Inequality, Redistribution and Optimal Trade Policy: A Public Finance Approach*

Roozbeh Hosseini
University of Georgia
roozbeh@uga.edu

Ali Shourideh
Carnegie Mellon University
ashourid@andrew.cmu.edu

March 27, 2018

Abstract

In this paper, we explore the relationship between optimal trade and redistributive policies. In particular, in an environment where international trade affects the relative wages and the reallocation of workers across various sectors is frictional, we study how income taxes and trade policies should be designed in order to balance the efficiency gains from trade with the costs associated with the resulting increased inequality. We show that when personal income taxes can depend on workers sectoral choices, free trade is optimal. However, when personal income tax only depend on income, production must be distorted. We show that these distortions have two general properties: first, they are independent of the network structure of trade and trade elasticities; second, they must take the form of value-added taxes (VAT). Finally, in a quantitative version of our model, we show that sector-specific transfers must be large and absent such taxes, distortion to trade are significant. These results have implications for design of trade assistance programs and trade agreements.

*We would like to thank Laurence Ales, Andy Atkeson, Ariel Burstein, Pablo Fajgelbaum, Chris Sleet, Jonathan Vogel and various conference and seminar participants.
1 Introduction

Gains from international trade are unequal. International trade often rewards productive workers, firms while punishing workers and firms in less productive industries. In this paper, we discuss the implications of inequality and trade on optimal trade and redistributive policies.

Recent evidence, as for example illustrated by Autor et al. (2013), suggests that international trade and global reallocation of production has significantly changed the allocation of labor and inequality of income in the United States.¹ According to Autor et al. (2013), the import competition created by China’s entrance into the World Trade Organization (WTO) has created a significant decline in manufacturing employment in the United States. Under this narrative, the import competition stemming from China’s cheap labor together with the inability of workers to move to sectors/locations with comparative advantage is at the root of this problem. Given such distributional effects of free trade and modern governments’ desire for redistribution, an important question arises: How should redistributive policy be designed to distribute these unequal gains from trade? Perhaps more importantly: Should global trade be distorted in order to achieve this redistribution? In this paper, we set out to answer these questions from a theoretical and quantitative perspective.

We study a multi-country Ricardian model of trade where global production occurs through input-output linkages and workers exhibit imperfect mobility. In this environment, we study the joint determination of optimal trade policy and income taxes under cooperation among countries. Theoretically, we show that when personal income taxes depend on workers’ sectoral choices, free trade is optimal. However, absent sector-specific income taxes, it is optimal to distort trade by using tax and subsidies targeting producers. We show that these tax and subsidies have two key properties: 1. they are independent of the structure of international trade, i.e., input-output linkages and trade elasticities; 2. they must be Value Added Taxes (VAT) thereby allowing firms to deduct cost of intermediate inputs. Finally, we use the theory to quantitatively investigate the importance of such policies. Our main result is that absent sector-specific taxes, optimal distortions to free trade are large and essential.

In our model, production is done competitively across countries where goods produced in one country can be used for production in others. Thus production occurs through a network of input-output linkages across locations. We assume that workers are heterogeneous in terms of their ability to move across production units. The structure of mobility has two layers: 1. set of goods are partitioned into classes across which mobility is imperfect and costly - which we refer to as sectors, 2. within each sector movement of labor is free. We model costly mobility by assuming that workers receive non-pecuniary costs and benefits from working in each sector. We assume that these benefits have a type 1 extreme value distribution. This allows us to tractably characterize the elasticity matrix of workers sectoral choice in response to changes in sectoral wages.²

¹There is a large literature trying to understand the effect of trade on inequality in developed and developing countries. Papers include but are not limited to: Goldberg and Pavcnik (2007), Verhoogen (2008), Helpman et al. (2010), Lagakos and Waugh (2013), Helpman et al. (2017), Artuç et al. (2010), and Caliendo et al. (2017a) among many others.

²This is the so-called compensating differential model of labor mobility a la Rosen (1986). An alternative model of imperfect labor mobility is the model in Roy (1951) where movements of workers are driven by productivity differentials. While we believe that qualitative results between the two models are similar, the calibration of such a
In this model, we study the joint determination of optimal trade policy and redistributive income taxes. We do this by considering two types of personal income taxes: 1. by allowing taxes to depend on a workers’ sectoral choice and their income; 2. by allowing taxes to depend only on workers’ income. In order to be able to fully realize the global gains from trade, we consider the determination of trade under cooperation among countries. That is, we characterize the world pareto frontier when the objective function within each country exhibits redistributive motives. In the first part of the paper, we theoretically investigate the determinants of taxes and trade policy.

When sector-specific income taxes are available, it is never optimal to distort trade. In other words, consumer taxes should be equal for all goods consumed in a country while producer taxes must also be equal across all sectors. Intuitively, sector-specific tax and transfers can allow the government to achieve its desired redistribution. Note that since government cannot have taxes that depend on workers mobility costs, these taxes do affect workers sectoral decision. Nevertheless, any redistribution through consumer and producer taxes can be achieved via direct dependence of income taxes on a worker’s sectoral choice. We provide formulas that characterize these taxes. Due to the tractable nature of imperfect mobility, we can show that transfers must partially compensate workers for changes in their-after tax income adjusted by a measure of the elasticity of sectoral choice.

Absent sector-specific income taxes, distortions to trade are optimal. The idea behind this result is somewhat similar to the tagging idea discussed by Akerlof (1978). When income taxes cannot depend on workers’ sectoral choice, the classic production efficiency result of Diamond and Mirrlees (1971) does not hold and consumer and producer taxes are used to deliver partial redistribution. We go further by characterizing the key properties of these optimal taxes. First, we show that the main tool for redistribution is differential taxes on producers. We show that these taxes are only dependent on the elasticity matrix of labor supply and distribution of income in each country. They are, in fact, independent of the general equilibrium forces affecting prices. In other words, optimal producer taxes are directly independent of the structure of global trade (world input-output linkages) as well as elasticity of trade to trade costs. Intuitively, this is because when the set of consumer and producer taxes are rich enough and are all at their optimum levels, prices are optimal. As a result any changes in them is second order for welfare. In other words, producer taxes can be found by fully ignoring the general equilibrium effect of changes in producer taxes.

Second, we show that optimal producer taxes take the form of a Value Added Tax (VAT) under which firms must be allowed to deduct the cost of intermediate input from their tax bill. VAT taxes leave the intermediate good choices of firms undistorted since such distortions do not help government redistribute resources across workers. This has few stark implications. First, tariffs are never optimal and should not be allowed. Second, policy proposals such as border tax adjustments (see Auerbach (2017) for a description) whereby firms deduct exports while they are taxed for imports are not optimal.

Finally, we use a quantitative version of our model to investigate the magnitude of the optimal taxes mentioned above. To do so, we consider a version of Eaton and Kortum (2002) model ex-
tended by Caliendo and Parro (2015) to allow for trade in intermediate goods and an input-output structure. There are iceberg trade costs that we treat as technological parameters. In this model, within each sector there is a continuum of varieties whose output can be purchased by firms in each country. Similar to Caliendo et al. (2017a), we assume that workers are perfectly mobile across varieties within a sector while their mobility across sectors is modeled as the compensating differentials model. We use the World Input-Output table to calibrate this model using a sample of 39 countries and 10 broad sectors. In this model, VAT taxes increase unit cost of production while they reduce wages by reducing labor income. However, since labor is imperfectly elastic, the decline in wages does not offset the increase in unit costs. In order to correctly measure the resulting distortions to trade, we extend the analysis in Caliendo et al. (2017b) to measure trade distortions in presence of imperfect mobility and elastic labor supply.

We consider two quantitative exercises. First, we allow for income taxes to depend on workers’ sectoral choice. We show that when this is the case, such transfers must be large; up to 10 times workers’ incomes. This is mainly because the incentive effect of these transfers are limited due to a low elasticity of labor supply. This exercise highlights the importance of programs such as Trade Adjustment Assistance in redistributing gains from trade.

Next, we consider optimal trade and redistributive tax policies absent sector-specific income taxes. We highlight three main results. First, we show that trade distortions are large. As we show, this is mainly due to the high variation in wages within a sector and across countries. One of the main forces that determine the degree to which sectors should be taxed or subsidized is their level of income within the country. As we show, due to the significant variation of this across countries, optimal VAT taxes vary significantly and as a result they significantly distort trade.

Second, we show that trade wedges in manufacturing are lower than other sectors. Similar to the above, this is mainly driven by the fact that variation in manufacturing wages are lower than those of other sectors. Input-output linkages and low trade costs are a key behind this result. Note that since trade costs are low in manufacturing, the variation of manufacturing prices across countries is small. This means that the variation in unit costs of production are smaller across countries and so is the variation in trade shares and as a result labor income or wages. Finally, we show that when optimal VAT taxes are compared to the solution of optimal income taxation under free trade, the welfare gains for all countries are positive and significant, from 1%-5% in the aggregate. For low income tradable sectors, particularly Hotels & Restaurants, gains are much larger – average around 15% - while manufacturing often has welfare losses – around 1%-5%. This suggests that gains from optimal VAT taxes are large.

We think of our exercise as having two main substantive implications. First, it highlights the importance of tax and transfer programs that try to compensate workers for losing their jobs to import competition. While these programs exists; Trade Adjustment Assistance program (and to some extent the disability insurance program) in the United States and European Globalization Adjustment Fund in the European Union. They are extremely limited in terms of their coverage, magnitude and funding. Our paper highlights the importance of such programs and the need to provide a comprehensive design – a task that we are undertaking in an accompanying paper. Second, our paper highlights the possibility of adjusting trade agreements in order to allow for VAT subsidies. While implementation of such rules would be difficult, our theoretical result offer a roadmap for designing these policies. In particular, since at the optimum VAT subsidies should only depend on the income distributions across the sectors, perhaps this can be viewed as a
guideline for allowing such subsidies.

1.1 Related Literature

Our paper builds on several strands of literatures in public finance and international trade. The public finance literature has classically studied the problem of the optimal design of direct and indirect taxes. This dates back to the work by Ramsey (1927) as well as later work by Diamond and Mirrlees (1971), Atkinson and Stiglitz (1976) and Deaton (1981). The seminal work by Diamond and Mirrlees (1971) has established the optimality of production efficiency under constant returns to scale and a rich tax structure. While later studies such as that of Naito (1999) have shown that departures from production efficiency is necessary when the tax structure is incomplete, the precise determinant of optimal distortions to production are still unexplored. In fact, Naito (1999) argues that tariffs can be welfare improving.\footnote{Saez (2004) argues that in the long-run workers switch to sectors with higher wages and thus there is no need for distortions to production. In a world where the technology frontier keeps expanding, it is hard to imagine that such differentials diminish even in the long-run.} To our knowledge, our paper is one of the few papers that studies the determinants of optimal producer taxes in an economy with incomplete direct taxes - when income taxes do not depend on workers’ sectoral choice. We contribute to this literature by showing that at the optimum, production tax/subsidies are independent of the structure of trade, and the technology frontier. They, however, depend on labor supply elasticities and the distribution of income.

Our paper is also related to a more recent literature on optimal taxation in trade and spatial models: papers such as Lyon and Waugh (2017), Fajgelbaum and Gaubert (2018), Ales and Sleet (2017). Lyon and Waugh (2017) study an optimal income taxation problem in a small open economy where prices move exogenously and workers move across sectors/islands. Our focus instead is on a more general class of policy instruments namely we allow for policies that distort trade. In Fajgelbaum and Gaubert (2018) and Ales and Sleet (2017) the role of taxes are to correct externalities. We relate to this literature by characterizing optimal policy when the tax function is incomplete - as opposed to the existence of spatial externalities.

While our paper is somewhat related to the literature on optimal trade policy - see papers by Bagwell and Staiger (1999), Opp (2010), Costinot et al. (2015), Beshkar and Lashkaripour (2017), and among others, our paper is different in that this literature often assumes strategic motives between the government as the main determinant of trade policy. As a result, trade policy helps governments manipulate their terms of trade in their favor. In our setup, this motive is shut down intentionally to focus on optimal policy under cooperation. Thus our paper is mainly related to the design of trade agreements and discussion of the type industrial policies allowed under such agreements. In the trade literature, our paper is also related to the work by Dixit and Norman (1980) and Dixit and Norman (1986) who show how to redistribute gains from trade using proportional taxes. The key assumption in their work is the factors are supplied inelastically and as a result such taxes are non-distortionary. In our model, labor is elastic – due to choice of effort and imperfect mobility. This gives rise to a meaningful trade-off between efficiency cost of tax policies and redistributive gains.

Finally, our paper is also related to the growing literature on characterizing allocation of production, prices and distortions in models with input-output linkages. Examples include but are
not limited to Caliendo et al. (2017b), Baqaee (2016) and Liu (2017). Our contribution to this literature is to define the magnitude of distortions to trade in the presence of imperfect factor mobility as well as characterizing optimal wedges that balance these distortions with redistribution.

The rest of the paper is organized as follows: in section 2 we describe a general model of trade with imperfect mobility, section 3 describes and characterize the optimal policy problem in such a model and in section 4, we discuss the determinants of such policies. In section 5, we work with an applied framework and describe how to measure wedges in this environment. In section 6, we present our quantitative results and we conclude in section 7.

2 A Model of Trade with Imperfect Mobility

In this section, we provide the basic framework of analysis in this paper. This framework is based on an economy where movement of labor across sectors is costly, i.e., workers need to incur costs for choosing the sector. As we show, we use a fairly general model of trade that can encompass many of the trade models used in the literature as its special case.

Geography. We consider a model where production and consumption occur in $N_C$ countries; each country is represented by $c \in \{1, \cdots, N_C\}$. In each country, a unit continuum of workers live who are potentially heterogeneous with respect to their working opportunities, as we clarify below.

Production. There are a total of $N$ goods produced globally. In particular, suppose that each good, $i \in \{1, \cdots, N\}$ can be produced in country $c$ according to the production function

$$Y^c_i = G^c_i \left( L^c_i, \{Q^c_{ij}\}_{j=1}^{N} \right)$$

where $L^c_i$ is total effective units of labor in this sector, and $Q^c_{ij}$ is the total amount of good $j$ used in production of good $i$ in country $c$. We assume that the above production function exhibits constant returns to scale with respect to all factors. Moreover, producers of each good are price takers - they take as given the prices of their products, their inputs and wages of their workers.

Workers and structure of mobility. In each country $c$ there is a unit continuum of workers. They are assumed to have preferences over a vector of consumption goods $x = (x_1, \cdots, x_N)$ where $x_i$ is the consumption of good $i$. Aside from choosing a consumption bundle $x$, workers can work in firms that produce a good $j$ and choose how much effort to put in their work. The structure of mobility in country $c$ is as follows: The set of goods are partitioned into $N^c_j$ classes - what we refer to as sectors - with each class in country $c$ given by $J^c_j$ for $j \in \{1, \cdots, N^c_j\}$, i.e., $\bigcup J^c_j = \{1, \cdots, N\}$ and $J^c_j \cap J^c_j' = \emptyset$. We assume that within each sector workers can move freely across production units. However, movements of workers across sectors are costly and imperfect. In particular, we assume that each worker draws a vector $\nu = (\nu_1, \cdots, \nu_{N^c_j})$ where $\nu_j \in \mathbb{R}$ is the utility benefit (if negative it is a cost) if a worker works in class $j$. Specifically, the

---

4We can also work with a version of our model where countries are heterogeneous in their population in which case, $\sum_i \mu^c_i = M^c$, $\forall c \in \{1, \cdots, N_C\}$. 
utility of a worker that consumes a bundle \( \mathbf{x} \) and supplies \( \ell \) units of labor and works in sector \( j \) is given by
\[
[U^c(\mathbf{x}) - v^c(\ell)] e^{\nu}, \forall j \in \{1, \cdots, N^j\}
\] (1)

We assume that workers’ preferences satisfy the following properties:

**Assumption 1.** Preferences of workers in country \( c \) satisfy the following:

1. The utility from consumption, \( U^c(\mathbf{x}) \) is strictly concave, increasing and positive.

2. Disutility of effort \( v^c(\cdot) \) is a strictly convex and strictly increasing function and satisfies
\[
(v^c)'(0) = 0, v^c(0) = 0,
\]

3. Sectoral preferences \( \nu \) are independently drawn across sectors and have a Gumbel distribution, i.e.,
\[
\nu_j \sim \text{Gumbel}\left(\frac{\log \mu^j}{\sigma}, \frac{1}{\sigma}\right),
\]
where \( \sigma, \mu^j > 0 \). Moreover, \( \sum_{j=1}^{N^j} \mu^j = 1 \).

The above assumptions about the structure of worker mobility is made to allow for a somewhat general structure yet maintain tractability. Since mobility across production units in each sector \( J^c_j \) is free, wages of workers must be equal. Moreover, as Assumption 1 states, \( \nu_j \) is distributed according to a type-1 extreme value distribution and satisfies
\[
\Pr(\nu_j \leq z) = e^{-\mu^j_j e^{-\sigma z}} = H^c_j(z)
\]

As a result, many of the tractability results from extreme value theory carries through to our analysis. For example, the fraction of workers in sector \( j \) is given by
\[
\Lambda^c_j = \frac{\mu^c_j (\mu^c_j)^\sigma}{\sum_{k=1}^{N^j_k} (\mu^c_k)^\sigma},
\] (2)

where
\[
u^c_j = U^c(\mathbf{x}^c_j) - v^c(\ell^c_j)
\] (3)
is the utility from consumption and labor supply (intensive margin) for a worker who works in sector \( j \). This tractability together with the separability of workers’ sectoral choice and consumption and labor supply implies that allocations of consumption and labor supply are independent of \( \nu \). Moreover, sectoral choice is simply represented by (2).

The parameter \( \sigma \) captures the degree of mobility of workers across sectors. When \( \sigma = 0 \), the fraction of workers in sector \( j \) is given by \( \mu^c_j \). That is, workers are attached to sectors and cannot move across them. At the other extreme where \( \sigma = \infty \), only sectors \( j \) in which \( u^c_j = \max_k u^c_k \) holds will have non-zero measure of workers. That is, wages must be equated across sectors - those that produce.
Markets. We assume that for each good $i \in \{1, \ldots, N\}$, there is a competitive market where producer of this good in country $c$ sell their product while producers of other goods and consumers purchase their demand from this market. The market price of good $i$ is given by $\hat{p}_i$.

Seemingly, the assumption above suggests that all goods are traded in all countries and trade is free and therefore, our specification cannot handle trade costs. As we will clarify later, this literal interpretation of the model is at best incomplete. In particular, we argue that our specification is general enough and can encompass a model with iceberg trade costs a la Samuelson, non-tradable goods, autarky, etc.

In addition to markets for goods and services, we also assume that in each country there are competitive labor markets. Since labor mobility within in each sectors is not costly, there is a wage associated with working in sector $j$ in country $c$ which is given by $w_j^c$, $\forall j \in \{1, \ldots, N^c\}$, $c \in \{1, \ldots, N_C\}$.

Governments and Policies. We assume that in each country $c$, there is a government that has access to income taxes, producer taxes and consumer taxes. In particular, we assume that government in country $c$ imposes an affine tax on income given by $T^c(z) = \tau^c z - T^c$ where $z$ is a worker’s income. Moreover, upon consumption of good $i$ by a consumer in country $c$, the government imposes an ad-valorem tax rate given by $t_{x,c}^i$. Finally, a producer of good $i$ in country $c$ faces a tax rate of $t_{p,c}^i$ on its revenue while it faces a tax rate $t_{p,c}^{ij}$ on its purchases of intermediate inputs of good $j$. For now, we assume that governments do not have any expenditure and therefore must have a balanced budget.

Given the above structure for the global economy, a competitive equilibrium given governments’ policies can be defined as: (i) consumption and leisure allocations by each type of worker, $\{x_j^c, \ell_j^c\}_{j=1}^{N^c}$, and sectoral choices $\Lambda_j^c$, (ii) production in each sector, $L_i^c, \{Q_{ij}\}_{j=1}^{N^c}$, (iii) vector of prices, $\hat{p} = (\hat{p}_1, \ldots, \hat{p}_N)$, and wages, $\{w_i^c\}_{i=1}^{N}$, such that:

1. Workers in each country $c$ maximize - given prices and government policies:
   \[
   x_j^c, \ell_j^c \in \arg \max_{x,\ell} U^c(x) - v^c(\ell)
   \]
   subject to
   \[
   \sum_{i=1}^{N} \hat{p}_i (1 + t_{x,c}^i) x_i \leq w_j^c \ell (1 - \tau^c) + T^c
   \]

2. Fraction of workers in sector $j$ is given by
   \[
   \Lambda_j^c = \frac{\mu_j^c (u_j^c)^\sigma}{\sum_{k=1}^{N_j} \mu_k^c (u_k^c)^\sigma}
   \]

\footnote{The use of affine income tax instead of an arbitrary income tax function is for ease of exposition. As we show in the Appendix, including an arbitrary non-linear income tax will significantly complicate the analysis while the main insights remain the same.}
3. Firms maximize - given prices and government policies:

$$L^c_i, \{Q^c_{ij}\} \in \arg \max_{L, Q_{ij}} \hat{p}_i \left(1 - t^p_{c,i}\right) G^c_i \left(L, \{Q_{ij}\}_{j=1}^N\right) - w^c_i L - \sum_{j=1}^N \left(1 + t^p_{ij}\right) \hat{p}_j Q_{ij}$$

4. Government budget constraint holds:

$$\sum_{c=1}^C C^c = \sum_{c=1}^C \sum_{i} \hat{p}_i t^x_{c,i} Y^c_i + \sum_{c=1}^C \sum_{j=1}^N \sum_{i} \hat{p}_j t^p_{ij} Q^c_{ij} + \sum_{c=1}^C \sum_{j=1}^N \sum_{i=1}^{N^c_j} \hat{p}_i t^x_{c,i} \Lambda^c_j x^c_{ji} + \sum_{c=1}^C \sum_{j=1}^N \sum_{i=1}^{N^c_j} \Lambda^c_j w^c_j \ell^c_j$$

5. Markets clear:

$$\sum_{c=1}^C \sum_{j=1}^{N^c_j} \Lambda^c_j x^c_{ji} + \sum_{c=1}^C \sum_{k=1}^N Q^c_{kj} = \sum_{c=1}^C G^c_i \left(L^c_i, \{Q^c_{ij}\}\right), \forall i \in \{1, \ldots, N\}, \forall c \in \{1, \ldots, C\}$$

$$\sum_{i \in \mathcal{I}_j^c} L^c_i = \Lambda^c_j \ell^c_j, \forall j \in \{1, \ldots, N^c_j\}, \forall c \in \{1, \ldots, C\}$$

Note that in our definition of equilibrium, we have a single budget constraint for the governments. In other words, our equilibrium concept allows for transfers across countries. As a result and in general trade can be unbalanced. In our environment, imposing budget balance introduces additional distortions which we want to abstract from. In our quantitative exercise, we study the role of these inter-government transfers.

**Generality of the Model** As we have claimed before, our model encompasses various versions of the models that are popular in the international trade literature. Here, we provide some remarks to illustrate this:

1. Iceberg Trade Costs: While we have assumed that markets are competitive, it is possible to map a trade model with iceberg costs into the setup above. In particular, consider a model with iceberg trade costs. In order to map such a model into our setup, we first extend the set of goods so that each good is produced in a different country. That is, we define an extended set of goods \(\{1, \ldots, N\}\) and partition it into goods produced in each country given by \(\mathcal{I}^c\) where

$$\bigcup_{c=1}^C \mathcal{I}^c = \{1, \ldots, N\}, \mathcal{I}^c \cap \mathcal{I}^{c'} = \emptyset.$$ 

Let \(c^* (i)\) be the country in which good \(i\) is produced. Suppose that the iceberg cost of shipping good \(i\) from country \(c\) to \(c'\) is given by \(d_{i,c,c'}\) with \(d_{i,c,c'} = 1\) for all \(c\) and \(i\) in \(\mathcal{I}^c\). We then refer to \(x^c_{i,j}\), consumption of good \(i\) in country \(c\), as this consumption including the iceberg trade cost. Similarly, \(Q^c_{j,i}\) is defined by intermediate input demand of sector \(j\).
in country $c'$ of good which includes the iceberg cost $d_i^{c,c'}$. Given this relabelling, utility functions and production functions in any country $c'$ are given by

$$U^{c'}(x) = U^{c'}\left(\frac{x_1}{d_1^{c'(1),c'}}, \frac{x_2}{d_2^{c'(2),c'}}, \cdots, \frac{x_N}{d_N^{c'(N),c'}}\right)$$

$$Y_i^{c'} = G_i^{c'}\left(L_i^{c'}, \left\{ \frac{Q_{ij}^{c'}}{d_{ij}^{c'(j),c'}} \right\}\right), \forall i \in I^{c'}$$

Note that in the above we relied heavily on the extension of the set of goods so that each good is produced in a unique country. This is often possible in neoclassical models of trade. For example, in an Armington model, each country produces a set of differentiated products that are imperfect substitutes and hence, no extension is necessary. In Ricardian models, for example that of Dornbusch et al. (1977) or Eaton and Kortum (2002), goods produced in each country are perfect substitutes and thus the set of goods can be easily relabeled so that certain class of goods are perfectly substitutes in utility and production functions. For example, if the set of goods are cloth and wine and the set of countries are England and Portugal, we can extend the set of goods to be English Wine, Portuguese Wine, English Cloth and Portuguese Cloth. In this case, the wine and cloth produced in each country are perfect substitute for all consumers and producers.

Note that it is important the type of trade costs are technological as opposed to government imposed. In particular, the goal of the paper is to find the optimal government imposed trade costs given the technological ones.

2. Non-tradable goods: It is fairly straightforward to see that non-tradable goods are captured in our model by requiring certain goods to be produced and used (by consumers or producers) in one country. That is, they do not enter the utility of workers and production function of firms in other countries.

Given the above discussion, it is easy to see that various neoclassical models of trade such as the Armington model, Ricardian models of Dornbusch et al. (1977) and Eaton and Kortum (2002) are special cases of our general model.

**Generality of the Policies** Despite the focus of our paper on trade policy, we have not explicitly introduced tariffs in the above environment. This is in part because tariffs are a special case of consumer and producer taxes. To see this, suppose that the government in country $c$ imposes an ad-valorem tariff, $\tau_i^c$, on good $i$ which is on net being imported into country $c$. This implies that if the (international) market price of good $i$ is $\hat{p}_i$, the price faced by producers and consumers in country $c$ is $\hat{p}_i (1 + \tau_i^c)$. In other word, a tariff imposes a tax on the use of a good - by consumer and producers, while at the same time, it imposes a subsidy on the production/making of good $i$ in country $c$. That is, under this tariff we have:

$$t_i^{x,c} = \tau_i^c, t_i^{p,c} = -\tau_i^c, t_{ji}^{p,c} = \tau_j^c.$$

The above example illustrates that the indirect (commodity) tax policies considered in our model include the possibility of tariff. Another way to see this is to realize that in the presence
of tariffs, $\tau^c_i$, the consumer price vector in country $c$ is given by
\[
\hat{p}^c = ((1 + \tau^c_1)(1 + t^c_1,\hat{p})_1, \ldots, (1 + \tau^c_N)(1 + t^c_N,\hat{p})_N)
\]
while the set of after tax/tariff prices faced by producers in sector $i$ is given by
\[
(1 - t^p,c_i)(1 + \tau^c_i)(\hat{p}_i, (1 + t^p,c_j)(1 + \tau^c_j)^N_{j=1}.
\]
Since consumers and producers only care about after tax/tariff prices, it is possible to redefine consumer and producer taxes so that the resulting equilibrium allocations and prices will be the same as the ones in the economy with tariffs. Because of this equivalence of tax systems, we focus our attention to indirect tax policies that only involve consumer and producer taxes.

Finally, while we often focus on the case where transfers and taxes targeted to workers in a specific sectors are absent - as they are in the above specification - we also consider the case where such transfers/taxes are present. When lump-sum taxes/transfers are available we represent them by $T^c_i$. As we will show, the presence of such taxes and transfers are required for optimality of free trade.

### 2.1 Trade and Inequality under Laissez-Faire

Before turning to optimal trade policy, it is useful to illustrate the effect of opening up to trade on inequality. To do so, we describe the distribution of income with and without free trade in a Laissez-faire economy - an economy without government. We do this by assuming that the utility function is CES in final goods in each country, there are no intermediate inputs and all goods are tradable without trade costs, i.e.,

\[
U^c(x) = \left[ \sum_{i=1}^{N} \alpha_i x_i^{1-\gamma} \right]^{\frac{\gamma}{\gamma-1}}
\]
for $\gamma > 0$ and

\[
Y^c_i = A^c_i L^c_i.
\]
In addition, we assume that $v^c(\ell) = \frac{\varepsilon}{1+\varepsilon}\ell^{1+1/\varepsilon}$ where $\varepsilon$ is the Frisch elasticity of labor supply. Furthermore, we assume that each sector is consisted of only one good. The rest of the setup is identical to the one described above.

We first start with a closed economy. In the appendix, we show that equating supply and demand for good $i$ in country $c$ implies that

\[
z^c_{i,AUT} \propto \left( \alpha_i^\gamma (A^c_i)^{\gamma-1} \right)^{\frac{1+\varepsilon}{\gamma+\varepsilon+\sigma+\varepsilon}} \mu^c_i
\]
where $z^c_{i,AUT}$ is the income for a worker in sector $i$ and country $c$ under autarky.

In an open economy, a similar argument can be used to show that

\[
z^c_{i,T}(\theta) \propto \left( \frac{\alpha_i^\gamma (A^c_i)^{\gamma+\varepsilon+\sigma+\varepsilon} \mu^c_i}{\sum_{c'=1}^{N} \zeta^c_i \mu^c_i (A^c_i)^{(1+\varepsilon)(1+\sigma)}} \right)^{\frac{1+\varepsilon}{\gamma+\varepsilon+\sigma+\varepsilon}}
\]
\footnote{This particular way of defining tariffs works as a tax on imports and subsidy on exports.}
where $\zeta_c$ can be thought of as the weight of the country in average global productivity of good $i$. We can take log of the formula for income under free trade and autarky (closed economy), we can write

$$\log z_{i,T}^c - \log z_{i,AUT}^c = \kappa - \frac{1 + \varepsilon}{\gamma + \varepsilon + \sigma + \varepsilon \sigma} \log \left( \sum_{c' = 1}^{N_C} \frac{\zeta_{c'} \mu_{i'} (A_{c'}^i)^{(1+\varepsilon)(1+\sigma)}}{\mu_i (A_{i}^c)^{(1+\varepsilon)(1+\sigma)}} \right).$$

If we assume $N_C = 2$, then the above can be written as

$$\log z_{i,T}^c - \log z_{i,AUT}^c = \kappa - \frac{1 + \varepsilon}{\gamma + \sigma + \varepsilon + \varepsilon \sigma} \log \left( \frac{\zeta_c + \zeta_{c'} \mu_{i'} (A_{c'}^i)^{(1+\varepsilon)(1+\sigma)}}{\mu_i (A_{i}^c)^{(1+\varepsilon)(1+\sigma)}} \right), \ c' \neq c.$$

In words, the change in income between autarky and free trade is an increasing function of $\mu_i (A_{i}^c)^{(1+\varepsilon)(1+\sigma)}$. This is effectively a measure of comparative advantage for sector $i$ in country $c$. In words, it measures the productivity of sector $i$ in country $c$ relative to country $c'$ adjusted by the relative size of the labor force in that sector. Hence, workers in sectors that have a high modified comparative advantage experience the highest increase in their income. If these sectors happen to be sectors with high initial income, then this implies that free trade leads to higher inequality. In other words, if comparative advantage is positively correlated with income, then income inequality is higher under free trade.

This example illustrates how the model captures the increase in income inequality coming from opening to free trade. The effect of trade on inequality is particularly magnified when the competing sector in the opposite country has a large mass of workers attached to it - captured by $\mu_i$. In essence, this model has features similar to a Heckscher-Ohlin model where workers in each sector are effectively a different factor of production and this result is a version of Stopler-Samuelson Theorem. In what follows, we study optimal trade policy to mitigate the effect of trade on inequality.

### 3 The Optimal Taxation Problem

The model developed in the previous section clarifies how trade can lead to higher income inequality. It, furthermore, makes precise the type of policies that can be used by the governments in each country. In this section, we describe the optimal policy problem faced by these governments under cooperation and derive formulas that govern the behavior of optimal trade policy.

Starting with the government in each country $c$, we assume that the government evaluates the allocation of resources across workers according to a social welfare function given by

$$\int W \left( \max_{j \in \{1, \ldots, N_j^c\}} u_j^c e^{v_j} \right) dH^c (\nu)$$

where $W (\cdot)$ is a concave and increasing function and $u_j^c$ is the utility of a worker in sector $j$ associated with consumption and effort as described in evaluated according to preferences in (3). The curvature of $W (\cdot)$ represent the government’s desire for redistribution. Moreover, $H^c (\nu)$ is
the joint distribution of $\nu$. In the appendix, we show that the above social welfare function can be written as

$$\sigma U^c \int e^{-u^c z - \sigma} z^{-1 - \sigma} W(z) \, dz$$

where $U^c = \sum_j \mu_j^c (u_j^c)^\sigma$. As the above shows, $U^c$ is a sufficient statistic for welfare. The above also implies that somewhat surprisingly - and due to the properties of the extreme-value distribution. As a result, redistributive motives of the government within the country is independent of the function $W$ and solely depends on $\sigma$.

Our main assumption about the determination of policy is that it is determined under cooperation. In other words, we want to consider trade agreements whereby governments negotiate with each other on coordination of their tax policy. We assume that the outcome of this negotiation process is efficient in that it is equivalent to maximizing a weighted average of aggregate welfare in each country. In other words, the optimal trade policy problem is given by

$$\max_{\{p_{ij}^c, t_{ij}^c, T^c, \tau^c\}_{j=1}^{N_c}} \sum_{c=1}^{N_c} \lambda_c \int W \left( \max_j u_j^c e^{\nu_j^c} \right) \, dH^c(\nu)$$

where $u_j^c$ is a utility profile arising from a competitive equilibrium of the economy described above given the policy choices $\{p_{ij}^c, t_{ij}^c, T^c, \tau^c\}_{j=1}^{N_c}$.

Given this objective for the government, the optimal taxation problem for the government is to maximize social welfare given that fact that the utility profile $u_j$ arises from a competitive equilibrium as defined above. As in Atkinson and Stiglitz (1976) or Diamond and Mirrlees (1971), it is often useful to write this problem in terms of (after-tax) consumer and producer prices defined by

$$q_i^c = \hat{p}_i (1 + t_i^{x,c})$$

$$p_i^c = \hat{p}_i (1 - t_i^{p,c}), p_{ij}^c = \hat{p}_j (1 + t_{ij}^{p,c})$$

where $q_i^c$ is the after-tax price paid by a worker/consumer for good $i$ in country $c$ while $p_i^c$ and $\{p_{ij}^c\}_{j=1}^{N_c}$ are the after-tax prices faced by producers of good $i$.

Although the task of characterizing optimal taxes seem somewhat cumbersome and difficult, the notion of after-tax prices defined above can help us simplify the problem. To see this, note first that in any equilibrium and for a good $i$ in sector $j$, wages must satisfy $w_{ij}^c = p_{ij}^c e^{G_{ij}} - \text{from the optimality condition of the firms. Additionally, we can realize that the vector of consumer prices $q^c$ solely determines the choice of consumption allocation by each consumer. To see this consider the problem for a hypothetical consumer in country $c$ that has a utility function $U^c(x)$, wealth $I$ and faces a price vector $q$. Let $x^c(q; I)$ be the vector of consumption demand for an individual in country $c$ with wealth $I$ who faces a price vector $q$ and let $V^c(q; I)$ be the associated indirect utility function. Then, in any competitive equilibrium as defined above, we must have that $x_j^c = x^c(q; (1 - \tau) w_{ij}^c \ell_j + T^c)$.

Finally, producer prices affect the production decision of firms. However, since the set of taxes on producers is general enough, we can show that any demand vector for intermediate inputs by a firm $i$ in country $c$, $\{Q_{ij}^c\}_{j=1}^{N_c}$, can be implemented by carefully choosing the vector of producer taxes, $\{t_{ij}^{p,c}\}_{j=1}^{N_c}$. Using these observations, the optimal policy problem can be simplified as it is stated in the following proposition:
**Proposition 2.** Any solution of the optimal taxation problem - and its associated consumer and producer prices - must solve the following optimization problem

$$
\max_{p^c,q^c,\{\ell^c\},w^c,\tau^c,T^c,\{Q^c_{ij}\}} \sum_{c=1}^{N_C} \lambda^c \int \max_j W \left( [V^c(q^c; (1 - \tau) w^c_j \ell^c_j + T^c) - v^c(\ell^c_j)] e^{\nu_j} \right) dH^c(\nu)
$$

(P)

subject to

$$
\sum_{c=1}^{C} \sum_{j=1}^{N_j} \Lambda^c_j \ell^c_j \left( q^c_j; (1 - \tau^c) w^c_j \ell^c_j (\theta) + T^c \right) + \sum_{c=1}^{C} \sum_{k=1}^{N} Q^c_{ki} = \sum_{c=1}^{C} G_i^c \left( L^c_i, \{Q^c_{ij}\}_{j=1}^{N_i} \right), \forall i
$$

$$
\ell^c_j \in \arg \max_{\ell} \left( q^c_j; (1 - \tau^c) w^c_j \ell + T^c \right) - v^c(\ell)
$$

$$
\Lambda^c_j \ell^c_j = \sum_{i \in J^c_j} L^c_i
$$

$$
\frac{\mu^c_j \left[ V (q^c_j; (1 - \tau) w^c_j \ell^c_j + T^c) - v^c(\ell^c_j) \right]^{\sigma}}{\sum_{k=1}^{N_j} \mu^c_k \left[ V (q^c_k; (1 - \tau) w^c_k \ell^c_k + T^c) - v^c(\ell^c_k) \right]^{\sigma}} = \Lambda^c_j
$$

$$
p_i^c \frac{\partial G_i^c \left( L^c_i, \{Q^c_{ij}\} \right)}{\partial L} = w^c_j, \forall i \in J^c_j
$$

Conversely, any solution of the above optimization problem can be used to construct a solution of the optimal taxation problem.

Proof of this proposition is relegated to the Appendix.

As the above proposition establishes, the only restrictions that competitive equilibrium imposes on allocations is feasibility and optimal choice of labor supply. The idea behind the above proposition is straightforward. Given that consumption is calculated using the demand function of households, households budget constraints must be satisfied. Furthermore, since feasibility is satisfied, these conditions imply that government budget constraint must hold. Finally, given that consumption comes from demand function and optimality of choice of labor supply, the consumption and labor supply allocation must satisfy optimality for workers.

An investigation of the above problem clarifies the role of commodity taxes. In particular, as it can be seen, producer prices affect the labor supply incentive of workers. As a result, it seems somewhat intuitive that producer taxes can be used to affect incentives to work. It is also important to note that since consumer and producer prices are sufficient to characterize allocations, this implies that actual producer and consumer taxes are indeterminate. This result is standard as it is typically irrelevant whether to tax consumer or producers. Optimal allocations however determine the ratio $\frac{1 - \ell^p_i}{1 + \ell^p_i} = \frac{p_i}{q_i}$.

As a side note, the above result does not change much in the presence of sector-specific tax and transfers. We can replace the term $T^c$ with $T^c_j$ where $j$ represent the sector in which each worker works.
4 Properties of Optimal Policies

In this section, we discuss qualitative features of optimal redistributive and trade policies. We start our analysis under the assumption that sector-specific income tax and transfers are available. We then assume that such policies are unavailable and derive the implications for tax and trade policy.

4.1 Optimal Policy with Sector-Specific Income Taxes

When governments can use taxes and transfers targeted at workers in a sector, they can achieve better redistribution by relying on such taxes. Note, however, that since a government cannot make such taxes and transfers dependent on workers’ utilities – the realization $\nu$ – such tax and transfers are not lump-sum in that they will affect the allocations of workers among sectors. Nevertheless, we can show that in presence of such tax and transfers, free trade is optimal. That is, producers and consumers must face the same relative prices in all countries. We thus have the following proposition:

**Proposition 3.** When sector-specific tax and transfers are available, optimal taxes must satisfy

$$t_{p,c}^i = -t_{p,c}^j = t^c, t^x_c = t^c.$$

In particular, $t^c = t^x = 0$ is part of a solution to the optimal taxation problem.

The proof of the above proposition is relegated to the Appendix. We show this by considering the relaxed problem where government can allocate goods and effort without any constraint to workers in each sector while it allocates inputs and labor freely to firms. We then show that the solution to this problem equates marginal rate of substitution between consumption of all goods and effort across all countries. It thus can be implemented as an equilibrium allocation with sector-specific income taxes.

The above proposition states that households’ and firms’ decisions are undistorted. To the extent that constant taxes are equivalent to lump-sum transfers for workers and fully non-distortionary for firms, we can simply assume that they are zero.

This proposition is reminiscent of the celebrated production efficiency result by Diamond and Mirrlees (1971). As they show, production must always be efficient under the requirement that the tax structure is rich enough - all goods enjoyed by households can be taxed. In our environment, this requirement manifests itself as existence of sector-specific tax and transfers. In effect, one can think about workers working in different sectors as consumption and to impose a tax on this decision, sector-specific income taxes are required. At the same time, sector-specific taxes allow the government to achieve redistribution without distorting households’ consumption decision and firms’ production decision.

Given that production and trade decisions are undistorted, the redistribution solely occurs through sector-specific taxes and transfers. While determinants of optimal tax and transfers are in general difficult to characterize, it is possible to focus on a simple case often used in the literature. In particular, suppose that the utility function is CES in consumption goods and that

$$v^c(\ell) = \frac{\ell^c}{1+\varepsilon_c} \ell^{1+1/\varepsilon_c}.$$  

In this case, since there are no income effects on labor supply and using the fact that $U^c = \sum_{j=1}^{N_j} \mu^j_c \left( u^j_c \right)^{\sigma}$ is a sufficient statistic for social welfare in country $c$ can be used to characterize optimal transfers. We have the following lemma:
Lemma 4. If consumer and producer taxes are set to zero, optimal sector-specific transfers must satisfy

\[ T_j^c - T_{j'}^c = \frac{1}{\sigma} \left[ \frac{u_j}{V_i^c (\hat{p}; z_j^c + T_j^c)} - \frac{u_{j'}}{V_i^c (\hat{p}; z_{j'}^c + T_{j'}^c)} \right]. \]

Moreover, when \( U^c (x) \) is homothetic and \( v^c (\ell) = \frac{x^c}{1 + \epsilon c} \ell^{1+\frac{\epsilon}{1+\epsilon}} \),

\[ T_j^c - T_{j'}^c = -\frac{z_j^c - z_{j'}^c}{(1 + \epsilon c)(1 + \sigma)} \]

where \( z_j^c \) is the before-tax labor income of a worker in sector \( j \).

The above formula is quite intuitive. Absent workers behavioral response, transfers must fully compensate workers for income differentials, i.e. \( T_j^c - T_{j'}^c = -\frac{z_j^c - z_{j'}^c}{1 + \epsilon c} \). Note that due to disutility of leisure, a unit of income is worth \( \frac{1}{1 + \epsilon c} \) unit of utility. However, this compensation leads to a behavioral response in that workers will switch to sectors with high transfers. Thus, it needs to be divided by \( 1 + \sigma \) in order to adjust for the behavioral response.

One way to reinterpret the above formula is that optimal transfers must partially compensate workers for the differential gains from trade. In fact, this compensation is simply the difference in income under free trade adjusted by elasticity of labor supply – combined intensive and extensive margin. That is the change in inequality caused by trade - or other technological sources - must be compensated but only partially.

While the above formula sheds light on a possible response by a redistributive government to changes in inequality caused by trade, it is somewhat unrealistic. In reality, there is very little sector dependence embedded in personal income tax and transfers in most developed countries. In the United States, for example, there are few Federal programs that compensate workers who lost their jobs to import competition from abroad and indirectly make transfers dependent on workers’ sectors: namely the TAA (the Trade Adjustment Assistance) and the Disability Insurance program.\(^7\) As noted by Autor et al. (2013), disability insurance benefits far outweigh those of the TAA. As Autor et al. (2014) emphasize, the benefits from disability insurance are around 5% of the workers’ previous salary which is much smaller than what the above formula implies for reasonable values of elasticities - we will discuss this in more detail in section **. This evidence points to very little dependence of the personal income tax and transfer system on a worker’s sectoral choice. In what follows, we assume that the personal income tax code is independent of workers sectoral choice and derive properties of optimal taxes.

### 4.2 Optimal Policy without Sector-Specific Income Taxes

When government cannot use taxes and transfers to target workers in a certain sector, they can rely on affine income taxes or use consumer and producer taxes to deliver some redistribution. In this section, we discuss which of these tools will be used in the optimal policy problem. Our main result is that producer prices and consumer prices are not equal. In addition, we provide formulas

\(^7\)TAA provides monetary and training assistance to workers who lost their jobs due to trade while disability insurance allows
for both producer taxes assuming that market prices are given by the lagrange multipliers on the feasibility constraint.

In order to understand the trade-offs involved in determination of producer taxes, consider a perturbation of producer prices $p^c_i$ in the above problem. One the one hand, an increase in producer prices for sector $i$ increases their welfare and output - through its effect on labor supply. On the other hand, an increase in producer prices increases the demand for all goods produced - from workers in sector $i$. Therefore at the optimum, the cost of the increase in resources must be equal to the welfare and output benefit of the increase in producer prices.

In order to understand the response of various allocations to such perturbations, it is useful to define the elasticities of the sectoral choice with respect to changes in wages. Note that in presence of workers sectoral choice, sectoral choice elasticity takes a matrix form since a change in wage of any sector affects the decision of all workers in all sectors. We can thus define

$$\xi_{j,j'}^c = \frac{\partial \Lambda_j^c}{\partial w_{j'}^c} \frac{w_{j'}^c}{\Lambda_j^c}$$

which is the (extensive margin) elasticity of labor supply (number of workers) in sector $j$ to an increase in the wage in sector $j'$. To characterize these elasticities in a tractable manner, we make the following assumption about the workers’ utility functions:

**Assumption 5.** Workers’ utility functions must satisfy $U^c(x) = \left[ \sum_{i=1}^N \alpha_i^c x_i^{1-\frac{1}{\gamma^c_i}} \right]^{\frac{1}{\gamma^c_i-1}}$ and $v^c(\ell) = \frac{\ell^{1+1/\varepsilon^c}}{1+1/\varepsilon^c}$ where $\alpha_i^c \geq 0$ and $\varepsilon^c > 0$.

Assumption 5 implies that workers’ intensive margin of labor supply, $\ell$, does not exhibit any income effect. It, thus, significantly simplifies the exposition of the problem. While it is possible to provide formulas for general utility functions, it is much more cumbersome and is thus relegated to the appendix.

Using Assumption 5, we have (derivation given in the Appendix) that we have

$$\xi_{j,j'}^c = \frac{\sigma}{(1-\tau_j^c) \frac{z_{j'}^c}{1+\varepsilon_j^c} + T_c} \left[ -\Lambda_j^c + 1 [j = j'] \right]$$

(7)

where $z_{j'}^c$ is the (before-tax) income of a worker in sector $j'$ in country $c$. As (7) illustrates, an increase in the wage in sector $j'$ decreases the number of workers in sector $j \neq j'$. Moreover, the extensive margin elasticity is independent of the sector $j \neq j'$. In essence, this is very similar to a CES demand system where own-price elasticities are different from cross-price elasticities but cross-price elasticities are all equal, i.e., $\xi_{j,j'}^c$ is independent of $j$ when $j \neq j'$. Using the above elasticities, we can characterize the response of demand and supply in country $c$ to a perturbation of producer taxes. We thus have the following proposition:
Proposition 6. Optimal producer tax for good $i$ only depends on its sector $J^c_i$ and is given by

$$
\frac{1}{(1 - t^{p,c}_j)} \varepsilon^c = (1 + \varepsilon^c) (1 - \tau^c) + \frac{\sigma (1 - \tau^c)}{z_j(1 - \tau^c)} + T^c \sum_{j'} \Lambda^c_{j'} \left( z^c_{j'} - z^c_j \right) \left( \frac{z^c_j}{1 + \varepsilon^c} \right) + T^c \sum_{j'} \Lambda^c_{j'} \left( \frac{z^c_{j'} - z^c_j}{1 - t^{p,c}_{j'}} \right), \forall i \in J^c_i
$$

where $W^c_j(\nu)$ is the social marginal welfare weight of a worker with utility shock $\nu$ who works in sector $j$. The expectation operator takes the average value among workers in sector $j$, country $c$.

Moreover, optimal producer taxes satisfy $t^{p,c}_{ij} = -t^{p,c}_i$. The above formula describes the determinants of optimal producer taxes in terms of their impact on supply, demand and welfare. An increase in producer taxes of good $i$ in sector $j$ reduces wages in this sector and as a result the intensive margin of labor supply. The percent decline in supply of good $i$, holding other prices constant, is given by the left hand side of (8). This decline is accompanied by a decline in demand coming from two sources: a change in income coming from intensive margin and a change in income coming from movements of workers or the extensive margin. These effects are captured by the the first line in the RHS of (8). At the same time, this increase in taxes leads to a welfare loss, captured by $E^{j,c} [W^c] (1 - \tau^c)$ and an increase in supply of all goods captured by the last term in (9).

Note that in this model, producer taxes are not pinned down separately from marginal labor income tax, $\tau^c$. To see this, consider an increase in $\tau^c$ by a small amount $\delta$. We can adjust producer taxes by $dt^{p,c}_j$ so that

$$
(1 - t^{p,c}_j - dt^{p,c}_j) (1 - \tau^c - \delta) = (1 - t^{p,c}_j) (1 - \tau^c)
$$

This perturbation leaves labor supply unchanged and adjusts income so that after tax income of workers does not change. As a result, equilibrium outcomes are unchanged. In other words, what matters is the combined distortion $(1 - t^{p,c}_j) (1 - \tau^c)$. Due to this, in most of our analysis, we assume that $\tau^c = 0$ and focus on variations of $t^{p,c}_j$ across sectors and countries.

The basic idea of the role of producer taxes in this model is reminiscent of the idea of tagging as proposed by Akerlof (1978). In particular, since the sector in which a worker is working is potentially correlated with her productivity and income, this information can be used to reduce the dead-weight loss of taxation. Since sector-dependent income taxes are unavailable, producer commodity taxes and subsidies can be used to make after tax income sector-dependent. In other words, producer taxes play the role of tags - in Akerlof (1978) language.

There are multiple lessons that can be drawn from (8):

**Optimal taxes are independent of the pattern of trade.** As the above formula shows, trade or more generally parameters of technology across countries only affect optimal taxes through their effect on the distribution of income. In other words, in this model, the source of inequality is irrelevant for optimal taxes. Moreover, trade elasticities, the network structure of trade do not affect optimal taxes directly. A stark illustration of this can be done by considering the model in
which labor is fully fixed, i.e., \( \sigma = 0 \). Under this restriction, the formula for optimal producer taxes is given by

\[
\frac{1}{1 - \rho^{p,c}_j} \varepsilon^c = (1 - \tau^c) \left( 1 + \varepsilon^c - W^c_j \right)
\]

Note that this formula also holds when we completely ignore general equilibrium effects and assume that producer taxes do not affect the structure of wages. In fact, when we set \( W^c_j = 0 \), i.e., when the government does not care about workers in a sector, the above formula is simply the tax rate at the peak of Laffer curve,

\[
\frac{1}{(1 - \rho^{p,c}_j) (1 - \tau^c)} = \frac{1 + \varepsilon^c}{\varepsilon^c}.
\]

In other words, general equilibrium effects (and as a result the global process of production) are completely irrelevant for optimal producer taxes. The key question then is why general equilibrium effects do not play a role in the taxes faced by producers. The main reason behind this result is that even though we have assumed away sector-specific income taxes, our tax structure is quite rich, i.e., it has enough flexibility. A rich tax schedule implies that the value of each good to the government in each country is simply price of that good. In other words, the government in each country \( c \) can perturb its tax schedule (consumer taxes and producer taxes) and increase only supply of one particular good (that is either used or produced in each country) by a small amount. Since all taxes are at the optimum, this perturbation of taxes has no first order effect and thus the valuation of the government is proportional to the price of this good. That each government marginal valuation of each good is simply its price, implies that the government can simply ignore the general equilibrium effect and use current prices to evaluate the change in demand and supply coming from perturbation of \( \rho^{p,c}_j \). We make this argument more precise in the appendix, by constructing one such perturbation.

As the above argument illustrates having a rich set of taxes, so that each government can affect all margins of substitution and transformation for its consumers and producers, is critical. Moreover, another important facet of this result is the existence of inter-governmental transfers so that we can ignore the budget constraint of the government in each country. In other words, we do not have to worry about the effect of a perturbation of taxes on the each government’s budget constraint. When we discuss our computational results, we relax this assumption and impose budget balance at the country level and study its implications.

Optimal taxes depend on inequality across sectors As the formula illustrates distribution of income across sectors, captured by \( z^c_i \), affects producer taxes through two sources: 1. directly through its effect on social marginal welfare weights, 2. through its effect on the change in composition of demand and supply in country \( c \).

Optimal producer taxes and subsidies take the form of a VAT tax As stated in Proposition 6, the tax on inputs in sector \( i \) is equal to the negative of the tax on sales in sector \( i \). In other words, producers should be allowed deduct the cost of intermediate inputs from their revenue. Thus, they are simply paying taxes on their value added. Intuitively, when taxes are not VAT, the marginal rate of transformation between output of the firm and its inputs is not equated to prices, and thus firms’ input decisions are distorted. This distortion does not generate any
redistributional gains and thus cannot be optimal. This result implies that the so-called border
tax adjustments are suboptimal. In fact, border tax adjustments create a wedge between the price
paid for domestic versus foreign inputs. 8

At last, we must also address optimal taxes on consumption goods. The determinants of op-
timal consumption taxes are very much similar to the literature on optimal commodity taxation;
Diamond and Mirrlees (1971), Atkinson and Stiglitz (1976), and Deaton (1981) among many oth-
ers. In fact, since in the above, we have assumed that consumer preferences are homothetic in
consumption goods, optimal consumer taxes must be equated in each country and thus we have
a version of the uniform commodity taxation:

**Lemma 7.** Under Assumption 5, optimal consumption taxes are equal for all consumed goods in
country $c$.

In the appendix, we provide formulas for consumption taxes without homothetic utility func-
tion. As usual, the dependence of expenditure shares on income is the key determinant of optimal
consumption taxes.

5 An Applied Framework

In this section, we work with a particular version of the above model that is often used in the
international trade literature. We will show that our optimal tax formulas identically apply to this
framework. Moreover, we use this framework to measure the distortionary effect of our policy
recommendations (VAT taxes and subsidies) on trade.

5.1 A Model of Trade with Government Policy

Our applied framework closely follows Caliendo and Parro (2015) and Caliendo et al. (2017a). The
world economy is consisted of $N_c$ countries in each of which production occurs in $N_J$ sectors.
There are two types of goods in each sector $j$ of country $c$: intermediate goods, denoted by $q^j_c$, and
final goods, denoted by $Q^j_c$. Firms in each sector $j$ in country $c$ can produce many varieties
of intermediate goods and one variety of composite final good. The final goods in each sector
can be either consumed or can be used as material input in production of intermediate goods as
we describe below.

**Intermediate goods**

There is a continuum of intermediate goods variety $\omega_j \in [0, 1]$ within each sector. Varieties differ
in the efficiency of production $z^c_j(\omega_j)$. Intermediate goods are produced using labor and material
inputs which are the final goods from all other sectors and countries. Output of a variety $\omega_j$ in
sector $j$ in country $c$ is given by

$$q^j_c(\omega_j) = z^c_j(\omega_j)(L^c_j(\omega_j))^{x_j} \prod_{k=1}^{N_J} (M^c_{j,k}(\omega_j))^{\gamma^c_{j,k}}$$

---

8This point is also shown originally by Grossman (1980) and later generalized by Costinot and Werning (2017).
\[ \sum_{k=1}^{N_j} \gamma_{j,k}^c = 1 - \chi_j^c, \]

Here \( L_j^c (\omega_j) \) is labor input and \( M_{j,k}^c (\omega_j) \) is intermediate input from sector \( k \) that is used in production of \( \omega_j \) variety in sector \( j \), country \( c \). Value added share \( \chi_j^c \) and intermediate good share \( \gamma_{j,k}^c \) can vary across countries and sectors. We follow Eaton and Kortum (2002) we assume \( z_j^c \) has a Frechet distribution with CDF \( e^{-\lambda_j^c z - \nu} \). Parameter \( \lambda_j^c \) is the average efficiency of production in sector \( j \), country \( c \). We refer to this parameter as sector TFP. Parameter \( \nu \) is the dispersion of productivity across varieties.

Each firm pays a value added tax at rate \( t_{j,c,p}^c \). Let \( w_{j,c}^c \) be wages in sector \( j \) of country \( c \) and \( P_{c,k}^c \) be the price of the final good in sector \( k \) in country \( c \). The unit cost of producing variety \( \omega_j \) in sector \( j \), country \( c \) is given by

\[
\psi_{j,c}^c = \left( \frac{w_{j,c}^c}{(1 - t_{j,c,p}^c) \chi_j^c} \right)^{\chi_j^c} \prod_{k=1}^{N_j} \left( \frac{P_{c,k}^c}{\gamma_{j,k}^c} \right)^{\gamma_{j,k}^c}, \tag{10}
\]

**Final goods**

Final goods in sector \( j \), country \( c \) are produced by combining lowest cost intermediate goods \( \omega_j \) across all countries. Let \( Q_{c,j}^c \) be quantity of final good in sector \( j \), country \( c \) and \( r_{j,c}^c (\omega_j) \) be the quantity of variety \( \omega_j \) used in production. The production function is given by

\[
Q_{c,j}^c = \left[ \int r_{j,c}^c (\omega_j)^{1 - \frac{1}{\eta_j}} d\omega_j \right]^{\frac{\eta_j}{1 - \eta_j}},
\]

where \( \eta_j > 0 \) is the elasticity of substitution across intermediate goods within sector \( j \), and \( r_{j,c}^c (\omega_j) \) is the demand for intermediate \( \omega_j \) from the lowest cost supplier. We can solve for the demand for intermediate good \( r_{j,c}^c (\omega_j) \)

\[
r_{j,c}^c (\omega_j) = Q_{c,j}^c \left( \frac{p_{j,c}^c (\omega_j)}{P_{j}^c} \right)^{-\eta_j},
\]

where \( P_{j}^c \) must satisfy

\[
P_{j}^c = \left[ \int p_{j,c}^c (\omega_j)^{1 - \eta_j} d\omega_j \right]^{\frac{1}{1 - \eta_j}},
\]

and \( p_{j,c}^c (\omega_j) \) is the lowest price of intermediate good of variety \( \omega_j \) purchased in country \( c \).

**Trade costs and prices**

We assume trade across countries is costly and we explicitly model this as an iceberg cost a la ?. We allow iceberg cost to be sector specific and depend on origin and destination countries.

\[ \text{Given discussion in Section on optimality of zero optimal tax on consumption and intermediate good, we ignore those taxes for our quantitative analysis here.} \]

21
More specifically, we denote the cost of shipping an intermediate good produced in sector \( j \) from country \( c' \) to country \( c \) as \( d^c_{j} \geq 1 \) with \( d^c_{j} = 1 \). This means for one unit of good \( j \) to arrive at country \( c \), \( d^c_{j} \) unit must be shipped from country \( c' \). Given this notation and description of the the lowest price of intermediate good of variety \( \omega_j \) purchased in country \( c \) is given by

\[
\tilde{p}^c_j (\omega_j) = \min_{c'} \left\{ \frac{d^c_{j} \psi^c_{j}}{z^c_{j} (\omega_j)} \right\}.
\]

Following Eaton and Kortum (2002), the distributions of productivities are independent across goods, sectors and countries. Furthermore, \( 1 + \nu > \eta_j \). With these assumptions, we can solve for price of final good \( j \) in country \( c \):

\[
P^c_j = \Gamma \left( 1 + \frac{1 - \eta_j}{\nu} \right)^{\frac{1}{1 - \eta_j}} \left[ \sum_{c'} \lambda^c_{j} \left( d^{c,c'}_{j} \psi^{c'}_{j} \right)^{-\nu} \right]^{\frac{1}{\nu}}.  \tag{11}
\]

**Trade shares**

Let \( X^c_j = P^c_j Q^c_j \) be the total expenditure on sector \( j \)’s varieties in country \( c \). Also, let \( X^{c,c'}_{j} \) be expenditure in country \( c \) on good \( j \) produced in country \( c' \). The share of expenditure in sector \( j \) of country \( c \) that is produced in country \( c' \) is given by \( \pi^{c,c'}_{j} = X^{c,c'}_{j} / X^c_j \). Using properties of Frechet distribution, we can derive the following formula for \( \pi^{c,c'}_{j} \):

\[
\pi^{c,c'}_{j} = \frac{\lambda^{c'}_{j} \left( d^{c,c'}_{j} \psi^{c'}_{j} \right)^{-\nu}}{\sum_{c''} \lambda^{c''}_{j} \left( d^{c,c''}_{j} \psi^{c''}_{j} \right)^{-\nu}}.  \tag{12}
\]

**Workers and mobility**

There is a mass of \( M^c \) workers in country \( c \). Their choice set and preferences are exactly the same as that described in section 2 with the following assumption on utility over consumption and leisure

\[
U^c (x) - v^c (\ell) = \left[ \sum_{j=1}^{N} \alpha^c_j x^c_j^{1 - \frac{1}{\gamma}} \right] \frac{1}{\gamma} - \frac{\ell^{1 + 1/\varepsilon^c}}{1 + 1/\varepsilon^c}.
\]

Note that this utility function implies that disposable income and utility of a worker in sector \( j \) or country \( c \) are given by

\[
I^c_j = \left( \frac{w^c_j}{P^c} \right)^{1 + \varepsilon^c} + T^c
\]

and

\[
u^c_j = \left( \frac{w^c_j}{P^c} \right)^{1 + \varepsilon^c} + T^c.
\]

\footnote{The derivation is standard and it omitted here. See, for example, appendix in Caliendo and Parro (2015) or Caliendo et al. (2017a) for details.}
where \( P^c = \prod_{j=1}^{N_j} \left( \frac{p_j^c}{o_j^c} \right)^{\alpha_j^c} \) is the aggregate price. The choice of sector is given by the equation (2). When sector-specific taxes are available, we can simply replace \( T^c \) in the above by \( T_j^c \).

**Market Clearing**

In equilibrium, total expenditure on sector \( j \) in country \( c \) must equal the sum of households’ expenditure and intermediate good producers’ expenditure. Therefore, the market clearing condition is given by

\[
\alpha_j^c \sum_{i=1}^{N_j} M_j^c \Lambda_i^c I_i^c + \sum_{k=1}^{N_j} \gamma_{k,j}^c \sum_{\sigma=1}^{N_c} X_k^\sigma \pi_k^\sigma c = X_j^c, \tag{13}
\]

where \( I_i^c \) is defined above. Note that here \( X_k^\sigma \pi_k^\sigma c \) is country \( c \)'s export of good \( j \) from country \( c \). The first term in equation (13) is the value of domestic consumption and the second term is the total value of export. Note also that the right hand side of this equation is the gross output of industry \( j \) in country \( c \).

Finally, labor market clearing in each sector \( j \), county \( c \) is given by

\[
M_j^c \Lambda_j^c (w_j^c \ell_j^c) = (1 - t_{j,p}^c) \chi_j^c \sum_{c'=1}^{N_c} \pi_j^{c',c} X_{j}^{c'}, \tag{14}
\]

where the right hand side is labor share of after tax income and the left hand side is the total labor income in sector \( j \), county \( c \).

**Governments and Optimal Policies**

The governments’ objective is the same as in (6) and is determined under cooperation. As it turns out, since this model is a special case of the model in section (2), the implications of the model for optimal taxes is the same. Moreover, the formulas describing optimal taxes are also identical. We, thus, refer the proof and derivation of these formulas to the appendix.\(^{12}\)

### 5.2 Distortions to Trade

While we have extensively discussed the implications of the model for optimal policies, it is still unclear how such policies, namely VAT taxes, distort trade. In this section, we discuss how to understand the distortionary effects of such policies on trade. Note that the analysis done by Caliendo et al. (2017b) characterizes distortions to trade - although without taking a stance on its source (trade costs vs tariffs vs other tax and subsidies). While informative, their definition of distortions is for an economy where labor is perfectly mobile and inelastically supplied. As we illustrate below endogeneity of labor supply as well as its imperfect mobility, changes the measurement of distortions to trade.

\(^{11}\)Given that there is indeterminacy between income tax \( \tau^c \) and producer taxes, we normalize income tax rate in each country to zero. See the discussion related to \( \tau^c \) in section

\(^{12}\)The formulas in section (4) are derived under the assumption that production of all goods are interior. However, since the structure of the model in section within each sector is Ricardian, not all goods are produced in each country. Nevertheless, in the appendix, we show that this does not change the implication of the model for optimal taxes.
A Simple Example without Intermediate Goods

In this section, we provide an intuition for the way VAT taxes distort trade in our model. In particular, we consider a two country version of our model where there are no intermediate inputs to production, i.e., labor is the only input of production. Moreover, we assume that $\sigma = 0$ and that the fraction of workers in sector $i$ in country $c$ is given by $\mu_{ci}$. We further assume that there are no trade costs, i.e., $d_{ic,c'j} = 1$. Suppose that each country imposes a value-added tax $t_{cj}$ on producers in the country and the resulting revenue is rebated to workers in a lump-sum fashion. Finally, we assume that workers’ preferences are identical across countries.

Given this particular assumption, the unit cost of production of intermediate goods in sector $j$ country $c$ is given by $\frac{w_{cj}}{1-t_{cj}}$. Hence, trade shares are given by

$$\pi_{c,c'j} = \frac{\lambda_{c'j}}{\sum_{c''} \lambda_{c''j} \left( \frac{w_{c''j}}{1-t_{c''j}} \right)^{-\nu}}$$

where in the above, we are using the fact that price of sector $j$ production is identical across countries and given by $P_j$; this is because there are no trade costs. Note that because of this, $\pi_{c,c'j}$ is independent of $c$ and only depends on $c'$. In other words, the share of expenditure on sector $j$ output from country $c'$ is the same in all countries. We, thus, refer to it $\pi_{c'j}$.

We can use (15) and write

$$\frac{\pi_{c,c'j}^{c,c'}}{\pi_{c,cj}^{c,c'}} = \frac{\lambda_{c'j} \left( \frac{w_{c'j}}{1-t_{c'j}} \right)^{-\nu}}{\lambda_{c'j} \left( \frac{w_{cj}}{1-t_{cj}} \right)^{-\nu}} \left( \frac{1-t_{c'j}}{1-t_{cj}} \right)^{\nu}$$

Note that given the above, one candidate for a trade distortion is simply

$$\hat{\tau}_{c,c'j} = \frac{\pi_{c,c'j}^{c,c'}}{\pi_{c,cj}^{c,c'}} \left( \frac{w_{c'j}}{w_{cj}} \right)^{\nu}$$

The above definition of the wedge is measures the deviation of the trade shares from those implied by the true unit costs. While this is potentially a useful benchmark, in what follows, we show that it can be biased measure of the wedge. The main problem is that in this formulation, wages are endogenous and they could potentially change with a change in VAT taxes. In fact, as we show below, it is possible that according to the above definition, $\hat{\tau}_{c,c'j}$ is different from one yet trade is undistorted.

To see this, we try to solve for the trade shares. The assumption of no trade cost and immobility of labor allows us to achieve this. If $X_{cj}$ is the expenditure in sector $j$ country $c$, the demand for country $c'$’s production in sector $j$ is given by

$$\sum_{c'} \pi_{c'j}^{c,c} X_{c'j} = \pi_{c'j}^{c} \sum_{c'} X_{c'j}^{c}$$

In equilibrium, this demand must equate the supply of intermediate goods in sector $j$ country $c$. Given technologies exhibit constant returns to scale, the value of supply after taxes paid must be
equal to labor income in sector \( j \) country \( c \). Hence, in equilibrium

\[(1 - t_j^c) \pi_j^c \sum_{c'} X_j^{c'} = w_j^c \mu_j^c \ell_j^c\]

Optimality of effort by workers implies that

\[\ell_j^c = \left(\frac{w_j^c}{P}\right)^{\varepsilon_c}\]

where \( P \) is the aggregate consumer price index in each country. Combining the above, we have

\[(1 - t_j^c) \pi_j^c \sum_{c'} X_j^{c'} = \mu_j^c \left(\frac{w_j^c}{P}\right)^{1+\varepsilon_c}\]

We can replace the above in (15) and have

\[\pi_j^{c,c'} = \frac{\lambda_j^c}{\lambda_j^c} \left(\frac{1 - t_j^{c'}}{1 - t_j^c}\right)^{\nu} \left(\frac{1 - t_j^{c'} \pi_j^{c,c'} \mu_j^{c'}}{1 - t_j^c \pi_j^{c,c'} \mu_j^{c'}}\right)^{-\frac{1}{1+\varepsilon_c}}\]

After some algebra, we have

\[\pi_j^{c,c'} = \left[\frac{\lambda_j^c}{\lambda_j^c} \left(\frac{\mu_j^c}{\mu_j^{c'}}\right)^{\frac{1+\varepsilon_c}{1+\varepsilon_c+\nu}} \left(\frac{1 - t_j^{c'}}{1 - t_j^c}\right)^{\frac{\varepsilon_c \nu}{1+\varepsilon_c+\nu}}\right]\]

(17)

The above equation ties the ratio of trade shares to fundamentals of the model as well as VAT taxes. Note that when \( \varepsilon_c = 0 \), VAT taxes do not cause a deviation from efficient trade shares. In essence, in this model an increase in VAT taxes affects trade in two ways: First, it increases the unit cost of production in a country, thereby leading to a lower trade share for that country. Second, it lowers labor income, reduces wages and discourages workers from supplying labor. When labor supply is inelastic, i.e., \( \varepsilon_c = 0 \), the increase in unit cost exactly offsets the decline in wages and trade remains undistorted. When, however, labor supply is elastic, the indirect effect of the decline in wages is smaller than the direct increase in unit costs. As a result, trade will be distorted as reflected in (17). Therefore, an alternative candidate for trade wedge would be the following:

\[\tau_j^{c,c'} = \left(\frac{1 - t_j^{c'}}{1 - t_j^c}\right)^{\frac{\varepsilon_c \nu}{1+\varepsilon_c+\nu}}\]

This example clarifies the role of VAT taxes in affecting trade in two ways: 1. reducing trade shares, 2. lowering unit costs. In what follows, we extend this discussion to the general model. As we will see, since the general model is more complicated, we cannot exactly apply the same formulas as above.
**Trade Distortions in the General Model**

We will use the general model to describe two notions of distortions: first, one that takes into account the indirect effect of VAT taxes on wages - through its effect on labor income and labor supply; second, one that simply captures the deviation of trade shares from those implied unit costs.

To this end, recall that in the general model, ratio of trade shares are described by

\[
\frac{\pi_{j'}^{c,c}}{\pi_{j}^{c,c}} = \frac{\lambda_{j}^{c}}{\chi_{j}^{c}} \left[ \left( \frac{w_{c}'}{\chi_{j}^{c}} \right)^{\nu \chi_{j}^{c'}} \prod_{k=1}^{N} \left( \frac{P_{k}^{c}}{\gamma_{j,k}^{c}} \right)^{\gamma_{j,k}^{c}} \right]^{-\nu} \left( 1 - t_{p,c}^{j} \right)^{\nu \chi_{j}^{c'}}
\]

Moreover, if \( Y_{j}^{c} \) is the market value of output of sector \( j \) in country \( c \), then we have

\[
(1 - t_{p,c}^{j}) Y_{j}^{c} = \left( u_{j}^{c} \right)^{1 + \varepsilon_{c}} (P^{c})^{-\varepsilon_{c}} \Lambda_{j}^{c}
\]

where

\[
\Lambda_{j}^{c} = \frac{\mu_{j}^{c} (u_{j}^{c})^{\sigma}}{\sigma \sum_{k=1}^{N} \mu_{k} (u_{k}^{c})^{\sigma}}.
\]

We can use the above equations to define our notions of wedges as follows:

**Definition 8.** Given any set of VAT tax policies across countries and states, we define the general trade wedge to be

\[
\tau_{j}^{c,c'} = \frac{(1 - t_{p,c}^{j})^{\nu \chi_{j}^{c'}} (1 + \varepsilon_{c})^{1 + \sigma - 1}}{(1 + \varepsilon_{c})^{1 + \sigma}}
\]

Moreover, partial trade wedge is defined as

\[
\tilde{\tau}_{j}^{c,c'} = \frac{(1 - t_{p,c}^{j})^{\nu \chi_{j}^{c'}}}{(1 - t_{p,c}^{j})^{\nu \chi_{j}^{c'}}}.
\]

In the above definitions, general trade wedge takes into account that as VAT tax changes, labor income is reduced in sector \( j \) and workers are discouraged to provide effort and move away from sector \( j \). Note that this definition of the wedge, keeps output in sector \( j \) constant. The partial trade wedge simply captures the deviation of trade shares from units costs. In our quantitative exercise, we use the above notions to provide a better understanding of distortionary effect of optimal policies on trade.

### 6 Quantitative Analysis

In this section, we apply our results to a calibrated version of the model in section 5.
6.1 Calibration

We calibrate the model in section 5 by taking a Laissez-Faire version of this economy - one without government policies - to the data. We choose some of the parameters of the model similar to those chosen in the literature - mostly parameters related to the behavioral responses: elasticities of labor supply (intensive and extensive margin), trade elasticities, etc. The parameters of technology, productivities as well as production functions are chosen to match the data.\footnote{We do this mainly because finding comprehensive data on government policies is difficult and is in line with the rest of the international trade literature.}

**Data.** The data that we use to calibrate the model come from the following two sources: 1) World Input-Output Database (WIOD) which includes World Input-Output Tables (WIOT) and Socio Economic Accounts (SEA),\footnote{http://www.wiod.org/} and 2) Groningen Growth and Development Centre (GGDC) Productivity Level Database which includes data on relative prices and labor productivity across countries for 42 major economies and up to 35 industries in 2005.\footnote{https://www.rug.nl/ggdc/productivity/pld/}

We use the World Input-Output Tables to construct bilateral trade shares as well as intermediate input and labor shares for each sector and country. Data for employment by sector comes from Socio Economic Accounts. Finally, we use price data from GGDC together with data from WIOT to back out parameters of preferences and technology.

**Sectors and countries.** We calibrate the model to a sample of 39 countries and 10 sectors. To construct sectors we aggregate 34 industries classified according to the International Standard Industrial Classification (ISIC Rev. 3.1) into 10 categories. Table 1 contains a complete list of industries and aggregated 10 sectors. Our sample consists of the following countries: Australia, Austria, Belgium, Bulgaria, Brazil, Canada, China, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, India, Indonesia, Italy, Ireland, Japan, Latvia, Lithuania, Luxembourg, Malta, Mexico, the Netherlands, Poland, Portugal, Romania, Russia, Spain, Slovak Republic, Slovenia, South Korea, Sweden, Turkey, the United Kingdom and the United States.

**Parameters chosen independently.** In line with many estimates of the intensive elasticity of labor supply, we choose $\epsilon_c = 0.5$ – see Chetty et al. (2011) for an extensive discussion. We choose the dispersion of productivity to be $\nu = 4$. This is based on estimated trade elasticities in Simonovska and Waugh (2014). The elasticity of substitution in final goods production function is $\eta_j = 4$. Moreover, we choose $\sigma = 0.2$. This parameter was chosen to match migration-elasticity of 0.2 that is estimated in Caliendo et al. (2017a). Finally, we assume that the elasticity of substitution across different consumption goods is 1, i.e., $\gamma = 1$.

**Bilateral trade shares, labor shares, and intermediate input shares.** We use the World Input-Output Tables in 2005 to construct bilateral trade flows $X_{j,c'}^{c,d'}$, gross output $X_j^c$ and bilateral trade shares $\pi_{j,c,d'}$. These quantities are used in calibrating consumption shares and trade costs as we describe below.
<table>
<thead>
<tr>
<th>Sectors</th>
<th>Industries in each sector</th>
<th>ISIC rev.3 code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agriculture</td>
<td>Agriculture, hunting, forestry &amp; fishing</td>
<td>AtB</td>
</tr>
<tr>
<td>Mining</td>
<td>Mining &amp; quarrying</td>
<td>C</td>
</tr>
<tr>
<td></td>
<td><strong>Food, beverage &amp; tobacco</strong></td>
<td>15t16</td>
</tr>
<tr>
<td></td>
<td><strong>Textile &amp; textile products</strong></td>
<td>17t18</td>
</tr>
<tr>
<td></td>
<td><strong>Leather, leather products &amp; footwear</strong></td>
<td>19</td>
</tr>
<tr>
<td></td>
<td><strong>Wood products &amp; products of wood and cork</strong></td>
<td>20</td>
</tr>
<tr>
<td></td>
<td><strong>Paper, printing &amp; publishing</strong></td>
<td>21t22</td>
</tr>
<tr>
<td></td>
<td><strong>Coke, refined petroleum &amp; nuclear fuel</strong></td>
<td>23</td>
</tr>
<tr>
<td></td>
<td><strong>Chemicals &amp; chemical products</strong></td>
<td>24</td>
</tr>
<tr>
<td></td>
<td><strong>Rubber &amp; plastics</strong></td>
<td>25</td>
</tr>
<tr>
<td></td>
<td><strong>Other non-metallic mineral products</strong></td>
<td>26</td>
</tr>
<tr>
<td></td>
<td><strong>Basic &amp; fabricated metal</strong></td>
<td>27t28</td>
</tr>
<tr>
<td></td>
<td><strong>Machinery, not elsewhere classified</strong></td>
<td>29</td>
</tr>
<tr>
<td></td>
<td><strong>Electrical &amp; optical equipment</strong></td>
<td>30t33</td>
</tr>
<tr>
<td></td>
<td><strong>Transport Equipment</strong></td>
<td>34t35</td>
</tr>
<tr>
<td></td>
<td><strong>Other manufacturing</strong></td>
<td>36t37</td>
</tr>
<tr>
<td></td>
<td><strong>Electricity, gas &amp; water supply</strong></td>
<td>E</td>
</tr>
<tr>
<td></td>
<td><strong>Construction</strong></td>
<td>F</td>
</tr>
<tr>
<td></td>
<td><strong>Sale and repair of motor vehicles &amp; motorcycles; retail sale of fuel</strong></td>
<td>50</td>
</tr>
<tr>
<td></td>
<td><strong>Wholesale trade, except of motor vehicles and motorcycles</strong></td>
<td>51</td>
</tr>
<tr>
<td></td>
<td><strong>Retail trade and repair, except of motor vehicles &amp; motorcycles</strong></td>
<td>52</td>
</tr>
<tr>
<td>Hotels &amp; restaurants</td>
<td><strong>Hotels &amp; restaurants</strong></td>
<td>H</td>
</tr>
<tr>
<td>Transport &amp; communication</td>
<td><strong>Inland transport</strong></td>
<td>60</td>
</tr>
<tr>
<td></td>
<td><strong>Water transport</strong></td>
<td>61</td>
</tr>
<tr>
<td></td>
<td><strong>Air transport</strong></td>
<td>62</td>
</tr>
<tr>
<td></td>
<td><strong>Other supporting transport activities</strong></td>
<td>63</td>
</tr>
<tr>
<td></td>
<td><strong>Post &amp; telecommunications</strong></td>
<td>64</td>
</tr>
<tr>
<td>Financial &amp; business services</td>
<td><strong>Financial intermediation</strong></td>
<td>J</td>
</tr>
<tr>
<td></td>
<td><strong>Real estate activities</strong></td>
<td>70</td>
</tr>
<tr>
<td></td>
<td><strong>Renting of machinery &amp; equipment and other business activities</strong></td>
<td>71t74</td>
</tr>
<tr>
<td>Other services</td>
<td><strong>Public administration and defense</strong></td>
<td>L</td>
</tr>
<tr>
<td></td>
<td><strong>Education</strong></td>
<td>M</td>
</tr>
<tr>
<td></td>
<td><strong>Health &amp; social work</strong></td>
<td>N</td>
</tr>
<tr>
<td></td>
<td><strong>Other services</strong></td>
<td>O</td>
</tr>
</tbody>
</table>

Table 1: List of aggregated sector and corresponding industries
We also use WIOD to construct share of value added in gross output, $\chi_{cj}$, and the intermediate input shares $\gamma_{jk}$ across countries and sectors using data on value added, gross output and intermediate consumption.

**Final consumption shares.** Using value added data from WIOD together with sectoral employment data in SEA, we compute value added per worker in each sector and each country. We assume that national income in each country is equal to the sum of all value added. We also assume trade balance at the country level.\(^{16}\) With these two assumptions, we can use equation (13) (goods market clearing) to back out final consumption shares as

$$\alpha_{j}^{c} = \frac{X_{j}^{c} - \sum_{k=1}^{N_{j}} \gamma_{k,j}^{c} \sum_{c'=1}^{N_{c}} X_{k}^{c',c}}{I_{c}^{c}},$$

where $I_{c}^{c}$ is the sum of all value added in country $c$.

**Sectoral choice parameters.** Another important ingredient of the model is the set of parameters $\mu_{cj}^{c}$. For this, we use employment and value-added data to back out wages and use the optimal sectoral choice formula to back out $\mu_{cj}^{c}$. We impute wages by assuming that workers are optimizing. To do so, we start by using sectoral relative price data in GGDC’s Productivity Level Database to construct price indices in each country using

$$P^{c} = \prod_{j=1}^{N} \left( \frac{P_{j}^{c}}{\alpha_{j}^{c}} \right)^{\alpha_{j}^{c}}.$$

Our parametric assumption about labor supply implies that the intensive margin of labor supply in sector $j$ is given by $\ell_{j}^{c} = \left( \frac{w_{j}^{c}}{P^{c}} \right)^{\epsilon}$. Therefore, total labor income in sector $j$, country $c$ is

$$M_{j}^{c} \Lambda_{j}^{c} \left( \frac{w_{j}^{c}}{P^{c}} \right)^{1+\epsilon} \frac{1}{(P^{c})^{\epsilon}}.$$

This has to be equal to the total value added in sector $j$ country $c$ which we observe in WOID. Given that we observe $M_{j}^{c} \Lambda_{j}^{c}$ in SEA, we can solve for $w_{j}^{c}$ by setting the above equal to the observed value added.

Given the observed wages, we can solve the maximization problem for workers in sector $j$ of country $c$ and derive their utility $u_{j}^{c}$. Using these imputed utilities, equation (2) can be used to solve for the vector of Gumbel sectoral parameters $\mu_{j}^{c}$ such that sectoral employment in the model matches the employment data in all countries and all sectors.

**Trade costs and efficiency parameter.** The last set of parameters to be calculated are trade costs, $d_{j}^{c,c'}$, and sectoral productivities, $\lambda_{j}^{c}$.\(^{17}\) To calculate these values, we use the price data

---

\(^{16}\)Note that this calculation assume that trade balances are 0. An alternative approach is to include trade deficit and rebate it to workers in a lump-sum fashion as in Caliendo and Parro (2015). As it turns out, the resulting parameters are very close to what we compute under balanced trade assumption and as a result we have decided to proceed with balanced trade.

\(^{17}\)Our calibration procedure closely follows that of Lewis et al. (2018).
and computed wages to calculate the unit cost $\psi_j$ (assuming value added taxes are zero in the benchmark). The iceberg trade costs can be calculated using the equation (12). In particular, we have

$$\pi_{j}^{c,c'} = \frac{\lambda_j^{c'} \left( \frac{\psi_j^{c'} \psi_j^c}{P_j^c} \right)^{-\nu}}{(P_j^c)^{-\nu} \Gamma \left( 1 + \frac{1 - \eta_j}{\nu} \right)^{1-\eta_j}},$$

where in the above we have used the definition of $P_j^c$ in (11). Therefore

$$\frac{\pi_{j}^{c,c'}}{\pi_{j}^{c',c'}} = \left( \frac{d_{j}^{c,c'}}{d_{j}^{c',c'}} \right)^{-\nu} \left( \frac{P_{j}^{c'}}{P_{j}^{c}} \right)^{-\nu},$$

since $d_{j}^{c,c'} = 1$. Given our assumed value for $\nu$, the equation above can be solve to back out the trade cost $d_{j}^{c,c'}$.

Finally, note that we can combine equations (12) and (11) to arrive at the following equation:

$$\pi_{j}^{c,c'} \left( P_j^c \right)^{-\nu} = \lambda_j^{c'} \left( \psi_j^{c'} \right)^{-\nu} \Gamma \left( 1 + \frac{1 - \eta_j}{\nu} \right)^{1-\eta_j}.$$

We can then use this equation to solve for sectoral TFP parameters $\lambda_j^c$. This concludes the calibration.

Figure 1 shows the heat map of logarithm of sectoral TFPs in each country. Lighter shades represent higher TFP. Few of patterns stand out. First, ‘finance and business services’ and ‘utilities’ have high TFP in almost all countries. Second, ‘construction’, ‘hotels & restaurants’ and ‘other services (which includes government, health and education)’ are low TFP sectors. Third, there are more variations in TFP in ‘agriculture’, ‘manufacturing’ and ‘trade’ (whole sale and retail), i.e., the are the tradable sectors.

Figure 1: Log of sector TFP for all sectors and countries if the sample. Lighter shades represent higher TFP.
Figure 2 shows heat map of trade costs in four sectors ‘agriculture’, ‘manufacturing’, ‘trade’ and ‘finance and business services’ across all countries. The rows (y-axis) represent origin country and columns (x-axis) represent destination. Here, white cells represent no trade costs or very small trade costs and darker cells represent higher trade costs. Note that manufacturing stands out as the sector with lowest trade costs, followed by finance and business services. Moreover, the US and many other advanced countries face lowers trade cost when exporting goods in all sectors. This is consistent with findings in Waugh (2010) among others.

(a) Agriculture, forestry & fishing
(b) Manufacturing
(c) Wholesale & retail trade
(d) Financial & business services

Figure 2: Iceberg trade cost for select industries. Row (y-axis) countries represent origin. Column (x-axis) represent destination.
6.2 Optimal Policies

Using the parametrization described in the previous section, we can solve the optimal taxation problem. In doing so we assume social welfare function within each country satisfies \( W(z) = \log z \). Moreover, we assume Pareto weights on each country’s social welfare is proportional to the value added per worker in that country (in the observed data). We make this assumption to avoid massive redistribution from rich to poor countries under optimal policies. We choose policies (with and without sector-specific income taxes) in order to maximize the following social welfare function

\[
\sum_{c=1}^{N_c} \lambda^c M^c \int_{\nu \in \mathbb{R}^N} \log \left( \max_j u^c_j e^{\nu_j} \right) dH^c(\nu).
\]

Note that we can show that the above objective is given by

\[
\sum_{c=1}^{N_c} \lambda^c M^c \left[ \log \left( \sum_j \mu^c_j \left( u^c_j \right)^\sigma \right) + \gamma \right]
\]

where \( \gamma \) is the Euler-Mascheroni constant.

6.2.1 Optimal Policy with Sector-Specific Income Taxes

In order to provide a benchmark, we start by solving for optimal policies under the assumption that income tax and transfers in each country can depend on workers’ sector. As we have demonstrated in section 4.1, when income taxes are flexible and can depend on workers’ sector there is no need for distortionary taxes on production. Therefore, we only need to solve for sector specific tax/transfers \( T^c_j \) in each sector \( j \) and country \( c \). Figure 3 shows the heat map for ratio tax/transfers to pre tax income in each sector and each country. Positive values represent transfer and negative values represent taxes. We make three observations. First, sectors that receive transfer are almost the same across countries (all cells with darker shade). These are ‘Agriculture’, ‘Construction’, ‘Trade (Retail and Wholesale)’, ‘Hotels and Restaurants’ and ‘Other services’. Second, magnitude of transfers are very large relative to pre-tax income. Third, tax payer sectors are always the same across countries (these are cells with lighter shades).
The particular pattern of tax and transfers can be understood using the formula provided in Lemma goes back to the result in Lemma 4 which implies that the sectors with higher real income (pre tax) receive lower transfers (pay more taxes). Moreover, the degree of this compensation, depends on the labor supply elasticity \((1 + \sigma)(1 + \varepsilon)\). In our model, this elasticity is 1.8. Hence, the difference between transfers paid to workers in each sector must be equal to half of the difference in their income. This implies that sectors with low income, ‘Hotels and Restaurants’ must receive high transfers and this is the case in all countries because their income is low in all countries. This is demonstrated in Figure 4 which plots the ratio of transfers to pre-tax income in each sector and each country.
Figure 4: Ratio of optimal sector specific transfers (taxes are negative) to pre tax income plot against log of pre tax income. Each circle represents one sector in one country.

The large magnitude of these transfers relative to income highlights the importance of these transfer policies in compensating workers who are affected by trade. The large magnitude of these transfers hint that in the absence of sector specific transfer, we should expect to see large distortionary policies that try to achieve the goal of income redistribution. This is what we do next.

6.2.2 Optimal Policy without Sector-Specific Income Taxes

In this section, we focus on optimal policies absent sector-specific income taxes, i.e. $T_j^c = T^c$ for all $j$ and all $c$. As we saw in section 4.2, under this assumption, the optimal policy is a set of sector and country specific value added taxes on producers, $\nu_j^{p,c}$ together with a lump-sum transfer to all workers.

VAT Taxes

Figure 5 depicts the distribution of optimal VAT taxes across countries and sectors. As it is shown, there is significant heterogeneity across sectors suggesting significant redistribution through VAT taxes.
We start our analysis of the optimal policies by investigating optimal VAT taxes implied by the model. Recall that optimal taxes can be decomposed into four components - as shown in Proposition 6:

\[
\frac{1}{1 - t_{j}^{p,c}} = -\frac{1}{\varepsilon} E^{j,c} \left[ \mathcal{W}^{c} \right] + \frac{1 + \varepsilon}{\varepsilon} \\
+ \frac{1}{\varepsilon} \frac{\sigma}{1 + \varepsilon} \sum_{j'} \Lambda_{j,j'}^{c} \left( z_{j}^{c} - z_{j'}^{c} \right) \\
+ \frac{1}{\varepsilon} \frac{\sigma}{1 + \varepsilon} \sum_{j'} \Lambda_{j,j'}^{c} \left( \frac{z_{j}^{c}}{1 - t_{j}^{p,c}} - \frac{z_{j'}^{c}}{1 - t_{j'}^{p,c}} \right)
\]

The forces that determine optimal VAT taxes are: 1. welfare effect – the first term on the RHS which captures the social marginal welfare weight associated with workers in sector \( j \); 2. intensive margin elasticity; 3. composition effect on demand - captured by the third term; 4. composition effect on supply - the fourth term.

Figure 6 shows the left hand side of the formula i.e. \( \frac{1}{1 - t_{j}^{p,c}} \) and terms on the right hand side for 4 countries. As we see in all countries and most sectors the supply and demand composition terms are smaller and often offset each other. Therefore, pattern of optimal taxes almost closely follows the pattern of marginal social welfare. In other words, changes in optimal VAT taxes across countries are to a first order driven by changes in social marginal welfare weight of workers. Recall that \( \mathcal{W}_{j}^{c} = \lambda_{j}^{c} \frac{1}{z_{j}^{c}/(1 + \varepsilon) + T^{c}} \). Since \( \lambda_{j}^{c} \) is proportional to a country’s GDP per capita, \( \mathcal{W}_{j}^{c} \) is related to the deviation of income of workers in sector \( j \) from its average in the country. Therefore, changes in optimal VAT taxes mainly capture the changes in the relative position of a sector’s income per capita relative to its national average. This insight will help us greatly in understanding the behavior of trade wedges.
Figure 6: Optimal tax formula (equation 8). The black solid line represents the optimal tax in sector \( j \) country \( c, \frac{1}{1-t_{p,c}^j} \). Blue dashed line is the marginal social welfare term. Gray line is the elasticity (intensive margin) term. Red line is the effect on composition of demand from sector \( j \) relative to other sectors. Green line is the effect on composition of supply to sector \( j \) relative to other sectors.

**Trade Wedges**

We, next, turn to the wedges to trade. As we have defined them in section 5.2, wedges to trade can be calculated using VAT taxes as follows:

\[
\hat{\tau}_{c,c'}^j = \left1\right (1 - t_{p,c}^j) \nu x_{c'}^{j'} \left \{ \frac{c + \sigma + \sigma}{1 + \epsilon} \right \}^{1+\epsilon} (1+\epsilon)
\]

\[
\hat{\widehat{\tau}}_{c,c'}^j = \left \{ \frac{t_{p,c}^j}{1+\epsilon} \right \}^{1+\epsilon} \left \{ \frac{c + \sigma + \sigma}{1 + \epsilon} \right \}^{1+\epsilon}
\]

\[
\hat{\tau}_{c,c'}^j = \left \{ \frac{1}{1+\epsilon} \right \}^{1+\epsilon} \frac{c + \sigma + \sigma}{1 + \epsilon}
\]

36
where $\tau_{j}^{c,c'}$ is the trade wedge that takes into account the change in wages and labor supply resulting from distorting trade while $\hat{\tau}_{j}^{c,c'}$ only measures the of trade shares from those implied by the (undistorted) unit costs. As it can be seen, these wedges are mainly determined by VAT taxes across countries and the share of labor in production.

In Figure 7, we plot the distribution of general trade wedges in four sectors across the world; those with highest export shares across countries. These include: Manufacturing, Mining, Hotels & Restaurant, and Transport & Communication. Each cell represents the wedge to trade between a destination country – on the x-axis – and a source country – on the y-axis. That is, if the cell associated with manufacturing with U.S. as its x-axis and China as its y-axis displays a number greater than 1, it implies that optimal policies lead to an increase in the share of imports from China to the United States.

We draw two main conclusion from the pattern of these wedges: First, the wedges are significantly different from 1 implying that it is optimal to significantly distort trade. Second, among sectors with significant trade, distortions to trade are lowest for manufacturing.

The fact that optimal trade wedges significantly deviate from 1 mechanically comes from two sources: 1. variations in optimal VAT taxes, 2. variations in labor share (or alternatively share of intermediate goods in production). Since variations in marginal social welfare weight is the main source of variations in optimal VAT taxes, significant deviation of wedges from 1 simply reflect the fact that sectoral income relative to average income varies across countries. This can be seen by comparing the Manufacturing sector in the United States and Germany, as an example. In this case, due to a higher social welfare weight for manufacturing workers in Germany, optimal VAT taxes are higher in the United States which then leads to a trade wedge higher than 1. Note that variations in labor share can also significantly affect wedges. To see this, note that optimal Manufacturing VAT taxes are larger in China yet it is optimal to increase Manufacturing imports from China to the United States. This is mainly because of the manufacturing sector in China has a higher share of intermediate inputs than that of the United States – 75.9% for China vs. 65.0% for the U.S.

Figure 8 depicts the share of domestic production in each sector and country for the calibrated economy. As we can see, the sectors with significant share of outside production are: Manufacturing, Mining, Hotel & Restaurant, Transport & Communication. Among these sectors, as shown in Figure 7, Manufacturing has the lowest level of distortions, i.e., trade wedges are close to 1. This result is mainly driven by cross-sectional variation in wages across countries. Note that to the extent that optimal VAT taxes are different across countries due to changes in the social marginal welfare weights, an important determinant of trade wedges would be the variation of wages across countries within one sector. Figure 9 shows the variance of log of wages across countries for each sector and as we see variation in log wages for manufacturing is among the lowest. A potential explanation of this can be explained by low trade and costs and input-output linkages. In particular, since for manufacturing trade costs are low, variation of prices across countries are smaller. Since the share of each sector in its production is the largest, this implies that unit costs are closer to each other and thus the variations in trade shares and wages are smaller.
Figure 7: Optimal bilateral distortions for select industries, $\tau_{j}^{c,c'}$. Each cell represents the trade wedge $\tau_{j}^{c,c'}$ between a destination country $c$ on the x-axis and a source country $c'$ on the y-axis.
Figure 8: Domestic share of production across sectors and country, $\pi_{j,c}^{c,c}$, under Laissez-Faire

Figure 9: Cross-country variance of log wages for each sector under optimal VAT taxes

Welfare Gains
Here, we discuss the implications of our optimal policies for welfare. While in our benchmark calibration, very much in line with the international trade literature, we have assumed a Laissez-Faire economy, perhaps this cannot be a good benchmark for welfare since in reality taxes are not zero. To address this issue, we use as benchmark the solution of the optimal policy problem assuming that trade is free and production is efficient. In other words, we solve the optimal policy problem where in each country marginal tax rate $\tau_c$ and lump-sum transfer $T^c$ are chosen.

Given this benchmark, we depict the welfare gains of the model with optimal VAT taxes relative to those in the benchmark in Figure 10. As it can be seen, these welfare gains are large, from around 1% of consumption to around 5% of consumption in the aggregate while within the
country, the sectoral gains and losses could be much larger. Among sectors with significant trade, manufacturing always experiences welfare loss - due its high wages - while Hotels & Restaurants often experience welfare gains.

We view these large welfare gains as suggesting that the prescriptions of our model for optimal policy are significant and important. They thus identify large costs to the lack of existence of sector-specific taxes and transfers.

![Welfare Gains from Optimal Policy by Country](image)

Figure 10: Consumption equivalent measure of welfare gains; Optimal VAT taxes vs. optimal linear income taxes.

7 Conclusion

In this paper, we have provided a theoretical and quantitative analysis of optimal trade policies and income taxation when inequality is affected by trade. We have shown that when income taxes can be sector-specific, free trade is optimal. Absent such income taxes, trade is distorted using VAT taxes and subsidies. Theoretically, we show that these taxes are independent of the trade structure and only dependent on income distribution and labor supply elasticities. Quantitatively, we show that the implied distortions to trade are large and that the aggregate welfare gains from sector-specific VAT taxes are significant.

References


Appendix

A Proofs

A.1 Proof of Proposition 2

Proof. That any allocations and consumer and producer price vectors must satisfy the constraints in problem (P) is straightforward from the definition of indirect utility function, demand function, and optimality condition by firms of intermediate goods.

Now suppose that allocations \( \{x_j^c(\theta), \ell_j^c(\theta)\}_{c \in \{1, \ldots, N_C\}, j \in \{1, \ldots, N\}} \) together with consumer/producer prices and wage \( \{p^c, q^c, w^c\}_{c \in \{1, \ldots, N_C\}} \) and policies \( \{\tau^c, T^c\}_{c \in \{1, \ldots, N_C\}} \) satisfy the constraints in \((P)\). We construct a set of commodity taxes and show that under these policies the resulting allocations and prices are a competitive equilibrium.

Consider an arbitrary vector of international prices \( \hat{\mathbf{p}} = (\hat{p}_1, \cdots, \hat{p}_N) \) where \( \hat{p}_i > 0 \). For each good \( i \) and country \( c \), we can define

\[
\ell_i^{x,c} = \frac{q_i^c}{\hat{p}_i} - 1
\]

and

\[
\ell_i^{p,c} = \frac{p_i^c}{\hat{p}_i} - 1
\]

\[
\ell_{ij}^{p,c} = \frac{p_i^c}{\hat{p}_j} \frac{\partial G_i^c}{\partial Q_{ij}} \left( \{Q_{ij}^c\}, L_i^c \right) - 1
\]

Obviously, at prices given by \( \hat{\mathbf{p}} \) and taxes as defined above the firms’ first order conditions are satisfied. Since the production functions are assumed to be concave, then we must have that firms’ optimality conditions are satisfied. Moreover, by definition of the demand function and indirect value function, the workers’ optimality conditions are also satisfied. This implies that we only need to check that government budget constraints are satisfied. Since all workers budget constraints hold - given the definition of the demand function, all feasibility constraints are satisfied - the first constraint in problem \((P)\), and transfers between governments are allowed, Walras law implies that government budget constraints are satisfied. This completes the proof. \(\square\)
A.2 Proof of Proposition 3

Proof. We first consider a relaxed problem where the planner chooses the allocation of consumption, effort and sectoral choices across countries. We then show that the solution to this relaxed problem satisfies the constraints in program P.

Consider the following relaxed problem:

\[
\max \left\{ x^c_j, \{\ell^c_j\}, \{Q^c_{ij}\} \right\}
\]

subject to

\[
\sum_{c=1}^{N_C} \sum_{j=1}^{N^c_j} \Lambda^c_j x^c_{j,i} + \sum_{c=1}^{N_C} \sum_{k=1}^{N} Q^c_{ki} = \sum_{c=1}^{N_C} G^c_i \left( L^c_i \left\{ Q^c_{ij} \right\}^{N^c_j}_{j=1} \right), \forall i
\]

\[
\Lambda^c_j \ell^c_{j,i} = \sum_{i \in J^c_j} L^c_i
\]

\[
\mu^c_{j} \left[ U^c_c(x^c_j) - v^c_c(\ell^c_{j,i}) \right]^\sigma
\sum_{k=1}^{N^c_j} \mu^c_k \left[ U^c_c(x^c_k) - v^c_c(\ell^c_{k}) \right]^\sigma = \Lambda^c_j
\]

Note the above is a more relaxed version of (P) since we have allowed the for arbitrary allocation of consumption to workers and arbitrary choice of effort, \( \ell \). Hence, if we show that its solution satisfies the claim in the proposition and

Let \( \rho_i \) be the lagrange multiplier associated with the resource constraint for good \( i \). Now, consider a perturbation of \( x^c_{j,i} \) and \( x^c_{j,i} \) by \( dx_i \) and \( dx_i \) so that utility of workers in sector \( j \) country \( c \) are unchanged. That is,

\[
\frac{\partial U^c_c(x^c_j)}{\partial x_i} dx_i + \frac{\partial U^c_c(x^c_j)}{\partial x_{i'}} dx_{i'} = 0.
\]

This perturbation leaves the sectoral choice \( \Lambda^c_j \) as well as aggregate welfare unchanged. Hence, at the optimum, we must have that

\[
\rho_i dx_i + \rho_{i'} dx_{i'} = 0.
\]

Hence,

\[
\frac{\partial U^c_c(x^c_j)}{\partial x_i} = \frac{\rho_i}{\rho_{i'}}, \forall i, i'
\]

where the above holds if the optimal allocation of consumption is interior. The above implies that marginal rates of substitutions are equated across all workers and all goods. Let the price vector be given by \( \hat{p} = (\rho_1, \rho_2, \cdots, \rho_N) \). Then, the above implies that if \( \hat{u}^c_j = U^c_c(x^c_j) \), then

\[
\mathbf{x}^c_j \in \arg \max_{x: U^c_c(x) \geq \hat{u}^c_j} \sum_i \hat{p}_i x_i
\]
Let $C^c(\hat{u}; \hat{p})$ be the expenditure function associated with the above problem. Moreover, if $h^c_i(\hat{u}; \hat{p})$ is the Hicksian demand for good $i$ in country $c$, then $x^c_{j,i} = h^c_i(\hat{u}^c_j; \hat{p})$.

Next, consider a perturbation of $x^c_j$ by $dx_i$ and a perturbation of $\ell^c_j$ by $d\ell$ where

$$dx_i = \frac{\partial h^c_i(\hat{u}^c_j; \hat{p})}{\partial \hat{u}^c_j} \left( v^c \right)' (\ell^c_j) d\ell.$$  

Note that the resulting change in utility of workers in sector $j$ is given by

$$du^c_j = \sum_i \frac{\partial U^c(x^c_j)}{\partial x_i} \frac{\partial h^c_i(\hat{u}^c_j; \hat{p})}{\partial \hat{u}^c_j} \left( v^c \right)' (\ell^c_j) d\ell - \left( v^c \right)' (\ell^c_j) d\ell = 1 \times \left( v^c \right)' (\ell^c_j) d\ell - \left( v^c \right)' (\ell^c_j) d\ell = 0$$  

Hence, this perturbation leave sectoral choice and aggregate welfare unchanged. Therefore, if $\zeta^c_j$ is the Lagrange multiplier associated with labor market clear in sector $j$ with $i \in J^c_j$, then, at the optimum, we must have

$$- \sum_i \rho_i dx_i + \zeta^c_j d\ell = 0$$  

Note that the FOC associated with $L^c_i$ implies that

$$\rho_i \frac{\partial G^c_i}{\partial L_i} - \zeta^c_j = 0$$  

Combining the above, we have

$$\sum_i \rho_i dx_i = \rho_i \frac{\partial G^c_i}{\partial L_i} d\ell = \sum_i \rho_i \frac{\partial h^c_i(\hat{u}^c_j; \hat{p})}{\partial \hat{u}^c_j} \left( v^c \right)' (\ell^c_j) = \rho_i \frac{\partial G^c_i}{\partial L_i}$$  

Using the definition, of the expenditure function, we have

$$\frac{\partial}{\partial \hat{u}^c} C^c(\hat{u}; \hat{p}) \left( v^c \right)' (\ell^c_j) = \rho_i \frac{\partial G^c_i}{\partial L_i} (20)$$  

Now, given the above, we can define a competitive equilibrium where all consumer and producer indirect taxes are zero and show that the solution to the above relaxed problem constitutes a competitive equilibrium.

To see this, let

$$w^c_j = \zeta^c_j$$

and let $T^c_j$ be defined as

$$T^c_j = C^c(\hat{u}^c_j; \hat{p}) - w^c_j \ell^c_j$$

Note that by the definition of the expenditure function and indirect utility function $V^c(I; \hat{p})$, we have

$$V^c(C^c(\hat{u}; \hat{p})) = \hat{u} \Rightarrow V^c(I; \hat{p}; C^c(\hat{u}; \hat{p})) \frac{\partial C^c(\hat{u}; \hat{p})}{\partial \hat{u}} = 1$$
Hence, equation (20) becomes

$$(v^c)'(\ell_j^c) = w^c_{j,i}V_i^c(\tilde{p}; w_j; \ell_j^c + T_j^c) \Rightarrow \ell_j^c \in \arg\max_{\ell} V_i^c(\tilde{p}; w_j; \ell_j^c + T_j^c) - v^c(\ell)$$

Note that by definition of Hicksian demand,

$$x_j^c = x^c(\tilde{p}; w_j^c; \ell_j^c + T_j^c)$$

This implies that the solution of $(P')$ satisfies the constraints in $(P)$. This concludes the proof. □

### A.3 Proof of Lemma 4

**Proof.** As we have shown in Appendix B.2, aggregate social welfare in $(P)$ is only a function of $\sum_j \mu_j^c(u_j^c)^\sigma$. Now, consider a perturbation of $T_j^c$ by $dT_j$ and $T_{j'}^c$ for any two sectors $j, j'$ in $(P)$ such that

$$\mu_j^c(\tilde{u}_j^c)^{\sigma-1} V_{i,j}^c dT_j + \mu_{j'}^c(\tilde{u}_{j'}^c)^{\sigma-1} V_{i,j'}^c dT_{j'} = 0$$

or

$$\Lambda_j^c \frac{V_{i,j}^c}{\tilde{u}_j^c} dT_j + \Lambda_{j'}^c \frac{V_{i,j'}^c}{\tilde{u}_{j'}^c} dT_{j'} = 0$$

Note that this perturbation leaves welfare the same. Moreover, the fraction of workers in every sector but $j, j'$ remains unchanged. Hence, the change in $\Lambda_j^c$ and $\Lambda_{j'}^c$ is given by

$$d\Lambda_j^c = \sigma \Lambda_j^c \frac{V_{i,j}^c}{\tilde{u}_j^c} dT_j, d\Lambda_{j'}^c = \sigma \Lambda_{j'}^c \frac{V_{i,j'}^c}{\tilde{u}_{j'}^c} dT_{j'}$$

Since we know that the solution to $(P)$ coincides with $(P')$, optimality condition associated with this perturbation is given by

$$- \sum_i \rho_i \left( \Lambda_j^c \frac{\partial x_{j,i}^c}{\partial T_j^c} dT_j + \Lambda_{j'}^c \frac{\partial x_{j',i}^c}{\partial T_{j'}^c} dT_{j'} + x_{j,i}^c d\Lambda_j^c + x_{j',i}^c d\Lambda_{j'}^c \right) + \zeta_j^c \ell_j d\Lambda_j^c + \zeta_{j'}^c \ell_{j'} d\Lambda_{j'}^c = 0$$

Note that from our construction in section (A.2), we have that $\zeta_j^c = w_j^c$. Hence, we can write the above as

$$-dT_j \Lambda_j^c \sum_i \rho_i \frac{\partial x_{j,i}^c}{\partial T_j^c} - dT_{j'} \Lambda_{j'}^c \sum_i \rho_i \frac{\partial x_{j',i}^c}{\partial T_{j'}^c} + d\Lambda_j^c \left( z_{j,i}^c - \sum_i \rho_i x_{j,i}^c \right) + d\Lambda_{j'}^c \left( z_{j',i}^c - \sum_i \rho_i x_{j',i}^c \right) = 0$$

Since $\rho_i$’s are the prices that consumers face, we must have

$$-\Lambda_j^c dT_j - \Lambda_{j'}^c dT_{j'} - \sigma \Lambda_j^c \frac{V_{i,j}^c}{\tilde{u}_j^c} T_j^c dT_j - \sigma \Lambda_{j'}^c \frac{V_{i,j'}^c}{\tilde{u}_{j'}^c} T_{j'}^c dT_{j'} = 0$$

Hence,

$$-\Lambda_j^c + \lambda_j^c \frac{V_{i,j}^c}{\tilde{u}_j^c} + \lambda_{j'}^c \frac{V_{i,j'}^c}{\tilde{u}_{j'}^c} = 0$$
This implies that
\[ T_j^c - T_j'^c = -\frac{1}{\sigma} \left[ \frac{u_j^c}{V'\left(\hat{\mathbf{p}}; z_j^c + T_j^c\right)} - \frac{u_j'^c}{V'\left(\hat{\mathbf{p}}; z_j'^c + T_j'^c\right)} \right] \]

When utility is homothetic and \( v^c(\ell) = \frac{\ell^{1+\varepsilon}}{1+\varepsilon} \), then \( V_I = f(\hat{\mathbf{p}}) \) is only a function of \( \hat{\mathbf{p}} \) and \( u_j^c = f(\hat{\mathbf{p}}) \left( \frac{z_j^c}{1+\varepsilon} + T_j^c \right) \). Then the above implies that
\[ T_j^c - T_j'^c = -\frac{z_j^c - z_j'^c}{(1+\sigma)(1+\varepsilon)} \]

\[ \square \]

A.4 Proof of Proposition 6

Proof. We prove this using the properties of the compensating differential model shown in section B.2. Based on the derivation in section B.2, the optimal taxation problem is given by
\[ \max \sum c \lambda^c \int W(z) d \left( e^{-\sum_j \mu_j^c \left[ \frac{1}{1+\varepsilon} \left( \frac{w_j^c}{q^c} \right)^{1+\varepsilon} + \frac{T^c}{q^c} \right]^{\sigma}} z^{-\sigma} \right) \]

subject to
\[ \sum c \left( \frac{\alpha_i^c}{q^c} \right)^{\gamma^c} (q_i^c)^{-\gamma^c} \sum_j \Lambda_j^c \left[ (q_i^c)^{-\varepsilon} (w_j^c)^{1+\varepsilon} + T^c \right] + \sum c \sum Q_{ki} = \sum c G_i^c \left( L_i^c, \{ Q_{ik}^c \} \right) \]
\[ \mu_i^c \frac{\partial}{\partial L} G_i^c \left( L_i^c, \{ Q_{ik}^c \} \right) = w_j^c, \forall i \in J^c \]
\[ \Lambda_j^c \left( \frac{w_j^c}{q^c} \right)^{\varepsilon^c} = L_j^c \]
\[ \Lambda_j^c = \frac{\mu_j^c \left[ 1 \left( \frac{w_j^c}{q^c} \right)^{1+\varepsilon^c} + \frac{T^c}{q^c} \right]^{\sigma}} {\sum j' \mu_j^{c'} \left[ 1 \left( \frac{w_j^{c'}}{q^{c'}} \right)^{1+\varepsilon^{c'}} + \frac{T^{c'}}{q^{c'}} \right]^{\sigma}} \]
\[ q^c = \left[ \sum_i \left( c_i^c \right)^{\gamma^c} (q_i^c)^{1-\gamma^c} \right]^{\frac{1}{1-\gamma^c}} \]

Consider a perturbation of \( w_i^c \) by \( \delta w_i^c \). This perturbation leads to a change in composition of workers and distribution of their utilities in each sector, a change in demand (through a direct change coming from wages, behavior respons of hours and another behavioral effect coming from a change in occupational choice), and a change in supply coming from the intensive margin of labor supply and occupational choice.

Here, we describe each effect:
1. Welfare effect of a change in \( w^c_i \): Recall that

\[
\Lambda^c_j = \frac{\mu^c_j}{\sum_k \mu^c_k} \left[ \frac{1}{1+\varepsilon^c} \left( \frac{w^c_i}{q^c} \right)^{1+\varepsilon^c} + \frac{T^c}{q^c} \right]^{\sigma^c}
\]

Hence,

\[
\frac{\delta \Lambda^c_j}{\Lambda^c_j} \frac{w^c_i}{\delta w^c_i} = \begin{cases} 
-\sigma (1 + \varepsilon^c) \Lambda^c_j \frac{1}{1+\varepsilon^c} \frac{\partial^c_i (q^c) - \varepsilon^c (w^c_i)^{1+\varepsilon^c}}{q^c} + T^c = -\sigma \Lambda^c_j \frac{z^c_i}{q^c u^c_i}, & j \neq i \\
\sigma (1 + \varepsilon^c) \frac{1}{1+\varepsilon^c} \frac{\partial^c_i (q^c) - \varepsilon^c (w^c_i)^{1+\varepsilon^c}}{q^c} (1 - \Lambda^c_j) = \sigma \frac{z^c_i}{q^c u^c_i} (1 - \Lambda^c_j) & j = i
\end{cases}
\]

where in the above \( z^c_i \) is average before tax income for workers in sector \( i \) and \( u^c_i \) is the expected utility of workers in sector \( i \). In other words

\[
z^c_i = \frac{(w^c_i)^{1+\varepsilon^c}}{(q^c)^{\varepsilon^c}}, u^c_i = \frac{1}{1 + \varepsilon^c} \left( \frac{w^c_i}{q^c} \right)^{1+\varepsilon^c} + \frac{T^c}{q^c}.
\]

The effect of this perturbation on welfare is given by – we let \( U^c = \sum_j \mu^c_j \left[ \frac{1}{1+\varepsilon^c} \left( \frac{w^c_i}{q^c} \right)^{1+\varepsilon^c} + \frac{T^c}{q^c} \right]^{\sigma^c} \)

\[
\lambda^c \sigma \Lambda^c_j \frac{z^c_i}{q^c u^c_i} \int [W(z) - W(z) U^c z^{-\sigma}] U^c z^{-1-\sigma} e^{-U^c z^{-\sigma}} dz \frac{\delta w^c_i}{w^c_i}
\]

Note that the conditional distribution of \( \nu \) among workers in sector \( j \) is given by \( \hat{H}^c_j (\nu) = e^{-\frac{\nu^c_i}{\Lambda^c_j} e^{-\varepsilon^c}} \) – see Appendix B.2. Therefore, the expectation of marginal utility of income among the workers in sector \( i \) is given by

\[
\frac{1}{q^c} \int_{-\infty}^{\infty} W'(u^c_i e^\nu) e^\nu dH^c_i (\nu) = \frac{1}{q^c} \int_{-\infty}^{\infty} W'(u^c_i e^\nu) e^\nu d \left( e^{-U^c(u^c_i e^\nu)^{-\sigma}} \right)
\]

\[
= \frac{1}{u^c_i q^c} \int_{0}^{\infty} W'(z) z d \left( e^{-U^c z^{-\sigma}} \right)
\]

\[
= \frac{1}{u^c_i q^c} \int_{0}^{\infty} W'(z) \sigma z^{-\sigma} U^c e^{-U^c z^{-\sigma}} d z
\]

Using integration by parts and assuming that \( W'(\infty) = 0 \), we can write the above as

\[
\frac{1}{u^c_i q^c} \int_{0}^{\infty} W'(z) \sigma z^{-\sigma} U^c e^{-U^c z^{-\sigma}} d z = \frac{\sigma}{u^c_i q^c} \int_{0}^{\infty} z^{-\sigma} U^c e^{-U^c z^{-\sigma}} d (W(z))
\]

\[
= -\frac{\sigma}{u^c_i q^c} \int W(z) \left[ -\sigma z^{-1-\sigma} + U^c z^{-1-2\sigma} \right] U^c e^{-U^c z^{-\sigma}} d z
\]

\[
= \frac{\sigma}{u^c_i q^c} \int W(z) \left[ 1 - U^c z^{-\sigma} \right] d \left( e^{-U^c z^{-\sigma}} \right)
\]

48
which is the same as the derivative of welfare with respect to \( w^c_i \) - multiplied by \( \lambda^c \Lambda^c_i z^c_i \). Therefore the welfare effect of an increase in wage \( w^c_i \) is given by
\[
\Lambda^c_i z^c_i \varepsilon^c \left[ \mathcal{W}^c \right] \delta w^c_i \]
where \( \mathcal{W}^c \) is the marginal utility of income and the expectation is taken over workers in sector \( i \) country \( c \).

2. The demand effect of a perturbation in \( w^c_i \): The effect on demand is given by
\[
\sum_k \rho_k s_k \left[ (1 + \varepsilon^c) \frac{\delta w^c_i}{w^c_i} z^c_i + \sum_j \Lambda^c_j \frac{\delta \Lambda^c_j}{\Lambda^c_j} I^c_j \right] = \\
\sum_k \rho_k s_k \left[ (1 + \varepsilon^c) z^c_i \Lambda^c_i + \sigma \frac{z^c_i}{Q^c u^c_i} \Lambda^c_i \sum_j \Lambda^c_j (I^c_j - I^c_j) \right] \frac{\delta w^c_i}{w^c_i}
\]
where \( I^c_j \) is the average disposable income for workers in sector \( j \) given by
\[
I^c_j = (q^c)^{-\varepsilon^c} \left( w^c_j \right)^{1+\varepsilon^c} + T^c
\]

3. The supply effect of a perturbation in \( w^c_i \): The effect on supply is given by
\[
\rho_i \frac{\partial G^c_i}{\partial L_i} \frac{\partial L_i}{\partial w^c_i} \delta w^c_i + \sum_j \rho_j \frac{\partial G^c_j}{\partial L_j} \frac{\partial L_j}{\partial \Lambda^c_j} \delta \Lambda^c_j = \\
\varepsilon^c \rho_i \frac{\partial G^c_i}{\partial L_i} \frac{L^c_i}{w^c_i} \delta w^c_i + \sum_j \rho_j \frac{\partial G^c_j}{\partial L_j} \frac{L^c_j}{w^c_i} \delta \Lambda^c_j = \\
\varepsilon^c \rho_i \frac{\partial G^c_i}{\partial L_i} L^c_i \frac{z^c_i}{w^c_i} \sum_{j \neq i} \rho_j \frac{\partial G^c_j}{\partial L_j} L^c_j + \sigma \rho_i (1 - \Lambda^c_i) \frac{z^c_i}{Q^c u^c_i} \frac{\partial G^c_i}{\partial L_i} L^c_i \frac{\delta w^c_i}{w^c_i}
\]

Note that in a competitive equilibrium
\[
w^c_i = (1 - t^{p,c}_i) \rho_i \frac{\partial G^c_i}{\partial L_i} \\
L^c_i = \frac{(w^c_i)^{\varepsilon^c}}{(Q^c)^{\varepsilon^c}} \sigma_i \\
\Lambda^c_i z^c_i = L^c_i w^c_i \\
\frac{z^c_i}{1 - t^{p,c}_i} \Lambda^c_i = \rho_i \frac{\partial G^c_i}{\partial L_i} L^c_i
\]
We can write the optimality condition
\[
\frac{z^c_i \Lambda^c_i \varepsilon^c \left[ \mathcal{W}^c \right]}{w^c_i} \delta w^c_i = 0
\]
We can write the above as:

\[
- \sum_k \rho_k s_k \left[ (1 + \varepsilon^c) + \frac{1}{Q^c u_i^c} \sum_j \Lambda_j^c \left( z_i^c - z_j^c \right) \right] + \varepsilon^c \frac{1}{1 - t_i^{p.c}} = 0
\]

Note that

\[
\sum_k \rho_k s_k = \sum_k \frac{q_k}{1 + t_k^c} s_k = 1 - \sum_k \frac{q_k t_k^c}{1 + t_k^c} s_k = 1 - r
\]

where \( r \) is the marginal revenue of consumption taxes from giving one unit of wealth to any individual. Since \( t_k^c \)'s are normalized to zero, \( r = 0 \).

We can thus write

\[
\begin{aligned}
\varepsilon^c \frac{1}{1 - t_i^{p.c}} &= \left[ 1 + \varepsilon^c + \frac{1}{Q^c u_i^c} \sum_j \Lambda_j^c \left( z_i^c - z_j^c \right) \right] \\
&\quad - \mathbb{E}_z^{i,c} [W^c] + \frac{\sigma}{Q^c u_i^c} \sum_j \Lambda_j^c \left( \frac{z_j^c}{1 - t_j^{p.c}} - \frac{z_i^c}{1 - t_i^{p.c}} \right)
\end{aligned}
\]

which concludes the proof. \( \square \)

B Other Derivations

B.1 A Quasi-Linear Example

Suppose that utility function in country \( c \) is given by

\[
x_i^c + \sum_{k \geq 2} \alpha_k^c x_k^{1 - \gamma_k^c} - v(\ell)
\]

with \( \alpha_k^c \geq 0 \). Note that under this utility function for any good \( k \geq 2 \), we must have that

\[
\alpha_k^c x_k^{-1/\gamma_k^c} = \frac{q_k^c}{q_i^c}
\]

Note that under these preferences, \( \varepsilon_{q,i,j} = 0 \) for all \( i \geq 2 \). Moreover, for all \( k \geq 2 \), \( \zeta_{i,k} = 0 \) if \( i \neq k \) while \( \zeta_{k,k} = -\gamma_k^c \). Hence, (??) becomes

\[
\gamma_i^c \frac{t_i^{p.c}}{1 + t_i^{p.c}} = \sum_j \int \left( 1 + W_j^c - r_j^c \right) dF_j^c.
\]
B.2 The Derivations of Sectoral Choice and Social Welfare

The fraction of workers in sector $i$ is given by

$$\Lambda_i = \int_{\nu \in \mathbb{R}^N} 1 \left[ \log u_i^\nu + \nu_i \geq \log u_j^\nu + \nu_j, \forall j \neq i \right] dH_1^\nu \cdots dH_{N_j}^\nu =$$

$$\int_{-\infty}^{\infty} \prod_{k \neq i} e^{-\mu_k e^{-\sigma(\nu_i + u_i^\nu - u_k^\nu)}} \sigma \mu_k e^{-\sigma \nu_i} e^{-\sigma u_i^\nu} d\nu_i =$$

$$\int_{-\infty}^{\infty} \prod_{k} e^{-\mu_k e^{-\sigma(\nu + \log u_i^\nu - \log u_k^\nu)}} \sigma \mu_k e^{-\sigma \nu} d\nu =$$

$$\int_{-\infty}^{\infty} e^{-\sum_{k} \mu_k (u_k^\nu)^\sigma (u_i^\nu)^{-\sigma e^{-\sigma \nu}}} \sigma \mu_i e^{-\sigma \nu} d\nu =$$

$$\int_{-\infty}^{\infty} \mu_i (u_i^\nu)^\sigma d \left( e^{-\sum_{k} \mu_k (u_k^\nu)^\sigma (u_i^\nu)^{-\sigma e^{-\sigma \nu}}} \right) =$$

$$\sum_{\nu \in \mathbb{R}^N} \left[ \log u_i^\nu + \nu_i \geq \log u_j^\nu + \nu_j, \forall j \neq i \right] W (u_i^{\nu,\nu}) dH_1^\nu \cdots dH_{N_j}^\nu$$

The above algebra also establishes that the conditional distribution of $\nu$ among workers who choose sector $j$ is the c.d.f. $H_j^\nu (\nu) = e^{-\frac{\nu^\nu_j}{\nu_j}}$.

As defined in the text, welfare is given by

$$\sum_{i=1}^{N} \int_{\nu \in \mathbb{R}^N} 1 \left[ \log u_i^\nu + \nu_i \geq \log u_j^\nu + \nu_j, \forall j \neq i \right] W (u_i^{\nu,\nu}) dH_1^\nu \cdots dH_{N_j}^\nu$$

We have

$$\sum_{i} \int_{\nu \in \mathbb{R}^N} 1 \left[ \log u_i^\nu + \nu_i \geq \log u_j^\nu + \nu_j, \forall j \neq i \right] W (u_i^{\nu,\nu}) dH_1^\nu \cdots dH_{N_j}^\nu =$$

$$\sum_{i} \int_{-\infty}^{\infty} \prod_{k \neq i} e^{-\mu_k e^{-\sigma(\nu_i + u_i^\nu - u_k^\nu)}} \sigma \mu_k e^{-\sigma \nu_i} e^{-\sigma u_i^\nu} W (u_i^{\nu,\nu_i}) d\nu_i =$$

$$\sum_{i} \int_{-\infty}^{\infty} e^{-\sum_{k} \mu_k (u_k^\nu)^\sigma (u_i^\nu)^{-\sigma e^{-\sigma \nu}}} \sigma \mu_i e^{-\sigma \nu} W (u_i^{\nu,\nu_i}) d\nu =$$

$$\sum_{i} \int_{-\infty}^{\infty} e^{-\frac{\mu_i (u_i^\nu)^\sigma}{\mu_i} e^{-\sigma \nu}} \sigma \mu_i e^{-\sigma \nu} W (u_i^\nu) d\nu =$$

$$\sum_{i} \Lambda_i W (u_i^\nu) \int_{-\infty}^{\infty} e^{-\left( \frac{\mu_i}{\mu_i} e^{-\sigma \nu} \right)} d\nu =$$

$$\sum_{i} \int_{0}^{\infty} \Lambda_i W (z) d \left( e^{-\nu e^{-\sigma \nu}} \right)$$

51
Finally, the elasticity of sectoral choice with respect to wages can be calculated as follows:

\[ \Lambda_i^c = \frac{\mu_i^c}{\sum_j \mu_j^c} \left[ \frac{1}{1+\varepsilon^c} \left( \frac{w_i^c}{q^c} \right)^{1+\varepsilon^c} + \frac{T^c}{q^c} \right] \]

where we have assumed that \( \varepsilon^c = 0 \) and used the fact that \( \ell_i^c(\theta) = \left( \frac{w_i^c \theta}{q^c} \right)^{\varepsilon^c} \). We thus have that

\[
\frac{\partial \Lambda_i^c}{\partial w_j^c} = -\frac{1}{w_j^c} \sigma \frac{\mu_j^c}{\sum_k \mu_k^c} \left[ \frac{1}{1+\varepsilon^c} \left( \frac{w_j^c}{q^c} \right)^{1+\varepsilon^c} + \frac{T^c}{q^c} \right] \sum_k \mu_k^c \left[ \frac{1}{1+\varepsilon^c} \left( \frac{w_k^c}{q^c} \right)^{1+\varepsilon^c} + \frac{T^c}{q^c} \right]^{\sigma} \\
+ 1[i = j] \frac{1}{w_j^c} \sigma \frac{1}{1+\varepsilon^c} \left( \frac{w_j^c}{q^c} \right) \left[ \frac{1}{1+\varepsilon^c} \left( \frac{w_j^c}{q^c} \right)^{1+\varepsilon^c} + \frac{T^c}{q^c} \right] \sum_k \mu_k^c \left[ \frac{1}{1+\varepsilon^c} \left( \frac{w_k^c}{q^c} \right)^{1+\varepsilon^c} + \frac{T^c}{q^c} \right]^{\sigma} \\
= \sigma \frac{1}{1+\varepsilon^c} \left( \frac{w_j^c}{q^c} \right)^{1+\varepsilon^c} + \frac{T^c}{q^c} \left( -\Lambda_j^c \Lambda_i^c + 1[i = j] \Lambda_i^c \right)
\]

This implies that (7) holds.

### B.3 Trade and Inequality in Laissez-Faire

In the closed economy, total demand for good \( i \) in country \( c \) is given by

\[ \frac{\alpha_i^c}{(P_i^c)^{1-\gamma}} I^c \]

where \( I^c \) is total income and \( P^c \) is the consumer price index in the country. The supply of good \( i \) is given by

\[ A_i^c \Lambda_i^c \ell_i^c \]

Using the fact that wages satisfy \( w_i^c = A_i^c p_i^c \) and labor supply satisfies \( \ell_i^c = \left( \frac{w_i^c}{P_i^c} \right)^{\varepsilon^c} \) and \( \Lambda_i^c = \mu_i^c \left( \frac{w_i^c}{1+\varepsilon^c} \right)^{\sigma} / \sum_k \mu_k^c \left( \frac{w_k^c}{1+\varepsilon^c} \right)^{\sigma} \). Therefore, total supply of \( i \) is given by

\[
\frac{\mu_i^c (p_i^c A_i^c)^{(1+\varepsilon^c)(1+\sigma)}}{\sum_k \mu_k^c (p_k^c A_k^c)^{(1+\varepsilon^c)(1+\sigma)}} \frac{1}{(P_i^c)^{-\varepsilon}}
\]

In equilibrium this must be equal to demand and therefore

\[ \mu_i^c (A_i^c)^{(1+\varepsilon^c)(1+\sigma)} (p_i^c)^{\gamma-1+(1+\varepsilon^c)(1+\sigma)} \propto \alpha_i^c \]

52
or
\[ (p_i A_i^c)^{\gamma+\varepsilon+\sigma+\varepsilon_0} \propto \frac{\alpha_i^\gamma}{\mu_i^c} (A_i^c)^{\gamma-1} \Rightarrow z_{i, AUT}^c = \frac{(p_i A_i^c)^{1+\varepsilon}}{(Pc)^{\varepsilon}} \propto \left[ \frac{\alpha_i^\gamma}{\mu_i^c} (A_i^c)^{\gamma-1} \right]^{1+\varepsilon\gamma+\varepsilon+\sigma+\varepsilon_0} \]
which implies (4).

In the open economy with no trade cost, total demand for good \( i \) is given by
\[ \frac{\alpha_i^\gamma}{P^{1-\gamma}} \sum_c I^c \]
where \( P \) is the world price index.

Using similar formulas as above, the total supply of \( i \) is given by
\[ \sum_c \mu_i^c (p_i A_i^c)^{(1+\varepsilon)(1+\sigma)} \frac{1}{\sum_k \mu_k^c (p_k A_k^c)^{\sigma(1+\varepsilon)}} p_i \]
If we let \( \zeta^c \) be given by
\[ \zeta^c = \frac{1}{\sum_k \mu_k^c (p_k A_k^c)^{\sigma(1+\varepsilon)}} \]
Then total supply of good \( i \) is proportional to
\[ \sum_c \zeta^c \mu_i^c (p_i A_i^c)^{(1+\varepsilon)(1+\sigma)} p_i^{-1} \]
Hence,
\[ p_i^{\gamma+\varepsilon+\sigma+\varepsilon_0} \sum_c \zeta^c \mu_i^c (A_i^c)^{(1+\varepsilon)(1+\sigma)} \propto \alpha_i^\gamma \]
As a result,
\[ z_{i, T}^c \propto \left( \frac{\alpha_i^\gamma (A_i^c)^{\gamma+\varepsilon+\sigma+\varepsilon_0}}{\sum_{c'} \zeta^{c'} \mu_i^{c'} (A_i^{c'})^{(1+\varepsilon)(1+\sigma)}} \right)^{\frac{1+\varepsilon\gamma+\varepsilon+\sigma+\varepsilon_0}{\gamma+\varepsilon+\sigma+\varepsilon_0}} \]
which is the same as (5).


Consider the extended version of Eaton and Kortum (2002) in section (5). Here, we describe the formulas that characterize optimal taxes.

In this model, the optimal taxation without sector-specific income taxes are given by
\[ \max_{\psi_j^c, X_j^c, T^c, \tilde{w}_j^c} \frac{1}{\sigma} \sum_{c=1}^{N_c} \lambda^c M^c \log \left( \sum_{j=1}^{N_j} \mu_j^c \left( \frac{(w_j^c)^{1+\varepsilon}}{(Pc)^{1+\varepsilon}} + \frac{T_j^c}{Pc} \right)^{\sigma} \right) \]
s.t.

\[ \alpha_j^c \sum_{i=1}^{N_d} M^c \Lambda_i^c \left( \frac{(w_i^c)^{1+\epsilon}}{(P_c)^{1+\epsilon}} + T^c \right) + \sum_{k=1}^{N_j} \gamma_{k,j}^c \sum_{c'=1}^{N_c} \pi_{k}^{c',c} X_{k}^{c'} = X_{j}^c \]

\[ M^c \Lambda_j^c \left( \frac{(w_j^c)^{1+\epsilon}}{(P_c)^{1+\epsilon}} \right) = (1 - t_{j}^{p,c}) \chi_j^c \sum_{c'=1}^{N_c} \pi_{j}^{c,c} X_{j}^{c'} \]

\[ \Lambda_j^c = \frac{\mu_j^c (w_j^c)^{\sigma(1+\epsilon)}}{\sum_{i=1}^{N_j} \mu_i^c (w_i^c)^{\sigma(1+\epsilon)}} \]

\[ \psi_j^c = \left( \frac{w_j^c}{\chi_j^c (1 - t_{j}^{p,c})} \right) \prod_{k=1}^{N_j} \left( \frac{P_k^c}{\gamma_{j,k}^c} \right) \]

\[ \pi_{j}^{c,c'} = \frac{\chi_j^c (\tau_{j}^{c,c'} \psi_j^c)^{-\nu}}{\sum_{c''=1}^{N_c} \lambda_j^{c''} (\tau_{j}^{c,c''} \psi_{j}^{c''})^{-\nu}} \]

\[ P_j^c = \Gamma \left( 1 + \frac{1 - \eta_j}{\nu} \right)^{1-\eta_j} \left[ \sum_{c'=1}^{N_c} \chi_j^c (\tau_{j}^{c,c'} \psi_j^c)^{-\nu} \right]^{-\frac{1}{\nu}} \]

\[ P_c = \prod_{j=1}^{N} \left( \frac{P_j^c}{\alpha_j^c} \right) \]

In the above problem, using the definition of unit cost, we can write

\[ \frac{\chi_j^c (1 - t_{j}^{p,c})}{w_j^c} = \left[ \psi_j^c \prod_{k=1}^{N_j} \left( \frac{P_k^c}{\gamma_{j,k}^c} \right)^{-\gamma_{j,k}^c} \right]^{\frac{1}{\gamma_j^c}} \]

replacing this in the labor market clearing condition, we have the following optimal policy problem

\[ \max_{\psi_j^c, \chi_j^c, T^c, \alpha_j^c} \frac{1}{\sigma} \sum_{c=1}^{N_c} \lambda^c M^c \log \left( \sum_{j=1}^{N_j} \mu_j^c \left( \frac{(w_j^c)^{1+\epsilon}}{(P_c)^{1+\epsilon}} + T^c \right)^{\sigma} \right) \]

subject to
\begin{align*}
\alpha_j \sum_{i=1}^{N_j} M^c \Lambda_i^c \left( \frac{(w^c_i)^{1+\epsilon}}{(P^c)^{1+\epsilon}} + T^c \right) + \sum_{k=1}^{N_j} \gamma_{k,j} \sum_{c'=1}^{N_c} \pi^{c,c'}_{k,j} X^c_{k,j} &= X^c_j \\
M^c \Lambda_j^c \left( \frac{(w^c_j)^{1+\epsilon}}{(P^c)^{1+\epsilon}} + T^c \right) &= (1 - \mu_j^c) \chi_j^c \sum_{c'=1}^{N_c} \pi^{c,c'}_{j,c'} X^c_{j,c'} \\
\Lambda_j^c &= \frac{\mu_j^c \left( \frac{(w^c_j)^{1+\epsilon}}{(P^c)^{1+\epsilon}} + T^c \right)}{\sum_{i=1}^{N_j} \mu_i^c \left( \frac{(w^c_i)^{1+\epsilon}}{(P^c)^{1+\epsilon}} + T^c \right)} \sigma \\
\pi^{c,c'}_{j,c'} &= \frac{\lambda_j^{c'} \left( \tau_{c,c'}^j \psi_{c'}^j \right)^{1-\nu}}{\sum_{c''=1}^{N_c} \lambda_j^{c''} \left( \tau_{c,c''}^j \psi_{c''}^j \right)^{1-\nu}} \\
P_j^c &= \Gamma \left( 1 + \frac{1 - \eta_j}{\nu} \right) \frac{1}{\nu} \left[ \sum_{c'=1}^{N_c} \chi_j^{c'} \left( \tau_{c,c'}^j \psi_{c'}^j \right)^{1-\nu} \right] \frac{1}{\nu} \\
P^c &= \prod_{j=1}^{N} \left( \frac{P^c_j}{\alpha_j^c} \right)^{\alpha_j^c} \text{FOC w.r.t. } Q^c \\
\sum_j \rho_j^c \alpha_j^c \frac{I^c}{Q^c} - \zeta^c &= 0 \\
\text{FOC w.r.t. } \hat{w}^c_i: \\
\lambda^c M^c \Lambda_i^c \frac{(\hat{w}^c_i)^{1+\epsilon}}{\hat{w}^c_i} \frac{\delta \hat{w}^c_i}{\hat{w}^c_i} - \epsilon \sum_j \rho_j^c \alpha_j^c Q^c M^c \Lambda_i^c \frac{(\hat{w}^c_i)^{1+\epsilon}}{\hat{w}^c_i} \frac{\delta \hat{w}^c_i}{\hat{w}^c_i} - \sum_j \rho_j^c \alpha_j^c Q^c M^c \sum_k \frac{\delta \Lambda_k^c}{\Lambda_k^c} I_k^c Q^c \\
-\epsilon \xi_j^c M^c \Lambda_j^c \frac{(\hat{w}^c_j)^{1+\epsilon}}{\hat{w}^c_j} \frac{\delta \hat{w}^c_i}{\hat{w}^c_j} - M^c \sum_j \xi_j^c \frac{\delta \Lambda_j^c}{\Lambda_j^c} \frac{\hat{w}^c_j}{\hat{w}^c_i} = 0 \\
\text{Note that} \\
\frac{\delta \Lambda_j^c \hat{w}^c_i}{\Lambda_j^c \delta \hat{w}^c_i} &= -\sigma \frac{(\hat{w}^c_i)^{1+\epsilon}}{u^c_i} - \Lambda_j^c + \sigma \frac{(\hat{w}^c_i)^{1+\epsilon}}{u^c_i} \mathbf{1}[i = j] = \sigma \frac{(\hat{w}^c_i)^{1+\epsilon}}{u^c_i} \left( \mathbf{1}[i = j] - \Lambda_j^c \right) \\
&= \sigma \frac{z_i^c}{Q^c u^c_i} \left( \mathbf{1}[i = j] - \Lambda_j^c \right) 
\end{align*}
Hence,

\[ \Lambda_i^c W_i^c z_i^c - (1 + \varepsilon) \sum_j \rho_j^c \alpha_j^c \Lambda_i^c z_i^c \]

\[ - \sum_j \alpha_j^c \rho_j^c \sum_k \sigma \frac{z_i^c}{Q^c u_i^c} (1 [i = k] - \Lambda_i^c) \Lambda_k^c I_k^c \]

\[ -\varepsilon \xi_i^c \Lambda_i^c (\hat{w}_i^c)^\varepsilon \sum_j \sigma \frac{z_i^c}{Q^c u_i^c} (1 [i = j] - \Lambda_i^c) \Lambda_j^c (\hat{w}_j^c)^\varepsilon = 0 \]

\[ \Lambda_i^c W_i^c z_i^c - (1 + \varepsilon) \sum_j \rho_j^c \alpha_j^c \Lambda_i^c z_i^c \]

\[ - \sum_j \alpha_j^c \rho_j^c \sigma \frac{z_i^c}{Q^c u_i^c} \Lambda_i^c \Lambda_j^c (I_i^c - I_j^c) \]

\[ -\varepsilon \xi_i^c \Lambda_i^c (\hat{w}_i^c)^\varepsilon - \sigma \frac{z_i^c}{Q^c u_i^c} \left( \xi_i^c \Lambda_i^c (\hat{w}_i^c)^\varepsilon - \Lambda_i^c \sum_j \Lambda_j^c \xi_j^c (\hat{w}_j^c)^\varepsilon \right) = 0 \]

\[ W_i^c - (1 + \varepsilon) \sum_j \rho_j^c \alpha_j^c \]

\[ -\sigma \frac{1}{Q^c u_i^c} \sum_j \alpha_j^c \rho_j^c \Lambda_j^c (I_i^c - I_j^c) \]

\[ -\varepsilon \xi_i^c \frac{1}{Q^c \hat{w}_i^c} - \sigma \frac{1}{Q^c u_i^c} \sum_j \Lambda_j^c \left[ \xi_i^c (\hat{w}_i^c)^\varepsilon - \xi_j^c (\hat{w}_j^c)^\varepsilon \right] = 0 \]

Suppose

\[ \xi_i^c \Phi_i^c = \Gamma_i^c \]

Suppose we have \( \Gamma_i^c \), then

\[ W_i^c - (1 + \varepsilon) \sum_j \rho_j^c \alpha_j^c \]

\[ -\sigma \frac{1}{Q^c u_i^c} \sum_k \alpha_k^c \rho_k^c \sum_j \Lambda_j^c (I_i^c - I_j^c) \]

\[ -\varepsilon \frac{\Gamma_i^c}{\Phi_i^c w_i^c} \frac{1}{Q^c u_i^c} \sum_j \Lambda_j^c \left[ \xi_i^c (\hat{w}_i^c)^\varepsilon - \xi_j^c (\hat{w}_j^c)^\varepsilon \right] = 0 \]

Note that

\[ \frac{1}{\Phi_i^c w_i^c} = \frac{1}{\chi_i^c (1 - n_i^c)} \]
Then, we can write the above as

\[
\varepsilon \frac{1}{1 - t_{i}^{p,e}} \Gamma_{i}^{e} \chi_{i}^{e} = \hat{W}_{i}^{e} - (1 + \varepsilon) \sum_{j} \rho_{j}^{c} \alpha_{j}^{e} \tag{21}
\]

where in the above

\[
\hat{W}_{i}^{e} = \frac{\chi_{i}^{e}}{Q_{c} u_{i}^{e}}
\]

We can show that

\[
\rho_{i}^{e} = \rho, \forall i, c \quad \xi_{i}^{c} \Phi_{i}^{e} = -\chi_{i}^{e} \rho
\]

Note that this implies that

\[
\xi_{i}^{c} = -\frac{\chi_{i}^{e}}{\Phi_{i}^{e}} \rho = -\frac{w_{i}^{c}}{1 - t_{i}^{p,e}} \rho \rightarrow \xi_{i}^{c} (\hat{w}_{i}^{e})^{e} = -\rho \frac{z_{i}^{c}}{1 - t_{i}^{p,e}}
\]

We can thus write the above formula as

\[
-\varepsilon \frac{1}{1 - t_{i}^{p,e}} \rho = \hat{W}_{i}^{e} - (1 + \varepsilon) \rho
\]

\[
- \sigma \frac{1}{Q_{c} u_{i}^{e}} \rho \sum_{j} \Lambda_{j}^{c} (I_{i}^{e} - I_{j}^{c})
\]

\[
+ \sigma \frac{1}{Q_{c} u_{i}^{e}} \rho \sum_{j} \Lambda_{j}^{c} \left( \frac{z_{i}^{c}}{1 - t_{i}^{p,e}} - \frac{z_{j}^{c}}{1 - t_{j}^{p,e}} \right)
\]

\[
\varepsilon \frac{1}{1 - t_{i}^{p,e}} = 1 + \varepsilon - \hat{W}_{i}^{e}
\]

\[
+ \sigma \frac{1}{Q_{c} u_{i}^{e}} \sum_{j} \Lambda_{j}^{c} (I_{i}^{e} - I_{j}^{c})
\]

\[
- \sigma \frac{1}{Q_{c} u_{i}^{e}} \sum_{j} \Lambda_{j}^{c} \left( \frac{z_{i}^{c}}{1 - t_{i}^{p,e}} - \frac{z_{j}^{c}}{1 - t_{j}^{p,e}} \right)
\]

This formula is identical to the one shown in Proposition (6).