1), with $\Gamma^a_\phi(\phi_0) > (<)\Gamma^b_\phi(\phi)$.

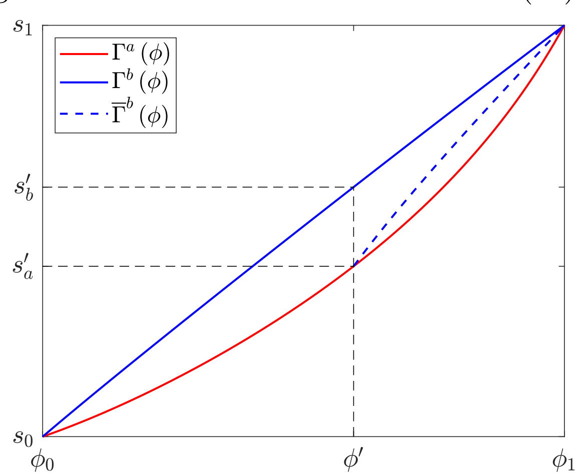
(vii) Suppose that $\alpha^i(\phi) = \alpha(\phi)$, $K_0^i, K_1^i \in \mathbb{R}_{++}$ and $K_i^a = \lambda K_i^b$ for $\lambda > 1$. If the function $\eta_1(t, \lambda) \equiv \frac{F_y(t\lambda)\lambda^2 K_1^b}{\left[1+F(t\lambda)\lambda K_1^b\right]}$ is strictly increasing (decreasing) in λ on $[1, \infty)$ for $t \in [K_0^b, K_0^b x^b(\phi_1)]$, then $\Gamma^a(\phi) < (>)\Gamma^b(\phi)$ on (ϕ_0, ϕ_1) with $\Gamma^a_\phi(\phi_0) < (>)\Gamma^b_\phi(\phi_0)$.

Proof. Lemma 4.i. I start by showing that there is some $\phi' \in (\phi_0, \phi_1)$ such that $\Gamma^a(\phi') < \Gamma^b(\phi')$. First, note that $\Gamma^a(\phi) \leq \Gamma^b(\phi)$ for all $\phi \in (\phi_0, \phi_1)$. To see this, suppose to the contrary that there is a $\phi' \in (\phi_0, \phi_1)$ such that $\Gamma^a(\phi') > \Gamma^b(\phi')$, and let ϕ_- and ϕ_+ be the first time the functions Γ^a and Γ^b intersect to the left and to the right of ϕ', respectively. As $\Gamma^a(\phi) > \Gamma^b(\phi)$ for $\phi \in (\phi_-, \phi_+)$, then $\frac{\Gamma^a_\phi(\phi_+)/\Gamma^a_\phi(\phi_-)}{\Gamma^b_\phi(\phi_+)/\Gamma^b_\phi(\phi_-)} \leq 1$. Differentiating the right-hand side of (24) to obtain an expression for $\Gamma^i_\phi(\phi)$ yields

$$\frac{\Gamma^a_\phi(\phi_+)/\Gamma^a_\phi(\phi_-)}{\Gamma^b_\phi(\phi_+)/\Gamma^b_\phi(\phi_-)} = e^{\sigma \int_{\phi_-}^{\phi_+} \left[\frac{\partial \ln A(\Gamma^a(u),u)}{\partial \phi} - \frac{\partial \ln A\left(\Gamma^b(u),u\right)}{\partial \phi}\right] du} \frac{\alpha^a(\phi_+)/\alpha^a(\phi_-)}{\alpha^b(\phi_+)/\alpha^b(\phi_-)}. \tag{33}$$

The strict log-supermodularity of A and $\Gamma^a(\phi) > \Gamma^b(\phi)$ for $\phi \in (\phi_-, \phi_+)$ imply that the first term of the last expression is strictly greater than one. As the assumptions about α^a and α^b imply that the second term is weakly greater than one, then $\frac{\Gamma^a_\phi(\phi_+)/\Gamma^a_\phi(\phi_-)}{\Gamma^b_\phi(\phi_+)/\Gamma^b_\phi(\phi_-)} > 1$. The last inequality contradicts our previous result, implying that $\Gamma^a(\phi) \leq \Gamma^b(\phi)$ for all $\phi \in (\phi_0, \phi_1)$. Moreover, a similar argument shows that assuming $\Gamma^a(\phi) = \Gamma^b(\phi)$ on any nondegenerate interval $I \subseteq [\phi_l, \phi_h]$ also yields contradictory implications about the relative slopes of these functions, so there must be some $\phi' \in [\phi_0, \phi_1]$ such that $\Gamma^a(\phi') < \Gamma^b(\phi')$.

Figure 9: Solutions to the General BVP (20), Γ



Note: The figure depicts solutions to alternative parametrizations of the general BVP (20). The BVPs corresponding to Γ^a and Γ^b differ only in the parameter function $\alpha(\phi)$ as indicated in lemma 4.i Restricted to $[\phi', \phi_1]$,the BVPs corresponding to Γ^b and $\overline{\Gamma}^b$ differ only in their initial conditions.

Next, I show that the previous result implies $\Gamma^a(\phi) < \Gamma^b(\phi)$ for all $\phi \in (\phi_0, \phi_1)$. First, with ϕ' defined as above, note that the restriction of Γ^b to $[\phi', \phi_1]$ is part of the unique solution to a properly defined version of BVP (20) on $[\phi', \phi_1]$. Now, let $\{\overline{z}^b, \overline{x}^b, \overline{\Gamma}^b\}$ be the unique solution to the BVP (20) on $[\phi', \phi_1]$ with the same parameters but with boundary conditions $\overline{x}^b(\phi') = 1$, $\overline{\Gamma}^b(\phi') = s'_a < s'_b$ and $\overline{\Gamma}^b(\phi_1) = s_1$. The situation is depicted in figure 9. It is readily seen that these two versions of BVP (20) satisfy the conditions of the no-crossing result in lemma 2.ii with $\overline{\Gamma}^b(\phi') < \Gamma^b(\phi')$, so $\overline{\Gamma}^b(\phi) < \Gamma^b(\phi)$ on $[\phi', \phi_1)$. In addition, it is easy to check that the BVPs associated to $\overline{\Gamma}^b(\phi)$ and to the restriction of $\Gamma^a(\phi)$ to $[\phi', \phi_1]$ satisfy the premises of Lemma 4.i, so the result in the previous paragraph implies $\Gamma^a(\phi) \leq \overline{\Gamma}^b(\phi)$ on $[\phi', \phi_1]$. Accordingly, $\Gamma^a(\phi) \leq \overline{\Gamma}^b(\phi) < \Gamma^b(\phi)$ for all $\phi \in [\phi', \phi_1)$. This argument can be easily adapted to show that there is a function $\underline{\Gamma}^b : [\phi_0, \phi'] \to S$, such that $\Gamma^a(\phi) \leq \underline{\Gamma}^b(\phi) < \Gamma^b(\phi)$ for all $\phi \in (\phi_0, \phi']$, completing the proof of the result. Of note, this second part of the argument requires a slightly different version of the no-crossing result in proposition 2.ii.

Finally, I show that $\Gamma^a_\phi(\phi_0) < \Gamma^b_\phi(\phi_0)$ and $\Gamma^a_\phi(\phi_1) > \Gamma^b_\phi(\phi_1)$. The relative positions of Γ^a, Γ^b and $\overline{\Gamma}^b$ imply $\Gamma^a_\phi(\phi_1) \geq \overline{\Gamma}^b_\phi(\phi_1) \geq \Gamma^b_\phi(\phi_1)$. Moreover, if $\overline{\Gamma}^b_\phi(\phi_1) = \Gamma^b_\phi(\phi_1)$, then Γ^b and $\overline{\Gamma}^b$ would be solutions to the same IVP with the same initial condition at ϕ_1, which would imply $\overline{\Gamma}^b = \Gamma^b$ on $[\phi', \phi_1]$. As this contradicts our earlier results, it must be the case that $\overline{\Gamma}^b_\phi(\phi_1) > \Gamma^b_\phi(\phi_1)$. Putting together these results we get $\Gamma^a_\phi(\phi_1) \geq \overline{\Gamma}^b_\phi(\phi_1) > \Gamma^b_\phi(\phi_1)$. The other part of the claim can be proved making only minor adjustments to this argument.

Lemma 4.ii. As a first step, note that there is no $\phi' \in (\phi_0, \phi_1]$ such that $\Gamma^a(\phi) \geq \Gamma^b(\phi)$ for all $\phi \in (\phi_0, \phi']$. Indeed, following the line of argument in the proof of lemma 4.i, one can show that assuming the existence of such a ϕ' leads to contradictory implications about the relative slopes of the functions Γ^a and Γ^b.

Next, I show $\Gamma^a_\phi(\phi_0) < \Gamma^b_\phi(\phi_0)$. The previous result yields $\Gamma^a_\phi(\phi_0) \leq \Gamma^b_\phi(\phi_0)$, so suppose for a moment that $\Gamma^a_\phi(\phi_0) = \Gamma^b_\phi(\phi_0) = \gamma_0$. As both BVPs have the same boundary conditions, equations (20a)-(20b) imply $x^i_\phi(\phi) = \frac{(\sigma-1)\partial \ln A(s_0,\phi_0)}{\partial \phi}$ and $\frac{z^i_\phi(\phi_0)}{z^i(\phi_0)} = -\frac{\partial \ln A(s_0,\phi_0)}{\partial \phi}$. Log-differentiating both sides of equation (20c) and evaluating at ϕ_0 yields

$$\Gamma^a_{\phi\phi}(\phi_0) - \Gamma^b_{\phi\phi}(\phi_0) = \frac{F_y(K_0)K_0}{F(K_0)} \frac{(\sigma-1)\partial \ln A(s_0,\phi_0)}{\partial \phi} \gamma_0 \left\{ \frac{F(K_0)K_1^a}{\left[1+F(K_0)K_1^a\right]} - \frac{F(K_0)K_1^b}{\left[1+F(K_0)K_1^b\right]} \right\} > 0,$$

where the inequality follows from $K_1^a > K_1^b$. The last expression implies that there is some $\phi' \in (\phi_0, \phi_1]$ such that $\Gamma_\phi^a(\phi) > \Gamma_\phi^b(\phi)$ on $(\phi_0, \phi']$, contradicting our earlier result.

Finally, $\Gamma_\phi^a(\phi_0) < \Gamma_\phi^b(\phi_0)$ implies $\Gamma^a(\phi) < \Gamma^b(\phi)$ on some (small enough) interval (ϕ_0, ϕ''), so ϕ^+ is the first intersection of Γ^a and Γ^b to the right of ϕ''.

Lemma 4.iii. The idea of the proof is to show that Γ^a and Γ^b can be thought of as the inverses of the matching functions of two artificial economies, and then use this additional information to prove the result. Let $\{z^i, x^i, \Gamma^i\}$ be the solution to the BVP in the statement of the lemma and consider the following artificial economy. In this economy, there are no fixed costs of production and no fixed costs to export but the set of active firms and the set of exporters are fixed. In particular, the set of active firms are those with productivity in the range $[\phi_0, \phi_1]$, while the fraction of firms that export at each productivity level is fixed and given by $FX^i(\phi) \equiv F(K_0^i x^i(\phi))$. The set of available workers are those with skills in the range $[s_0, s_1]$. The distribution of skills is given by the restriction of V to $[s_0, s_1]$ and the mass of workers is $\int_{s_0}^{s_1} V(s)\, dsL$. The total mass of firms with productivity ϕ is given by $g(\phi)\alpha^i(\phi)\overline{M}$, so the total mass of firms is $\int_{\phi_0}^{\phi_1} g(\phi)\alpha^i(\phi)\overline{M}$. Finally, τ_i is set to satisfy $K_1^i \equiv \tau_i^{1-\sigma}$.

Now I show that if p^i, $r^{d,i}$ and H^i denote the price, domestic revenue and inverse-matching functions of the economy described above, then $H^i = \Gamma^i$. An argument similar to the one in section 4 implies that $\{p^i, r^{d,i}, H^i\}$ satisfy the differential equations (11), (12) and

$$H_\phi^i(\phi) = \frac{r^{d,i}(\phi)\left[1 + FX^i(\phi) K_1^i\right] g(\phi)\alpha^i(\phi)\overline{M}}{A(H^i(\phi), \phi) V(H^i(\phi)) p^i(\phi) L}, \tag{34}$$

with boundary conditions $H^i(\phi_0) = s_0$ and $H^i(\phi_1) = s_1$. Note that there is no boundary condition on the domestic revenue function $r^{d,i}$, as the zero-profit condition for firms with productivity ϕ_0 is no longer an equilibrium condition (no fixed costs of production). As a result, the levels of the functions $r^{d,i}$ and p^i cannot be determined without an additional condition (provided below). However, these conditions are enough to pin down H^i. To see this, let $\{p^i, r^{d,i}, H^i\}$ be any triplet of functions satisfying the equilibrium conditions described above, and define $\delta^i(\phi) \equiv \left[1 + FX^i(\phi) K_1^i\right]\alpha^i(\phi)$, $v^i(\phi) \equiv r^{d,i}(\phi)/r^{d,i}(\phi_0)$ and $y^i(\phi) \equiv p^i(\phi) L / r^{d,i}(\phi_0)\overline{M}$. Then, it is readily seen that $\{y^i, v^i, H^i\}$ is the unique solution to the BVP (20) with parameter $K_1 = 0$ and $\alpha = \delta^i$.[50] However, note that, by

[50] With $K_1 = 0$, the value of K_0 is irrelevant.

construction, $\{z^i, x^i, \Gamma^i\}$ is also a solution to this parametrization of the BVP (20), so it must be the case that $H^i = \Gamma^i$.

Let us now derive an additional condition to pin down the revenue function in this artifical economy. In equilibrium, the total revenue of firms with productivity less or equal than ϕ' equals a constant fraction of the total wages paid to workers employed at those firms,

$$r^{d,i}(\phi_0)\,\alpha^i(\phi_0)\left[1 + F\left(K_0^i\right)K_1^i\right]\int_{\phi_0}^{\phi'} x^i(\phi)\,\frac{\left[1 + FX^i(\phi)\,K_1^i\right]}{\left[1 + F\left(K_0^i\right)K_1^i\right]}\frac{\alpha^i(\phi)}{\alpha^i(\phi_0)}\,g(\phi)\,\overline{M}d\phi = \qquad (35)$$

$$\cdots\frac{\sigma}{\sigma - 1}L\int_{s_0}^{H^i(\phi')} w^i\left(H^i(\phi)\right)V\left(H^i(\phi)\right)ds, \text{ for } i = a, b.$$

Differentiating the left- and right hand sides of the last expression with respect to ϕ', and evaluating the resulting expressions at $\phi' = \phi_0$ yields

$$r^{d,i}(\phi_0)\,\alpha^i(\phi_0)\left[1 + F\left(K_0^i\right)K_1^i\right]g(\phi_0)\overline{M} = \frac{\sigma}{\sigma - 1}Lw^i(s_0)\,V(s_0)\,H_\phi^i(\phi_0) \text{ for } i = a, b. \qquad (36)$$

The last expression, together with the numeraire assumption, $\int_{s_0}^{s_1} w^i(s)\,V(s)\,ds = 1$, and the inverse matching function H^i, can be used to pin down the value of $r_d^i(\phi_0)$. To see this, note that H^i determines the growth rate of wages along the skill dimension (condition 10), while the numeraire assumption pins down their levels, so the wage schedule is fully determined. Then, equation (36) can be used to pin down $r_d^i(\phi_0)$, the only remaining endogenous variable.

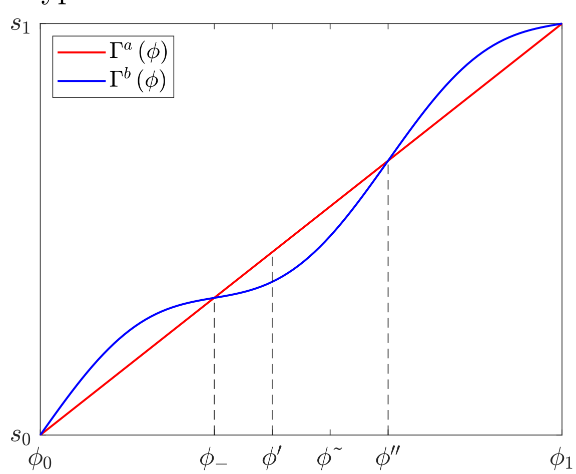
With previous results we are ready to prove the lemma. As $H^a(\phi) < H^b(\phi)$ for $\phi \in [\phi_0, \phi_1]$ by assumption, wages grow faster along the skill dimension in economy a than in economy b, so the numeraire assumption implies $w^a(s_0) < w^b(s_0)$. In addition, $H^a(\phi) < H^b(\phi)$ for $\phi \in [\phi_0, \phi_1]$ also implies that $H_\phi^a(\phi_0) \leq H_\phi^b(\phi_0)$. These observations and (36) imply $r^{d,a}(\phi_0)\,\alpha^a(\phi_0)\left[1 + F\left(K_0^a\right)K_1^a\right] < r^{d,b}(\phi_0)\,\alpha^b(\phi_0)\left[1 + F\left(K_0^b\right)K_1^b\right]$. Finally, the last inequality, expression (35) evaluated at $\phi' = \phi_1$ for $i = a, b$, and the numeraire assumption yield the desired result.

Lemma 4.iv. Differentiating the right-hand side of (24) yields,

$$\Gamma_{\phi}^{i}(\phi)=[s_1-s_0]\frac{h^i(\phi,\Gamma^i(\phi))\,x^i(\phi)^{\frac{\sigma}{\sigma-1}}\left[1+F\left(K_0^i x^i(\phi)\right)K_1^i\right]}{\int_{\phi_0}^{\phi_1}h^i(t,\Gamma^i(t))\,x^i(t)^{\frac{\sigma}{\sigma-1}}\left[1+F\left(K_0^i x^i(t)\right)K_1^i\right]dt}. \tag{37}$$

I prove the claim by showing that assuming that the claim does not hold leads to contradictory implications about the relative sizes of the denominators on the right-hand side of (37), Dem^i for $i=a,b$. I will use $Num^i(\phi)$ for the numerator on the right-hand side.

Figure 10: Hypotetical Solutions to the General BVP (20), Γ



Note: The figure depicts hypothetical solutions to the general BVP (20) with the features implied by the assumption $x^b(\phi)\lambda \leq x^a(\phi)$ given the conditions in lemma 4.iv. as described in the proof. Of note, said assumption implies $\phi^{\sim} \in (\phi', \phi'']$, with the figure showing one of many possibilities.

Suppose the claim of the lemma is not true and $x^b(\phi)\lambda \leq x^a(\phi)$ for some $\phi \in [\phi_0, \phi_1]$. Noting that $x^a(\phi_0) < \lambda x^b(\phi_0)$, let $\phi^{\sim} > \phi_0$ be the lowest productivity value at which $x^b(\phi)\lambda = x^a(\phi)$.[51] Clearly, $x^a(\phi)$ must be catching up to $x^b(\phi)\lambda$ to the left of $\phi^{\sim}$, so $\Gamma^b(\phi) < \Gamma^a(\phi)$ on some interval (ϕ', ϕ''), with $\phi' < \phi^{\sim} \leq \phi''$, $\Gamma^b(\phi'') = \Gamma^a(\phi'')$ and $\Gamma_{\phi}^b(\phi'') \geq \Gamma_{\phi}^a(\phi'')$. This situation is depicted in figure 10.

As $x^b(\phi'')\lambda \leq x^a(\phi'')$ and $h^a(\phi'', \Gamma^a(\phi'')) = h^b\left(\phi'', \Gamma^b(\phi'')\right)$, then $Num^b(\phi'') < Num^a(\phi'')$. This observation, expression (37), and $\Gamma_{\phi}^b(\phi'') \geq \Gamma_{\phi}^a(\phi'')$ yield $Dem^b < Dem^a$. In addition, $x^b(\phi'')\lambda \leq x^a(\phi'')$, and expression (37) imply $\frac{\Gamma_{\phi}^b(\phi'')/\Gamma_{\phi}^b(\phi_0)}{\Gamma_{\phi}^a(\phi'')/\Gamma_{\phi}^a(\phi_0)} < 1$, which, to-

[51] Note that $\phi^{\sim}$ is well defined due to the continuity of the functions x^a and x^b.

gether with $\Gamma^b_\phi(\phi'') \geq \Gamma^a_\phi(\phi'')$, implies $\Gamma^b_\phi(\phi_0) > \Gamma^a_\phi(\phi_0)$. This last observation implies $\Gamma^b(\phi) > \Gamma^a(\phi)$ on some neighborhood of ϕ_0 (excluding ϕ_0), so let ϕ_- be the lowest productivity value to the right of ϕ_0 such that $\Gamma^b(\phi^-) = \Gamma^a(\phi^-)$. As $\Gamma^b(\phi) > \Gamma^a(\phi)$ on (ϕ_0, ϕ_-), we have $\Gamma^b_\phi(\phi_-) \leq \Gamma^a_\phi(\phi_-)$ and $x^b(\phi_-) > x^a(\phi_-)$. Using these results and (37) yields $Dem^b > Dem^a$, contradicting our previous finding. Then it must be the case that $x^b(\phi)\lambda > x^a(\phi)$ for all for all $\phi \in [\phi_0, \phi_1]$.

Lemma 4.v. As in the case of lemma 4.iii, the idea of the proof is to show that Γ^a and Γ^b can be thought of as the inverse matching functions of two artificial economies, and then use this additional information to prove the result. These artificial economies are defined as in the proof of lemma 4.iii.

The same argument used in the proof of lemma 4.iii implies that if p^i, $r^{d,i}$ and H^i are the price, domestic revenue and inverse-matching functions of the economy described above, then $H^i = \Gamma^i$. In addition, equation (35) also holds in this economy, which can be differentiated with respect to the limit of integration to get

$$r^{d,i}(\phi)\,\delta^i(\phi)\,g(\phi)\overline{M} = \frac{\sigma}{\sigma-1} L w^i\left(H^i(\phi)\right) V\left(H^i(\phi)\right) H^i_\phi(\phi) \text{ for } i = a, b, \tag{38}$$

where $\delta^i(\phi)$ was defined in the statement of the lemma. Below, I use this expression to assess the relative sizes of $r^{d,i}(\phi_0)$ for $i = a, b$, which delivers them main result as an immediate corollary.

STEP 1: *Let Φ^* be the set of productivity levels given by $\Phi^* = \{\phi \in [\phi_0, \phi_1] : H^b(\phi) = H^a(\phi),\ H^b_\phi(\phi) \leq H^a_\phi(\phi)\}$, and let S^* denote the set of corresponding skill levels, $S^* \equiv \{s \in [s_0, s_1] : s = H^i(\phi)$ for some $\phi \in \Phi^*\}$. Then, $w^b(s) < w^a(s)$ for some $s \in S^*$.*

Suppose that this is not the case and $w^b(s) \geq w^a(s)$ for all $s \in S^*$ and let N^i be the matching function of the artificial economy described above, that is, N^i is the inverse function of H^i. The gist of the argument can be laid out with the help of figure 10, in which $\Phi^* = \{\phi_-, \phi_1\}$. Letting $s_- = H^i(\phi_-)$ and $s'' = H^i(\phi'')$, note that the relative postion of N^a and N^b to the left and right of s_-, the strict log-supermodularity of A, and equation (10) imply that $w^b(s)$ grows faster than $w^a(s)$ on (s_-, s'') and slower on (s_0, s_-). This observation, together with $w^b(\phi_-) \geq w^a(\phi_-)$, implies $w^b(\phi) > w^a(\phi)$ on $[s_0, s_-) \cup (s_-, s'']$. A similar argument about the relative positions of N^a and N^b to the left of s_1 yields $w^b(\phi) > w^a(\phi)$ on$[s'', s_1)$. However, these results are inconsistent with the same numeraire condition in both economies, so the claim in step 1 must be true.

STEP 2: *Economies a and b satisfy* $r^{d,b}(\phi_0) < r^{d,a}(\phi_0)$.

Let $s_- \equiv \inf S^* \in S^*$ and note that $N^b(s) \leq N^a(s)$ for $s \in [s_0, s_-]$. If this was not the case and $N^b(s') > N^a(s')$ for some $s' \in (s_0, s_-)$, then the first intersection of N^a and N^b to the left of s' would belong to S^*, constradicting the definition of s_-. Next, I show that $r^{d,b}(\phi_0) < r^{d,a}(\phi_0)$ regardless of the relative sizes $w^b(s_-)$ and $w^a(s_-)$.

Suppose $w^b(s_-) \leq w^a(s_-)$. This condition and the definition of s_- imply that the right-hand side of (38) at s_- is greater in the a economy, so $r^{d,b}(\phi_0)\, x^b(\phi_-)\, \delta^b(\phi_-) \leq r^{d,a}(\phi_0)\, x^a(\phi_-)\, \delta^a(\phi_-)$. As $N^b(s) \leq N^a(s)$ on $[s_0, s_-]$ implies $x^b(\phi_-) \geq x^a(\phi_-)$, while $\delta^b(\phi_-) > \delta^a(\phi_-)$ by assumption, the last inequality implies $r^{d,b}(\phi_0) < r^{d,a}(\phi_0)$.

Now suppose $w^b(s_-) > w^a(s_-)$. As $w^a(s)$ is growing faster that $w^b(s)$ on $s \in [s_0, s_-]$ and $w^b(s_-) > w^a(s_-)$, then it must be the case that $w^b(s_0) > w^a(s_0)$. Per step 1, there is a $\phi_+ \in S^*$ with $w^b(s_+) < w^a(s_+)$, so equation (38) implies $r^{d,b}(\phi_0)\, x^b(\phi_+)\, \delta^b(\phi_+) < r^{d,a}(\phi_0)\, x^a(\phi_+)\, \delta^a(\phi_+)$. In addition, note that $\frac{w^b(s_+)}{w^b(s_0)} < \frac{w^a(s_+)}{w^a(s_0)}$ implies that the first term in the right-hand side of the following expression is larger in economy a,

$$\ln \frac{A(s_+, \phi_+)}{A(s_0, \phi_0)} = \int_{s_0}^{s_+} \frac{\partial \ln A(u, N^i(u))}{\partial s} du + \int_{\phi_0}^{\phi_+} \frac{\partial \ln A(H^i(t), t)}{\partial \phi} dt, \text{ for } i = a, b.$$

As the left-hand side of the last expression is the same in both economies, this observation implies that the second term in the right-hand side must be larger in economy b. With the latter term proportional to $\ln x^i(\phi_+)$, we have $x^b(\phi_+) > x^a(\phi_+)$. These observations, toghether with $\delta^b(\phi_+) > \delta^a(\phi_+)$, imply $r^{d,b}(\phi_0) < r^{d,a}(\phi_0)$.

Step 2, the numeraire assumption, and equation (35) evaluated at $\phi' = \phi_1$ yield inequality (32).

Lemma 4.vi. I consider the case in which $\eta_0(t, \lambda)$ is strictly decreasing in λ. As a first step, note that there is no $\phi' \in (\phi_0, \phi_1]$ such that $\Gamma^a(\phi) \leq \Gamma^b(\phi)$ for all $\phi \in (\phi_0, \phi']$. Suppose to the contrary that there is such a value $\phi' \in (\phi_0, \phi_1]$ and let ϕ_+ be the first time the functions Γ^a and Γ^b intersect to the right of ϕ'. By assumption, $\Gamma^a(\phi) \leq \Gamma^b(\phi)$ for $\phi \in (\phi_0, \phi_+)$, so $\frac{\Gamma^a_\phi(\phi_+)/\Gamma^a_\phi(\phi_0)}{\Gamma^b_\phi(\phi_+)/\Gamma^b_\phi(\phi_0)} \geq 1$. In addition, differentiating (24) yields

$$\frac{\Gamma^a_\phi(\phi_+)/\Gamma^a_\phi(\phi_0)}{\Gamma^b_\phi(\phi_+)/\Gamma^b_\phi(\phi_0)} \leq \left[\frac{x^a(\phi_+)}{x^b(\phi_+)}\right]^{\frac{\sigma}{\sigma-1}} \times \frac{\exp\{\int_{\phi_0}^{\phi_+} \eta^0(K_0^b x^b(\phi), \lambda) K_0^b x^b_\phi(\phi) d\phi\}}{\exp\{\int_{\phi_0}^{\phi_+} \eta^0(K_0^b x^b(\phi), 1) K_0^b x^b_\phi(\phi) d\phi\}}.$$

The strict log-supermodularity of A and $\Gamma^a(\phi) \leq \Gamma^b(\phi)$ for $\phi \in (\phi_0, \phi_+)$ yield $x^b(\phi_+) \geq x^a(\phi_+)$, while the fact that η^0 is strictly decreasing in λ implies that the second term on the right-hand side of the last expression is strictly less than one—i.e., $\frac{\Gamma^a_\phi(\phi_+)/\Gamma^a_\phi(\phi_0)}{\Gamma^b_\phi(\phi_+)/\Gamma^b_\phi(\phi_0)} < 1$. However, this contradicts our previous result, so the initial assumption must be false.

Next, I show that $\Gamma^a_\phi(\phi_0) > \Gamma^b_\phi(\phi_0)$. Per the previous result, $\Gamma^a_\phi(\phi_0) \geq \Gamma^b_\phi(\phi_0)$, so suppose for a moment that $\Gamma^a_\phi(\phi_0) = \Gamma^b_\phi(\phi_0) = \gamma_0$. As both BVPs have the same boundary conditions, equations (20a)-(20b) imply $x^i_\phi(\phi) = \frac{(\sigma-1)\partial \ln A(s_0,\phi_0)}{\partial\phi}$ and $\frac{z^i_\phi(\phi_0)}{z^i(\phi_0)} = -\frac{\partial \ln A(s_0,\phi_0)}{\partial\phi}$. Log-differentiating both sides of equation (20c) and evaluating at ϕ_0 yields

$$\Gamma^a_{\phi\phi}(\phi_0) - \Gamma^b_{\phi\phi}(\phi_0) = K^b_0 \frac{(\sigma-1)\partial \ln A(s_0,\phi_0)}{\partial\phi}\gamma_0 \left\{ \frac{F_y\left(K^b_0\lambda\right)\lambda K_1}{\left[1+F\left(K^b_0\lambda\right)K_1\right]} - \frac{F_y\left(K^b_0\right)K_1}{\left[1+F\left(K^b_0\right)K_1\right]} \right\} < 0,$$

where the inequality follows from $\eta^0\left(K^b_0,\lambda\right) < \eta^0\left(K^b_0,1\right)$. The last expression implies that there is some $\phi' \in (\phi_0,\phi_1]$ such that $\Gamma^a_\phi(\phi) < \Gamma^b_\phi(\phi)$ on $(\phi_0,\phi']$, which contradicts our first result. Accordingly, we must have $\Gamma^a_\phi(\phi_0) > \Gamma^b_\phi(\phi_0)$.

Note that $\Gamma^a_\phi(\phi_0) > \Gamma^b_\phi(\phi_0)$ implies $\Gamma^a(\phi) > \Gamma^b(\phi)$ on $\left(\phi_0,\phi_+\right)$, where ϕ_+ is the first intersection of Γ^a and Γ^b to the right of ϕ_0. Now I show that $\phi_+ = \phi_1$, which is the desired result. Suppose for a moment that $\phi_+ < \phi_1$. If we define on $[\phi_+,\phi_1]$, $w^i(\phi) \equiv x^i(\phi)/x^i\left(\phi_+\right)$ and $y^i(\phi) = z^i(\phi)/x^i\left(\phi_+\right)$, then it is readily seen that $\{y^i, w^i(\phi), \Gamma^i\}$ solve BVP (20) in said interval, with $\{\alpha^i(\phi), K^i_1\} = \{\alpha(\phi), K_1\}$ and parameter $\overline{K}^i_0 = K^i_0 x^i\left(\phi_+\right)$. Note that $\Gamma^a(\phi) > \Gamma^b(\phi)$ on $\left(\phi_0,\phi_+\right)$ yields $x^a\left(\phi_+\right) > x^b\left(\phi_+\right)$, so $\overline{K}^a_0 > \overline{K}^b_0$. Then, the BVPs associated to $\{y^i, w^i(\phi), \Gamma^i\}$ satisfy the conditions of lemma 4.vi, so the result in the previous paragraph yields $\Gamma^a_\phi\left(\phi_+\right) > \Gamma^b_\phi\left(\phi_+\right)$. However, $\Gamma^a(\phi) > \Gamma^b(\phi)$ on $\left(\phi_0,\phi_+\right)$ implies $\Gamma^a_\phi\left(\phi_+\right) \leq \Gamma^b_\phi\left(\phi_+\right)$, so assuming $\phi_+ < \phi_1$ yields a contradiction.

Lemma 4.vii. I consider the case in which $\eta_1(t,\lambda)$ is strictly increasing in λ. First, note that there is no $\phi' \in (\phi_0,\phi_1]$ such that $\Gamma^a(\phi) \geq \Gamma^b(\phi)$ for all $\phi \in (\phi_0,\phi']$. To see this, suppose to the contrary that there is such a value $\phi' \in (\phi_0,\phi_1]$ and let ϕ_+ be the first time the functions Γ^a and Γ^b intersect to the right of ϕ'. As in the proof of lemma 4.vii, the relative positions of Γ^a and Γ^b imply $\frac{\Gamma^a_\phi(\phi_+)/\Gamma^a_\phi(\phi_0)}{\Gamma^b_\phi(\phi_+)/\Gamma^b_\phi(\phi_0)} \leq 1$, but computing Γ^i_ϕ from equation (24) yields

$$\frac{\Gamma^a_\phi(\phi_+)/\Gamma^a_\phi(\phi_0)}{\Gamma^b_\phi(\phi_+)/\Gamma^b_\phi(\phi_0)} \geq \left[\frac{x^a\left(\phi_+\right)}{x^b\left(\phi_+\right)}\right]^{\frac{\sigma}{\sigma-1}} \times \frac{\exp\{\int_{\phi_0}^{\phi_+}\eta^1\left(K^b_0 x^b(\phi),\lambda\right)K^b_0 x^b_\phi(\phi)d\phi\}}{\exp\{\int_{\phi_0}^{\phi_+}\eta^1\left(K^b_0 x^b(\phi),1\right)K^b_0 x^b_\phi(\phi)d\phi\}} > 1,$$

where the second strict inequality follows from the fact that η^1 is strictly increasing in λ.

Second, I show $\Gamma^a_\phi(\phi_0) < \Gamma^b_\phi(\phi_0)$. Per the previous result, $\Gamma^a_\phi(\phi_0) \leq \Gamma^b_\phi(\phi_0)$, so suppose for a moment that $\Gamma^a_\phi(\phi_0) = \Gamma^b_\phi(\phi_0) = \gamma_0$. As both BVPs have the same boundary conditions,equations (20a)-(20b) imply $x^i_\phi(\phi) = \frac{(\sigma-1)\partial \ln A(s_0,\phi_0)}{\partial\phi}$ and $\frac{z^i_\phi(\phi_0)}{z^i(\phi_0)} = -\frac{\partial \ln A(s_0,\phi_0)}{\partial\phi}$. Log-differentiating both sides of equation (20c) and evaluating at ϕ_0 yields

$$\Gamma^a_{\phi\phi}(\phi_0) - \Gamma^b_{\phi\phi}(\phi_0) = K^b_0 \frac{(\sigma-1)\partial \ln A(s_0,\phi_0)}{\partial\phi}\gamma_0 \left\{ \frac{F_y\left(K^b_0\lambda\right)K^b_1\lambda^2}{\left[1+F\left(K^b_0\lambda\right)\lambda K^b_1\right]} - \frac{F_y\left(K^b_0\right)K^b_1}{\left[1+F\left(K^b_0\right)K^b_1\right]} \right\} > 0,$$

where the inequality follows from $\eta^1\left(K^b_0,\lambda\right) > \eta^1\left(K^b_0,1\right)$. The last expression implies that there is some $\phi' \in (\phi_0,\phi_1]$ such that $\Gamma^a_\phi(\phi) > \Gamma^b_\phi(\phi)$ on $(\phi_0,\phi']$, which contradicts our first result. Accordingly, we must have $\Gamma^a_\phi(\phi_0) < \Gamma^b_\phi(\phi_0)$.

Note that $\Gamma^a_\phi(\phi_0) < \Gamma^b_\phi(\phi_0)$ implies $\Gamma^a(\phi) < \Gamma^b(\phi)$ on $\left(\phi_0,\phi_+\right)$, where ϕ_+ is the first intersection of Γ^a and Γ^b to the right of ϕ_0. Now I show that $\phi_+ = \phi_1$, which is the desired result. Suppose for a moment that $\phi_+ < \phi_1$. If we define on $[\phi_+,\phi_1]$, $w^i(\phi) \equiv x^i(\phi)/x^i\left(\phi_+\right)$ and $y^i(\phi) = z^i(\phi)/x^i\left(\phi_+\right)$, then it is readily seen that $\{y^i, w^i(\phi), \Gamma^i\}$ solve BVP (20) in said interval, with $\alpha^i(\phi) = \alpha(\phi)$ and parameter $\overline{K}^i_0 = K^i_0 x^i\left(\phi_+\right)$. That is, $\overline{K}^a_0 = \lambda_1\overline{K}^b_0$, where $\lambda_1 \equiv \frac{\lambda x^a\left(\phi_+\right)}{x^b\left(\phi_+\right)} > 1$, as the BVPs associated with Γ^i satisfy the conditions of lemma 4.iv on $[\phi_0,\phi_+]$. In addition, $\Gamma^a(\phi) < \Gamma^b(\phi)$ on $\left(\phi_0,\phi_+\right)$ implies $x^a\left(\phi_+\right) < x^b\left(\phi_+\right)$, so $\lambda_1 < \lambda$.

The previous discussion implies that the BVPs that $\{y^i, w^i(\phi), \Gamma^i\}$ for $i = a, b$ solve on $\left[\phi_+,\phi_1\right]$ differ only in the parameters $\{\overline{K}^i_0, K^i_1\}$, with $\overline{K}^a_0 = \lambda_1\overline{K}^b_0$ and $K^a_1 = \lambda K^b_1$. To understand the implication of this difference, it is convenient to consider a third BVP on $\left[\phi_+,\phi_1\right]$ differing from the previous two only in the parameters $\{\overline{K}^c_0, K^c_1\}$, with $\overline{K}^c_0 = \overline{K}^a_0 = \lambda_1\overline{K}^b_0$ and $K^c_1 = \lambda_1 K^b_1$. Given these definitions, note that the BVPs associated with $\left\{y^b, w^b(\phi), \Gamma^b\right\}$ and $\{y^c, w^c(\phi), \Gamma^c\}$ satisfy the conditions in lemma 4.vii, so our previous results imply $\Gamma^c_\phi\left(\phi_+\right) < \Gamma^b_\phi\left(\phi_+\right)$. In addition, the BVPs associated to $\{y^a, w^a(\phi), \Gamma^a\}$ and $\{y^c, w^c(\phi), \Gamma^c\}$ satisfy the assumptions of 4.ii with $K^a_1 > K^c_1$, so $\Gamma^a_\phi\left(\phi_+\right) < \Gamma^c_\phi\left(\phi_+\right)$. These inequalities yield $\Gamma^a_\phi\left(\phi_+\right) < \Gamma^b_\phi\left(\phi_+\right)$. However, $\Gamma^a(\phi) < \Gamma^b(\phi)$ on $\left(\phi_0,\phi_+\right)$ implies $\Gamma^a_\phi\left(\phi_+\right) \geq \Gamma^b_\phi\left(\phi_+\right)$, which is a contradiction. Then it must be the case that $\phi_+ = \phi_1$. ■

B.4 Section 5

B.4.1 Proof of Proposition 2

Let us start with the proof of $\phi^*_a < \phi^*_\tau$. For any $\phi^* \in [\underline{\phi},\overline{\phi}]$, let $\left\{\overline{p}(.;\phi^*), \overline{r}^d(.;\phi^*), \overline{H}(.;\phi^*)\right\}$ and $\left\{p(.;\phi^*), r^d(.;\phi^*), H(.;\phi^*)\right\}$ denote, respectively, the solution to the BVPs of the open and closed economies with activity cutoff ϕ^*, where the notation emphasizes the dependence of the solution on ϕ^*. In the sense described in section 4.2, these BVPs are equivalent to BVP (20) with $K_1 = 0$ and $\alpha(\phi;\phi^*) = 1$ for the closed-economy BVP and $\overline{K}_1 = 0$ and $\overline{\alpha}(\phi;\phi^*) = \left[1 + F\left(\frac{\tau^{1-\sigma}}{\sigma f_x}\overline{r}^d(\phi;\phi^*)\right) n\tau^{1-\sigma}\right]$ for the open-economy BVP. As $\overline{\alpha}(\phi;\phi^*)$ is increasing, these two parameterizations of BVP (20) satisfy the conditions of lemma 4.i, so $\overline{H}(\phi;\phi^*) < H(\phi;\phi^*)$ for all $\phi \in (\phi^*, \overline{\phi})$. In turn, this result implies that these BVPs satisfy the assumptions of lemma 4.iii, so

$\int_{\phi^*}^{\overline{\phi}} \overline{x}(\phi;\phi^*)\frac{\overline{\alpha}(\phi;\phi^*)}{\overline{\alpha}(\phi^*;\phi^*)} g(\phi)\, d\phi > \int_{\phi^*}^{\overline{\phi}} x(\phi;\phi^*) g(\phi)\, d\phi,$

where $\left\{\overline{z}(.;\phi^*), \overline{x}(.;\phi^*), \overline{H}(.;\phi^*)\right\}$ and $\{z(.;\phi^*), x(.;\phi^*), H(.;\phi^*)\}$ are the respective solutions to the paramerizations of BVP (20) discussed above. After some algebraic manipulation, the last inequality yields $\beta(\overline{r}(.;\phi_a^*), \phi_a^*) > \beta^a\left(r^d(.;\phi_a^*), \phi_a^*\right) = L$, where β and β^a are the functions defined by the left-hand sides of equations (19) and (14) described in proposition 1. Per the discussion leading to proposition 1, these functions are strictly decreasing in the value of the parameter ϕ^*, so we must have $\phi_a^* < \phi_\tau^*$ for equation (19) to hold in the open economy.

Let us now prove $N^\tau(s) > N^a(s)$ for all $s \in [\underline{s}, \overline{s})$ and proposition 2.ii. Let $N(s;\phi^*)$ be the inverse function of $H(\phi;\phi^*)$. Following the discussion above, these results can be easily proved by decomposing the total effect on the matching function into that of the increase in the exit cutoff (intensive-margin channel) and that of having an increasing share of exporters at each productivity level in the open economy (extensive-margin channel). Starting with the former, the no-crossing result in lemma 2.i and $\phi_a^* < \phi_\tau^*$ imply $N^a(s) = N(s;\phi_a^*) < N(s;\phi_\tau^*)$ on $[\underline{s}, \overline{s})$. Bringing the effects of exporters into the picture, lemma 4.i implies that $H(\phi;\phi_\tau^*) > \overline{H}(\phi;\phi_\tau^*) = H^\tau(\phi)$ on $\left(\phi_\tau^*, \overline{\phi}\right)$, or $N(s;\phi_\tau^*) < \overline{N}(s;\phi_\tau^*) = N^\tau(s)$ on $(\underline{s}, \overline{s})$. Combining these observations yields the desired result.

B.4.2 Proof of Proposition 3

Proposition 3.i. Let us start with the proof of $\phi_h^* < \phi_l^*$. For any $\phi^* \in [\underline{\phi}, \overline{\phi}]$ and $i = l, h$, let $\{\overline{p}^i(.;\phi^*), \overline{r}^{d,i}(.;\phi^*), \overline{H}^i(;\phi^*)\}$ denote the solution to the BVP of the open economy described in lemma 3.iii with variable trade costs τ_i and activity cutoff ϕ^*. In the sense described in section 4.2, these BVPs are equivalent to BVP (20) with $K_0^i = \frac{f}{f_x}\tau_i^{1-\sigma}$, $K_1^i = n\tau_i^{1-\sigma}$, and $\alpha^i(\phi;\phi^*) = 1$. As $K_0^l = \lambda K_0^h$ and $K_1^l = \lambda K_1^h$ with $\lambda =$

$(\tau_l/\tau_h)^{1-\sigma} > 1$, these parameterization of BVP (20) satisfy the conditions of lemma 4.iv, so $\overline{x}^l(\phi;\phi^*)\,\lambda > \overline{x}^h(\phi;\phi^*)$ or $\overline{r}^{d,l}(\phi;\phi^*)\,\tau_l^{1-\sigma} > \overline{r}^{d,h}(\phi;\phi^*)\,\tau_h^{1-\sigma}$ for all $\phi \in [\phi^*, \overline{\phi}]$, where $\{\overline{z}^i(.;\phi^*), \overline{x}^i(.;\phi^*), \overline{H}^i(.;\phi^*)\}$ for $i = l, h$ are the solutions to the parameterizations of BVP (20) discussed above.

Defining $\delta^i(\phi) \equiv \left[1 + F\left(K_0^i \overline{x}^i(\phi;\phi^*)\right) K_1^i\right]$, the previous result yields $\delta^l(\phi) > \delta^h(\phi)$. As such, these BVPs satisfy the assumptions of lemma 4.v, so $R^l(\phi^*) > R^h(\phi^*)$, where $R^i(\phi^*) \equiv \int_{\phi^*}^{\overline{\phi}} \overline{r}^{d,i}(\phi;\phi^*)\,\delta^i(\phi)\,g(\phi)\,d\phi\overline{M}$. After some algebraic manipulation, these results yield $\beta^l\left(\overline{r}^{d,l}(.;\phi_h^*), \phi_h^*\right) > \beta^h\left(\overline{r}^{d,h}(.;\phi_h^*), \phi_h^*\right) = L$, where β^i is defined as in proposition 1. Per the discussion leading to proposition 1, β^l is strictly decreasing in the value of the parameter ϕ^*, so we must have $\phi_h^* < \phi_l^*$ for equation (19) to hold for $\tau = \tau^l$.

Finally, the continuity of the matching functions and $\phi_h^* < \phi_l^*$ imply that there is a skill level $s' \in (\underline{s}, \overline{s}]$ such that $N^l(s) > N^h(s)$ on $[\underline{s}, s')$—i.e., inequality necessarily increases among the least-skilled workers in the economy after a trade liberalization.

Proposition 3.ii. As discussed in the text, the distributional effects of the extensive-margin channel are theoretically ambiguous, so here I derive the impact on relative wages of the other two channels, the selection-into-activity and the intensive-margin channels. Let $\{z(\phi;\phi^*,\alpha),\ x(\phi;\phi^*,\alpha),\ H(\phi;\phi^*,\alpha)\}$ denote the unique solution to BVP (20) with constant $K_1 = 0$, parameter function α, and boundary conditions $\{x(\phi^*) = 1, H(\phi^*) = \underline{s}, H(\overline{\phi}) = \overline{s}\}$, where the notation emphasizes the dependence of the solution on $\{\phi^*, \alpha\}$. In addition, $N(\phi;\phi^*,\alpha)$ denotes the inverse of $H(\phi;\phi^*,\alpha)$. For $i = l, h$, let $\left\{\phi_i^*, p^i, r^{d,i}, H^i\right\}$ be the activity cutoff, price, domestic revenue and inverse-matching functions of the two open economies in the statement of the proposition (these economies differ only in the variable trade costs they face, with $\tau_l < \tau_h$). Defining $\alpha^i(\phi) \equiv \left[1 + F\left(\frac{\tau_i^{1-\sigma}}{\sigma f_x} r^{d,i}(\phi)\right) n\tau_i^{1-\sigma}\right]$ for $i = l, h$, the open-economy BVPs associated with each H^i are equivalent to parameterizations of BVP (20) with $K_1 = 0$ and $\alpha = \alpha^i$.[52] We are ready to prove the main results.

Let us start with the **selection-into-activity channel**. As discussed in the text, the matching functions N^0 and N^h in figure 2 differ only in their activity cutoffs—i.e., $N^0 = N\left(\phi;\phi_l^*,\alpha^h\right)$ and $N^h = N\left(\phi;\phi_h^*,\alpha^h\right)$. Accordingly, the no-crossing result in lemma 2.i implies $N^0(s) > N^h(s)$ on $[\underline{s}, \overline{s})$. As the economies associated with N_0 and N^h have the same fraction of exporters at each productivity and face the same variable costs (same α),

[52] See the proof of proposition 2.

their difference isolates the effects of the selection-into-activity channel on relative wages.

Let us now turn to the **intensive-margin channel**. Define $\alpha^1(\phi) \equiv 1 + F\left(\frac{\tau_h^{1-\sigma}}{\sigma f_x} r^{d,h}(\phi)\right) n\tau_l^{1-\sigma}$ and note that $\alpha^1(\phi)$ differs from $\alpha^h(\phi)$ only in the value of the variable trade cost outside the function F. As discussed in the text, the matching functions N_0 and N_1 in figure 2 differ only in their parameter function α—i.e., $N^0 = N\left(\phi; \phi_l^*, \alpha^h\right)$ and $N^1 = N\left(\phi; \phi_l^*, \alpha^1\right)$. In addition, note that for any pair $\phi'', \phi' \in [\phi^*, \overline{\phi}]$ such that $\phi'' > \phi'$ and $F\left(\tau_h^{1-\sigma} r^{d,h}(\phi'')/\sigma f_x\right) > 0$, we have $\alpha^1(\phi'')/\alpha^1(\phi') > \alpha^h(\phi'')/\alpha^h(\phi')$. Accordingly, the BVPs associated with N^0 and N^1 satisfy the conditions of lemma 4.i, so $N^1(s) > N^0(s)$ on $[\underline{s}, \overline{s})$.

Proposition 3.iii. To prove the result, it is convenient to break the changes in the BVP of the open economy introduced by the liberalization in two parts, the change associated to the decline in variable trade costs and the change associated to the rise in the activity cutoff (allowing the set of exporters to adjust in each case). Starting with the former, let N^0 be the matching function resulting from reducing τ_h to τ_l in the BVP of the open economy before the liberalization, keeping the activity cutoff unchanged. If the assumption on η_1^F is satisfied, then it is readily seen that F and the open-economy BVPs associated with N^h and N^0 satisfy the conditions in lemma 4.vii with $K_0^h = f\tau_h^{1-\sigma}/f_x$, $K_1^h = n\tau_h^{1-\sigma}$, $K_i^0 = \lambda K_i^h$, and $\lambda = (\tau_l/\tau_h)^{1-\sigma} > 1$. Accordingly, $N^0(s) > N^h(s)$ on $(\underline{s}, \overline{s})$.

Now consider the change in the matching function associated with the rise in the activity cutoff—i.e., the difference between N^0 and N^l. Suppose that N^0 and N^l intersect on $(\underline{s}, \overline{s})$ with the first intersection occurring at s', namely $N^0(s') = N^l(s') = \phi'$. If for $i = 0, l$, we define on $\left[\phi', \overline{\phi}\right]$ the functions $w^i(\phi) \equiv \frac{r^{d,i}(\phi)}{r^{d,i}(\phi')}$ and $y^i(\phi) \equiv \frac{p^i(\phi)}{r^{d,i}(\phi')\overline{M}}[L - fM^i - \int_{\phi_i^*}^{\overline{\phi}} nf_x \int_{\underline{y}}^{\frac{r^{d,i}(\phi)\tau_l^{1-\sigma}}{\sigma f_x}} ydF(y)\, g(\phi)\, \overline{M} d\phi]$, then $\{w^i, y^i, H^i\}$ is the unique solution to BVP (20) with parameters $\alpha^i(\phi) = 1$, $K_0^i = \frac{r^{d,i}(\phi')\tau_l^{1-\sigma}}{\sigma f}$, $K_1^i = n\tau_l^{1-\sigma}$ and boundary conditions $w^i(\phi') = 1$, $H^i(\phi') = s'$ and $H^i\left(\overline{\phi}\right) = \overline{s}$. In addition, note that the log-supermodularity of A and $H^l(\phi) < H^0(\phi)$ on $[\phi_l^*, \phi')$ implies $r^{d,0}(\phi') > r^{d,l}(\phi')$, so $K_0^0 > K_0^l$. Accordingly, if the assumption on η_0^F is satisfied, then its is readily seen that F and the open-economy BVPs associated with N^l and N^0 satisfy the conditions of lemma 4.vi on $\left[\phi', \overline{\phi}\right]$, so $H_\phi^l(\phi') < H_\phi^0(\phi')$. However, $H^l(\phi) < H^0(\phi)$ on $[\phi_l^*, \phi')$ implies $H_\phi^l(\phi') \geq H_\phi^0(\phi')$, which is a contradiction. Then it must be the case that N^l and N^0 do not intersect on $(\underline{s}, \overline{s})$, so N^l lies strictly above N^0 on $[\underline{s}, \overline{s})$ as shown in the picture.

Combining the last two results we get $N^l(s) > N^h(s)$ on $[\underline{s}, \overline{s})$, so inequality is pervasively higher after the liberalization. This concludes the proof of the proposition.

B.5 Section 6

B.5.1 Proof of Proposition 4

In the free-entry model the activity cutoff may rise or fall when the economy starts trading. The reasons behind this ambiguity are discussed in the text. In addition, as stated in the text, proposition 4.i considers essentially the same case as proposition 2. As such, here I focus on Proposition 4.ii.

Proposition 4.ii. Let $\phi^*_\tau < \phi^*_a$. If $N^\tau(s) < N^a(s)$ for all $s \in [\underline{s}, \overline{s})$, then the strict log-supermodularity of A and equation (12) implies that $r^{d,\tau}(\phi) > r^{d,a}(\phi)$ for all $\phi \geq \phi^*_a$, so domestic profits in the open economy are necessarily higher than in autarky. With strictly positive export profits, this observation implies that total average profits must be higher in the open economy, violating the free entry condition (22). Accordingly, $N^\tau(s)$ must lie above $N^a(s)$ for some values of s, implying that $N^\tau(s)$ and $N^a(s)$ must intersect at least once on $(\underline{s}, \overline{s})$.

Next, I show that $N^\tau(s)$ and $N^a(s)$ intersect exactly once on $(\underline{s}, \overline{s})$. The argument is more easily stated in terms of the inverse functions H^τ and H^a. Let ϕ_0 be the first time that H^τ and H^a intersect on $(\phi^*_a, \overline{\phi})$. Note that the restrictions of H^τ and H^a on $[\phi_0, \overline{\phi}]$ are the unique solutions to parameterizations of BVP (20) that differ only in the parameter function α^i, with $K^i_1 = 0$ for $i = \tau, a$, $\alpha^\tau(\phi) = 1 + F\left(\frac{r^{d,\tau}(\phi)}{\sigma f_x}\tau^{1-\sigma}\right)$ and $\alpha^a(\phi) = 1$. Then, an immediate application of lemma 4.i yields $H^\tau(\phi) < H^a(\phi)$ on $(\phi_0, \overline{\phi})$, so H^τ and H^a ($N^\tau(s)$ and $N^a(s)$) intersect exactly once on $(\phi^*_a, \overline{\phi})$ ($(\underline{s}, \overline{s})$) at ϕ_0 ($s_0 = H^i(\phi_0)$).

The last result implies that, in the open economy, inequality is lower among workers with skill levels below s_0, but higher among workers with skill level above s_0. Put another way, opening to trade leads to wage polarization. The effects of the intensive- and extensive-margin channels can be proved by adapting the arguments in proposition 2.ii.