International Trade and Labor Markets: Unemployment, Inequality and Redistribution

A dissertation presented

by

Oleg Itskhoki

to

The Department of Economics

in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the subject of Economics

Harvard University
Cambridge, Massachusetts

April 2009
International Trade and Labor Markets: 
Unemployment, Inequality and Redistribution

International trade is typically believed to lead to aggregate welfare gains for trading countries. However, it is also often viewed as a source of growing social disparity—by causing unemployment and greater inequality within countries—which calls for an offsetting policy response. This dissertation consists of three theoretical essays studying these issues.

The first chapter develops a model of international trade with labor market frictions that differ across countries. We show that differences in labor market institutions constitute a source of comparative advantage and lead to trade between otherwise similar countries. Although trade ensures aggregate welfare gains for both countries, the more flexible country stands to gain proportionately more. An increase in the country’s labor market flexibility leads to welfare gains at home, but causes welfare losses in the trading partner via decreased competitiveness of foreign firms. Trade can increase or decrease unemployment by inducing an intersectoral labor reallocation generating rich patterns of unemployment.

The second chapter proposes a new framework for thinking about the distributional consequences of trade that incorporates firm and worker heterogeneity, search and matching frictions in the labor market, and screening of workers by firms. Larger firms pay higher wages and exporters pay higher wages than non-exporters. The opening of trade enhances wage inequality and raises unemployment, but expected welfare gains are ensured if workers are risk neutral. And while wage inequality is larger in a trade equilibrium than in autarky, reductions of trade impediments can either raise or reduce wage inequality.

Conventional wisdom suggests that the optimal policy response to rising income inequality is greater redistribution. The final chapter studies an economy in which trade is
associated with a costly entry into the foreign market, so that only the most productive agents can profitably export. In this model, trade integration simultaneously leads to rising income inequality and greater efficiency losses from taxation, both driven by the extensive margin of trade. As a result, the optimal policy response may be to reduce the marginal taxes, thereby further increasing inequality. In order to reap most of the welfare gains from trade, countries may need to accept increasing income inequality.
# TABLE OF CONTENTS

*Acknowledgments* ........................................................................ vii

1. *Labor Market Rigidities, Trade and Unemployment* ....................... 1
   1.1 Introduction .......................................................................... 1
   1.2 The Model ........................................................................... 5
      1.2.1 Preferences and Demand .................................................. 6
      1.2.2 Technologies and Market Structure ................................. 7
      1.2.3 Wages and Profits .......................................................... 9
      1.2.4 Labor Market ............................................................... 13
   1.3 Equilibrium Structure .......................................................... 17
   1.4 Trade, Welfare and Productivity .............................................. 23
      1.4.1 Welfare ......................................................................... 23
      1.4.2 Trade Structure ............................................................ 26
      1.4.3 Productivity ................................................................... 27
   1.5 Unemployment ..................................................................... 30
      1.5.1 Symmetric Countries ..................................................... 30
      1.5.2 Asymmetric Countries .................................................... 35
   1.6 Firing Costs and Unemployment Benefits .................................. 41
   1.7 Concluding Comments ........................................................ 43

2. *Inequality and Unemployment in a Global Economy* ....................... 45
   2.1 Introduction .......................................................................... 45
   2.2 Sectoral Equilibrium ............................................................ 50
      2.2.1 Model Setup ................................................................. 50
      2.2.2 Firm’s Problem ............................................................ 54
      2.2.3 Sectoral Variables ........................................................ 58
      2.2.4 Firm-specific Variables ................................................. 62
   2.3 Sectoral Wage Inequality ....................................................... 65
   2.4 Sectoral Unemployment ......................................................... 73
   2.5 Sectoral Income Inequality ...................................................... 77
   2.6 General Equilibrium ............................................................ 79
      2.6.1 Outside Sector and Risk Neutrality ................................. 80
      2.6.2 Single Differentiated Sector and Risk Neutrality .............. 83
      2.6.3 Outside Sector and Risk Aversion ................................... 86
   2.7 Conclusion .......................................................................... 88

3. *Optimal Redistribution in an Open Economy* ................................ 91
   3.1 Introduction .......................................................................... 91
   3.2 Closed Economy ................................................................... 97
ACKNOWLEDGMENTS

I am extremely fortunately to have had the honor of working very closely with Elhanan Helpman throughout my studies at Harvard: first as a student, then as a Teaching Fellow, and finally as a co-author. His contribution to this dissertation is evident – two of the three chapters of this dissertation are based on my joint work with Elhanan. It is impossible to overstate how much I have learned from working with Elhanan over the past three years. Learning-by-doing is extremely important in our profession, and I was privileged to witness in real time how the highest quality academic research is done – something which I believe is an absolutely essential part of academic training for any doctoral student.

My professional development would not be complete had it not been for Gita Gopinath, another teacher and co-author. I have not included the research we have done with Gita into this dissertation, but I am equally excited about it. Working with her helped me acquire a very important complementary set of skills. I am also grateful to Steve Redding and Roberto Rigobon, my two other co-authors, from whom I also learned a great deal.

I am very indebted to my advisor, Aleh Tsyvinski, for his guidance, support and all his effort he put into me throughout my student years. Pol Antràs’ advice and support was also extremely important, especially his encouragement during the tough times of the job market year.

I was very lucky to have come to Harvard at the time when a group of young professors joined the faculty, including Aleh Tsyvinski, Gita Gopinath, Pol Antràs and Manuel Amador. I have learned a great deal from them, both how to do outstanding economic research and how to survive in the highly competitive environment of academic economics. This experience is truly indispensable.

I also wish to thank Ken Rogoff and John Campbell whose exciting and very broad research and most importantly their ability to link state-of-the-art theory with practice and real-life problems provided inspiration in the beginning of my doctoral studies. I have
also learned a lot from interactions with Philippe Aghion, Alberto Alesina, Robert Barro, Emmanuel Farhi, Jordi Galí, Oliver Hart, Larry Katz, David Laibson, Greg Mankiw, Nathan Nunn, Ariel Pakes, Andrei Shleifer, Jeremy Stein and Jim Stock. Finally, I was very fortunate to have the invaluable experience of working as a Research Assistant for Daron Acemoglu at MIT.

Part of the work on this dissertation was done at the Tel Aviv University, the Centre for Economic Performance at the London School of Economics and at the Minneapolis Fed. I thank these institutions for hospitality and productive academic environment.

I have had the privilege of studying at Harvard with a number of great future economists. I wish to thank my TFs Ian Martin and Parag Pathak who provided a great example to follow; my teaching colleagues Filipe Campante, Giacomo Ponzetto and Elias Albagli; my office mates at different times Mihai Manea, Anthony Niblett and Thomas Baranga, and many of my classmates from whom I learned a lot. I wish to separately acknowledge and thank Loukas Karabarbounis, Matt Weinzierl and Elias Albagli for the very valuable comments on my research.

Harvard has given me the opportunity to meet hundreds of interesting people and make many great friends. Without their support the completion of this dissertation would not have been possible. I should mention separately Lena and Sasha, Erdin and Ellen, Artashes, and Keyu.

Finally, I thank my family—Mom and Dad, Olga, Dave and Max, and Grandma—for their support, their belief in me, and their joy for my achievements.

To my Albina, for her love and patience!
1. LABOR MARKET RIGIDITIES, TRADE AND UNEMPLOYMENT

(with Elhanan Helpman)

1.1 Introduction

International trade and international capital flows link national economies. Although such links are considered to be beneficial for the most part, they produce an interdependence that occasionally has harmful effects. In particular, shocks that emanate in one country may negatively impact trade partners. On the trade side, links through terms-of-trade movements have been studied extensively, and it is now well understood that, say, capital accumulation or technological change can worsen a trade partner’s terms of trade and reduce its welfare. On the macro side, the transmission of real business cycles has been widely studied, such as the impact of technology shocks in one country on income fluctuations in its trade partners.

Although a large literature addresses the relationship between trade and unemployment, we fall short of understanding how these links depend on labor market rigidities. Indeed, measures of labor market flexibility developed by Botero et al. (2004) differ greatly across countries. The rigidity of employment index, which is an average of three other indexes—difficulty of hiring, difficulty of firing, and rigidity of hours—shows wide variation in its range between zero and one hundred (where higher values represent larger rigidities). Importantly, countries with very different development levels may have similar labor market rigidities. For example, Chad, Morocco and Spain have indexes of 60, 63 and 63, respectively, which are about twice the average for the OECD countries (which

---

1 Their original data has been updated by the World Bank and is now available at http://www.doingbusiness.org/ExploreTopics/EmployingWorkers/. The numbers reported in the text come from this site, downloaded on May 20, 2007. Other measures of labor market characteristics are available for OECD countries; see Nickell (1997) and Blanchard and Wolfers (2000).
is 33.3) and higher than the average for sub-Saharan Africa. The United States has the lowest index, equal to zero, while Australia has an index of three and New Zealand has an index of seven, all significantly below the OECD average. Yet some of the much poorer countries also have very flexible labor markets, e.g., both Uganda and Togo have an index of seven.\footnote{There is growing awareness that institutions affect comparative advantage and trade flows. Levchenko (2007), Nunn (2007) and Costinot (2006) provide evidence on the impact of legal institutions, while Cuñat and Melitz (2007) and Chor (2006) provide evidence on the impact of labor market rigidities.}

We develop in this paper a two-country model of international trade in order to study the effects of labor market frictions on trade flows, productivity, welfare and unemployment. We are particularly interested in the impact of a country’s labor market rigidities on its trade partner, and the differential impact of lower trade impediments on countries with different labor market rigidities. Blanchard and Wolfers (2000) emphasize the need to allow for interactions between shocks and differences in labor market characteristics in order to explain the evolution of unemployment in European economies. They show that these interactions are empirically important. On the other side, Nickell et al. (2002) emphasize changes over time in labor market characteristics as important determinants of the evolution of unemployment in OECD countries. We focus the analysis on search and matching frictions in Sections 1.2-1.5, and discuss in Section 1.6 how the results generalizations to economies with firing costs and unemployment benefits.\footnote{While we use a static specification of labor market frictions, our analysis is consistent with a steady state of a dynamic model as we show in Helpman and Itskhole (2009).}

The literature on trade and unemployment is large and varied. One strand of this literature considers economies with minimum wages, of which Brecher (1974) represents an early contribution.\footnote{His approach has been extended by Davis (1998) to study how wages are determined when two countries trade with each other, one with and one without a minimum wage.} Another approach, due to Matusz (1986), uses implicit contracts. A third approach, exemplified by Copeland (1989), incorporates efficiency wages into trade models. Yet another line of research uses fair wages. Agell and Lundborg (1995) and Kreickemeier and Nelson (2006) illustrate this approach. The final approach uses search and matching in labor markets. While two early studies (Davidson, Martin and Matusz,}
1988, and Hosios, 1990) extended the two-sector model of Jones (1965) to economies with this type of labor market friction, Davidson, Martin and Matusz (1999) provide a particularly valuable analysis of international trade with labor markets that are characterized by Diamond-Mortensen-Pissarides-type search and matching frictions (see Pissarides, 2000, for the theory of search and matching in labor markets). In their model differences in labor market frictions, both across sectors and across countries, generate Ricardian-type comparative advantage.

Our two-sector model incorporates Diamond-Mortensen-Pissarides-type frictions into both sectors; one producing homogenous goods, the other producing differentiated products. In both sectors wages are determined by bargaining. There is perfect competition in homogeneous goods and monopolistic competition in differentiated products. In the differentiated-product sector firms are heterogeneous, as in Melitz (2003). These firms exercise market power in the product market on the one hand, and bargain with workers over wages on the other. Moreover, there are fixed and variable trade costs in the differentiated sector. We focus the analysis on the differentiated sector and think about the homogeneous sector as the rest of the economy. The analysis generalizes in a straightforward way to the case of multiple differentiated sectors.

We develop the model in stages. The next section describes demand, product markets, labor markets, and the determinants of wages and profits. In the following section, Section 1.3, we discuss the structure of equilibrium, focusing on the case in which both countries are incompletely specialized, and—as in Melitz (2003)—only a fraction of firms export in the differentiated-product industry and some entrants exit this industry. This is followed by an analysis of the impact of labor market frictions on trade, welfare, and productivity in Section 1.4. We allow the labor market frictions to vary both across countries and sectors. There we also study the differential impact of lower trade impediments on countries with

---

different labor market frictions. Importantly, we show that both countries gain from trade in welfare terms and in terms of total factor productivity, independently of trade costs and differences in labor market rigidities. The lowering of labor market frictions in the differentiated sector of one country raises its welfare, but harms the trade partner. Nevertheless, both countries benefit from simultaneous proportional reductions of labor market frictions in the differentiated sector across the world.

By lower frictions in its differentiated sector’s labor market a country gains a competitive advantage in this sector, which is reminiscence of a productivity improvement (but not identical). As a result, it attracts more firms into this sector while the foreign country attracts fewer firms. The entry and exit of firms overwhelms the terms of trade movement, leading to welfare gains in the country with improved labor market frictions and welfare losses in its trade partner.

In Section 1.4 we also show that labor market flexibility is a source of comparative advantage. The country with relatively lower labor market frictions in the differentiated sector (i.e., lower relative to the homogeneous sector) has a larger fraction of exporting firms and it exports differentiated products on net. Moreover, the share of intra-industry trade is smaller while the volume of trade is larger the larger is the difference in relative sectoral labor market rigidities across countries.

In Section 1.5 we take up unemployment. We show that the relationship between unemployment and labor market rigidity in the differentiated sector is hump-shaped when the countries are symmetric. A decline in labor market frictions in the differentiated sector decreases the sectoral rate of unemployment and induces more workers to search for jobs in the differentiated-product sector. When the differentiated sector has the lower sectoral rate of unemployment, which happens when labor market frictions are relatively lower in this sector, the reallocation of workers across sectors, i.e., the composition effect, reduces the aggregate rate of unemployment. Under these circumstances the aggregate rate of unemployment declines, because both the shift in the sectoral rate of unemployment and the reallocation of workers across sectors reduce the aggregate rate of unemployment. On the other hand, when labor market frictions are higher in the differentiated sector, these
two effects impact unemployment in opposite directions, with the composition effect dom-
inating in a highly rigid labor market and the sectoral unemployment effect dominating
in a mildly rigid labor market. As a result, unemployment initially increases and then
decreases as labor market frictions decline, starting from high levels of rigidity.

We also discuss the transmission of shocks across asymmetric countries, using numerical
examples to illustrate various patterns. In particular, we show that in the absence
of unemployment in the homogenous sector, if a single country reduces its labor market
frictions in the differentiated sector this reduces unemployment in the country’s trading
partner by inducing a labor reallocation from the differentiated-product sector to the
homogeneous-product sector. We also show that lowering trade impediments can increase
unemployment in one or both countries, despite its positive welfare effect, and that the
interaction between trade impediments and labor market rigidities produces rich patterns
of unemployment. Specifically, differences in rates of unemployment across countries do
not necessarily reflect differences in labor market frictions; the more flexible country can
have higher or lower unemployment, depending on the height of trade impediments and
the levels of labor market frictions.

In Section 1.6 we discuss the impacts of firing costs and unemployment benefits as
additional sources of labor market frictions. In particular, we describe conditions under
which the previous results remain valid, as well as how they change when these conditions
are not satisfied. The last section summarizes some of the main insights from this analysis.

1.2 The Model

We develop in this section the building blocks of our analytical model. They consist of
a demand structure, technologies, product and labor market structures, and determinants
of wages and profits. After describing these ingredients in some detail, we discuss in the
next three sections equilibrium interactions in a two-country world. In order to focus on
labor market rigidities, we assume that the two countries are identical except for labor
market frictions. This means that the demand structure and the technologies are the same
in both countries. They can differ in the size of their labor endowment, but this difference
is not consequential for the type of equilibrium we discuss in the main text.

1.2.1 Preferences and Demand

Every country has a representative agent who consumes a homogeneous product $q_0$ and a continuum of brands of a differentiated product whose real consumption index is $Q$. The real consumption index of the differentiated product is a constant elasticity of substitution aggregator:

$$Q = \left[ \int_{\omega \in \Omega} q(\omega)^\beta d\omega \right]^{\frac{1}{\beta}}, \quad 0 < \beta < 1, \quad (1.1)$$

where $q(\omega)$ represents the consumption of variety $\omega$, $\Omega$ represents the set of varieties available for consumption, and $\beta$ is a parameter that controls the elasticity of substitution between brands.

Consumer preferences between the homogeneous product, $q_0$, and the real consumption index of the differentiated product, $Q$, are represented by the quasi-linear utility function

$$U = q_0 + \frac{1}{\zeta} Q^\zeta, \quad 0 < \zeta < \beta.$$

The restriction $\zeta < \beta$ ensures that varieties are better substitutes for each other than for the outside good $q_0$. We also assume that the consumer has a large enough income level to always consume positive quantities of the outside good, in which case it is convenient to choose the outside good as numeraire, so that its price equals one, i.e., $p_0 = 1$. Under the circumstances $p(\omega)$, the price of brand $\omega$, and $P$, the price index of the brands, are measured relative to the price of the homogeneous product.

The utility function $U$ implies that a consumer with spending $E$ who faces the price index $P$ for the differentiated product chooses $Q = P^{-1/(1-\zeta)}$ and $q_0 = E - P^{-\zeta/(1-\zeta)}$.

---

6 Alternatively, we could use a homothetic utility function in $q_0$ and $Q$ (see Chapter II).

7 The assumption that consumer spending on the outside good is positive is equivalent to assuming $E > P^{-\zeta/(1-\zeta)}$. Since $\zeta > 0$, the demand for $Q$ is elastic and total spending $PQ$ rises when $P$ falls.
As a result, the demand function for brand $\omega$ can be expressed as

$$q(\omega) = Q - \beta - \zeta - \beta p(\omega) - 1$$

and the *indirect* utility function as

$$V = E + \frac{1 - \zeta}{\zeta} P - \frac{1}{\zeta} = E + \frac{1 - \zeta}{\zeta} Q^\zeta.$$

As usual, the indirect utility function is increasing in spending and declining in price. A higher price index $P$ reduces the demand for $Q$, and—holding expenditure $E$ constant—reduces welfare. This decline in welfare results from the fact that consumer surplus,

$$(1 - \zeta) P^{-\zeta/(1-\zeta)} / \zeta = (1 - \zeta) Q^\zeta / \zeta,$$

declines as $P$ rises and $Q$ falls. In what follows, we characterize equilibrium values of $Q$, from which we infer welfare levels.

### 1.2.2 Technologies and Market Structure

All goods are produced with labor, which is the only factor of production. The market for the homogeneous product is competitive, and this good serves as numeraire, so that $p_0 = 1$. When a firm is matched with a worker, they produce one unit of the homogenous good.

The market for brands of the differentiated product is monopolistically competitive. A firm that seeks to supply a brand $\omega$ bears an entry cost $f_e$ in terms of the homogeneous good, which covers the technology cost and the cost of setting up shop in the industry. After bearing this cost, the firm learns how productive its technology is, as measured by $\theta$; a $\theta$-firm requires $1/\theta$ workers per unit output. In other words, if a $\theta$-firm employs $h$ workers it produces $\theta h$ units of output. Before entry the firm expects $\theta$ to be drawn from a known cumulative distribution $G_\theta (\theta)$.

After entry the firm has to bear a fixed production cost $f_d$ in terms of the homogeneous good; without it no manufacturing is possible. Following Melitz (2003), we assume that the differentiated-product sector bears a fixed cost of exporting $f_x$ in terms of the homogeneous product. In addition, it bears a variable cost of exporting of the melting-iceberg type:
τ > 1 units have to be exported for one unit to arrive in the foreign country.\(^8\)

We label the two countries A and B. If a country-\(j\) firm, \(j = A, B\), with productivity \(\theta\) hires \(h_j\) workers and chooses to serve only the domestic market, then (1.2) implies that its revenue equals

\[
R_j = Q_j^{-(\beta - \zeta)} \Theta^{1-\beta} h_j^{\beta},
\]

where \(\Theta \equiv \theta^{\beta/(1-\beta)}\) is a transformed measure of productivity that is more convenient for our analysis. Higher \(Q_j\) implies tighter competition in the differentiated product market of country-\(j\) and proportionately reduces revenues for all firms serving this market.

If, instead, this firm chooses also to export, then it has to allocate output \(\theta h_j\) across the domestic and foreign markets, i.e., \(\theta h_j = q_{dj} + q_{xj}\), where \(q_{dj}\) represents the quantity allocated to the domestic market and \(q_{xj}\) represents the quantity allocated to the export market.\(^9\) With an optimal allocation of output across markets, the resulting total revenue is

\[
R_j = \left[ Q_j^{\frac{\beta - \zeta}{1-\beta} \mu} + \tau^{\frac{\beta}{1-\beta} Q_{(-j)}^{\frac{\beta - \zeta}{1-\beta} \mu}} \right]^{1-\beta} \Theta^{1-\beta} h_j^{\beta},
\]

where \((-j)\) is the index of the country other than \(j\). In general, the revenue function of

---

\(^8\) As is common in models with home market effects, we assume that there are no trade frictions in the homogeneous-product sector. We show in our working paper Helpman and Itskhoki (2008) that adding trade costs to the homogenous sector does not affect the results when these costs are not too large. In that paper there are no labor market frictions in the homogenous sector, but the same arguments can be adapted to our framework.

\(^9\) From (1.2) these quantities have to satisfy

\[
q_{dj} = Q_j^{\frac{\beta - \zeta}{1-\beta} \mu} \mu_\nu \rho_{dj} \text{ and } q_{xj} = \tau Q_{(-j)}^{\frac{\beta - \zeta}{1-\beta} \mu} (\tau p_{xj})^{\frac{\beta}{1-\beta} \mu}.
\]

In this specification \(\rho_{dj}\) and \(p_{xj}\) are producer prices of home and foreign sales, respectively. Note that when exports are priced at \(p_{xj}\), consumers in the foreign country pay an effective price of \(\tau p_{xj}\) due to the variable export costs. Under the circumstances they demand \(Q_{(-j)}^{(-\zeta)/(-1-\beta)} (\tau p_{xj})^{-1/(1-\beta)}\) consumption units. To deliver these consumption units the supplier has to manufacture \(q_{xj}\) units, as shown above. Such a producer maximizes total revenue when marginal revenues are equalized across markets. In the case of constant elasticity of demand functions this requires equalization of producer prices, which implies that the optimal allocation of output satisfies

\[
q_{xj}/q_{dj} = \tau^{\frac{\beta}{1-\beta} (Q_{(-j)}/Q_j)} \rho_{dj}^{\frac{\beta - \zeta}{1-\beta} \mu}.
\]
country-\(j\) firm with productivity \(\theta\) can therefore be represented by

\[
R_j (\Theta, h_j) = \left[ Q_j^{\frac{\beta-\zeta}{1-\beta}} + I_{xj} (\Theta) \tau^{-\frac{\beta}{1-\beta}} Q_{(-j)}^{\frac{\beta-\zeta}{1-\beta}} \right]^{1-\beta} \Theta^{1-\beta} h_j^\beta, \tag{1.4}
\]

where \(I_{xj} (\Theta)\) is an indicator variable that equals one if the firm exports and zero otherwise.

### 1.2.3 Wages and Profits

There are search and matching frictions in every sector and firms post vacancies in order to attract workers. The cost of posting vacancies and the matching process generate hiring costs. Moreover, search and matching frictions generate bilateral monopoly power between a worker and his firm, as a result of which they engage in wage bargaining.\(^\text{10}\)

We assume that in the homogeneous-product sector every firm employs one worker. This assumption is common in the search and matching literature (see Pissarides, 2000) and in our case leads to no loss of generality. Since firms in this sector are homogenous in terms of productivity and produce a homogenous good, our analysis does not change if we allow firms to hire multiple workers, as long as they remain price takers.

When a firm and a worker match, they bargain over the surplus from the relationship. Since the outside option of each party equals zero at this stage, the surplus—which consists of the revenue from sales of one unit of the homogeneous product—equals one. Assuming equal weights in the bargaining game then implies that the worker gets a wage \(w_0 = 1/2\) and the firm gets a profit \(\pi_0 = 1/2\), and these payoffs are the same in every country. We discuss additional details of the labor market equilibrium in this sector in the following section.

In the differentiated-product industry firms are heterogeneous in terms of productivity but face the same cost of hiring in the labor market. A \(\Theta\)-firm from country \(j\) that seeks to employ \(h_j\) workers bears the hiring cost \(b_j h_j\) in terms of the homogeneous good, where \(b_j\) is exogenous to the firm yet it depends on sectoral labor market conditions, as we

---

\(^{10}\) In the earlier working paper version (Helpman and Itskhoki, 2008), we focused on the case in which there are no labor market frictions in the homogeneous-product sector. The current framework incorporates it as a special case (see footnote 15).
discuss below. It follows that a worker cannot be replaced without cost. Under these circumstances, a worker inside the firm is not interchangeable with a worker outside the firm, and workers have bargaining power after being hired. Workers exploit this bargaining power in the wage determination process.

We assume that the $h_j$ workers and the firm engage in strategic wage bargaining with equal weights in the manner suggested by Stole and Zwiebel (1996a,b), which is a natural extension of Nash bargaining to the case of multiple workers. The revenue function (1.4) then implies that the firm gets a fraction $1/(1 + \beta)$ of the revenue and the workers get a fraction $\beta/(1 + \beta)$.

Recall that $\beta$ determines the concavity of the revenue function in the number of workers; a lower $\beta$ makes the revenue more concave and reduces the revenue loss from the departure of a marginal worker. Therefore, lower $\beta$ reduces the equilibrium share of the workers in the division of revenue. This bargaining outcome is derived under the assumption that at the bargaining stage a worker’s outside option is unemployment, and the value of unemployment is zero because there are no unemployment benefits and the model is static. In Section 1.6 we discuss unemployment benefits, and in Helpman and Itskhoki (2009) we show that our bargaining solution carries over to the steady state of a dynamic model.

Anticipating the outcome of this bargaining game, a $\Theta$-firm that wants to stay in the industry chooses an employment level, $h_j$, and whether to serve the foreign market, $I_{xj} \in \{0, 1\}$, that maximize profits. That is, it solves the following problem:

$$\pi_j(\Theta) \equiv \max_{I_{xj} \in \{0, 1\}, h_j \geq 0} \left\{ \frac{1}{1 + \beta} \left[ Q_j^{\beta-\varsigma} + I_{xj} \tau^{\beta-\varsigma} Q_j^{\beta-\varsigma} \right]^1 - \beta \cdot \Theta^{1-\beta} h_j^{\beta} - b_j h_j - f_d - I_{xj} f_x \right\}. \quad (1.5)$$

11 In the solution to the Stole and Zwiebel bargaining game the firm and a worker equally divide the marginal surplus from their relationship, i.e.,

$$\frac{\partial}{\partial h} \left[ R_j(\Theta, h) - w_j(\Theta, h) h \right] = w_j(\Theta, h),$$

where $w_j(\Theta, h)$ is the bargained wage rate in a $\Theta$-firm in country-$j$ which employs $h$ workers. Therefore, the left-hand side represents the surplus of the firm from employing the marginal worker, accounting for the fact that his departure will impact the wage rate of the remaining workers. The wage on the right-hand side is the worker’s surplus. Using the expression for revenue (1.4), the above condition represents a differential equation for the wage schedule which yields the solution $w_j(\Theta, h) = \beta/(1 + \beta) \cdot R_j(\Theta, h)/h$. 
The solution to this problem implies that the employment level of a $\Theta$-firm in country $j$ can be decomposed into

$$h_j (\Theta) = h_{dj} (\Theta) + I_{xj} (\Theta) h_{xj} (\Theta),$$

where $h_{dj} (\Theta)$ represents employment for domestic sales, $h_{xj} (\Theta)$ represents employment for export sales, and

$$h_{dj} (\Theta) = \phi_1 \beta_j b_j^{\frac{1}{1-\beta}} Q_j^{\beta-\frac{\alpha-\xi}{1-\beta}} \Theta,$$

$$h_{xj} (\Theta) = \phi_1 \beta_j b_j^{\frac{1}{1-\beta}} \tau^{-\frac{\beta}{1-\beta}} Q_j^{\beta-\frac{\alpha-\xi}{1-\beta}} (-j) \Theta,$$

where

$$\phi_1 = \left( \frac{\beta}{1+\beta} \right)^{\frac{\alpha}{1-\beta}}.$$

Furthermore, a country-$j$ firm with productivity $\Theta$ pays wages

$$w_j (\Theta) = \frac{\beta}{1+\beta} R_j (\Theta) h_j (\Theta) = b_j,$$

where the first equality is the outcome of the bargaining game and the second equality follows from the optimal employment condition (1.6). Firms find it optimal to increase their employment up to the point at which the bargaining outcome is a wage rate equal to the cost of replacing a worker, $b_j$. Since this hiring cost is common across all firms, in equilibrium country-$j$ firms of all productivity levels, exporters and non-exporters alike, pay equal wages, $w_j = b_j$.\(^ {12}\)

Finally, the operating profits of a $\Theta$-firm in country-$j$ are

$$\pi_j (\Theta) = \pi_{dj} (\Theta) + I_j (\Theta) \pi_{xj} (\Theta),$$

\(^ {12}\)This equilibrium outcome generalizes to other revenue functions and bargaining concepts as long as firms are allowed to vary their employment and the marginal hiring costs are equal across the firms. In Chapter II we develop a richer model, in which there is unobserved worker heterogeneity in addition to firm heterogeneity, wages are higher in more productive firms, and exporters pay a wage premium. Bernard and Jensen (1995) and Farriños and Martín-Marcos (2007) provide evidence to the effect that exporting firms pay higher wages.
where \( \pi_{dj} (\Theta) \) represents operating profits from domestic sales, \( \pi_{xj} (\Theta) \) represents operating profits from export sales, and

\[
\begin{align*}
\pi_{dj} (\Theta) &= \phi_1 \phi_2 b_j^{1-\beta} Q_j^{\frac{\beta - \zeta}{1-\beta}} \Theta - f_d, \\
\pi_{xj} (\Theta) &= \phi_1 \phi_2 b_j^{1-\beta} Q_j^{\frac{\beta - \zeta}{1-\beta}} \Theta - f_x,
\end{align*}
\]

(1.8)

where

\[
\phi_2 = \frac{1 - \beta}{1 + \beta}.
\]

Note that higher labor market rigidity, reflected in a higher \( b_j \), reduces proportionately gross operating profits (i.e., not accounting for fixed costs) in the domestic and foreign market. Therefore, an increase in \( b_j \) is similar to a proportional reduction in the productivity of all country \( j \)'s firms.

The profit functions in (1.8) imply that exporting is profitable if and only if

\[
\pi_{xj} (\Theta) \geq 0,
\]

i.e., there exists a cutoff productivity level, \( \Theta_{xj} \), defined by

\[
\pi_{xj} (\Theta_{xj}) = 0,
\]

(1.9)

such that all firms with productivity above this cutoff export (provided they choose to stay in the industry) and all firms with productivity below it do not export. Firms with low productivity that do not export may nevertheless make money from supplying the domestic market. For this to be the case, their productivity has to be at least as high as \( \Theta_{dj} \), implicitly defined by

\[
\pi_{dj} (\Theta_{dj}) = 0.
\]

(1.10)

We shall consider equilibria in which \( \Theta_{xj} > \Theta_{dj} > \Theta_{min} \equiv \Theta_{min}^{\beta/(1-\beta)} \), where \( \Theta_{min} \) is the lowest productivity level in the support of the distribution \( G_\theta (\theta) \). That is, equilibria in which high-productivity firms profitably export and supply the domestic market, intermediate-productivity firms cannot profitably export but can profitably supply the domestic market, and low-productivity firms cannot make money and exit. Anticipating this outcome, a prospective firm enters the industry only if expected profits from entry
are at least as high as the entry cost \( f_e \). Therefore the free-entry condition is

\[
\int_{\Theta_{dj}} \pi_{dj} (\Theta) dG (\Theta) + \int_{\Theta_{xj}} \pi_{xj} (\Theta) dG (\Theta) = f_e, \tag{1.11}
\]

where \( G (\Theta) \) is the distribution of \( \Theta \) induced by \( G_\theta (\theta) \). The first integral represents expected profits from domestic sales, while the second integral represents expected profits from foreign sales. In equilibrium expected profits just equal entry costs.

### 1.2.4 Labor Market

A country is populated by families. Each family has a fixed supply of \( L \) workers, and the family is the representative consumer whose preferences were described in Section 1.2.1. We assume that there is a continuum of identical families in every country, and the measure of these families equals one in every country.\(^{13}\)

A family in country \( j \) allocates workers to sectors—\( N_j \) workers to the differentiated-product sector and \( N_{0j} = L - N_j \) workers to the homogeneous-product sector—which determines in which sector every worker searches for work. Once committed to a sector, a worker cannot switch sectors. Thus, there is perfect intersectoral mobility \( ex \ ante \) and no mobility \( ex \ post \).

Let the matching function in the homogeneous sector be Cobb-Douglas, so that \( H_{0j} = m_{0j} V_{0j}^\chi N_{0j}^{1-\chi} \) is the number of matches when the number of vacancies in the sector equals \( V_{0j} \) and the number of workers searching for jobs in the sector equals \( N_{0j} \), where \( 0 < \chi < 1 \).

We allow the efficiency of the matching process, as measured by \( m_{0j} \), to vary across countries. It follows that output of homogenous products equals \( H_{0j} \), the probability of a worker finding a job in this sector equals \( x_{0j} \equiv H_{0j}/N_{0j} = m_{0j} (V_{0j}/N_{0j})^\chi \), and the probability of a firm finding a worker equals \( H_{0j}/V_{0j} = m_{0j} (N_{0j}/V_{0j})^{1-\chi} = m_{0j}^{1+\alpha} x_{0j}^{-\alpha} \), where \( \alpha \equiv (1-\chi)/\chi > 0 \).\(^{14}\) We shall use \( x_{0j} \) as our measure of tightness in the sector’s

\(^{13}\) When preferences are homothetic rather than quasi-linear, the family interpretation is useful but not essential (see Section 2.6 of Chapter II).

\(^{14}\) Below we impose parameter restrictions which ensure that matching probabilities are between zero and one. In a dynamic model with continuous time these probabilities are replaced by hazard rates which can take arbitrary positive values including the limiting case of frictionless labor market. We show in
labor market.

Next assume that the cost of posting vacancies equals $v_{0j}$ per worker in country $j$, measured in terms of the homogenous good. Then a firm’s entry cost into the industry equals $v_{0j}$. After paying this cost the firm is matched with a worker with probability $m_{0j}^{1+\alpha}x_{0j}^{-\alpha}$ and not matched otherwise. When the firm is matched with a worker they bargain over the surplus from the relationship, as described in the previous section; the worker gets a wage $w_0 = 1/2$ and the firm gets a profit $\pi_0 = 1/2$. Under these circumstances expected profits equal $m_{0j}^{1+\alpha}x_{0j}^{-\alpha}/2$ and firms enter up to the point at which these expected profits cover the entry cost $v_{0j}$. In other words, in equilibrium tightness in the labor market equals\(^\dagger\)

$$x_{0j} = a_{0j}^{-1/\alpha}, \quad a_{0j} \equiv \frac{2v_{0j}}{m_{0j}^{1+\alpha}} > 1. \quad (1.12)$$

The derived parameter $a_{0j}$ summarizes labor market frictions in the homogeneous sector; it rises with the cost of vacancies and declines with the efficiency of the matching process. Evidently, tightness in the labor market declines with $a_{0j}$.

The expected income of a worker searching for a job in the homogenous sector is $\omega_{0j} = x_{0j}w_{0j}$, which together with (1.12) yields

$$\omega_{0j} = \frac{1}{2}a_{0j}^{-1/\alpha}. \quad (1.13)$$

That is, the expected income of this worker rises with the efficacy of matching in the homogeneous sector and declines with the cost of vacancies. Finally, note that as a result of free entry of firms, the cost of hiring per worker, $b_{0j} \equiv v_{0j}V_{0j}/H_{0j} = (v_{0j}/m_{0j}^{1+\alpha})x_{0j}^\alpha$, equals one half in the homogeneous sector in both countries:

$$b_{0j} = \frac{1}{2}a_{0j}x_{0j}^\alpha = \frac{1}{2}. \quad (1.14)$$

Helpman and Itskhoki (2009) that this type of dynamic specification yields steady state outcomes which are similar to our static specification.

\(^\dagger\) We assume that $m_{0j}^{1+\alpha} < 2v_{0j} < 1$, which ensures that the probability of a worker finding a job and the probability of a firm finding a worker are both smaller than one. Alternatively, when $m_{0j}^{1+\alpha} = 2v_{0j} = 1$, workers and firms are matched with probability one and there is full employment in the homogenous sector, as in Helpman and Itskhoki (2008).
We now turn to the differentiated sector. Let $H_j$ be aggregate employment in the differentiated sector. An individual searching for work in the differentiated-product sector expects to find a job with probability $x_j = H_j/N_j$, where $x_j$ measures the degree of tightness in the sector’s labor market. Conditional on finding a job an individual expects to be paid a wage $w_j = b_j$ (see (1.7)). Therefore the expected income from searching for a job in the differentiated sector is $x_j b_j$.

A family allocates workers to sectors so as to maximize the family’s aggregate income. A worker allocated to the homogeneous sector earns an expected income of $\omega_{0j}$, given in (1.13). On the other hand, a worker allocated to the differentiated sector earns an expected income of $x_j b_j$. In an equilibrium with employment in both sectors the two expected incomes have to be equal. That is, a family chooses $0 < N_j < L_j$ only if

$$x_j b_j = \omega_{0j}. \quad (1.15)$$

Unemployment in the differentiated sector is an equilibrium outcome when $x_j < 1$. We provide below parameter restrictions that ensure this condition.

We now interpret the parameter $b_j$ of the cost-of-hiring function $b_j h$; this variable is exogenous to the firm but endogenous to the industry. As in the homogeneous sector, workers in the differentiated sector are randomly matched with firms. The number of successful matches is $H_j = m_j V_j^{\chi} N_j^{1-\chi}$ in country $j$, where $V_j$ is the number of vacancies and $N_j$ is the number of individuals searching for jobs in this sector. Note that $\chi$ is the same here as in the matching function of the homogenous sector, but $m_j$—which measures the efficiency of matching—is allowed to differ across countries and sectors. It follows that when the cost per vacancy is $v_j$ in the differentiated sector of country $j$, then the cost of hiring is $b_j = v_j V_j/H_j$ per worker, and $b_j$ can be related in a simple way to tightness in the labor market $x_j$:

$$b_j = \frac{1}{2} a_j x_j^a, \quad a_j \equiv \frac{2v_j}{m_j^{1+a}}, \quad (1.16)$$

16 This is similar to the indifference between staying in the countryside and migrating to the city in the Harris and Todaro (1970) model.

17 See Blanchard and Gali (2008) for a similar specification.
where $a_j$ is our measure of frictions in the differentiated sector’s labor market, which is increasing in the cost of vacancies and decreasing in the productivity of matching. Note the symmetry in the modeling of hiring costs in the homogenous and differentiated sectors (compare (1.16) with (1.14)).

Next note that (1.12)-(1.16) uniquely determine the hiring cost $b_j$ and tightness in the labor market $x_j$:

\[
\begin{align*}
x_j &= x_{0j} \left( \frac{a_{0j}}{a_j} \right) ^{\frac{1}{1+\alpha}} = \left( \frac{1}{a_{0j}/a_j} \right) ^{\frac{1}{1+\alpha}}, \\
w_j &= b_j = b_{0j} \left( \frac{a_j}{a_{0j}} \right) ^{\frac{1}{1+\alpha}} = \frac{1}{2} \left( \frac{a_j}{a_{0j}} \right) ^{\frac{1}{1+\alpha}}.
\end{align*}
\]  

(1.17)

Note that $x_j < 1$ and there is unemployment in the differentiated sector if and only if $a_{0j}/a_j^\alpha > 1$, which we assume to be satisfied. It follows from this characterization that whenever a country has the same labor market frictions in both sectors, so that $a_{0j} = a_j$, it has the same labor market tightness in both sectors and the same cost of hiring in both sectors. Yet while the cost of hiring is independent in this case from the common level of labor market frictions because $b_{0j} = b_j = 1/2$, tightness in the sectoral labor markets declines with the level of frictions. This implies that when $a_{0j} = a_j$ in both countries no country has comparative advantage in one of the sectors (see the discussion in the next section), even when the level of labor market frictions varies across countries. In Helpman and Itskhoki (2009) we show that similar patterns arise in the steady state of a dynamic model.

In what follows we assume that $a_A/a_{0A} > a_B/a_{0B}$, so that country $B$ has relatively lower labor market frictions in the differentiated sector. This implies $b_A > b_B$, i.e., country $A$ has a larger hiring cost in the differentiated sector, and $x_A/x_{0A} < x_B/x_{0B}$, i.e., the sectoral labor market tightness is relatively lower in the differentiated sector of country $A$. Note, however, that our assumption on relative sectoral labor market frictions has no implications for whether the labor market is tighter in one sector or the other. When $a_j/a_{0j} > 1$ in both countries, sectoral tightness is higher in the homogeneous sector in both countries; when $a_j/a_{0j} < 1$ in both countries, sectoral tightness is higher in the differentiated sector in both countries; and when $a_A/a_{0A} > 1 > a_B/a_{0B}$, sectoral tightness
is higher in the homogenous sector in country $A$ and higher in the differentiated sector in country $B$. Sectoral labor market frictions can differ due to the fact that it may be more difficult to match workers with firms in some sectors than in other, and labor market frictions can differ across countries due to differences in matching efficiency or differences in costs of posting vacancies.\(^{18}\) We allow these possibilities in order to accommodate variation in sectoral rates of unemployment, which feature in the data.\(^{19}\)

Evidently, the model is bloc recursive, in the sense that the equilibrium wage rate and tightness in the labor market are uniquely determined by labor market frictions. We show in Section 1.6 that this property also holds with firing costs and unemployment benefits. The implication is that labor market frictions determine $(b_j, x_j)$ in country $j$, and these in turn impact other endogenous variables, such as trade, welfare and unemployment.

The sectoral rates of unemployment are $1 - x_{0j}$ in the homogenous sector and $1 - x_j$ in the differentiated sector. As a result, the economy-wide rate of unemployment can be expressed as

$$u_j = \frac{N_{0j}}{L} (1 - x_{0j}) + \frac{N_j}{L} (1 - x_j),$$

which is a weighted average of the sectoral rates of unemployment, where the weights are the fractions of workers seeking jobs in every sector. It follows that the unemployment rate can rise either because it rises in one or both sectors or because more individuals search for work in the sector with a higher rate of unemployment.

### 1.3 Equilibrium Structure

We focus on equilibria with incomplete specialization, in which every country produces homogeneous and differentiated products. Conditions for incomplete specialization are

\(^{18}\) In a dynamic model sectors may differ in separation rates, which leads to $b_{0j} \neq b_j$; see Helpman and Itskhoki (2009). Specifically, we show that if the differentiated sector has a higher separation rate it leads to greater turnover in this sector and a country with more efficient matching technology has a comparative advantage in this sector. Furthermore, policy differences can be a source of cross-country variation in labor market frictions, which we discuss in Section 1.6.

\(^{19}\) To illustrate, the BLS reports that in 2007 Mining had an unemployment rate of 3.4%, Construction had 7.4%, and Manufacturing had 4.3% (see http://www.bls.gov/cps/cpsaat26.pdf, accessed on April 25, 2008).
described in the Appendix, and in our earlier working paper (Helpman and Itskhoki, 2008) we discuss properties of equilibria with complete specialization. This section is devoted to a description of equilibria and some of their properties. More substantive results, which build on this section, are developed and discussed in subsequent sections.

Equations (1.8)-(1.10) yield the following expressions for the domestic market and export cutoffs:

\[
\begin{align*}
\Theta_{dj} &= \frac{1}{\phi_1 \phi_2} f_d b_j^{\frac{1-\beta}{\beta}} Q_j^{\frac{\beta - \zeta}{1-\beta}}, \\
\Theta_{xj} &= \frac{1}{\phi_1 \phi_2} f_x b_j^{\frac{1-\beta}{\beta}} \tau^{\frac{\beta - \zeta}{1-\beta}} (-j).
\end{align*}
\]  

(1.19)

Now substitute these expressions into (1.8) and the resulting profit functions into the free-entry condition (1.11) to obtain

\[
 f_d \int_{\Theta_{dj}}^{\infty} \left( \frac{\Theta}{\Theta_{dj}} - 1 \right) dG(\Theta) + f_x \int_{\Theta_{xj}}^{\infty} \left( \frac{\Theta}{\Theta_{xj}} - 1 \right) dG(\Theta) = f_e, \quad j = A, B.
\]  

(1.20)

This form of the free-entry condition generates a curve in \((\Theta_{dj}, \Theta_{xj})\) space on which every country’s cutoffs have to be located, because this curve depends only on the common cost variables and on the common distribution of productivity. Moreover, this curve is downward-sloping, as depicted by FF in Figure 1.1, and each country has to be located above the 45° line for the export cutoff to be higher than the domestic cutoff.\(^{20}\)

Also note that as the export cutoff goes to infinity, the domestic cutoff approaches the cutoff of a closed economy, which is represented by \(\Theta_d^c\) in the figure. It therefore follows that if the cutoff \(\Theta_d\) in the closed economy is larger than \(\Theta_{\text{min}}\), so is \(\Theta_d\) in the open

\(^{20}\) Note, from (1.19), that in a symmetric equilibrium, in which \(Q_j = Q_{(-j)}\), the export cutoff is higher if and only if \(\tau^{\beta/(1-\beta)} f_x > f_d\), which is the condition required for exporters to be more productive in Melitz (2003). We assume for convenience that this condition is satisfied for all \(\tau \geq 1\) which requires \(f_x > f_d\).
Finally note that (1.19) yields

\[
\frac{\Theta_{xj}}{\Theta_{d(-j)}} = \frac{f_x^{\frac{1}{1-\beta}}}{f_d^{\frac{1}{1-\beta}}} \left[ \frac{b_j}{b_{(-j)}} \right]^{\frac{\beta}{1-\beta}}, \quad j = A, B. \tag{1.21}
\]

Equations (1.20) and (1.21) can be used for solving the four cutoffs as functions of labor market frictions and cost parameters. As is evident, the cutoffs do not depend on the levels of the hiring costs \(b_j\), only on their relative size. And once the cutoffs have been solved, they can be substituted into (1.19) to obtain solutions for the real consumption indexes \(Q_j\).

Our primary interest is in the influence of trade and labor market frictions on the trading countries. We therefore use (1.20) and (1.21) to calculate the impact of these

---

21 The autarky production cutoff is the solution to

\[
f_d \int_{\Theta_d^*}^{\infty} \left( \frac{\Theta}{\Theta_d^*} - 1 \right) dG(\Theta) = f_x,
\]

which does not depend on labor market frictions. Note also that \(\Theta_d^* > \Theta_{\text{min}}\) if and only if \((\Theta/\Theta_{\text{min}}) > 1 + f_x/f_d\), where \(\Theta\) is the mean of \(\Theta\), which we assume to be satisfied. This results from the fact that the integral on the left-hand side of the above equation is decreasing in \(\Theta_d^*\) and assumes its largest value of \((\Theta/\Theta_{\text{min}}) - 1\) when \(\Theta_d^* = \Theta_{\text{min}}\).
variables on the cutoffs, obtaining

\[
\hat{\Theta}_{dj} = \frac{\delta_{xj}}{\Delta} \left[ - (\delta x_{(j)} + \delta d_{(j)}) \left( \hat{b}_j - \hat{b}_{(-j)} \right) - (\delta d_{(j)} - \delta x_{(j)}) \hat{\tau} \right],
\]

\[
\hat{\Theta}_{xj} = \frac{\delta_{dj}}{\Delta} \left[ (\delta x_{(j)} + \delta d_{(j)}) \left( \hat{b}_j - \hat{b}_{(-j)} \right) + (\delta d_{(j)} - \delta x_{(j)}) \hat{\tau} \right],
\]

where

\[
\delta_{dj} = \frac{f_d}{\Theta_{dj}} \int_{\Theta_{dj}}^{\infty} \Theta dG(\Theta), \quad \delta_{xj} = \frac{f_x}{\Theta_{xj}} \int_{\Theta_{xj}}^{\infty} \Theta dG(\Theta), \quad \Delta = \frac{1 - \beta}{\beta} (\delta_{dA} \delta_{dB} - \delta_{xA} \delta_{xB}).
\]

Note that \(\delta_{dj}/\phi_2\) is average revenue per entering firm from domestic sales in country \(j\) and \(\delta_{xj}/\phi_2\) is average revenue per entering firm from export sales.\(^{22}\) Moreover, \(\delta_{dj}\) equals average gross operating profits (not accounting for fixed costs) per entering firm from domestic sales and \(\delta_{xj}\) equals average gross operating profits per entering firm from exporting.

Building on these insights, we prove in the Appendix the following lemmas:

**Lemma 1.1.** Let \(b_A > b_B\). Then \(\Theta_{dA} < \Theta_{dB}\) and \(\Theta_{xA} > \Theta_{xB}\).

**Lemma 1.2.** An increase in \(\tau\) raises the export cutoff \(\Theta_{xj}\) and reduces the domestic cutoff \(\Theta_{dj}\) in both countries.

**Lemma 1.3.** Let \(b_A > b_B\). Then \(Q_A < Q_B\).

The first lemma shows that in the country with the relatively higher labor market frictions in the differentiated sector exporting requires higher productivity at the firm level and that firms with lower productivity at the bottom of the productivity distribution break even. The former result is quite intuitive; a disadvantage in labor costs needs to be compensated with a productivity advantage to make exporting profitable. The latter stems

\(^{22}\) To see this, note that profit maximization (1.5) implies \(\pi_z(\Theta) = \phi_2 R_{zj}(\Theta) - f_z\) for \(z = d, x\), where \(R_{dj}(\Theta)\) is revenue from domestic sales and \(R_{xj}(\Theta)\) is revenue from foreign sales. Then, from the zero profit conditions (1.9)-(1.10), we have \(R_{zj}(\Theta) = f_z/\phi_2 \cdot \Theta/\Theta_{zj}\). As a result, the average revenues per entering firm from domestic sales and exports equal

\[
\int_{\Theta_{xj}}^{\infty} R_{xj}(\Theta) dG(\Theta) = \frac{f_x}{\phi_2} \int_{\Theta_{xj}}^{\infty} \Theta dG(\Theta) = \frac{\delta_{xj}}{\phi_2}, \quad z = d, x.
\]
from the fact that in a country with higher \( b_j \) expected profits from exporting are lower at the entry stage, which has to be offset by higher expected profits from domestic sales in order for the free entry condition to be satisfied. This implies that lower-productivity firms find it profitable to serve the domestic market. The second lemma just restates a well known result from Melitz (2003) which also holds in our framework: Higher variable trade costs cut into export profits, enabling only more productive firms to profitably export. Under the circumstances lower productivity firms need to survive entry in order to be able to cover the entry cost. The third lemma states that the country with higher relative labor market frictions in the differentiated sector has lower real consumption of differentiated products. This stems from the home market effect. Due to the presence of trade costs, a country suffers a disadvantage in the local supply of differentiated products when its \( b_j \) is higher in the differentiated industry.

For our equations to describe an equilibrium with incomplete specialization, it is necessary to ensure positive entry of firms in both countries, i.e., \( M_j > 0 \) for \( j = A, B \), where \( M_j \) is the number of firms that enter the differentiated sector in country \( j \). This places restrictions on the permissible difference across countries in labor market rigidities. To derive the implications of these restrictions, first recall that \( Q^\zeta_j = P_j Q_j \) is total spending on differentiated products in country \( j \), and \( M_j \delta_{zj}/\phi_2 \) is total revenue from domestic sales when \( z = d \) and from foreign sales when \( z = x \). Since aggregate spending has to equal aggregate revenue in market \( j \), we have

\[
Q^\zeta_j = M_j \frac{\delta_{dj}}{\phi_2} + M_{(-j)} \frac{\delta_{x(-j)}}{\phi_2},
\]

where the first term on the right-hand side is revenues of domestic firms and the second term is revenues of foreign firms from sales in country \( j \)'s market. Having solved for the cutoffs, which uniquely determine the \( \delta_{zj} \)s, and the real consumption indexes, \( Q_j \)s, these equations for \( j = A, B \) yield the following solutions for the number of entrants:

\[
M_j = \frac{(1 - \beta) \phi_2}{\beta \Delta} \left[ \delta_{d(-j)} Q^\zeta_j - \delta_{x(-j)} Q^\zeta_{(-j)} \right]. \tag{1.23}
\]
We show in the Appendix that they imply:

**Lemma 1.4.** In an equilibrium with incomplete specialization: (i) $\delta_{dj} > \delta_{xj}$ in both countries; (ii) if $b_A > b_B$, then $\delta_{dA} > \delta_{dB}$ and $\delta_{xA} < \delta_{xB}$.

**Lemma 1.5.** Let $b_A > b_B$. Then $M_A < M_B$.

Lemma 1.4 is a technical lemma, which describes conditions that hold in an equilibrium with incomplete specialization. The economic implication of part (i) is that average revenue per entering firm from domestic sales exceeds average revenue per entering firm from export sales in each one of the countries, and the economic implication of part (ii) is that in country $A$, which has the relatively higher labor market frictions in the differentiated sector, average revenue per entering firm from domestic sales is higher and average revenue per entering firm from export sales is lower than in country $B$. And the last lemma states that there is less entry of firms in the differentiated product industry in the country in which labor market frictions are relatively higher in this sector; a result which is quite intuitive.

Finally, consider the determinants of the number of workers searching for jobs in the differentiated sector, $N_j$, and aggregate employment in that sector, $H_j$. On the one hand, the wage bill in the differentiated sector, $w_j H_j$, equals $\omega_{0j} N_j$, because the wage rate is $w_j = b_j = \omega_{0j} / x_j$ (see (1.17)) and $x_j = H_j / N_j$ by definition. This implies that aggregate income equals $\omega_{0j} L$, where income $\omega_{0j} N_j$ is derived from the differentiated sector and income $\omega_{0j} N_0 = \omega_{0j} (L - N_j)$ is derived from the homogeneous sector. On the other hand, the wage bill in the differentiated sector equals the fraction $\beta / (1 + \beta)$ of revenue (a result from the bargaining game). Therefore

\[
\omega_{0j} N_j = \frac{\beta}{1+\beta} M_j \left( \frac{\delta_{dj}}{\phi_2} + \frac{\delta_{xj}}{\phi_2} \right), \tag{1.24}
\]

where $M_j (\delta_{dj} + \delta_{xj}) / \phi_2$ is total revenue of country-$j$ firms from domestic sales and exporting. It follows that, once the cutoffs and the numbers of firms are known, this equation determines the number of workers searching for jobs in the differentiated-product indus-
Having solved for \(N_j\), aggregate employment in the differentiated sector is

\[H_j = x_j N_j.\]  

(1.25)

The remaining \(N_{0j} = L - N_j\) workers search for jobs in the homogenous-good sector, with \(H_{0j} = x_{0j} N_{0j}\) of them finding employment and generating \(H_{0j}\) units of output of the homogenous good. This completes the description of an equilibrium with incomplete specialization.

1.4 Trade, Welfare and Productivity

In this section we explore channels through which the two countries are interdependent. For this purpose we organize the discussion around two main themes: the impact of a country’s labor market frictions on its trade partner, and the differential effects of trade impediments on countries with different labor market frictions.

1.4.1 Welfare

We are interested in knowing how labor market rigidities and trade frictions affect welfare, and in particular their differential impact on the welfare of the two countries. Since aggregate spending in country \(j\), \(E_j\), equals aggregate income, and aggregate income equals \(\omega_{0j} L = a_{0j}^{-1/\alpha} L/2\) (see (1.13)), the indirect utility function (1.3) implies that welfare is higher the larger the real consumption index of differentiated products \(Q_j\) is and the lower the friction in the homogeneous sector \(a_{0j}\) is. We have already shown that \(Q_j\) is higher in country \(B\) (see Lemma 1.3). It follows that welfare is also higher in \(B\) as long as the labor market friction in its homogenous sector, \(a_{0B}\), is not too high relative to that in country \(A\), \(a_{0A}\).

Now combine the formulas for change in the cutoffs (1.22) with the log-differential of

\[\omega_{0j}\] is determined by labor market frictions in the homogeneous sector; see (1.13).
the first equation in (1.19) to obtain

\[
\frac{\beta - \zeta}{1 - \beta} \hat{Q}_j = \frac{1}{\Delta} \left[ -\delta_{d(-j)} (\delta_{xj} + \delta_{dj}) b_j + \delta_{xj} (\delta_{x(-j)} + \delta_{d(-j)}) \hat{b}_{(-j)} - \delta_{xj} (\delta_{d(-j)} - \delta_{x(-j)}) \hat{\tau} \right].
\]

(1.26)

This equation has a number of implications. First, it shows that a reduction in a country’s labor market frictions in the differentiated sector, i.e., a decline in \(a_j\) which reduces \(b_j\) (see (1.17)), raises its real consumption index \(Q_j\) and therefore its welfare, but it reduces the trade partner’s welfare. On the other hand, a simultaneous reduction in labor market frictions in the differentiated sectors of both countries at a common rate \(\hat{a}_A = \hat{a}_B\), which implies \(\hat{b}_A = \hat{b}_B\), raises everybody’s welfare.\(^{24}\) Second, a reduction in a country’s labor market frictions at a common rate in both sectors, i.e., a decline in \(a_{0j}\) and \(a_j\) such that \(\hat{a}_{0j} = \hat{a}_j\), does not impact its real consumption index \(Q_j\) nor the real consumption index of its trade partner \(Q_{(-j)}\). As a result, the trade partner’s welfare does not change, yet \(j\)’s welfare rises because expected income of a worker, \(\omega_0\), rises (see (1.3) and (1.13), and recall that expenditure equals income, \(E = \omega_{0j}L\)). Third, in view of Lemma 1.2, a reduction in trade impediments raises welfare in both countries. We summarize these findings in\(^{25}\)

**Proposition 1.1.** (i) A reduction in labor market frictions in country \(j\)’s differentiated sector raises its welfare and reduces the welfare of its trade partner. (ii) A simultaneous proportional reduction in labor market frictions in the differentiated sectors of both countries raises welfare in both of them. (iii) A reduction in labor market frictions in country \(j\) at a common rate in both sectors raises its welfare and does not affect the welfare of its trade partner. (iv) A reduction of trade impediments raises welfare in both countries and \(Q_j\) rises proportionately more in country \(B\), which has relatively lower labor market frictions in the differentiated sector.

The first part of this proposition is intriguing: it states that a country harms the trade

\(^{24}\) This follows from the fact that \(-\delta_{d(-j)} (\delta_{xj} + \delta_{dj}) + \delta_{xj} (\delta_{x(-j)} + \delta_{d(-j)}) = -\beta \Delta / (1 - \beta) < 0.\)

\(^{25}\) The very last part of the proposition follows from the fact that (1.26) implies

\[
\frac{\beta - \zeta}{1 - \beta} \left[ \hat{Q}_j - \hat{Q}_{(-j)} \right] = -\frac{1}{\Delta} \left[ (\delta_{d0} + \delta_{xj}) (\delta_{d(-j)} + \delta_{x(-j)}) \left( \hat{b}_j - \hat{b}_{(-j)} \right) + (\delta_{d(-j)} \delta_{xj} - \delta_{x(-j)} \delta_{dj}) \hat{\tau} \right].
\]

Under these circumstances \(\hat{Q}_j > \hat{Q}_{(-j)}\) in response to \(\hat{\tau} < 0\), when \(\hat{b}_j < \hat{b}_{(-j)}\) (by Lemma 1.4).
partner by reducing its labor market frictions. Moreover, this happens despite the fact that the trade partner \((-j)\) enjoys better terms of trade as a result of improved labor market conditions in country \(j\), because \((-j)\) pays lower prices for imported varieties from \(j\) and gets access to a larger set of foreign varieties. The reason for this result is that lower labor market frictions in country \(j\)'s differentiated sector act like a productivity improvements in this country, which makes foreign firms less competitive and therefore *crowds them out* from this sector. As a result, fewer foreign firms enter the industry. The entry of domestic firms does not fully compensate foreign consumers for the exit of foreign firms due to the home market effect, so that foreign welfare declines, and this negative welfare effect is larger than the welfare gain from improved terms of trade.\(^{26}\)

The last part of this proposition establishes that both countries gain from trade, because autarky is attained when \(\tau \to \infty\).\(^{27}\) To emphasize this conclusion, we restate it in:

**Proposition 1.2.** Both countries gain from trade.

This proposition is interesting, because it is well known that gains from trade are not ensured in economies with nonconvexities and distortions (see Helpman and Krugman, 1985), and in addition to the standard nonconvexities and distortions that exist in models of monopolistic competition our model contains frictions in labor markets. The intuition is that every country gains access to a larger variety choice, and in addition, the differentiated sector—which is too small relative to its first-best size—expands. Together, these effects of trade opening dominate the welfare outcome.

---

\(^{26}\) Demidova (2008) studies a full employment model with exogenous differences in productivity distributions across countries, and finds that: (a) productivity improvements in one country hurt its trade partner; and (b) falling trade costs benefit disproportionately the more productive country, and may even hurt the less productive country. Our results on labor market frictions are similar to these (except that in our case both countries necessarily gain from falling trade costs), because— notwithstanding unemployment—labor market frictions have analogous effects to productivity. We stress that these effects are analogous but not identical, because our cross-country differences in relative labor market frictions are not identical to the differences in productivity in Demidova's paper.

\(^{27}\) The following is a direct proof of the gains-from-trade argument: We have seen that the domestic cutoff is higher in every country in the trading equilibrium than in autarky. The first equation in (1.19) then implies that \(Q_j\) is higher in every country in the trading equilibrium, because this equation also holds in autarky. In addition, in the earlier working paper Helpman and Itskhoki (2008), we show that both countries gain from trade when the difference in labor market institutions is large enough to cause the relatively rigid country to specialize in the production of the homogeneous good. Interestingly, in this case the gains from trade accrue disproportionately to the relatively rigid country, although its level of welfare is always lower than that of the relatively flexible country.
1.4.2 Trade Structure

From Lemma 1.1 we know that the country with lower relative labor market frictions in the differentiated sector has a lower export cutoff and a higher domestic cutoff; therefore it also has a larger fraction of exporting firms in the differentiated-product sector. In addition, we know that exports per entering firm equal $\frac{\delta x_j}{\phi_2}$. Therefore exports of differentiated products from $j$ to $(-j)$ are

$$X_j = M_j \frac{\delta x_j}{\phi_2}.$$  

Lemma 1.4 states that country $B$ has a larger $\delta x_j$ and Lemma 1.5 states that it has more firms. Therefore $X_B > X_A$, which implies that $B$ exports differentiated products on net. As a consequence, country $A$ exports homogeneous goods.

As in the standard Helpman-Krugman model of trade in differentiated products, there is intra-industry trade. We can therefore decompose the volume of trade into intra-industry and intersectoral trade. Because trade is balanced, the total volume of trade equals $2X_B$, the volume of intra-industry trade equals $2X_A$, and the share of intra-industry trade equals

$$\frac{X_A}{X_B} = \frac{\delta x_A M_A}{\delta x_B M_B}.$$  

Using this representation we show in the Appendix that the share of intra-industry trade declines in $b_A/b_B$. These results are summarized in

**Proposition 1.3.** Let $b_A > b_B$. Then: (i) A larger fraction of differentiated-sector firms export in country $B$. (ii) Country $B$ exports differentiated products on net and imports homogeneous goods. (iii) The share of intra-industry trade is smaller the larger $b_A/b_B$ is.

That is, as in Davidson, Martin and Matusz (1999), labor market frictions impact comparative advantage, and in our case they also impact the share of intra-industry trade. In addition, under Pareto-distributed productivity, the model also implies that the volume of trade is larger the larger is the difference in relative hiring costs across countries, $b_A/b_B$, and the smaller are the trade impediments (see Appendix). These are testable
implications of our model.

1.4.3 Productivity

Alternative measures of total factor productivity (TFP) can be used to characterize the efficiency of production. We choose to focus on one such measure—the employment-weighted average of firm-level productivity—which is commonly used in the literature.\textsuperscript{28} In the differentiated sector this measure is

\[
TFP_j = \frac{M_j}{H_j} \left[ \int_{\Theta_{dj}}^{\infty} \Theta^{1/\beta} h_{dj}(\Theta) dG(\Theta) + \int_{\Theta_{xj}}^{\infty} \Theta^{1/\beta} h_{xj}(\Theta) dG(\Theta) \right].
\]

(1.27)

Recall that \( q_{zj}(\Theta) = \Theta^{(1-\beta)/\beta} h_{zj}(\Theta) \) for \( z = d, x \). Therefore, \( TFP_j \) equals the output of differentiated products divided by employment in the differentiated sector.\textsuperscript{29}

Using (1.6) and (1.8)-(1.10), we can express (1.27) as

\[
TFP_j = \frac{\delta_{dj} \varphi_{dj} + \delta_{xj} \varphi_{xj}}{\delta_{dj} + \delta_{xj}} = \varpi_{dj} \varphi_{dj} + \varpi_{xj} \varphi_{xj},
\]

(1.28)

where \( \varpi_{dj} = \delta_{dj}/(\delta_{dj} + \delta_{xj}) \) is the share of domestic sales in revenue and \( \varpi_{xj} \) is the share of exports, i.e., \( \varpi_{xj} = 1 - \varpi_{dj}, \ j = A, B \). Moreover,

\[
\varphi_{zj} \equiv \varphi(\Theta_{zj}) = \frac{\int_{\Theta_{zj}}^{\infty} \Theta^{1/\beta} dG(\Theta)}{\int_{\Theta_{xj}}^{\infty} \Theta dG(\Theta)}, \quad z = d, x,
\]

where \( \varphi_{dj} \) represents the average productivity of firms that serve the home market and \( \varphi_{xj} \) represents the average productivity of exporting firms. It follows that aggregate productivity equals the weighted average of the productivity of firms that serve the domestic market

\textsuperscript{28} This corresponds to the measure analyzed by Melitz (2003) in the appendix. Note that Melitz uses revenue to weight firm productivity levels. However, in equilibrium, revenue is proportional to employment, in which case his and our productivity indexes are the same.

\textsuperscript{29} An alternative, and potentially more desirable, measure of productivity, would divide output by the number of workers searching for jobs in the differentiated-product sector, \( N_j \). This measure is always smaller than \( TFP_j \) by the factor \( x_j \). It follows that labor market liberalization has an additional positive effect on this measure of productivity as compared to the measure used in the main text. Also note that \( TFP_j \) measures productivity in the differentiated-product sector only, rather than in the entire economy, and productivity in the homogeneous-product sector is constant given \( a_{0j} \). We discuss in the Appendix a productivity measure that accounts for the compositional effects across sectors.
and the productivity of firms that export, with the revenue shares serving as weights. We show in the Appendix that $\varphi(\cdot)$ is an increasing function. Therefore average productivity is higher among exporters, i.e., $\varphi_{xj} > \varphi_{dj}$.

Expression (1.28) implies that the cutoffs $\{\Theta_{dj}, \Theta_{xj}\}$ uniquely determine the $TFP_j$s, because $\varpi_{zj}$ and $\varphi_{zj}$ depend only on the cutoffs. Moreover, since the two cutoffs are linked by the free-entry condition (1.20), $TFP_j$ can be expressed as a function of the domestic cutoff $\Theta_{dj}$. This implies that in a closed economy $TFP_j$ is not responsive to changes in labor market frictions, because $\Theta_{dj}$ is uniquely determined by the fixed costs of entry and production and the ex ante productivity distribution.

Productivity $TFP_j$ is higher in the trade equilibrium than in autarky, however, because $\varphi(\Theta_{xj}) > \varphi(\Theta_{dj}) > \varphi(\Theta_{d(-j)})$, and in autarky $\varpi_{zj} = 0$. That is, the average productivity of exporters and nonexporters alike is higher in the trade equilibrium than is the average productivity of firms in autarky. In addition, trade reallocates revenue to the exporting firms, which are on average more productive. For both these reasons trade raises $TFP_j$. We summarize these results in

**Proposition 1.4.** (i) In the closed economy, $TFP_j$ does not depend on labor market frictions. (ii) $TFP_j$ is higher in any trade equilibrium than in autarky.

Next recall that in an open economy a reduction of trade costs raises the domestic cutoff and reduces the export cutoff. In addition, a reduction in country $j$’s labor market frictions in the differentiated sector raises $\Theta_{dj}$ and $\Theta_{x(-j)}$ and reduces $\Theta_{d(-j)}$ and $\Theta_{xj}$. Finally, a simultaneous and proportional decline in both countries’ labor market frictions in the differentiated sector (i.e., $\hat{b}_A = \hat{b}_B < 0$) leaves all these cutoffs unchanged (see (1.22)).

How do changes in labor market frictions impact productivity? In the case in which both countries’ labor market frictions decline by the same factor of proportionality, the answer is simple: the $TFP_j$s do not change. As long as productivity is measured with regard to the number of employed workers rather than the number of workers searching for jobs, measured sectoral productivity levels are not sensitive to the absolute levels of $b_j$s; only the relative levels matter. This result points to a shortcoming of this TFP measure.
We nevertheless continue the analysis with this measure, because it is commonly used in the literature.

A shock that raises the domestic cutoff $\Theta_{dj}$ and reduces the export cutoff $\Theta_{xj}$ affects $TFP_j$ through three channels. First, the reallocation of revenue from firms that serve the home market to exporters raises the weight on the productivity of exporters, $\varpi_{xj}$, which raises in turn $TFP_j$. Second, some least-efficient firms exit the industry, thereby raising the average productivity of firms that sell only in the home market, $\varphi_{dj}$, which raises $TFP_j$. Finally, some firms with productivity below $\Theta_{xj}$ begin to export, thereby reducing the average productivity of exporters, $\varphi_{xj}$, which reduces $TFP_j$. 30

The presence of the third effect, which goes against the first two, does not enable us to sign the impact of single-country reductions of labor market frictions on productivity; in general, productivity may increase or decrease. The sharp result for the comparison of autarky to trade derives from the fact that, in a move from autarky to trade, the third effect is nil. In the Appendix, we provide sufficient conditions for productivity to be monotonically rising with $\Theta_{dj}$, and therefore declining with $b_j$ and $\tau$ and rising with $b_{(-j)}$. In this section, however, we limit our discussion to the case of Pareto-distributed productivity draws, which yields sharp predictions.

Under the assumption of Pareto-distributed productivity, that is, $G(\Theta) = 1-(\Theta_{\text{min}}/\Theta)^k$ for $\Theta \geq \Theta_{\text{min}}$, (1.28) results in (see Appendix):

$$\hat{TFP}_j = \frac{\delta_{dj}(\varphi_{xj} - \varphi_{dj})(1 + k - 1/\beta)}{\delta_{dj}\varphi_{dj} + \delta_{xj}\varphi_{xj}} \Theta_{dj},$$  

(1.29)

where $k > 1/\beta$ is required for $TFP_j$ to be finite, and we therefore assume that it holds, and an increase in $\Theta_{dj}$ is accompanied by a corresponding decrease in $\Theta_{xj}$ in order to satisfy the free-entry condition. As a result, $TFP_j$ is higher the higher $\Theta_{dj}$ is (and the lower $\Theta_{xj}$ is). It follows that productivity is higher in country $B$, and a reduction in a country’s labor market frictions in the differentiated sector raises its productivity and reduces the productivity of its trade partner. An implication of this result is that the

30 Formally, this decomposition can be represented as $\hat{TFP}_j = \hat{\varphi}_{xj}(\varphi_{xj} - \varphi_{dj}) + (1 - \varphi_{xj})\varphi_{dj} + \varphi_{xj}\hat{\varphi}_{xj}$ with $\hat{\varphi}_{xj} > 0$, $\varphi_{dj} > 0$ and $\varphi_{xj} < 0$. 

29
gap in productivity between countries $B$ and $A$ is increasing in $b_A/b_B$ and therefore in $a_A/a_B$, their relative labor market frictions in the differentiated sector. These results are summarized in

**Proposition 1.5.** Let $b_A > b_B$ and let $\Theta$ be Pareto-distributed with shape parameter $k > 1/\beta$. Then: (i) $\text{TFP}_j$ is higher in $B$; (ii) a decline in $a_j$ raises $\text{TFP}_j$ and reduces $\text{TFP}_{(-j)}$; (iii) a reduction of trade costs $\tau$ raises $\text{TFP}_j$ in both countries.

In other words, total factor productivity is higher in the country with relatively lower labor market frictions in the differentiated sector, and while a reduction of labor market frictions in this sector in any country raises its own total factor productivity, this hurts the total factor productivity of the country’s trade partner.

### 1.5 Unemployment

Before discussing the variation of unemployment across countries with different labor market frictions in Section 1.5.2, we first examine the determinants of unemployment in a world of symmetric countries.

#### 1.5.1 Symmetric Countries

We study in this section countries with $a_{0A} = a_{0B} = a_0$ and $a_A = a_B = a$, so that $b_A = b_B = b$, in order to understand how changes in the common levels of labor market frictions and the common level of variable trade cost affect unemployment. In such equilibria, the cutoffs $\Theta_d$ and $\Theta_x$, the consumption index $Q$, the number of entrants $M$, the number of individuals searching for jobs in the differentiated-product sector $N$, the number of workers employed in that sector $H$, and the rate of unemployment $u$ are the same in both countries. We therefore drop the country index $j$ for convenience. From Section 1.3 we know that two symmetric economies are at the same point on the $FF$ curve in Figure 1.1 (point $S$), the location of this point is invariant to the common level of labor market frictions, and this point is higher the larger $\tau$ is. Moreover, (1.26) implies that $Q$ is lower the higher are either $b$ or $\tau$. When $b$ is higher as a result of higher frictions
in the labor market of the differentiated sector, welfare is lower because $Q$ is lower while aggregate income $E = \omega_0 L$ is not affected (recall that welfare is given by (1.3)).

![Figure 1.2: Unemployment in a world of symmetric countries](image)

In order to assess the impact of labor market rigidities on unemployment, we need to know their quantitative impact on $Q$. For this reason we use (1.26) to obtain

$$\dot{Q} = -\frac{\beta}{\beta - \zeta} \left( b + \frac{\delta_x \delta_d + \delta_x \hat{\tau}}{\delta_d + \delta_x \hat{\tau}} \right).$$

Next combine (1.23) and (1.24) to obtain $\omega_0 \dot{N} = \beta Q^\zeta / (1 + \beta)$, which together with the previous equation yields

$$\dot{N} = -\frac{\beta \zeta}{\beta - \zeta} \left( b + \frac{\delta_x \delta_d + \delta_x \hat{\tau}}{\delta_d + \delta_x \hat{\tau}} \right)$$

under the assumption that changes in $b$ are driven by changes in $a$, our measure of labor market frictions in the differentiated sector. In other words, in this analysis we keep constant the level of labor market frictions in the homogeneous sector, $a_0$ (below we discuss the case of simultaneous reductions in labor market frictions in both sectors).
Finally, from (1.17) and (1.18) together with the formula for $\hat{N}$ we obtain

$$\text{sign}\{\hat{u}\} = \text{sign}\left\{ \left[ 1 - (2b - 1) \frac{\beta \zeta}{\beta - \zeta} \right] \hat{b} - (2b - 1) \frac{\beta \zeta}{\beta - \zeta} \frac{\delta x}{\delta x + \hat{\tau}} \right\}. $$

It is evident from this formula that lower frictions in the differentiated sector’s labor market (lower $b$) reduce unemployment if and only if

$$2b = \left( \frac{a}{a_0} \right)^{1/\alpha} < 1 + \frac{\beta - \zeta}{\beta \zeta},$$

i.e., if and only if labor market frictions are low in this sector to begin with. This condition is always satisfied when labor market frictions are higher in the homogeneous sector, i.e. $a_0 > a$. If labor market frictions in the differentiated sector are high, however, and the above inequality is reversed, then a reduction in $a$—and hence in $b$—may raise the rate of unemployment. In fact, the relationship between $b$ and the rate of unemployment has an inverted U shape as depicted in Figure 1.2.

To understand this result, note that changes in $a$ impact unemployment through two channels: the rate of unemployment in the differentiated sector $1 - x$, and the fraction of individuals searching for jobs in this sector $N/L$. Reductions in these labor market frictions raise $x$ and thereby reduce the sectoral rate of unemployment. On the other hand, such reductions attract more workers to the differentiated-product sector and thereby reduce the rate of unemployment if and only if the sectoral rate of unemployment is higher in the homogeneous sector (i.e., $x < x_0$). When $a_0 > a$ the sectoral rate of unemployment is higher in the homogenous sector and both channels lead to a reduction in the rate of unemployment. On the other hand, when $a > a_0$, the two channels conflict, and the latter, i.e., the reallocation of labor toward the differentiated sector, dominates when labor

---

31 In this derivation we use $bx = b_0 x_0 = \omega_0$, where $\omega_0$ is given in (1.13) and it does not vary with $a$, the measure of labor market frictions in the differentiated sector. Therefore, $\hat{b} = -\hat{x}$. Also note that $\hat{x}_0 = 0$ since $b_0 \equiv 1/2$. We have

$$uL\hat{u} = dN(x_0 - x) - Ndx = xN \left[ \left( \frac{x_0}{x} - 1 \right) \hat{N} + \hat{b} \right].$$

From (1.17), $x_0/x = (a/a_0)^{1/(1+\alpha)} = 2b$. Finally, combining these results with the expression for $\hat{N}$, we obtain the result in the text.
market frictions are high.\footnote{It can also be shown that in the symmetric case lower $a$ lead to increased entry of firms $M$, an increase in $N$ proportionately to $M$, and a more than proportional increase in employment $H$.}

Next consider a proportional reduction in both sectors’ labor market frictions, i.e., $\hat{a}_0 = \hat{a} < 0$. This has no effect on the search cost $b$ and does not impact the real consumption index $Q$ (see (1.17) and (1.26)). However, it reduces expected income $\omega_0$: from (1.13), $\hat{\omega}_0 = -\hat{a}/\alpha$. It therefore follows from $\omega_0N = \beta Q^k/(1 + \beta)$ that $\hat{N} = \hat{a}/\alpha$, and it follows from (1.17) that $\hat{x} = \hat{x}_0 = -\hat{a}/\alpha$. Using these expressions, and $N_0 + N = L$, the unemployment formula (1.18) implies that $\text{sign} \{\hat{u}\} = \text{sign} \{\hat{a}\}$. In other words, a reduction of labor market frictions at a common rate in both sectors reduces the rate of unemployment. Note that this sort of change in labor market frictions impacts unemployment through two channels, which may operate in opposite directions. On one hand, it raises tightness in each sector’s labor market, thereby reducing both sectoral rates of unemployment. On the other hand, it leads to a reallocation of workers from the differentiated to the homogeneous sector. If the sectoral rate of unemployment is higher in the differentiated sector, this reduces the rate of unemployment. But if the sectoral rate of unemployment is lower in the differentiated sector, this raises the rate of unemployment. Nevertheless, the composition effect is dominated by the sectoral effects.\footnote{From (1.18) and $N_0 + N = L$ we obtain $\hat{u}L\hat{u} = -N_0\hat{x}_0 - N\hat{x} + (\hat{x}_0 - \hat{x}) N\hat{N}$, where the first two expressions on the right-hand sided represent the sectoral effects and the third represents the composition effect. Since $\hat{x} = \hat{x}_0 = -\hat{N} = -\hat{a}/\alpha$, the sectoral effects dominate.}

Finally, consider changes in trade impediments. As the formula for the sign of changes in the rate of unemployment shows, a lower trade cost $\tau$ raises the rate of unemployment if and only if $b > 1/2$ (i.e., $a > a_0$).\footnote{The effect of a reduction in trade costs on unemployment is larger the larger is the share of trade in the sector’s revenue, i.e., the larger is $\delta_{xj}/(\delta_{ij} + \delta_{xj})$. When the economies are nearly closed, this effect is very small.} In this case the impact on unemployment operates only through the reallocation of labor across sectors, because sectoral unemployment rates do not change. In particular, more workers search for jobs in the differentiated sector when $\tau$ declines, and therefore aggregate unemployment rises when the differentiated sector has higher sectoral unemployment and aggregate unemployment falls when the differentiated sector has lower sectoral unemployment. Since the lowering of trade costs raises welfare,
this means that welfare and unemployment may respond in opposite directions to changes in trade costs.

We summarize the main findings of this section in

**Proposition 1.6.** In a symmetric world economy: (i) reductions in labor market frictions in the differentiated sectors at the same rate in both countries reduce aggregate unemployment if and only if \( a < a_0 \cdot [1 + (\beta - \zeta) / \beta \zeta]^{1+\alpha} \); (ii) reductions in labor market frictions at a common rate in both sectors and both countries reduce aggregate unemployment; (iii) reductions in trade impediments raise aggregate unemployment if and only if \( a > a_0 \).

An intriguing result is that lower trade barriers may raise unemployment. Lower trade costs make exporting more profitable in the differentiated-product sector. Moreover, the tightness in its labor market is not affected by falling trade costs. This increases demand for labor in the differentiated sector and leads to reallocation of workers towards this sector. Under these circumstances, the sectoral unemployment rates remain the same, but the aggregate unemployment rate may increase or decrease due to the compositional effect across sectors.\(^{35}\) The direction of this effect depends on whether the differentiated sector has a higher or lower unemployment rate.

Also note that unemployment can increase or decrease when welfare rises. That is, depending on the nature of the disturbance and the initial labor market frictions, unemployment and welfare can move in the same or in opposite directions. For this reason changes in unemployment do not necessarily reflect changes in welfare. This results from the standard property of search and matching models, in which unemployment is a productive activity which leads to creation of productive matches. Under these circumstances an expansion of the high-wage\(\backslash\)high-unemployment sector results in higher unemployment, but may also raise welfare.

\(^{35}\) See Felbermayr, Prat and Schmerer (2008) for a one-sector search model in which trade causes an increase in sectoral labor market tightness by reducing the real cost of vacancies, but naturally has no compositional effect.
1.5.2 Asymmetric Countries

We address in this section the impact of trade and labor market frictions on unemployment when the two countries are not symmetric. We first discuss some analytical results and then turn to numerical examples to illustrate the key mechanisms and various special cases.

In our working paper, Helpman and Itskhoki (2008), we provide analytical results for countries that are nearly symmetric, in the sense that they have no labor market frictions in the homogeneous sector and the difference between their labor market frictions in the differentiated sector is very small. Under these circumstances \( b_A > b_B \) implies that: (i) a reduction in a country’s labor market frictions reduces the rate of unemployment in its trade partner, yet it reduces home unemployment if and only if the initial levels of friction in the labor markets are low; and (ii) country \( B \) has a lower rate of unemployment if and only if the levels of labor market frictions are low to begin with. Evidently, a country’s level of unemployment depends not only on its own labor market frictions but also on those of its trade partner. Moreover, lower domestic labor market frictions do not guarantee lower unemployment relative to the trade partner, unless the frictions in both labor markets are low. As a result, one cannot infer differences in labor market rigidities from observations of unemployment rates. Richer results obtain with large labor market frictions, as we show below.

For our numerical illustrations we use a Pareto distribution of productivity levels,

\[
G(\Theta) = 1 - \left( \frac{\Theta_{\text{min}}}{\Theta} \right)^k, \quad \text{for } \Theta \geq \Theta_{\text{min}} \text{ and } k > 2.
\]

As is well known, the shape parameter \( k \) controls the dispersion of \( \Theta \), with smaller values of \( k \) representing more dispersion. It has to be larger than two for the variance of productivity to be finite. We show in the Appendix how the equilibrium conditions are simplified when productivity is distributed Pareto, and these equations are used to generate our numerical examples. One convenient implication of the Pareto assumption is that condition (1.11) implies \( \delta_{dj} + \delta_{xj} = kf_e \), and therefore aggregate revenue in the differentiated
sector is independent of labor market frictions and is the same in both countries. For the simulations we also assume that \( a_{0A} = a_{0B} = a_0 \), so that labor market frictions in the homogenous sector are the same in both countries, as a result of which expected income of workers, \( \omega_0 \), is also the same in both countries, i.e., \( \omega_{0A} = \omega_{0B} = \omega_0 \). In addition, we assume that \( a_A > a_B > a_0 \), so that labor market frictions are larger in the differentiated sectors of both countries than in their homogeneous sectors, and particularly so in country \( A \). This implies \( b_A > b_B > 1/2 \).

Combining (1.23) and (1.24), we obtain the following expression for global revenues generated in the differentiated sector:

\[
Q_A^c + Q_B^c = \frac{1}{\phi_2} \left[ M_A(\delta d_A + \delta x_A) + M_B(\delta d_B + \delta x_B) \right] = \frac{1 + \beta}{\beta} \omega_0 (N_A + N_B).
\]

Therefore, whenever \( Q_A^c + Q_B^c \) rises, the world-wide allocation of workers to the differentiated sector, \( N_A + N_B \), must also increase.\(^{36}\) Next note that Proposition 1.1 establishes that a reduction in trade costs raises \( Q_j \) in both countries. Therefore, the above discussion implies that a reduction in trade costs increases \( N_A + N_B \). In the Appendix we also show that \( N_A/N_B \) declines with reductions in \( \tau \) when \( b_A > b_B \). This then implies that \( N_B \), the number of job-seekers in the differentiated sector of country \( B \), necessarily increases. Since a fall in \( \tau \) does not affect sectoral labor market tightness, we conclude that a reduction in trade costs increases unemployment in country \( B \), which has lower labor market frictions in the differentiated sector. The effect on \( N_A \) and hence on the unemployment rate in country \( A \) is ambiguous, as we illustrate below.

The intuition behind this result is the following. Lower trade impediments increase the global size of the differentiated sector, which features increasing returns to scale and love of variety. As a result, the country with a more flexible labor market, which has a competitive edge in this sector, becomes more specialized in differentiated products.

\(^{36}\) Note that this result does not rely on the Pareto assumption. Under the Pareto assumption, however, we additionally have \( Q_A^c + Q_B^c = kf_e(M_A + M_B)/\phi_2 \), so that the total number of entrants into the differentiated sector must also increase. Moreover, in the Appendix we show that in this case \( \omega_0 N_j/M_j = \beta k f_e/(1 - \beta) \). That is, the number of workers searching for jobs in the differentiated sector relative to the number of firms depends on expected income \( \omega_0 \), but does not depend on the trade cost or labor market frictions in the differentiated sector.
That is, the number of entering firms, employment, and the number of job-seekers in the differentiated sector, all increase in country B. This compositional shift leads to a higher rate of unemployment in this country, because the sectoral rate of unemployment is higher in the differentiated sector. Finally, the reallocation of labor in country A may shift in either direction, depending on how strong the comparative advantage is (see below).

Figure 1.3: Unemployment as a function of $b_A$ when $b_B$ is low ($b_B = 0.55$)

Figure 1.3 depicts the response of unemployment rates to variation in country A's labor market frictions $a_A$, which changes monotonically $b_A$; the rising broken-line curve represents country B and the hump-shaped solid-line curve represents country A. Country B has $b_B = 0.55 > 1/2$, and therefore the two countries have the same rate of unemployment when $b_A = 0.55$. As $b_A$ rises, country A becomes more rigid. This raises initially the rate of unemployment in both countries, but B's rate of unemployment remains lower for a while. At some point, however, the rate of unemployment reaches a peak in country A, and it falls for further increases in $b_A$. As a result, the two rates of unemployment become equal again, after which further increases in rigidity in country A raise the rate of unemployment in country B and reduce it in country A, so that the rate of unemployment is higher in country B thereafter. The mechanism that operates here is that once the labor market frictions become high enough in country A, the contraction of the differentiated-

---

37 In Figures 1.3-1.4 we use the following parameters: $m_0 = 2v_0 = 1$, $f_s = 3$, $f_d = 1$, $f_c = 0.5$, $k = 2.5$, $\beta = 0.75$, $\zeta = 0.5$, $L = 0.1$ and $\tau = 1.1$ when it is held constant.
product sector leads to overall lower unemployment in $A$ despite the fact that its sectoral
unemployment rate is high. When $b_A$ is very high the sectoral unemployment rate is very
high, but no individuals search for jobs in this sector, as a result of which there is no
unemployment at all. This explains the hump in $A$’s curve. Note that in the range in
which the rate of unemployment falls in country $A$ the rate of unemployment keeps rising
in country $B$. The reason is that there is no change in market tightness in country $B$
and its differentiated-product sector becomes more competitive the more rigid the labor
market becomes in $A$. As a result the differentiated sector attracts more and more workers
in country $B$, which raises its rate of unemployment. The monotonic impact of country
$A$’s labor market rigidities on the unemployment rate in $B$ holds globally, and not only
around the symmetric equilibrium.\textsuperscript{38}

![Figure 1.4](image)

**Figure 1.4**: Unemployment as a function of $b_A$ when $b_B$ is high ($b_B = 0.65$)

Figure 1.4 is similar to Figure 1.3, except that now the level of labor market frictions
in country $B$ is higher, i.e., $b_B = 0.65 > 1/2$, and therefore the two curves intersect at
$b_A = 0.65$. Moreover, starting with a symmetric world that has these higher labor market
rigidities, increases in $b_A$ always raise unemployment in $B$ and reduce unemployment in $A$.
As a result, country $A$ has lower unemployment when $b_A > b_B$ and higher unemployment
when $b_A < b_B$. That is, in this case a more rigid country always has a lower unemployment

\textsuperscript{38} In Figures 1.3-1.4, country $A$ specializes in the homogeneous good when $b_A \geq b'$; in Figure 1.4, country $B$ specializes in the homogeneous good when $b_A \leq b''$; in Figure 1.3 country $B$ specializes in the differentiated good for $b_A \geq b''$. 

38
rate when it specializes (incompletely) in the low-unemployment sector.

A comparison between Figures 1.3 and 1.4 demonstrates the importance of the overall level of labor market rigidities for unemployment outcomes. When labor market frictions are high, a relatively more flexible country always has a higher rate of unemployment. Moreover, the rates of unemployment in the two countries move in opposite directions as labor market frictions change in either one of the countries. In contrast, when labor market rigidities are low and the difference in labor market frictions across countries is not large, the rate of unemployment is lower in a more flexible country and the rates of unemployment in both countries co-move in response to changes in labor market frictions.

![Figure 1.5: Unemployment as a function of τ when b_A and b_B are low (b_A = 0.6 and b_B = 0.56)](image)

The next three figures depict variations in unemployment in response to trade frictions, in the form of variable trade costs τ: Figure 1.5 for the case of low frictions in labor markets, Figure 1.6 for the case in which frictions are low in country B but high in A, and Figure 1.7 for the case in which frictions are high in both countries. In all three cases unemployment rises in B and falls in A when trade frictions decline. Nevertheless, the

---

39 In Figures 1.5-1.7 we use the following parameters: \(m_0 = 2v_0 = 1, f_s = 5, f_d = 1, f_c = 0.5, k = 2.5, \beta = 0.75, \zeta = 0.5, \text{ and } L = 0.1.\)

40 This pattern is not general. As we know, in the symmetric case lower trade impediments raise unemployment in both countries, which is also the case when countries are nearly symmetric. We can also provide examples in which the rigid country has a hump in its rate of unemployment as trade frictions vary.
rate of unemployment is not necessarily higher in A. In particular, unemployment is always higher in A when frictions in labor markets are low in both countries, yet unemployment is always higher in B when frictions in labor markets are high in both countries. In between, when labor market frictions are low in B and high in A, the relative rate of unemployment depends on trade impediments; it is lower in A when the trade frictions are low and lower in B when the trade frictions are high. This shows that labor market frictions interact with trade impediments in shaping unemployment.
Our analysis has focused on search and matching as the main frictions in labor markets, and we used $a_{0j} = 2v_{0j}/m_{0j}^{1+\alpha}$ and $a_j = 2v_{0j}/m_{0j}^{1+\alpha}$ as measures of labor market rigidity. Evidently, in this specification rigidity in a sector’s labor market is higher if either it is more costly to post vacancies in this sector or the matching process is less efficient in it.

We can also incorporate firing costs and unemployment benefits as additional sources of labor market rigidity. These labor market policies are widespread and they differ greatly across countries. But note that governments can also influence search and matching costs by facilitating the flow of information about job vacancies and about unemployed workers. Moreover, in some countries there are government agencies that directly assign unemployed workers to firms, and workers need to try these jobs in order to be eligible for unemployment benefits. In other words, government policies can influence not only firing costs and unemployment benefits, but also our measures of labor market frictions, $a_{0j}$ and $a_j$, which were analyzed above.

In order to save space, we briefly describe in this section results of a formal analysis conducted in our working paper, Helpman and Itskhoki (2008), under the simplifying assumption that there is full employment in the differentiated sector. This analysis can be extended to allow for labor market frictions in the homogenous-good sector, as in the earlier sections of the current paper.

With firing costs and unemployment benefits, $(x_j, b_j)$ remains a sufficient statistic for labor market frictions, with $b_j$ reinterpreted to represent the overall effective labor cost for a differentiated-sector firm, while the definition of $x_j$ does not change; it remains the same measure of labor market tightness in the differentiated sector. Importantly, the effects of $x_j$ and $b_j$ on the equilibrium outcomes described in Sections 3–5 do not change, except for the qualification of welfare effects to be discussed below.

Firing costs operate similarly to matching frictions, yielding a type of equivalence between the hiring and firing costs. Specifically, higher firing costs reduce labor market tightness $x_j$ and increase the effective labor cost $b_j$. Moreover, as long as unemployment benefits are not too high (see below), the effects of firing costs on welfare, trade patterns,
productivity and unemployment in trading economies, are the same as those of matching frictions. That is, all the earlier results of this paper extend to the case in which there are positive firing costs in addition to matching frictions.

Higher unemployment benefits always reduce equilibrium labor market tightness $x_j$, but they may increase or decrease the effective labor cost $b_j$. The intuition for this result is that unemployment benefits provide unemployment insurance to the workers on the one hand and a better outside option in the wage bargaining game on the other. Because higher unemployment benefits provide better unemployment insurance, workers are willing to search for jobs in a less tight labor market, with a higher sectoral rate of unemployment. This effect reduces the cost of hiring for firms. On the other side the better outside option of workers at the wage bargaining stage improves their bargaining position and increases the effective cost of labor to firms. Either one of these effects can dominate. Therefore $b_j$ may rise or decline in response to higher unemployment benefits. When $b_j$ decreases, it leads to an expansion of the differentiated sector, which raises welfare. But because unemployment benefits need to be financed by (lump-sum) taxes, the additional taxes required to finance higher unemployment benefits reduce disposable income and hurt welfare. Therefore on net welfare may rise or decline, but it definitely rises in response to a small rise in unemployment benefits that reduces $b_j$ when the initial level of these benefits is small.\footnote{Severance pay affects labor costs similarly to unemployment benefits, except that it has no impact on disposable income.}

We also show that firing costs and unemployment benefits not withstanding, international trade may raise unemployment in both countries. The reason is that trade attracts more workers to the differentiated sector without affecting sectoral labor market tightness. Therefore, when this sector has the lower labor market tightness, trade increases aggregate unemployment.
1.7 Concluding Comments

We have studied the interdependence of countries that trade homogeneous and differentiated products, and whose labor markets are characterized by search and matching frictions. Variation in labor market frictions and the interactions between trade impediments and labor market rigidities generate rich patterns of unemployment. For example, lower frictions in a country’s labor markets do not ensure lower unemployment, and unemployment and welfare can both rise in response to a policy change.

Contrary to the complex patterns regarding unemployment, the model yields sharp predictions about welfare. In particular, both countries gain from trade. Moreover, changes in one country’s labor market frictions can differentially impact welfare of the trade partners. For example, reducing a country’s frictions in the labor market of the differentiated sector raises competitiveness of its firms. This improves the foreign country’s terms of trade, but also crowds out foreign firms from the differentiated-product sector. As a result, welfare rises at home and declines abroad, because the terms-of-trade improvement in the foreign country is outweighed by the decline in the competitiveness of its firms. Nevertheless, a common reduction in labor market frictions in the differentiated sectors raises welfare in both countries. These results contrast with the implications of models of pure comparative advantage, in which movements in the terms of trade dominate the outcomes.\footnote{See, for example, Brügemann (2003) and Alessandria and Delacroix (2008). The former examines the support for labor market rigidities in a Ricardian model in which the choice of regime impacts comparative advantage. The latter analyzes a two-country model with two goods, in which every country specializes in a different product and governments impose firing taxes. The authors find that a coordinated elimination of these taxes yields welfare gains for both counties, yet no country on its own has an incentive to do it.}

We also show that labor market frictions confer comparative advantage, and that differences in these labor market characteristics shape trade flows. In particular, the country with relatively lower labor market frictions in the differentiated sector exports differentiated products on net and imports homogeneous goods. Moreover, the larger the difference in these relative frictions, the lower is the share of intra-industry trade. These are testable implications about trade flows and international patterns of specialization.
In addition, we show that trade raises total factor productivity in the differentiated-product sectors of both countries, while productivity does not change in the homogeneous sector. And productivity is higher in the country with relatively lower labor market frictions in the differentiated sector.

An important conclusion from our analysis is that simple one-sector macro models that ignore compositional effects may be inadequate for assessing labor market frictions, and especially so in a world of integrated economies. Moreover, a focus on terms-of-trade as the major channel of the international transmission of shocks misses the impact of competitiveness, which can dominate economic outcomes.
2. INEQUALITY AND UNEMPLOYMENT IN A GLOBAL ECONOMY

(with Elhanan Helpman and Stephen J. Redding)

2.1 Introduction

Two core issues in international trade are the allocation of resources across economic activities and the distribution of incomes across factors of production. Recent research has emphasized the allocation of resources across heterogeneous firms, but has largely concentrated on heterogeneity in the product market (productivity and size) rather than the labor market (workforce composition and wages). Developing trade models that incorporate both product and labor market heterogeneity is therefore important for explaining firm data and understanding the consequences of trade liberalization. To the extent that wages vary across firms within sectors, reallocations of resources across firms provide an additional channel for international trade to influence income distribution.

In this paper, we develop a new framework for examining the distributional consequences of trade that incorporates this channel and captures three plausible features of product and labor markets. First, there is heterogeneity in firm productivity, which generates differences in firm profitability. Second, search and matching frictions in the labor market imply that workers outside a firm are imperfect substitutes for those inside the firm, which gives rise to multilateral bargaining between each firm and its workers. Third, workers are heterogeneous in terms of match-specific ability, which can be imperfectly observed by firms. Together these three components of the model generate variation in wages across firms within industries and imply that trade liberalization affects income distribution.

In the closed economy, we derive a sufficient statistic for wage inequality, which deter-
mines all scale-invariant measures of wage inequality, such as the Coefficient of Variation, Gini Coefficient and Theil Index. This sufficient statistic depends on the dispersion parameters for worker ability and firm productivity as well as other product and labor market parameters that influence workforce composition. Greater dispersion of worker ability has ambiguous effects on wage inequality, because it affects both relative wages and employment levels across firms. In contrast, greater dispersion of firm productivity raises wage inequality, because more productive firms pay higher wages. This close relationship between wage and productivity dispersion within industries receives empirical support from studies using micro data, including Davis and Haltiwanger (1991) and Faggio, Silvanes and Van Reenen (2007).

In the open economy, only the most productive firms export; firms of intermediate productivity serve only the domestic market; and the least productive firms exit without producing because they cannot cover fixed production costs. Exporting firms have higher revenue than non-exporting firms, screen workers more intensively, employ workforces of higher average ability, and pay higher wages. The model therefore is consistent with empirical findings that exporters pay higher wages than non-exporters (see for example Bernard and Jensen, 1997) and that a substantial component of this wage difference is explained by workforce composition (see for example Kaplan and Verhoogen, 2006, Schank, Schnabel and Wagner, 2007, and Munch and Skaksen, 2008).

The open economy wage distribution is a mixture of the wage distributions for employees of domestic and exporting firms, with exporters paying higher wages than non-exporters. Therefore the open economy wage distribution depends on the fraction of exporters and the exporter wage premium, as well as on the sufficient statistic for wage inequality from the closed economy. Opening closed economies to trade increases wages and employment at high-productivity exporters relative to low-productivity domestic firms. As a result the opening of trade raises wage inequality for any measure of wage inequality that respects second-order stochastic dominance.

Once the economy is open to trade, the relationship between wage inequality and the fraction of exporting firms is non-monotonic. In particular, in the limiting case in which all
firms export, there is the same level of wage inequality in the open and closed economies. When all firms export, a small reduction in the share of exporting firms increases wage inequality, because of the lower wages paid by domestic firms. Similarly, when no firm exports, a small increase in the share of exporting firms raises wage inequality, because of the higher wages paid by exporters.

These results for wage inequality hold for each country and for arbitrary asymmetries between countries. Our analysis is therefore consistent with empirical findings of increased wage inequality in both developed and developing countries following trade liberalization (see for example the survey by Goldberg and Pavcnik, 2007). These predictions contrast with those of the Stolper-Samuelson Theorem of the Heckscher-Ohlin model, which implies rising wage inequality in developed countries and declining wage inequality in developing countries. As changes in wage inequality in our framework are driven by reallocations across firms, our analysis is also consistent with empirical evidence that the vast majority of the reallocation observed following trade liberalization takes place within rather than between industries. Finally, as wage inequality in our model arises from heterogeneity in unobserved match-specific ability, our results are also compatible with the observation that changes in the return to observed skills typically account for a relatively small share of the overall increase in wage inequality following trade liberalization (see for example Attanasio, Goldberg and Pavcnik, 2004, and Menezes-Filho et al., 2008).

The presence of labor market frictions gives rise to equilibrium unemployment, which introduces a distinction between the distribution of income across all workers and the distribution of wages across employed workers. Labor market frictions have symmetric effects on firms of all productivities, and hence do not affect the sufficient statistic for wage inequality in the closed economy, but do affect unemployment and hence income inequality. Opening closed economies to trade generally raises unemployment, because it reduces the share of matched workers that are hired, but under some conditions it can also raise the share of job-seekers that are matched, which reduces unemployment.

Together these predictions for wage inequality and unemployment imply that the distributional consequences of trade liberalization are quite different from those in neoclassical
trade theory. Workers employed by high-productivity exporting firms receive higher real wages in the open economy than in the closed economy. In contrast, workers employed by low-productivity domestic firms may receive lower or higher real wages in the open economy than in the closed economy. Finally, because unemployment is typically higher in the open economy than in the closed economy, there are more workers with the lowest real income in the open economy.

In addition to these distributional consequences for \textit{ex post} welfare, the opening of trade also has implications for \textit{ex ante} expected welfare. \textit{Ex ante} workers face income risk because of unemployment and wage dispersion across firms. With incomplete insurance, the increase in unemployment and wage inequality induced by the opening of trade increases income risk. Nonetheless, as long as workers are risk neutral, expected welfare gains are ensured.

Our paper is related to recent research on firm heterogeneity in international trade building on the influential framework of Melitz (2003).\footnote{For alternative approaches to firm heterogeneity and trade, see Bernard et al. (2003) and Yeaple (2005).} While this literature yields rich predictions for the product market, firms pay workers with the same characteristics the same wage irrespective of firm productivity, which sits awkwardly with a large empirical literature that finds an employer–size wage premium and rent–sharing within firms.

This study is also related to the literature on international trade and labor market frictions. One strand of this literature assumes that firm wages are related to productivity, revenue or profits because of “efficiency wage” or “fair wage” concerns, including Amiti and Davis (2008) and Egger and Kreickemeier (2008, 2009).\footnote{A related trade literature examines efficiency wages and unemployment, including Davis and Harrigan (2007).} In contrast, the relationship between firm wages and revenue in our framework is derived from worker heterogeneity and labor market frictions. As a result, our model implies quite different determinants of wage inequality and unemployment, which include the dispersion of worker ability and the other product and labor market parameters that influence workforce composition, as well as the dispersion of firm productivity.
Another strand of this literature, more closely related to our own work, examines the implications of search frictions for trade, including Davidson, Martin and Matusz (1988, 1999), Felbermayr, Prat and Schmerer (2008, 2009) and Helpman and Itskhoki (Chapter I of this dissertation). Our main point of departure from this literature is the introduction of worker heterogeneity and imperfect screening of workers by firms, which generates wage inequality that is influenced by both trade liberalization and labor market frictions. While Davidson, Matusz and Schevchenko (2008) also develop a model of firm and worker heterogeneity with an exporter wage premium, they assume one-to-one matching between firms and workers and only two types of firms and workers. In contrast, our framework allows for an endogenous measure of matches for each firm and continuous distributions of firm productivity and worker ability. As a result, the opening of trade changes both employment and wages across firms of heterogeneous productivity, which changes the wage distribution and generates the non-monotonic relationship between wage inequality and trade openness.

Our paper is also related to the broader economics literature on matching. One strand of this literature is concerned with competitive assignment models, and investigates the conditions under which there is assortative matching, including Heckman and Honore (1990), Ohnsorge and Trefler (2007), Legros and Newman (2007), and Costinot and Vogel (2009). In contrast, another strand of this literature considers search frictions in the labor market (see Pissarides, 2000, for a review of this literature). Within the search literature, several approaches have been taken to explaining wage differences across workers. One influential line of research follows Burdett and Mortensen (1998) and Mortensen (2003) in analyzing wage dispersion in models of wage posting and random search. Another important line of research examines wage dispersion when both firms and workers are heterogeneous, including models of pure random search such as Shimer and Smith (2000), Albrecht and Vroman (2002), and models incorporating on-the-job-search such as Postel-Vinay and Robin (2002) and Lentz (2007).

Our framework combines search frictions with firm screening of worker ability, which results in both unemployment and residual wage inequality. Search frictions generate equi-
librium unemployment and ensure that workers with the same observed and unobserved characteristics can be matched with firms of different productivity. Screening induces variation in firm workforce composition despite random search, because more productive firms screen more intensively, hire workers of higher average ability and hence pay higher wages.\(^3\) As a result, workers matched with firms of different productivity receive different wages.

The remainder of the paper is structured as follows. Section 2 outlines the model and its sectoral equilibrium. Section 3 presents our results on sectoral wage inequality, Section 4 presents our results on sectoral unemployment, and Section 5 presents our results on sectoral income inequality. Section 6 examines alternative ways of closing the model to study the feedback from general equilibrium outcomes to the sectoral equilibrium. Section 7 concludes. The Appendix contains technical details, including proofs of various results.

### 2.2 Sectoral Equilibrium

The key predictions of our model relate to the distribution of wages and employment across firms and workers within sectors. As these predictions hold for any given values of expected worker income, prices in other sectors and aggregate income, we begin in this section by characterizing sectoral equilibrium for given values of these variables, before determining these variables in general equilibrium in Section 2.6 below. Throughout the analysis of sectoral equilibrium, all prices, revenues and costs are measured in terms of a numeraire, where the choice of numeraire is specified in the analysis of general equilibrium in Section 2.6.

#### 2.2.1 Model Setup

We consider a world of two countries, home and foreign, where foreign variables are denoted by an asterisk. In each country there is a continuum of workers who are \textit{ex ante} identical. Initially, we assume workers are risk neutral, but we extend the analysis to

\(^3\) An alternative potential approach to generating variation in workforce composition is directed search, as considered for example in Moen (1997) and Mortensen and Wright (2002).
introduce risk aversion in Section 2.6. The supply of workers to the sector is endogenously determined by expected income. Demand within the sector is defined over the consumption of a continuum of horizontally differentiated varieties and takes the constant elasticity of substitution (CES) form. The real consumption index for the sector \( Q \) is therefore defined as follows:

\[
Q = \left[ \int_{j \in J} q(j)^\beta \, dj \right]^{1/\beta}, \quad 0 < \beta < 1,
\]

(2.1)

where \( j \) indexes varieties; \( J \) is the set of varieties within the sector; \( q(j) \) denotes consumption of variety \( j \); and \( \beta \) controls the elasticity of substitution between varieties. To simplify notation, we suppress the sector subscript except where important, and while we display expressions for home, analogous relationships hold for foreign. The price index dual to \( Q \) is denoted by \( P \) and depends on the prices \( p(j) \) of individual varieties \( j \). Given this specification of sectoral demand, the equilibrium revenue of a firm is:

\[
r(j) = p(j)q(j) = Aq(j)^\beta,
\]

(2.2)

where \( A_i \) is a demand-shifter for sector \( i \), which depends on the dual price index for the sector \( P_i \), prices in other sectors \( (P_{-i}) \) and aggregate income \( (\Omega) \).\(^4\) The precise functional form for the demand-shifter, \( A_i = \tilde{A}_i (P, \Omega) \), depends on the specification of demand across sectors, as discussed further when we analyze general equilibrium below.

Each firm takes the demand shifter as given when making its decisions, because it supplies one of a continuum of varieties within the sector, and is therefore of measure zero relative to the sector as a whole.

The product market is modeled in the same way as in Melitz (2003). There is a competitive fringe of potential firms who can choose to enter the differentiated sector by

\[^4\text{We use bold typeface to denote vectors, so that } P_{-i} \text{ is a vector of all price indexes } P_j \text{ other than } i \text{ and } P \text{ is a vector of all price indexes. As is well known, the demand function for a variety } j \text{ in sector } i \text{ can be expressed as:}
\]

\[
q_i(j) = A_i^\alpha p_i(j)^{-\frac{\beta}{\alpha}}, \text{ where } A_i = \left( E_i / \int_{j \in J} p_i(j)^{-\frac{\beta}{\alpha}} \, dj \right)^{1-\beta},
\]

and \( E_i \) is total expenditure on varieties within the sector, which depends on aggregate income \( (\Omega) \), the price index for the sector \( (P_i) \) and prices for all other sectors \( (P_{-i}) \).
paying an entry cost of \( f_e > 0 \). Once a firm incurs the sunk entry cost, it observes its productivity \( \theta \), which is independently distributed and drawn from a Pareto distribution \( G_\theta (\theta) = 1 - (\theta_{\min}/\theta)^z \) for \( \theta \geq \theta_{\min} > 0 \) and \( z > 1 \). The Pareto distribution is not only tractable, but together with our other assumptions implies a Pareto firm-size distribution, which provides a reasonable approximation to observed data (see Axtell, 2001). Since in equilibrium all firms with the same productivity behave symmetrically, we index firms by \( \theta \) from now onwards.

Once firms observe their productivity, they decide whether to exit, produce solely for the domestic market, or produce for both the domestic and export market. Production involves a fixed cost of \( f_d > 0 \) units of the numeraire. Similarly, exporting involves a fixed cost of \( f_x > 0 \) units of the numeraire and an iceberg variable trade cost, such that \( \tau > 1 \) units of a variety must be exported in order for one unit to arrive in the foreign market.

Output of each variety \((y)\) depends on the productivity of the firm \((\theta)\), the measure of workers hired \((h)\), and the average ability of these workers \((\bar{a})\):

\[
y = \theta h^\gamma \bar{a}, \quad 0 < \gamma < 1.
\]  

(2.3)

This production technology can be interpreted as capturing either human capital complementarities (e.g., production in teams where the productivity of a worker depends on the average productivity of her team) or a managerial time constraint (e.g., a manager with a fixed amount of time who needs to allocate some time to each worker). In Appendix, we derive the production technology under each of these interpretations. A key feature of the production technology is complementarities in worker ability, where the productivity of a worker is increasing in the abilities of other workers employed by the firm.\(^5\)

The labor market features heterogeneity in worker ability and search and matching frictions. Worker ability is assumed to be match-specific, independently distributed and drawn from a Pareto distribution, \( G_a (a) = 1 - (a_{\min}/a)^k \) for \( a \geq a_{\min} > 0 \) and \( k > 1 \). Since

\(^5\) The existence of these production complementarities is the subject of a long line of research in economics, including Lucas (1978), Rosen (1982), and Garicano (2000). For empirical evidence see for example Moretti (2004).
worker ability is match-specific and independently distributed, a worker’s ability draw for a given match conveys no information about ability draws for other potential matches. Search and matching frictions are modeled following the standard Diamond-Mortensen-Pissarides approach. A firm that pays a search cost of \( bn \) units of the numeraire can randomly match with a measure of \( n \) workers, where the search cost \( b \) is endogenously determined by the tightness of the labor market as discussed below.

Consistent with a large empirical literature in labor economics, we assume that match-specific worker ability cannot be costlessly observed when firms and workers are matched.\(^6\) Instead, we assume that firms can undertake costly investments in worker screening to obtain an imprecise signal of worker ability, which is in line with a recent empirical literature on firm screening and other recruitment policies.\(^7\) To capture the idea of an imprecise signal in as tractable a way as possible, we assume that by paying a screening cost of \( ca^\delta/\delta \) units of the numeraire, where \( c > 0 \) and \( \delta > 0 \), a firm can identify workers with an ability below \( a_c \).\(^8\) Screening costs are increasing in the ability threshold \( a_c \) chosen by the firm, because more complex and costlier tests are required for higher ability cutoffs.\(^9\)

This specification of worker screening is influenced by empirical evidence that more productive firms not only employ more workers but also screen more intensively, have workforces of higher average ability and pay higher wages. Each of these features emerges naturally from our specification of production and screening, as demonstrated below, because production complementarities imply a greater return to screening for more productive firms and the costs of screening are the same for all firms. Our formulation also ensures that the multilateral bargaining game between firms and workers over the surplus

---

\(^6\) For example, Altonji and Pierret (2003) find that as employers learn about worker productivity the wage equation coefficients on easily observed characteristics, such as education, fall relative to the coefficients on hard-to-observe correlates of worker productivity.

\(^7\) For empirical evidence on the resources devoted by firms to the screening of job applicants, see for example Barron, Black and Loewenstein (1987), Pellizari (2005), and Autor and Scarborough (2007).

\(^8\) In this formulation, there is a fixed cost of screening, even when the screening is not informative, i.e., when \( a_c = a_{\text{min}} \). We focus on interior equilibria in which firms of all productivities choose screening tests that are informative, \( a_c > a_{\text{min}} \), and so the fixed cost of screening is always incurred. As we show below, this is the case when the screening cost, \( c \), is sufficiently small.

\(^9\) There are therefore increasing returns to scale in screening. All results generalize immediately to the case where the screening costs are separable in \( a_c \) and \( n \) and linear in \( n \).
from production remains tractable. As the only information revealed by screening is which workers have match-specific abilities above and below $a_c$, neither the firm nor the workers know the match-specific abilities of individual workers, and hence bargaining occurs under conditions of symmetric information.

### 2.2.2 Firm’s Problem

The complementarities between workers’ abilities in the production technology provide the incentive for firms to screen workers. By screening and not employing workers with abilities less than $a_c$, a firm reduces output (and hence revenue and profits) by decreasing the measure of workers hired ($h$), but raises output by increasing average worker ability ($\bar{a}$). Since there are diminishing returns to the number of workers hired ($0 < \gamma < 1$), output can be increased by screening as long as there is sufficient dispersion in worker ability (sufficiently low $k$).\(^{10}\) With a Pareto distribution of worker ability, a firm that chooses a screening threshold $a_c$ hires a measure $h = n (a_{\min}/a_c)^k$ of workers with average ability $\bar{a} = ka_c/(k - 1)$. Therefore the production technology can be re-written as follows:

$$y = \kappa_y \theta n^{\gamma} a_c^{1-\gamma k}, \quad \kappa_y \equiv \frac{k}{k-1} a_{\min}^{\gamma k},$$

(2.4)

where we require $0 < \gamma k < 1$ for a firm to have an incentive to screen.\(^{11}\)

Given consumer love of variety and a fixed production cost, no firm will ever serve the export market without also serving the domestic market. If a firm exports, it allocates its output ($y(\theta)$) between the domestic and export markets ($y_d(\theta)$ and $y_x(\theta)$, respectively) to equate its marginal revenues in the two markets, which from (2.2) implies

$$\left[\frac{y_d(\theta)}{y_x(\theta)}\right]^{\beta-1} = \tau^{-\beta} (A^*/A).$$

Therefore a firm’s total revenue can be expressed as

\(^{10}\) Since production complementarities provide the incentive for firms to screen, the marginal product of workers with abilities below $a_c$ is negative, as shown in Appendix. While worker screening is a key feature of firms’ recruitment policies, and production complementarities provide a tractable explanation for it, other explanations are also possible, such as fixed costs of maintaining an employment relationship (e.g. in terms of office space or other scarce resources).

\(^{11}\) In contrast, when $\gamma > 1/k$, no firm screens and the model reduces to the model of Helpman and Itskhoki (see Chapter I), which has no screening or worker heterogeneity. We do not discuss this case here.
follows:

\[ r(\theta) \equiv r_d(\theta) + r_x(\theta) = \Upsilon(\theta)^{1-\beta}Ay(\theta)^\beta, \quad (2.5) \]

where \( r_d(\theta) \equiv Ay_d(\theta)^\beta \) is revenue from domestic sales and \( r_x(\theta) \equiv A^*[y_x(\theta)/\tau]^\beta \) is revenue from exporting. The variable \( \Upsilon(\theta) \) captures a firm’s “market access,” which depends on whether it chooses to serve both the domestic and foreign markets or only the domestic market:

\[ \Upsilon(\theta) \equiv 1 + I_x(\theta)\tau^{-\frac{\beta}{1-\beta}} \left( \frac{A^*}{A} \right)^{\frac{1}{1-\beta}}, \quad (2.6) \]

where \( I_x(\theta) \) is an indicator variable that equals one if the firm exports and zero otherwise.\(^{12}\)

After having observed its productivity, a firm chooses whether or not to produce, whether or not to export, the measure of workers to sample, and the screening ability threshold (and hence the measure of workers to hire). Once these decisions have been made, the firm and its hired workers engage in strategic bargaining with equal weights over the division of revenue from production in the manner proposed by Stole and Zwiebel (1996a,b). The only information known by firms and workers at the bargaining stage is that each hired worker has an ability greater than \( a_c \). Therefore, the expected ability of each worker is \( \bar{a} = k/(k-1)a_c \), and each worker is treated as if they have an ability of \( \bar{a} \).

Combining (2.3) and (2.5), firm revenue can be written as \( r = \Upsilon(\theta)^{1-\beta}A(\theta\bar{a})^\beta h^{\beta\gamma} \), which is continuous, increasing and concave in \( h \). As the fixed production, fixed exporting, search and screening costs have all been sunk before the bargaining stage, all other arguments of firm revenue are fixed. Furthermore, the outside option of hired workers is unemployment, whose value we normalize to zero. Therefore, the solution to the bargaining game is that the firm receives the fraction \( 1/(1+\beta\gamma) \) of revenue (2.5), while each worker receives the fraction \( \beta\gamma/(1+\beta\gamma) \) of average revenue per worker (see Appendix).

Anticipating the outcome of the bargaining game, the firm maximizes its profits. Com-

\(^{12}\) Note that \([y_d(\theta)/y_x(\theta)]^{\beta-1} = \tau^{-\beta}(A^*/A)\) and \(y_d(\theta) + y_x(\theta) = y(\theta)\) imply \(y_d(\theta) = y(\theta)/\Upsilon(\theta)\) and \(y_x(\theta) = y(\theta)[\Upsilon(\theta) - 1]/\Upsilon(\theta)\), and hence \(r_d(\theta) = r(\theta)/\Upsilon(\theta)\) and \(r_x(\theta) = r(\theta)[\Upsilon(\theta) - 1]/\Upsilon(\theta)\).
bining (2.4), (2.5), and (2.6) this profit maximization problem can be written as:

$$
\pi(\theta) \equiv \max_{n \geq 0, \ a_c \geq a_{\min}, \ I_x \in \{0,1\}} \left\{ \frac{1}{1 + \beta \gamma} \left[ 1 + I_x \tau^{-\frac{1}{1-\beta}} \left( \frac{A^*}{A} \right)^{\frac{1}{1-\beta}} \right]^{1-\beta} A \left( \kappa \theta n^{\gamma} a_c^{1-\gamma k} \right)^{\beta} \right. \\
\left. - bn - \frac{c}{\delta} a_c^{\delta} - f_d - I_x f_x \right\},
$$

(2.7)

where $I_x$ is the export status indicator, which equals 1 when the firm exports and 0 otherwise.

The firm’s decision whether or not to produce and whether or not to export takes a standard form. The presence of a fixed production cost implies that there is a zero-profit cutoff for productivity, $\theta_d$, such that a firm drawing a productivity below $\theta_d$ exits without producing. Similarly, the presence of a fixed exporting cost implies that there is an exporting cutoff for productivity, $\theta_x$, such that a firm drawing a productivity below $\theta_x$ does not find it profitable to serve the export market. Given that a large empirical literature finds evidence of selection into export markets, where only the most productive firms export, we focus on values of trade costs for which $\theta_x > \theta_d > \theta_{\min}$. The firm market access variable is therefore determined as follows:

$$
\Upsilon(\theta) = \begin{cases} 
1, & \theta < \theta_x, \\
\Upsilon_x, & \theta \geq \theta_x,
\end{cases} \quad \Upsilon_x \equiv 1 + \tau^{-\frac{1}{1-\beta}} \left( \frac{A^*}{A} \right)^{\frac{1}{1-\beta}} > 1.
$$

(2.8)

The firm’s first-order conditions for the measure of workers sampled ($n$) and the screening ability threshold ($a_c$) are:

$$
\frac{\beta \gamma}{1 + \beta \gamma} r(\theta) = bn(\theta), \\
\frac{\beta (1 - \gamma k)}{1 + \beta \gamma} r(\theta) = ca_c(\theta)^{\delta}.
$$

These conditions imply that firms with larger revenue sample more workers and screen

---

13 For empirical evidence of selection into export markets, see for example Bernard and Jensen (1995) and Roberts and Tybout (1997).
to a higher ability threshold. While the measure of workers hired, \( h = n (a_{\text{min}}/a_c)^k \), is increasing in the measure of workers sampled, \( n \), it is decreasing in the screening ability threshold, \( a_c \). Under the assumption \( \delta > k \), firms with larger revenue not only sample more workers but also hire more workers. Finally, from the division of revenue in the bargaining game, the total wage bill is a constant share of revenue, which implies that firm wages are monotonically increasing in the screening ability cutoff:

\[
  w(\theta) = \frac{\beta \gamma r(\theta)}{1 + \beta \gamma h(\theta)} = b \frac{n(\theta)}{h(\theta)} = b \left[ \frac{a_c(\theta)}{a_{\text{min}}} \right]^k .
\]  

(2.9)

Thus, firms with larger revenue have higher screening ability cutoffs and pay higher wages, but the expected wage conditional on being sampled is the same across all firms:

\[
  \frac{w(\theta)h(\theta)}{n(\theta)} = b,
\]

which implies that workers have no incentive to direct their search.\(^{14}\) Combining the measure of workers hired, \( h = n (a_{\text{min}}/a_c)^k \), with the first-order conditions above yields the following relationship between firm wages and the measure of workers hired:

\[
  \ln w(\theta) = \text{constant} + \frac{k}{\delta - k} \ln h(\theta).
\]

Therefore, under the assumption \( \delta > k \), the model exhibits an employer-size wage premium, where firms that employ more workers pay higher wages.

Using the firms’ first-order conditions, firm revenue (2.5) and the production technology (2.4), we can solve explicitly for firm revenue as a function of firm productivity (\( \theta \)), the demand shifter (\( A \)), the search cost (\( b \)) and parameters:

\[
  r(\theta) = \kappa_r \left[ c^{-\beta(1-\gamma k)/\delta} b^{-\beta\gamma Y(\theta)(1-\beta)} A^{\theta^\beta} \right]^{1/\Gamma},
\]

(2.10)

\(^{14}\) We note that search frictions and wage bargaining alone are not enough to generate wage variation across firms in our model. From the firm’s first-order condition for the number of workers sampled, each firm equates workers’ share of revenue per sampled worker to the common search cost. In the special case of our model without worker heterogeneity and screening, all sampled workers are hired, which implies that each firm’s wage is equal to the common search cost (see Chapter I).
where $\Gamma \equiv 1 - \beta \gamma - \beta (1 - \gamma k)/\delta > 0$, and the constant $\kappa_r$ is defined in Appendix. An implication of this expression is that the relative revenue of any two firms depends solely on their relative productivities and relative market access: $r(\theta') = (\theta'/\theta'')^{\beta/\Gamma} (\Upsilon(\theta')/\Upsilon(\theta''))^{\beta/\Gamma} r(\theta'')$.

Finally, using the two first-order conditions in the firm’s problem (2.7), firm profits can be expressed solely in terms of firm revenue and the fixed production and exporting costs:

$$
\pi(\theta) = \frac{\Gamma}{1 + \beta \gamma} r(\theta) - f_d - I_x(\theta) f_x.
$$

(2.11)

2.2.3 Sectoral Variables

To determine sectoral equilibrium, we use the recursive structure of the model. In a first bloc of equations, we solve for the tightness of the labor market ($x, x^*$) and search costs ($b, b^*$) in each country. In a second bloc of equations, we solve for the zero-profit productivity cutoffs ($\theta_d, \theta_d^*$), the exporting productivity cutoffs ($\theta_x, \theta_x^*$), and sectoral demand shifters ($A, A^*$). A third and final bloc of equations determines the remaining components of sectoral equilibrium: the dual price index ($P, P^*$), the real consumption index ($Q, Q^*$), the mass of firms ($M, M^*$), and the size of the labor force ($L, L^*$). As discussed above, we solve for sectoral equilibrium for given values of expected income in the sector ($\omega, \omega^*$), prices in other sectors ($P_{-i}, P_{-i}^*$) and aggregate incomes ($\Omega, \Omega^*$), which are determined in general equilibrium below.

Labor Market Tightness and Hiring Costs

Following the standard Diamond-Mortensen-Pissarides search model, the search cost ($b$) is assumed to be increasing in labor market tightness ($x$):

$$
b = \alpha_0 x^{\alpha_1}, \quad \alpha_0 > 1, \quad \alpha_1 > 0,
$$

(2.12)

where labor market tightness equals the ratio of workers sampled ($N$) to workers searching for employment in the sector ($L$): $x = N/L$.\textsuperscript{15} Under the assumption of risk neutrality,
the supply of workers searching for employment in the sector depends on expected worker income, which equals the probability of being sampled \((x)\) times the expected wage conditional on being sampled \((w(\theta) h(\theta)/n(\theta) = b)\) from the analysis above:

\[
\omega = xb, \tag{2.13}
\]

where we discuss in Section 2.6 how this condition is modified under the assumption of risk aversion. Together (2.12) and (2.13) determine the search cost and labor market tightness \((b, x)\) for a given value of expected income \((\omega)\):

\[
b = \alpha_0^{\frac{1}{1+\alpha_1}} \omega^{\frac{\alpha_1}{1+\alpha_1}} \quad \text{and} \quad x = \left(\frac{\omega}{\alpha_0}\right)^{\frac{1}{1+\alpha_1}}, \tag{2.14}
\]

where we assume \(\alpha_0 > \omega\) so that \(0 < x < 1\), as discussed in Section 2.6 below. Analogous relationships determine search costs and labor market tightness \((b^*, x^*)\) for a given value of expected income \((\omega^*)\) in foreign. The search cost in (2.14) depends solely on parameters of the search technology \((\alpha_0, \alpha_1)\) and expected income \((\omega)\). In particular, we have

**Lemma 2.1.** The search cost \(b\) and the measure of labor market tightness \(x\) are both increasing in expected worker income \(\omega\).

**Proof:** The lemma follows immediately from equation (2.14).

When we characterize general equilibrium in Section 2.6 below, we specify conditions under which expected income \((\omega)\) is constant and those under which it changes with other endogenous variables. While we treat \(\omega\) as given in solving for sectoral equilibrium, our results for sectoral inequality and unemployment continue to hold when it responds in general equilibrium to other endogenous variables, except where otherwise discussed.

vacancies and decreasing in the productivity of the matching technology, while \(\alpha_1\) depends on the weight of vacancies in the Cobb-Douglas matching function.
Productivity Cutoffs and Demand

The two productivity cutoffs can be determined using firm revenue (2.10) and profits (2.11). The productivity cutoff below which firms exit ($\theta_d$) is determined by the requirement that a firm with this productivity makes zero profits:

$$\frac{\Gamma}{1 + \beta \gamma \kappa_r} \left[ c^{\frac{\beta (1 - \gamma k)}{\delta}} b^{-\beta \gamma A \theta_d^\beta} \right]^{1/\Gamma} = f_d. \quad (2.15)$$

Similarly, the exporting productivity cutoff above which firms export ($\theta_x$) is determined by the requirement that at this productivity a firm is indifferent between serving only the domestic market and serving both the domestic and foreign markets:

$$\frac{\Gamma}{1 + \beta \gamma \kappa_r} \left[ c^{\frac{\beta (1 - \gamma k)}{\delta}} b^{-\beta \gamma A \theta_x^\beta} \right]^{1/\Gamma} \left[ \Upsilon_x^{(1-\beta)/\Gamma} - 1 \right] = f_x. \quad (2.16)$$

These two conditions imply the following relationship between the productivity cutoffs:

$$\left[ \Upsilon_x^{(1-\beta)/\Gamma} - 1 \right] \left( \frac{\theta_x}{\theta_d} \right)^{\beta/\Gamma} = \frac{f_x}{f_d}. \quad (2.17)$$

In equilibrium, we also require the free entry condition to hold, which equates the expected value of entry to the sunk entry cost. Using the zero profit and exporting cutoff conditions, (2.15) and (2.16) respectively, and the relationship between variety revenues for firms with different productivities, the free entry condition can be written as:

$$f_d \int_{\theta_d}^\infty \left[ \left( \frac{\theta}{\theta_d} \right)^{\beta/\Gamma} - 1 \right] dG_\theta + f_x \int_{\theta_x}^\infty \left[ \left( \frac{\theta}{\theta_x} \right)^{\beta/\Gamma} - 1 \right] dG_\theta = f_e. \quad (2.18)$$

Equations (2.15), (2.16) and (2.18) can be used to solve for home’s productivity cutoffs and the demand shifter ($\theta_d, \theta_x, A$) for a given value of the foreign demand shifter ($A^*$), which only influences home sectoral equilibrium through exporter market access ($\Upsilon_x > 1$).\(^{16}\) Three analogous equations can be used to solve for foreign variables ($\theta^*_d, \theta^*_x, A^*$) for a given value of $A$. Together these six equations allow us to solve for the productivity

\(^{16}\) In a symmetric equilibrium $A = A^*$ and $\Upsilon_x = 1 + \frac{\tau}{1-\gamma}$, which implies that the ratio of the two productivity cutoffs is pinned down by (2.17) alone.
cutoffs and demand shifters in the two countries ($\theta_d, \theta_x, A, \theta_d^*, \theta_x^*, A^*$) for given values of search costs ($b, b^*$), which were determined in the previous bloc of equations. Having solved for the productivity cutoffs and demand shifters, firm market access in each country ($Y(\theta), Y^*(\theta)$) follows immediately from (2.6).

The productivity cutoffs and demand shifter depend on two dimensions of trade openness in (2.15), (2.16) and (2.18). First, both depend on an extensive margin of trade openness, as captured by the ratio of the productivity cutoffs $\rho \equiv \theta_d/\theta_x \in [0,1]$, which determines the fraction of exporting firms $[1 - G_{\theta}(\theta_x)]/[1 - G_{\theta}(\theta_d)] = \rho z$. Second, both depend on an intensive margin of trade openness, as captured by the market access variable, $Y_x > 1$, which determines the ratio of revenues from domestic sales and exporting, as discussed in footnote 12. These two dimensions of trade openness are linked through the relationship between the productivity cutoffs (2.17).

**Expenditure, Mass of Firms and the Labor Force**

Having solved for the demand shifter for sector $i$ ($A_i$), the price index for that sector ($P_i$) can be determined from consumer optimization given prices in all other sectors ($P_{-i}$) and aggregate income ($\Omega$):

$$A_i = \tilde{A}_i (P, \Omega).$$  \hfill (2.19)

Having solved for the demand shifter ($A$) and the price index ($P$) for the sector (we now drop the sectoral subscript $i$), the real consumption index ($Q$) follows from consumer optimization, which from (2.2) implies:

$$Q = (A/P)^{1/\beta},$$  \hfill (2.20)

and yields total expenditure within the sector $E = PQ$. Similar relationships determine the foreign price index, real consumption index and total expenditure ($P^*, Q^*, E^*$).

The mass of firms within the sector ($M$) can be determined from the market clearing condition that total domestic expenditure on differentiated varieties equals the sum of the
revenues of domestic and foreign firms that supply varieties to the domestic market:

\[ E = M \int_{\theta_d}^{\infty} r_d (\theta) \, dG_{\theta} (\theta) + M^* \int_{\theta_x^*}^{\infty} r_x^* (\theta) \, dG_{\theta} (\theta). \] (2.21)

From \( r_d (\theta) = r (\theta) / \Upsilon (\theta), \) \( r_x (\theta) = r (\theta) (\Upsilon (\theta) - 1) / \Upsilon (\theta), \)\(^{17}\) and total firm revenue (2.10), domestic and foreign revenue can be expressed in terms of variables that have already been determined \((\theta_d, \theta_x, \Upsilon (\theta))\). Therefore we can solve for the mass of firms in each country \((M, M^*)\) from (2.21) and a similar equation for foreign.

The mass of workers searching for employment in the sector \((L)\) can be determined by noting that total labor payments are a constant fraction of total revenue from the solution to the bargaining game:

\[ \omega L = M \int_{\theta_d}^{\infty} w (\theta) \, h (\theta) \, dG_{\theta} (\theta) = M \beta \gamma \int_{\theta_d}^{\infty} r (\theta) \, dG_{\theta} (\theta), \] (2.22)

where we have solved for the mass of firms \((M)\) and total firm revenue \((r(\theta))\) above, and where a similar equation determines the sectoral labor force in foreign \((L^*)\). Finally, we also require that the sectoral labor force is less than or equal to the supply of labor \((L \leq \bar{L})\), as discussed in Section 2.6 below. This completes our characterization of sectoral equilibrium.

### 2.2.4 Firm-specific Variables

Having characterized sectoral equilibrium, we can solve for all firm-specific variables using the following two properties of the model. First, from firm revenue (2.10), the relative revenue of any two firms depends solely on their relative productivities and relative market access. Second, from firm profits (2.10), the lowest productivity firm with productivity \(\theta_d\) makes zero profits. Combining these two properties with the firm’s first-order conditions above, all firm-specific variables can be written as functions of firm productivity \((\theta)\), firm market access \((\Upsilon (\theta))\), the zero-profit productivity cutoff \((\theta_d)\), search costs \((b)\) and

\(^{17}\) See footnote 12.
parameters:

\[ r(\theta) = \Upsilon(\theta)^{1-\beta} \cdot r_d \cdot \left( \frac{\theta}{\tau_d} \right)^{\frac{\beta}{\tau}}, \]
\[ n(\theta) = \Upsilon(\theta)^{1-\beta} \cdot n_d \cdot \left( \frac{\theta}{\nu_d} \right)^{\beta \tau}, \]
\[ a_c(\theta) = \Upsilon(\theta)^{1-\beta} \cdot a_d \cdot \left( \frac{\theta}{\tau_d} \right)^{\beta \tau}, \]
\[ h(\theta) = \Upsilon(\theta)^{1-\beta (1-k/\delta)} \cdot h_d \cdot \left( \frac{\theta}{\nu_d} \right)^{\frac{\beta (1-k/\delta)}{\tau}}, \]
\[ w(\theta) = \Upsilon(\theta)^{k(1-\beta)} \cdot w_d \cdot \left( \frac{\theta}{\nu_d} \right)^{\frac{k \beta}{\tau}}, \]

\[ r_d \equiv \frac{1+\beta\gamma}{1-\delta}, \quad n_d \equiv \frac{\beta \gamma}{1-\delta}, \quad a_d \equiv \left[ \frac{\beta (1-k) \gamma}{1-\delta} \right]^{1/\delta}, \quad h_d \equiv \frac{\beta \gamma}{1-\delta} \left[ \frac{\beta (1-k) \gamma}{1-\delta} \right]^{-k/\delta}, \quad w_d \equiv b \left[ \frac{\beta (1-k) \gamma}{1-\delta} \right]^{-k/\delta}, \]

where market access (\( \Upsilon(\theta) \)) is determined as a function of firm productivity in (2.8). Note that firm-specific variables only depend on sectoral and general equilibrium through the zero-profit cutoff productivity (\( \theta_d \)), firm market access (\( \Upsilon(\theta) \)) and hence the exporting cutoff productivity (\( \theta_x \)), and search costs (\( b \)).

The solutions for firm-specific variables (2.23) capture a number of key features of the heterogeneity observed across firms within sectors. More productive firms not only have higher revenue, profits and employment, as in the benchmark model of firm heterogeneity of Melitz (2003), but also pay higher wages as shown in Figure 2.1. These results are consistent with empirical evidence of rent-sharing whereby higher firm revenue and profits are shared with workers through higher wages (e.g., van Reenen, 1996) and with the large empirical literature that finds an employer size-wage premium (see the survey by Oi and
Additionally, the differences in wages across firms are driven by differences in workforce composition. More productive firms have workforces of higher average ability, which are more costly to replace in the bargaining game, and therefore pay higher wages. These features of the model are consistent with findings from matched employer-employee datasets that the employer-size wage premium is largely explained by the positive correlation between employment and a firm’s average worker fixed effect. The reason more productive firms have workforces of higher average ability in the model is that they screen more intensively, which also receives empirical support. An emerging literature on firm recruitment policies provides evidence of more intensive screening policies for larger firms and higher-wage matches.

Our framework features residual wage inequality in the sense that workers with the same observed and unobserved characteristics receive different wages depending on the firm with which they are matched. While this feature of the model is consistent with recent empirical evidence (see for example Autor, Katz and Kearney, 2008, Lemieux, 2006, and Mortensen, 2003), competitive models without labor market frictions can in principle generate residual wage inequality if there are worker characteristics observed by the firm but not by the econometrician. While the firm may observe additional worker characteristics, several features of the data suggest that this is not the only explanation for residual wage inequality. In competitive frictionless models, arbitrage typically eliminates differences in wages across firms for workers with the same characteristics. Yet the empirical literature using matched employer-employee datasets finds that firm fixed effects make a substantial contribution towards wage variation after controlling for time-varying worker

---

18 Combining the solutions for firm revenue and employment in (2.23) with the Pareto productivity distribution, firm revenue and employment are also Pareto distributed, with shape parameters that depend on the dispersion of firm productivity, the dispersion of worker ability, and product and labor market parameters that influence workforce composition. See Helpman et al. (2008a) for further discussion.

19 See Figure 3 in Abowd, Kramarz and Margolis (1999) and the discussion in Abowd and Kramarz (1999).

20 For example, Barron, Black and Loewenstein (1987) find that expenditures on screening workers are positively and significantly related to employer size, while Pellizari (2005) finds that matches created through more intensive screening pay higher wages.
characteristics and worker fixed effects (see for example Abowd, Creecy and Kramarz, 2002). The contribution of the firm fixed effects arises naturally in our framework from labor market frictions.\footnote{In our model, expected wages vary with both firm productivity and a worker’s match-specific ability draw. As more productive firms screen to higher ability thresholds, only workers with ability draws above those thresholds receive the higher wages paid by more productive firms. It follows that both the average wage and wage dispersion are increasing in a worker’s match-specific ability draw, as discussed in Helpman, Itskhoki and Redding (2008a).}

Finally, our solutions for firm-specific variables are also consistent with findings from the recent empirical literature on exports and firm performance. As a result of fixed costs of exporting, there is a discrete jump in firm revenue at the productivity threshold for entry into exporting ($\theta_x$), where $\Upsilon(\theta)$ jumps from 1 to $\Upsilon_x > 1$, which implies a discrete jump in all other firm variables. Therefore, exporters not only have higher revenue and employment than non-exporters, as in the benchmark model of firm heterogeneity of Melitz (2003), but also pay higher wages, as found empirically by Bernard and Jensen (1995, 1997) and many subsequent studies. These differences in wages between exporters and non-exporters are associated with differences in workforce composition, which is also consistent with recent empirical findings using matched employer-employee datasets (see for example Kaplan and Verhoogen, 2006, Schank, Schnabel and Wagner 2007, and Munch and Skaksen, 2008).

### 2.3 Sectoral Wage Inequality

While workers are \textit{ex ante} identical and have the same expected income, there is \textit{ex post} wage inequality because workers receive different wages depending on the employer with whom they are matched. In this section, we consider the within-sector distribution of wages across employed workers. This sectoral wage distribution is a weighted average of the distributions of wages for workers employed by domestic firms, $G_{w,d}(w)$, and for workers employed by exporters, $G_{w,x}(w)$, with weights equal to the shares of employment
in the two groups of firms.

\[
G_w (w) = \begin{cases} 
S_{h,d}G_{w,d}(w) & \text{for } w_d \leq w \leq w_d/\rho^{\beta k/\delta} \\
S_{h,d} & \text{for } w_d/\rho^{\beta k/\delta} \leq w \leq w_d\Gamma_x^{k(1-\beta)/\delta}/\rho^{\beta k/\delta} \\
S_{h,d} + (1 - S_{h,d})G_{w,x}(w) & \text{for } w \geq w_d\Gamma_x^{k(1-\beta)/\delta}/\rho^{\beta k/\delta},
\end{cases} \tag{2.24}
\]

where \( \rho \) and \( \Gamma_x \) are the extensive and intensive margins of trade openness defined above; \( w_d = w(\theta_d) \) is the wage paid by the least productive firm in (2.23); \( w_d/\rho^{\beta k/\delta} = w(\theta_x^-) \) is the wage paid by the most productive non-exporter; and \( w_d\Gamma_x^{k(1-\beta)/\delta}/\rho^{\beta k/\delta} = w(\theta_x^+) \) is the wage paid by the least productive exporter. Note that \( w_d \) depends on general equilibrium variables only through search costs \((b)\). The share of workers employed by domestic firms, \( S_{h,d} \), can be evaluated using the Pareto productivity distribution and the solution for firm-specific variables (2.23) as:

\[
S_{h,d} = \frac{1 - \rho^z - \beta(1-k/\delta)}{1 + \rho^z - \beta(1-k/\delta)} \left[ \frac{\Gamma_x^{(1-\beta)(1-k/\delta)}}{\rho^{\beta k/\delta}} - 1 \right].
\]

The distributions of wages across workers employed by domestic and exporting firms can also be derived from the solutions for firm-specific variables (2.23). Given that productivity is Pareto distributed and both wages and employment are power functions of productivity, the distribution of wages across workers employed by domestic firms is a truncated Pareto distribution:

\[
G_{w,d}(w) = \frac{1 - \left( \frac{w_d}{w} \right)^{1+1/\mu}}{1 - \rho^z - \beta(1-k/\delta)} \quad \text{for } w_d \leq w \leq w_d/\rho^{\beta k/\delta}. \tag{2.25}
\]

Similarly, the distribution of wages across workers employed by exporters, \( G_{w,x}(w) \), is an untruncated Pareto distribution:

\[
G_{w,x}(w) = 1 - \left[ \frac{w_d}{w} \Gamma_x^{k(1-\beta)/\delta}/\rho^{\beta k/\delta} \right]^{1+1/\mu} \quad \text{for } w \geq w_d\Gamma_x^{k(1-\beta)/\delta}/\rho^{\beta k/\delta}. \tag{2.26}
\]

The wage distributions for workers employed by domestic firms and exporters have the
same shape parameter, \( 1 + 1/\mu \), where \( \mu \) is defined as:

\[
\mu \equiv \frac{\beta k/\delta}{z\Gamma - \beta}, \quad \text{where}\; \Gamma \equiv 1 - \beta \gamma - \frac{\beta}{\delta}(1 - \gamma k).
\] (2.27)

For the mean and variance of the sectoral wage distribution to be finite, we require \( 0 < \mu < 1 \) and hence \( z\Gamma > 2\beta \), which is satisfied for sufficiently large \( z \) (a sufficiently skewed firm productivity distribution).\(^{22}\)

**Sectoral Wage Inequality in the Closed Economy**

The closed economy wage distribution can be obtained by taking the limit \( \rho \to 0 \) in the open economy wage distribution (2.24). In the closed economy, the share of employment in domestic firms is equal to one, and the sectoral wage distribution across workers employed by domestic firms is an untruncated Pareto distribution with lower limit \( w_d \) and shape parameter \( 1 + 1/\mu \). Given an untruncated Pareto distribution, all scale-invariant measures of inequality, such as the Coefficient of Variation, the Gini Coefficient and the Theil Index, depend solely on the distribution’s shape parameter. None of these measures depends on the lower limit of the support of the wage distribution \( (w_d) \), and they therefore do not depend on search costs \( (b) \) or expected worker income \( (\omega) \). While these variables affect the mean of the wage distribution, they do not affect its dispersion. An important implication of this result is that the model’s predictions for wage inequality are robust to alternative ways of closing the model in general equilibrium to determine expected income \( (\omega) \).

**Proposition 2.1.** In the closed economy, \( \mu \) is a sufficient statistic for sectoral wage inequality. In particular: (i) The Coefficient of Variation of wages is \( \mu/\sqrt{1 - \mu^2} \); (ii) The Lorenz Curve is represented by \( s_w = 1 - (1 - s_h)^{1/(1+\mu)} \), where \( s_h \) is the fraction of workers and \( s_w \) is the fraction of their wages when workers are ordered from low to high wage earners; (iii) The Gini Coefficient is \( \mu/(2 + \mu) \); and (iv) The Theil Index is \( \mu - \ln(1 + \mu) \).

\(^{22}\) While we concentrate on the wage distribution, as this is typically the subject of the economic debate over the impact of trade liberalization, the income distribution could also be influenced by profits. As discussed in footnote 18, the model can be also used to determine the distribution of revenue (and hence profits) across firms.
Proof: See Appendix.

Evidently, sectoral wage inequality is monotonically increasing in $\mu$ (the lower the shape parameter of the wage distribution $1 + 1/\mu$, the greater wage inequality). Using this result, we can analyze the relationship between sectoral wage inequality and the dispersion of firm productivity and worker ability.

**Proposition 2.2.** In the closed economy, inequality in the sectoral distribution of wages is increasing in firm productivity dispersion (lower $z$), and increasing in worker ability dispersion (lower $k$) if and only if $z^{-1} + \delta^{-1} + \gamma > \beta^{-1}$.

Proof: The proof follows immediately from Proposition 2.1 and the definition of $\mu$.

Since more productive firms pay higher wages, greater dispersion in firm productivity (lower $z$) implies greater sectoral wage inequality. In contrast, greater dispersion in worker ability (lower $k$) has an ambiguous effect on sectoral wage inequality because of two countering forces. On the one hand, a reduction in $k$ increases relative employment in more productive firms (from (2.23)) that pay higher wages, which increases wage inequality. On the other hand, a reduction in $k$ decreases relative wages paid by more productive firms (from (2.23)), which reduces wage inequality. When the parameter inequality in the proposition is satisfied, the change in relative employment dominates the change in relative wages, and greater dispersion in worker ability implies greater sectoral wage inequality. Additionally, the sectoral wage distribution depends on the other product and labor market parameters that influence workforce composition. These include the concavity of revenue ($\beta$) and production ($\gamma$), and the convexity of screening costs ($\delta$), as can be seen from the definition of $\mu$ in (2.27).

The model’s prediction that sectoral wage inequality is closely linked to the dispersion of firm productivity receives strong empirical support. In particular, Davis and Haltiwanger (1991) show that wage dispersion across plants within sectors accounts for a large share of overall wage dispersion, and is responsible for more than one third of the growth in overall wage dispersion in U.S. manufacturing between 1975 and 1986. Additionally, they find that between-plant wage dispersion is strongly related to between-plant size.
dispersion, which in our model is driven by productivity dispersion. Similarly, Faggio, Salvanes and van Reenen (2007) show that a substantial component of the increase in individual wage inequality in the United Kingdom in recent decades has occurred between firms within sectors and is linked to increased productivity dispersion between firms within sectors.

While greater firm productivity dispersion (associated for example with innovations such as Information and Communication Technologies, ICTs) is one potential source of increased wage inequality in the model, another potential source is international trade as considered in the next section. Indeed, both greater firm productivity dispersion and international trade raise wage inequality through the same mechanism of greater dispersion in firm revenue and wages within industries, and both raise measured productivity at the industry level through reallocations of resources across firms.

Open Versus Closed Economy

The sectoral wage distribution in the open economy depends on the sufficient statistic for wage inequality in the closed economy ($\mu$) and the extensive and intensive measures of trade openness ($\rho$ and $\Upsilon_x$, respectively). In the two limiting cases of $\rho = 0$ (no firm exports) and $\rho = 1$ (all firms export), the open economy wage distribution is an untruncated Pareto with shape parameter $1 + 1/\mu$. From Proposition 2.1, all scale-invariant measures of inequality for an untruncated Pareto distribution depend solely on the distribution’s shape parameter. Therefore there is the same level of wage inequality in the open economy when all firms export as in the closed economy.

To characterize sectoral wage inequality in the open economy when $0 < \rho < 1$ (only some firms export), we compare the actual open economy wage distribution ($G_w (w)$) to a counterfactual wage distribution ($G^c_w (w)$). For the counterfactual wage distribution, we choose an untruncated Pareto distribution with the same shape parameter as the wage distribution in the closed economy ($1 + 1/\mu$) but the same mean as the wage distribution in the open economy. An important feature of this counterfactual wage distribution is that it has the same level of inequality as the closed economy wage distribution. Therefore, if
we show that there is more inequality with the open economy wage distribution than with
the counterfactual wage distribution, this will imply that there is more wage inequality in
the open economy than in the closed economy.

The counterfactual wage distribution has two other important properties, as shown
formally in Appendix. First, the lowest wage in the counterfactual wage distribution \( (w^c_w) \)
lies strictly in between the lowest wage paid by domestic firms \( (w_d) \) and the lowest wage
paid by exporters \( (w(\theta^+_x)) \) in the actual open economy wage distribution. Otherwise,
the counterfactual wage distribution would have a mean either lower or higher than the
actual open economy wage distribution, which contradicts the requirement that the two
distributions have the same mean. Second, the counterfactual wage distribution has a
smaller slope than the actual wage distribution at \( w(\theta^+_x) \). Otherwise, the counterfactual
wage distribution would have a greater density than the actual wage distribution for
\( w \geq w(\theta^+_x) \), and would therefore have a higher mean than the actual wage distribution.

Figure 2.2: Cumulative distribution function of wages

Together, these two properties imply that the relative location of the cumulative dis-
bution functions for actual and counterfactual wages is as shown in Figure 2.2. The
actual and counterfactual cumulative distributions intersect only once and the actual dis-
tribution lies above the counterfactual distribution for low wages and below it for high
wages.\textsuperscript{23} This pattern provides a sufficient condition for the counterfactual wage distribution to second-order stochastically dominate the wage distribution in the open economy. Therefore, for all measures of inequality that respect second-order stochastic dominance, the open economy wage distribution exhibits greater inequality than the counterfactual wage distribution. It follows that the wage distribution in the open economy exhibits more inequality than the wage distribution in the closed economy. This result holds independently of whether the opening of trade affects expected worker income ($\omega$), because $\omega$ affects the lower limit of the actual open economy wage distribution (and hence the lower limit of the counterfactual wage distribution), but does not affect the comparison of levels of inequality between the two distributions.

**Proposition 2.3.** (i) Sectoral wage inequality in the open economy when some but not all firms export is strictly greater than in the closed economy; and (ii) Sectoral wage inequality in the open economy when all firms export is the same as in the closed economy.

**Proof:** The proof follows from the discussion above, as shown formally in Appendix.

Proposition 2.3 holds for asymmetric countries and irrespective of which of the model’s parameters are the source of the asymmetry across countries. While for simplicity we focus on the case of two countries, extending the analysis to a world of many countries is straightforward. The proposition identifies an alternative mechanism for trade to influence wage inequality from the Stolper-Samuelson Theorem of traditional trade theory. While the Stolper-Samuelson theorem emphasizes reallocations of resources across sectors that change the relative rewards of skilled and unskilled workers, the changes in wage inequality in Proposition 2.3 arise from changes in wages and employment across firms within sectors. This prediction that the opening of trade can increase wage inequality for asymmetric countries, the emphasis on reallocation across firms within sectors, and the focus on residual wage inequality find support in the recent empirical literature on trade and wage inequality reviewed in Goldberg and Pavcnik (2007).

\textsuperscript{23} Note that the actual and counterfactual distributions can intersect either above the wage at the most productive non-exporter, $w(\sigma_{\gamma})$, as shown in Figure 2.2, or below it. In both cases, the actual and counterfactual distributions have the properties discussed in the text.
While the opening of closed economies to trade raises wage inequality, it also increases average wages conditional on being employed (in terms of the numeraire). Under the conditions discussed in Section 2.6, expected worker income \( \omega \) is constant in general equilibrium, and hence so are search costs \( b \) and the lower limit of the wage distribution \( w_d \). As a result, the discrete increase in wages at the productivity threshold for exporting implies that the open economy wage distribution first-order stochastically dominates the closed economy wage distribution, as can be seen from (2.24). To the extent that expected worker income \( \omega \) is increased by the opening of trade, as discussed in Section 2.6, this raises the lower limit of the open economy wage distribution \( w_d \) and further reinforces the first-order stochastic dominance result.

Since sectoral wage inequality when all firms export is the same as in the closed economy, but sectoral wage inequality when only some firms export is higher than in the closed economy, the relationship between sectoral wage inequality and the fraction of exporters is non-monotonic. An increase in the share of firms that export can either raise or reduce sectoral wage inequality depending on the initial share of firms that export \( \rho \), which in turn depends on the relative productivity cutoffs \( \rho \). As \( \rho \to 0 \), no firm exports, and a small increase in the share of firms that export raises sectoral wage inequality, because of the higher wages paid by exporters. As \( \rho \to 1 \), all firms export, and a small reduction in the share of firms that export increases sectoral wage inequality, because of the lower wages paid by domestic firms. Therefore the model points to the initial level of trade openness as a relevant control in examining the empirical relationship between wage inequality and trade openness.

While fixed and variable trade costs \( f_x \) and \( \tau \), respectively) both influence sectoral wage inequality, they do so through slightly different mechanisms, because they have different effects on the extensive and intensive margins of trade openness \( \rho \) and \( \Upsilon_x \), respectively). This can be seen most clearly for symmetric countries, where the intensive margin depends on variable trade costs alone \( \Upsilon_x = 1 + \tau \frac{u_d}{1 - \tau} \), and changes in the fixed costs of exporting affect only the extensive margin \( \rho \). More generally, for asymmetric countries, the intensive margin depends on variable trade costs and relative sectoral demand levels
\[ Y_x = 1 + \tau^{\frac{1}{1-\beta}} \left( \frac{A^*}{A} \right)^{1-\beta} \], and fixed and variable trade costs affect both margins of trade openness.

To illustrate the non-monotonic relationship between sectoral wage inequality and trade openness, Figure 2.3 graphs the variation in the Theil Index of wage inequality with symmetric countries as we vary the fixed costs of exporting \((f_x)\) and hence the extensive margin of trade openness \((\rho)\). Similar non-monotonic relationships are observed as we vary variable trade costs \((\tau)\) and for other measures of wage inequality such as the Gini Coefficient.

### 2.4 Sectoral Unemployment

The presence of labor market frictions generates equilibrium unemployment. Workers can be unemployed either because they are not matched with a firm or because their match-specific ability draw is below the screening threshold of the firm with which they are matched. Therefore the sectoral unemployment rate \(u\) can be expressed as one minus the product of the hiring rate \(\sigma\) and the tightness of the labor market \(x\):

\[
u = \frac{L - H}{L} = 1 - \frac{H N}{N L} = 1 - \sigma x,
\]

Figure 2.3: Theil index of sectoral wage inequality
where $\sigma \equiv H/N$, $H$ is the measure of hired workers, $N$ is the measure of matched workers, and $L$ is the measure of workers seeking employment in the sector.

The sectoral tightness of the labor market ($x$) in (2.14) depends on the search friction parameter ($\alpha_0$) and expected worker income ($\omega$). Therefore the tightness of the labor market is not directly affected by trade openness and is only indirectly affected in so far as trade openness influences $\omega$. While in this section we examine the comparative statics of unemployment for a given value of $\omega$, in Section 2.6 we determine its value in general equilibrium. As part of that analysis, we provide conditions under which $\omega$ is unaffected by trade openness, and examine how the comparative statics of unemployment change when it responds to trade openness.

In contrast, the sectoral hiring rate ($\sigma$) depends directly on trade openness, which influences firm revenues and hence screening ability thresholds. Using the Pareto productivity distribution, the sectoral hiring rate can be expressed as a function of the extensive and intensive margins of trade openness ($\rho$ and $\Upsilon_x$ respectively), the sufficient statistic for wage inequality ($\mu$), and other parameters, as shown in Appendix:

$$
\sigma = \varphi(\rho, \Upsilon_x) \cdot \frac{1}{1 + \mu} \cdot \left[ \frac{\Gamma \frac{c_0^\delta}{f_d}}{\beta (1 - \gamma k) \Gamma} \right]^{k/\delta},
$$

(2.29)

where the term in square parentheses is the hiring rate of the least productive firm ($h_d/n_d$) and:

$$
\varphi(\rho, \Upsilon_x) \equiv \frac{1 + \left[ \frac{\Upsilon_x^{(1-\beta)(1-k/\delta)}}{1 - 1} - 1 \right] \rho^{z-\beta(1-k/\delta)/\Gamma}}{1 + \left[ \frac{\Upsilon_x^{(1-\beta)}}{1 - \gamma k} - 1 \right] \rho^{z-\beta/\Gamma}}.
$$

Evidently, we have $\varphi(0, \Upsilon_x) = 1$ and $0 < \varphi(\rho, \Upsilon_x) < 1$ for $0 < \rho \leq 1$, since $\Upsilon_x > 1$ and $\delta > k$.

**Sectoral Unemployment in the Closed Economy**

The closed economy hiring rate can be obtained by taking the limit $\rho \to 0$ in (2.29); it depends solely on model parameters for a given value of expected worker income ($\omega$). The hiring rate depends on the screening cost ($c$) but not on the search friction parameter...
(α₀), and it depends on both the dispersion of firm productivity (z) and the dispersion of worker ability (k). Combining the closed economy hiring rate (σ) from (2.29) with labor market tightness (x) from (2.14), we can examine the determinants of the sectoral unemployment rate (u). This yields

**Proposition 2.4.** Let ω be constant. Then the closed economy sectoral unemployment rate u is increasing in the search friction α₀, decreasing in the screening cost c, increasing in the dispersion of firm productivity (lower z), and can be either increasing or decreasing in the dispersion of worker ability (lower k).

**Proof:** The proof follows immediately from equations, (2.14), (2.27), (2.28) and (2.29).

It is clear from this proposition that search and screening costs have quite different effects on sectoral unemployment. As the search cost (b) rises in response to a rise in α₀, the sectoral tightness of the labor market (x) falls, which increases the sectoral unemployment rate. In contrast, as the screening cost (c) increases, firms screen less intensively, which increases the sectoral hiring rate (σ), and thereby reduces the sectoral unemployment rate.

It is also clear that dispersion of firm productivity has a different effect on sectoral unemployment from the dispersion of worker ability. Since more productive firms screen more intensively, an increase in the dispersion of firm productivity (lower z) reduces the sectoral hiring rate (σ), which increases sectoral unemployment. In contrast, an increase in the dispersion of worker ability (lower k) has an ambiguous effect on the sectoral hiring rate (σ) and hence on the sectoral unemployment rate. On the one hand, more dispersion in worker ability increases the probability of being hired conditional on being sampled (([aₘᵢₙ/aₖ()]^k) for a given screening threshold (aₖ(θ) > aₘᵢₙ), which reduces sectoral unemployment. On the other hand, more dispersion in worker ability induces firms to screen more intensively (lower k raises aₖ(θ) from (2.23)), which increases sectoral unemployment.²⁴ Like sectoral wage inequality, the sectoral unemployment rate also depends on

---

²⁴ There is an additional compositional effect of greater dispersion in worker ability. From (2.23), lower k increases n(θ) for all firms, but has a larger effect for more productive firms. Since more productive firms screen more intensively, this change in sectoral composition increases sectoral unemployment.
other product and labor market parameters that influence workforce composition ($\beta$, $\gamma$ and $\delta$, which enter $\Gamma$ and $\mu$).

**Open Versus Closed Economy**

For a given value of expected worker income ($\omega$), the opening of trade only affects the sectoral unemployment rate ($u$) through the hiring rate ($\sigma$). Furthermore, the open economy hiring rate (2.29) equals the closed economy hiring rate times the fraction $\varphi(\rho, \Upsilon_x)$, which depends on both the extensive and intensive margins of trade openness. This fraction is strictly less than one in an equilibrium where some firms export ($0 < \rho \leq 1$), and hence the sectoral hiring rate ($\sigma$) is strictly lower in the open economy than in the closed economy. We therefore have

**Proposition 2.5.** Let $\omega$ be invariant to trade. Then the sectoral unemployment rate $u$ is strictly higher in the open economy than in the closed economy.

*Proof:* The proof follows immediately from equation (2.29).

The opening of trade results in an expansion in the revenue of exporters and a contraction in the revenue of non-exporters, which changes industry composition towards more productive firms that screen more intensively, and thereby increases sectoral unemployment. While the sectoral unemployment rate is higher in the open economy than in the closed economy, once the economy is open to trade, the relationship between sectoral unemployment and trade openness (like the relationship between sectoral wage inequality and trade openness) can be non-monotonic. In particular, the sectoral unemployment rate can be either monotonically decreasing in trade openness, or it can exhibit an inverted U-shape, where sectoral unemployment is initially increasing in trade openness before decreasing in trade openness, as discussed further in Helpman, Itskhoki and Redding (2008b).

As our analysis in this section considers the case where $\omega$ is invariant to trade, it focuses solely on changes in unemployment due to changes in firms’ screening policies ($\sigma$). In our analysis of general equilibrium in Section 2.6 below, we show that the opening of trade can also affect $\omega$ and have offsetting effects on unemployment through labor market
tightness \((x)\).

2.5 Sectoral Income Inequality

The sectoral distribution of income depends on both the sectoral distribution of wages and the unemployment rate, where unemployed workers all receive the same income of zero. Since there is greater wage inequality and a higher unemployment rate in the open economy than in the closed economy, it follows that there is also greater income inequality. As shown in Appendix, the Theil Index of income inequality \((T_i)\) can be expressed as the following function of the Theil Index of wage inequality \((T_w)\) and the unemployment rate \((u)\),

\[
T_i = T_w - \ln(1 - u).
\] (2.30)

A similar result holds for the Gini Coefficient of income inequality \((G_i)\), which can be expressed in terms of the Gini Coefficient of wage inequality \((G_w)\) and the unemployment rate:

\[
G_i = (1 - u)G_w + u.
\] (2.31)

Sectoral Income Inequality in the Closed Economy

The comparative statics for sectoral income inequality in the closed economy follow from those for sectoral wage inequality and unemployment above.

**Proposition 2.6.** Let \(\omega\) be constant. In the closed economy sectoral income inequality, as measured by either the Theil Index or the Gini Coefficient, is increasing in the search friction \(\alpha_0\), decreasing in the screening cost \(c\), increasing in the dispersion of firm productivity (lower \(z\)), and can be either increasing or decreasing in the dispersion of worker ability (lower \(k\)).

\[25\] The Theil Index of inequality allows an exact decomposition of overall inequality into within and between-group inequality (Bourguignon, 1979), where the groups are here employed and unemployed workers. In general, the Gini Coefficient does not allow such a decomposition, but in the present case all unemployed workers receive the same income of zero, which is strictly less than the lowest income of an employed worker. Therefore, a similar decomposition can be undertaken for the Gini Coefficient, as shown in Appendix.
Proof: The proposition follows immediately from Propositions 2.2 and 2.4 and the expressions for the Theil Index (2.30) and Gini Coefficient (2.31).

While a rise in the search friction $\alpha_0$ (which raises the search cost $b$) or a reduction in the screening cost $c$ leaves sectoral wage inequality unchanged, it raises sectoral unemployment and hence increases sectoral income inequality. On the other hand, a rise in the dispersion of firm productivity (lower $z$) increases sectoral income inequality through both higher wage inequality and higher unemployment. In contrast, a rise in the dispersion of worker ability (lower $k$) has an ambiguous effect on sectoral wage inequality, unemployment and income inequality. Furthermore, greater dispersion of worker ability can raise sectoral wage inequality while at the same time reducing sectoral income inequality (and vice versa) as shown in Helpman, Itskhoki and Redding (2008a), so that conclusions based on wage inequality can be misleading if the ultimate concern is income inequality.

Open Versus Closed Economy

The effect of the opening of trade on sectoral income inequality also follows from its effects on sectoral wage inequality and unemployment above.

Proposition 2.7. Let $\omega$ be invariant to trade. Then sectoral income inequality, as measured by the Theil Index or the Gini Coefficient, is higher in the open economy than in the closed economy.

Proof: The proposition follows immediately from Propositions 2.3 and 2.5 and the expressions for the Theil Index (2.30) and Gini Coefficient (2.31).

The opening of trade raises sectoral income inequality through two channels. First, the partitioning of firms by productivity into exporters and non-exporters, and the discrete increase in wages at exporters relative to non-exporters, raises sectoral wage inequality. Second, the reallocation of employment towards more productive firms that screen more intensively reduces the hiring rate and increases sectoral unemployment.

While sectoral income inequality in the open economy is higher than in the closed economy, once the economy is open to trade, sectoral income inequality (like wage inequality
and unemployment) has a non-monotonic relationship with trade openness. Therefore a further increase in trade openness can either increase or decrease sectoral income inequality depending on the initial level of trade openness.

2.6 General Equilibrium

Up to this point, we have analyzed sectoral equilibrium in the closed and open economy, taking as given expected worker income ($\omega$), prices in other sectors ($P_{-i}$) and aggregate income ($\Omega$). In this section, we examine the determination of these variables in general equilibrium and the relationship between them.

We begin by assuming that workers are risk neutral and consider two alternative ways of closing the model in general equilibrium. First, we introduce an outside good, which is homogeneous and produced without search frictions. This approach is particularly tractable, as with risk neutrality expected income in the differentiated sector is pinned down by the wage in the outside sector when both goods are produced. Therefore expected worker income is invariant to the opening of trade in equilibria where both goods are produced.\(^{26}\) Second, we consider a single differentiated sector and solve for endogenous expected worker income. While endogenizing expected worker income complicates the determination of general equilibrium, all of our results for sectoral wage inequality are unchanged, and we obtain a new general equilibrium effect for unemployment, since labor market tightness responds endogenously to the opening of trade.

Closing the model in general equilibrium also enables us to examine the effect of the opening of trade on workers’ ex ante expected and ex post welfare. In the first two sub-sections, we characterize these effects under risk neutrality. In a final sub-section, we introduce risk aversion, which enables us to address the issue of globalization and income risk. We show that risk aversion introduces a new general equilibrium effect, which works against the expected welfare gains from the increase in average productivity induced by

\(^{26}\) While we assume no search frictions in the outside sector, Helpman and Itskhoki (2009) show in a model without worker heterogeneity or screening that introducing search frictions in the outside sector generates an expected income $\omega_0$ that is independent of features of the differentiated sector. Augmenting the model here to incorporate search frictions in the outside sector would generate a similar result.
the opening of trade.

Individual workers in the differentiated sector experience idiosyncratic income risk as a result of the positive probability of unemployment and wage dispersion. In each of the alternative ways of closing the model, we assume that preferences are defined over an aggregate consumption index \((C)\) and exhibit Constant Relative Risk Aversion (CRRA):

\[
U = \frac{\mathbb{E} C^{1-\eta}}{1-\eta}, \quad 0 \leq \eta < 1,
\]

where \(\mathbb{E}\) is the expectations operator. Expected indirect utility is therefore:

\[
V = \frac{1}{1-\eta} \mathbb{E} \left( \frac{w}{p} \right)^{1-\eta}.
\] (2.32)

2.6.1 Outside Sector and Risk Neutrality

We begin closing the model using an outside sector under the assumption of risk neutrality \((\eta = 0)\). The aggregate consumption index \((C)\) is defined over consumption of a homogeneous outside good \((q_0)\) and a real consumption index of differentiated varieties \((Q)\):

\[
C = \left[ \vartheta^{1-\zeta} q_0^\zeta + (1 - \vartheta)^{1-\zeta} Q^\zeta \right]^{1/\zeta}, \quad 0 < \zeta < \beta,
\]

where \(Q\) is modelled as in Section 2.2 above, and \(\vartheta\) determines the relative weight of the homogeneous and differentiated sectors in consumer preferences.\(^{27}\) While for simplicity we consider a single differentiated sector, the analysis generalizes in a straightforward way to the case of multiple differentiated sectors.

In the homogeneous sector, the product market is perfectly competitive and there are no labor market frictions. In this sector, one unit of labor is required to produce one unit of output and there are no trade costs. Therefore, as we choose the homogeneous good as the numeraire \((p_0 = 1)\), the wage in this sector is equal to one in both countries.

To determine expected worker income in the differentiated sector, we use an indif-

\(^{27}\) While in the analysis here we assume that workers have CRRA-CES preferences and experience income risk, Helpman, Itskhoki and Redding (2008a,b) consider an alternative specification with quasi-linear preferences and income insurance within families.
ference condition between sectors analogous to that in Harris and Todaro (1970), which equates the expected utility of entering each sector in an equilibrium where both goods are produced. Under risk neutrality, this Harris-Todaro condition implies that expected worker income in the differentiated sector equals the certain wage of one in the homogeneous sector:

\[ xb = \omega = 1, \]  

(2.33)

where incomplete specialization can be ensured by appropriate choice of labor endowments \((\bar{L}, \bar{L}^\star)\) and relative preferences for the homogeneous and differentiated goods \((\theta)\). Positive unemployment occurs in the differentiated sector for a sufficiently large search friction \(\alpha_0\), such that \(\alpha_0 > \omega = 1\) and hence \(0 < x < 1\) in (2.14). Given an expected income of one in each sector, each country’s aggregate income is equal to its labor endowment:

\[ \Omega = \bar{L}. \]  

(2.34)

To determine the price index in the differentiated sector \((P)\), we use the functional relationship (2.19) introduced above, which with CES preferences between the homogeneous and differentiated sector takes the following form:

\[
A^{\frac{1}{1-\beta}} = \frac{(1 - \theta) P^{\varphi - \frac{\beta - \gamma}{(1 - \beta)(1 - \gamma)}} \Omega}{\theta + (1 - \theta) P^{\frac{\gamma}{1-\gamma}}},
\]  

(2.35)

where the right-hand side is monotonically increasing in \(P\). Therefore this relationship uniquely pins down \(P\) given the demand shifter \((A)\) and aggregate income \((\Omega)\).

To determine general equilibrium, we use the conditions for sectoral equilibrium in Section 2.2 above (where (2.35) replaces (2.19)), and combine them with the Harris-Todaro condition (2.33) and aggregate income (2.34). Together these relationships determine the equilibrium vector \((x, b, \theta_d, \theta_x, A, Q, P, M, L, \omega, \Omega)\). Having determined this equilibrium vector, the price index \(P\)—dual to the aggregate consumption index \(C\)—and consumption of the homogeneous good \(q_0\) follow from CES demand. Finally, equilibrium employment in the homogeneous sector follows from labor market clearing \((L_0 = \bar{L} - L\), where incomplete
specialization requires \( L < \bar{L} \).

Having characterized general equilibrium, we are now in a position to examine the impact of the opening of trade on \textit{ex ante} expected and \textit{ex post} welfare. To characterize the impact on \textit{ex ante} welfare, note that differentiated sector workers receive the same expected indirect utility as workers in the homogenous sector when both goods are produced:

\[
V = \frac{1}{P} \quad \text{for } \eta = 0. \tag{2.36}
\]

Therefore the change in expected welfare as a result of the opening of trade depends solely on the change in the aggregate price index \((P)\), which with our choice of numeraire depends solely on the change in the price index for the differentiated sector \((P)\). These comparative statics are straightforward to determine. From the free entry condition (2.18), the opening of trade raises the zero-profit productivity cutoff \((\theta_d)\). Using the Harris-Todaro condition (2.33) and labor market tightness (2.14), search costs \((b)\) remain constant as long as both goods are produced, because expected worker income equals one. Therefore, from the zero-profit cutoff condition (2.15), the rise in \(\theta_d\) implies a lower value of the demand shifter \((A)\). Given constant aggregate income \((\Omega)\) and a lower value of \(A\), CES demand (2.35) implies that the opening of trade reduces the price index for the differentiated sector \((P)\), which implies higher expected welfare in the open than in the closed economy.

While \textit{ex ante} welfare is the same for all workers, the opening of trade has distributional consequences for \textit{ex post} welfare. In the homogeneous sector, there is no uncertainty, and \textit{ex post} and \textit{ex ante} welfare are the same. In contrast, in the differentiated sector, the opening of trade raises the zero-profit productivity cutoff \((\theta_d)\) and induces selection into export markets \((\theta_x > \theta_d)\), which from the solutions for firm-specific variables (2.23) implies higher wages in exporting firms and lower wages in domestic firms (in terms of the numeraire). Additionally, there is a higher unemployment rate in the differentiated sector in the open economy than in the closed economy, since expected worker income \((\omega)\) is invariant to the opening of trade as long as both goods are produced. To the extent that there are workers who are unemployed in the open economy, but would be employed in the closed economy, these workers experience lower income in the open economy than
in the closed economy. Using these results for the incomes of employees of exporters, employees of domestic firms and the unemployed, as well as the lower aggregate price index in the open economy established above, we can compare welfare in the open and closed economies as follows:

**Proposition 2.8.** Let \( \eta = 0 \). Then (i) Every worker’s ex ante welfare is higher in the open economy than in the closed economy; (ii) A homogeneous sector worker’s ex post welfare is higher in the open economy than in the closed economy; (iii) In the differentiated sector: (a) The ex post welfare of a worker employed by an exporting firm with productivity \( \theta \) is higher in the open economy than in the closed economy; (b) The ex post welfare of workers who are unemployed in the open economy, but who would be employed in the closed economy, is lower than in the closed economy; (c) The ex post welfare of a worker employed by a domestic non-exporting firm with productivity \( \theta \) can be either higher or lower in the open economy than in the closed economy.

**Proof:** The proposition follows from the indirect utility function, the free entry condition (2.18), the zero-profit cutoff condition (2.15), and CES demand (2.35), as shown in Appendix.

As there is no unemployment or income inequality in the outside sector, aggregate unemployment and income inequality depend on the differentiated sector’s share of the labor force as well as unemployment and income inequality within this sector. As a result, the opening of trade has additional compositional effects on aggregate unemployment (as discussed in Chapter I) and aggregate income inequality (as discussed in Helpman, Itskhoki and Redding, 2008a,b).

### 2.6.2 Single Differentiated Sector and Risk Neutrality

We next consider a single differentiated sector under the assumption of risk neutrality \( (\eta = 0) \). The aggregate consumption index \( (C) \) is defined over consumption of a continuum
of horizontally differentiated varieties:

\[ C = Q, \]

where \( Q \) again takes the same form as in Section 2.2 above. While for simplicity we again assume a single differentiated sector, the analysis generalizes in a straightforward way to the case of multiple differentiated sectors.

General equilibrium can be determined in the same way as sectoral equilibrium in Section 2.2, while also solving for expected worker income (\( \omega \)) and aggregate income (\( \Omega \)). We choose the dual price index (\( P \)) in one country as the numeraire, and assume for simplicity throughout this sub-section that countries are symmetric, which implies \( P = P^* = 1 \). Having determined \( P \), the differentiated sector’s real consumption index (\( Q \)) follows immediately from the demand shifter (\( A \)) in (2.20):

\[ Q = A^{1/(1-\beta)}. \]

To determine expected worker income (\( \omega \)), we combine the zero-profit cutoff condition (2.15), the search technology (2.12) and expected worker income (2.13), which together yield the following upward-sloping relationship between \( Q \) and \( \omega \):

\[ Q = \left( \frac{f_d}{\kappa_r} \right) \frac{1}{\Gamma} \frac{\beta(1-\gamma_k)}{(1-\beta)^{\gamma_k}} \frac{\beta\gamma}{\alpha_0 (1-\beta)(1+\alpha_1)} \theta_d \frac{1}{\Gamma} \frac{\beta\gamma}{\alpha_1} \frac{\omega}{1-\beta} \frac{1}{1+\alpha_1}. \tag{2.37} \]

A second upward-sloping relationship between \( Q \) and \( \omega \) is provided by equilibrium labor payments (2.22), which with country symmetry can be written as:

\[ \omega = \frac{1}{L} \frac{\beta\gamma}{1+\beta\gamma} Q, \tag{2.38} \]

where we have used labor market clearing: \( L = \bar{L} \).

Having determined the zero-profit productivity cutoff (\( \theta_d \)) from the first bloc of equations (2.15), (2.16) and (2.18), the two equations (2.37) and (2.38) can be solved in closed form for \( Q \) and \( \omega \). Having solved for \( \omega \), aggregate income is given by \( \Omega = \omega \bar{L} \), and all remaining endogenous variables of the model can be solved for in closed form, as shown.
in Appendix.\footnote{As discussed in Appendix, the stability of the equilibrium requires $\frac{\beta \gamma}{1 - \beta} \frac{\sigma_1}{1 + \alpha_1} > 1$, which is satisfied for sufficiently convex search costs (sufficiently high $\alpha_1$) and sufficiently high elasticities of substitution between varieties ($\beta$ sufficiently close to but less than one).}

Under the assumption of risk neutrality, and noting $P = P^* = 1$, \textit{ex ante} expected welfare equals expected income ($\omega$), which is now endogenous and responds to the opening of trade. We are therefore in a position to determine the comparative statics of opening closed economies to trade.

**Proposition 2.9.** Let $\eta = 0$. Then the opening of trade: (i) Increases expected worker income ($\omega$) and hence expected welfare; (ii) Increases labor market tightness ($x$) and search costs ($b$).

\textit{Proof:} See Appendix for the formal derivation of these results.

The predictions of the model without the outside sector are similar to those of the model with the outside sector. The opening of trade increases \textit{ex ante} expected welfare and has distributional consequences for \textit{ex post} welfare depending on whether workers are employed or unemployed and depending on whether they are employed by exporters or domestic firms. One new general equilibrium effect is that the increase in average productivity in the differentiated sector following the opening of trade increases expected worker income ($\omega$), which in turn increases the tightness of the labor market ($x$), and hence raises equilibrium search costs ($b$).

The model’s predictions for sectoral wage inequality do not depend on expected worker income ($\omega$) and search costs ($b$), and are therefore the same with or without the outside sector. In contrast, the endogenous determination of expected worker income ($\omega$) opens up a new channel for trade to affect sectoral unemployment (2.28). As shown in Section 2.2, the opening of closed economies to trade raises sectoral unemployment for a given value of expected worker income ($\omega$), because it reduces the hiring rate ($\sigma$). In the model without the outside sector, the opening of trade now also increases expected worker income ($\omega$). This “income effect” reduces sectoral unemployment through increased labor
market tightness \((x)\). Depending on parameter values, this increase in labor market tightness can dominate the reduction in the hiring rate, so that sectoral unemployment can fall rather than rise following the opening of trade. Finally, in the model with a single differentiated sector, there are no changes in sectoral composition, so that our results for sectoral inequality and unemployment extend immediately to aggregate inequality and unemployment.

### 2.6.3 Outside Sector and Risk Aversion

To introduce risk aversion \((0 < \eta < 1)\), we return to the model with the outside sector, where we can explore the implications of uncertainty for the allocation of resources between the riskless homogeneous sector and risky differentiated sector.\(^{29}\) Introducing risk aversion changes the equilibrium share of revenue received by workers in the bargaining game, but does not affect any of the comparative statics of sectoral equilibrium considered above.\(^{30}\) General equilibrium can be determined in the same way as in Section 2.6.1, but with appropriate modifications for risk aversion to the Harris-Todaro condition (2.33) and aggregate income (2.34).

Under the assumption of risk aversion, the Harris-Todaro condition equates expected utility in the differentiated sector to the certain wage of one in the homogeneous sector, and therefore takes the following form:

\[
x\sigma \mathbb{E} w^{1-\eta} = x\sigma \int_{w_d}^{\infty} w^{1-\eta} dG_w (w) = 1,
\]

where expected utility in the differentiated sector equals the probability of being matched \((x)\) times the probability of being hired conditional on being matched \((\sigma)\) times expected

---

\(^{29}\) Introducing risk aversion in the model with a single differentiated sector has little effect, because there is no riskless activity to or from which resources can move.

\(^{30}\) In Appendix, we derive the solution to the Stole and Zwiebel (1996a,b) bargaining game when workers are risk averse. We show that with CRRA-CES preferences the solution takes a similar form as when there are differences in bargaining weight between the firm and its workers.
utility conditional on being hired. This condition can be expressed as:

$$
\Lambda (\rho, \Upsilon_x) \frac{b^{1-\eta}x}{\phi_w (1 + \mu \eta)} = 1,
$$

(2.40)

where \( \phi_w \) is a derived parameter defined in Appendix and:

\[
\Lambda (\rho, \Upsilon_x) \equiv \frac{1 + \rho z - \beta (1 - \eta k/\delta) \Gamma \left[ \Upsilon_x (1 - \beta)(1 - \eta k/\delta) \Gamma \right]^{\frac{1}{1+(1-\eta)\alpha_1}} - 1}{1 + \rho z - \beta \Gamma \left[ \Upsilon_x (1 - \beta) \Gamma \right]^{\frac{1}{1+(1-\eta)\alpha_1}} - 1}.
\]

Evidently, we have \( \Lambda (0, \Upsilon_x) = 1 \) and \( 0 < \Lambda (\rho, \Upsilon_x) < 1 \) for \( 0 < \rho \leq 1 \), since \( \Upsilon_x > 1 \), \( \delta > k \) and \( 0 < \eta < 1 \).

There is income risk in the differentiated sector, because of unemployment and wage inequality, which imply that risk averse workers require a risk premium to enter this sector. To determine expected worker income in the differentiated sector \( (\omega = xb) \), we combine (2.40) with (2.12) to obtain:

$$
\omega = (\alpha_0)^{\frac{\eta}{1+(1-\eta)\alpha_1}} [(1 + \mu \eta) \phi_w^{\frac{\alpha_1+1}{1+(1-\eta)\alpha_1}} \Lambda (\rho, \Upsilon_x)^{-\frac{1+\alpha_1}{1+(1-\eta)\alpha_1}}],
$$

(2.41)

where a sufficiently large search friction \( (\alpha_0) \) ensures positive unemployment \((0 < x < 1 \) in (2.14)) and a positive risk premium in the differentiated sector \((\omega - 1 > 0)\). Aggregate income is the sum of worker income in the homogeneous sector and the differentiated sector:

$$
\Omega = \bar{L} + (\omega - 1) L,
$$

(2.42)

As shown in the analysis of sectoral equilibrium in Sections 2.2–2.4 above, the opening of trade increases sectoral wage inequality and unemployment for a given value of \( \omega \). This increase in wage inequality and unemployment enhances income risk in the differentiated sector, which implies that risk averse workers require a higher risk premium to enter the differentiated sector. This “risk effect” raises expected worker income \( (\omega) \) following the

---

31 The terms in the price index \( (P) \) and \( 1/(1-\eta) \) cancel from the Harris-Todaro condition equating expected indirect utility (2.32) in the two sectors.
opening of trade (since $\Lambda(0, \Upsilon_x) = 1$ and $0 < \Lambda(\rho, \Upsilon_x) < 1$ for $0 < \rho \leq 1$ in (2.41)), which increases labor market tightness ($x$) and search costs ($b$). We are now in a position to state the following comparative statics for the opening of closed economies to trade.

**Proposition 2.10.** Let $0 < \eta < 1$. Then the opening of trade: (i) Increases expected worker income ($\omega$); (ii) Increases labor market tightness ($x$) and search costs ($b$).

**Proof:** The proposition follows from the free entry condition (2.18), expected worker income (2.41) and labor market tightness (2.14), as shown in Appendix.

Under risk aversion, the opening of trade has two offsetting effects on expected welfare. On the one hand, it raises the zero-profit productivity cutoff, which increases average productivity, expands the size of the differentiated sector and reduces the differentiated sector price index. On the other hand, it increases the risk premium in the differentiated sector, which increases search costs, contracts the size of the differentiated sector and increases the differentiated sector price index. In addition, there are distributional consequences of the opening of trade for ex post welfare depending on a worker’s sector and firm of employment, as discussed in the case of risk neutrality above.

The model’s predictions for sectoral wage inequality are unchanged by the introduction of risk aversion, because they do not depend on expected worker income ($\omega$). In contrast, as in the model without the outside sector, the increase in expected worker income as a result of the opening of trade modifies the model’s predictions for sectoral unemployment.

While the reduction in the hiring rate ($\sigma$) established in Section 2.2 above increases unemployment, the increase in labor market tightness ($x$) induced by higher expected worker income reduces unemployment. As in the risk neutral case discussed above, aggregate inequality and unemployment depend on their sectoral values and sectoral composition.

### 2.7 Conclusion

The relationship between international trade and earnings inequality is one of the most hotly-debated issues in economics. Traditionally, research has approached this topic from the perspective of neoclassical trade theory with its emphasis on specialization across
industries and changes in the relative rewards of skilled and unskilled labor. In this paper we propose a new framework that features variation in employment, wages and workforce composition across firms within industries, and equilibrium unemployment. These features are explained by firm heterogeneity, worker heterogeneity, search frictions and screening of workers by firms.

We characterize the distribution of wages across workers and the determinants of unemployment. In the closed economy, there is a single sufficient statistic for wage inequality, which is increasing in the dispersion of firm productivity, and can be either increasing or decreasing in the dispersion of worker ability. Opening closed economies to trade raises wage inequality, but once economies are open to trade, further increases in trade openness can either raise or reduce wage inequality. The unemployment rate depends on the fraction of workers that are matched (the tightness of the labor market) and the fraction of these matched workers that are hired (the hiring rate). While opening closed economies to trade reduces the hiring rate, it leaves labor market tightness unchanged except for general equilibrium effects through expected worker income. We provide conditions under which expected income remains constant in general equilibrium, in which case the opening of closed economies to trade raises income inequality through both greater wage inequality and higher unemployment.

Since trade affects wage inequality and unemployment, it influences both ex ante expected welfare and ex post welfare once firms and workers are matched. When workers are risk-neutral, welfare gains from trade are ensured. When workers are risk averse, the reduction in the consumer price index as a result of the productivity gains induced by the opening of trade is counterbalanced by greater income risk in the differentiated sector. As compensation for this greater income risk, workers receive higher expected income in the open economy than in the closed economy, which increases labor market tightness. As a result, the increase in unemployment from a lower hiring rate is offset by a reduction in unemployment from a tighter labor market.

Our model provides a framework which can be used to analyze the complex interplay between wage inequality, unemployment and income risk, and their relation to interna-
tional trade. In emphasizing wage inequality across firms within industries, it is compatible with trade-related changes in income inequality, even in the absence of large observed reallocations of resources across sectors.
3. OPTIMAL REDISTRIBUTION IN AN OPEN ECONOMY

3.1 Introduction

The idea that trade leads to distributional conflict is one of the most widely accepted views among economists. The 1980s and 90s saw rapid globalization and a concomitant increase in wage and income inequality in the U.S., the U.K. and many other developed countries.\(^1\) The contribution of trade, as opposed to skill-biased technical change, to rising inequality is intensely debated. Feenstra and Hanson’s (1999) estimates suggest that outsourcing alone could account for as much as 40% of the increase in the U.S. skill premium in the 1980s. Other studies, summarized in Krugman (2008), arrive at more conservative estimates suggesting that trade accounted for about 15-20% of the increase in income inequality.\(^2\) Even more striking is the evidence for developing countries. Goldberg and Pavcnik (2007) summarize a body of literature studying the consequences of trade liberalization across a number of developing countries after 1970s, all finding a significant increase in inequality.

The above evidence raises the question as to how optimal national redistribution policies should respond to increasing world trade. Optimal redistribution policy has to balance equity and efficiency considerations. The conventional view is that greater inequality in a closed economy should be offset by more progressive taxation and higher marginal tax rates. This view is often extended to trade-induced inequality: when trade causes rising inequality, the optimal policy response should be to increase redistribution, rather than to

\(^1\) See, for example, Burtless and Jencks (2003) and Machin and van Reenen (2007).

\(^2\) A recent study by Bloom, Draca and van Reenen (2008) finds that trade with China is an important force behind differential technology adoption across British firms which in turn leads to growing wage inequality. Another recent study by Broda and Romalis (2008) suggests, however, that the same trade with China has likely reduced inequality on the consumption side in the US by reducing the price of the consumption bundle for the poor relative to that for the rich.
limit trade. This conventional intuition does not hold, however, when the original cause of rising inequality also intensifies the efficiency margin in the economy. In this paper we show that in a class of models, trade-induced inequality is intricately linked with a more sensitive efficiency response of economy to taxation, both being caused by an extensive margin of trade. Increasing marginal tax rates, therefore, is not necessarily the optimal response to trade-induced inequality. In fact, under some circumstances, it is optimal to reduce marginal tax rates and taxation progressivity in response to trade liberalization, which further worsens the inequality outcome. That is, countries might need to accept increasing inequality in order to reap the most welfare gains from trade.

To address the issue of optimal redistribution in an open economy, one needs a particular modeling framework that allows for analyzing the distributional consequences of trade. Traditionally this has been the Stolper-Samuelson Theorem of the Heckscher-Ohlin (HO) model. Recently however the empirical limitations of this framework have become apparent. As already mentioned, trade liberalizations led to a sharp increase in inequality in unskilled-labor abundant developing countries, a phenomenon at odds with the prediction of the HO model. In addition, the contribution of the residual component of wage inequality within groups of workers with similar observable characteristics appears to be at least as important as the growing skill premium across groups, as emphasized by the HO model. Finally, contrary to the main mechanism of adjustment in the HO model, the reallocation within sectors appears to be more important than across sectors for both adjustment to trade and inequality dynamics.

---

3 For example, Irwin (2008, pp. 142–3) summarizes this intuition: “I suspect the policy advice would be the same regardless of whether trade was found to have been responsible for 4 percent or 40 percent of a given amount of wage inequality. That response would probably be as follows: inequality may be undesirable, but it should be addressed not by closing markets through greater protectionism, but by more progressive income taxation, a stronger social safety net, and more assistance for displaced workers.”

4 A related observation is that the movements in relative prices of skilled to unskilled goods, which are at the core of the Stolper-Samuelson mechanism, tended to be small (e.g., see Lawrence and Slaughter, 1993).

5 For example, see Autor, Katz and Kearney (2008) for the evidence for US and Attanasio, Goldberg and Pavcnik (2004) for the evidence for a developing country (Colombia).

6 For example, Faggio, Salvanes and van Reenen (2007) show that most of the increase in wage inequality in UK happened within industries, while Levinsohn (1999) shows the relative importance of within-industry reallocation in response to trade liberalization in Chile.
In this paper we propose a simple modeling framework for thinking about the distributional effects of trade and analyzing the optimal redistribution policies in an open economy. Although this framework is highly stylized, its predictions for the relationship between trade and inequality are consistent with those arising from a more detailed model of product and labor markets, developed in Chapter II of this dissertation, where we propose an alternative to HO framework consistent with a number of empirical regularities.\(^7\)

The key ingredients of the current model are unobservable agent heterogeneity and fixed costs of exporting, as in Melitz (2003) and Yeaple (2005), which allow only the most productive agents to participate in international trade.\(^8\) Consequently, trade disproportionately benefits the most productive agents within sectors and occupations, leading to greater income inequality in a trading equilibrium than in autarky. In addition, selection into the exporting activity generates an extensive margin of trade, which is sensitive to national redistribution policies and contributes to the overall efficiency margin of taxation.

The unobservable agent heterogeneity points towards the Mirrlees (1971) framework for analyzing the optimal redistribution policy. In our analysis we deviate as little as possible from the baseline Mirrlees closed economy. The departure that we consider is the introduction of imperfect substitutability between individual varieties of the differentiated final good.\(^9\) Imperfect substitutability results in love-of-variety which is the source of trade in our model as in Krugman (1980) and Helpman and Krugman (1985).\(^10\) Agents in our economy are worker-entrepreneurs each producing a distinct variety of the final good. Unobservable productivity heterogeneity across agents generates income inequal-

---

\(^7\) In Chapter II we constructed a model of firm heterogeneity, unobservable worker heterogeneity, random search and endogenous screening which allows to account for size and exporter wage premium, residual component of wage inequality and patterns of inter- and intra-sectoral reallocation in response to trade liberalization.

\(^8\) Empirically only a small fraction of firms export even in the most tradable sectors (Bernard and Jensen, 1999) and fixed costs of trade appear to be quantitatively very important (Das, Roberts and Tybout, 2007; Bernard and Jensen, 2004).

\(^9\) Imperfect substitutability between different types of labor in the Mirrlees model were studied by Feldstein (1973) and Stiglitz (1982) in a two-class economy.

\(^10\) Love-of-variety models are the principal explanation for intra-industry trade which currently accounts for the majority of world trade (e.g., see Helpman, 1999). Moreover, Broda and Weinstein (2006) show the empirical importance of the love-of-variety effect and estimate large welfare gains from increased variety through international trade.
ity. Inequality averse society designs incentive-compatible redistribution policies in order to partly offset equilibrium inequality by optimally balancing the equity-efficiency considerations. For the purposes of tractability, we restrict the analysis to a limited set of policy instruments, which however allows for differential marginal taxes on exporters and non-exporters, as well as a subsidy for entry into the foreign market.

The main result of our analysis is that selection into foreign trade activities leads to both greater inequality and greater efficiency losses from redistribution in an open economy relative to autarky. In more technical language, both inequality and efficiency margins intensify in an open economy causing the equity-efficiency trade-off to become more tight. As a result, the optimal redistribution policy response to trade liberalization is in general ambiguous. On the one hand, selection into exporting activity leads to an increase in the dispersion of relative revenues across the groups of exporters and non-exporters. This causes greater income inequality and calls for more redistribution. On the other hand, selection into exporting activity also results in an active extensive margin of trade, which is sensitive to the stance of the redistribution policy.\(^\text{11}\) Therefore, the efficiency losses from redistribution also increase in an open economy as they now combine both intensive and extensive margins. On net, the optimal policy response to increasing trade can be both to raise or to reduce marginal taxes and the progressivity of the tax schedule.

Societies might have to accept growing income inequality as a necessary outcome in order to capture the welfare gains from trade. If redistribution policies are determined by highly inequality-averse agents, society will be bound to forgo a large fraction of the welfare gains from trade. It is important to note here that this argument does not rely on whether there are losers from trade or whether every agent gains from trade. It only requires that gains from trade are not equally distributed, with the most gains concentrated among highly productive agents who can take the advantage of opening up by exporting their products. In fact, in the present model, all agents gain from trade in absolute terms.

\(^{11}\) Higher marginal taxes do not necessarily affect the extensive margin directly, however, they have an indirect effect through the response of the optimal scale of production. Fixed costs activities require a certain scale in order to be justified. Since higher marginal taxes reduce the optimal scale for all firms they also negatively affect the extensive margin.
provided that there is no redistribution policy response to increasing trade.\textsuperscript{12}

In Section 3.2 we start our analysis by considering the case of the closed economy. We show that greater dispersion of revenues coming from the underlying ability distribution robustly leads to higher marginal tax rates and more progressive average taxes. In Section 3.3 we study the baseline model of an open economy without fixed costs of trade in which agents of all ability levels participate in trade and export. We show, quite surprisingly, that in this environment neither inequality nor efficiency margins respond to trade.\textsuperscript{13} The reason is that trade increases revenues proportionally for all agents, and hence, leads to no distributional conflict. In addition, this model does not have an active extensive margin of trade, and hence, the efficiency margin is not affected. It follows that trade in a model without selection into the export market induces no inequality response and leaves the optimal redistribution policy unaltered.

In Section 3.4 we study the case of the open economy with fixed costs of exporting. We start by showing that inequality of both revenues and utilities is greater in an open economy than in autarky even though trade is beneficial for all agents. Fixed costs of trade allow only the most productive agents to profitably engage in exporting. These are the agents who benefit the most from trade liberalization, and this exacerbates the distribution conflict. We then show that the presence of an extensive margin of trade also magnifies the efficiency consequences of taxation. As a result, the optimal progressivity of taxation can either increase or decrease.

We start our analysis with a single linear tax instrument and then introduce additional tax instruments such as export market entry subsidy and differential tax brackets for exporters and non-exporters. The key insight here is that a limited set of tax instruments has to balance the goals of redistributing income and inducing optimal entry. When some

\textsuperscript{12} When domestic redistribution policy responds to trade some agents may lose. Therefore, in this model there are no losers from trade per se, but trade may induce a domestic policy response which harms some of the agents in absolute terms relative to autarky.

\textsuperscript{13} This is the case with symmetric countries and cooperative policy determination (defined as the Nash bargaining solution between the two countries). We also study the case of non-cooperative policy determination. In this case terms of trade partly shield the country from distortionary domestic taxation and result in inefficiently high level of taxes in both countries. A similar effect is studied by Epifani and Gancia (2008) in a different model of taxation.
instruments are not feasible, the other instruments have to adjust in order to replicate the optimal allocation as close as possible. For example, if entry subsidy is infeasible, marginal taxes will be regressive in order to encourage the right amount of entry at the cost of providing less income redistribution. In contrast, when entry subsidy is available, the optimal taxation scheme can be strongly progressive.

Finally, Section 3.5 shows how the insights of our analysis can be extended to other areas such as technology adoption. There is substantial evidence that technology adoption is another activity which requires considerable fixed costs so that only the most productive agents can take advantage of new technologies. Therefore, the results of our analysis can be readily applied to studying the relationship between technology adoption and inequality, as well as the optimal redistribution policy response. Section 3.5 also provides our concluding remarks. The technical details of the derivations and proofs are relegated to the Appendix.

**Related literature**

The issue of optimal redistribution in an open economy has received little attention in the literature. Following Dixit and Norman (1980, 1986) the literature mainly focused on the possibility of compensating losers from trade, typically in the context of a HO model, rather than on the design of the optimal redistribution policies. A few studies in this body of literature are most closely related to this paper. Spector (2001) introduces Mirrlees incentive constraints into an otherwise standard HO model and shows that trade may lead to welfare losses by endogenously limiting the set of instruments available to the government for redistributional purposes. Ichida (2003) uses a two-sector model with unobservable agent heterogeneity to study the possibility of attaining a Pareto improvement from trade without overcompensating some of the agents. Davidson and Matusz (2006) design the lowest cost compensation policies for the losers from trade in a two-sector economy with heterogenous agents and participation decisions, but fixed labor supply. A

---

14 There exists, however, a vast literature on optimal capital and profit taxation in an open economy surveyed in Gordon and Hines (2002). Another strand of literature, started by Cameron (1978) and Rodrik (1998), studies the optimal size of the government in an open economy. Finally, there recently has emerged an active literature on optimal dynamic redistribution in closed economy macro models, as surveyed in Golosov, Tsyvinski and Werning (2006).
recent paper by Egger and Kreickemeier (2008) analyzes a model of firm heterogeneity and fair wages, and studies the possibility of ensuring both welfare gains from trade and a more equal wage distribution when only a linear profit tax is allowed.

In the closed economy literature, the closest paper is by Saez (2002), who studies optimal redistribution in a model with both intensive and extensive margins of labor supply responses. He finds that Negative Income Tax programs are no longer optimal in a model with labor force participation decisions. The presence of an extensive margin of labor supply leads to the optimality of negative marginal taxes for the agents who are most likely to drop out of labor force.

3.2 Closed Economy

In this section we lay out the setup of the baseline closed economy and derive a general formula for the optimal tax rate which, as we show later, extends to an open economy environment. In the closed economy, our main emphasis is on the effect of income dispersion on marginal tax rates.

3.2.1 Economic Environment

The economic environment departs minimally from the original Mirrlees (1971) economy in order to allow for a meaningful analysis of international trade in later sections. We consider a static one sector economy. The economy is populated by a measure $L$ of worker-entrepreneurs heterogenous in their ability $n$, which is distributed on $[n_{\text{min}}, n_{\text{max}}] \subset \mathbb{R}^+$ according to a cumulative distribution function $H(n)$. Each agent can produce his own variety of the final good according to a linear production technology:

$$y_n = n\ell_n,$$  \hfill (3.1)

where $y_n$ is output and $\ell_n$ is labor input of an agent with productivity $n$. Since all agents with productivity $n$ are symmetric, we use $n$ to index the agents. Note that the amount

\footnote{In particular, we allow both $n_{\text{min}} = 0$ and $n_{\text{max}} \to \infty$. For convenience, we write 0 and $\infty$ as the limits of integration implying that $H(n) \equiv 0$ for $n \in [0, n_{\text{min}})$ and $H(n) \equiv 1$ for $n \in (n_{\text{max}}, \infty)$.}
of product variety in the economy is determined by the measure of agents who choose to produce and service the market in equilibrium.

The final good is a Dixit and Stiglitz (1977) CES aggregator of individual varieties:

$$Q = \left[ L \int_0^{\infty} y_n^\beta dH(n) \right]^{1/\beta}, \quad 0 < \beta \leq 1,$$

(3.2)

where $1/(1 - \beta) > 1$ is the elasticity of substitution between the varieties. Most of the public finance literature following Mirrlees (1971) assumes perfect substitutability between the produced varieties ($\beta = 1$). The baseline case of this paper is the case of imperfect substitutability between varieties ($\beta < 1$), which will be the source of international trade in the following sections.

The assumption of imperfect substitutability between varieties has a number of implications. First, it leads to well-defined boundaries of the firms as opposed to the original Mirrlees (1971) model in which the boundaries of the firms are indeterminant. With perfect substitutability among varieties, the equilibrium allocations are identical for any number of firms that hire any fraction of the total labor supply in the competitive labor market. With imperfect substitutability, each agent produces his own variety and hence is an entrepreneur operating a separate firm, while the labor market is effectively non-existent. We view this feature of the model as a useful abstraction for the purposes of this paper. Second, imperfect substitutability results in monopolistic power and monopolistic competition among the producers of individual varieties. This introduces additional distortions to the equilibrium allocation that an optimal redistribution policy will have to take into account.

Consumption aggregator in (3.2) leads to the following (real) revenue function for each

16 Important exceptions to this are Feldstein (1973) and Stiglitz (1982) who study the case of imperfect substitutability between different types of labor input in a two class economy which results in an optimal income subsidy for high productivity types. This result is driven by the attempt of the optimal income taxation to manipulate equilibrium relative prices of different varieties. Similar effects will be at play in the present paper.

17 In Chapter II we developed a detailed model of product and labor market equilibrium which has similar implications for the effects of trade on inequality.
individual variety:\textsuperscript{18} 
\[ r_n = Q^{1-\beta} y_n^\beta. \] (3.3)

The revenue is increasing and concave in the agent’s output and shifts out with an increase in the economy-wide real consumption. CES preferences imply that tighter product market competition, causing higher real consumption, increases revenues for every producer in the market. The opposite prediction arises in a two-sector model (e.g., see Melitz and Ottaviano, 2008; Helpman and Itskhoki, Chapter I of this dissertation), although the implications for the relative revenues of different agents, central for the analysis of optimal redistribution, are robust across these models.

Utility of every agent in the economy is given by $U(c, \ell)$, where $c$ is consumption and $\ell$ is labor effort. To make the analysis tractable, we adopt the GHH preferences (due to Greenwood, Hercowitz and Huffman, 1988), featuring no income effects on labor supply, a constant wage elasticity of labor supply and a constant relative risk aversion:

\[ U(c, \ell) = \frac{1}{1-\rho_a} \left( c - v(\ell) \right)^{1-\rho_a}, \quad v(\ell) = \frac{1}{\gamma} \ell^\gamma, \quad \gamma = 1 + \frac{1}{\varepsilon}. \] (3.4)

Here $\rho_a$ denotes constant relative risk aversion of the agents and $\varepsilon$ is the constant labor supply elasticity.

All agents face the same tax schedule $T(r)$, which is a function of only their income (revenues), assuming that their labor effort is unobservable. As a result, the budget constraint of agent $n$ is given by:

\[ c_n = r_n - T(r_n), \quad r_n = Q^{1-\beta} (n\ell_n)^\beta, \] (3.5)

\textsuperscript{18} CES aggregator implies constant elasticity demand, $y_n = Q \cdot (p_n/P)^{-1/(1-\beta)}$, where $p_n$ is the variety’s price and $P = \left[ L \int_0^\infty p_n^{-\beta/(1-\beta)} dH(n) \right]^{-(1-\beta)/\beta}$ is the ideal price index associated with consumption aggregator (3.2). Nominal revenue is then given by $p_n y_n = PQ^{1-\beta} y_n^\beta$ and real revenue is $r_n \equiv p_n y_n / P$. The price level, $P$, plays no role in the main text analysis and can be normalized to 1 without loss of generality.

\textsuperscript{19} The key simplifying assumption here is the absence of income effects on labor supply, which is a common benchmark in the public finance literature (e.g., Diamond, 1998, and Saez, 2002). Since our focus in this paper is on the effects of trade on inequality and optimal redistribution, we choose to shut down the additional effects operating through the utility function in the baseline analysis.
where real consumption equals real after tax revenues and the production function (3.1) is substituted into the expression for real revenues (3.3). Therefore, each agent maximizes utility (3.4) subject to his budget constraint (3.5). We denote the resulting utility by $U_n$.

Finally, the government chooses the tax schedule $T(\cdot)$ to maximize the social welfare function

$$W = L \int_{0}^{\infty} G(U_n) dH(n)$$

subject to the government budget constraint

$$L \int_{0}^{\infty} T(r_n) dH(n) \geq 0$$

and behavioral responses of the agents (i.e., incentive compatibility constraints) which require that $\{c_n, \ell_n, r_n, U_n\}$ are the outcomes of agent optimization given a tax schedule $T(\cdot)$. Government budget constraint (3.7) implicitly assumes that the only purpose of taxation is redistribution. In general, $G(\cdot)$ is a strictly increasing and weakly concave function. We restrict it to the case of constant relative inequality aversion:

$$G(u) = \frac{1}{1 - \rho_g} u^{1 - \rho_g}, \quad \rho_g \geq 0,$$

where $\rho_g$ is the constant relative inequality aversion parameter of the government. The case of $\rho_g = 0$ corresponds to the utilitarian planner and the other limiting case $\rho_g \to \infty$ corresponds to the Rawlsian planner. Note that the overall measure of inequality aversion in the economy is given by $\rho = \rho_a + \rho_g$, which is what matters for the optimal redistribution policy rather than its decomposition into the individual and aggregate components ($\rho_a$ and $\rho_g$).\(^{20}\) Therefore, without loss of generality, we assume for notational convenience that $\rho_a = 0$ and $\rho = \rho_g \geq 0$.

We start our analysis with the case of linear taxation, so that $T(r) = -\Delta + tr$, where

\(^{20}\) The optimal amount of redistribution depends on the cross-sectional distribution of

$$G'(U_n) \cdot \frac{\partial U(c_n, \ell_n)}{\partial c} = U_n^{-\rho_g} \cdot U_n^{-\rho_a} = \left( c_n - v(\ell_n) \right)^{-(\rho_a + \rho_g)}.$$ 

Since there is no uncertainty and the model is static, the value of $\rho_a$ also does not affect the optimal individual allocations $(c_n, \ell_n)$.\(^{20}\)
is the marginal tax rate common for all agents and \( \Delta \) is the uniform transfer. The government budget constraint in this case becomes \( \Delta = tR \), where

\[
R \equiv \int_{0}^{\infty} r_n dH(n)
\]

is the average revenue which, from (3.2) and (3.3), can be seen to equal the per capita consumption \( (R = Q/L) \). In the open economy case we allow for the introduction of additional tax instruments, including differential marginal tax rates for exporters and non-exporters, as well as an entry subsidy into the exporting activity.\(^{21}\)

### 3.2.2 Optimal Redistribution in Closed Economy

In this section we derive a general formula for the optimal linear tax which generalizes to the case of open economy studied in later sections. We lay out in detail only the steps of the derivation that prove to be useful in developing intuition for the results that follow, relegating the formal proofs to the Appendix. We then provide the main result of this section: the level of marginal tax increases in the equilibrium dispersion of relative revenues, which is pinned down by the exogenous distribution of abilities in the closed economy.

Using the government budget constraint (3.7) and substituting the agent’s budget constraint (3.5) into the utility function (3.4), agent \( n \)’s optimization yields:

\[
U_n = \max_y \left\{ \Delta + (1 - t)Q^{1-\beta}y^{\beta} - \frac{1}{\gamma} \left( \frac{y}{n} \right)^{\gamma} \right\} = tR + (1 - \beta/\gamma)(1 - t)r_n,
\]

where \( r_n = Q^{1-\beta}y_n^{\beta} \) and \( y_n \) is the optimal output of agent \( n \) that satisfies the first order condition

\[
\beta(1 - t)Q^{1-\beta}y_n^{\beta-1} = v' \left( \frac{y_n}{n} \right) \frac{1}{n} = \frac{y_n^{\gamma-1}}{n^{\gamma}}.
\]

\(^{21}\) A numerical investigation of an unconstrained Mirrlees (1971) optimal policy is intended in the next versions of the paper.
This optimality condition also implies $v(y_n/n) = \beta/\gamma(1-t)r_n$, which we have used in the second line of (3.8).

Tax rate $t$ affects agent $n$’s utility both directly and indirectly. There is a positive direct effect through the amount of transfer proportional to the average revenue in the economy $R$ and a negative direct effect from taxing away a fraction of individual revenue $r_n$. The indirect effect of a tax rate operates through the equilibrium impact of $t$ on aggregate consumption and average revenue ($Q$ and $R$): An increase in the marginal tax rate reduces the per capita taxation base ($R$) and, in addition, reduces revenue of every agent by reducing $Q$. Therefore, the overall effect of taxes on utility is:

$$dU_n = \left[ \frac{\partial U_n}{\partial t} + \frac{\partial U_n}{\partial Q} \frac{dQ}{dt} \right] dt = (R - r_n)dt + \frac{dQ}{Q} \left[ tR + (1-t)(1-\beta)r_n \right]. \quad (3.9)$$

The first term is the redistributional component from agents with above-average revenues towards agents with below-average revenues. The second term is the efficiency component which is proportional to the effect of taxes on real consumption. Note that agents with lower revenue, and hence, lower utility always gain more (or lose less) from an increase in taxes.

Finally, the optimality condition for the marginal tax rate is given by

$$\int_0^\infty G'(U_n) \frac{dU_n}{dt} dH(n) = 0. \quad (3.10)$$

The government searches for a tax rate which equalizes to zero the cross-sectional weighted average utility gains from a marginal increase in the tax rate. The agents are weighted by their marginal contribution to the social welfare function, $G'(U_n)$. With positive inequality aversion ($\rho > 0$), the government puts more weight on the agents with lower utility.

---

22 The indirect effect of taxes on utility operating through their impact on the scale of individual production ($y_n$) is nil by the Envelop Theorem. The effect on utility through $Q$ is not nil since individual optimization does not internalize the CES externality or, in other words, agents exercise their monopolistic power and underproduce relative to the efficient allocation.
It proves useful to introduce the following notation. Denote by
\[
\bar{\varepsilon} \equiv \frac{d \ln Q}{d \ln (1 - t)} = -\frac{(1 - t) dQ}{Q dt}
\]
the elasticity of output (or consumption) with respect to the marginal tax rate. This elasticity quantifies the efficiency loss from taxation and is the first key object that will shape optimal taxes throughout this paper. We refer to it as the \textit{efficiency margin} of taxation. With this definition we can combine (3.9) and (3.10) to rewrite the optimality condition for the marginal tax rate as:
\[
\int_0^\infty G'(U_n) \left[ \left( 1 - \frac{r_n}{R} \right) - \bar{\varepsilon} \left( \frac{t}{1 - t} + (1 - \beta) \frac{r_n}{R} \right) \right] dH(n) = 0.
\]
Already this expression suggests that what matters for the optimal tax rate is the dispersion of relative revenues.

Rearranging the expression above, we obtain the general formula which determines the optimal linear tax rate:
\[
\frac{t}{1 - t} = \frac{1}{\bar{\varepsilon}} \cdot \alpha - (1 - \beta)(1 - \alpha),
\]
(3.11)
where
\[
\alpha = \int_0^\infty \lambda^{-1} G'(U_n) \left( 1 - \frac{r_n}{R} \right) dH(n) = -\text{cov} \left( \lambda^{-1} G'(U), \frac{r}{R} \right)
\]
(3.12)
and \( \lambda = \int_0^\infty G'(U_n) dH(n) \) is the average marginal utility (or, more precisely, average marginal contribution of agents to the social welfare), which is also equal to the shadow value of one real dollar (unit of consumption) in the hands of the government. Observe from (3.12) that \( \alpha \) is the cross-sectional covariance between two normalized variables, both with a mean of 1. It is the second key object affecting the level of optimal taxation. We refer to \( \alpha \) as the \textit{inequality margin} of taxation since it is a measure of dispersion of relative utilities resulting from the dispersion of relative revenues.

We prove in the Appendix the following:

\textbf{Lemma 3.1.} \( 0 \leq \alpha \leq 1. \) \( \alpha = 0 \) if and only if either \( \rho = 0 \) or \( r_n \equiv R \) for all agents. \( \alpha \to 1 \) if and only if \( \rho \to \infty \) and \( r_{n_{\text{min}}} = 0. \)
Formula (3.11) and Lemma 3.1 provide a general characterization of the optimal linear tax rate. Since \( \alpha \in [0, 1] \), \( t/(1-t) \) equals a convex combination of \( 1/\bar{\varepsilon} \) and \( -(1-\beta) \). When \( \alpha = 0 \), which from Lemma 3.1 occurs either when there is no inequality or when the society does not care about inequality, the optimal marginal tax rate is \( t = -(1-\beta)/\beta \). This optimal subsidy offsets the monopolistic distortion – a constant mark-up equal to \( 1/\beta \) – and implements an efficient allocation. When \( \alpha \to 1 \), which happens under the Rawlsian planner and provided that the least able agent does not produce, the optimal linear tax rate \( t \to 1/(1+\bar{\varepsilon}) \), which is the revenue maximizing tax rate at the peak of the Laffer curve.\(^{23}\)

More generally, when \( \alpha \in (0, 1) \), the optimal marginal tax rate trades off efficiency for equity. Finally, note that when \( \beta \in (0, 1) \), optimal taxation has to balance redistributional considerations with offsetting the monopolistic distortion. If agents were price takers, the second term would be absent from (3.11) and \( \beta < 1 \) would have no effect on the optimal taxes.

We summarize the discussion above in:

**Proposition 3.1.** The optimal linear income tax rate satisfies

\[
\frac{t}{1-t} = \frac{1}{\bar{\varepsilon}} \cdot \alpha - (1-\beta)(1-\alpha), \quad 0 \leq \alpha \leq 1. \quad (3.11')
\]

Therefore,

\[
\frac{1-\beta}{\beta} \leq t \leq \frac{1}{1+\bar{\varepsilon}},
\]

where \( (1-\beta)/\beta \) is the efficiency maximizing subsidy which is optimal if and only if there is either no inequality or no inequality aversion; \( 1/(1+\bar{\varepsilon}) \) is the revenue maximizing tax which is optimal if and only if there is extreme inequality aversion (Rawlsian planner) and the least productive agent does not produce.

The optimal tax rate is determined by the interaction between the efficiency margin \( \bar{\varepsilon} \) and the inequality margin \( \alpha \): greater efficiency margin reduces the optimal tax rate, while greater inequality margin increases it. Note that (3.11') characterizes the optimal tax rate

\(^{23}\) Similar results were obtained by Sheshinski (1972) and Hellwig (1986) for a general utility function, but with perfectly substitutable labor supply of different agents.
only implicitly since $\alpha$ is an endogenous object which in particular depends on the tax rate $t$. Throughout the paper we assume that $\partial \alpha / \partial t < 0$, which intuitively implies that a higher tax rate reduces the inequality margin. This assumption is also sufficient for the concavity of the welfare function in $t$ and hence for the uniqueness and optimality of the tax rate determined by the first order condition (3.11'). In the Appendix we discuss sufficient conditions for $\partial \alpha / \partial t < 0$ (see also Lemma 3.3 below).

The discussion so far has been general in the sense that all previous results extend to the open economy environment that we consider in the following sections. We now characterize the closed economy values of $\tilde{\varepsilon}$ and $\alpha$. We have the following:

**Lemma 3.2.** In the closed economy $\tilde{\varepsilon} = \varepsilon$.

That is, the elasticity of final good production with respect to the marginal tax rate (the efficiency margin) is equal to the labor supply elasticity, which is also the intensive margin of agents’ responses to an increase in the marginal tax rate. The intuition is that an increase in $t$ leads all agents to reduce their output proportionally with elasticity $\varepsilon$. This in turn leads to a proportional reduction in final output which is a homothetic aggregator of individual outputs.

We now turn to the value of $\alpha$, which belongs to the interval $[0, 1]$ by Lemma 3.1. From (3.12), $\alpha$ is a covariance between $r/R$ and $U^{-\rho}/E\{U^{-\rho}\}$. From (3.8), note that $U_n$ is a linear transformation of $r_n/R$. This suggests that $\alpha$ should increase in the dispersion of relative revenues and in the value of $\rho$ which makes marginal utility steeper. We prove in the Appendix the following lemma based on the first order approximation of $U^{-\rho}$ around $(E, U)^{-\rho}$, which becomes exact as dispersion of ability $n$ decreases towards zero:24

**Lemma 3.3.** In the closed economy, the second order approximation to $\alpha$ around $r_n = R$ for all $n$ is given by:

$$\alpha \approx \rho \cdot \frac{1}{1 + \frac{1}{1 - t} \frac{1}{1 - \beta / \gamma}} \cdot \text{var}\left(\frac{r}{R}\right).$$

(3.13)

---

24 Exact results for the comparative statics of $\alpha$ under a general skill distribution are unavailable. Therefore, we rely on this approximation in the analytical discussion and verify its implications numerically in Section 3.4 for a special distribution of skills.
Consistent with the intuition, the approximate solution implies that \( \alpha \) increases in the inequality aversion \( \rho \) and in the variance of relative revenues which also equals the square of the coefficient of variation of revenues. Importantly, from the approximation in Lemma 3.3 it follows that \( \partial \alpha / \partial t < 0 \), i.e. the inequality margin is decreasing in the level of marginal tax. As discussed above, this is sufficient to ensure that the second order condition is satisfied and the first order condition \((3.11')\) identifies the unique optimal marginal tax rate.

Finally, we characterize the distribution of relative revenues in equilibrium. From \((3.3)\) we have \( r_n/R = L(y_n/Q)^\beta \) and from the optimality condition for the agent’s problem \((3.8)\) it follows that

\[
y_n = \left[ \beta (1 - t) \right] ^{1/\gamma} Q^{1-\beta \gamma} n^{-\gamma / \gamma - \beta}.
\]

Therefore, relative revenues are pinned down exclusively by the distribution of underlying ability:

**Lemma 3.4.** *In the closed economy equilibrium,*

\[
\frac{r_n}{R} = \frac{n^{\beta \gamma}}{\int_0^\infty n^{\beta \gamma} dF(n)}.
\]

Intuitively, relative revenues do not depend on the equilibrium level of taxes which reduce revenues of all agents proportionally. This has an immediate implication that \( \text{var}(r/R) \) is determined uniquely by the distribution of underlying ability \( n \) and primitive elasticities \( \beta \) and \( \gamma \). More dispersion in the underlying ability translates into more dispersion in the relative revenues, which in turn leads to a higher \( \alpha \) and higher optimal linear tax rate.

We summarize the implications of the above analysis in the following:

**Proposition 3.2.** *Assuming the approximation in Lemma 3.3 to be accurate, the optimal linear tax rate in the closed economy (i) increases in the inequality aversion \( \rho \) and in the dispersion of relative revenues exogenously determined by the dispersion of ability; (ii) does not depend on the size of the economy \( L \); (iii) depends ambiguously on the labor supply elasticity \( \varepsilon \).*
The central result is that in the closed economy marginal tax rate increases in the dispersion of relative revenues, which from Lemma 3.4 is exogenously determined by the dispersion of underlying ability distribution.\textsuperscript{25} Greater dispersion of relative revenues increases the inequality margin, $\alpha$, which leads to a higher optimal tax rate.

Interestingly, the size of the economy does not affect the level of taxes, although the set of available varieties and the per capita consumption increase with $L$.\textsuperscript{26} The intuition is that greater $L$ increases the scale of production and revenues proportionally for all agents without affecting $\alpha$, and also does not affect the elasticity of labor supply and hence $\tilde{\varepsilon}$. This result is useful for thinking about the effects of trade in the following section. Finally, a change in the elasticity of labor supply $\varepsilon$ has an ambiguous effect on the optimal tax rate since it not only affects the efficiency margin ($\tilde{\varepsilon}$), but also changes the equilibrium dispersion of relative revenues (see Lemma 3.4 and recall that $\gamma = 1 + 1/\varepsilon$).

3.3 Open Economy I: No Fixed Costs

In this section we open our economy to trade and study how this affects the optimal taxation. The source of trade in this economy is love-of-variety, as in Helpman and Krugman (1985) style models, generated by imperfect substitutability ($\beta < 1$) between different varieties produced at home and abroad.\textsuperscript{27} Love-of-variety models are the principal explanation for intra-industry trade which currently accounts for the majority of world trade (e.g., see Helpman, 1999). Moreover, Broda and Weinstein (2006) show the empirical importance of the love-of-variety effect and estimate large welfare gains from increasing variety through international trade.

\textsuperscript{25} This conclusion can be shown to be robust to a number of extensions, including the introduction of income effects and non-constant elasticity of labor supply. Similar insights for the marginal tax on a median agent can be obtained in a model with general non-linear taxes.

\textsuperscript{26} Specifically, $d \ln(Q/L) = (1 + \varepsilon)(1 - \beta)/\beta \cdot d \ln L$.

\textsuperscript{27} Increasing returns, another common ingredient of Helpman-Krugman style models, are not needed in this framework as the amount of variety is restricted by the number of agents in the two countries. As a result, the free entry condition is missing as well, which in particular has implications for equilibrium response of the size of the individual firms to trade. We chose this modeling approach to depart minimally from the original Mirrlees economy. Free entry introduces additional interesting public finance considerations, which are arguably more relevant in the studies of optimal capital and profit taxation.
In the baseline open economy, each agent can sell part of his output at home and export the other part abroad. Exporting involves no fixed cost, but is subject to an iceberg-type variable trade cost. Specifically, for one unit of good to arrive abroad, $\tau > 1$ units have to be shipped out. Costs of trade is what distinguishes international trade from trade within national borders. Without fixed cost, every agent will participate in trade since trade allows the agents to partly escape decreasing demand and concavity of the revenue function in the domestic market. We study how opening up to trade in this economy affects income distribution and optimal income taxation. Throughout the analysis we assume no international labor mobility.

Although in principle there can be interesting interactions between optimal taxes and tariffs, we do not address the issue in this paper. Rather, we focus our analysis on the national income taxation policies in an open economy. One can rationalize this assumption in the following way: while tariffs are bound to be low by WTO agreements, national redistribution policies are set unilaterally by sovereign states.

3.3.1 Properties of Open Economy Equilibrium

For simplicity of exposition, here we describe the case of two symmetric trading economies. The case of asymmetric countries is studied in the Appendix. We characterize equilibrium allocations in the home economy with analogous characterizations for the foreign. Foreign variables are denoted with an asterisk.

An agent producing $y$ units of his variety supplies $y_d$ to the domestic market and exports the remaining $y_x = y - y_d$. Domestic sales generate revenues $Q^{1-\beta}y_d^\beta$, as explained in footnote 18. Export revenues are given by $Q^{*1-\beta}(y_x/\tau)^\beta$ since only $y_x/\tau$ units of the exported good reach the foreign market. In the symmetric equilibrium $Q = Q^*$.

The optimal allocation of output for domestic sales and exports results in the following

---

28 The discussion here assumes that the price levels in both countries are normalized to 1, $P = P^* = 1$. This is a valid assumption in a symmetric equilibrium or when $\tau \rightarrow 1$. In the Appendix we consider a general non-symmetric case in which we explicitly allow for $P \neq P^*$. 

108
where
\[ \Upsilon_x \equiv 1 + \tau^{-\beta} \frac{Q^*}{Q} > 1 \]
is the foreign market access variable which quantifies how much the access to the foreign consumers would boost producer’s revenues. \( \Upsilon_x \) decreases in the variable trade cost \( \tau \) and increases in the relative size of demand \( Q^*/Q \). Note that the revenue of every agent in the economy is magnified proportionally by the access to international trade.

As in the closed economy, agents choose output to maximize their utility:

\[ y_n = \arg\max_y \left\{ \Delta + (1 - t)r(y) - \frac{1}{\gamma} \left( \frac{y}{n} \right)^\gamma \right\} = [\beta(1 - t)]^{\frac{1}{\gamma-\beta}} \Upsilon_x^{\frac{1-\beta}{\gamma-\beta}} Q^{\frac{1-\beta}{\gamma-\beta}} n^{-\frac{\gamma}{\gamma-\beta}}, \]

where \( \Delta \) is the transfer from the government, which satisfies the government budget constraint \( \Delta = tR \). Maximized utility of agent \( n \) is again given by \( U_n = tR + (1 - \beta/\gamma)(1 - t)r_n \), where his revenues equal

\[ r_n = \Upsilon_x^{1-\beta} Q^{1-\beta} y_n^\beta. \]

Consequently, average revenues are given by

\[ R = \int_0^\infty r_n dH(n) = \Upsilon_x^{1-\beta} Q^{1-\beta} \int_0^\infty y_n^\beta dH(n). \]

Finally, balanced trade implies that aggregate revenues of domestic agents equal aggregate consumption, \( LR = Q \). This concludes the description of symmetric open economy equilibrium.

Without further characterization we can prove an important result about inequality

---

29 Optimal allocation across markets requires equalization of marginal revenues in the two destinations which implies \( y_x/y_d = \tau^{-\beta/(1-\beta)} = \Upsilon_x - 1 \) or equivalently \( y_d/y = 1/\Upsilon_x \) and \( y_x/y = (\Upsilon_x - 1)/\Upsilon_x \).
Proposition 3.3. In an open economy without fixed costs of trade, the dispersion of relative revenues is the same as in the closed economy and is determined by the dispersion of underlying exogenous ability distribution, as described in Lemma 3.4. Therefore, trade does not lead to an increase in income inequality, provided that all agents participate in the exporting activity.

As we show in the Appendix, this proposition does not require the countries to be symmetric and does not depend on how the marginal tax rates in the two countries respond to trade. The result follows from the fact that international trade in this economy increases revenues proportionally for all agents, and given constant labor supply elasticity, all agents respond by proportionally increasing their outputs. As a result, the distribution of relative revenues is unaffected and the amount of income inequality is unchanged. Changes in taxes have similar proportional effects and do not alter the distribution of relative revenues (cf. Lemma 3.4).

To summarize, the model without fixed costs predicts no inequality effects of globalization. In Section 3.4, we study how this prediction is altered in a model with fixed costs and selection into export activity.

Gains from Trade

We now discuss who gains from trade in this economy. This discussion allows for arbitrary asymmetries across countries and the formal details are found in the Appendix. Define the social welfare function by

\[ W(t, t^*) = \int_0^\infty G(U_n) \, dH(n), \]

(3.15)

where \( U_n \) is the equilibrium utility of agent \( n \) in an open economy with domestic marginal tax rate \( t \) and foreign marginal tax rate \( t^* \). We denote by \( W^a(t) \) the social welfare function in autarky. It is immediate to prove a general gains from trade result in this economy:

Proposition 3.4. Welfare is higher in an open economy than in the closed economy.
Proof: For any $t^*$, $W(t, t^*) \geq W^a(t)$ since $r_n$ increases proportionally in an open economy for all agents, given that taxes are unchanged. Open economy welfare is equal to $\max_{t'} W(t', t^*) \geq W(t, t^*)$ for any $t$. Therefore, open economy welfare is greater than closed economy welfare $\max_{t'} W^a(t')$. A more formal argument is developed in the Appendix.

This proposition implies that regardless of the taxation policy response at home and abroad, opening up to trade leads to welfare gains in both countries in the aggregate. The intuition is that trade in this model unambiguously increases the choice set of the country irrespective of the taxation policy abroad. This result stands in contrast to Spector (2001) and Epifani and Gancia (2008), who demonstrate a possibility of welfare loses in related types of economies.\(^{30}\)

Aggregate gains from trade do not guarantee, however, that every agent in the economy gains from trade. As was mentioned above, every agent gains proportionally from trade if the open economy taxes are the same as in autarky. If opening up to trade leads to a change in the tax policy, however, in principle some agents may lose from trade. Specifically, if taxes go up in the open economy, the most productive agents might end up losing from opening up to trade; by contrast, if taxes go down in the open economy, the least productive agents in the economy might end up losing.\(^{31}\) The Appendix provides formal conditions for high (low) productivity agents to lose from trade. It is important to note that in this economy agents lose not from trade per se, but from the endogenous response of the redistributional policy to trade. Therefore, a source of protectionism might not be the distributional consequences of trade per se, but rather the expectation that trade will cause unfavorable changes in the domestic policies.

\(^{30}\) In our economy every agent produces a distinct variety. In Spector (2001), every variety produced at home is also produced abroad. As a result, after opening up to trade, the government loses the ability to manipulate prices of the varieties and hence its choice set is restricted. Epifani and Gancia (2008) consider a love-of-variety model of trade without agent heterogeneity, but with a public good provided by the government.

\(^{31}\) Note the implication that if there is tax rate convergence across trading partners, the poor may lose in the country with initially high taxes and the rich may lose in the country with initially low taxes.
3.3.2 Optimal Redistribution

The immediate implication of Proposition 3.3 is that the dispersion of relative revenues does not depend on the size of variable trade cost $\tau$. Moreover, from the expression for equilibrium utility we observe that, given that the marginal tax rate remains unchanged, the distribution of relative utilities across agents is unchanged, and every agent gains from trade proportionally. This immediately implies that $\alpha$, as defined in (3.12), also remains the same as long as marginal taxes are not adjusted. In other words, opening up this economy to trade does not have distributional consequences, and only the potential change in the efficiency margin can call for tax policy adjustment. This is the focus of the remainder of this section.

In the text, we focus again on the case of symmetric countries and cooperative policy determination. Asymmetric countries and non-cooperative policies are characterized in the Appendix and briefly discussed below. We define cooperative policies as the outcome of a Nash bargaining solution between the two countries, where the non-cooperative Nash equilibrium is taken as the status quo point. Cooperative solution ensures that countries do not forgo any possible Pareto gains from policy coordination.

In a symmetric world equilibrium $t^* = t$ and $W(t,t) = W^*(t,t)$. In the Appendix we show that optimal tax rates are determined by the following first order condition:

$$\frac{dW(t,t)}{dt} = \frac{\partial[W(t,t) + W^*(t,t)]}{\partial t} = \frac{\partial W(t,t)}{\partial t} + \frac{\partial W(t,t)}{\partial t^*} = 0,$$

which takes into account the effect of the domestic tax rate on welfare both at home and abroad. We further show that this condition again implies

$$\frac{t}{1-t} = \frac{1}{\varepsilon} \cdot \alpha - (1 - \beta)(1 - \alpha),$$

$32$ Since aggregate revenue increases in an open economy, the size of the transfer $\Delta = tR$ increases as well even with constant marginal tax rate $t > 0$. This ensures that even very low productivity agents gain from trade proportionally in the utility terms.
as in the closed economy. The inequality margin, \( \alpha \), is still defined by (3.12) and

\[
\bar{\epsilon} \equiv \left. \frac{d \ln Q}{d \ln(1 - t)} \right|_{t = t^*} = \varepsilon.
\]

Therefore, neither efficiency margin (\( \bar{\epsilon} \)), nor inequality margin (\( \alpha \)) change in a symmetric open economy. Consequently, in a symmetric open economy, cooperatively-set marginal taxes are the same as in autarky. The intuition is straightforward: Opening up to trade in this economy does not induce redistributional effects as discussed in Section 3.3.1, and the output response to a coordinated change in taxes is still given by the labor supply elasticity \( \varepsilon \). This is the case as output of each variety for domestic sales and exports is reduced proportionally in response to a global increase in the income tax rate. In other words, the efficiency margin is still determined by the intensive margin of production, which equals the labor supply elasticity.

We summarize this in:

**Proposition 3.5.** Cooperatively-set taxes in a symmetric world economy without fixed costs of trade are the same as in the closed economy. Both the efficiency and the inequality margins of optimal taxation are the same in the open economy as in autarky.

Opening up to trade in this economy is similar to increasing the measure of agents, \( L \), in the closed economy, which according to Proposition 3.2 has no effect on the optimal tax rate. In the next section we show how this result contrasts with the implications of a model with fixed costs of trade and selection into exporting activity.

Consider now what happens when policies are determined non-cooperatively and countries are asymmetric. With non-cooperative policy determination, the countries are shielded by endogenous terms-of-trade from distortionary domestic policies. Specifically, a fraction of welfare losses from distortionary domestic taxation is born by the consumers in the trade partner, who now have to face higher import prices and depreciated terms of trade. As a result, both trading countries will tend to set taxes inefficiently high in the sense that a coordinated reduction in taxes would induce a Pareto improvement.\(^{33}\) Finally,

\[^{33}\text{The terms-of-trade externality is predicted to be stronger for smaller and more open economies, which}\]
when countries are asymmetric, Proposition 3.5 no longer holds, even cooperatively-set optimal taxes in the open economy are no longer the same as in autarky, and under certain condition the model predicts convergence in the tax rates across countries. The Appendix provides a detail discussion of these issues.

### 3.4 Open Economy II: Fixed Costs of Trade

We now introduce fixed costs of entering the export market into the trade model of the previous section. This follows a vast theoretical literature started by Melitz (2003). A model with fixed costs of trade leads to a selection of agents into foreign trade with only the most productive of them being able to profitably export.³⁴ This is consistent with the empirical facts that only a small fraction of firms participate in international trade and these firms appear to be more productive (Bernard and Jensen, 1999; Eaton, Kortum and Kramarz, 2004). Moreover, empirical estimates of the fixed costs of entering the foreign market appear to be large quantitatively (Das, Roberts and Tybout, 2007; Bernard and Jensen, 2004).

Fixed costs of trade and selection into exporting have two important effects. First, they lead to discontinuously higher revenues for exporters relative to non-exporters, even controlling for their productivity. This is because revenues need to be discreetly higher for exporters in order for them to cover the fixed cost of trade. Second, selection into exporting creates an additional extensive margin of trade, which is sensitive to domestic redistribution policies. In this section, we study the interaction between these two effects and their implications for optimal taxation. We restrict our attention to the case of symmetric countries each with unit continuum of agents \( L = 1 \) and to cooperative policy determination. We start by considering the optimal linear tax and then add additional empirically tend to have larger governments (Rodrik, 1998; Alesina and Wacziarg, 1998). The importance of terms-of-trade externality on the level of optimal taxation was recently studied by Epifani and Gancia (2008), both theoretically and empirically. Finally, Mendoza and Tesar (2005) in a quantitative model of taxation in the open economy find the gains from policy coordination to be small.

³⁴ Alternative modeling strategies can generate selection into exporting without fixed costs. For example, Bernard, Eaton, Jensen and Kortum (2003) consider a model of Bertrand competition and Melitz and Ottaviano (2008) develop a model with non-CES demand.
policy instruments.

3.4.1 Equilibrium Properties

We start our analysis by characterizing an open economy equilibrium holding the tax policy \( t \) constant. Given fixed costs of trade, each agent now has to decide whether to serve the domestic market only at no fixed cost or to pay the fixed cost and both serve the domestic market and export. In the former case his real revenue is \( Q^{1-\beta}y^\beta \) as in the closed economy, while in the latter case his revenue is \( Y_x^{1-\beta}Q^{1-\beta}y^\beta \) as in the open economy of the previous section. Since the countries are symmetric, \( Y_x = 1 + \tau^{-\beta/(1-\beta)} > 1 \). The fixed cost equals \( f_x \) in the units of the final good. For an agent to profitably engage in foreign trade the increase in revenues from the access to the foreign market should exceed both the fixed cost and the variable cost of extra production for the export market. As we show below, only the most productive agents participate in the exporting activity, provided that the fixed cost is high enough.

We assume that the tax is levied on the revenues of the agents gross of fixed cost, or in other words, that the fixed cost of trade is not tax deductible. We make this assumption to preserve the tractability of the analysis. In the Appendix we show that the qualitative predictions remain unchanged if the fixed costs are tax deductible, while in the numerical investigation of Section 3.4.3 we show that the predictions of the specification with non-tax-deductible fixed costs are robust to making the fixed costs tax-deductible.

Under these circumstances, the problem of the agent with productivity \( n \) is given by:

\[
U_n = \max_{y, I_x \in \{0, 1\}} \left\{ \Delta + (1 - t) \left( 1 + I_x^{\tau^{-\beta}} \right)^{1-\beta} Q^{1-\beta}y^\beta - v(y/n) - I_x f_x \right\}
\]

\[
= tR + (1 - t)(1 - \beta/\gamma)r_n - I_n f_x,
\]

where \( I_n \) is the indicator variable of the agent’s export status. The appendix derives the optimality conditions for the choice of \( y_n \) and \( I_n \): the condition for \( y_n \) is similar to the
ones in the previous sections, while the entry cutoff condition is given by

\[(1 - t)(1 - \beta/\gamma) \left[ r_{n_x^+} - r_{n_x^-} \right] = f_x, \]

where \(n_x\) denotes the productivity cutoff, i.e. the productivity of an agent who is indifferent between serving and not serving the export market; \(n_x^+ (n_x^-)\) denotes the right (left) limit as the revenues jump discontinuously for an agent who decides to be an exporter. This condition implies that the increase in after-tax revenues net of the cost of effort should be large enough to cover the fixed cost for an agent to export profitably. The fixed cost \(f_x\) is assumed to be large enough so that \(n_x > n_{\text{min}}\). All agents with \(n > n_x\) serve the foreign market, while agents with \(n < n_x\) serve only the domestic market.

Equilibrium in the product market implies that aggregate output (which is split between private consumption and covering fixed costs of exporting) equals aggregate revenues, \(Q = R\). This equilibrium condition also implies balanced trade. Aggregate revenues are given by

\[ R = \int_0^\infty r_n dH(n) = Q^{1-\beta} \int_0^\infty \left( 1 + I_n \tau^{-\beta/\gamma} \right)^{1-\beta} y_n^\beta dH(n). \]  \hspace{1cm} (3.17)

Together with the agent’s optimal choice of \(\{y_n, I_n\}\), this condition allows us to characterize the symmetric world equilibrium.

In the Appendix we show that the relative revenue of agent \(n\) in the economy with fixed costs is given by:

\[
\frac{r_n}{R} = \frac{\left( 1 + I_n \tau^{-\beta/\gamma} \right)^{1-\beta} y_n^\beta}{\int_0^\infty \left( 1 + I_n \tau^{-\beta/\gamma} \right)^{1-\beta} y_n^\beta dH(n)} \begin{cases} 
\frac{n^{\beta/\gamma}}{n^{\gamma-\beta}}, & n \in [n_{\text{min}}, n_x), \\
\left( 1 + \tau^{\beta/\gamma} \right)^{1-\beta} n^{\gamma-\beta}, & n \in (n_x, n_{\text{max}}].
\end{cases}
\]  \hspace{1cm} (3.18)

Therefore, the distribution of relative revenues depends on the exporting cutoff \(n_x\) and on the foreign market access variable \(\Upsilon_x\). The former depends among other things on the

---

35 If the fixed cost is tax deductible, agents will compare before-tax net revenues with the fixed cost. However, since taxes reduce the scale of production for all agents, an increase in taxes will have the same effect on the exporting cutoff \(n_x\) independently of whether the fixed cost is tax deductible or not.
fixed cost of trade $f_x$, while the latter depends on the variable cost of trade $\tau$. Contrast this result with Lemma 3.4, which characterizes the distribution of relative revenues for the closed economy, as well as for the open economy without fixed costs. In closed economy and in open economy without selection into the export market, the distribution of relative revenues is uniquely pinned down by the underlying distribution of ability and does not respond to growing trade driven, for example, by a fall in $\tau$.

Figure 3.1: Relative Revenues (left) and Utility (right) in Open Economy with Fixed Costs

Note: Although revenues and utilities increase for all agents in open economy relative to autarky, relative revenues fall for non-exporters and rise for exporters. Revenues increase discontinuously for exporters, while utility is continuous and experiences a kink at $n_x$. Parameters for this figure are chosen as discussed in Section 3.4.3.

While revenues increase discontinuously for exporters, utility increases continuously with productivity $n$, but experiences a kink at $n_x$, as illustrated in Figure 3.1. Therefore, the access to foreign trade benefits disproportionately high-productivity agents who can easily cover the fixed cost and profitably export. This lies at the core of the distributional conflict in the open economy (cf. Helpman, Itskhoki and Redding, Chapter II of this dissertation). While in this model all agents gain from trade in absolute terms, trade worsens the inequality of relative revenues and utilities, as we discuss below. However, before we turn to the analysis of inequality, we summarize the comparative statics in the open economy with fixed costs:

\begin{proposition} \textit{(i) A reduction in the cost of trade (variable $\tau$ or fixed $f_x$) leads to an}

\end{proposition}

\footnote{This proposition requires a mild stability condition which is satisfied when $\beta$ is not too small. Appendix contains a formal proof.}
increase in real output $Q$ and a reduction in exporting cutoff $n_x$. (ii) All agents gain from greater trade integration, although these gains are not proportionally distributed.

We now characterize the patterns of income and utility inequality in the open economy with fixed costs using (3.18) and (3.16). First, observe from (3.18) that when either no agent participates in trade ($n_x = n_{\text{max}}$) or all agents export ($n_x = n_{\text{min}}$), the distribution of relative revenues is determined by the underlying ability distribution, as described in Lemma 3.4. This is not surprising since the former case corresponds to the closed economy and the latter case is similar to an open economy with no (or negligibly small) fixed costs, as we studied in previous sections.

![Figure 3.2](image.png)

Figure 3.2: Inequality of Revenues in Open Economy with Fixed Costs relative to Autarky: against variable trade costs (left) and fraction of exporting agents (right)

Note: As variable cost of trade $\tau$ falls, the export productivity cutoff $n_x$ decreases and the fraction of exporting agents $1 - H(n_x)$ increases. The two limiting cases ($n_x = n_{\text{max}}$ and $n_x = n_{\text{min}}$) are attained respectively when $\tau \to \infty$ and when $\tau$ is just above 1. In these two limiting cases, the inequality of revenues is the same in the open economy as in autarky. Baseline parameters as in Figure 3.1.

Our central result, which we prove in the Appendix, is that whenever some but not all agents export, the dispersion of relative revenues is strictly greater than in autarky. Intuitively, when $n_x \in (n_{\text{min}}, n_{\text{max}})$, the ranking of agents by revenues is the same as in autarky, but agents with high revenues ($n > n_x$) earn relatively more, while agents with low revenues ($n < n_x$) earn relatively less. Further, since inequality is the same in autarky and in the fully open economy, falling trade costs have a non-monotonic effect on inequality, first increasing it and later decreasing it, as illustrated in Figure 3.2.\footnote{One can show that decreasing fixed costs of trade have an inverted U-shape effect on inequality.} These
results replicate the findings of Helpman, Itskhoki and Redding (2008b) within a richer product and labor market equilibrium framework.

Finally, the implications for the inequality of utilities are largely the same as those for inequality of revenues. The only difference is that inequality of utilities is higher in an open economy with positive fixed cost than in autarky even when all agents export \( n_x = n_{\text{min}} \). This is because fixed costs of trade constitute a disproportionately higher burden for low productivity agents with relatively low revenues.

We summarize these findings in (see the Appendix for a formal proof):

**Proposition 3.7.** The inequality of relative revenues and utilities is higher in an open economy than in autarky when some, but not all agents export. Falling trade costs first increase and then decrease income inequality.

Contrast this proposition with the result of Section 3.3.1 that opening up to trade in an economy without fixed costs has no effect on inequality. This emphasizes that what generates inequality response to trade in this type of models is not agent heterogeneity per se, by rather the selection into the exporting activity.

We now study how these distributional consequences of trade affect optimal taxation.

### 3.4.2 Optimal Linear Tax Rate

Through what channels do taxes affect welfare in the open economy with fixed costs? The effect of a linear tax on agents’ utilities in the open economy with fixed costs is still given by (3.9). The effect of taxes on exporting cutoff \( n_x \) does not influence agent \( n_x \)’s utility directly since on the margin this agent is indifferent between exporting and not exporting. The only indirect effect of taxes on utility is still coming through their effect on aggregate output \( Q \). Further, in the cooperative solution with symmetric countries, the first order condition for the optimal tax is still given by (3.10). Therefore, optimal tax rate is still determined by (3.11'):

\[
\frac{t}{1-t} = \alpha \cdot \frac{1}{\hat{\varepsilon}} - (1 - \beta)(1 - \alpha),
\]
where $\tilde{\epsilon}$ and $\alpha$ are defined as before. What changes from the closed economy environment is the output elasticity with respect to the tax rate, $\tilde{\epsilon}$, and the equilibrium distribution of revenues which affects the inequality margin, $\alpha$.

We start with the inequality margin. Proposition 3.7 states that dispersion of relative revenues and utilities increases in the open economy with fixed costs. This intensifies the inequality margin of the optimal taxation. In the Appendix we discuss conditions under which we can unambiguously predict that $\alpha$ increases in the open economy. The numerical illustration provided in the left panel of Figure 3.3 demonstrates that $\alpha$ tracks very closely the dispersion of relative revenues.

![Figure 3.3: Equity Margin (left) and Efficiency Margin (right)](image_url)

Note: The left panel shows that inequality margin, $\alpha$, tracks closely the dispersion of relative revenues, $\text{var}(r/R)$. Note two differences: (i) for high values of $\tau$, $\alpha$ is much closer to $\alpha^a$ than $\text{var}(r/R)$ to $\text{var}(r^a/R^a)$. This is because with high $\tau$ and low $Y_x$, increased revenues of high-productivity exporters are almost entirely consumed by the fixed cost of trade; (ii) $\alpha > \alpha^a$ even when $\tau$ is low enough so that all agents participate in exporting. This is because fixed costs of trade constitute a disproportional burden on low-productivity agents with relatively low revenues, and as a result, the inequality of utilities (unlike the inequality of revenues) is still higher in this case than in autarky. The right panel illustrates how the efficiency margin $\tilde{\epsilon}$ increases in the open economy over and above the labor supply elasticity $\epsilon$ due to the extensive margin of trade; $\tilde{\epsilon}/\epsilon$ tracks closely $h(n_x)n_x$ until $n_x = n_{\text{min}}$ (here the productivity distribution is chosen to be Pareto). Baseline parameters as in Figure 3.1; in autarky, $\alpha^a = 0.204$ and $\text{var}(r^a/R^a) = 0.84$; linear tax rate held constant at the optimal autarky level ($t^a = 17.3\%$).

Consider now the efficiency margin. In the Appendix we prove the following:

**Lemma 3.5.** Assume the density $h(n)$ associated with cdf $H(n)$ exists at $n_x$. In an open economy equilibrium with fixed costs,

$$\tilde{\epsilon} = \frac{\text{d} \ln Q}{\text{d} \ln (1-t)} = \tilde{\epsilon} \cdot \frac{1 + \nu_x}{1 - \tilde{\epsilon}(1 - \beta)\nu_x},$$

where $\nu_x \equiv \frac{g_x}{\beta(1-t)R}h(n_x)n_x \geq 0$.

In addition, stability of the equilibrium requires $\tilde{\epsilon}(1 - \beta)\nu_x < 1$, which is always satisfied
for high enough $\beta$.

Note that whenever $h(n_x) > 0$, $\bar{\varepsilon} > \varepsilon$, i.e. trade with fixed costs and selection into the export market magnifies the efficiency margin of taxation. The intuition is straightforward: Given selection into the export market, taxation negatively affects not only the intensive margin, but also the extensive margin of trade. Therefore, the overall efficiency loss is greater than in the case when the extensive margin does not respond. The behavior of the efficiency margin in an open economy with fixed costs is illustrated in the right panel of Figure 3.3. The conclusion here is that the same feature of the model which leads to greater inequality in an open economy, i.e. the fixed costs of trade and selection into the exporting activity, also necessarily magnifies the efficiency margin.

To summarize, efficiency and inequality margins are linked tightly together, and both are driven by selection into the exporting activity. As a result, the optimal marginal tax rate can both go up or down in the open economy relative to autarky despite the fact that inequality increases. In other words, in order to capture the welfare gains from trade, a country has to accept an increase in income inequality, since trying to minimize income inequality resulting from trade brings about excessive efficiency losses.

![Figure 3.4: Optimal Linear Tax](image)

**Figure 3.4: Optimal Linear Tax**

*Note: This figure plots the optimal linear tax rate as a function of variable trade cost $\tau$, for both the specifications with tax-deductible and non-tax-deductible fixed costs of trade. The behavior of the optimal tax rate reflects the balance between the efficiency and inequality margins in an open economy. For this calibration, the efficiency margin dominates, and the optimal linear tax rate is lower in an open economy in which only a fraction of agents export than in autarky. Note that when all agents export, the tax rate in an open economy can be greater than in autarky. This is because $\alpha > \alpha^a$ in an open economy with fixed costs even when all agents export (see Figure 3.3). Baseline parameters as in Figure 3.1; the optimal autarky tax rate: $t^a = 17.3\%$.***
We illustrate these findings in Figure 3.4, where we plot the optimal linear tax rate as a function of variable trade cost $\tau$, both for the case of non-tax-deductible and tax-deductible fixed cost. Qualitatively, the results are the same in both cases, although the quantitative response of the tax rate is larger when the fixed cost is not tax-deductible. Note that for our calibration, the efficiency margin dominates the inequality margin. As a result, the optimal linear tax is lower in an open economy in which only a fraction of agents export than in autarky.\(^{38}\) A reduction in the marginal tax rate further exacerbates the inequality effects of trade. In the following section we study how these conclusions hold to the introduction of additional tax instruments.

Contrast this policy response to the policy response in the closed economy to exogenously growing income inequality driven by the dispersion of underlying ability (see Proposition 3.2). In the closed economy, an increase in the dispersion of relative revenues leads to an increase in the inequality margin $\alpha$, leaving the efficiency margin $\bar{\varepsilon}$ unchanged (Lemmas 3.2 and 3.3). Therefore, the optimal tax rate necessarily increases. We illustrate this in Figure 3.5. The left panel plots the optimal tax rates in the closed and open economies for the same equilibrium values of $\alpha$; in an open economy $\alpha$ is driven by a

\(^{38}\) It is also easy to develop an example in which the optimal tax is higher in an open economy equilibrium than in autarky. This simply requires choosing $h(n_x) \approx 0$. However, even in this case, the optimal tax rate will not be high enough to reduce the inequality margin $\alpha$ to its autarky level.
reduction in variable trade costs $\tau$, while in the closed economy the dispersion of ability distribution is adjusted to induce the same values for $\alpha$. As variable trade cost falls, the optimal tax rate in an open economy decreases despite an increase in the inequality margin $\alpha$ (see left panel of Figure 3.3), while the same increase in $\alpha$ would lead to a higher optimal tax rate in the closed economy. The right panel of Figure 3.5 plots the response of the optimal tax rates against the inequality margin $\alpha$. Again in the open economy case, $\alpha$ is driven by a reduction in variable trade cost, while in the closed economy it is driven by a corresponding increase in the dispersion of underlying ability distribution. The conclusion here is that the optimal policy response to growing income inequality can be very different depending on the original source of inequality increase.

3.4.3 Additional Tax Instruments

The restriction to a single tax rate appears to be particularly restrictive in the open economy case with fixed costs since there are two well-separated groups of agents – exporters and non-exporters – and an entry decision in addition to a baseline intensive margin. Therefore, we study here how our conclusions change when we introduce additional policy instruments such as an entry subsidy and differential marginal tax rates for exporters and non-exporters.

Denote by $t_d$ and $t_x$ the marginal tax rates on gross revenues for non-exporters and exporters respectively. Denote by $s$ the export market entry subsidy. With this notation, the problem of an agent becomes:

$$U_n = \max_{y, I_x} \left\{ \Delta + \left[ 1 - t_d (1 - I_x) - t_x I_x \right] \left( 1 + I_x \tau^{\frac{\beta}{1-\beta}} \right)^{1-\beta} Q^{1-\beta} y^\beta - v \left( \frac{y}{n} \right) - (f_x - s) I_x \right\}$$

(3.19)

and the government budget constraint implies

$$\Delta = \int_0^\infty \left\{ \left[ t_d (1 - I_n) + t_x I_n \right] r_n - s I_n \right\} dH(n),$$

(3.20)

Note that with these three instruments available, it is without loss of generality (for the set of feasible allocations) to assume that the fixed cost is not tax deductible. One can think alternatively about this setup as a two-bracket tax system.
where \( r_n = \left(1 + I_n \tau^{1-\beta}\right)^{1-\beta} Q^{1-\beta} y_n^\beta \). The expression for aggregate output (3.17) remains unchanged; however, agents’ optimality conditions determining intensive and extensive margins are now different, as described in the Appendix.

Analytical characterization of the optimal policy with three instruments is intractable. We thereby rely on the numerical methods to characterize the optimal policy in this case. But before we do so, we characterize the optimal entry subsidy taking the marginal tax rates in the two brackets as given. This exercise allows us to develop intuition for interpreting the numerical results to follow.

**Optimal Entry**

First, consider the case of no inequality aversion \( (\rho = 0) \) which implies a utilitarian planner and no risk-aversion at the individual level. In this case we have the following result (see the Appendix):

**Proposition 3.8.** With no inequality aversion \( (\rho = 0) \), the first best allocation requires \( t_d = t_x = -(1 - \beta)/\beta \) and \( s = 0 \), i.e. a constant negative marginal tax rate to offset the monopolistic distortion and no entry subsidy.

This proposition implies that entry is efficient provided that taxes fully offset the monopolistic distortion.\(^{40}\) At the individual level the fixed cost is traded off for additional revenues from entry into the export market. At the aggregate level entry increases variety at the cost of allocating a greater share of output to cover the fixed costs. It turns out that when the monopolistic distortions are offset, private incentives are perfectly aligned with the public welfare. Finally, note that under no inequality aversion a single linear tax is sufficient to fully restore efficiency and attain the first best allocation.

We now characterize the optimal utilitarian entry subsidy with exogenously set marginal tax rates on exporters and non-exporters (see the Appendix):

**Proposition 3.9.** Let the marginal tax rates \( t_d \) and \( t_x \) be given exogenously. Then: (i)

\(^{40}\) Interestingly, in the Melitz (2003) setup the allocation is efficient without any government intervention as the monopolistic profits are fully used to cover the entry costs as in Dixit and Stiglitz (1977).
With no inequality aversion \((\rho = 0)\) and given \(t_d = t_x = t\), the optimal entry subsidy is

\[
s^o = f_x \frac{1 - \beta (1 - t)}{1 - \frac{\beta}{\gamma} \beta (1 - t)}.
\]

Optimal utilitarian subsidy \(s^o\) is increasing in \(t\). When \(t_d \neq t_x\), \(\partial s^o / \partial t_d < 0\) and \(\partial s^o / \partial t_x > 0\). (ii) With positive inequality aversion \((\rho > 0)\), the optimal subsidy is strictly smaller than \(s^o\) given the same level of marginal tax rates \(t_d\) and \(t_x\).

This proposition is useful in developing intuition for the numerical results below. First, distortionary taxes \((t > -(1 - \beta) / \beta)\) negatively affect the entry decision and need to be offset by an entry subsidy. Second, this effect is stronger, the larger is the difference between the marginal tax rates on exporters and on non-exporters (formally, the smaller is \((1 - t_x) / (1 - t_d)\)). Third, setting \(t_d > t_x > -(1 - \beta) / \beta\) may partly or fully offset the need to use an entry subsidy.

The overall lessons here are twofold: (1) when marginal taxes are used to reduce inequality or for other exogenous reasons, entry has to be subsidized to avoid significant efficiency losses due to adjustment on the extensive margin; (2) proper entry incentives can be generated via different combinations of policy instruments (e.g., low \(t_x\) relative to \(t_d\) instead of positive \(s\)), which can be useful when entry subsidy is unavailable.\(^{41}\)

Finally, since entry has distributional consequences, under inequality aversion optimal entry subsidy will be lower than the utilitarian entry subsidy, and as a result entry will be less than efficient (from the point of view of a society that does care about inequality). With these results in mind, we turn to our numerical exercise.

### Numerical Solution

Before discussing the results of our numerical analysis, we briefly comment on the calibration of the parameters of the model. We choose the Pareto productivity distribution with a shape parameter of 2.2, consistent with the findings in Saez (2001) for the upper tail of the productivity distribution, where the entry decision is particularly relevant in our calibration. In the baseline case, the inequality aversion \(\rho\) is set to 2, which can be

---

\(^{41}\) As we discuss in the numerical section, the tax system will be adjusted in a way to both reduce inequality and provide entry incentives if some of the instruments are not available.
viewed as a combination of logarithmic individual preferences ($\rho_a = 1$) and logarithmic social welfare function ($\rho_g = 1$). The demand parameter, $\beta$, is calibrated to 0.75, which implies an elasticity of substitution of 4, consistent with the estimates in Bernard, Eaton, Jensen and Kortum (2003) and Broda and Weinstein (2006). For labor supply elasticity $\varepsilon$, we use the value of 0.5, consistent with the empirical estimates surveyed in Tuomala (1990, Chapter 3) and used in Saez (2001, 2002).

The fixed costs of trade $f_x$ are calibrated such that 40% of output is produced by exporting agents and exports account for 15% of consumption, also consistent with the evidence. Finally, $\tau$ is set in the baseline case to 1.5, corresponding to a variable trade cost of 50% in line with the estimates in Anderson and van Wincoop (2004). In our analysis we solve numerically for the optimal tax policy (given different sets of available policy instruments) for different values of the variable trade cost $\tau \in [1, 2]$.

We consider the following sets of policy instruments: (1) a single linear tax rate ($t$) with both deductible and non-deductible fixed costs; (2) a single linear tax rate with an entry subsidy ($t$ and $s$); (3) two separate tax rates for exporters and non-exporters ($t_d$ and $t_x$); (4) all three policy instrument ($t_d$, $t_x$ and $s$). In some cases we also solve for the optimal linear tax rate given a utilitarian subsidy, $s^\circ$.

The left panel of Figure 3.6 plots our first results regarding the optimal amount of entry in the export market ($n_x$). The dotted line characterizes the first-best entry under a utilitarian planner ($n_x^\circ$). With positive inequality aversion and for any set of available policy instrument, the optimal entry is strictly less than under the utilitarian planner ($n_x > n_x^\circ$). Note that this gap is greater when only a single tax instrument is available, but it becomes very small when two or more policy instruments are available.

The right panel of Figure 3.6 plots the fraction of welfare gains attained with a given set of policy instruments, where 100% corresponds to the welfare gains with all three policy instruments relative to no taxation. As variable trade costs fall and the trade flows increase, having a richer set of policy instruments becomes increasingly important. Having only a single tax rate may allow to capture as little as 50% of the welfare gains when $\tau$ is small, while having a single tax rate and an entry subsidy always guarantees at least 85%
of the welfare gains. The case with the marginal tax rates but no entry subsidy falls in between.

The overall lesson from the two panels of Figure 3.6 is that having a single linear tax in an environment with entry into the export market is very restrictive, both for encouraging the right amount of entry and for capturing the welfare gains. A single tax rate is not very successful in balancing the entry decision with redistribution. However, as soon as two or more policy instruments are available, nearly optimal entry can be ensured and a large fraction of the welfare gains can be captured.

Figure 3.7 plots the optimal entry subsidy (as a fraction of fixed cost $f_x$) as a function of variable trade costs and for different sets of policy instruments. First, note that when only a single tax rate is available (two lines in the middle), the optimal subsidy $s$ is very
close to the utilitarian subsidy $s^*$, but $s < s^*$ in all cases, consistent with Proposition 3.9. Second, the lower line corresponds to a hypothetical utilitarian subsidy in the case when only two marginal tax rates are available. Note that this subsidy is relatively small, suggesting that the tax system with two marginal tax rates will be adjusted in order to encourage entry (i.e., $t_x < t_d$ as we show below). Finally, the upper line is the optimal subsidy in the case when all three policy instruments are available. Note that in this case, $s \approx f_x$, i.e. almost all fixed costs of entry are reimbursed by the government. As we show below, this is the case because the government uses the available marginal tax rates to aggressively redistribute away from the high productivity exporters (i.e., $t_x \gg t_d$).

Further, Figure 3.8 plots the marginal tax rates as a function of variable trade costs with different sets of policy instruments. The left panel compares the cases with and without an entry subsidy when only a single marginal tax rate is available. The result is quite intriguing, although very intuitive. When an entry subsidy is unavailable, the marginal tax rate has to go down in the open economy in order to encourage more entry (cf. Figure 3.4 in Section 3.4.2). As a result, equality has to be traded off to ensure moderately efficient entry, which leads to a poor welfare performance of a single linear tax rate (Figure 3.6). In contrast, when an entry subsidy is available, the marginal tax rate goes up. The entry subsidy ensures the right amount of entry, while a higher marginal tax rate allows to moderate the increased inequality. As Figure 3.6, this combination of policy instruments is very successful from the welfare point of view.
The right panel of Figure 3.8 plots the marginal taxes on exporters (dashed lines) and non exporters (solid lines) both when an entry subsidy is and is not available. The results here are even more intriguing. When the entry subsidy is unavailable, the tax system becomes increasingly *regressive* as the trade costs fall, i.e. the marginal tax rate on poor low-productivity non-exporters ($t_d$) increases, while the marginal tax rate on rich high-productivity exporters ($t_x$) falls. At first glance counter-intuitive, this tax system adjustment is necessary to encourage the right amount of entry. However, it only further exacerbates increased inequality, which is why this tax system does not perform that well on the welfare account (right panel of Figure 3.6). Finally, when the entry subsidy is available, the tax system becomes strongly *progressive*, with a marginal tax rate on exporters almost twice as high as on non-exporters for low values of $\tau$. Here again, the entry subsidy ensures the right amount of entry, while the marginal taxes allow to aggressively redistribute away from the relative winners towards relative losers from trade.

The overall lesson from both panels of Figure 3.8 is that once a policy instrument (in this case an entry subsidy) is restricted, the optimal choice for other instruments may dramatically change in order to replicate the effects of the missing instrument. In particular, the tax schedule may revert from very progressive to strongly regressive.

Finally, in Figure 3.9 we plot the welfare and inequality outcomes of optimal policy for different sets of instruments. The right panel plots welfare as a function of trade costs – the dashed line in case with no redistribution policy and solid lines when optimal policies are in place. Note that the gains from trade are present even when $t = 0$ or $t = t^a$ is held fixed. However, having three tax instruments in place, allows to magnify the welfare gains from trade quite substantially when trade costs are small.

Lastly, the right panel of Figure 3.9 plots the resulting inequality of revenues the optimal policies are in place (the dashed line again corresponds to $t = 0$). Note that in all cases, even when the policy responds optimally to trade, trade results in higher inequality relative to autarky. Moreover, the optimal policy response often exacerbates inequality relative to a passive redistribution policy. For example, this is the case when three instruments are available and trade costs are high ($\tau > 1.5$), and this is the case when
only a linear tax is available and costs are low ($\tau < 1.4$). This confirms our conclusion in the end of Section 3.4.2 that in order to capture the most gains from trade, the countries may need to accept increased income inequality.

### 3.5 Discussion

Selection into export markets has consequences for both income inequality and efficiency losses from taxation. Trade openness intensifies both sides of the equity-efficiency trade-off, making the optimal redistribution policy response to trade ambiguous. More generally, when the original source of increasing income inequality is also inherently connected with the extent of welfare losses from taxation, the optimal policy response should not necessarily be to offset rising inequality.

The results of the paper immediately extend to any activity that features fixed costs and a non-trivial entry decision so that only the most productive agents can profitably participate. Specifically, consider the case of technology adoption. If new technology adoption requires paying a fixed cost, this will lead to similar implications for inequality and optimal policy response. Note that even with one heterogenous factor of production and no built-in skill-bias of technology, fixed costs and participation decision are enough to cause rising inequality in response to skill-neutral technological change.

The prediction of the model for the tax system is the following. In economic activities with extensive margin response, it is optimal to subsidize entry by covering a fraction of...
fixed costs and then redistribute revenues away from high-productivity agents who take advantage of the fixed cost activity using a progressive tax schedule. When, however, certain policy instruments are unavailable, other policy instruments should adjust in order to replicate as close as possible the optimal allocation. Specifically, when an entry subsidy is infeasible, it might well become optimal to use a regressive tax schedule, which encourages entry at the cost of less redistribution.

In the next versions of this paper we plan to solve numerically for the fully unrestricted Mirrlees policy, and study how closely a two-brackets tax system replicates the unrestricted allocation. Finally, an interesting avenue for further research is to contrast the optimal taxation of activities that are and are not subject to free entry conditions. This may shed light on the differential response of profit and income (or capital and labor) taxation in an open economies.
A. APPENDICES FOR CHAPTER I

A.1 Conditions for Incomplete Specialization

We derive here a limit on $b_A/b_B$ which secures an equilibrium in which both countries are incompletely specialized. When this condition is violated, the country with a relatively more rigid labor market (higher $b$) specializes in the production of the homogenous good. Throughout we assume for concreteness that $A$ is the relatively more rigid country, so that $b_A/b_B \geq 1$. We assume that $L$ is large enough in both countries so that both countries always produce the homogenous good. Following the main text, we analyze only equilibria with $\Theta_{xj} > \Theta_{dj} > \Theta_{\text{min}}$, so that not all producing firms export and there are also firms that exit. As shown in the text, this requires $f_x > f_d$ which we assume holds.

Given $b_A > b_B$, incomplete specialization implies that there is positive entry of firms in the differentiated sector of country $A$, i.e., $M_A > 0$. Equation (1.23) in the text implies that $M_A = 0$ whenever

$$
\delta_{dB} \left( \frac{Q_A}{Q_B} \right) ^{\zeta} \leq \delta_{xB}.
$$

When this condition is satisfied with equality we also find, using (1.19), that

$$
\delta_{dB} \left[ \frac{\Theta_{xB} f_d}{\Theta_{dB} f_x} \right] ^{\frac{1-\beta}{\beta}} = \delta_{xB}.
$$

(A.1)

Note that this relationship is a (generally nonlinear) upward-sloping curve in $(\Theta_{dB}, \Theta_{xB})$-space, lying between the 45°-line and $\Theta_{xB} = \Theta_{dB} \tau^{\beta/(1-\beta)} f_x/f_d$ (i.e., the equilibrium condition when $b_A = b_B$).

We can now prove the following

**Lemma A.1.** Let $\tau > 1$ and $b_A > b_B$. Then there exists a unique $\bar{b}(\tau) > 1$, with $\bar{b}'(\tau) > 0$, such that (A.1) holds for $b_A/b_B = \bar{b}(\tau)$. For $b_A/b_B < \bar{b}(\tau)$, there is incomplete specialization in equilibrium so that $M_A > 0$. For $b_A/b_B \geq \bar{b}(\tau)$, country $A$ specializes in the homogenous good so that $M_A = 0$.

**Proof:** Recall that $\Theta_{dB}$ is decreasing and $\Theta_{xB}$ is increasing in $\tau$. This implies that $\delta_{dB}/\delta_{xB}$ is increasing in $\tau$. (1.22) implies that $\tau^{1-\beta} \Theta_{xB}/\Theta_{dB}$ is increasing in $\tau$. Next, $\Theta_{xB}/\Theta_{dB}$ and $\delta_{dB}/\delta_{xB}$ are decreasing in $b_A/b_B$. These considerations, together with (A.1), imply that $\bar{b}(\tau)$ is unique and

1 In the special case of a Pareto distribution, (A.1) is a ray through the origin.
increasing in $\tau$ whenever it is finite.\footnote{Note that $\bar{b}(\tau) > 1$ by construction, since $\Theta_{x_B} = \Theta_{d_B} \tau^{3/(1-\beta)} f_x/f_d$ when $b_A = b_B$.} Finally, $Q_A/Q_B$ is decreasing in $b_A/b_B$. Therefore, from (1.23), $M_A > 0$ whenever $b_A/b_B < \bar{b}(\tau)$ and $M_A = 0$ whenever $b_A/b_B \geq \bar{b}(\tau)$.

Evidently, Lemma A.1 implies that there is an upper bound on how different the relative labor market frictions can be in the two countries for complete specialization not to occur in equilibrium. As we show in the numerical examples of Section 1.5.2, a wide range of $b_A/b_B > 1$ is consistent with incomplete specialization equilibrium. See the working paper version, Helpman and Itskhoki (2008), for the analysis of equilibria with complete specialization.

### A.2 Proof of Lemmas 1.1–1.5 and Proposition 1.3

**Proof of Lemma 1.1** follows immediately from (1.22). First note that in equilibria with $\Theta_{d_j} < \Theta_{x_j}$, we have $\Delta = \frac{1-\beta}{\beta} \left( \delta_{dA} \delta_{dB} - \delta_{xA} \delta_{dB} \right) > 0$. Indeed, $\Theta_{d_j} < \Theta_{x_j}$ implies $\frac{\delta_{dj}}{\delta_{xj}} > \frac{f_d \Theta_{xj}}{f_x \Theta_{dj}}$. Using these inequalities for $j = A, B$ together with (1.21) implies $\frac{\delta_{dA} \delta_{dB}}{(\delta_{xA} \delta_{dB})} > \tau^{2\beta/(1-\beta)} > 1$, in which case $\Delta > 0$.\footnote{This also implies $\delta_{dj} > \delta_{xj}$ in at least one country and in both countries in the vicinity of a symmetric equilibrium.} Then an increase in $b_A/b_B$ reduces $\Theta_{dA}$ and $\Theta_{xB}$ and increases $\Theta_{dB}$ and $\Theta_{xA}$ (see (1.22)). Therefore, $b_A > b_B$ implies $\Theta_{dA} < \Theta_{dB}$ and $\Theta_{xA} > \Theta_{xB}$ since in a symmetric equilibrium these relationships hold with equality.

**Proof of Lemma 1.2** also follows immediately from (1.22) and the fact that $\delta_{dj} > \delta_{xj}$, which we prove below (Lemma 1.4).

**Proof of Lemma 1.3** follows from (1.19) and Lemma 1.1. Note that (1.19) implies:

$$\left( \frac{Q_A}{Q_B} \right)^{\frac{\beta-\zeta}{\beta}} = \frac{\Theta_{dA}}{\Theta_{dB}} \left( \frac{b_B}{b_A} \right)^{\frac{\beta}{\beta}}.$$\(^2\)

When $b_A > b_B$, Lemma 1.1 implies $\Theta_{dA} < \Theta_{dB}$ and hence we have $Q_A < Q_B$.

**Proof of Lemma 1.4** follows from (1.23), the incomplete specialization requirement $M_j > 0$ and Lemmas 1.1 and 1.3. Specifically, when $b_A > b_B$, $M_A > 0$ together with (1.23) imply

$$\frac{\delta_{dB}}{\delta_{xB}} > \left( \frac{Q_B}{Q_A} \right)^{\frac{\zeta}{\beta}} > 1,$$

where the last inequality follows from Lemma 1.3. Lemma 1.1 implies that $\delta_{dA} > \delta_{dB}$ and $\delta_{xA} < \delta_{xB}$ since $\delta_{xj}$ is a decreasing function of $\Theta_{xj}$ ($z = d, x$ and $j = A, B$). Therefore, $\delta_{dA}/\delta_{xA} > \delta_{dB}/\delta_{xB} > 1$.\(^2\)
Proof of Lemma 1.5 follows from (1.23) and Lemmas 1.3 and 1.4. Specifically, (1.23) implies

\[ M_A - M_B = \frac{(1 - \beta)\phi_2}{\beta \Delta} \left[ (\delta_{dB} + \delta_{xA})Q_A^\zeta - (\delta_{dA} + \delta_{xB})Q_B^\zeta \right]. \]

When \( b_A > b_B \), Lemma 1.3 implies \( Q_A < Q_B \) and Lemma 1.4 implies \( \delta_{dA} > \delta_{dB} > \delta_{xB} > \delta_{xA} \). Therefore, in this case \( M_A < M_B \).

Proof of Proposition 1.3 follows from Lemmas 1.1, 1.4 and 1.5 and the definition of intra-industry trade. When \( b_A > b_B \), Lemma 1.1 states that \( \Theta_{xA} > \Theta_{xB} \) and \( \Theta_{dA} < \Theta_{dB} \) which implies that a larger fraction of firms export in country \( B \): \[ \frac{1 - G(\Theta_{xj})}{1 - G(\Theta_{dj})} > \frac{1 - G(\Theta_{xj})}{1 - G(\Theta_{dj})}. \]

In the text the share of intra-industry trade is shown to equal \( X_A/X_B = \delta_{xA}M_A/(\delta_{xB}M_B) \). Using (1.23), we have:

\[ \frac{X_A}{X_B} = \frac{\delta_{dA}}{\delta_{xA}} \frac{Q_B}{Q_A} \frac{\zeta}{\zeta - 1}. \]

From (1.22) and (1.26) and using Lemma 1.4, an increase in \( b_A/b_B \) leads to a decrease in \( \delta_{dB}/\delta_{xB} \) and to increases in \( \delta_{dA}/\delta_{xA} \) and \( Q_B/Q_A \). Therefore, an increase in \( b_A/b_B \) reduces \( X_A/X_B \).

In Appendix A.4, we prove additionally that under Pareto-distributed productivity the total volume of trade increases in the proportional gap between relative labor market frictions, \( b_A/b_B \).

A.3 Derivation of results on productivity for Section 1.4.3

We first show that \( \varphi_{zj} = \varphi(\Theta_{zj}) \) is monotonically increasing in \( \Theta_{zj} \). The log-derivative of \( \varphi(\Theta_{zj}) \) is

\[ \hat{\varphi}_{zj} = \Theta_{zj}G'(\Theta_{zj}) \left[ \frac{\Theta_{zj}}{\int_{\Theta_{zj}}^{\infty} \theta dG(\theta) - \int_{\Theta_{zj}}^{\infty} \Theta_{1/\beta}^{1/\beta} dG(\theta)} - \frac{\Theta_{zj}^{1/\beta}}{\int_{\Theta_{zj}}^{\infty} \Theta_{1/\beta}^{1/\beta} dG(\theta)} \right] \hat{\Theta}_{zj} \quad \text{for } z = d, x. \]

The term in the square brackets is positive since

\[ \frac{\Theta_{zj}^{1/\beta}}{\int_{\Theta_{zj}}^{\infty} \Theta_{1/\beta}^{1/\beta} dG(\theta)} < \left( \frac{\Theta_{zj}}{\int_{\Theta_{zj}}^{\infty} \theta dG(\theta)} \right)^{1/\beta} < \int_{\Theta_{zj}}^{\infty} \Theta_{1/\beta}^{1/\beta} dG(\theta), \]

where the first inequality follows from Jensen’s inequality and the second inequality comes from the fact that \( \beta < 1 \) and \( \Theta_{zj} < \int_{\Theta_{zj}}^{\infty} \theta dG(\theta)/1 - G(\Theta_{zj}) \).
Next we provide the general expression for a log-change in aggregate productivity:

\[
\hat{\text{TFP}}_j = \left\{ 1 + \frac{\kappa_{dj}}{\kappa_{xj}} \left[ \frac{\kappa_{dj}}{\kappa_{xj}} \left( \Theta_{xj}^{1-\beta} - \text{TFP}_j \right) + \left( \text{TFP}_j - \Theta_{dj}^{1-\beta} \right) \delta_{dj} \left( \varphi_{xj} - \varphi_{dj} \right) \right] \right\} \delta_{dj} \delta_{dj} \delta_{xj} \varphi_{xj} + \delta_{xj} \varphi_{xj} \hat{\Theta}_{dj},
\]

(A.2)

where

\[
\kappa_{zj} \equiv \kappa(\Theta_{zj}) = \frac{f_z(\Theta_{zj})}{G'(\Theta_{zj})} = \frac{\Theta_{zj} G'(\Theta_{zj})}{\int_{\Theta_{zj}}^{\infty} \Theta dG(\Theta)}.
\]

A series of sufficient conditions can be suggested for the terms in curly brackets to be positive. Since \( \text{TFP}_j \geq \Theta_{dj}^{1-\beta}/\beta \) is always true, it is sufficient to require that

\[
\text{TFP}_j \leq \Theta_{xj}^{1-\beta}/\beta,
\]

which holds for large enough \( \Theta_{xj} \), i.e., when the economy is relatively closed. However, this inequality fails to hold when \( \Theta_{xj} \) approaches \( \Theta_{dj} \). If this condition fails, it is sufficient to have

\[
\left[ \frac{\text{TFP}_j - \Theta_{xj}^{1-\beta}/\beta}{\text{TFP}_j - \Theta_{dj}^{1-\beta}/\beta} \right] \geq \kappa_{xj}/\kappa_{dj},
\]

which is, in particular, satisfied when \( \kappa_{dj} \geq \kappa_{xj} \). This latter condition is always satisfied if \( \kappa(\cdot) \) is a non-increasing function and is equivalent to

\[
-\Theta G''(\Theta) G'(\Theta) \geq 2 + \Theta G'(\Theta) \int_{\Theta}^{\infty} \xi dG(\xi);
\]

that is, \( G''(\cdot) \) has to be negative and large enough in absolute value. This condition is satisfied for the Pareto distribution since in this case \( \kappa(\cdot) \) is constant and \( \kappa_{dj} \equiv \kappa_{xj} \). However, it is not satisfied, for example, for the exponential distribution.

Finally, the necessary and sufficient condition is

\[
\left( \kappa_{xj} - \kappa_{dj} \right) \text{TFP}_j - \left( \kappa_{xj} \Theta_{xj}^{1-\beta}/\beta - \kappa_{dj} \Theta_{dj}^{1-\beta}/\beta \right) \leq \varphi_{xj} - \varphi_{dj}
\]

which is satisfied when

\[
\frac{\left( \kappa_{xj} - \kappa_{dj} \right) \left( \text{TFP}_j - \Theta_{dj}^{1-\beta}/\beta \right)}{\varphi_{xj} - \varphi_{dj}} = \left( \kappa_{xj} - \kappa_{dj} \right) \left[ \varphi_{xj} + \frac{\varphi_{dj} - \Theta_{dj}^{1-\beta}/\beta}{\varphi_{xj} - \varphi_{dj}} \right] \leq 1.
\]

This condition also does not hold in general; however, it is certainly satisfied for large enough \( \Theta_{xj} \).

Now we provide the derivation of equation (1.29) under the assumption of Pareto-distributed productivity draws. When \( \Theta \) is distributed Pareto with the shape parameter \( k > 1/\beta \), there is a straightforward way of computing the change in \( \text{TFP}_j \). Taking the log derivative of (1.28), we have

\[
\hat{\text{TFP}}_j = \left[ \frac{\delta_{dj} \varphi_{dj} + \delta_{xj} \varphi_{xj} \hat{\Theta}_{dj}}{\delta_{dj} \varphi_{dj} + \delta_{xj} \varphi_{xj}} + \frac{\delta_{dj} \varphi_{dj} + \delta_{xj} \varphi_{xj}}{\delta_{dj} \varphi_{dj} + \delta_{xj} \varphi_{xj}} \right] + \frac{\delta_{dj} \varphi_{dj} \varphi_{xj} + \delta_{xj} \varphi_{xj} \hat{\Theta}_{dj}}{\delta_{dj} \varphi_{dj} + \delta_{xj} \varphi_{xj}}.
\]

Under the Pareto assumption, the free-entry condition (1.20) can be written as \( \delta_{dj} + \delta_{xj} = kf_e \),
which implies \( \delta dj \hat{\delta} dj + \delta xj \hat{\delta} xj = 0 \). We use this to simplify

\[
\text{TFP}_j = \frac{\delta dj \hat{\varphi} dj (\hat{\delta} dj + \hat{\varphi} dj) + \delta xj \varphi xj (\hat{\delta} xj + \hat{\varphi} xj)}{\delta dj \hat{\varphi} dj}.
\]

Next note that \( \delta zj = f_z \frac{k}{k-1} (\Theta_{\min}/\Theta_{zj})^k \) so that \( \hat{\delta} zj = -k \hat{\Theta} zj \) and \( \varphi xj = \frac{k-1}{k-1/\beta} \Theta_{zj}^{(1-\beta)/\beta} \), implying \( \hat{\varphi} xj = (1-\beta)/\beta \hat{\Theta} xj \). Thus, the log-derivative of the free-entry condition can also be written as \( \delta dj \hat{\Theta} dj + \delta xj \hat{\Theta} xj = 0 \). Therefore,

\[
\delta dj (\hat{\delta} dj + \hat{\varphi} dj) = -\delta xj (\hat{\delta} xj + \hat{\varphi} xj) = -[k - (1-\beta)/\beta] \delta dj \Theta dj.
\]

Using this, we obtain our result (1.29) in the text.

Finally, we discuss an alternative measure of productivity which takes into account the sectoral composition of resource allocation:

\[
\text{TFP'}_j = L - N_j H_{0j} + \frac{N_j H_j}{L} \cdot \text{TFP}_j,
\]

which is a weighted average of \( x_0j = H_{0j}/N_{0j} \) (the productivity in the homogenous sector) and \( \text{TFP''}_j = x_j \cdot \text{TFP}_j \) (productivity in the differentiated-product sector). The weights are the respective fractions of the two sectors in the labor force. Note that both sectoral productivity measures take into account unemployment of labor. Further, note that \( \text{TFP'}_j = \text{TFP}_j + \hat{x}_j \). If \( \text{TFP''}_j > x_0j \), an extensive margin increase in the size of the differentiated sector improves productivity. Reduction in trade costs or labor market frictions in the differentiated sector (decreases in \( a_j/a_{0j} \)) shift resources towards the differentiated sector by increasing \( N_j \). Moreover, decreases in labor market frictions improve sectoral labor market tightness and hence increase productivity. These are the additional effects captured by this alternative measure of aggregate productivity.

### A.4 Solution under Pareto assumption for Section 1.5.2

We characterize here the solution of the model under the assumption that productivity draws \( \Theta \) are distributed Pareto with the shape parameter \( k > 2 \). That is, \( G(\Theta) = 1 - (\Theta_{\min}/\Theta)^k \) defined for \( \Theta \geq \Theta_{\min} \). We use this characterization in Section 1.5.2 in order to solve numerically for the equilibrium response of unemployment to different shocks. In the end of this appendix we provide some analytical results under Pareto-distributed productivity referred to in the text.

Pareto-distributed productivity leads to the following useful functional relationship:

\[
\delta zj = f_z \frac{k}{k-1} \int_{\Theta_{zj}}^{\infty} \Theta dG(\Theta) = f_z \frac{k}{k-1} \left( \frac{\Theta_{\min}}{\Theta_{zj}} \right)^k, \quad z = d, x,
\]

so that \( \hat{\delta} zj = -k \hat{\Theta} dj \). As a result, we can rewrite the free entry condition (1.20) as

\[
f_d \Theta dj^k + f_x \Theta xj^k = (k-1)f_e \Theta_{\min}^{-k} \quad \Leftrightarrow \quad \delta dj + \delta xj = kf_e.
\]
Manipulating cutoff conditions (1.19) and the free entry condition above, we can obtain two equations to solve for \( \{\Theta_{dj}, \Theta_{xj}\} \). For concreteness, consider \( j = A \):

\[
fd \Theta_{dA}^{-k} + fx \Theta_{xA}^{-k} = (k - 1) f_e \Theta_{\min}^{-k},
\]

\[
fx \left[ \tau \frac{\beta k}{fd} \frac{fx}{fd} \psi \frac{\beta k}{fx} \right]^{-k} \Theta_{dA}^{-k} + fd \left[ \tau \frac{\beta k}{fd} \frac{fx}{fd} \psi \frac{\beta k}{fx} \right]^{k} \Theta_{xA}^{-k} = (k - 1) f_e \Theta_{\min}^{-k},
\]

where \( \psi \equiv b_A / b_B \) is the relative labor market rigidity of country \( A \). This is a linear system in \( \{\Theta_{dA}^{-k}, \Theta_{xA}^{-k}\} \) and there are similar conditions for country \( B \), with \( \psi^{-1} \) replacing \( \psi \). The solution to this system is given by

\[
\Theta_{dA}^{-k} = \frac{fx}{\Delta \Theta} \left[ \tau \frac{\beta k}{fd} \left( \frac{fx}{fd} \right)^{k-1} \psi \frac{\beta k}{fx} - 1 \right],
\]

\[
\Theta_{xA}^{-k} = \frac{fd}{\Delta \Theta} \left[ 1 - \tau \frac{\beta k}{fd} \left( \frac{fx}{fd} \right)^{-(k-1)} \psi \frac{\beta k}{fx} \right],
\]

where

\[
\Delta \Theta = \frac{f_x^2 \Theta_{\min}^{-k}}{(k - 1) f_e} \tau \frac{\beta k}{fx} \left( \frac{fx}{fd} \right)^{-k} \psi \frac{\beta k}{fx} \left[ \tau \frac{\beta k}{fd} \left( \frac{fx}{fd} \right)^{2(k-1)} - 1 \right] > 0.
\]

Using this result we can derive a condition on primitive parameters for \( \Theta_{dj} < \Theta_{xj} \) to hold in equilibrium:

\[
\frac{fd}{fd + fx} \left( \tau \frac{\beta k}{fd} \frac{fx}{fd} \right)^{k} + \frac{fx}{fd + fx} \left( \tau \frac{\beta k}{fx} \frac{fx}{fd} \right)^{-k} > \max \left\{ \psi \frac{\beta k}{fd}, \psi \frac{\beta k}{fx} \right\}, \tag{A.3}
\]

which is satisfied for large \( \tau \) and for \( \psi \equiv b_A / b_B \) not very different from one. Next note that as \( \tau \to \infty \), \( \Theta_{xA} \to \infty \) and \( \Theta_{dA} \to \left[ \frac{fx}{(k - 1) f_e} \right]^{1/k} \Theta_{\min} = \Theta_{dA}^c \). Therefore, the condition for \( \Theta_{dA} > \Theta_{\min} \) is \( k < 1 + fd/fx \) which is equivalent to the condition in the text. One can also show that \( \Theta_{dA} \) decreases in \( \tau \) in the range \( \tau \in (\tau^*, \infty) \) where

\[
\tau^* = \tau^*(\psi, fx/fd) : \quad (\tau^*)^{\frac{\beta k}{fx}} = \left( \frac{fd}{fx} \right)^{k-1} \left( \psi \frac{\beta k}{fx} + \sqrt{\psi \frac{2\beta k}{fx} - 1} \right).
\]

The first cutoff condition in (1.19) allows to solve for \( Q_j \) once \( \Theta_{dj} \) is known; \( Q_j \) is also decreasing in \( \tau \) in the range \( (\tau^*, \infty) \). It is straightforward to show that \( Q_j \) decreases in \( b_A \) and increases in \( b_B \). Using the Pareto assumption and the equation for \( M_j \) (1.23), we get

\[
M_j = \phi_2 \frac{k - 1}{k} \Theta_{\min}^{-k} \frac{fd Q_j^2 \Theta_{dA}^{-k} - fx Q_j^2 \Theta_{xA}^{-k}}{f_x^2 \Theta_{dA}^{-k} \Theta_{dB}^{-k} - f_x^2 \Theta_{xA}^{-k} \Theta_{xB}^{-k}}. \tag{A.4}
\]

The condition for \( M_A > 0 \) can then be written as

\[
\left( \Theta_{xB} \Theta_{dB} \right)^k \left( \frac{Q_A}{Q_B} \right) > 1.
\]
One can show that this inequality imposes a restriction on parameters \( \{ \tau, \psi, f_x/f_d \} \) such that \( \tau > \tau^*(\psi, f_x/f_d) \), which implies that \( Q_j \) is decreasing in \( \tau \) whenever there is no complete specialization (\( M_j > 0 \) for both \( j \)). This is consistent with Lemma 1.2 in the text.

Finally, using the condition for \( N_j \) (1.24) and the free entry condition under the Pareto assumption, we get:

\[
\omega_{0j} N_j = \phi_1 \frac{1 - \tau}{\phi_2} M_j \left[ \delta_{d_j} + \delta_{x_j} \right] = \phi_1 \frac{1 - \tau}{\phi_2} k f_e M_j.
\]

That is, under the Pareto assumption, \( N_j \) is always proportional to \( M_j \). The remaining equilibrium condition is \( H_j = x_j N_j \) and the expression for unemployment (1.18). Labor market tightness in the two sectors and the expected income \( \omega_{0j} \) are still determined by (1.12), (1.13) and (1.17).

We use the equations above to solve for equilibrium comparative statics numerically. Additional analytical results can also be obtained under the Pareto assumption for \( M_j, N_j \) and \( u_j \) departing from (A.4).

**Volume of Trade (remark for Section 4.3)** Under the Pareto assumption we can get a simple prediction about the response of the trade volume to \( \tau, b_A \) and \( b_B \). Recall that the total volume of trade (when \( b_A > b_B \) equals \( 2X_B \) where we have

\[
X_B = \phi_2^{-1} M_B x_B = \frac{\delta_{d_A} Q_A^\psi - Q_A^\psi}{\delta_{d_A} x_A} - 1 = \frac{\frac{\delta_{d_A} x_A}{a_A}}{\delta_{x_A} x_B} - 1 = \frac{\frac{\delta_{d_A} x_A}{a_A}}{\delta_{x_A} x_B} - 1
\]

As \( b_A \) increases or \( b_B \) falls, the denominator remains unchanged while \( \Theta_{x_A}/\Theta_{d_A} \) and \( Q_B \) increase and \( Q_A \) decreases. As a result the volume of trade unambiguously rises. Finally, one can also show that \( X_B \) decreases in \( \tau \). Substitute the expression for \( \Theta_{x_A}/\Theta_{d_A} \) (derived from (1.19)) in the expression for \( X_B \) to get

\[
X_B = Q_A^\psi \frac{\tau^{\frac{\beta k}{\beta}}}{\tau^{\frac{\beta k}{\beta}} - 1}.
\]

Now note that \( X_B \) decreases in \( \tau \) since \( Q_A \) and \( Q_B/Q_A \) decrease in \( \tau \) and \( Q_B > Q_A \).

**Proof that \( N_A/N_B \) decreases in \( \tau \) when \( b_A > b_B \)** In the text we show that \( N_A + N_B \) increases as \( \tau \) falls. We show now that when \( b_A > b_B \), \( N_A/N_B \) decreases as \( \tau \) falls, which implies that \( N_B \) necessarily increases. Under the Pareto assumption \( \delta_{d_j} + \delta_{x_j} = k f_e \). Therefore, (1.23) and (1.24) imply

\[
\frac{\omega_{0A} N_A}{\omega_{0B} N_B} = \frac{M_A}{M_B} = \frac{\delta_{dA} Q_B^\psi - \delta_{xA} Q_A^\psi}{\delta_{dA} Q_B^\psi - \delta_{xA} Q_A^\psi} = \frac{1 - \frac{\delta_{dA} Q_A^\psi}{\delta_{dA} Q_B^\psi}}{1 + \frac{Q_A^\psi}{Q_B^\psi}} < 1,
\]

where the last inequality comes from Lemma 1.5 under the assumption that \( b_A > b_B \). Recall that \( \omega_{0A}/\omega_{0B} \) depends only on \( a_{0A}/a_{0B} \) and does not depend on \( \tau \). From Proposition 1.1, \( Q_B/Q_A \) increases as \( \tau \) falls. Taking this and the fact that \( N_A < N_B \) into account, it is sufficient to show
that $d\delta_{x_B} - d\delta_{d_A} = \delta_{x_B}\hat{\delta}_{x_B} - \delta_{d_A}\hat{\delta}_{d_A} > 0$ in response to a fall in $\tau$, to establish that $N_A/N_B$ declines in this case. Under Pareto-distributed productivity,

$$\frac{\delta_{x_B}\hat{\delta}_{x_B} - \delta_{d_A}\hat{\delta}_{d_A}}{-\hat{\tau}} = k\left(\delta_{d_A}\hat{\theta}_{d_A} - \delta_{x_B}\hat{\theta}_{x_B}\right) = k\left(\delta_{d_B}\delta_{x_B} - \delta_{d_A}\delta_{x_A}\right)\left(\delta_{d_A} - \delta_{x_A}\right) > 0,$$

where the second equality comes from (1.22) and the inequality is obtained by Lemma 1.4 and the fact that under the Pareto assumption $\delta_{d_A} + \delta_{x_A} = \delta_{d_B} + \delta_{x_B} = kf_e$. This proves that $N_B$ increases as $\tau$ falls when $b_A > b_B$. Since changes in $\tau$ do not affect labor market tightness, $x_0B$ and $x_B$, the only effect on the unemployment rate $u_B$ is through $N_B$, and hence the unemployment rate in the flexible country increases in response to trade liberalization.
B. APPENDICES FOR CHAPTER II

B.1 Sectoral Equilibrium

Given a Pareto distribution of worker ability, $G_a(a) = 1 - (a_{\text{min}}/a)^k$, a firm that chooses a screening threshold $a_c$ hires a measure $h = n(a_{\text{min}}/a_c)^k$ of workers with average ability $\bar{a} = ka_c/(k-1)$. Therefore the production technology (2.3) can be written as:

$$y = \kappa y a^{\gamma} a_c^{1-\gamma k}, \quad \kappa_y = \frac{k}{k-1} a_c^{\gamma k}$$ \hspace{1cm} (B.1)

Since in equilibrium all firms with the same productivity behave symmetrically, we index firms by $\theta$ from now onwards. Exporters choose output to supply to the domestic market ($y_d(\theta)$) and the export market ($y_x(\theta)$) to equate marginal revenues in the two markets, which from (2.2) implies:

$$\left[\frac{y_d(\theta)}{y_x(\theta)}\right]^{\beta - 1} = \tau^{-\beta} (A*/A) .$$ \hspace{1cm} (B.2)

Total firm output is:

$$y(\theta) = y_d(\theta) + y_x(\theta).$$ \hspace{1cm} (B.3)

Together (B.2) and (B.3) imply:

$$y_d(\theta) = y(\theta) / \Upsilon(\theta), \quad y_x(\theta) = y(\theta) \frac{\Upsilon(\theta) - 1}{\Upsilon(\theta)},$$ \hspace{1cm} (B.4)

where:

$$\Upsilon(\theta) = 1 + I_x(\theta) \tau^{-\frac{\beta}{1-\beta}} \left( \frac{A^*}{A} \right)^{\frac{1}{1-\beta}}$$ \hspace{1cm} (B.5)

and $I_x(\theta)$ is an indicator variable that equals one if the firm exports and zero otherwise. Note that for non-exporters we have $y_d(\theta) = y(\theta)$ and $\Upsilon(\theta) = 1$, which implies that the allocation rule in (B.4) also holds for non-exporters. Total firm revenue is:

$$r(\theta) = r_d(\theta) + r_x(\theta).$$ \hspace{1cm} (B.6)

Together (2.2), (B.2) and (B.6) imply:

$$r_d(\theta) = r(\theta) / \Upsilon(\theta), \quad r_x(\theta) = r(\theta) \frac{\Upsilon(\theta) - 1}{\Upsilon(\theta)},$$
where again this allocation rule also holds for non-exporters. Additionally, (2.2), (B.2) and (B.4) imply:

\[ r(\theta) \equiv r_d(\theta) + r_x(\theta) = \Upsilon(\theta)^{1-\beta} A y(\theta)^{\gamma}, \quad (B.7) \]

Therefore the firm’s problem can be written as:

\[
\pi(\theta) \equiv \max_{n \geq 0, a \geq a_{\min}, I_x \in \{0, 1\}} \left\{ \frac{1}{1 + \beta_\gamma} \left[ 1 + I_x \tau - \frac{\gamma}{\beta} \left( \frac{A^*}{A} \right) \right] ^{1-\beta} A \left( k_\gamma a \gamma a_c^{-\gamma k} \right)^{\beta} - bn - \frac{c}{\beta_\gamma} - f_d - I_x f_x \right\},
\]

where \(1/(1 + \beta_\gamma)\) is the equilibrium share of revenue received by the firm as the outcome of the bargaining game with workers, which is modelled as in Stole and Zweibel (1996a,b). We discuss the bargaining game and derive its solution in Subsection B.6 below.

The firm’s first-order conditions for the measure of workers sampled \((n)\) and the screening ability threshold \((a_c)\) are:

\[
\frac{\beta_\gamma}{1 + \beta_\gamma} r(\theta) = bn(\theta), \tag{B.8}
\]

\[
\frac{\beta(1 - \gamma_k)}{1 + \beta_\gamma} r(\theta) = ca_c(\theta)^{\delta}. \tag{B.9}
\]

Combining the two first-order conditions (B.8) and (B.9) we obtain the following relationship between \(n(\theta)\) and \(a_c(\theta)\):

\[(1 - \gamma_k)bn(\theta) = c a_c(\theta)^{\delta}.\]

Using the expression for \(r(\theta)\) in (B.7), we can solve explicitly for

\[
n(\theta) = \phi_1 \phi_2 \left( \frac{k a_c}{k - 1} \right)^{\beta_\gamma} \quad \text{and} \quad \phi_2 \equiv \left( \frac{1 - \gamma_k}{\beta_\gamma} \right)^{\frac{1}{\gamma}}
\]

and, as in the text of the paper, \(\Upsilon(\theta) = \Upsilon_d = 1\) for \(\theta \in [\theta_d, \theta_x)\) and \(\Upsilon(\theta) = \Upsilon_x = 1 + \tau^{-\beta/(1-\beta)}(Q^*/Q)^{-\alpha/(1-\beta)}\) for \(\theta \geq \theta_x\). Also we solve for

\[
\frac{\beta_\gamma}{1 + \beta_\gamma} r(\theta) = bn(\theta) = \phi_1 \phi_2 \left( \frac{k a_c}{k - 1} \right)^{\beta_\gamma} \left( \frac{\tau(1 - \gamma_k)}{\Upsilon(\theta)} \right) ^{\frac{1}{1-\gamma}} \Upsilon(\theta) \frac{1-\beta}{1-\gamma} Q^{-\frac{\beta-\alpha}{\gamma} \theta^\frac{\beta}{\gamma}}, \tag{B.10}
\]

\[
\pi(\theta) + f_d + I_x(\theta)f_x = \frac{\Gamma}{1 + \beta_\gamma} r(\theta) = \frac{\Gamma}{1 + \beta_\gamma} \phi_1 \phi_2 \left( \frac{k a_c}{k - 1} \right)^{\beta_\gamma} \left( \frac{\tau(1 - \gamma_k)}{\Upsilon(\theta)} \right) ^{\frac{1}{1-\gamma}} \Upsilon(\theta) \frac{1-\beta}{1-\gamma} Q^{-\frac{\beta-\alpha}{\gamma} \theta^\frac{\beta}{\gamma}}, \tag{B.11}
\]

\[
h(\theta) = n(\theta) \left( \frac{a_{\min}}{a_c(\theta)} \right)^k = \left( \frac{k a_c}{k - 1} \right)^{\beta_\gamma} \phi_1 \phi_2 \left( \frac{k a_c}{k - 1} \right)^{\beta_\gamma} \left( \frac{\tau(1 - \gamma_k)}{\Upsilon(\theta)} \right) ^{\frac{1}{1-\gamma}} \Upsilon(\theta) \frac{1-\beta}{1-\gamma} Q^{-\frac{\beta-\alpha}{\gamma} \theta^\frac{\beta}{\gamma}}.
\]

142
so that \( \kappa \equiv \phi_1 \phi_2^{\beta(1-\gamma k)} \). Finally, we solve for the wage rate:

\[
w(\theta) = \frac{\beta \gamma}{1 + \beta \gamma} \frac{r(\theta)}{h(\theta)} = b \frac{n(\theta)}{h(\theta)} = b \left( \frac{a_c(\theta)}{\theta_{\text{min}}} \right)^k
\]

\[
= a_{\text{min}} \phi_1 \phi_2^{\beta(1-\gamma k)} c^{-(1-\gamma k)} b^{-\frac{\beta}{\gamma}} (1-\gamma k) \theta^{\frac{\beta}{\gamma}}.
\] \hspace{1cm} \text{(B.12)}

Note that we have the following relationship, which proves useful in further derivations:

\[
w(\theta)h(\theta) = bn(\theta) = \frac{\beta \gamma}{1 + \beta \gamma} r(\theta).
\]

Now, using the zero-profit cutoff condition,

\[
\pi(\theta_d) = \frac{\Gamma}{1 + \beta \gamma} r(\theta) = \frac{\Gamma}{\beta \gamma} \phi_1 \phi_2^{\beta(1-\gamma k)} c^{-\frac{\beta(1-\gamma k)}{\gamma}} b^{-\frac{\beta}{\gamma}} Q^{-\frac{\beta}{\gamma}} - f_d = 0,
\]

we can express all firm-level variables solely as the following functions of \( \theta/\theta_d \), \( b \) and \( \Upsilon(\theta) \) reported in the paper:

\[
\begin{align*}
  r(\theta) &= \Upsilon(\theta)^{1-\beta} \cdot r_d \cdot \left( \frac{\theta}{\theta_d} \right)^{\frac{\beta}{\gamma}}, \\
n(\theta) &= \Upsilon(\theta)^{1-\beta} \cdot n_d \cdot \left( \frac{\theta}{\theta_d} \right)^{\frac{\beta}{\gamma}}, \\
o_c(\theta) &= \Upsilon(\theta)^{1-\beta} \cdot a_d \cdot \left( \frac{\theta}{\theta_d} \right)^{\frac{\beta}{\gamma}}, \\
h(\theta) &= \Upsilon(\theta)^{1-\beta} \cdot h_d \cdot \left( \frac{\theta}{\theta_d} \right)^{\frac{\beta(1-k/\delta)}{\gamma}}, \\
w(\theta) &= \Upsilon(\theta)^{1-\beta} \cdot w_d \cdot \left( \frac{\theta}{\theta_d} \right)^{\frac{\beta}{\gamma}},
\end{align*}
\]

\hspace{1cm} \text{(B.13)}

From (B.11), the zero-profit productivity is determined by:

\[
\frac{\Gamma}{1 + \beta \gamma} \kappa_r \left[ c^{-\frac{\beta(1-\gamma k)}{\gamma}} b^{-\beta \gamma} A \theta_d^{\gamma} \right]^{1/T} = f_d.
\]

\hspace{1cm} \text{(B.14)}

Similarly, from (B.11), the exporting productivity cutoff is determined by:

\[
\frac{\Gamma}{1 + \beta \gamma} \kappa_r \left[ c^{-\frac{\beta(1-\gamma k)}{\gamma}} b^{-\beta \gamma} A \theta_x^{\gamma} \right]^{1/T} \left[ \Upsilon_x^{(1-\beta)/T} - 1 \right] = f_x.
\]

\hspace{1cm} \text{(B.15)}

These two conditions imply the following relationship between the productivity cutoffs:

\[
\left[ \Upsilon_x^{(1-\beta)/T} - 1 \right] \left( \frac{\theta_x}{\theta_d} \right)^{\beta/T} = \frac{f_x}{f_d}.
\]

\hspace{1cm} \text{(B.16)}

where in a symmetric equilibrium, \( A = A^* \) and hence \( \Upsilon_x = 1 + \tau^{\frac{\beta}{\gamma}} \). Therefore the ratio of the two productivity cutoffs in a symmetric equilibrium is pinned down by (B.16) alone.

The free entry condition that equates the expected value of entry to the sunk entry cost is:

\[
\int_{\theta_d}^{\infty} \pi(\theta) dG(\theta) = f_e.
\]
where from (B.11) and (B.14)-(B.15), we have:

\[
\pi(\theta) = \pi_d(\theta) + \pi_x(\theta) = f_d \left[ \left( \frac{\theta}{\theta_d} \right)^{\beta/\Gamma} - 1 \right] + I_x(\theta)f_x \left[ \left( \frac{\theta}{\theta_x} \right)^{\beta/\Gamma} - 1 \right],
\]

where \( I_x(\theta) = 1 \) for \( \theta \geq \theta_x \) and \( I_x(\theta) = 0 \) for \( \theta < \theta_x \). Using these relationships we can rewrite the free entry condition as:

\[
f_d \int_{\theta_d}^{\infty} \left[ \left( \frac{\theta}{\theta_d} \right)^{\beta/\Gamma} - 1 \right] dG_\theta + f_x \int_{\theta_x}^{\infty} \left[ \left( \frac{\theta}{\theta_x} \right)^{\beta/\Gamma} - 1 \right] dG_\theta = f_e. \tag{B.17}
\]

The mass of firms within the sector \((M)\) is determined by:

\[
E = M \int_{\theta_d}^{\infty} r_d(\theta) dG_\theta(\theta) + M^* \int_{\theta_x}^{\infty} r_x^*(\theta) dG_\theta(\theta). \tag{B.18}
\]

Using the expressions for equilibrium revenue from domestic sales and exports derived above (see (B.7) and (B.10)), we can rewrite (B.18) as:

\[
E = \frac{1 + \beta \gamma}{\Gamma} \left[ M f_d \int_{\theta_d}^{\infty} \left( \frac{\theta}{\theta_d} \right)^{1-\beta/\Gamma} \left( \frac{\theta}{\theta_d} \right)^{\beta/\Gamma} dG_\theta(\theta) + M^* f_x \frac{1 - \beta \gamma}{\beta \gamma} \frac{\theta}{\theta_x^*} - 1 \int_{\theta_x^*}^{\infty} \left( \frac{\theta}{\theta_x^*} \right)^{1-\beta/\Gamma} dG_\theta(\theta) \right].
\]

The sectoral labor force \((L)\) can be determined from the equality between the total sectoral wage bill and workers’ share of total sectoral revenue:

\[
\omega L = M \int_{\theta_d}^{\infty} w(\theta) h(\theta) dG_\theta(\theta) = M \frac{\beta \gamma}{1 + \beta \gamma} \int_{\theta_d}^{\infty} r(\theta) dG_\theta(\theta), \tag{B.19}
\]

Using the expressions for equilibrium revenue from domestic sales and exports derived above, we can also rewrite (B.19) as:

\[
L = \frac{\beta \gamma}{\Gamma} M \left[ f_d \int_{\theta_d}^{\infty} \left( \frac{\theta}{\theta_d} \right)^{\beta/\Gamma} dG_\theta(\theta) + f_x \int_{\theta_x}^{\infty} \left( \frac{\theta}{\theta_x} \right)^{\beta/\Gamma} dG_\theta(\theta) \right] = z \gamma f_e M,
\]

where we have evaluated the integrals in the square brackets using the Pareto distribution and applied the free entry condition (B.17). This condition implies that \(L/M\) is constant in any equilibrium and \(L\) and \(M\) are equivalent measures of the size of the differentiated sector. Finally, observe that in a symmetric case the expression for \(E\) above can be considerably simplified and becomes identical to the expression for \(L\) above up to a factor of \(\beta \gamma/(1 + \beta \gamma)\).

### B.1.1 Symmetric Countries Closed Form Solutions

In this section, we characterize sectoral equilibrium for the case of symmetric countries, where a number of the expressions above simplify further. Evaluating the integrals in the free entry
condition (B.17) using a Pareto productivity distribution, we obtain:

\[
\left( \frac{\beta}{z\Gamma - \beta} \right) f_d \left( \frac{\theta_{\text{min}}}{\theta_d} \right)^z \left[ 1 + \frac{f_x}{f_d} \left( \frac{\theta_d}{\theta_x} \right)^z \right] = f_e.
\]

Using (B.16), we can rewrite this as

\[
\left( \frac{\beta}{z\Gamma - \beta} \right) f_d \left( \frac{\theta_{\text{min}}}{\theta_d} \right)^z \left[ 1 + \left( \frac{f_d}{f_x} \right)^{\frac{\Gamma - \beta}{\beta}} \left( \frac{f_{\text{e}}}{\theta_d} \right) \Gamma - \beta \right]^{z\Gamma/\beta} = f_e. \tag{B.20}
\]

Note that these two expressions do not themselves rely on symmetry. Under the assumption of a symmetric equilibrium, \( A = A^* \), and hence we have \( Y_x = 1 + \frac{\tau}{\beta/1 - \beta} \). Therefore, in a symmetric equilibrium, (B.20) defines \( \theta_d \), and after solving for \( \theta_d \), \( \theta_x \) can be obtained from (B.16).

Using (B.16) and (B.20), we can now derive the conditions on the parameters that ensure \( \theta_x > \theta_d > \theta_{\text{min}} \) in a symmetric equilibrium. Note that the square bracket in (B.20) is always greater than 1. Therefore, it is enough to require that

\[ f_d > f_e \frac{z\Gamma - \beta}{\beta} \]

To ensure that \( \theta_d > \theta_{\text{min}} \) in any symmetric equilibrium. Therefore a high enough \( f_d \) ensures \( \theta_d > \theta_{\text{min}} \) in a symmetric equilibrium. Next note from (B.16) that since \( \left[ Y_x^{(1-\beta)/\Gamma} - 1 \right] < 1 \), it is enough to require that \( f_x \geq f_d \) to ensure \( \theta_x > \theta_d \) in any symmetric equilibrium. Therefore a high enough \( f_x \) ensures \( \theta_x > \theta_d \) in a symmetric equilibrium. Note that the same condition applies in Melitz (2003). Numerical simulations suggest that a much weaker condition is generally sufficient in this model.

Once we have established the equilibrium value of \( \theta_d \), we can determine the equilibrium demand shifter \( A \) from the domestic productivity cutoff condition (B.14):

\[
A = \left( \frac{1 + \beta\gamma}{\kappa \Gamma} \right)^\Gamma f_d c^{\beta(1-\gamma k)/\delta} b^{\beta} \theta_d^{-\beta}. \tag{B.21}
\]

Note that \( b \) and \( c \) do not affect \( \theta_d \) in the symmetric equilibrium. However, these parameters do affect the demand shifter \( A \). In contrast, trade costs do not alter the relationship between \( A \) and \( \theta_d \) in (B.21), but higher trade costs do reduce \( \theta_d \) and hence increase \( A \) (see (B.20)).

Finally, knowing \( A \) and following the steps outlined in Section 2.2.3, we can solve for sectoral price index \( P \), real consumption index \( Q \) and expenditure \( E = PQ \). Then the mass of firm entrants \( M \) and the measure of workers searching for a job in the differentiated sector \( L \) should satisfy:

\[
\frac{\beta\gamma}{1 + \beta\gamma} E = L = z\gamma f_e M.
\]

This completes the solutions for the case of symmetric countries.
B.2 Derivations and Proofs for Section 2.3

The share of workers employed by firms that serve only the domestic market is from (B.13) and the Pareto productivity distribution:

\[ S_{h,d} = 1 - \frac{\int_{\theta_d}^{\infty} h(\theta) \, dG_\theta(\theta)}{\int_{\theta_d}^{\infty} h(\theta) \, dG_\theta(\theta)} = \frac{1 - \rho^{-\beta(1-k/\delta)}/\Gamma(1-\beta(1-k/\delta))}{1 + \rho^{-\beta(1-k/\delta)} [ \Gamma(1-\beta(1-k/\delta)) - 1]}, \]

(B.22)

where \( \rho \equiv \theta_d/\theta_x \). To compute the distribution of wages across workers employed by non-exporting firms, note that the fraction of workers receiving a particular wage \( w(\theta) \in [w_d, w_d/\rho^{\beta k/\delta}] \) is proportional to \( h(\theta)dG_\theta(\theta) \). In other words, we have:

\[ G_{w,d}(w) = \frac{M \int_{\theta_d}^{\theta_w,d(w)} h(\theta) \, dG_\theta(\theta)}{M \int_{\theta_d}^{\theta_x} h(\theta) \, dG_\theta(\theta)} = 1 - \frac{\int_{\theta_d}^{\theta_w,d(w)} h(\theta) \, dG_\theta(\theta)}{\int_{\theta_d}^{\theta_x} h(\theta) \, dG_\theta(\theta)} \quad \text{for} \quad w \in [w_d, w_d/\rho^{\beta k/\delta}], \]

where \( \theta_{w,d}(\cdot) \) is the inverse of \( w(\cdot) \) and equal to \( \theta_{d}(w/w_d)^{\delta \Gamma/(\beta k)} \). Finally, for \( w < w_d \), \( G_{w,d}(w) = 0 \), and for \( w > w_d/\rho^{\beta k/\delta} \), \( G_{w,d}(w) = 1 \). Using the Pareto productivity distribution, the distribution of wages across workers employed by domestic firms is the following truncated Pareto distribution:

\[ G_{w,d}(w) = \frac{1 - \left( \frac{\theta_d}{\theta_{w,d}(w)} \right) z^{-\beta(1-k/\delta)}}{1 - \left( \frac{\theta_d}{\theta_x} \right) z^{-\beta(1-k/\delta)}} = 1 - \left( \frac{w_d}{w} \right)^{1+1/\mu} \quad \text{for} \quad w \in [w_d, w_d/\rho^{\beta k/\delta}], \]

where \( \mu \equiv \beta k/\delta(z\Gamma - \beta) \).

The distribution of wages across workers employed by exporters can be computed in the same way:

\[ G_{w,x}(w) = \frac{M \int_{\theta_x}^{\theta_{w,x}(w)} h(\theta) \, dG_\theta(\theta)}{M \int_{\theta_x}^{\theta_x} h(\theta) \, dG_\theta(\theta)} = 1 - \frac{\int_{\theta_x}^{\theta_{w,x}(w)} h(\theta) \, dG_\theta(\theta)}{\int_{\theta_x}^{\theta_x} h(\theta) \, dG_\theta(\theta)} \quad \text{for} \quad w \in \left[ w_d \frac{k^{\beta(1-\beta)}}{\Gamma(1-\beta)} / \rho^{\beta k/\delta}, \infty \right), \]

where \( \theta_{w,x}(\cdot) \) is the inverse of \( w(\cdot) \) and equal to \( \theta_{d}(w/w_d)^{\delta \Gamma/(\beta k) \frac{k^{\beta(1-\beta)}}{\Gamma(1-\beta)}} / \rho^{\beta k/\delta} \). Finally, for \( w < w_d \frac{k^{\beta(1-\beta)}}{\Gamma(1-\beta)} / \rho^{\beta k/\delta} \), \( G_{w,x}(w) = 0 \). Using the Pareto productivity distribution, the distribution of wages across workers employed by exporters is the following untruncated Pareto distribution:

\[ G_{w,x}(w) = 1 - \left( \frac{\theta_x}{\theta_{w,x}(w)} \right) z^{-\beta(1-k/\delta)/(1-\beta)} = 1 - \left( \frac{w_d}{w} \frac{k^{\beta(1-\beta)}}{\Gamma(1-\beta)} \rho^{-\beta k/\delta} \right)^{1+1/\mu} \quad \text{for} \quad w \in \left[ w_d \frac{k^{\beta(1-\beta)}}{\Gamma(1-\beta)} / \rho^{\beta k/\delta}, \infty \right). \]

Combining \( S_{h,d} \), \( G_{w,d}(\cdot) \) and \( G_{w,x}(\cdot) \) together we obtain the unconditional wage distribution
among workers employed in the differentiated sector, \(G_w(w)\), as defined in the paper:

\[
G_w(w) = \begin{cases} 
S_{h,d}G_{w,d}(w) & \text{for } w_d \leq w \leq w_d/\rho^{\beta_k}, \\
S_{h,d} & \text{for } w_d/\rho^{\beta_k} \leq w \leq w_dY_x^{\frac{1+\alpha_0}{2}}/\rho^{\beta_k}, \\
S_{h,d} + (1-S_{h,d})G_{w,x}(w) & \text{for } w \geq w_dY_x^{\frac{1+\alpha_0}{2}}/\rho^{\beta_k},
\end{cases} \tag{B.23}
\]

**Proof of Proposition 2.1**

Consider the closed economy wage distribution:

\[
G^a_w(w) = 1 - \left(\frac{w_d}{w}\right)^{1+1/\mu}, \quad \text{for } w \geq w_d.
\]

It is a Pareto distribution with a shape parameter \((1 + 1/\mu) > 2\) and lower bound \(w_d\) defined in (B.13). We now show that with a Pareto distribution a large class of inequality measures depends only on the shape parameter and not on the lower bound of the distribution.

(i) With this distribution, the mean and variance of wages are given by:

\[
\bar{w}^a = (1 + \mu)w_d \quad \text{and} \quad \text{Var}^a(w) = \frac{1 + \mu}{1 - \mu^2}w_d^2.
\]

Recall that we require \(\mu \in (0, 1)\) so that the variance of the wage distribution is finite. Therefore, the coefficient of variation is given by:

\[
CV^a_w = \frac{\sqrt{\text{Var}^a(w)}}{\bar{w}^a} = \frac{\mu}{\sqrt{1 - \mu^2}}.
\]

Clearly, \(\mu\) is a sufficient statistic for the coefficient variation.

(ii) We now characterize the Lorenz Curve for the sectoral wage distribution. In order to do so, we compute the share in sectoral employment and the share in the sectoral wage bill of firms with productivity below \(\theta'\):

\[
s_h(\theta) = \frac{M \int_{\theta_d}^{\theta} h(\theta)dG_\theta(\theta)}{M \int_{\theta_d}^{\infty} h(\theta)dG_\theta(\theta)} = 1 - \left(\frac{\theta_d}{\theta}\right)^{z-\beta(1-k/\delta)/\Gamma},
\]

\[
s_w(\theta) = \frac{M \int_{\theta_d}^{\theta} w(\theta)h(\theta)dG_\theta(\theta)}{M \int_{\theta_d}^{\infty} w(\theta)h(\theta)dG_\theta(\theta)} = 1 - \left(\frac{\theta_d}{\theta}\right)^{z-\beta/\Gamma},
\]

where we have used the solution for firm-level variables (B.13) and the Pareto productivity distribution. From these expressions we can solve for the wage share \(s_w\) as a function of the employment share \(s_h\):

\[
s_w = \mathcal{L}^a(s_h) = 1 - (1-s_h)^{1/(1+\mu)}, \quad \mu = \frac{\beta k}{\delta(\delta \Gamma - \beta)}, \quad s_h, s_w \in [0, 1]. \tag{B.24}
\]

This relationship represents the Lorenz Curve, because workers’ wages are increasing in firm productivity and hence this procedure ranks workers by their wages. Note that \(\mu\) is the only parameter
which determines the position of the Lorenz Curve. A higher $\mu$ makes the Lorenz Curve more convex which implies greater wage inequality.

(iii) The Gini Coefficient is determined uniquely by the shape of the Lorenz Curve and, therefore, $\mu$ is also a sufficient statistic for the Gini Coefficient. Specifically, the Gini Coefficient is defined as:

$$G^a_w = 1 - 2 \int_0^1 L^a(s_h) ds_h = \frac{\mu}{2 + \mu}.$$ 

(iv) The Theil Index of wage inequality is defined as:

$$T^a_w = \int_{w_d}^{\infty} \frac{w}{w^a} \ln \left( \frac{w}{w^a} \right) dG^a_w(w),$$

where the properties of the Theil Index are discussed in Bourguignon (1979). Using the autarky wage distribution, we can compute this integral to obtain:

$$T^a_w = \mu - \ln(1 + \mu).$$

Note that $\mu$ is a sufficient statistic for the Theil Index, which is decreasing in $\mu$ and equals to zero when $\mu = 0$.

Note that all discussed measures of inequality depend on the parameters of the model only through their effect on $\mu$; all measures of inequality are increasing in $\mu$. This completes the proof of Proposition 2.1.

**Proof of Proposition 2.2**

From Proposition 1 we know that $\mu$ is a sufficient statistic for sectoral wage inequality in the closed economy. Recall that

$$\mu = \frac{\beta k}{\delta(z\Gamma - \beta)}, \quad \Gamma = 1 - \beta \gamma - \frac{\beta}{\delta} (1 - \gamma k).$$

Evidently, $\partial \mu / \partial z < 0$. The effect of $k$ is more subtle as it increases both the numerator and denominator in the expression for $\mu$. Taking the derivative with respect to $k$ and rearranging we obtain:

$$\text{sign} \left\{ \frac{\partial \mu}{\partial k} \right\} = \text{sign} \{1 - z\gamma \mu \} = \text{sign} \{\beta^{-1} - \gamma - \delta^{-1} - z^{-1}\}.$$ 

This finishes the proof of Proposition 2.2.

**Proof of Proposition 2.3**

We first prove the second part of the proposition. The two limiting cases for the open economy wage distribution (B.23) are (1) autarky (as $\rho = \theta_d / \theta_x \to 0$ and no firms export) and (2) when all firms export ($\rho \to 1$). In the first case, the wage distribution $G^a_w(w)$ is an untruncated Pareto with shape parameter $(1 + 1/\mu)$ and lower bound $w_d$. In the second case, the wage distribution is again an untruncated Pareto with shape parameter $(1 + 1/\mu)$, but now with a higher lower bound given by $w_d \Upsilon_x^{k(1-\beta)/(\delta \Gamma)}$. However, from Proposition 1, a broad class of inequality measures depend
only on the shape parameter of the Pareto distribution and do not depend on its lower bound. Therefore, wage inequality is the same in autarky and when all firms export.

Now consider the first part of the proposition. Define the following notation:

\[ \eta_1 \equiv \Upsilon(1-\beta)(1-k/\delta) - 1, \quad \eta_2 \equiv \Upsilon^{1-\beta} - 1, \]
\[ \vartheta_1 \equiv z - \beta(1-k/\delta) \Gamma x - 1, \quad \vartheta_2 \equiv z - \delta. \]

Using this notation, the lowest wage paid by exporters and the highest wage paid by domestic firms can be written as:

\[ w(\theta^+_x) = w(\theta^-_x) \frac{1+\eta_2}{1+\eta_1} \quad \text{and} \quad w(\theta^-_x) = w_d \rho^{\delta_2 - \delta_1}. \]

Similarly, using this notation, the actual wage distribution (B.23) can be written as:

\[ G_w(w) = \begin{cases} 
\frac{1}{1+\eta_1 \rho^{\delta_1}} \left[ 1 - \left( \frac{w_d}{w} \right)^{1+1/\mu} \right], & w_d \leq w \leq w(\theta^-_x), \\
(1-\rho^{\delta_1})/(1+\eta_1 \rho^{\delta_1}), & w(\theta^-_x) \leq w \leq w(\theta^+_x), \\
\frac{1-\rho^{\delta_1}}{1+\eta_1 \rho^{\delta_1}} + \frac{1}{1+\eta_1 \rho^{\delta_1}} \left[ 1 - \left( \frac{w(\theta^+_x)/w}{1+1/\mu} \right) \right], & w \geq w(\theta^+_x) 
\end{cases} \tag{B.25} \]

and the mean of this distribution can be written as

\[ \bar{w} = (1+\mu) w_d \frac{1+ \eta_2 \rho^{\delta_2}}{1+ \eta_1 \rho^{\delta_1}}. \]

The counterfactual wage distribution is defined as:

\[ G_w^c(w) = 1 - \left( \frac{w_d}{w} \right)^{1+1/\mu}, \quad w \geq w_d^c, \tag{B.26} \]

where, in order for the mean of the counterfactual distribution to equal \( \bar{w} \), its lower limit must satisfy:

\[ w_d^c = \frac{1+ \eta_2 \rho^{\delta_2}}{1+ \eta_1 \rho^{\delta_1}} w_d. \]

Therefore we can establish the following result:

\[ w_d^c > w_d \quad \text{since} \quad \frac{1+ \eta_2 \rho^{\delta_2}}{1+ \eta_1 \rho^{\delta_1}} > 1 \quad \text{for} \quad 0 < \rho < 1, \]

as \( 0 < \eta_1 < \eta_2 < 1 \) and \( 1 < z/2 < \vartheta_2 < \vartheta_1 < z \). Similarly, we can establish:

\[ w_d^c < w(\theta^+_x) \quad \text{since} \quad \frac{1+ \eta_2 \rho^{\delta_2}}{1+ \eta_1 \rho^{\delta_1}} < \frac{1+ \eta_2 \rho^{\delta_2}}{1+ \eta_1 \rho^{\delta_1}} \rho^{\delta_1 - \delta} = \frac{(1+ \eta_2 \rho^{\delta_2}}{(1+ \eta_1 \rho^{\delta_1})}, \]

with the inequality being satisfied since:

\[ \frac{1+ \eta_2 \rho^{\delta_2}}{1+ \eta_1 \rho^{\delta_1}} = \frac{(1+ \eta_2 \rho^{\delta_2}}{(1+ \eta_1 \rho^{\delta_1})} + (1- \rho^{\delta_2}) = \frac{(1+ \eta_2 \rho^{\delta_2}}{(1+ \eta_1 \rho^{\delta_1})} + \frac{1- \rho^{\delta_2}}{(1+ \eta_1 \rho^{\delta_1})} < \frac{(1+ \eta_2 \rho^{\delta_2}}{(1+ \eta_1 \rho^{\delta_1})}. \]
as $\rho^{\phi_1} < \rho^{\phi_2} < 1$ and $(1 + \eta_2) > (1 + \eta_1)$. Note that in general we can have either $w_d^c > w(\theta_x^-)$ or $w_d^c < w(\theta_x^-)$, but the same arguments apply in each case.

We can also show that the slope of the counterfactual wage distribution is smaller than the slope of the actual wage distribution at $w(\theta_x^+)$: $g_w(w(\theta_x^+)) > g_w^c(w(\theta_x^+))$. Since the truncations of $G_w(w)$ and $G_w^c(w)$ at $w(\theta_x^+)$ are both Pareto with shape parameter $(1 + 1/\mu)$, we can show that $g_w(w(\theta_x^+)) > g_w^c(w(\theta_x^+))$ by establishing that $1 - G_w(w(\theta_x^+)) > 1 - G_w^c(w(\theta_x^+))$. From (B.25) and (B.26), this implies:

$$1 - \frac{1 - \rho^{\phi_1}}{1 + \eta_1 \rho^{\phi_1}} > \left(\frac{w_d^c}{w(\theta_x^+)}\right)^{1+1/\mu} \Leftrightarrow \frac{(1 + \eta_1) \rho^{\phi_1}}{1 + \eta_1 \rho^{\phi_1}} > \left(\frac{1 + \eta_2 \rho^{\phi_2} (1 + \eta_1)}{1 + \eta_2 \rho^{\phi_2} (1 + \eta_1)}\right)^{1/2} \rho^{\phi_1}$$

$$\Leftrightarrow \phi(\rho) \equiv \left(\frac{1 + \eta_1 \rho^{\phi_1}}{1 + \eta_1} - \frac{1 + \eta_2 \rho^{\phi_2}}{1 + \eta_2}\right) > 0.$$ 

To show that $\phi(\rho) > 0$ for all $\rho \in [0, 1)$, note that:

$$\phi(0) \equiv \left(\frac{1}{1 + \eta_1} - \frac{1}{1 + \eta_2}\right) > 0,$$

as $\eta_1 < \eta_2$ and $\phi_1 > \phi_2$. Note also that $\phi(1) = 1 - 1 = 0$. Consider now the derivative of $\phi(\rho)$ for $\rho \in (0, 1)$:

$$\phi'(\rho) = \frac{\eta_1 \eta_2}{\rho} \left[\frac{1 + \eta_1 \rho^{\phi_1}}{1 + \eta_1} \rho^{\phi_2} - \frac{1 + \eta_2 \rho^{\phi_2}}{1 + \eta_2} \rho^{\phi_1}\right].$$

Note that

$$\frac{\eta_2 \rho^{\phi_2}}{1 + \eta_2 \rho^{\phi_2}} > \frac{\eta_1 \rho^{\phi_1}}{1 + \eta_1 \rho^{\phi_1}},$$

since $\eta_1 < \eta_2$ and $\rho^{\phi_1} < \rho^{\phi_2}$. As a result, whenever $\phi(\rho) \leq 0$, we also necessarily have $\phi'(\rho) < 0$. Therefore, if there exists $\rho'$ such that $\phi(\rho') = 0$, then $\phi(\rho) < 0$ for all $\rho > \rho'$. But since $\phi(1) = 0$, this implies that $\phi(\rho) > 0$ for all $\rho \in (0, 1)$.

We now establish that $G_w^c(w)$ second-order stochastically dominates $G_w(w)$ for $\rho \in (0, 1)$. Using the fact that the truncations of $G_w(w)$ and $G_w^c(w)$ at $w(\theta_x^+)$ are both Pareto with shape parameter $(1 + 1/\mu)$, and the result above that $g_w(w(\theta_x^+)) > g_w^c(w(\theta_x^+))$, we know that this inequality holds for all $w > w(\theta_x^+)$. We have two cases:

1. $w(\theta_x^-) \leq w_d^c < w(\theta_x^+)$:

$$g_w(w) - g_w^c(w) = \begin{cases} > 0, & w_d \leq w < w(\theta_x^-), \\ = 0, & w(\theta_x^-) \leq w < w_d^c, \\ < 0, & w_d^c \leq w < w(\theta_x^+), \\ > 0, & w \geq w(\theta_x^+). \end{cases}$$
2. \( w_d^c < w(\theta^-) \):

\[
g_w(w) - g_w^c(w) = \begin{cases} 
> 0, & w_d \leq w < w_d^c, \\
\leq 0, & w_d^c \leq w < w(\theta^-), \\
< 0, & w(\theta^-) \leq w < w(\theta^+_x), \\
> 0, & w \geq w(\theta^+_x). 
\end{cases}
\]

Importantly, \( g_w(w) - g_w^c(w) \) takes either only positive or only negative values in the range \([w_d^c, w(\theta^-)]\), since for this range

\[
g_w(w) - g_w^c(w) = (C - C^c) w^{-(2+1/\mu)}
\]

where \( C \) and \( C^c \) are positive constants.

Note that in both cases the above characterization of \( g_w(w) - g_w^c(w) \) implies that this difference of density functions is positive for low values of \( w \), negative for intermediate values of \( w \), and again positive for larger values of \( w \). This immediately implies that the cumulative distribution functions intersect only once in the range where the difference of density functions is negative (see Figure 2 in the text), which is a sufficient condition to establish that indeed \( G_w^c(w) \) second-order stochastically dominates \( G_w(w) \) (see, for example, Mas-Colell, Whinston and Green 1995, p.195).

Therefore, for all measures of inequality that respect second-order stochastic dominance, wage inequality in the open economy when some but not all firms export is strictly greater than in the closed economy. This finishes the proof of Proposition 2.3.

B.3 Derivations and Proofs for Section 2.4

The sectoral unemployment rate is given by

\[
u = \frac{L - H}{L} = 1 - \frac{H N}{N L} = 1 - \sigma x,
\]

where \( \sigma = H/N \) is the hiring rate and \( x = N/L \) is labor market tightness. From (2.14), labor market tightness is given by \( x = (\omega/\alpha_0)^{1/(1+\alpha_1)} \).

We now derive the expression for the sectoral hiring rate:

\[
s = M \int_{\theta_d}^{\infty} h(\theta) dG_\theta(\theta) = \frac{h_d \int_{\theta_d}^{\theta^c} \left( \frac{\theta}{\theta_d} \right)^{\beta(1-k/\delta)} dG_\theta(\theta) + h_x \int_{\theta_d}^{\infty} \left( \frac{\theta}{\theta_d} \right)^{\beta/\Gamma} dG_\theta(\theta) + n_x \int_{\theta^c}^{\infty} \left( \frac{\theta}{\theta_d} \right)^{\beta/\Gamma} dG_\theta(\theta)} {n_d \int_{\theta_d}^{\theta^c} \left( \frac{\theta}{\theta_d} \right)^{\beta/\Gamma} dG_\theta(\theta) + n_x \int_{\theta_d}^{\infty} \left( \frac{\theta}{\theta_d} \right)^{\beta/\Gamma} dG_\theta(\theta)}
\]

where \( h_x = \Upsilon_x \left( 1 - \beta \right) \left( 1 - k/\delta \right) \left( \theta_x/\theta_d \right)^{\beta(1-k/\delta)} h_d \), \( n_x = \Upsilon_x^{1-\beta} \left( \theta_x/\theta_d \right)^{\beta} n_d \) and we have used the solution for firm-specific variables (B.13). Evaluating the integrals using the Pareto productivity distribution yields

\[s = s^a \cdot \varphi(\rho, \Upsilon_x),\]
where
\[ \sigma^a \equiv \frac{1}{1 + \mu} \left[ \frac{\Gamma}{\beta(1 - \gamma k)} \frac{ca_{min}^{\delta}}{f_d} \right]^{k/\delta} \]
is the autarky hiring rate and
\[ \varphi(\rho, \Upsilon_x) \equiv \frac{1 + \left[ \Upsilon_x^{(1 - \beta)(1 - k/\delta)} - 1 \right] \rho^{\gamma - \beta(1 - k/\delta)/\Gamma}}{1 + \left[ \Upsilon_x^{(1 - \beta)} - 1 \right] \rho^{\gamma - \beta/\Gamma}} = 1 + \eta_1 \rho^{\beta_1} \frac{1}{1 + \eta_2 \rho^{\beta_2}}, \]
where the last equality uses the notation introduced in Section B.2 of this appendix. Note that \( \varphi(0, \Upsilon_x) = 1 \) and \( \varphi(\rho, \Upsilon_x) < 1 \) for \( 0 < \rho \leq 1 \) since \( \eta_1 < \eta_2 \) and \( \beta_1 > \beta_2 \). Therefore, the hiring rate is lower in any open economy equilibrium than in autarky.

Opening up to trade does not affect sectoral labor market tightness \( \omega \) directly, but does affect it indirectly through \( \omega \). In case when \( \omega \) is invariant to trade, sectoral unemployment is affected by trade only through hiring rate. As a result, in this case sectoral unemployment is strictly higher in any open economy equilibrium than in autarky. This constitutes a proof of Proposition 2.5.

Finally, we look at the determinants of the hiring rate and sectoral unemployment in the autarky equilibrium. From the expression for \( u, x \), and \( \sigma^a \), it is evident that holding \( \omega \) constant \( c, z \) and \( k \) affect the unemployment rate only through their effects on the hiring rate \( \sigma^a \), while \( \alpha_0 \) affects the unemployment rate only through its effect on labor market tightness \( x \). Specifically, \( \alpha_0 \) decreases \( x \) and increases \( u; z \) and \( c \) increase \( \sigma^a \) and decrease \( u; \) the effect of \( k \) on \( \sigma^a \) and hence \( u \) is ambiguous. This constitutes a proof of Proposition 2.4.

### B.4 Derivations and Proofs for Section 2.5

First, consider the Lorenz Curve for the sectoral income distribution. When the Lorenz Curve for the wage distribution is given by \( s_w = L_w(s_h) \) and the unemployment rate is \( u \), the Lorenz Curve for the income distribution can be written as:
\[ s_t = L_t(s_t) = \begin{cases} 
0, & 0 \leq s_t \leq u, \\
L_w \left( \frac{s_t - u}{1 - u} \right), & u \leq s_t \leq 1 
\end{cases} \]
and from (B.24) \( L_w(s) = 1 - (1 - s)^{1/(1 + \mu)} \). We can now compute the Gini Coefficient of sectoral income inequality:
\[ G_t = 1 - 2 \int_0^1 L_t(s_t) ds_t = u + (1 - u)G_w = u + (1 - u) \left[ 1 - \frac{\int_0^1 L_w \left( \frac{s_t - u}{1 - u} \right) ds_t}{1 - u} \right]. \]
Substitution of variables \( s_h = (s_t - u)/(1 - u) \) results in
\[ G_t = u + (1 - u)G_w, \]
where \( G_w \) is the Gini Coefficient of sectoral wage inequality as stated in the text of the paper. This is the expression we state in the text. Therefore, sectoral income inequality as measured by the
Gini coefficient is increasing in the sectoral unemployment rate and sectoral wage inequality.

Now consider the Theil Index of sectoral income inequality. The general definition of the Theil index is

$$T_ι = \int \frac{ϖ}{\bar{ϖ}} \ln \frac{ϖ}{\bar{ϖ}} dG_ϖ(ϖ),$$

where $ϖ$ is a measure of income distributed according to the cumulative distribution function $G_ϖ(ϖ)$ and $\bar{ϖ}$ is the mean of this measure. Note that $ϖ \ln ϖ = 0$ at $ϖ = 0$. Then for the sectoral income distribution we have $G_ι(ι) = u$ for $ι \in [0, w_d)$ and $G_ι(ι) = u + (1 - u)G_w(ι)$ for $ι > w_d$, where $G_w(w)$ is the sectoral wage distribution. The mean of the sectoral income distribution is $\bar{ι} = (1 - u)\bar{w} + u \cdot 0 = (1 - u)\bar{w}$, where $\bar{w}$ is the mean of the sectoral wage distribution. We can now compute the Theil Index of sectoral income inequality:

$$T_ι = u \cdot 0 + (1 - u) \int_{w_d}^{∞} \frac{w}{ι} \ln \frac{w}{ι} dG_w(w) = \frac{(1 - u)\bar{w}}{ι} \int_{w_d}^{∞} \frac{w}{\bar{w}} \left[ \ln \frac{w}{\bar{w}} - \ln \frac{ι}{\bar{w}} \right] dG_w(w).$$

Since $ι/\bar{w} = (1 - u)$, we can rewrite this as:

$$T_ι = T_w - \ln(1 - u),$$

where $T_w$ is the Theil Index of sectoral wage inequality as stated in the text of the paper. The same expression can be derived using the decomposition of the Theil Index into within and between-group components, as studied in Bourguignon (1979) and applied to this setting in Helpman, Itskhoki and Redding (2008a,b). Finally, note that sectoral income inequality as measured by Theil Index is increasing in the sectoral unemployment rate and sectoral wage inequality, consistent with the results for the Gini Coefficient above. Combining these results with the Propositions in Sections 2.3 and 2.4 we arrive at Propositions 2.6 and 2.7.

B.5 Derivations and Proofs for Section 2.6

Outside Sector and Risk Neutrality

As noted in the paper, the Harris-Todaro condition that equates *ex ante* expected indirect utility in the two sectors takes the following form:

$$V = \frac{1}{1 - \eta} \mathbb{E} \left( \frac{w}{P} \right)^{1 - \eta} = \frac{Ew}{P} = \frac{1}{P}, \text{ for } η = 0,$$

which, given the probability of matching $x$ and the expected wage conditional on matching of $w(θ)h(θ)/h(θ) = b$, implies:

$$xb = \omega = 1.$$  

From this expression, a sufficient condition for $α_0 > \omega$, and hence $0 < x < 1$ in (2.14), is $α_0 > 1$.

*Ex post* indirect utility depends on the aggregate price index ($P$) and *ex post* wages, which
depend on a worker’s sector and firm of employment:

\[
V = \begin{cases} 
1/\mathcal{P} & \text{if employed in the outside sector} \\
 w(\theta)/\mathcal{P} & \text{if employed by a } \theta\text{-firm in the differentiated sector} \\
0 & \text{if unemployed} 
\end{cases} \tag{B.28}
\]

where \(\mathcal{P}\) is the dual of the aggregate consumption index (\(C\)):

\[
\mathcal{P} = \left[\vartheta + (1 - \vartheta) \frac{P_{\text{diff}}}{\mathcal{P}} - \zeta\right] - 1 - \zeta, \tag{B.29}
\]

and depends on the differentiated sector price index (\(P\)), which can be determined from the expression for the demand-shifter (\(A\)) with CES preferences:

\[
A^{\frac{1}{1-\beta}} = \frac{(1 - \vartheta) \frac{P_{\text{diff}}}{\mathcal{P}} - \zeta}{\vartheta + (1 - \vartheta) \mathcal{P}^{-\frac{1}{1-\beta}}} \tag{B.30}
\]

where we have used \(\Omega = \bar{L}\). From the solutions for firm-specific variables (B.13), \textit{ex post} wages in the differentiated sector in the closed and open economies are:

\[
w^a(\theta) = w_d(\frac{\theta}{\bar{\theta}_d})^{-\frac{\beta k}{\gamma}} \text{ for } \theta \geq \bar{\theta}_d, \\
w^t(\theta) = \begin{cases} 
w_d(\frac{\theta}{\bar{\theta}_d})^{-\frac{\beta k}{\gamma}} \text{ for } \theta \in [\theta_d, \theta_t] \\
\Upsilon x^{-\beta} w_d(\frac{\theta}{\bar{\theta}_d})^{-\frac{\beta k}{\gamma}} \text{ for } \theta \geq \theta_t 
\end{cases}
\]

where the superscripts \(a\) and \(t\) denote closed and open economy variables respectively.

To characterize the impact of the opening of trade on \textit{ex post} wages at domestic and exporting firms, note that \(w^t(\theta) < w^a(\theta)\) for \(\theta \in [\theta_t, \theta_x] \) because \(\theta_t > \theta_d\). Note also that \(w^t(\theta) > w^a(\theta)\) for \(\theta \geq \theta_x\) whenever \(\Upsilon x^{-\beta} > (\theta_d/\theta_t)^{\beta}\). To show that this inequality must be satisfied, note that the free entry conditions in the closed and open economy (B.20) together imply:

\[
\left(\frac{\theta_t}{\theta_d}\right)^{\frac{\beta}{\gamma}} = \left[1 + \left(\frac{f_d}{f_x}\right)^{\frac{z}{1-\beta}} \frac{\Upsilon^{\frac{1-\beta}{\beta}}}{x^{\frac{1-\beta}{\beta}} - 1}\right]^{\frac{1-\beta}{\beta}}.
\]

We can rewrite this condition as:

\[
\left(\frac{\theta_t}{\theta_d}\right)^{\frac{\beta}{\gamma}} - 1 = \left[\Upsilon^{\frac{1-\beta}{\beta}} - 1\right]^{\frac{z}{1-\beta}} \left(\frac{f_d}{f_x}\right)^{\frac{z}{1-\beta}} \left[\Upsilon^{\frac{1-\beta}{\beta}} - 1\right]^{\frac{z}{1-\beta}} - 1,
\]

where the above inequality holds because \(\Upsilon x > 1\), \(z\Gamma > \beta\) and \(f_d/f_x < 1\). This immediately implies \((\theta_d/\theta_t)^{\beta} < \Upsilon^{\frac{1-\beta}{\beta}}\).

Proposition 2.8 can be proved as follows. The Harris-Todaro condition under risk neutrality implies \(xb = \omega = 1\), which together with the search cost \(b = \alpha_0 x^{\alpha_1}\) implies that search costs and labor market tightness are invariant to trade: \(b = \alpha_0^{1/(1+\alpha_1)}\) and \(x = \alpha_0^{-1/(1+\alpha_1)}\). From the free
entry condition (B.20), \( \theta_d \) is higher in the open economy than in the closed economy. From the zero-profit productivity cutoff condition (B.14), a higher value of \( \theta_d \) and an unchanged value of \( b \) implies a lower value of \( A \) in the open economy than in the closed economy. From (B.30) above, a lower value of \( A \) implies a lower value of \( P \) in the open economy than in the closed economy. From (B.29), this reduction in \( P \) in turn implies a lower value of \( \mathcal{P} \) and higher \textit{ex ante} expected welfare (B.27) in the open economy than in the closed economy, which establishes part (i) of the proposition. Parts (ii)-(iii) of the proposition follow immediately from the lower value of \( \mathcal{P} \) in the open economy than in the closed economy, the changes in \textit{ex post} wages and welfare (established in (B.28) above), and the analysis of unemployment in Section B.3 (noting that in this general equilibrium environment \( \omega \) is invariant to trade).

**Single Differentiated Sector and Risk Neutrality**

With a single differentiated sector, the conditions for sectoral equilibrium remain the same. With a single The expression for the demand shifter becomes:

\[
A \equiv Q^{1-\beta} P = Q^{1-\beta},
\]

(B.31)

where we have used the choice of numeraire: \( P = 1 \). Therefore, total revenues are \( E = Q \).

The solutions for \((\theta_d, \theta_x)\) under the assumption of country symmetry follow from (B.20) and (B.16), as discussed above. Using symmetric countries, the equilibrium labor force (B.19) and labor market clearing \((L = \bar{L})\), we obtain:

\[
\frac{\beta \gamma}{1 + \beta \gamma} Q = \bar{L} \omega.
\]

Using the search cost \( b = \alpha_0 x^{\alpha_1} \), expected worker income \( bx = \omega \) and the zero-profit cutoff condition (B.14), we obtain:

\[
Q = \left( \frac{f_d}{\kappa r} \right) \frac{\Gamma}{\Gamma - \beta} c^{\frac{\beta (1-\gamma)}{\gamma}} \theta_d^{-\frac{\beta \gamma}{1 + \alpha_1}} \omega^{-\frac{\beta}{1 + \alpha_1}} \alpha_0^{\frac{\beta \gamma}{1 + \alpha_1}} \theta_x^{-\frac{\beta \gamma}{1 + \alpha_1}} \omega^{-\frac{\beta \gamma}{1 + \alpha_1}}
\]

These equations define two upward-sloping relationships in \((Q, \omega)\) space that determine the equilibrium values of \( Q \) and \( \omega \). One can show that a necessary and sufficient condition for the stability of the equilibrium is:

\[
\frac{\beta \gamma}{1 - \beta} \frac{\alpha_1}{1 + \alpha_1} > 1,
\]

(B.32)

which is satisfied for sufficiently large \( \alpha_1 \) (sufficiently convex hiring costs) and sufficiently large \( \beta \) (a sufficiently high elasticity of substitution between varieties). As \( 0 < \alpha_1 / (1 + \alpha_1) < 1 \), a necessary condition for this parameter restriction to hold is:

\[
\frac{\beta \gamma}{1 - \beta} > 1,
\]

which is satisfied for values of \( \beta \) sufficiently close to but less than 1. Assuming parameter values satisfying these inequalities, the conditions for sectoral equilibrium can be solved to yield the
following solutions for the other endogenous variables of the model:

\[
Q = Q^* = \frac{1 + \beta \gamma}{\beta \gamma} \alpha_0^{\frac{\beta(1 - \gamma)(1 + \alpha_1)}{\beta(1 - \gamma)(1 + \alpha_1) - \beta \gamma \alpha_1}} e^{\frac{\beta(1 + \alpha_1)}{\beta(1 - \gamma)(1 + \alpha_1) - \beta \gamma \alpha_1} \theta_d \frac{\beta(1 + \alpha_1)}{\beta(1 - \gamma)(1 + \alpha_1) - \beta \gamma \alpha_1} L^{\frac{1 + \alpha_1}{\beta(1 - \gamma)(1 + \alpha_1) - \beta \gamma \alpha_1}},
\]

\[
\omega = \omega^* = \alpha_0^{\frac{\beta(1 - \gamma)(1 + \alpha_1)}{\beta(1 - \gamma)(1 + \alpha_1) - \beta \gamma \alpha_1}} e^{\frac{\beta(1 + \alpha_1)}{\beta(1 - \gamma)(1 + \alpha_1) - \beta \gamma \alpha_1} \theta_d \frac{\beta(1 + \alpha_1)}{\beta(1 - \gamma)(1 + \alpha_1) - \beta \gamma \alpha_1} L^{\frac{1 + \alpha_1}{\beta(1 - \gamma)(1 + \alpha_1) - \beta \gamma \alpha_1}},
\]

\[
b = b^* = \alpha_0^{\frac{\beta(1 - \gamma)(1 + \alpha_1)}{\beta(1 - \gamma)(1 + \alpha_1) - \beta \gamma \alpha_1}} e^{\frac{\beta(1 + \alpha_1)}{\beta(1 - \gamma)(1 + \alpha_1) - \beta \gamma \alpha_1} \theta_d \frac{\beta(1 + \alpha_1)}{\beta(1 - \gamma)(1 + \alpha_1) - \beta \gamma \alpha_1} L^{\frac{1 + \alpha_1}{\beta(1 - \gamma)(1 + \alpha_1) - \beta \gamma \alpha_1}},
\]

\[
x = x^* = \alpha_0^{\frac{\beta(1 - \gamma)(1 + \alpha_1)}{\beta(1 - \gamma)(1 + \alpha_1) - \beta \gamma \alpha_1}} e^{\frac{\beta(1 + \alpha_1)}{\beta(1 - \gamma)(1 + \alpha_1) - \beta \gamma \alpha_1} \theta_d \frac{\beta(1 + \alpha_1)}{\beta(1 - \gamma)(1 + \alpha_1) - \beta \gamma \alpha_1} L^{\frac{1 + \alpha_1}{\beta(1 - \gamma)(1 + \alpha_1) - \beta \gamma \alpha_1}},
\]

\[
M = M^* = \left( \frac{2 \Gamma - \beta}{z} \right) \theta_d^z \alpha_0^{\frac{\beta(1 - \gamma)(1 + \alpha_1)}{\beta(1 - \gamma)(1 + \alpha_1) - \beta \gamma \alpha_1}} e^{\frac{\beta(1 + \alpha_1)}{\beta(1 - \gamma)(1 + \alpha_1) - \beta \gamma \alpha_1} \theta_d \frac{\beta(1 + \alpha_1)}{\beta(1 - \gamma)(1 + \alpha_1) - \beta \gamma \alpha_1} L^{\frac{1 + \alpha_1}{\beta(1 - \gamma)(1 + \alpha_1) - \beta \gamma \alpha_1}},
\]

where

\[
\Delta \equiv \frac{(1 - \beta)(1 + \alpha_1) - \beta \gamma \alpha_1}{(1 - \beta)} < 0,
\]

\[
\kappa_b \equiv \frac{\beta \gamma}{(1 + \beta \gamma)} \left[ \frac{\theta_d (1 + \beta \gamma)}{\kappa_r \Gamma} \right]^{\frac{1}{1 - \beta}} > 0.
\]

From the above, a sufficient condition for \( \alpha_0 > \omega \), and hence \( 0 < x < 1 \) in (2.14), is:

\[
\frac{(1 + \alpha_1)(1 - \beta)}{(1 + \alpha_1)(1 - \beta) - \beta \gamma \alpha_1} > \frac{1}{\alpha_0^{\frac{\beta(1 - \gamma)(1 + \alpha_1)}{\beta(1 - \gamma)(1 + \alpha_1) - \beta \gamma \alpha_1}}} e^{\frac{\beta(1 + \alpha_1)}{\beta(1 - \gamma)(1 + \alpha_1) - \beta \gamma \alpha_1} \theta_d \frac{\beta(1 + \alpha_1)}{\beta(1 - \gamma)(1 + \alpha_1) - \beta \gamma \alpha_1} L^{\frac{1 + \alpha_1}{\beta(1 - \gamma)(1 + \alpha_1) - \beta \gamma \alpha_1}}\]

Finally, under the assumption of risk neutrality, expected indirect utility is:

\[
\mathbb{V} = \mathbb{E} \left( \frac{w}{\mathbb{F}} \right) = \omega \quad \text{for} \quad \eta = 0.
\]

(B.34)

The proof of Proposition 2.9 follows immediately from the closed form solutions for \((\theta_d, \theta_x)\) in (B.20) and (B.16) and from the closed form solutions for \((Q, \omega, b, x, M, \mathbb{V})\) in (B.33) and (B.34).

**Outside Sector and Risk Aversion**

We first show how the introduction of worker risk aversion affects the equilibrium share of revenue received by workers in the bargaining game. Specifically, with CRRA-CES preferences, the solution to the bargaining game under risk aversion takes a similar form as when there are differences in bargaining weight between the firm and its workers. Given workers’ revenue share, we next show that the determination of sectoral equilibrium remains unchanged except for the Harris-Todaro condition equating expected utility in the two sectors.

**Stole and Zweibel Bargaining**

The bargaining game takes the same form as discussed in the paper for risk neutrality, except that workers are assumed to be risk averse \((0 < \eta < 1)\), and instead of assuming equal bargaining weights we allow for differences in bargaining weight between the firm and its workers. The bargaining power of workers is denoted by \(\lambda \in (0,1)\) and their outside option of unemployment involves zero income.
We start with the case of a discrete number of workers $h$ and then take the limit of continuous divisibility of the workforce. Denote by $R(h)$ the revenue function when a firm employs $h$ workers and denote by $w(h)$ the equilibrium wage schedule that the firm pays to each worker when the employment level is $h$. Now consider that the firm is separated with $\delta > 0$ workers (the special case is $\delta = 1$ and we will later take the limit $\delta \to 0$). This will reduce revenue to $R(h-\delta)$ and the wage rate to $w(h-\delta)$.

Now consider the firm’s bargaining with the marginal $\delta$ workers when employment is $h$. Denote by $t$ the wage that the firm pays to these $\delta$ workers, while other workers receive $w(h)$. The utility of these marginal workers is $t\gamma - \eta/(1-\eta)$. If the bargaining breaks down and they leave, the revenues will fall by $R(h) - R(h-\delta)$ and the firm will be paying the remaining $h-\delta$ workers the wage rate $w(h-\delta)$. Therefore, the incremental payoff to the firm from employing these marginal $\delta$ workers is:

$$\left[R(h) - R(h-\delta)\right] - (h-\delta)\left[w(h) - w(h-\delta)\right] - \delta t.$$

Assuming zero outside option for workers, the payoff from employment to each of the marginal workers is equal to $t\gamma - \eta/(1-\eta)$. Therefore, we can write Stole and Zweibel Bargaining solution as:

$$t(h) = \arg \max_t \left(\left[R(h) - R(h-\delta)\right] - (h-\delta)\left[w(h) - w(h-\delta)\right] - \delta t\right)^{1-\lambda} \left(\frac{1}{\lambda(1-\eta)}\right)\lambda$$

The equilibrium requirement is that $t(h) \equiv w(h)$. That is, for every employment level $h$, as a result of bargaining with marginal workers the firm pays them accordingly to the equilibrium wage schedule.

Denote by

$$\phi \equiv \frac{1 - \lambda}{\lambda(1-\eta)}$$

the effective relative bargaining weight of the firm. Then we can implicitly write the bargaining solution, taking into account the equilibrium requirement, as:

$$(1 + \phi)\delta w(h) = \left[R(h) - R(h-\delta)\right] - (h-\delta)\left[w(h) - w(h-\delta)\right].$$

This holds for every $h > 0$. We now take the limit as $\delta \to 0$ to obtain the differential equation for the wage schedule:

$$(1 + \phi)w(h) = R'(h) - w'(h)h \quad \Rightarrow \quad w(h) = \frac{1}{h^{1+\phi}} \int_0^h (R')^\phi \, dh.$$

This is the generalized Stole and Zweibel (1996a,b) condition for the case of asymmetric bargaining weights. Therefore risk aversion is equivalent to an adjustment in the bargaining weights (greater risk aversion of workers reduces their bargaining weight). When revenue is a power function of employment with power $\beta\gamma$ ($R(h) = Ah^{\beta\gamma}$), the solution to this differential equation is given by:

$$w(h) = \frac{\beta\gamma}{\phi + \beta\gamma} \frac{R(h)}{h}.$$
Note that the wage rate is again a constant fraction of average revenues. The fraction of revenues accruing to workers is increasing in $\beta \gamma$ (decreasing in the concavity of revenues) and decreasing in $\phi$ (the relative effective bargaining power of the firm). In turn, $\phi$ is decreasing in $\lambda$ (the primitive bargaining power of workers) and increasing in $\eta$ (the degree of risk aversion of workers). The case of symmetric bargaining weights and no risk aversion is a special case of this more general formulation. Specifically, when $\lambda = 1$ and $\eta = 0$, we have $\phi = 1$ and arrive at our baseline formulation in the paper.

**Sectoral Equilibrium**

After taking into account the change in workers’ revenue share, the only equilibrium condition that changes when risk aversion is introduced is the Harris-Todaro condition of worker indifference between searching for employment in the two sectors. Specifically, this condition now equates the utility from a certain income of one in the homogenous sector with the expected utility from being hired and receiving a wage drawn from the equilibrium wage distribution in the differentiated sector:

$$x \sigma E w^{1-\eta} = x \sigma \int_{w_d}^{\infty} w^{1-\eta} dG_w(w) = 1. \tag{B.35}$$

Evaluating the integral using the open economy wage distribution (B.23), this condition becomes:

$$x \sigma E w^{1-\eta} = x \sigma \frac{1 + \mu}{1 + \mu \eta} w_d^{1-\eta} \left[ \frac{1 + \rho z^{-\beta (1-\eta k/\delta)}}{1 + \rho z^{-\beta (1-k/\delta)}} \right]^{(1-\eta)(1-k/\delta)} - 1 = 1 \tag{B.36}$$

Recall that the open economy hiring rate can be expressed as:

$$\sigma = \frac{1 + \left[ \frac{\Upsilon_x^{(1-\beta)(1-k/\delta)}}{\Gamma} - 1 \right] \rho z^{-\beta (1-k/\delta) / \Gamma} \frac{1}{1 + \mu} \frac{1}{\phi_w}}{1 + \left[ \frac{\Upsilon_x^{(1-\beta)}}{\Gamma} - 1 \right] \rho z^{-\beta / \Gamma}}$$

where

$$\phi_w = \left[ \frac{\Gamma}{\beta (1 - \gamma k)} \frac{\sigma \delta_{\text{min}}}{f_d} \right]^{k/\delta}.$$ 

In addition, the lowest wage in the differentiated sector can be written as $w_d = b \phi_w$. Therefore, we can rewrite the Harris-Todaro condition in the form given in the paper:

$$\Lambda (\rho, \Upsilon_x) \frac{b^{1-\eta} x}{\phi_w (1 + \mu \eta)} = 1, \tag{B.36}$$

where

$$\Lambda (\rho, \Upsilon_x) = \frac{1 + \rho z^{-\beta (1-k/\delta)}}{1 + \rho z^{-\beta}} \left[ \frac{\Upsilon_x^{(1-\beta)(1-k/\delta)}}{\Gamma} - 1 \right]^{(1-\beta)(1-k/\delta)}.$$ 

Evidently, we have $\Lambda (0, \Upsilon_x) = 1$ and $0 < \Lambda (\rho, \Upsilon_x) < 1$ for $0 < \rho \leq 1$. 

158
Using the Harris-Todaro condition (B.36) and the hiring cost function \( b = \alpha_0 x^{\alpha_1} \), we obtain the following expressions for \((x, b)\):

\[
\begin{align*}
x &= \left( \frac{(1 + \mu \eta) \phi_w^\eta}{\alpha_0^{1-\eta} \Lambda(\rho, \Upsilon_x)} \right)^{1/(1+\eta\alpha_1)}, \\
b &= \alpha_0^{1/(1+\eta\alpha_1)} \left( \frac{(1 + \mu \eta) \phi_w^\eta}{\Lambda(\rho, \Upsilon_x)} \right)^{\alpha_1/(1+\eta\alpha_1)}.
\end{align*}
\]

(B.37)  (B.38)

Expected worker income is therefore:

\[
\omega = xb = \alpha_0^{\eta/(1+\eta\alpha_1)} \left( \frac{(1 + \mu \eta) \phi_w^\eta}{\Lambda(\rho, \Upsilon_x)} \right)^{1/(1+\eta\alpha_1)}.
\]

(B.39)

A sufficient condition for \(0 < x < 1\) is now given by:

\[
\alpha_0^{1-\eta} > \frac{(1 + \mu \eta) \phi_w^\eta}{\Lambda(\rho, \Upsilon_x)}.
\]

The proof of Proposition 2.10 follows from the above. In the closed economy \(\rho = 0\) and therefore \(\Lambda(0, \Upsilon_x) = 1\). Opening economy to trade leads to \(\rho = \theta_d/\theta_x > 0\) and therefore \(\Lambda(\rho, \Upsilon_x) < 1\). Consequently, equations (B.37)-(B.39) imply that \(x, b\) and \(\omega\) are higher in the open economy than in autarky.

### B.6 Supplementary Derivations

**Derivation of the Production Technology**

We assume the following production function:

\[
y = \theta h^\gamma a = \theta \left( \frac{1}{h} \right)^{1-\gamma} \int_0^h a_i di,
\]

(B.40)

where \(i \in [0, h]\) indexes the workers employed by the firm. One way to think about this technology is the following: A manager with productivity \(\theta\) has one unit of time which he allocates equally among his employees. Thus, the manager allocates \(1/h\) of his time to each worker and, as a result, a worker with match-specific ability \(a\) can contribute \(\theta(1/h)^{1-\gamma}a\) to the total output of the firm, where \((1 - \gamma)\) measures the importance of the managerial time input. Aggregating across workers yields the assumed production function. We further assume, following a large literature on moral hazard in teams, that the contributions of individual workers to total output are unobservable as production is done in teams and the production process is non-separable. As a result, the match-specific ability of workers cannot be deduced from observing output. This justifies the assumption that the manager splits his time equally among the workers since they all appear homogeneous because of the unobservable and non-verifiable nature of their match-specific ability. Alternatively, an equal managerial time allocation among workers can be rationalized by assuming that the productivity of each worker depends on average worker ability as a result of human capital
externalities across workers within firms.¹

**Marginal Product of Labor**

Given the production technology (B.40), the marginal product of a worker with match-specific ability \( a \) is:

\[
MP(a) \equiv MP(a|\bar{a}, h; \theta) = \theta h^{-(1-\gamma)} [a - (1 - \gamma)\bar{a}].
\]

Let match-specific ability in the pool of candidate employees be distributed according to cumulative distribution function \( G_a(a) \). Then, if the firm only hires workers with ability above \( a_c \), the mean ability of its workers is

\[
\bar{a}(a_c) = \frac{1}{1 - G_a(a_c)} \int_{a_c}^{\infty} a \, dG_a(a).
\]

The marginal product of the threshold–ability worker is thus

\[
MP(a_c) = \theta h^{-(1-\gamma)} [a_c - (1 - \gamma)\bar{a}(a_c)],
\]

which is negative whenever \((1 - \gamma)\bar{a}(a_c) > a_c \). Since \( \bar{a}(a_c) > a_c \) for any non–degenerate distribution, \( MP(a_c) < 0 \) can be guaranteed by choosing \( \gamma \) small enough.

Specifically, consider the case of Pareto–distributed ability as we assume in the paper. In this case \( \bar{a}(a_c) = ka_c/(k - 1) \) and the marginal product is

\[
MP(a_c) = -\theta h^{-(1-\gamma)} \frac{1 - \gamma k}{k - 1} a_c,
\]

which is negative whenever \( \gamma k < 1 \). Recall that this is the same parameter restriction which ensures that output is increasing in the cutoff ability. It also guarantees that the workers not hired by the firm have negative marginal product. The firm will not want to retain these workers even at a zero wage, because the resulting reduction in output from lower average worker ability would dominate the increase in output from a greater measure of hired workers.

Note further that given \( \gamma k < 1 \) the firm hires some workers with ability in the range \([a_c, \hat{a})\), where

\[
\hat{a} = (1 - \gamma)\bar{a} = \frac{(1 - \gamma)k}{k - 1} a_c > a_c,
\]

¹ For empirical evidence on human capital externalities within plants, see for example Moretti (2004). See also related work on O-ring production technologies following Kremer (1993).

² To define the marginal product rewrite the production function as

\[
y = \theta \left[ \int_0^{h} d_i \right]^{-(1-\gamma)} \int_0^{h} a_i d_i.
\]

Then the marginal product of adding worker \( h \) with productivity \( a_h \) is

\[
MP_h \equiv \frac{dy}{dh} = \theta h^{-(1-\gamma)} [a_h - (1 - \gamma)\bar{a}].
\]

Note that the production function does not depend on the ordering of workers and hence the marginal product depends only on the ability of the worker in the sense that \( MP_h = MP(a_h) = MP(a_h|\bar{a}, h; \theta) \).
and these workers have a negative marginal product. The firm would prefer not to hire these workers, but costly search and screening make it optimal to hire all workers with match-specific ability above $a_c$. Finally, note that the average marginal product of workers employed by a firm with productivity $\theta$ is always positive:

$$\overline{MP}(\theta) = \gamma \theta h(\theta)^{-(1-\gamma)} \bar{a}(a_c(\theta)) > 0.$$ 

\[ Division of Revenue in the Bargaining Game \]

In our model, the firm and workers bargain about the division of revenues, as the outside option of workers is normalized to zero. As in Stole and Zweibel (1996a,b), the firm bargains with every worker taking into account the effect of his departure on the bargaining game with remaining workers. All the firm’s other decisions — sampling, screening, production, exporting — are sunk by the bargaining stage. Therefore, from (B.7) and (2.3), revenues $r(\theta, h)$ is a continuous, increasing and concave function of employment $h$, and all other arguments of firm revenue ($\theta$, $\bar{a}(\theta)$, $\Upsilon(\theta)$, $A$) are fixed. Let $w(\theta,h)$ be the bargained wage rate that a $\theta$-firm pays as a function of the measure of workers hired $h$. This function has to satisfy the differential equation:

$$\frac{\partial}{\partial h} \left[ r(\theta,h) - w(\theta,h) h \right] = w(\theta,h)$$

(B.41)

so that the surplus of worker from employment (the wage rate) is equal to the marginal surplus of the firm from employing the worker.\(^3\) Using the assumed functional forms for revenues this differential equation yields the solution

$$w(\theta,h) = \frac{\beta \gamma}{1 + \beta \gamma} \frac{r(\theta,h)}{h},$$

so that each workers gets a fraction $\beta \gamma/(1 + \beta \gamma)$ of average revenues and the firm gets the remaining $1/(1 + \beta \gamma)$ share of revenues. The worker’s share of surplus is increasing in $\beta \gamma$, where $\beta$ captures the concavity of demand and $\gamma$ captures the concavity of the production technology. Therefore, the worker’s share of surplus is decreasing in the concavity of the revenue function in $h$, because a more concave revenue function implies a smaller effect of the departure of any given worker on firm revenue.

\(^3\) Since individual abilities are unobservable, the expressions on both sides of (B.41) are evaluated holding $\bar{a} = \bar{a}(\theta)$ constant for all $h$. Indeed, when a marginal worker departs, the productivity of the remaining workers is still $\bar{a}$ in expectation. The firm evaluates the expected loss from the departure of a worker as:

$$\frac{\partial r}{\partial h} + \frac{\partial r}{\partial \bar{a}} \frac{\partial \bar{a}}{\partial h},$$

where $\partial \bar{a}/\partial h$ equals zero in expectation and hence there remains only the direct effect of $h$ on revenues. As a result, the bargaining solution in this environment with symmetric uncertainty is the same as if all workers had a productivity of $\bar{a}$.
C. APPENDICES FOR CHAPTER III

C.1 Derivations and Proofs for Section 3.2

Consider the problem of the government:

\[ W(t) \equiv L \int_0^\infty G(U_n) dH(n) \rightarrow \max_t \]

subject to

\[ U_n = \max_y \{ tR + (1 - t)Q^{1-\beta}y^\beta - v(y/n) \}, \]

where \( LR = Q = [L \int_0^\infty y^\beta dH(n)]^{1/\beta} \) and \( y_n \) maximizes agent \( n \)'s utility, which with our function form \( v(x) = x^{\gamma}/\gamma \) yields

\[ y_n = \left[ \beta(1 - t) \right]^{\frac{1}{\beta + \gamma}} Q^{\frac{1}{\beta + \gamma}} n^{\frac{\gamma}{\beta + \gamma}} . \]

With this solution for output \( y_n \), we have the following individual revenue and utility:

\[ r_n = \left[ \beta(1 - t) \right]^{\frac{1}{\beta + \gamma}} Q^{\frac{1}{\beta + \gamma}} n^{\frac{\gamma}{\beta + \gamma}}, \]

\[ U_n = tR + (1 - t)(1 - \beta/\gamma) r_n. \]

The equilibrium differential of the agent’s utility with respect to the tax rate is

\[ \frac{dU_n}{dt} = (R - r_n) + t \frac{dR}{dt} + (1 - t)(1 - \beta) r_n \frac{1}{Q} \frac{dQ}{dt}, \quad r_n = Q^{1-\beta} y_n^\beta. \]

The term \( dy_n/dt \) enters into the differential with a coefficient of zero due to the Envelop Theorem since agents choose \( y_n \) optimally.\(^1\) Further, since \( R = Q/L \), we have \( d \ln R = d \ln Q \) and we can rewrite:

\[ \frac{dU_n}{dt} = (R - r_n) - \tilde{\varepsilon} \left[ \frac{t}{1 - t} R + (1 - \beta) r_n \right], \quad \tilde{\varepsilon} \equiv \frac{d \ln Q}{d \ln(1 - t)} = - \frac{(1 - t)Q}{Q dt}. \]

Consider now the first order condition for the government’s problem:

\[ \frac{\partial W(t)}{\partial t} = \int_0^\infty G'(U_n) \frac{dU_n}{dt} dH(n) = \int_0^\infty G'(U_n) \left\{ (R - r_n) - \tilde{\varepsilon} \left[ \frac{tR}{1 - t} + (1 - \beta) r_n \right] \right\} dH(n) = 0. \]

\(^1\) If agents were price takers, \( dy_n/dt \) would not be zero since agents are not fully optimizing in this case. Instead the sum \( [1 - \beta d \ln Q + \beta d \ln y_n]/dt \) would be zero. Therefore, both terms with \( dQ/dt \) and \( dy_n/dt \) drop out from the expression for \( dU_n/dt \). This then results in (3.11) becoming \( t/(1 - t) = \alpha/\tilde{\varepsilon} \), just as in the case with \( \beta = 1 \).
Expressing out \( t/(1 - t) \), we have:

\[
\frac{t}{1 - t} = \frac{1}{\tilde{\varepsilon}} \cdot \alpha - (1 - \beta)(1 - \alpha),
\]

where

\[
\alpha = \int_0^\infty G'(U_n)(R - r_n)dH(n) / \lambda R, \quad \lambda \equiv \int_0^\infty G'(U_n)dH(n).
\]

Note that \( \alpha \) is a cross-sectional covariance between two normalized variables, \( \lambda^{-1}G'(\mathcal{U}) \) and \( 1 - r/R \), both with a mean of 1. Alternatively, \( \alpha \) is minus the covariance between \( \lambda^{-1}G'(\mathcal{U}) \) and \( r/R \).

With this notation we can rewrite the first order condition as:

\[
\frac{\partial W(t)}{\partial t} = \lambda^{-1}R \left[ \alpha - \tilde{\varepsilon} \left( \frac{t}{1 - t} + (1 - \beta)(1 - \alpha) \right) \right] = 0.
\]

Then the second order condition can be written as:

\[
\frac{\partial^2 W(t)}{\partial t^2} = \lambda^{-1}R \left[ \frac{\partial \alpha}{\partial t} (1 + \tilde{\varepsilon}(1 - \beta)) - \frac{\tilde{\varepsilon}}{(1 - t)^2} \right],
\]

where we used the result of Lemma 3.2 that \( \tilde{\varepsilon} = \varepsilon \) and does not depend on \( t \) in the closed economy. Note that \( \partial \alpha / \partial t < 0 \) is sufficient, but is not necessary for the second order condition to be satisfied. Finally, one can show that \( \partial \alpha / \partial t < 0 \) is also sufficient for the concavity of \( W(t) \) on the relevant range (i.e., for \( t < 1/(1 + \varepsilon) \)).

We can now prove:

**Proof of Lemma 3.1:** Since \( G(\cdot) \) is increasing and concave, we have \( G'(\cdot) \geq 0 \) and \( G''(\cdot) \leq 0 \). Therefore, \( G'(U_n) \) is positive and decreasing with \( r_n \) since \( U_n \) increases with \( r_n \). This implies that \( \alpha \geq 0 \) as the negative of a covariance between two variables that move in the opposite directions.

Next note that we can write

\[
\alpha = 1 - \int_0^\infty \lambda^{-1}G'(U_n)\frac{r_n}{R}dH(n) \leq 1,
\]

since the second term is non-negative as an integral of a two non-negative function. Therefore, we conclude that \( 0 \leq \alpha \leq 1 \).

Further, \( \alpha = 0 \) if and only if either \( r_n/R \equiv 1 \) or \( G'(U_n) = \text{const} \). The former is the case when \( H(n) \) is degenerate with all mass at one value of \( n = n_{\text{min}} = n_{\text{max}} \) implying \( r_n \equiv R \) for all agents. The latter is the case when \( G'(\cdot) \equiv 1 \), i.e. \( \rho = 0 \) and there is no inequality aversion.

Finally, \( \alpha = 1 \) if and only if \( G'(U_n) \cdot r_n = 0 \) with probability 1. This is the case if and only if the planner has all weight on agents with no revenues. This, however, requires both that the planner is Rawlsian (\( \rho \to \infty \)) and that the least productive agent does not produce (\( r_{n_{\text{min}}} = 0 \)). The utility of this agent is given by \( U_n = tR \) and its maximization results in \( t/(1 - t) = 1/\tilde{\varepsilon} \), or \( t = 1/(1 + \tilde{\varepsilon}) \), which is at the peak of the Laffer curve.
Proof of Lemma 3.2: Log-differentiating the definition of $Q$, (3.2), we have

$$\frac{d \ln Q}{d \ln(1-t)} = L \int_0^\infty \left( \frac{y_n}{Q} \right)^\beta \frac{d \ln y_n}{d \ln(1-t)} dH(n).$$

From the expression for agents’ optimal output $y_n$, we have:

$$\frac{d \ln y_n}{d \ln(1-t)} = \frac{1}{\gamma - \beta} + \frac{1 - \beta}{\gamma - \beta} \frac{d \ln Q}{d \ln(1-t)}.$$

Combining these two expressions and observing that $L \int_0^\infty \left( \frac{y_n}{Q} \right)^\beta dH(n) = 1$ from the definition of $Q$, we obtain:

$$\delta \equiv \frac{d \ln Q}{d \ln(1-t)} = \frac{1}{\gamma - 1} = \varepsilon.$$

Comparative Statics for $\alpha$: Denote $x = r/R$. From Lemma 3.4, the distribution of $x$, $\varphi(x)$, is the normalized distribution of $n^{\beta\gamma}/(\gamma - \beta)$ with mean 1, and hence it does not depend on $t$ or $\rho$. Therefore, we can treat $\varphi(x)$ as being given exogenously. Denote $\delta = (1 - \beta/\gamma)(1 - t)/t$. Then we can write $\alpha$ as

$$\alpha = \text{cov} \left( 1 - x, \frac{(1 + \delta x)^{-\rho}}{\mathbb{E}(1 + \delta x)^{-\rho}} \right) = \frac{\int_0^\infty (1 - x)(1 + \delta x)^{-\rho} \varphi(x) dx}{\int_0^\infty (1 + \delta x)^{-\rho} \varphi(x) dx}.$$

One can show that for $\rho > 0$,

$$\text{sign} \left( \frac{\partial \alpha}{\partial t} \right) = -\text{sign} \left( \frac{\partial \alpha}{\partial \delta} \right) = \text{sign} \left( \text{cov} \left( 1 - x, \frac{x(1 + \delta x)^{-\rho - 1}}{\mathbb{E} \{x(1 + \delta x)^{-\rho - 1}\}} - \frac{(1 + \delta x)^{-\rho}}{\mathbb{E}(1 + \delta x)^{-\rho}} \right) \right).$$

The condition for $\partial \alpha/\partial t < 0$ can be written as follows:

$$\int_0^\infty \frac{x\varphi(x) dx}{(1 + \delta x)^{1+\rho}} \int_0^\infty \frac{x\varphi(x) dx}{(1 + \delta x)^{\rho + 1}} < \int_0^\infty \frac{x^2\varphi(x) dx}{(1 + \delta x)^{\rho + 1}} \int_0^\infty \frac{\varphi(x) dx}{(1 + \delta x)^{\rho}}.$$

Although this inequality is very close to Cauchy-Schwarz inequality, it does not hold for a general distribution $\varphi(x)$. However, this inequality can be shown numerically to hold for Uniform and Pareto distributions.

Proof of Lemma 3.3: Using the expression for $\mathcal{U}_n$, we have

$$G'(\mathcal{U}_n) = (tR + (1-t)(1-\beta/\gamma)r_n)^{-\rho} = \left( [t + (1 - \beta/\gamma)(1 - t)] R \right)^{-\rho} \left[ 1 + \frac{1}{1 + \frac{1}{1-t} \frac{1}{1-\beta/\gamma}} \left( \frac{r_n}{R} - 1 \right) \right]^{-\rho}.$$

Taking the first order Taylor approximation to $G'(\mathcal{U}_n)$ around $r_n/R = 1$ yields:

$$G'(\mathcal{U}_n) \approx \left( [t + (1 - \beta/\gamma)(1 - t)] R \right)^{-\rho} \left[ 1 - \rho \frac{1}{1 + \frac{1}{1-t} \frac{1}{1-\beta/\gamma}} \left( \frac{r_n}{R} - 1 \right) \right] + O \left( \frac{r_n}{R} - 1 \right)^2.$$
This implies
\[
\lambda = \int_0^\infty G'(U_n) \partial H(n) = \left(\left[ t + (1 - \beta/\gamma) (1 - t) \right] R \right)^{-\rho} + O_p \left( \frac{r}{R} - 1 \right)^2
\]
and, therefore,
\[
\lambda^{-1} G'(U_n) = 1 - \rho \frac{1}{1 + \frac{t}{1-t} \frac{1}{1-\beta/\gamma}} \left( \frac{r_n}{R} - 1 \right) + O_p \left( \frac{r}{R} - 1 \right)^2.
\]
Now using the definition of \(\alpha\), we have
\[
\alpha = -\text{cov} \left( \lambda^{-1} G'(U), \frac{r}{R} \right) = \rho \frac{1}{1 + \frac{t}{1-t} \frac{1}{1-\beta/\gamma}} \text{var} \left( \frac{r}{R} \right) + O_p \left( \frac{r}{R} - 1 \right)^3.
\]

**Proof of Lemma 3.4:** The solution to agent \(n\)'s problem yields:
\[
r_n = \left[ \beta(1-t) \right]^{\frac{\beta}{1-\gamma}} Q^{\frac{1-\beta}{1-\gamma}} n^{\frac{\beta\gamma}{1-\gamma}}.
\]
Since \( R = \int_0^\infty r_n \partial H(n) \), we have
\[
\frac{r_n}{R} = \frac{n^{\frac{\beta\gamma}{1-\gamma}}}{\int_0^\infty n^{\frac{\beta\gamma}{1-\gamma}} \partial H(n)}.
\]
Therefore, the distribution of \(r_n/R\) is the same as the normalized distribution of \(n^{\beta\gamma/(\gamma-\beta)}\). An increase in the dispersion of \(n^{\beta\gamma/(\gamma-\beta)}\) holding its mean constant leads to an increase in the dispersion of relative revenues.

**Proof of Proposition 3.2:** Assuming that the approximation in Lemma 3.3 is accurate, we have
\[
\frac{t}{1-t} = \alpha \left( \frac{1}{\varepsilon} + (1 - \beta) \right) - (1 - \beta), \quad \alpha \approx \rho \frac{1}{1 + \frac{t}{1-t} \frac{1}{1-\beta/\gamma}} \text{var} \left( \frac{r}{R} \right).
\]
This immediately implies \(\partial t / \partial \rho > 0\) and \(\partial t / \partial \text{var}(r/R) > 0\). Next, Lemma 3.4 implies that \(L\) does not affect the equilibrium distribution of \(r/R\), and hence, it also does not affect the optimal tax rate, \(t\). Finally, combining Lemmas 3.3 and 3.4, we have
\[
\alpha \approx \rho \frac{1}{1 + \frac{t}{1-t} \frac{1}{1-\beta/\gamma}} \left( \frac{\beta\gamma}{\gamma - \beta} \right)^2 \text{var}(\ln n).
\]
Choose \(\rho \text{var}(\ln n)\) such that optimal \(t = 0\). Then, using \(\gamma = 1 + 1/\varepsilon\), we have
\[
\text{sign} \left( \frac{\partial t}{\partial \varepsilon} \right) = \text{sign} \left( 2\beta - 1 - \frac{1}{\varepsilon} \right) \geq 0.
\]
Equilibrium with Asymmetric Countries

As discussed in footnote 18, the nominal revenues in the domestic market are given by $PQ_1^{1-\beta}y_d^{\beta}$. It is convenient to denote $Z \equiv QP^{1/(1-\beta)}$, so that nominal revenues can be written as $Z^{1-\beta}y_d^{\beta}$. Nominal revenues from exports are then given by $Z^{1-\beta}(y_x/\tau)^{\beta}$. The agent optimally splits his production $y$, to supply $y_d$ in the domestic and to export the remaining $y_x = y - y_d$ to the foreign market, in order to maximize real revenues $\left[Z^{1-\beta}y_d^{\beta} + Z^{1-\beta}(y_x/\tau)^{\beta}\right]/P$. This results in the following real revenue function:

$$r(y) = \Upsilon_x^{1-\beta}Q^{1-\beta}y^{\beta}, \quad \Upsilon_x \equiv 1 + \tau^{-\frac{\beta}{\gamma}} Z^*.$$

In addition, the split of output between the domestic and foreign market is given by $y_d = y/\Upsilon_x$ and $y_x = (\Upsilon_x - 1)y/\Upsilon_x$. Note the more general relative to (3.14) definition of the market access variable, $\Upsilon_x$, which now increases in the relative measure of nominal demand, $Z^*/Z$. $Z^*/Z$ can be high either because consumption is relatively high abroad or because the price level is relatively high abroad. Of course, in equilibrium $Q^*/Q$ and $P^*/P$ are closely connected.

The problem of an agent with productivity $n$ is given by

$$U_n = \max_y \left\{ \Delta + (1-t)\Upsilon_x^{1-\beta}Q^{1-\beta}y^{\beta} - \frac{1}{\gamma} \left( \frac{y}{n} \right) \right\}.$$ 

It results in

$$y_n = \left[\beta(1-t)\right]^{\frac{1}{\gamma}} \Upsilon_x^{\frac{1-\beta}{\gamma}} Q^{1-\beta} n^{-\frac{1}{\gamma}},$$

$$r_n = \Upsilon_x^{1-\beta}Q^{1-\beta}y_n^{\beta},$$

$$U_n = tR + (1-t)(1-\beta) r_n,$$

where the last expression utilizes the government budget constraint $\Delta = tR$ and $R = \int_0^\infty r_n dH(n)$ is the average revenue.

Denote by $Y$ the per capita total domestic production of the final good:

$$Y = \left[\int_0^\infty y_n^{\beta} dH(n)\right]^{1/\beta}.$$

Then we have

$$R = \Upsilon_x^{1-\beta}Q^{1-\beta}Y^{\beta},$$

and from agents’ optimization

$$Y^{\beta} = \left[\beta(1-t)\right]^{\frac{1}{\gamma}} \Upsilon_x^{\frac{1-\beta}{\gamma}} Q^{1-\beta} \Theta,$$

where $\Theta \equiv \int_0^\infty n^{\frac{\beta}{\gamma}} dH(n)$. 

166
Domestic consumption is given by

\[ Q = \left[ L \left( \frac{1}{1 - \alpha} \right) \right]^\beta Y^\beta + \tau^{-\beta} L^* \left( \frac{Y^* - 1}{Y^*} \right)^{\beta Y^\beta} \right]^{1/\beta}. \]

Balanced trade implies \( LR = Q \). Using the expressions for \( R \) and \( Q \), balanced trade condition can be rewritten as

\[ \frac{LY^\beta}{LY^*} = \left( \frac{Z}{Z^*} \right)^{1+\beta} \left( \frac{Y_x}{Y_x^*} \right)^\beta. \]

Using the expression for \( R \), this can be rewritten as

\[ \left( \frac{Q}{Q^*} \right)^\beta = \frac{Y_x}{Y_x^*} \left( \frac{Z}{Z^*} \right)^{1+\beta}, \quad Y_x = \frac{Z^* Z + \tau^{-\beta} Z^*}{Z^* + \tau^{-\beta} Z}. \tag{C.1} \]

This condition constitutes an equilibrium relationship between \( Q/Q^* \) and \( Z/Z^* \), and by consequence with \( P/P^* \). Next, combining the expression for \( Y \) and \( R \), we obtain the expression for domestic consumption

\[ Q^\beta \gamma^{-1} = \left[ \beta(1 - t) \right]^{\beta} L^X \gamma^\beta \gamma^{-1}. \tag{C.2} \]

Conditions (C.1), (C.2) and its counterpart for the foreign country are sufficient to solve for equilibrium \( Q, Z/Z^* \), \( Y_x \) and so forth. All the comparative statics can be done on these two conditions. Finally, (C.2) and its foreign counterpart result in

\[ \left( \frac{Q}{Q^*} \right)^{\frac{\delta_{x}}{\gamma - 1}} \left( \frac{Z}{Z^*} \right)^{1+\beta \frac{\delta_{x}}{\gamma - 1}} = \left( \frac{L^X}{L^X^*} \right)^{\frac{\delta_{x}}{\gamma - 1}} \left( \frac{1 - t}{1 - t^*} \right)^{\frac{1}{\gamma - 1}}. \tag{C.3} \]

Proof of Proposition 3.3: Combining the expressions for \( r_n \) and \( y_n \), we have

\[ r_n = \left[ \beta(1 - t) \right]^{\beta} \gamma^\beta \gamma^{-1} Q^\gamma \gamma^\beta \gamma^{-1} n^\gamma \gamma^{-1}. \]

Therefore, relative revenue is

\[ \frac{r_n}{R} = \frac{r_n}{\int_0^\infty r_n dH(n)} = \frac{n^\frac{\delta_{x}}{\gamma - 1}}{\int_0^\infty n^\frac{\delta_{x}}{\gamma - 1} dH(n)}, \]

as in the closed economy (Lemma 3.4). Note that this prove is invariant to the asymmetry of countries, the level of taxes in the two countries \( t \) and \( t^* \) and other details of the open economy equilibrium.

Proof of Proposition 3.4: From (C.2), we have that open economy equilibrium aggregate consumption is given by

\[ Q = \left[ \beta(1 - t) \right]^{\beta} (L^X)^{\frac{\delta_{x}}{\gamma - 1}} \gamma^\beta \gamma^{-1} n^\gamma \gamma^{-1}, \quad \varepsilon = \frac{1}{\gamma - 1}. \]
Note that in an open economy equilibrium with positive trade flows $\Upsilon_x > 1$ and the autarky obtains as a special case when $\tau \to \infty$ and hence $\Upsilon_x \to 1$. Therefore, we have

$$\frac{Q}{Q^a} = \left(\frac{1-t}{1-t^a}\right)^\varepsilon \Upsilon_x^{(1+\varepsilon) \frac{1-t^a}{1-t^a}},$$

where superscript $a$ stands for autarky. Since $LR = Q$ both in open economy and in autarky, we have $R/R^a = Q/Q^a$. Finally, since Proposition 3.3 implies $r_n/R = r_n^a/R^a$, and in equilibrium $U_n = tR + (1-t)(1-\beta/\gamma)r_n$, we have

$$\forall n \quad \frac{U_n}{U_n^a} = \frac{r_n}{r_n^a} = \frac{R}{R^a} = \frac{Q}{Q^a} = \left(\frac{1-t}{1-t^a}\right)^\varepsilon \Upsilon_x^{(1+\varepsilon) \frac{1-t^a}{1-t^a}}.$$

Therefore, setting $t = t^a$ in the open economy allows to increase utilities proportionally for all agents, and hence, increases welfare. This is independent of the tax rate in the trade partner and any asymmetries between countries since in $\Upsilon_x \geq 1$ and in any open economy equilibrium with positive trade flows the inequality is strict. Finally, by choosing $t$ optimally, the government will only further increase welfare. In other words, trade necessarily improves the choice set of the government. A similar proof allows to demonstrate welfare gains from any marginal reduction in trade costs $\tau$. \footnote{One additional step in this proof is to show that $\Upsilon_x$ necessarily increases as $\tau$ falls (see comparative statics below).

Losers from Trade: Consider agents with lowest skill $n_0 = 0$ and highest skill $n_\infty = \infty$ (even if these agents are imaginary and never occur in the model economy). The utilities of these agents are $U_0 = tR$ and $U_\infty = (1-t)(1-\beta/\gamma)r_\infty$ as the former agent does not produce and for the latter agent the government’s transfer is negligible relative to his after-tax revenues. From the discussion above we know that $r_n$ for all $n$ and $R$ increase proportionally in the open economy. Therefore, if both $n_0$ and $n_\infty$ gain from trade, then all agents in the economy gain from trade. Moreover, since there are aggregate gains from trade, $n_0$ necessarily gains when $t > t^a$ and $n_\infty$ necessarily gains when $t < t^a$. The conditions for these agents to lose from trade can be written as follows:

$$d \ln U_0 = -\frac{1-t}{t} d \ln(1-t) + d \ln Q < 0,$$

$$d \ln U_\infty = d \ln(1-t) + d \ln Q < 0,$$

where $t$ adjusts in response to trade and $Q$ changes in responds to both trade and a change in $t$. Comparative statics below allows to rewrite these conditions in terms of primitives. As discussed above, only one condition can be satisfied at a time. When neither of the conditions is satisfied, all agents in the economy gain from trade.

\footnote{One additional step in this proof is to show that $\Upsilon_x$ necessarily increases as $\tau$ falls (see comparative statics below).}
Comparative Statics with Asymmetric Countries: Take the total log-differential of (C.1)-(C.3) and the expression for $\Upsilon_x$:

$$\beta \left( \hat{Q} - \hat{Q}^* \right) = (1 + \beta) \left( \hat{Z} - \hat{Z}^* \right) + \left( \hat{\Upsilon}_x - \hat{\Upsilon}_x^* \right)$$

$$\hat{Q} = \varepsilon d \ln(1-t) + \frac{1 - \beta}{\beta} (1 + \varepsilon) \hat{\Upsilon}_x$$

$$\varepsilon (\gamma \beta - 1) \left[ \hat{Q} - \hat{Q}^* \right] + \frac{1 - \beta}{\beta} (1 + \varepsilon)(1 + \beta) \left[ \hat{Z} - \hat{Z}^* \right] = \varepsilon \left[ d \ln(1-t) - d \ln(1-t^*) \right]$$

$$\hat{\Upsilon}_x = \frac{1 + \beta}{\beta} \kappa \left[ -\frac{\beta}{1 - \beta} \hat{\tau} + (\hat{Z}^* - \hat{Z}) \right], \quad \kappa \equiv \frac{\beta}{1 + \beta} \frac{\Upsilon_x - 1}{\Upsilon_x} = \frac{\beta}{1 + \beta} \frac{\tau - \frac{\tau}{1 + \tau} (Z^*/Z)}{1 + \frac{\tau}{1 + \tau} - \frac{\tau}{1 + \tau} (Z^*/Z)} \in [0, 1/2),$$

where hats denote log-differentials. This is a system of four linear equations which allows to solve for $\hat{Q}, \hat{Z}, \hat{\Upsilon}_x$ and their foreign counterparts as functions of $\hat{\tau}, d \ln(1-t)$ and $d \ln(1-t^*)$.

We now provide this solution for the special case $\gamma \beta = 1$, which greatly simplifies the resulting expressions:

$$\hat{Q} = \varepsilon d \ln(1-t) + \hat{\Upsilon}_x,$$

$$\frac{1 + \beta}{\beta} \left[ \hat{Z} - \hat{Z}^* \right] = \varepsilon \left[ d \ln(1-t) - d \ln(1-t^*) \right],$$

$$\hat{\Upsilon}_x = -\frac{1 + \beta}{1 - \beta} \kappa \hat{\tau} - \kappa \varepsilon \left[ d \ln(1-t) - d \ln(1-t^*) \right],$$

which leads to

$$\hat{Q} = -\frac{1 + \beta}{1 - \beta} \kappa \hat{\tau} + \varepsilon (1 - \kappa) d \ln(1-t) + \varepsilon \kappa d \ln(1-t^*).$$

Note that real consumption increases as trade cost $\tau$ falls. Real consumption falls both in the levels of domestic and foreign tax rates. The elasticity with respect to domestic (foreign) tax rate is smaller (larger) the greater is the value of $\kappa$; $\kappa$ is high when trade cost $\tau$ is low or when the domestic country is relatively small (in terms of effective demand $Z^*/Z$). In the closed economy $\kappa = 0$. The results in the general case with $\gamma \beta \neq 1$ are qualitatively the same.

Finally, we evaluate the comparative statics for agents’ utilities and aggregate welfare. From agent $n$’s problem and using Envelop Theorem for the choice of $y_n$, we have the following utility differential:

$$d U_n = (R - r_n)dt + \left[ \frac{t}{1 - t} + (1 - \beta) \frac{r_n}{R} \left( 1 + \frac{\hat{\Upsilon}_x}{Q} \right) \right] (1 - t) R \hat{Q}.$$

Holding $\tau$ constant and using the above comparative statics, we obtain:

$$\frac{d U_n}{d t} = (R - r_n) - \varepsilon (1 - \kappa) \left[ \frac{t}{1 - t} + (1 - \beta) \frac{r_n}{R} \frac{1 - 2\kappa}{1 - \kappa} \right] R$$

and

$$\frac{d U_n}{d t^*} = -\varepsilon \kappa \left[ \frac{t}{1 - t} + 2(1 - \beta) \frac{r_n}{R} \right] R.$$
Therefore,
\[
\frac{dU_n}{dt} + \frac{dU_n}{dt^*} = (R - r_n) - \varepsilon \left[ t \frac{t}{1-t} + (1 - \beta) \frac{r_n}{R} \right] R,
\]
equivalent to \( \frac{dU_n}{dt} \) in the closed economy.

The differential of the welfare function is given by
\[
dW(t,t^*) = \int_0^\infty G'(U_n) dU_n dH(n).
\]
Therefore,
\[
\frac{\partial W(t,t^*)}{\partial t} = \lambda R \left[ \alpha - \varepsilon (1 - \kappa) \left( \frac{t}{1-t} + \frac{1-2\kappa}{1-\kappa} (1-\beta)(1-\alpha) \right) \right],
\]
\[
\frac{\partial W(t,t^*)}{\partial t^*} = -\lambda R \left[ \varepsilon \kappa \left( \frac{t}{1-t} + 2(1-\beta)(1-\alpha) \right) \right],
\]
where \( \alpha \) and \( \lambda \) are as defined in Section 3.2. Note that
\[
\frac{\partial W(t,t^*)}{\partial t} + \frac{\partial W(t,t^*)}{\partial t^*} = \lambda R \left[ \alpha - \varepsilon \left( \frac{t}{1-t} + (1-\beta)(1-\alpha) \right) \right],
\]
equivalent to \( \partial W^a(t)/\partial t \) in the closed economy. This result holds even for \( t \neq t^* \).

Non-cooperative Determination of Taxes: Non-cooperative solution is a Nash equilibrium in which both countries unilaterally choose their tax rate taking the tax rate in the trading partner as given. Formally, \((t_0, t_0^*)\) is a Nash equilibrium if \( t_0 \in \arg \max_t W(t; t_0^*) \) and \( t_0^* \in \arg \max_{t^*} W^*(t^*; t_0) \). The first order condition for the choice of \( t_0 \) is
\[
\frac{\partial W(t_0, t_0^*)}{\partial t} = \lambda R \left[ \alpha - \varepsilon (1 - \kappa) \left( \frac{t_0}{1-t_0} + \frac{1-2\kappa}{1-\kappa} (1-\beta)(1-\alpha) \right) \right] = 0,
\]
which results in
\[
\frac{t_0}{1-t_0} = \frac{1}{\varepsilon (1-\kappa)} \cdot \alpha - \frac{1-2\kappa}{1-\kappa} (1-\beta)(1-\alpha), \quad \kappa \in [0, 1/2).
\]
A symmetric condition holds for the foreign country. We also denote \( W_0 = W(t_0, t_0^*) \) and \( W_0^* = W^*(t_0^*, t_0) \).

Note the difference between this condition and (3.11') in the closed economy. Trade openness implies \( \kappa > 0 \), and as a result the efficiency margin of taxation decreases. There are two effects. First, note that \( \varepsilon d \ln Q/d \ln(1-t) = \varepsilon (1-\kappa) < \varepsilon \). That is, domestic consumption is less sensitive to the domestic taxes in the open economy when the foreign tax rate is held fixed. The reason is that agents can substitute away from more expensive local goods to the foreign goods for which prices do not change. In other words, terms of trade improve in response to an increase in domestic income taxes and partly shield the country from the efficiency losses. Moreover, from the discussion above (Proposition 3.5) it follows that world efficiency losses are still given by \( \varepsilon \), implying that the remainder of the efficiency losses are born by the trade partner through the deterioration of the
terms of trade.\(^3\)

Second, in the expression for \(t_0/(1 - t_0)\), the weight on the negative monopolistic distortion component \((1 - \beta)\) is reduced. This component is supposed to offset the monopolistic mark-up in the economy. There are less incentives to fully offset the monopolistic mark-up in an open economy for two related reasons: (1) higher monopolistic prices induce terms of trade improvement beneficial for the country and (2) inefficiency of monopolistic pricing harms domestic consumers only partially now as the effects are spilled over onto the foreign consumers who do not enter domestic welfare.

Overall, both of the two new effects operate in the direction to reduce equilibrium taxes. Both effects are stronger the larger is \(\kappa\). This is the case when the economy is smaller (in terms of its demand level \(Z\)) and more open (in terms of its ratio of trade to GDP which decreases in \(\tau\)). This is intuitive, since for more open and smaller countries trade provides a better shield from inefficient domestic policies. This may rationalize the finding in Rodrik (1998) and Alesina and Wacziarg (1998) that both smaller and more open economies tend to have larger governments.

We summarize these findings in: \(^4\)

**Proposition C.1.** In an open economy without fixed costs, non-cooperatively-set tax rate is higher than the optimal tax rate in the closed economy. The difference is larger the smaller is the country and the more open it is to trade. Non-cooperatively-set taxes are inefficient high.

Contrast this proposition with Proposition 3.5: there is a stark difference between cooperatively and non-cooperatively-set income taxes. Since higher tax rate at home harms the trade partner, and access to foreign trade partially shields the home economy from inefficiency, non-cooperative policy determination always results in higher taxes. This result is similar to findings of Epifani and Gancia (2008) in a different model of taxation in an open economy, who also provide empirical evidence for the importance of the terms of trade externality in affecting the size of the government in an open economy. At the same time, the quantitative calibration of Mendoza and Tesar (2005) suggests that the gains from tax policy coordination are small in the EU, which points towards small calibrated values of \(\kappa\)'s.

Finally, a direct implication of this discussion is that non-cooperatively set taxes are inefficient, in the sense that they leave the two countries strictly inside the Pareto frontier. Formally, there exist tax policies \(\tilde{t}\) and \(\tilde{t}^*\) such that \(W(\tilde{t}, \tilde{t}^*) > W_0\) and \(W^*(\tilde{t}^*, \tilde{t}) > W_0^*\). To see this, set \(\tilde{t} = t_0 - \delta\) and \(\tilde{t}^* = t_0^* - \delta\) for \(\delta\) small enough. In the Nash equilibrium, we have \(\partial W(t_0, t_0^*)/\partial t = 0\). Moreover, from the comparative statics above, we know that \(\partial W(t, t^*)/\partial t^* < 0\). In words, reducing \(t\) has a second-order negative effect on \(W\), while reducing \(t^*\) has a first-order positive effect. Therefore,

\[^3\] An opposite result obtains in a model with free entry (see Chapter I): there effects on entry dominate the terms of trade movements. As a result, more distortions at home improve welfare in the trade partner while they reduce efficiency at home by more than in the closed economy. Since models with free entry are more relevant for profit taxation rather than income taxation, this may have implications for the differential response of profit versus income taxes and capital versus labor taxes in the open economy (cf. Mendoza and Tesar, 2005).

\[^4\] Formally, the proof requires that \(\partial \alpha/\partial t < 0\) globally which we assume is the case (see Lemma 3.3 and the discussion that follows).

171
when $t$ and $t^*$ are both reduced by a small amount, $W$ necessarily increases. A similar argument works for the foreign country.

Cooperatively-set national policies make sure that these welfare gains from coordination are not forgone.

*Cooperative Determination of Taxes:* is a Nash bargaining solution in the game of the two countries, where the non-cooperative Nash equilibrium is taken as a status quo. Hence cooperatively-set taxes satisfy

$$
\max_{t, t^*} \left\{ \left( W(t, t^*) - W_0 \right) \left( W^*(t^*, t) - W^*_0 \right) \right\}.
$$

Therefore, the optimality condition for the tax rate is

$$
\frac{1}{W(t, t^*) - W_0} \frac{\partial W(t, t^*)}{\partial t} + \frac{1}{W^*(t^*, t) - W^*_0} \frac{\partial W^*(t^*, t)}{\partial t} = 0.
$$

A similar condition holds for $t^*$. When the two countries are symmetric, this condition reduces to $\partial W(t, t^*)/\partial t + \partial W^*(t^*, t)/\partial t = 0$. See the proof of Proposition 3.5.

We now consider the case of asymmetric countries. Using the comparative statics results, we can combine the two optimality conditions in this case to obtain:

$$
\eta \varepsilon \left[ -\frac{t}{1 - t} - \alpha / \varepsilon + (1 - \beta)(1 - \alpha) \right] + \eta^* \varepsilon^* \left[ -\frac{t^*}{1 - t^*} - \alpha^* / \varepsilon^* + (1 - \beta)(1 - \alpha^*) \right] = 0,
$$

where $\eta \equiv R/(W - W_0)$. This condition has two implications. First, condition (3.11') still applies, but only on average across the two economies. Generally, the tax rates in the open economy will be different from those in the closed economy, unless the two countries chose the same tax rate in autarky. Second, if one country has a higher tax rate than in autarky, the other country necessarily has a lower tax rate than in autarky. Moreover, under the assumption that the country that had a higher tax rate in autarky still has a higher tax rate in the open economy, there has to be a convergence of tax rates across countries when they start trading.

*Proof of Proposition 3.5:* As discussed above, the optimality condition for cooperatively-set taxes with symmetric countries reads as follows:

$$
\frac{\partial W(t, t^*)}{\partial t} + \frac{\partial W^*(t^*, t)}{\partial t} = 0,
$$

---

5 This statement requires the assumption that $\partial \alpha / \partial t < 0$ globally. Alternatively, it is a local result for small asymmetries across countries: there a relevant condition is satisfied due to the second order condition in the symmetric case.

6 One can show that this assumption is not automatically granted, however, it holds when $\beta$ is sufficiently close to one and the two countries are not too asymmetric.
with a symmetric condition holding for \( t^* \). We now make use of the comparative statics result to obtain:

\[
\frac{\partial W(t, t^*)}{\partial t} + \frac{\partial W^*(t^*, t)}{\partial t} = \lambda R \left[ \alpha - \varepsilon (1 - \kappa) \left( \frac{t}{1 - t} + \frac{1 - 2\kappa}{1 - \kappa} (1 - \beta)(1 - \alpha) \right) \right] \\
- \lambda^* R^* \left[ \varepsilon^* \kappa^* \left( \frac{t^*}{1 - t^*} + 2(1 - \beta)(1 - \alpha^*) \right) \right] = 0,
\]

and imposing the symmetry, it is equivalent to

\[
\lambda R \left[ \alpha - \varepsilon \left( \frac{t}{1 - t} + (1 - \beta)(1 - \alpha) \right) \right] = 0.
\]

This condition is the same as (3.11') in the closed economy. It directly implies that the efficiency margin is still the labor supply elasticity, \( \varepsilon \). Finally, Proposition 3.3 implies that the inequality margin, \( \alpha \), is also unchanged.

### C.3 Derivations and Proofs for Section 3.3

**Agent’s Problem, Open Economy Equilibrium and Comparative Statics:** Solution to agent \( n \)'s problem (3.16) leads to the following optimal allocation:

\[
y_n = \beta(1 - t) \gamma \left( 1 + I_n \tau^{-\frac{\beta}{1 - \gamma}} \right) \frac{1 - \beta}{1 - \gamma} Q \gamma n \gamma^{-\beta}
\]

and \( I_n = \mathbb{I}\{n > n_x\} \), where \( n_x \) is defined by

\[
(1 - t)(1 - \beta / \gamma) Q \gamma \left( 1 + I_n \tau^{-\frac{\beta}{1 - \gamma}} \right) = f_x,
\]

where \( \gamma_x = \left( 1 + \tau^{-\frac{\beta}{1 - \gamma}} \right) \). This results in the following revenue function:

\[
r_n = \beta(1 - t) \gamma \left( 1 + I_n \tau^{-\frac{\beta}{1 - \gamma}} \right) \frac{1 - \beta}{1 - \gamma} Q \gamma n \gamma^{-\beta}
\]

and aggregate revenues given by

\[
R = \beta(1 - t) Q \gamma \int_0^\infty \left( 1 + I_n \tau^{-\frac{\beta}{1 - \gamma}} \right) \frac{1 - \beta}{1 - \gamma} n \gamma^{-\beta} dH(n).
\]

Combining these two expressions yields (3.18) in the text. Finally, the maximized utility of agent \( n \) is given by

\[
U_n = tR + (1 - t)(1 - \beta / \gamma)r_n - I_n f_x,
\]

since agents’ optimality conditions imply \( v(y_n/n) = \beta(1 - t)/\gamma \cdot r_n \).

Balanced trade implies \( Q = R \). Using the expression for \( R \), we can solve for equilibrium \( Q \) as
a function of \( n_x \):

\[
Q^{\frac{\beta - 1}{\beta}} = \left[ \beta(1 - t) \right] \int_0^\infty \left( 1 + I_n \tau^{-\frac{\beta}{\gamma}} \right) \gamma \frac{1 - \beta}{\beta} n^{\frac{\beta}{\gamma} - 1} dH(n). 
\]  \( \text{(C.5)} \)

Together with (C.4), this expression characterizes open economy equilibrium with fixed costs of trade.

These expressions also allow to obtain comparative statics. The log-differentials of (C.4) and (C.5) can be written as:

\[
(1 - \beta) \hat{Q} + \beta \hat{n}_x = -d \ln(1 - t) - (1 - \beta) \zeta \hat{\Upsilon}_x + (1 - \beta/\gamma) f_x, \quad \text{(C.6)}
\]

\[
\hat{Q} + \varepsilon \nu_x \hat{n}_x = \varepsilon d \ln(1 - t) + (1 + \varepsilon) \frac{1 - \beta}{\beta} f_x \int_{n_x}^\infty \left( n/n_x \right)^{\frac{\beta}{\gamma} - 1} dH(n) - \hat{T}_x, \quad \text{(C.7)}
\]

where \( \zeta \equiv \frac{\Upsilon_x^{\frac{1 - \beta}{\beta}}}{\Upsilon_x^{\frac{1 - \beta}{\beta}} - 1} \) and \( \nu_x \equiv \frac{\gamma}{\beta(1 - t)} h(n_x)n_x \).

This fully describes the comparative statics of \( Q \) and \( n_x \) in response to changes in \( t, \tau \) and \( f_x \).

Note that the equilibrium system is stable when \( \varepsilon(1 - \beta) \nu_x < 1 \). For every distribution \( H(\cdot) \), there exists a \( \beta < 1 \) such that the stability condition above is satisfied for almost every \( n_x \). When the stability condition is not satisfied, the equilibrium is locally unstable. If \( \beta \) is too low (a very large CES externality) so that this condition is never satisfied, the model admits only two types of equilibria – when no agent exports or when every agent exports.

Finally, consider the change in agent \( n \)'s utility:

\[
dU_n = (R - r_n) dt + \hat{Q} [tR + (1 - t)(1 - \beta)r_n] + (1 - \beta) \hat{T}_x I_n r_n - \hat{f}_x I_n f_x. \quad \text{(C.8)}
\]

Note that changes in \( y_n \) and \( n_x \) do not have an effect on \( U_n \) by the Envelop Theorem. The reduction in trade costs (fixed \( f_x \) and variable through \( \Upsilon_x \)) directly affects only the agents who already participate in exporting. All indirect effects work exclusively through \( Q \).

\textbf{Proof of Proposition 3.6:} (i) A reduction in \( \tau \) leads to an increase in \( \Upsilon_x \). From (C.6)-(C.7) it follows that \( Q \) increases and \( n_x \) decreases as \( f_x \) falls or as \( \Upsilon_x \) increases. Stability condition \( \varepsilon(1 - \beta) \nu_x < 1 \) is required.

(ii) Greater trade integration implies either a fall in \( \tau \) (an increase in \( \Upsilon_x \)) or a fall in \( f_x \). Tax rate \( t \) is assumed to be held fixed. Therefore, from (C.8), there are three effects on utility – through \( Q \), through \( \Upsilon_x \) and through \( f_x \). By part (i) of this proposition, greater trade integration also results in higher \( Q \). Therefore, all effects on utility are positive, hence all agents gain from trade. Note, however, that only the effect of \( Q \) is common across all agents, while the direct effects of \( f_x \) and \( \Upsilon_x \) are experienced only by the pre-trade-liberalization exporters, who gain disproportionately more from trade liberalization.
Proof of Proposition 3.7: From (3.18) it is apparent that the distribution of revenues is equivalent to the one in Lemma 3.4 when \( n_x = n_{\min} \) or \( n_x = n_{\max} \). This implies that inequality in these two cases is the same as in autarky.

Now compare \( r_n/R \) defined by (3.18) with \( r^a_n/R^a \) defined in Lemma (3.4) in the case when \( n_{\min} < n_x < n_{\max} \). Note that one is a mean-preserving transformation of the other. Specifically, there exist \( \alpha_1 > 1 > \alpha_0 > 0 \) such that

\[
\frac{r_n}{R} = \begin{cases} 
\alpha_0 \frac{r^a_n}{R^a}, & n < n_x, \\
\alpha_1 \frac{r^a_n}{R^a}, & n \geq n_x,
\end{cases}
\]

and \( \int_0^\infty r_n/R \, dH(n) = \int_0^\infty r^a_n/R^a \, dH(n) = 1. \) This immediately implies that the cumulative distribution function of \( r_n/R \) is to the left (right) from that for \( r^a_n/R \) for \( n < n_x \) \( (n > n_x) \). Therefore, the distribution of \( r^a/R^a \) strictly second order stochastically dominates the distribution of \( r_n/R \), which implies

\[
\text{var} \left( \frac{r_n}{R} \right) > \text{var} \left( \frac{r^a_n}{R^a} \right).
\]

By consequence, this implies that income (revenue) inequality increases for \( n_x \lesssim n_{\max} \) (high trade costs) and decreases for \( n_x \gtrsim n_{\min} \) (low trade costs). In addition, one can show by direct computation that \( \text{var}(r/R) \) has an inverted U-shape as \( f_x \) fall from \( \infty \) towards 0.

Similar arguments can be developed for relative utilities when \( n_x < n_{\max} \). The difference is that when \( n = n_{\min} \), the dispersion of relative utilities is still greater than in autarky. To see this, consider

\[
\frac{U_n}{\int_0^\infty U_n \, dH(n)} = \begin{cases} 
\frac{t + (1-t)(1-\beta/\gamma)r_n/R}{t + (1-t)(1-\beta/\gamma)}, & \text{when no agent exports}, \\
\frac{t + (1-t)(1-\beta/\gamma)(r_n - f_x)/R}{t + (1-t)(1-\beta/\gamma) - f_x/R}, & \text{when all agents export}.
\end{cases}
\]

Since in both cases \( r_n/R = r^a_n/R^a \), the former distribution second order stochastically dominates the latter (strictly dominates when \( f_x > 0 \)).

Inequality Margin \( \alpha \): Following the same steps as in the proof of Lemma 3.3, we arrive at the following second order approximation, which becomes precise as the dispersion of \( n \) goes towards zero:

\[
\alpha \approx \rho \cdot \frac{1 - \frac{1}{(1-t)(1-\beta/\gamma)}}{1 + \frac{1}{1-1/\beta/\gamma} - \frac{\pi_x f_x}{(1-t)(1-\beta/\gamma)R}} \cdot \text{var} \left( \frac{r - f_x}{R - \pi_x f_x} \right),
\]

where \( \pi_x \equiv 1 - H(n_x) \) is the fraction of exporting agents. One can show that the variance term is strictly greater in open economy than the autarky variance of \( r^a/R^a \). However, the preceding term is strictly less than 1, which leads to ambiguity in the comparative statics of \( \alpha \) with respect to greater trade openness.

175
Proof of Lemma 3.5: From (C.6)-(C.7), it follows immediately that
\[ \tilde{\varepsilon} = \frac{d \ln Q}{d \ln (1 - t)} = \varepsilon \cdot \frac{1 + \nu_x}{1 - \varepsilon (1 - \beta) \nu_x}, \]
where \( \nu_x \) is as defined above. The stability condition \( \varepsilon (1 - \beta) \nu_x < 1 \) needs to be satisfied. \( \nu_x > 0 \) whenever \( h(n_x) > 0 \), which leads to \( \tilde{\varepsilon} > \varepsilon \).

Equilibrium Characterization with Additional Tax Instruments: The optimality conditions for the agent’s problem (3.19) are
\[
\begin{align*}
\text{for } n < n_x : & \quad \beta (1 - t_d) Q^{1 - \beta} y_n^\beta = \gamma v(y_n/n), \\
\text{for } n > n_x : & \quad \beta (1 - t_x) \Upsilon^{1 - \beta} Q^{1 - \beta} y_n^\beta = \gamma v(y_n/n),
\end{align*}
\]
where \( n_x \) is defined by
\[
(1 - \beta/\gamma) \beta^{\frac{\alpha}{\gamma - \alpha}} Q^{1 - \alpha} \left[ \Upsilon^{\frac{1 - \alpha}{\gamma - \alpha}} (1 - t_x) \gamma - (1 - t_d) \gamma \right] n_x^{\frac{\beta \gamma}{\gamma - \alpha}} = f_x - s. \tag{C.9}
\]
The equilibrium output is given by
\[
Q^{\frac{\gamma - \beta}{\gamma - \alpha}} = \beta^{\frac{\alpha}{\gamma - \alpha}} \left[ (1 - t_d) \gamma - t_x \gamma \int_0^{n_x} n^{\frac{\beta \gamma}{\gamma - \alpha}} dH(n) + (1 - t_x) \gamma \Upsilon \int_{n_x}^\infty n^{\frac{\beta \gamma}{\gamma - \alpha}} \right].
\]
The last two equations allow to solve for equilibrium and obtain comparative statics.

Optimal Entry: The utilitarian welfare is given by
\[
W^u = \int_0^\infty U_n dH(n),
\]
where \( U_n \) is equilibrium utility of agent \( n \) facing the tax system \( (t_d, t_x, s) \). Since \( U_n = c_n - v(y_n/n) \), we have
\[
W^u = \int_0^\infty [c_n - v(y_n/n)] dH(n) = Q - \pi_x f_x - \int_0^\infty v(y_n/n) dH(n),
\]
where we have used the fact total consumption is equal to total output of the final good minus the fraction of output that is spent on fixed costs. Here again \( \pi_x = 1 - H(n_x) \) is the fraction of agents that export, and hence, have to bear the fixed cost. From the definition of \( Q \), we have
\[
Q = \left[ \int_0^\infty \left( 1 + I_n \tau^{\frac{\beta}{\gamma - \alpha}} \right)^{1 - \beta} y_n^\beta \right]^{1/\beta}.
\]
Note that entry subsidy \( s \) does not enter directly the expression for utilitarian welfare, since the distributional effects induced by \( s \) do not impact \( W^u \). The only effect of \( s \) on \( W^u \) is indirect, through its effect on \( n_x \). Thus, we can solve for optimal entry \( n_x \) under the utilitarian welfare, and then recover what it implies for the level of the optimal utilitarian subsidy \( s^o \).
Consider the optimal utilitarian entry given exogenously-set tax rates \( t_d \) and \( t_x \):

\[
\frac{dW_u}{dn_x} = -h(n_x) \cdot \left[ \frac{1}{\beta} Q^{1-\beta} \left( R_x^{1-\beta} y_x^\beta - y_x^\beta \right) - v \left( \frac{y_x^+}{n_x} - v \left( \frac{y_x^-}{n_x} \right) \right) - f_x \right],
\]

\[
- h(n_x) \cdot \left[ Q^{1-\beta} \left[ \left( \frac{1}{\beta} - \frac{(1-t_x)}{\gamma} \right) R_x^{1-\beta} y_x^\beta - \left( \frac{1}{\beta} - \frac{(1-t_x)}{\gamma} \right) y_x^\beta \right] - f_x \right] = 0,
\]

where we have used agents’ optimality conditions which imply \( v(y_n/n) = \beta(1-t_d)/\gamma r_n \) for \( n < n_x \) and \( v(y_n/n) = \beta(1-t_x)/\gamma r_n \) for \( n > n_x \). We now use the optimal \( y_n \)’s to rewrite this optimality condition as

\[
\beta^\frac{\alpha}{\gamma-\beta} Q^{\frac{1}{\gamma-\beta}} \left[ \left( \frac{1}{\beta} - \frac{(1-t_x)}{\gamma} \right) R_x^{\frac{1-\beta}{\gamma-\beta}} - 1 \right] n_x^\frac{\alpha}{\gamma-\beta} = f_x - s,
\]

Contrast this optimality condition with the equilibrium condition for entry (C.9). This allows us to find the level of subsidy \( s^\circ \) which would ensure the optimal utilitarian entry for any given the tax rates \( t_d \) and \( t_x \).

**Proof of Proposition 3.8:** Compare (C.9) with (C.10) in the special case when \( t_d = t_x = -(1-\beta)/\beta \):

\[
\left( \frac{1}{\beta} - \frac{1}{\gamma} \right) \beta^\frac{\alpha}{\gamma-\beta} Q^{\frac{1}{\gamma-\beta}} \left[ R_x^{\frac{1-\beta}{\gamma-\beta}} - 1 \right] n_x^\frac{\alpha}{\gamma-\beta} = f_x - s,
\]

\[
\left( \frac{1}{\beta} - \frac{1}{\gamma} \right) \beta^\frac{\alpha}{\gamma-\beta} Q^{\frac{1}{\gamma-\beta}} \left[ R_x^{\frac{1-\beta}{\gamma-\beta}} - 1 \right] n_x^\frac{\alpha}{\gamma-\beta} = f_x.
\]

These two conditions imply the same amount of entry when \( s^\circ = 0 \), i.e. with \( t_d = t_x = -(1-\beta)/\beta \) the optimal utilitarian subsidy is zero.

Now consider the optimal allocation \( y_n \) for an agent \( n \leq n_x \) from the point of view of utilitarian welfare:

\[
\frac{dW_u}{dy_n} = h(n_x) y_n^{-1} \cdot \left[ (1 + I_n r^\frac{-\beta}{\gamma-\beta}) Q^{1-\beta} y_n^\beta - \gamma v \left( \frac{y_n}{n} \right) \right] = 0.
\]

This condition is equivalent to agents optimality if and only if \( \beta(1-t_d) = \beta(1-t_x) = 1 \). Therefore, the only combination of taxes and entry subsidy that implement the utilitarian optimum is \( t_d = t_x = -(1-\beta)/\beta \) and \( s = 0 \).

**Proof of Proposition 3.9:** Consider now the case when \( t_d = t_x = t \neq -(1-\beta)/\beta \). Then the comparison between (C.9) and (C.10) becomes:

\[
(1-\beta/\gamma)(1-t)[\beta(1-t)]^\frac{\alpha}{\gamma-\beta} Q^{\frac{1-\beta}{\gamma-\beta}} \left[ R_x^{\frac{1-\beta}{\gamma-\beta}} - 1 \right] n_x^\frac{\alpha}{\gamma-\beta} = f_x - s,
\]

\[
(1/\beta - \beta(1-t)/\gamma)[\beta(1-t)]^\frac{\alpha}{\gamma-\beta} Q^{\frac{1-\beta}{\gamma-\beta}} \left[ R_x^{\frac{1-\beta}{\gamma-\beta}} - 1 \right] n_x^\frac{\alpha}{\gamma-\beta} = f_x.
\]
They imply the same amount of entry when
\[
\frac{(1 - \beta / \gamma) (1 - t)}{(1/\beta - \beta(1 - t)/\gamma)} = \frac{f_x - s}{f_x} \quad \Rightarrow \quad \frac{s^0}{f_x} = \frac{1 - \beta(1 - t)}{1 - \beta^2(1 - t)/\gamma},
\]
as stated in the proposition. This expression implies \( \partial s^0 / \partial t > 0 \) since \( \gamma > \beta \).

More generally, when \( t_d \neq t_x \), the condition for \( s^0 \) is more tedious:
\[
\frac{f_x - s^0}{f_x} = \frac{(1 - \beta / \gamma) \left[ \left( \frac{1}{\beta} - \frac{\beta(1-t_d)}{\gamma} \right) \frac{1}{\gamma} - \left( \frac{1}{\beta} - \frac{\beta(1-t)}{\gamma} \right) \frac{1}{\gamma} \right]}{\left[ \frac{1}{\beta} - \frac{\beta(1-t_x)}{\gamma} \right] \left( 1-t_x \right)^{1/\gamma} - \left( \frac{1}{\beta} - \frac{\beta(1-t_d)}{\gamma} \right) \left( 1-t_d \right)^{1/\gamma}}.
\]

One can show that the right-hand side increases in \( t_d \) and decreases in \( t_x \). Therefore, \( \partial s^0 / \partial t_d < 0 \) and \( \partial s^0 / \partial t_x > 0 \).

Finally, marginally reducing \( s \) has only a second order effect on entry, while it has a first order effect on improving the income distribution since it increases the transfer \( \Delta \) and reduces the effective transfer to agents with \( n > n_{x} \). Therefore, with positive inequality aversion, optimal subsidy \( s < s^0 \).
BIBLIOGRAPHY


Nickell, Steven, Luca Nunziata, Wolfgang Ochel and Glenda Quintini (2002), “The Beveridge Curve, Unemployment and Wages in the OECD from the 1960s to the 1990s,” Center for Economic Performance, the LSE.


Tuomala, Matti (1990), Optimal income tax and redistribution, Oxford University Press.
