Non-homothetic preferences, parallel imports and the extensive margin of international trade

Reto Foellmi
Christian Hepenstrick
Josef Zweimüller

10-09

July 2010

DISCUSSION PAPERS

Schanzeneckstrasse 1
Postfach 8573
CH-3001 Bern, Switzerland
http://www.vwi.unibe.ch

Non-homothetic preferences, parallel imports and the extensive margin of international trade

Reto Foellmi[*], Christian Hepenstrick[†], Josef Zweimüller[‡]

July 20, 2010

Abstract

We study international trade in a model where consumers have non-homothetic preferences and where household income restricts the extensive margin of consumption. In equilibrium, monopolistic producers set high (low) prices in rich (poor) countries but a threat of parallel trade restricts the scope of price discrimination between countries. The threat of parallel trade allows differences in per capita incomes to have a strong impact on the extensive margin of trade, whereas differences in population sizes have a weaker effect. We also show that the welfare gains from trade liberalization are biased towards rich countries. We extend our model to more than two countries; to unequal incomes within countries; and to more general specifications of non-homothetic preferences. Our basic results are robust to these extensions.

JEL classification: F10, F12, F19

Keywords: Heterogenous markups, non-homothetic preferences, parallel imports, extensive margin of trade

[*]University of Bern and CEPR, Department of Economics, Schanzeneckstrasse 1, CH - 3001 Bern, Tel: ++41-31-631 55 98, Fax: ++41-31-631 39 92, e-mail: reto.foellmi@vwi.unibe.ch

[†]University of Zurich, Institute for Empirical Research in Economics, Muehlebachstrasse 86, CH - 8008 Zürich, Tel: ++41-44-634 37 25, Fax: ++41-44-634 49 07, e-mail: hepenstr@iew.uzh.ch

[‡]University of Zurich and CEPR, Institute for Empirical Research in Economics, Muehlebachstrasse 86, CH - 8008 Zürich, Tel: ++41-44-634 37 24, Fax: ++41-44-634 49 07, e-mail: zweim@iew.uzh.ch. Josef Zweimüller is also associated with CESifo and IZA.

1 Introduction

Theories of international trade typically assume that consumers have homothetic preferences, showing why product differentiation, increasing returns, and firm heterogeneity are crucial in explaining the extensive and intensive margins of international trade (e.g. Krugman, 1980, Melitz, 2003, Helpman, Melitz, Rubinstein, 2007, Chaney, 2008). While both casual observation and econometric analyses of consumer budgets suggest that homothetic preferences cannot be defended on empirical grounds, their nice aggregation properties and high tractability make them an ideal tool for studying settings in which technology rather than demand factors are the main driving force of aggregate outcomes.

The assumption of homothetic preferences, however, is clearly inappropriate for studying how the composition of aggregate income affects consumption and trade patterns. Consider two countries, Austria and Nigeria. In 2008, their PPP-adjusted national income was roughly of the same order of magnitude (311 bill US $ and 281 bill US $, respectively). While Austria is small and rich, with a population of 8.4 mill and per capita income of 37,680 US $, Nigeria is large and poor, with a population of 152 mill and per capita income of 1,940 US $. Should we expect the two countries to display similar economy-wide demands for a given set of consumer goods? Homothetic preferences predict that the representative Nigerian consumer purchases the same menu of goods as the representative Austrian consumer, but in quantities that are 95 percent lower. If this were so, trade patterns are unaffected by the composition of aggregate income and exclusively shaped by supply conditions such as comparative advantages, differences in factor endowments, trade costs, and other technological asymmetries.

In this paper, we explore the implications of non-homothetic preferences in the context of the "new" trade theory framework. While the supply side of our model is identical to the basic Krugman (1980) framework, we deviate from this framework by introducing non-homothetic preferences in a very stylized way: we assume that consumer goods are indivisible and either consumed in unit quantity or not consumed at all.[1] Prima facie, this assumption may seem overly simplistic. However it is a "natural" deviation from the standard CES-framework in the following sense. In our framework, consumption choices are about the number of goods – the extensive margin of consumption – while a choice along the intensive margin is ruled out by assumption. This is orthogonal to the standard CES framework where consumption choices only affect the *intensive* margin of consumption and a choice along the along the extensive margin is ruled out by infinite reservation prices (hence even the poorest household

[1] Preferences of this type were used, inter alia, by Murphy, Shleifer and Vishny (1989) to study how demand composition affects technology choices in the development process, by Matsuyama (2000) to study the role of non-homotheticities in Ricardian trade, and by Foellmi and Zweimüller (2006) to study the relationship between inequality and growth.

will consume all goods in positive, albeit tiny, amounts).

Adopting this stylized way of introducing non-homotheticities provides a simple and tractable framework that leads to equilibrium outcomes quite different from the standard model. To keep things simple and transparent we confine the basic analysis to the most simple case of two countries. *First*, we show that, when per capita endowments of the two countries are very similar, the world economy ends up in a *full trade* equilibrium. In such an equilibrium, all goods are internationally traded and consumed in both countries. In contrast, when per capita endowments of the two countries are sufficiently unequal, a *partial trade* equilibrium emerges. The reason is that only households in the rich country consume all goods produced worldwide whereas households in the poor country can afford only a subset of all goods. In a partial trade equilibrium, the fraction internationally traded goods increases in the similarity of per capita endowments. This result is reminiscent of the "Linder hypothesis", according to which more similar countries trade more intensively with each other.

Second, the partial trade regime provides us with a simple general equilibrium framework of *parallel trade*. The partial trade equilibrium is supported by the price setting behaviors of monopolistic firms. With indivisible goods, the highest price a firm can charge is the representative consumer's willingness to pay which is finite. But this may create arbitrage opportunities. Consider a US firm selling its good both in the US and in China. When price differences are sufficiently large, arbitrage traders will purchase the good cheaply on the Chinese market, ship it back and underbid local producers on the US market. Anticipating this, US firms either set a price such that the incentive for parallel trade vanishes; or US firms do not supply their product on the world market, but therefore can charge the high price in the US. The general equilibrium perspective of our model makes the latter fraction of firms endogenous. This effect is typically not considered in partial equilibrium settings of parallel trade but has a potentially important impact on trade patterns.

Third, we make precise the relative importance of population sizes and per capita endowments for trade patterns. The country with a large population is more productive because a smaller fraction of resources is wasted due to (iceberg) trade costs. When per capita endowment differences are small, a higher population can compensate a lower per capita endowment so that a world economy that is initially trapped in a partial trade equilibrium may switch to a full trade equilibrium as a result of population growth. However, when per capita endowment differences are high, a larger population can never fully compensate for a small per capita endowment. In that case, the world economy remains trapped in partial trade even when the population grows very large. Hence the impact of population size differs crucially from the impact of per capita endowments. A higher degree of similarity in per capita endowments lets the world economy always escape the partial trade regime, whatever the differences in population size of the two countries. In this sense per capita endowments are a more significant determinant of the trade regime than population sizes.

A *fourth* main result of our model concerns the welfare effects of trade. Comparing trade to autarky it turns out that in a full trade equilibrium the poor country gains relatively more from trade; in a partial trade equilibrium, however the rich country gains more. Moreover, a trade liberalization (a reduction in iceberg trade costs) increases welfare of consumers in both countries when the world economy is in a full trade equilibrium. However, in a partial trade equilibrium, a trade liberalization is beneficial for the rich country but actually *hurts* the poor country. The reason is that trade liberalizations deteriorate the poor country's terms of trade, because international price discrimination becomes more limited. Exporters of the poor country need to lower the prices they charge in the rich country to inhibit parallel trade, whereas exporters of the rich country have no need to adjust their export prices. However, they must adjust the prices they charge in their home market. This makes selling on the rich market at unrestricted high prices more attractive - thus, in the new partial trade equilibrium more firms of the rich country will concentrate their sales exclusively only on the home market and less products will be available on the world market. Consumers in the rich country benefit from the decreasing prices of the internationally traded goods, whereas consumers in the poor country are confronted with a lower range of import goods at unchanged prices.

Our basic model is simple enough to be extended in various directions. We first look at a world with more than two countries. It turns out that in a multi-country world our result that a trade liberalization decreases rather than increases trade needs to be qualified. With many rich and many poor countries, a trade liberalization stimulates trade and welfare due to more trade within the rich North and within the poor South. However, it decreases overall trade (and increases the welfare-gap) between the Northern and the Southern region. While the North gains for sure, also the South may gain when within-South trade increases more strongly than North-South trade falls.

Second, we allow for heterogenous trade costs. When trade costs differ for the various products (but are not too large to inhibit trade at all), trade liberalization implies that goods with high (low) trade costs will be traded more (less). The reason is that high-trade-cost producers can sufficiently price discriminate hence they have an incentive to sell their product also in the poor country. For a low-trade-cost producer a trade liberalization implies fiercer price competition on world markets. Therefore more low-trade cost producers will decide to sell their product exclusively on the rich home market and not to sell abroad. In such a situation, it depends on the relative importance of high-cost and low-cost producers whether a trade liberalization stimulates or dampens international trade.

In a third extension we look at the impact of policy-restrictions on parallel trade. Our basic model assumes "international exhaustion" in which case parallel imports are not legally forbidden. The holder of a product's property right (patent, trademark, copyright) can no longer exercise his property right once this product is sold either on the home market or on the world market - his property rights are exhausted. In contrast, many countries have implemented "national/regional exhaustion" in which case the property right runs out when the product is sold on the home market, but does not run out when the product is sold abroad. It turns out that restrictions to parallel trade help consumers in the poor country but hurt consumers in the rich country. The reason is a general equilibrium effect. Stronger parallel trade restrictions encourage producers of the rich country to sell their product abroad while being able to charge high prices at home. This tends to improve the terms of trade for the poor country.

The fourth extension explores the consequences of income inequality. In such a context the level of trade costs and the extent of within-country inequality determine the equilibrium outcome. For low trade costs there are producers charging high prices selling to the rich at home and abroad and other producers charging low prices selling to all households. Interestingly, in such a situation with low trade costs, where inequality arises within countries and not between countries, lower trade costs actually benefit the poor.

Finally, we show that our model extends to a more general class of preferences where consumers have a choice not only along the extensive margin but also along the intensive margin. We demonstrate that, with more general specifications of preferences, a partial trade equilibrium emerges provided that (i) the derived individual demand functions feature finite reservation prices (so that some consumers optimally choose not to buy certain goods when prices are too high) and (ii) demand elasticities decrease (and mark-ups increase) along the demand function. This makes an equilibrium possible where firms are indifferent between selling at high prices and small quantities in rich economies or low prices and large quantities on the world market so that a partial equilibrium (supported by a threat of parallel trade) emerges.

Several previous papers have incorporated non-homothetic preferences into the new trade theory framework. The classical contributions are Markusen (1986) and Bergstrand (1990) who stick to CES-preferences for differentiated products but introduce non-homotheticities through a homogenous product with a minimum consumption requirement.[2] Several recent

[2] Important empirical contributions include Hunter and Markusen (1988), Hunter (1991), Francois and Kaplan (1996), Dalgin, Mitra, and Trindade (2008), and Fieler (2010). Mitra and Trindade (2005) use nonhomothetic preferences over the industry aggregates to study how income inequality affects trade patterns. Chung (2005) used quasihomothetic preferences to address Trefler's (1995) missing trade puzzle. Falkinger (1990) uses non-homothetic preferences in a dynamic innovator-imitator model. Flam and Helpman (1987) consider qualitative product differentiation in a North-South model. This model has been extended by Choi, Hummels, and Xiang

papers abandon the CES-assumption and instead introduce variable elasticity of substitution (VES-) preferences. One approach, followed by Markusen (2010) and Simonovska (2010) aggregates differentiated consumer goods with a Stone-Geary subutility (with negative required consumption levels). This formulation implies that firms charge higher prices in richer markets, an outcome in line with empirical evidence (see e.g. Hsieh and Klenow 2007, Simonovska 2010, and Manova and Zhang 2009). Sauré (2009) also uses a Stone-Geary subutility and studies how heterogeneous trade costs affect trade patterns among symmetric countries. Behrens and Murata (2009) explore the pro-competitive effects of free trade when consumers have CARA-preferences. They find that trade reduces mark-ups and that low-income countries gain more from trade than high income countries. These papers focus on symmetric equilibria, i.e. equilibria where all goods are consumed by all households worldwide.[3] Their approaches differ from that of our paper, which focuses on the (asymmetric) partial trade equilibrium where some goods are consumed by all households worldwide, whereas other goods are only affordable to households in the rich country.

As mentioned above, our paper contributes to the literature by presenting a general equilibrium model of parallel trade. A large partial-equilibrium literature has explored the determinants and consequences of parallel trade (see Maskus 2000, and Ganslandt and Maskus 2007, for surveys). The empirical relevance and importance of parallel trade is undisputed. The question whether parallel imports should be permitted or not (or inhibited by appropriate policies) triggers hot political debates in many countries. While empirical evidence on the quantitative importance of parallel trade is hard to get, existing estimates suggest that parallel imports are quantitatively important. A large body of empirical evidence has looked at the pharmaceuticals market, where the pros and cons of parallel trade are most obvious (see Ganslandt and Maskus, 2004, for an interesting study of parallel trade on prices of pharmaceuticals in the EU). However, parallel trade is quantitatively important in many other industries. For instance, KPMG (2003) estimates that grey market sales of IT products could exceed USD 40 billion annually and that price advantages drive grey market activity. According to the estimates of the National Economics Research Association (NERA), parallel imports account for between 5% and 20% of trade within the EU for goods such as consumer electronics, cosmetics and perfumes, musical recordings, and soft drinks (NERA 1999). In other words, parallel

(2006), who focus on the role of income distribution in determining the trade patterns. Krishna and Yavas (2005) used consumption indivisibilities in combination with labor market imperfection to explain possible losses from trade in transition economies.

[3] Other papers that give up the standard CES framework have studied the role of income inequality on trade patterns. Fajgelbaum, Grossman, and Helpman (2009) use a nested logit demand system in which income distribution affects quality choice and patterns of trade. Desdoigts and Jaramillo (2009) adopt Lancaster's ideal variety approach to study the impact of inequality on trade patterns.

trade is an important phenomenon and relevant in many markets. It is therefore interesting to study the determinants and consequences of parallel trade (and/or the threat of it) in a general equilibrium framework.[4]

The remainder of this paper is organized as follows. In the next section, we present the basic assumptions and discuss the consumer behavior with non-homothetic preferences. Section 3 first discusses the closed economy case and then applies our basic framework to study patterns of trade among equally large but unequally rich countries. Section 4 discusses the role of population size versus per capita incomes. Section 5 extends the model to other relevant settings such as a multi-country world, restrictions to parallel trade, income inequality within countries. Section 6 discusses more general specifications of preferences. Section 7 concludes.

2 The Model

The economy is populated by $\mathcal{P}$ identical households. Each household is endowed with L units of labor, the only production factor. Labor is perfectly mobile within countries and immobile across countries. The labor market is competitive and the wage is W. Hence household income is $y = WL$. Production requires a fixed labor input F to set up a new firm and a variable labor input $1/a$ to produce one unit of output, the same for all firms.[5] Producing good j in quantity $x(j)$ thus requires a total labor input of $F + x(j)/a$. Product markets are imperfectly integrated in the sense that trade costs accrue when goods are traded internationally. Iceberg trade costs imply that $\tau \geq 1$ units have to be shipped to the other country in order for 1 unit to arrive at the destination.

Households spend their income on a continuum of differentiated goods, indexed by j. We assume that good j yields positive utility only for the first unit and zero utility for any additional units. Hence consumption is a binary choice: either you buy or you don't buy. Denote an indicator $x(j)$ that takes value 1 if good j is purchased and value 0 if not. Then utility takes the simple form

$$U = \int_0^\infty x(j)dj, \quad \text{where } x(j) \in \{0,1\}. \tag{1}$$

Notice that utility is additively separable and that the various goods enter symmetrically. Hence the household's utility is given by the number of consumed goods.

Now consider a household with income y which can choose among (a measure of) N goods

[4] Note that due to the static setting, we do not need to introduce patents. The design of patents is crucial for the outcomes in a dynamic setting. Grossman and Lai (2004, 2006) discuss these models.

[5] An extensive literature has documented the importance of productivity differences between firms. While relaxing the assumption of homogeneous firms is straightforward, we stick to it in order to keep the supply side of the model as simple as possible, allowing us to focus on the new effects due to the demand side.

that are supplied at prices $\{p(j)\}$.[6] The problem is to choose $\{x(j)\}$ to maximize the objective function (1) subject to the budget constraint $\int_0^N p(j)x(j)dj = y$. Denoting λ as household's marginal utility of income, the first order condition can be written as

$$\begin{aligned} x(j) &= 1 \text{ if } 1 \geq \lambda p(j) \\ x(j) &= 0 \text{ if } 1 < \lambda p(j) \end{aligned}$$

Rewriting this condition as $1/\lambda \geq p(j)$ yields the simple rule that the household will purchase good j if the household's willingness to pay $1/\lambda$ does not fall short of the price $p(j)$.[7] The resulting demand curve, depicted in Figure 1, is a step function which coincides with the vertical axis for $p(j) > 1/\lambda$ and equals unity for prices $p(j) \leq 1/\lambda$.

Figure 1

By symmetry, the household's willingness to pay is the same for all goods and equal to the inverse of λ, which itself is determined by the household's income and product prices. Intuitively, the demand curve shifts up when the income of the consumer increases (λ falls) and shifts down when the price level of all other goods increases (λ rises).

It is interesting to note the difference between consumption choices under 0-1 preferences and under the standard CES-case. With 0-1 preferences the household chooses how many goods to buy but there is no choice about the quantity in which a good is consumed.[8] Under CES preferences, a household has a choice about the (positive) quantities of the supplied goods, but essentially has no choice about how many different goods to buy (due to a reservation price of infinity it is optimal to purchase each product in positive amounts, whatever its price). In other words, the stylized case of 0-1 preferences is interesting because this assumption shifts the focus of consumer choice to the *extensive* margin, thus deviating from the CES case with its focus on the *intensive* margin. In Section 6 we show that our results generalize to more general preferences, allowing for adjustments on both the extensive and the intensive margin.

[6] Notice that the integral in (1) runs from zero to infinity. While preferences are defined over an infinitely large measure of potential goods, the number of goods actually supplied is limited by firm entry, i.e. only a subset of potentially producible goods can be purchased at a finite price.

[7] Strictly speaking, the condition $1 \geq \lambda p(j)$ is necessary but not sufficient for $c(j) = 1$ and the condition $1 < \lambda p(j)$ is sufficient but not necessary for $c(j) = 0$. The reason is that purchasing all goods for which $1 = \lambda p(j)$ may not be feasible given the consumer's budget. For when N different goods are supplied at the same price p but $y < pN$ the consumer picks at random which particular good will be purchased or not purchased. This case, however, never emerges in the general equilibrium.

[8] The discussion here rules out the case where incomes could be larger than pN, meaning that the consumer is subject to rationing (i.e. he would want to purchase more goods than are actually available at the available prices). While this could be a problem in principle, it will never occur in the equilibrium of the model.

3 Autarky equilibrium and the emergence of trade

Consider an economy living in autarky under monopolistic competition. After incurring the set-up costs, the various producers have a natural monopoly for their products. Since all monopolists have the same cost and demand curves and since there is a representative consumer, we can omit indices. The monopolistic firm faces a demand curve as depicted in Figure 1. This firm will charge a price equal to the representative consumer's willingness to pay $p = 1/\lambda$ and sell output of quantity 1 to each of the $\mathcal{P}$ households.

Without loss of generality, we choose labor as the numéraire and set $W = 1$. Two conditions characterize the autarky equilibrium. The *first* is the zero-profit condition, ensuring that operating profits cover the entry costs but do not exceed them to deter further entry. Entry costs are $FW = F$ and operating profits are $[p - W/a]\,\mathcal{P} = [p - 1/a]\,\mathcal{P}$. The zero-profit condition can be written as $p = (aF + \mathcal{P})\,/a\mathcal{P}$.[9] This implies a mark-up μ - the ratio of price to marginal cost – given by

$$\mu = \frac{aF + \mathcal{P}}{\mathcal{P}}.$$

which is determined by technology parameters a and F and the market size parameter $\mathcal{P}$. The *second* equilibrium condition is a resource constraint ensuring that there is full employment $\mathcal{P}L = FN + \mathcal{P}N/a$. From this equation, equilibrium product diversity in the decentralized equilibrium can be calculated[10]

$$N = \frac{a\mathcal{P}}{aF + \mathcal{P}}L.$$

Market size and technology influence mark-ups in our framework. We will show below that the mark-up channel is a crucial channel by which non-homothetic preferences affect patterns of trade and the international division of labor.

Now assume there are two countries, rich and poor, and consumers in both countries have the same preferences given by (1). Assume further that firms in the two countries produce different products. Under which condition will the two countries trade?

[9] Notice that we have argued that $p = 1/\lambda$ and $p = (aF + P)\,/aP$; it therefore seems that p is overdetermined, unless we have $\lambda = aP/\,(aF + P)$. To see that this is in fact the case, notice that increasing income by one unit approximates an increase in L (because income is $y = WL$ and we normalized $W = 1$). Hence we can write $\lambda = dU/dL = (\partial U/\partial N) \cdot (\partial N/\partial L)$. Since we have $U = N$, $\partial U/\partial N = 1$, and we have $\partial N/\partial L = aP/\,(aF + P)$ from equilibrium product diversity, this confirms the claim.

[10] Notice the difference between the 0-1 outcome and the standard CES-case. With CES, the mark-up is determined by the elasticity of substitution between differentiated goods; it is independent of technology and market size. In fact, the variability of the mark-up with 0-1 preferences will drive many of our results below. Moreover, with CES, equilibrium product diversity is independent of productivity a and proportional to set-up costs F and inversely proportional to market size P. We notice that with 0-1 preferences product diversity in the decentralized equilibrium is equal to the socially optimal product diversity.

Assumption 1 (Trade condition) $\tau \leq \sqrt{aF/\mathcal{P}+1}$.

The above assumption states a sufficient condition for the emergence of international trade. To see this, consider an entrepreneur shipping τ^2 units of his or her product to the other country so that, due to iceberg trade costs, τ units arrive there. The firm can exchange the remaining τ units for τ units of a (symmetric) foreign variety ship it back and sell it on the home market at price p. Thus autarky cannot be an equilibrium if the costs of producing τ^2 units falls short of the (local) autarky prize, i.e. $\tau^2/a < (aF+\mathcal{P})/a\mathcal{P}$. Solving for τ yields the trade condition. Note that the trade condition is independent of the other country's parameters such as population size, labor endowment, or technology parameters. The above trade condition is a sufficient but not a necessary condition for the emergence of international trade. We will assume throughout the paper that Assumption 1 holds.[11]

4 Trade between equally large but unequally rich countries

Let us now assume that Assumption 1 holds and consider a world economy with two countries with unequal wealth. We denote variables of the rich country with superscript R and variables of the poor country with superscript P. To highlight the importance of differences in per capita incomes as a source of international trade, we start by assuming that the two countries differ only in per capita endowments, but have equally large populations, hence $L^R > L^P$ and $\mathcal{P}^R = \mathcal{P}^P = \mathcal{P}$. We also assume that the two countries have identical production and transport technologies.

4.1 A full trade equilibrium

When the two countries are not very unequal, a possible equilibrium is one in which all producers sell on the world market, so that all goods are traded internationally. In such a *full trade equilibrium*, the price for a differentiated product in country $i = R, P$ equals the respective households' aggregate willingness to pay (see Figure 1), hence we have $p^R = 1/\lambda^R$ and $p^P = 1/\lambda^P$. Since country R is wealthier than country P, we have $\lambda^R < \lambda^P$ and $p^R > p^P$. By symmetry, the prices of imported and home-produced goods are identical within each country.

Solving for the full trade equilibrium is straightforward. Consider the resource constraint in the rich country. When N^R firms enter, $N^R F$ labor units are employed to set up these firms and $N^R \mathcal{P}(1+\tau)/a$ labor units are employed in the production to serve the world market. Since each of the $\mathcal{P}$ households inelastically supplies L^R units of labor, the resource constraint is $\mathcal{P}L^R = N^R F + N^R \mathcal{P}(1+\tau)/a$. This is analogous for the poor country P. Solving for N^i

[11] The case when Assumption 1 does not hold (high trade costs), is available upon request by the authors.

lets us determine the number of active firms in the two countries

$$N^i = \frac{a\mathcal{P}}{aF + (1+\tau)\,\mathcal{P}} L^i,\ i = R, P.$$

Now consider the zero-profit conditions in the two countries. An internationally active firm in the rich country generates total revenues equal to $\mathcal{P}(p^R + p^P)$ and has total costs $W^R\,[F + (1+\tau)\mathcal{P}/a]$. An internationally active firm in the poor country generates the same total revenues and has to incur the same labor requirement $F + (1+\tau)\mathcal{P}/a$. Hence, wages per efficiency unit have to equalize, $W^R = W^P$, for the zero profit conditions to hold in both countries. We use labor as the numéraire in the following, $W^R = W^P = 1$. The budget restrictions are therefore $p^i\left(N^R + N^P\right) = L^i$. Combining the zero profit condition with the budget restrictions and the number of firms lets us express the price charged in country i as

$$p^i = \frac{L^i}{L^R + L^P}\frac{aF + (1+\tau)\,\mathcal{P}}{a\mathcal{P}},\ i = R, P. \tag{2}$$

The prices of all differentiated products are the same within a country, irrespective of whether they are produced at home or abroad. Consequently, imported goods generate a lower mark-up than locally produced goods as exporters have to bear the trade costs fully.[12] The mark-ups (price over marginal cost) producers charge on their home market $\mu_D^i = ap^i$ and the mark-ups set in the export market $\mu_X^i = ap^i/\tau$ are given by

$$\mu_D^i = \frac{L^i}{L^R + L^P}\frac{aF + (1+\tau)\,\mathcal{P}}{\mathcal{P}},\ \text{and}\ \mu_X^i = \frac{1}{\tau}\frac{L^i}{L^R + L^P}\frac{aF + (1+\tau)\,\mathcal{P}}{\mathcal{P}},\ i = R, P.$$

In sum, the full trade equilibrium has a simple structure: the ratios of rich relative to poor country varieties, prices, and mark-ups are identical to the ratio in relative labor endowments (and nominal incomes), i.e. $N^P/N^R = p^P/p^R = \mu_D^P/\mu_D^R = \mu_X^P/\mu_X^R = W^P L^P/\left(W^R L^R\right) = L^P/L^R < 1$. The differences in per capita endowments and incomes translate one-to-one into differences in prices, hence international trade establishes an equilibrium such that real incomes and welfare levels equalize between the two countries. Under autarky, on the other hand, the poor country is clearly worse off than the rich country. As a result, international trade benefits the poor country more than it does the rich country.

4.2 Partial trade and the threat of parallel imports

Full trade cannot be an equilibrium outcome when per capita labor endowments and hence incomes between the two countries are very unequal, i.e. when L^P/L^R becomes small. The

[12] In this respect, 0-1 preferences differ strongly from CES preferences, as higher costs cannot be passed through to prices. With CES preferences, transportation costs are more than passed through to prices as exporters charge a fixed mark-up on marginal costs (including transportation). Notice limited cost pass-through is consistent with empirical evidence. A number of empirical studies document that marginal cost shocks are not fully passed through to prices at the firm level and that prices are substantially less volatile than costs. See Ravn et al. (2007) and the references quotes there.

reason is that if countries are sufficiently unequal, a threat of parallel trade emerges. To see the point most clearly, consider a US firm that sells its good both in the US and in China. The firm charges a price in China that equals a Chinese household's willingness to pay $p^P = 1/\lambda^P$ and a price in the US that equals the US households willingness to pay $p^R = 1/\lambda^R$. Because the difference between $1/\lambda^P$ and $1/\lambda^R$ is large, arbitrage opportunities emerge. Arbitrage traders purchase the good cheaply on the Chinese market, ship it back to the US, and underbid local producers on the US market. This threat of parallel trade also concerns Chinese firms which both produce for the local market and export to the US. When these firm charge prices in the US that exploit US households' high willingness to pay, arbitrage traders have an incentive to purchase the product in China cheaply and parallel export it to the US.

Clearly, firms anticipate this threat of parallel trade and adjust their international pricing accordingly. These firms thus take advantage of the large world market but are constrained in their pricing due to the threat of parallel trade. There is an alternative, potentially profitable, strategy: a rich-country firm could abstain from trading its product internationally and focus exclusively on its rich home market. This producer type has a smaller market but can exploit the rich country households' high willingness to pay because it is not subject to the threat of parallel imports. In equilibrium, both types of firms exist simultaneously and the relative popularity of the two strategies adjusts such that both yield the same profits. (We will see below that all firms in the poor country are strictly better off selling their product on the world market rather than limiting their sales to exports to the rich country and not selling on the local market.) This implies that only a subset of all available products is actually traded, which is why we call this equilibrium "partial trade" equilibrium.

Denote the price in the rich country of a good that is traded internationally by p_T^R ; the price in the rich country of a good that is not traded by p_N^R; and (as above) the price of a good in the poor country by p^P. When setting their prices, suppliers of goods traded internationally anticipate the threat of parallel trade and set a price that just prevents any incentive for arbitrage. This implies that the prices charged in the rich country for goods traded internationally may not exceed the corresponding prices for these goods in the poor country plus trade costs, i.e. $p_T^R \leq \tau/\lambda^P$, profit maximization implies that this condition holds with equality. Hence we must have $p_N^R = 1/\lambda^R$, and $p^P = 1/\lambda^P$.

The zero profit condition for a traded good is $(p_T^R + p^P)\mathcal{P} = [F + \mathcal{P}(1+\tau)/a]\, W^R$ for an internationally active rich-country producer and $(p_T^R + p^P)\mathcal{P} = [F + \mathcal{P}(1+\tau)/a]\, W^P$ for a poor-country producer. Both types of firms generate the same total revenues and have to incur the same labor input. As a result, the zero-profit condition requires the compensation per efficiency unit of labor to be the same in the two countries, $W^R = W^P = 1$. The prices of traded goods can be calculated straightforwardly from these zero profit conditions as

$$p_T^R = \frac{\tau}{1+\tau}\frac{aF + (1+\tau)\mathcal{P}}{a\mathcal{P}} \quad \text{and} \quad p^P = \frac{1}{1+\tau}\frac{aF + (1+\tau)\mathcal{P}}{a\mathcal{P}}.$$

The zero profit condition for a rich-country producer that sells his product exclusively on the home market is $p_N^R\mathcal{P} = F + \mathcal{P}/a$, from which we calculate the equilibrium price of a non-traded variety

$$p_N^R = \frac{aF + \mathcal{P}}{a\mathcal{P}}.$$

In a partial trade equilibrium, domestic and internationally active firms co-exist in equilibrium. To see why this is an equilibrium, consider the alternative situation in which all goods produced in the rich country are traded internationally. If all products were sold at a price that prevented parallel trade, all goods would be priced below the rich-country households' willingness to pay. However, this corresponds to a situation where the representative rich-country household is not able to spend all income. This, in turn, implies that country-R households have an infinitely large willingness to pay for additional products, which induces some country-R firms to switch strategy and sell only on their home markets.

In contrast to the rich country, do all producers in the poor country sell their product both at home and abroad? In principle one might think that country-P producers also have an incentive to sell their product exclusively in the rich country exploiting the country-R households' high willingness to pay (and not to sell their product on the home market to prevent parallel exports). While such a strategy generates the same total sales, it generates high overall costs as the country-P exporter also has to bear trade costs. Hence selling exclusively on the R market is not a profitable option for a P producer.

We are now ready to solve for the partial trade equilibrium. The resource constraint in the poor country is still given by $\mathcal{P}L^P = N^P\left(F + (1+\tau)\mathcal{P}/a\right)$ from which we calculate

$$N^P = \frac{a\mathcal{P}}{aF + (1+\tau)\mathcal{P}}L^P. \tag{3}$$

The resource constraint in country R is different from before because now we have to distinguish products that are exclusively sold domestically and those that are traded internationally. Denoting the traded and non-traded goods produced in the rich country by N_T^R and N_N^R, respectively, the resource constraint of country R is $\mathcal{P}L^R = N_T^R\left(F + (1+\tau)\mathcal{P}/a\right) + N_N^R\left(F + \mathcal{P}/a\right)$. Together with the trade balance condition $N_T^R p^P \mathcal{P} = N^P p_T^R \mathcal{P}$ and the terms of trade $p_T^R/p^P = \tau$ we can calculate

$$N_T^R = \frac{a\mathcal{P}}{aF + (1+\tau)\mathcal{P}}\tau L^P, \quad \text{and } N_N^R = \frac{a\mathcal{P}}{aF + \mathcal{P}}\left(L^R - \tau L^P\right). \tag{4}$$

4.3 Per capita incomes and patterns of trade

It is straightforward to see the condition under which the threat of parallel trade becomes a binding constraint on price setting in the rich country, allowing a partial trade equilibrium to emerge. In a full trade equilibrium, relative prices are $p^P/p^R = L^P/L^R$ and the threat of parallel trade is not binding as long as the price ratio satisfies $p^R/p^P \leq \tau$. It follows that the parallel trade constraint kicks in when

$$\tau = \frac{L^R}{L^P}. \tag{5}$$

In other words, a full trade equilibrium emerges when per capita incomes are sufficiently similar, $L^R/L^P \leq \tau$, and a partial trade equilibrium emerges when the gap in per capita incomes is high, $\tau < L^R/L^P$.

Figure 2 draws condition (5) in the $(L^P/L^R, \tau)$ space. Figure 2 is drawn for values of τ that satisfy the trade condition of Assumption 1. In region **F** (full trade), characterized by high values of L^P/L^R and intermediate values of τ, there is full trade.

In that region, consumers in the two countries have very similar incomes (and hence the differences in their willingness to pay are minor) so that the parallel trade constraint on prices in the rich market does not become binding and arbitrage opportunities do not emerge. In region **P** (partial trade), characterized by low trade costs and high differences in average incomes, a partial trade equilibrium emerges. When relative endowments L^P/L^R are low, the difference in willingness to pay between rich- and poor-country households is large, making the parallel-trade constraint binding.[13]

Figure 2

Let us highlight how the volume and structure of international trade depend on relative per capita incomes L^P/L^R. We define "trade intensity" ϕ as the fraction of traded goods, $N^P + N^R_T$, over the total goods produced worldwide, $N^P + N^R_T + N^R_N$. Using equations (3) and (4) calculating trade intensity in a partial trade equilibrium is straightforward.

$$\phi = \frac{(1+\tau)(1+aF/\mathcal{P})}{1 + aF/\mathcal{P} - \tau^2 + (1+\tau+aF/\mathcal{P})\, L^R/L^P} \text{ if } L^P/L^R < 1/\tau, \tag{6}$$

Alternatively, the world economy is in a full trade equilibrium with $\phi = 1$ if $L^P/L^R \geq 1/\tau$. Equation (6) reveals that a higher L^P/L^R, i.e. higher similarity between the two countries, is associated with a higher trade intensity ϕ. In Figure 3 we draw ϕ (vertical axis) against relative labor endowments L^P/L^R (horizontal axis) holding worldwide resources $\mathcal{P}\left(L^R + L^P\right)$ constant. (A decrease in L^P/L^R is then a mean-preserving spread in world endowments.)

Figure 3

[13] Notice that there is international trade even when income differences become extremely large and L^P/L^R becomes very small. The range of traded goods approaches zero, however, when L^P/L^R goes to zero.

A reduction in L^P/L^R leads to a lower intensity of trade: a decreasing range of traded goods $N^P + N_T^R$ is associated with an increasing range of non-traded goods N_N^R. In other words as the similarity of the two countries in per capita endowments (and per capita incomes) increases, the intensity of trade ϕ increases as well. The world economy reaches full trade when $L^P/L^R \geq 1/\tau$. We summarize this discussion in

Proposition 1 *a) When relative per capita endowments are sufficiently similar so that* $L^P/L^R \in [1/\tau, 1]$*, the general equilibrium features full trade. b) When per capita endowments become sufficiently dissimilar so that* $L^P/L^R \in (0, 1/\tau)$*, the general equilibrium is characterized by partial trade where a threat of parallel imports/exports constrains the prices charged for internationally traded goods in the rich country.*

Proof. In text. ■

It is worth noting that this simple model features the famous Linder hypothesis. Linder (1961) emphasized that the similarity of two countries, as measured by similarity in their per capita incomes, should be an important determinant of trade between them.

4.4 Welfare and the gains from trade

We proceed by studying welfare implications and the gains from trade. In particular, we are interested in how trade liberalizations (a reduction of τ) affect welfare and the distribution of trade gains between rich and poor countries. In a *full trade* equilibrium, households in both countries purchase all goods produced worldwide. Hence the welfare levels are identical in both countries despite their unequal endowment with productive resources

$$U^{R,f} = U^{P,f} = a\frac{\mathcal{P}\left(L^R + L^P\right)}{aF + (1+\tau)\mathcal{P}}.$$

Firms' price setting behavior drives this result. R-consumers are willing to pay higher prices than P-consumers because their *nominal* income is higher. In the full trade equilibrium, higher nominal incomes translate one to one into higher prices. *Real* incomes and welfare are therefore identical. To see the mechanism by which welfare is equalized even though the two countries have unequal welfare levels under autarky, consider firms' mark-ups. By assumption, all firms have identical production costs hence different prices reflect differences in mark-ups across countries. Since in equilibrium profits are zero, the markups are fully used to cover fixed costs and iceberg losses during transportation. Hence, the higher mark-ups in the rich country imply that the rich country households bear a larger share of these costs.

In a *partial trade* equilibrium, welfare levels of consumers in the two countries diverge. Country-P households purchase $N_T^R + N^P$ goods and country-R households purchase $N^P +$

$N_T^R + N_N^R$ goods. Using (3) and (4) we can calculate the welfare levels

$$\begin{aligned} U^{P,p} &= a\frac{\mathcal{P}(1+\tau)L^P}{aF+(1+\tau)\mathcal{P}} \quad \text{and} \\ U^{R,p} &= a\frac{\mathcal{P}(1+\tau)L^P}{aF+(1+\tau)\mathcal{P}} + a\frac{\mathcal{P}\left(L^R-\tau L^P\right)}{aF+\mathcal{P}}. \end{aligned}$$

Notice that while welfare in country R decreases in τ (lower trade costs or trade liberalization increases welfare), the opposite is true for country-P welfare. We are now able to state the following proposition.

Proposition 2 *a) Compared with autarky, country P gains more from trade than country R. b) R-consumers favor free trade, i.e. $\tau = 1$, whereas P-consumers derive their highest utility when there are trade barriers such that $\tau = \min\left\{\sqrt{aF/\mathcal{P}+1}, L^R/L^P\right\}$.*

Proof. The proposition can be readily demonstrated using Figure 4. Panel a) is drawn for the case when $L^R/L^P \leq \sqrt{aF/\mathcal{P}+1}$ so that a full trade equilibrium emerges with moderate trade costs. Panel b) is drawn for the case when $L^R/L^P > \sqrt{aF/\mathcal{P}+1}$ so that a full trade equilibrium is not feasible. Country-R welfare (the bold graph) is monotonically decreasing in τ in both panels of Figure 4. Hence the R-consumer reaches his maximum welfare when trade costs have reached their lowest possible level, at $\tau = 1$. However, the welfare of country P (the dotted graph) increases in τ in both panels of Figure 4 when trade costs are sufficiently low, i.e. in a situation where the world economy is in a partial trade equilibrium. A full trade regime emerges in panel a) when $\tau \in \left[L^R/L^P, (aF/\mathcal{P}+1)\, L^P/L^R\right]$ where welfare decreases in τ. The economies remain autarkic for even higher $\tau > (aF/\mathcal{P}+1)\, L^P/L^R$ where welfare obviously becomes independent of τ. Figure 4 also shows that the highest welfare for country-P consumers occurs at $\tau = L^R/L^P$ when $L^R/L^P \leq \sqrt{aF/\mathcal{P}+1}$ and at $\tau = \sqrt{aF/\mathcal{P}+1}$ when $L^R/L^P > \sqrt{aF/\mathcal{P}+1}$. Taken together, this yields the result in Proposition 2. ■

Figures 4a, 4b

Proposition 2 shows the crucial role of trade costs for welfare. Unequal countries have different preferred trade barriers (or different preferred degrees of trade liberalizations). Consumers in the rich country are essentially free-traders whereas consumers in the poor country only want liberalization up to a positive level of trade costs. What is the intuition behind this result? When the world economy has reached a partial trade equilibrium the threat of parallel imports constrains prices in the rich country to $p_T^R = \tau p^P$. Further trade liberalization forces country-P exporters to lower prices in country-R relative to prices in country P because price discrimination is limited by factor τ. Hence the terms of trade for the poor country deteriorate leading to the welfare loss.

5 Population sizes versus per capita endowments

In the previous section, differences in incomes across countries were due to differences in per capita endowments and the two countries had equally large populations. Let us now consider the case when countries differ along both dimensions. This is interesting because it allows us to gain insights on how the composition of aggregate income affects the extensive margin of international trade under non-homothetic preferences (as standard new trade theory also predicts that the two countries' sizes affect trade volume).

With 0-1 preferences it becomes most transparent how the composition of aggregate income affects the size of the home market if one considers a given product. As every household consumes exactly one unit of a given variety, a larger endowment of the representative home-consumer leaves the size of the home market unchanged. For the same reason, a larger population increases the home market one to one. For internationally active producers, having a relatively larger home market means that trade costs are a relatively smaller part of total costs. As firms located in a large country bear relatively fewer iceberg losses as a fraction of their total costs, labor in a large country is more productive than labor in a small country.

5.1 Relative wages and general equilibrium

To see how different population sizes affect relative wages, we need to check the zero-profit conditions of internationally active firms. Total revenues are given by $p^P\left(\tau\mathcal{P}^R+\mathcal{P}^{\mathcal{P}}\right)$ and do not differ by firm location. However, the amount of labor needed to serve the world market does differ. It is given by $F+\left(\mathcal{P}^R+\tau\mathcal{P}^{\mathcal{P}}\right)/a$ for country-R firms and by $F+\left(\tau\mathcal{P}^R+\mathcal{P}^{\mathcal{P}}\right)/a$ for country-P firms. We can calculate relative wages from the zero-profit conditions (note that the formula is the same under both full and partial trade)

$$\omega\equiv\frac{W^P}{W^R}=\frac{aF+\tau\mathcal{P}^P+\mathcal{P}^R}{aF+\mathcal{P}^P+\tau\mathcal{P}^R}. \tag{7}$$

It follows that $\omega \gtreqless 1$ if $\mathcal{P}^P/\mathcal{P}^R \gtreqless 1$. Hence the compensation per efficiency unit of labor is higher in the poor country when the poor country is larger and vice versa.

The discussion above suggests that a backward country in terms of per capita *endowment* can get ahead in terms of per capita *income* if it has a large population. This raises an interesting question. Could it be that a huge population raises incomes and willingness to pay in country P so much that the parallel-trade constraint on price setting becomes binding in country P rather than in country R? In other words, is it possible that producers in a poor country with a large population face a threat of parallel trade leading to a "reversed" partial trade equilibrium in which only a subset of poor-country varieties are internationally traded?

The answer is no. To see this, recall that the households' budget constraints in a full trade equilibrium are $W^i L^i = p^i(N^P + N^R)$, $i = R, P$ from which we can calculate relative prices under full trade

$$\frac{p^P}{p^R} = \frac{aF + \mathcal{P}^R + \tau\mathcal{P}^\mathcal{P}}{aF + \tau\mathcal{P}^R + \mathcal{P}^\mathcal{P}} \frac{L^P}{L^R}.$$

Verification that $\lim_{\mathcal{P}^P \to \infty} p^P/p^R = \tau L^P/L^R < \tau$ is straightforward. Even if the population in country P becomes extremely large, country-P households' willingness to pay – while eventually exceeding that of country-R households – will remain below τp^R. Hence arbitrage opportunities and therefore a threat of parallel trade do not exist. In sum, a "reversed" partial trade equilibrium will never emerge.

Figure 5 demonstrates that per capita incomes affect trade volumes for a given aggregate size of the economy. The bold line represents the combinations of relative per capita endowments and relative population sizes such that the world economy just enters the full trade regime (FP-boundary). More precisely, along this line the relative willingness to pay p^R/p^P are exactly equal to trade costs τ. Per capita incomes are more similar to the right of this curve (L^P/L^R closer to unity) so that the relative willingness to pay is strictly lower than τ. Per capita income differences are too dissimilar to the left of this curve, creating a threat of parallel trade so that a partial equilibrium emerges.[14] The figure shows that, whatever relative population sizes $\mathcal{P}^P/\mathcal{P}^R$, the world economy can reach a full trade equilibrium provided that relative per capita endowments L^P/L^R sufficiently approach unity. However, we cannot argue in the same way with increases in relative populations. When per capita incomes are sufficiently similar, an increase in population in the poorer country may push the world economy out of a partial trade into a full trade equilibrium. However, when per capita endowments are very dissimilar, $L^P/L^R < 1/\tau^2$, the world economy remains trapped in a partial trade equilibrium even when relative population size $\mathcal{P}^P/\mathcal{P}^R$ goes to infinity. In this sense, the model predicts that per capita incomes are more important than population sizes in shaping patterns of international trade.

To consider the distinct impact of per capita income, we draw a dotted 'iso-size' line, i.e. the combination of relative per capita endowments and relative population sizes for which aggregate endowments of the two economies are identical, $L^P\mathcal{P}^P = L^R\mathcal{P}^R$. (Recall that, under CES preferences, such a situation would feature a world equilibrium with perfect symmetry). Since the iso-size line is flatter than the FP boundary, the two curves cross when L^P/L^R becomes sufficiently low. Hence two countries with identical aggregate endowments end up in partial trade when one country is rich but small and the other country is large but poor.

Figure 5

[14] The FP boundary is defined by $\omega(\mathcal{P}^P/\mathcal{P}^R)L^P/L^R = \tau$. From equation (7) we have $\omega' > 0$, which implies a negative relationship between $\mathcal{P}^P/\mathcal{P}^R$ and L^P/L^R.

5.2 Welfare implications

When we allow populations to differ between the two countries, welfare implications remain qualitatively unchanged. In a *full trade* equilibrium, the welfare levels are

$$U^{i,f} = U^{i,f} = a\frac{\mathcal{P}^R L^R + \omega(\tau)\mathcal{P}^P L^P}{aF + \mathcal{P}^R + \tau\mathcal{P}^P}, \quad i = R, P$$

the same for both countries. Comparing full trade with autarky, it may be that the rich (rather than the poor) country gains more than the poor country. Welfare levels under autarky are $U^{i,a} = L^i / \left(F/\mathcal{P}^i + 1/a\right)$, $i = R, P$, which reveals that country-R households gain more from full trade when $L^P\mathcal{P}^P/(L^R\mathcal{P}^R) > (aF + \mathcal{P}^P)/(aF + \mathcal{P}^R)$. This situation arises when the rich country is very small so that access to the large world market generates a large gain in efficiency.

In a *partial trade* equilibrium we have

$$\begin{aligned} U^{P,p} &= a\frac{\left(\mathcal{P}^P + \tau\mathcal{P}^R\right) L^P}{aF + \tau\mathcal{P}^R + \mathcal{P}^P}, \text{ and} \\ U^{R,p} &= a\frac{\left(\mathcal{P}^P + \tau\mathcal{P}^R\right) L^P}{aF + \tau\mathcal{P}^R + \mathcal{P}^P} + a\frac{\mathcal{P}^R\left(L^R - \omega(\tau)\tau L^P\right)}{aF + a\mathcal{P}^R}. \end{aligned}$$

It can be shown that $\partial U^{P,p}/\partial\tau > 0$ and $\partial U^{R,p}/\partial\tau < 0$. Hence, allowing for unequal population sizes does not change the welfare implications of a trade liberalization. When the world economy is in partial trade equilibrium, a trade liberalization improves welfare of country-R consumers but hurts consumers in country P. The reason is the same as above. In a partial trade equilibrium, a trade liberalization deteriorates terms of trade for the poor country. The higher relative price of imported goods implies that the consumption basket poor country consumers can afford becomes smaller. The preferred level of openness is $\tau = 1$ in the rich country and $\tilde{\tau} > 1$ in the poor country, where $\tilde{\tau}$ satisfies $\omega(\tilde{\tau})L^P/L^R = \tilde{\tau}$. Notice that $\tilde{\tau}$ is also the critical level of trade costs that lets the world economy switch from a full trade to a partial trade equilibrium.

6 Extensions

6.1 More than two countries

The above analysis examined the case of bilateral trade. In a two-country context, we demonstrated that per capita incomes have a crucial impact on trade patterns and that the impact of the aggregate size of an economy depends on whether size comes from per capita income or from population. We also showed that trade liberalizations (a reduction in trade costs) increase trade in a full trade equilibrium; but may *decrease* the volume of international trade in a partial trade equilibrium. One could argue that the latter result is an unattractive feature of our model.

The above analysis has studied the case of bilateral trade. In a two-country context, we have shown that per capita incomes have a crucial impact on trade patterns and the impact of the aggregate size of an economy depends on whether size comes from per capita income or from population. We have also shown that trade liberalizations (a reduction in trade costs) increase trade in a full trade equilibrium; but may *decrease* the volume of international trade in a partial trade equilibrium. One could argue that the latter result is an unattractive feature of our model.

Two rich and two poor countries We show below that this result becomes much weaker or completely vanishes once we allow for more than two countries. To make the point, we consider the following interesting special case. Suppose there are four countries, identical in all dimensions except for their per capita endowments. Assume further that there are two rich countries, each of which has per capita endowment L^N, the "North"; and two poor countries with per capita endowment $L^S < L^N$, the "South". If $L^N/L^S < \tau$ there will be full trade, each country buys all domestically produced goods and imports all goods produced in all other countries. If $L^N/L^S > \tau$, however, there is partial trade. Partial trade now means that not all countries consume all goods, as the two poor (but not the two rich) countries consume only a subset of the global menu of goods. Denote by p^i and W^i the willingness to pay and the wage in country $i \in \{N, S\}$. Consider the partial trade equilibrium. The zero profit conditions become

$$2p^S\mathcal{P} + 2\tau p^S\mathcal{P} = \left(F + \frac{1+3\tau}{a}\mathcal{P}\right) W^i, \quad i = N, S$$

for goods that are traded between all four countries and

$$2p^N\mathcal{P} = \left(F + \frac{1+\tau}{a}\mathcal{P}\right) W^i$$

for goods that are traded only among the two rich countries. Exactly as in the two-country case, equal population sizes ensure factor price equalization, so we have $W^N = W^S = 1$. The resource constraints are

$$\begin{aligned} L^S\mathcal{P} &= N^S\left(F + \frac{1+3\tau}{a}\mathcal{P}\right) \text{ in type } S \text{ countries and} \\ L^N\mathcal{P} &= N_N^N\left(F + \frac{1+\tau}{a}\mathcal{P}\right) + N_T^N\left(F + \frac{1+3\tau}{a}\mathcal{P}\right) \text{ in type } N \text{ countries} \end{aligned}$$

where N^S is the range of goods produced in each S country; N_T^N is the range of goods supplied by firms in one of the N countries and traded worldwide; and N_N^N is the range of goods produced and traded only in the North. Since each Northern country imports all goods produced worldwide but each Southern country imports only a subset of these goods, the aggregate (regional) trade balance between North and South has to be balanced in equilibrium.[15] The value of aggregate Northern imports from the South are $2\tau p^S N^S$ and the value of aggregate Northern exports to the South are $2p^S N_T^N$ hence trade balance requires

$$\tau N^S = N_T^N.$$

Calculating the number of goods that are produced worldwide from the equations above is straightforward; it is equal to the level of welfare in the Northern country

$$U^N(\tau) = 2N^S + 2N_T^N + 2N_N^N = 2\frac{(1+\tau)\,L^S\mathcal{P}}{F + \frac{1+3\tau}{a}\mathcal{P}} + 2\frac{\left(L^N - \tau L^S\right)\mathcal{P}}{F + \frac{1+\tau}{a}\mathcal{P}}$$

From this equation, it can be shown that $\partial U^N(\tau)/\partial\tau < 0$, i.e. the welfare level of N-consumers increases as a result of a trade liberalization. Notice that, similar to the two-country case, this result arises because goods imported from the South become cheaper, creating demand for new goods. In the new equilibrium, more goods are produced and traded in the North. The deterioration (from the perspective of S countries) in the terms of trade leads to a situation where fewer goods are traded between the Northern and Southern world regions. This effect is similar to the two-country case. The increased threat of parallel trade induces Northern producers to withdraw their products from the world market and to sell their product exclusively in the rich Northern region.

The situation is somewhat different in the Southern countries. In particular, it may be that a trade liberalization also increases welfare in the South. We saw that the poorer country is strictly worse off as a result of lower trade costs in the two-country case. This need no longer be the case in the multi-country case. While North-South trade will unambiguously decrease due to a trade liberalization, lower trade costs will increase South-South trade. As a result, the impact of trade liberalizations on welfare in the South is unclear. To see this, calculate the welfare of the Southern consumer as

$$U^S(\tau) = 2N^S + 2N_T^N = 2\frac{(1+\tau)\,L^S\mathcal{P}}{F + \frac{1+3\tau}{a}\mathcal{P}}$$

[15] Due to the symmetry of our set-up, the volume of bilateral trade is undetermined. One of the Northern countries could produce predominantly (or exclusively) goods that are consumed only in the North while the other Northern country produces mainly (or exclusively) goods that are consumed worldwide. In that case, the first Northern country runs a trade surplus with the other Northern country and a trade deficit with both Southern countries taken together. Such trade imbalances cannot occur between the Southern countries, since each Southern country consumes all goods produced by the other Southern country, meaning that the South-South trade flows are of the same magnitude in either direction. However, each Southern country may run a surplus with one of the Northern countries that is balanced by a deficit with the other Northern country. Notice further that all bilateral trade flows are equalized in a full trade equilibrium. This is because all households in each country consume all goods that are produced worldwide.

from which $\partial U^S(\tau)/\partial\tau \leq 0$ follows when $aF/\mathcal{P} - 2 < 0$. The South is more likely to gain from lower trade costs if scale effects (lower $aF/\mathcal{P}$) grow in importance. A uniform global trade liberalization decreases N-S trade in this situation, but increased N-N and S-S trade dominate this effect in such a way that global trade increases, proving our initial claim.

Three unequal countries Now consider the alternative interesting case with three equally large but unequally rich countries, R, M, and P. The differences in endowments between any bilateral combination of countries is sufficiently large such that $\min\{L_R/L_M, L_M/L_P\} > \tau$. This latter assumption implies that the threat of parallel trade exists in all bilateral trade flows. The equilibrium then takes the following structure. Households in country R consume all goods produced worldwide; households in country M consume only a subset of these goods; and households in country P an even smaller subset. There are three groups of firms: firms that sell worldwide, firms that sell in M and R, and firms that only sell in R. The latter producers set a price p^R that equals the willingness to pay of country R households. The second group of producers sets a price p^M in country M and τp^M in country R. The first group sets a price p^P in country P, and a price τp^P in countries M and R. The zero profit conditions are given by [16]

$$
\begin{aligned}
\left(p^P + 2\tau p^P\right)\mathcal{P} &= W^i\left(F + \frac{1+2\tau}{a}\mathcal{P}\right) \quad i = P, M, R \text{ when selling worldwide} \\
\left(p^M + \tau p^M\right)\mathcal{P} &= W^i\left(F + \frac{1+\tau}{a}\mathcal{P}\right) \quad i = M, R \text{ when selling in } M \text{ and } R \\
p^R\mathcal{P} &= W^R\left(F + \frac{1}{a}\mathcal{P}\right) \text{ when selling in } R \text{ only.}
\end{aligned}
\tag{8}
$$

Prices p^P, p^M and p^R can be directly calculated from the zero profit conditions. To ensure that we have $\tau p^P < p^M$ the modified trade condition

$$\frac{(1+2\tau)\left(\tau^2 - 1\right)}{\tau^2 - \tau - 1} < \frac{aF}{\mathcal{P}}$$

has to be satisfied. This condition is stronger than Assumption 1, hence $\tau p^M < p^R$ follows. This condition ensures that all possible bilateral trade flows are strictly positive in equilibrium.

From equation (8) it is clear that equal population guarantees equal wages (per efficiency unit of labor); this also holds in the three-country case. To see this, notice that country P exports to both country M and country R. To ensure balanced trade, either country M or

[16] When specifying the zero-profit conditions, we have already implicitly assumed that goods are produced in a country that also consumes this good. In particular, goods that are consumed exclusively in country R are also produced in country R, and goods that are not consumed in country P are not produced in country P. This will be the case in equilibrium because equal population sizes lead to factor price equalization across the three countries $W^P = W^M = W^R$ and because larger profits margins (due to the absence of transport costs) let firms first serve the home market before selling abroad.

country R (or both) export to country P. When both countries export to country P, the first zero-profit condition has to hold in all three countries, from which factor prize equalization follows immediately. The situation is similar when only country M but not country R exports to country P. Then the first zero-profit condition ensures factor price equalization between countries M and P. In that case the country R runs a trade deficit with country P which has to be offset by a surplus with country M. Ruling out the knife-edge case where country M exports only to country P but does export to country R, we conclude that there must be some goods country M exports to country R which establishes factor price equalization between M and R. Taken together, we have $W^P = W^M = W^R = 1$.

Solving the equilibrium is somewhat tedious but straightforward. Using resource constraints and the trade balance conditions between each country and the rest of the world yields the following utility levels in the three countries (see Appendix 1)

$$
\begin{aligned}
U^P(\tau) &= \frac{1+2\tau}{F+\frac{1+2\tau}{a}\mathcal{P}} L^P \mathcal{P} \\
U^M(\tau) &= \frac{1+2\tau}{F+\frac{1+2\tau}{a}\mathcal{P}} L^P \mathcal{P} + \frac{(1+\tau)\left(L^M/L^P-\tau\right)}{F+\frac{1+\tau}{a}\mathcal{P}} L^P \mathcal{P} \\
U^R(\tau) &= \frac{1+2\tau}{F+\frac{1+2\tau}{a}\mathcal{P}} L^P \mathcal{P} + \frac{(1+\tau)\left(L^M/L^P-\tau\right)}{F+\frac{1+\tau}{a}\mathcal{P}} L^P \mathcal{P} + \frac{L^R/L^P-\tau L^M/L^P+\tau-\tau^2}{F+\frac{1}{a}\mathcal{P}} L^P \mathcal{P}
\end{aligned}
$$

These results allow us to determine the level of welfare in the three countries.

Proposition 3 *In a three country model with sufficient endowment differences across countries, $L_R > \tau L_M > \tau^2 L_P$, and sufficiently low trade costs so that the modified trade condition $(1+2\tau)\left(\tau^2-1\right)/\left(\tau^2-\tau-1\right) < aF/\mathcal{P}$ holds, the poor (rich) country loses (gains) from a trade liberalization (lower τ). The middle income country gains from a trade liberalization if L^M/L^P is not too large.*

Proof. see Appendix. ■

If L^P is sufficiently above zero, the middle income country gains from lower trade costs. As in the two-country case, changes in the terms of trade drive the results. The terms of trade improve for the rich country, both for trade with country P and country M. This improves the welfare of country R consumers. The terms of trade deteriorate for the poor country, both for trade with country M and country R. The situation is ambiguous for the middle income country. Here terms of trade improve against the poor country but deteriorate against the rich country. Hence it is not a prior clear whether country M will gain or lose. If the majority of goods is imported from the poor country, overall terms of trade will improve. In contrast, if the majority of goods is imported from rich country, overall terms of trade will deteriorate. The terms of trade effect will be negative if country P is very poor (and hence L^M/L^P large). In that case, the improvement in terms of trade with the poor country is negligible because the range of goods that can be imported from country P is small, and most goods will be imported from country R, with which the terms of trade deteriorate.

6.2 Heterogeneous trade costs

Another reason why the effect of trade liberalizations on volumes may be less clear are heterogeneous trade costs. For clarity, we go back to the case of two countries that are symmetric in all dimensions except for per capita endowments. Assume there are two product types, type 0 has low trade costs τ_0 and type 1 has high trade costs $\tau_1 > \tau_0$, but $\tau_1 \leq \sqrt{aF/\mathcal{P}+1}$ still holds. To keep things simple, we assume that a firm does not learn of the type of its product until the fixed setup investment has been made.[17] More precisely, a firm comes up with a product of type 0 with probability π, and type 1 with probability $1-\pi$. Assume further that firms can insure themselves perfectly against high-cost realizations, meaning that all firms make zero profits in equilibrium.

The interesting case is when the threat of parallel trade restricts firms of type 0 (low trade cost) in their price-setting, while this does not apply to firms of type 1 (high trade cost). In that case, there will be partial trade in low-cost varieties but full trade in high-cost varieties, and we have $p^R_{0T}/p^P_0 = \tau_0$ and $p^R_1/p^P_1 < \tau_1$. Notice that, since high-cost and low-cost varieties yield the same utility, the prices of these goods do not differ by type, $p^P_1 = p^P_0 = p^P$ and $p^R_1 = p^R_{0N} = p^R$. Denote by $E\tau \equiv \pi\tau_0 + (1-\pi)\,\tau_1$ the expected trade cost, and by $Ep^R \equiv \pi\tau_0 p^P + (1-\pi)\,p^R$ the expected price that an internationally active producer can charge in the rich country. The zero-profit condition of internationally active firms and of exclusive country-R producers now becomes

$$
\begin{aligned}
(Ep^R + p^P)\mathcal{P} &= W^i\left[F + \frac{\mathcal{P}\,(1+E\tau)}{a}\right] \text{ and} && (9)\\
p^R\mathcal{P} &= W^R\left[F + \frac{\mathcal{P}}{a}\right]. && (10)
\end{aligned}
$$

When producers of type 1 do not face a pricing constraint, while producers of type 0 do, the equilibrium has the following structure. All type 1 goods are traded internationally, while type 0 goods are only partially traded. This means country R produces πN^R goods of type 0 and $(1-\pi)\,N^R$ goods of type 1. The situation is analogous for country P.

The resource constraints for the two countries are now

$$
\begin{aligned}
L^R\mathcal{P} &= N^R_T\left(F + \frac{\mathcal{P}\,(1+E\tau)}{a}\right) + N^R_N\left(F + \frac{\mathcal{P}}{a}\right) && (11)\\
L^P\mathcal{P} &= N^P\left(F + \frac{\mathcal{P}\,(1+E\tau)}{a}\right). && (12)
\end{aligned}
$$

[17] This formulation avoids a situation where firms with low transport costs make positive profits in equilibrium. Allowing for profits would complicate the analysis but would not yield any substantial additional insights.

Finally, the balance of payments condition is given by

$$N^P E p^R = N_T^R p^P. \tag{13}$$

Equations (9) - (13) constitute a system of 6 equations in 7 unknowns: p^R, p^P, W^R, W^P, N^P, N_T^R, and N_N^R. We get the seventh equation by the choice of the numéraire $W^R = 1$.

We can solve for the general equilibrium in this regime and calculate welfare levels. An examination of the welfare level of country P – which is simply the sum of the worldwide traded varieties – suffices for checking the impact of a trade liberalization. This yields

$$U^P = N^P + N_T^R = a \frac{(1 + \tau_1 (1-\pi)) \mathcal{P}}{aF(1-\pi) + ((1-\pi) + \pi\tau_0 + (1-\pi)\tau_1) \mathcal{P}} L^P.$$

It is straightforward to check that $\partial U^P / \partial \tau_0 < 0$ whereas $\partial U^P / \partial \tau_1 > 0$. Hence whether a trade liberalization has a positive or a negative impact on the number of internationally traded varieties is ambiguous and depends on the relative importance of the types of goods constrained in price setting. When the proportion of high-trade cost products $(1-\pi)$ is sufficiently large, a general trade liberalization (a simultaneous reduction in τ_0 and τ_1) will increase the extensive margin of international trade.

6.3 National versus international exhaustion rules

Up to now we were working under the assumption of unrestricted parallel trade. The implicit assumption in the equilibria presented above was that there is "international exhaustion". This means that the intellectual property owner (i.e. a patent-, copyright-, and/or trademark-holder) loses its control of commercial exploitation when the product is sold on the national or international market. Hence, international arbitrageurs force firms to restrict their international price schedules to deter parallel trade. In many countries, however, parallel trade is restricted by law. For instance, the US applies "national exhaustion", meaning that a producer's patent or copyright expires when it is sold on the home market but not when sold on the international market. Similarly, the EU applies "regional exhaustion" which allows parallel trades only within the EU area. Parallel imports are restricted under national or regional exhaustion.

By introducing a new policy parameter we now investigate the role of exhaustion rules. Assume that ex ante there is an exogenous probability π that parallel trade is legally restricted for a particular good. Think of π as representing the share of industries for which "national exhaustion" applies (alternatively we can think of a world with "national exhaustion" rules, but the enforcement of these rules is uncertain. π then represents the probability that the rules are actually enforced). Firms learn only after paying the fixed costs F whether their product is subject to parallel trade. To keep things simple, we go back to the case where both countries have the same population sizes $\mathcal{P}^R = \mathcal{P}^P = \mathcal{P}$, so that we have wage equalization $W^R = W^P = 1$. The zero profit conditions of internationally active and domestic firms, respectively, become

$$Ep\mathcal{P} = F + \frac{(1+\tau)\mathcal{P}}{a} \text{ and } p^R\mathcal{P} = F + \frac{\mathcal{P}}{a}$$

where an internationally active firm's expected sales are given by $Ep\mathcal{P} \equiv \pi\left(p^P + p^R\right)\mathcal{P} + (1-\pi)\left(p^P + \tau p^P\right)\mathcal{P}$. (Notice that we implicitly assume that firms can perfectly insure themselves against low-price realizations in the case when no parallel trade restrictions apply. This assumption keeps things simple because it means that all firms make zero profits in equilibrium.) Using the zero-profit conditions of internationally active producers, we can solve for the price in the poor country

$$p^P = \frac{1}{1+\tau-\pi\tau}\left(\frac{F}{\mathcal{P}} + \frac{1+\tau}{a} - \pi\left(\frac{F}{\mathcal{P}} + \frac{1}{a}\right)\right).$$

Using Assumption 1, it is easy to show that stricter enforcement of parallel import restrictions or, equally, a higher share of products with "national exhaustion" rules (both represented by a higher π) is associated with lower prices p^P for internationally traded products in the poor country. The reason is that a higher π increases the incentive for rich-country firms to trade internationally. This generates a pro-competitive effect reducing prices in the poor country and thus increasing the welfare of the households there. The welfare of the rich country households falls. To see this, notice that the number of goods produced in the poor country is still given by equation (3) which does not depend on π. Using the budget constraint of poor consumers $L^P = p^P(N^P + N_T^R)$ lets us calculate

$$\frac{\partial N_T^R}{\partial \pi} = \frac{\partial N_T^R}{\partial p^P}\frac{\partial p^P}{\partial \pi} > 0$$

Since the number of goods produced in the poor country remains unchanged, welfare of households in the poor country increases. The pro-competitive effect on price in the poor country implies that poor consumers can afford more goods, improving their welfare. The opposite is true for the rich country. As the higher π induces more rich country firms to trade internationally, a larger fraction of resources is devoted to the production of these internationally traded goods. Hence a smaller range of goods is produced in the rich country and the total number of goods produced worldwide goes down. This establishes that welfare of rich-country consumers falls.

The effect of trade liberalizations (lower τ) on trade is now ambiguous. To see this consider first the case of "national exhaustion", i.e. $\pi = 1$. In such a situation producers can perfectly price discriminate between countries and hence a full trade equilibrium always prevails. Lower trade costs simply free resources to produce additional varieties. As all these additional varieties will be traded, trade will increase as a reaction to the liberalization. We have seen above that the converse holds for "international exhaustion", $\pi = 0$. Thus, in general the effect of a trade liberalization depends on the share of products with "national exhaustion" rules π, with positive effects more likely the closer π is to one.

6.4 Within-country inequality

Non-homotheticicities not only generate important effects of per capita endowments on trade, they also imply that within-country inequality may shape trade patterns in an important way. In our context, the case of within-country inequality can be best understood as a special case of a multi-country model, where the trade costs between some countries are zero. Between these countries, the parallel trade restrictions would immediately become binding as long as there is a slight difference in per capita endowments. In other words, with inequality within countries we always have some degree of exclusion, meaning that some firms will only sell to the rich charging a high price while other firms will sell to all consumers.

We highlight the role of inequality by looking at the most simple case of two identical countries that are both populated by rich and poor households. This simple example shows which mechanisms are present in the more general cases. In particular, we demonstrate that the threat of parallel trade under within country inequality affects trade patterns even when countries are completely symmetric.

Low trade costs We assume that β percent of the population in every country are poor owning an endowment of $\theta L < L$ and that the remaining $1 - \beta$ percent of the population are rich and own an endowment $[(1 - \beta\theta) / (1 - \beta)] L > 1$ (hence per capita endowment is still L). We index rich and poor households by r and p, respectively. Let us first consider the case when trade costs are low. In a fully integrated market when $\tau = 1$ there are two group of firms: "mass producers" selling to all consumers in both countries and "exclusive producers" selling only to the rich in both countries. In fact, this equilibrium holds true for small values of τ such that the scope for price discrimination between countries is limited. By symmetry, we have factor price equalization and set $W = 1$. Using countries' resource-constraints, firms' zero-profit and households' budget constraints,[18] it is straightforward to calculate the number of products, N^p and N^r, sold to everyone and to the rich, respectively, as

$$N^p = \frac{\theta a L}{aF/\mathcal{P} + 1 + \tau} \text{ and } N^r = \frac{(1-\theta)aL}{aF/\mathcal{P} + (1+\tau)(1-\beta)}.$$

[18] The resource constraint is given by $L\mathcal{P} = N^p [F + (1+\tau)\mathcal{P}/a] + N^r [F + (1-\beta)(1+\tau)\mathcal{P}/a]$ in both countries, the zero-profit conditions are $2p^p\mathcal{P} = F + (1+\tau)\mathcal{P}/a$ for firms that sell to all households worldwide and $2(1-\beta)p^r\mathcal{P} = F + (1-\beta)(1+\tau)\mathcal{P}/a$ for firms that sell only to rich households in both countries. Rich households' budget constraints are given by $[(1-\beta\theta)/(1-\beta)] L = 2p^r N^r + 2p^p N^p$in both countries, and poor households' budget constraints are $\theta L = 2p^p N^p$.

It is straightforward to study the impact of a trade liberalization. Using $U^p = N^p$ and $U^r = N^p + N^r$ it follows that both rich and poor consumers gain from a trade liberalization. While the rich gain more in absolute terms, the poor gain more in relative terms. Hence a trade liberalization even reduces consumption inequality $N^p / (N^r + N^p)$. This result generalizes to the case of more types of consumers. Notice that this is quite different from inequalities across countries where we saw that a fall in trade costs helps the richer but hurts poorer households.

High trade costs The above equilibrium arises if trade costs are sufficiently small. When trade costs are higher, an equilibrium with exclusive producers selling only to rich consumers and mass producers selling to all consumers worldwide no longer exists. The reason is that high trade costs make a new strategy attractive: sell to all consumers at home and only to rich consumers abroad. It can be shown (notes available from the authors upon request) that, depending on the extent of within-country inequality, either of two different equilibrium scenarios will emerge. If inequality is high, the equilibrium is characterized by some firms selling only to rich consumers in both countries while other firms follow a "separating" strategy: selling to both types of households in the home market; and only to the rich on the foreign market. With high within-country inequality, these strategies yield zero profit in equilibrium; and they strictly dominate the strategy of selling to rich and poor consumers in both markets.

In contrast, if inequality is low and trade costs are high, the equilibrium is characterized by the co-existence of firms selling to all households worldwide and firms selling to all consumer on the home market and to rich consumers on the foreign market. The exclusive strategy, i.e. selling only to the rich on both markets is not a profitable option. We find that trade liberalizations increase welfare for both types of consumers. The results that, in relative terms, the poor gain more than the rich from a trade liberalization continues to hold.

7 More general preferences

We assumed 0-1 preferences in the analysis above. On the one hand, this assumption yields a framework that is highly tractable and generates closed-form solutions. On the other hand, this assumption restricts households' adjustments to the extensive margin. The standard CES case with all adjustments happening on the intensive margin and our 0-1 case should thus be understood as two polar cases. We go beyond these polar cases in this section and introduce more general preferences that allow for adjustments on both margins. In particular, we will show that the qualitative characteristics of the equilibria we obtained with 0-1 preferences carry over to the case of more general preferences featuring both non-trivial intensive *and* extensive margins.

We take up the analysis of Section 4 where we study two equally large, but unequally rich countries. Within countries, households are identical. Trade patterns are therefore shaped by differences in per capita endowments across countries. Now replace 0-1 preferences by the following general utility function

$$U = \int_0^\infty v(c(j))dj,$$

where $c(j)$ denotes the consumed quantity of good j. It is assumed that the subutility $v()$ satisfies $v' > 0$, $v'' < 0$ and $v(0) = 0$. Beyond these standard assumptions on $v()$, we make two further crucial assumptions: (i) $v'(0) < \infty$ and (ii) $-v'(c)/[v''(c)c]$ is decreasing in c. The former assumption implies that reservation prices are finite and there is therefore a non-trivial extensive margin of consumption; the latter assumption implies that the price elasticity of demand is decreasing along the demand curve. Notice that monopolistic pricing implies $p = (1 + v'(c)/[v''(c)c])^{-1}W/a$. To ease notation, we denote by $\mu(c) \equiv (1 + cv''(c)/v'(c))^{-1}$ a monopolistic firm's mark-up. Since $-v'(c)/[v''(c)c]$ is decreasing in c, we have $\mu'(c) > 0$.

It is straightforward to see that finite reservation prices again make an autarky equilibrium possible. To ensure that there will be trade, we have to adjust the trade condition of Assumption 1 as follows

Proposition 4 *Denote by c_a^R consumption per variety under autarky in the rich country. If trade costs are sufficiently small $\tau < \mu(c_a^R)v'(0)/v'(c_a^R)$ where $aF/\mathcal{P} = c_a^R(\mu(c_a^R) - 1)$, trade occurs in equilibrium.*

Proof. The proof is based on the same tenet as in section 3 above. We determine the autarky equilibrium and ask under which conditions an entrepreneur has incentives to sell his products abroad. Setting $W = 1$, optimal monopolistic pricing implies $p = \mu(c)/a$. With free entry, profits $\mathcal{P}(p_a^R - 1/a)c_a^R$ must equal set up costs F

$$aF/\mathcal{P} = \left(\mu(c_a^R) - 1\right) c_a^R \tag{14}$$

The equilibrium is symmetric for all firms, hence the resource constraint reads

$$L^R = N_a^R \left(F + \mathcal{P}c_a^R/a\right) \tag{15}$$

Solving (14) and (15) for c_a^R and N_a^R, we see that c_a^R does not depend on L^R. Hence when the two countries differ only in L^i but have equal populations, intensive consumption levels under autarky are identical between the two countries, $c_a^R = c_a^P$. Selling one marginal unit abroad at price $v'(0)/\lambda_a^P$, allows the purchase of $v'(0)/(\lambda_a^P p_a^P)$ foreign goods. Since $\lambda_a^P = v'(c_a^P)/p_a^P$ and $c_a^R = c_a^P$ this is equal to $v'(0)/v'(c_a^R) > 1$. Reselling this (new) product at home, yields a price $v'(0)p_a^R/v'(c_a^R)$ minus trade costs. Hence, this strategy is profitable if

$\left[v'(0)p_a^R/v'(c_a^R)\right] \cdot \left[v'(0)/v'(c_a^R)\right] > \tau^2$. Expressing p_a^R in terms of c_a^R, we get the condition of the Proposition. ■

We are now able to discuss the *full trade* regime. Let us start from a full trade equilibrium where differences in per capita endowment L^P/L^R are sufficiently close to unity, so that firms do not face a threat of parallel trade. We denote the optimal price of a country-R firm on its home market by p_R^R and the corresponding price in country P by p_R^P. Monopoly pricing implies $p_R^R = \mu(c_R^R)/a$ and $p_R^P = \mu(c_R^P)\tau/a$, respectively. The firm does not face any threat of parallel trade if $p_R^R/p_R^P \leq \tau$, or if

$$\frac{\mu(c_R^R)}{\mu(c_R^P)} \leq \tau^2. \tag{16}$$

This is a sufficient condition for the existence of a full trade equilibrium.

Let us now consider the existence of a *partial trade* equilibrium. A necessary condition for such an equilibrium is that condition (16) is violated. This condition may be violated either if τ is close to unity, or if consumption levels c_R^R and c_R^P diverge strongly (recall that $\mu'(c) > 0$ and $\mu(0) = 1$).[19] If L^P approaches zero, c_R^P approaches zero as well, and the denominator in equation (16) approaches unity. There is thus a level of trade cost τ sufficiently close to unity and/or a country-P endowment L^P sufficiently small, so that we get $\mu(c_R^R) > \tau^2\mu(c_R^P)$.

When the inequality in (16) is violated, exporting firms will set prices $p_R^R/p_R^P = \tau$ to prevent parallel trade. Notice, however, that violation of condition (16) does not necessarily imply a partial trade equilibrium. The reason is that there is adjustment both along the extensive margin and along the intensive margin. Even when condition (16) is violated, all goods may be traded as country-P households may still consume all goods produced worldwide, but in lower quantities. In other words, violation of the condition (16) is necessary but not sufficient for a partial trade equilibrium.

To show under which conditions country-P households will not consume all goods produced worldwide, we need to look at incentives of country-R firms to sell their products exclusively to rich domestic consumers. The profit of a country R producer is given as follows (to ease notation let us write $p_R^R \equiv \tau p$ and $p_R^P \equiv p$)

$$\pi = \mathcal{P}\left(\tau p - 1/a\right) c_R^R + \mathcal{P}\left(p - \tau/a\right) c_R^P.$$

The demand curve of country-R consumers is given by $v'(c_R^R) = \lambda^R \tau p$ and the corresponding demand curve of country-P consumers is $v'(c_R^P) = \lambda^P p$. Hence we have, $dc_R^R/dp = (1/p)v'(c_R^R)/v''(c_R^R)$ and $dc_R^P/dp = (1/p)v'(c_R^P)/v''(c_R^P)$. The first order condition of the monopolistic firm's price

[19] By d'Hôpital's rule, noting that $v(0) = 0$ and $v'(0)$ finite, $\lim_{c\to 0} v'(c)c/v(c) = \lim_{c\to 0} 1 + v''(c)c/v'(c)$. However, $\lim_{c\to 0} v'(c)c/v(c) = v'(0) \cdot \lim_{c\to 0} c/v(c) = v'(0)/v'(0) = 1$. This implies $\lim_{c\to 0} v''(c)c/v'(c) = 0$ and $\lim_{c\to 0} \mu(c) = 1$.

setting choice is given by

$$\frac{\tau p-1/a}{\tau p}\left(-\frac{v'(c_R^R)}{v''(c_R^R)}\right)+\frac{p-\tau/a}{p}\left(-\frac{v'(c_R^P)}{v''(c_R^P)}\right)=\tau c_R^R+c_R^P.$$

To examine whether a partial trade equilibrium exists, let L^P and therefore c_R^P approach zero. The first order condition then becomes

$$\frac{\tau p-1/a}{\tau^2 p}\left(-\frac{v'(c_R^R)}{v''(c_R^R)c_R^R}\right)+\frac{p-\tau/a}{\tau p c_R^R}\left(-\lim_{c_R^P\to 0}\frac{v'(c_R^P)}{v''(c_R^P)}\right)=1 \tag{17}$$

Now consider the optimal decision of a country-R firm which decides to produce for domestic consumers only. Denoting by p^N and c_R^N price and quantity of non-traded goods, the first order condition for exclusive producers is

$$\frac{p^N-1/a}{p^N}\left(-\frac{v'(c_R^N)}{v''(c_R^N)c_R^N}\right)=1. \tag{18}$$

We now compare equations (17) and (18) for the case where τ is sufficiently close to 1 such that $p>\tau/a$. If $v'(0)/v''(0)$ is larger than zero - which is fulfilled if $v''(0)$ is finite - the price of a non-exporting firm p^N is strictly larger than the price of an exporting firm τp. Since $c_R^P\to 0$ when $L^P\to 0$, export revenues are zero, hence profits of the non-exporting firm must be higher because it sets the profit maximizing price $p^N>\tau p$. This implies that an outcome where all firms export cannot be an equilibrium – provided that L^P is sufficiently close to zero and τ is sufficiently close to one. We summarize our discussion in

Proposition 5 *If $v''(0)$ is finite, a partial trade equilibrium always exists.*

Proof. In text. ■

Note that $v''(0)$ all is always finite when preferences are HARA[20] with $v'(0)<\infty$.

8 Conclusions

This paper incorporates non-homothetic preferences into a standard "new" trade theory framework. We propose modeling non-homotheticities by indivisible consumer goods that are either consumed in unit quantity or not consumed at all. Such a specification implies that consumer choice is along the extensive margin whereas a choice along the intensive margin of consumption is ruled out by assumption. This is orthogonal to the standard CES-framework where households have infinite reservation prices and the allocation of expenditures relates solely to the intensive margin of consumption.

[20] Well known and often used special cases of the HARA class with $v'(0)<\infty$ are quadratic preferences or Stone-Geary with negative minimum consumption, for example.

We elaborate the *role of per capita incomes* in international trade patterns which, for given aggregate output, is absent in any homothetic model of international trade. Consider two countries with the same aggregate endowment, one country is small and rich and the other country is large and poor. Our model predicts that large differences in per capita endowments lead to a partial world trade equilibrium in which many goods produced in the rich country will not be traded and consumed in the poor country, while all goods produced in the poor country will be traded and consumed in the rich country. In contrast, when differences in per capita endowments are small, a full trade equilibrium emerges. In such an equilibrium, all goods produced in the two countries are traded internationally and consumed in both countries. Hence our model features the famous Linder hypothesis according to which countries that are more similar in per capita endowments trade more intensively with one another.

Our analysis provides us with a simple general equilibrium framework of *parallel trade*. The partial world trade equilibrium emerges when inequality across countries is high so that differences across countries in consumers' willingness to pay for differentiated products are very large. In that case, the threat of parallel trade limits the scope of price setting in the rich country. This is because arbitrage traders can purchase the good cheaply in the poor country, ship it back and underbid local producers in the rich country. To inhibit such parallel trade, internationally active firms have to set low prices in the rich country. In equilibrium, firms in the rich country are indifferent between selling their product on the world market and selling their product only on the home market. The general equilibrium perspective of our model makes the fraction of internationally active firms endogenous. This effect is typically not considered in partial equilibrium settings of parallel trade but has a potentially important impact on trade patterns.

Concerning the welfare effects of trade, we find that a trade liberalization (a reduction in iceberg trade costs) increases welfare of consumers in both countries when the world economy is in a full trade equilibrium, but hurts the poor country (and benefits the rich country) when the world economy is in a partial trade equilibrium. The reason for the latter result is that exporters of the poor country need to reduce prices of traded goods in the rich country to inhibit parallel trade, while exporters of the rich country have no such restrictions in the poor country. Consumers in the rich country face decreasing prices, and consumers in the poor country are confronted with a lower range of import goods because higher competition on the world market induces rich-country firms to concentrate their sales exclusively on the home market.

While our analysis is made under very specific assumptions, our model is simple enough to be extended in several directions. We extended our set-up to more than two countries, to heterogeneous trade costs, to commercial policies, and to within-country inequality. Finally, we showed that partial trade equilibria emerge for a broad class of more general preferences.

Our model is complementary to existing supply side approaches and potentially helpful in understanding the dynamics of world trade patterns that arise due to major changes in the distribution of world purchasing power. This is particularly relevant in the case of large emerging markets such as China, India, Brazil, etc. that have experienced high growth in per capita incomes over the past decades. From an empirical point of view, disentangling the demand effects emphasized in this paper from the supply/technology factors emphasized in the standard model is of particular interest.

References

[1] Behrens, Kristian and Yasusada Murata (2009) "Globalization and Individual Gains from Trade," CEPR Discussion Paper No. 7448.

[2] Bergstrand, Jeffrey H. (1990) "The Heckscher-Ohlin Samuelson Model, the Linder Hypothesis and the Determinants of Bilateral Intra-Industry Trade", *The Economic Journal* 100, 1216-1229.

[3] Chaney, Thomas (2008) "Distorted Gravity: The Intensive and Extensive Margins of International Trade," *American Economic Review* 98, 1707-21.

[4] Choi, Yo Chul, David Hummels, and Chong Xiang (2006) "Explaining Export Variety and Quality: The Role of the Income Distribution", NBER Working Paper No. 12531.

[5] Chung, Chul (2005) "Nonhomothetic Preferences as a Cause of Missing Trade and Other Mysteries", *mimeo.*

[6] Dalgin, Muhammed, Devashish Mitra, and Vitor Trindade (2008) "Inequality, Nonhomothetic Preferences, and Trade: A Gravity Approach", *Southern Economic Journal* 74, 747-774.

[7] Desdoigts Alain and Fernando Jaramillo (2009) "Trade, Demand Spillovers, and Industrialization: The Emerging Global Middle Class in Perspective," *Journal of International Economics* 79: 248-258

[8] Fajgelbaum, Pablo, Gene Grossman, and Elhanan Helpman (2009) "Income Distribution, Product Quality and International Trade." NBER Working Paper No. 15329.

[9] Falkinger, Josef (1990) "Innovator-Imitator Trade and the Welfare Effects of Growth", *Journal of Japanese and International Economies* 4: 157-172.

[10] Flam, Harry and Elhanan Helpman (1987) "Vertical Product Differentiation and North-South Trade", *American Economic Review* 77, 810-822.

[11] Fieler, Ana (2010) "Non-Homotheticity and Bilateral Trade: Evidence and a Quantitative Explanation", mimeo, University of Pennsylvania.

[12] Foellmi, Reto and Josef Zweimüller (2006) "Income Distribution and Demand-Induced Innovations", *Review of Economic Studies* 73, 941-960.

[13] Francois, Joseph F. and Seth Kaplan (1996) "Aggregate Demand Shifts, Income Distribution, and the Linder Hypothesis", *Review of Economics and Statistics* 78, 244-250.

[14] Ganslandt, Mattias and Keith E. Maskus (2004) "Parallel Imports and the Pricing of Pharmaceutical Products: Evidence from the European Union", *Journal of Health Economics* 23, 1035-1057.

[15] Ganslandt, Mattias and Keith E. Maskus (2007) "Intellectual Property Rights, Parallel Imports and Strategic Behavior," Working Paper Series 704, Research Institute of Industrial Economics

[16] Grossman, Gene M. and Edwin L.-C. Lai (2004) "International Protection of Intellectual Property", *American Economic Review*, Vol. 94(5), 1635-1653.

[17] Grossman, Gene M. and Edwin L.-C. Lai (2006) "Parallel Imports and Price Controls", NBER Working Paper No. 12432.

[18] Helpman, Elhanan, Marc Melitz, and Yona Rubinstein (2007) "Trading Partners and Trading Volumes", NBER Working Paper No. 12927.

[19] Hsieh, C. and P. Klenow (2007) "Relative prices and relative prosperity", *American Economic Review* 97, 562-585.

[20] Hunter, Linda C. (1991) "The Contribution of Nonhomothetic Preferences to Trade", *Journal of International Economics* 30, 345-358.

[21] Hunter, Linda C. and James R. Markusen (1988) "Per Capita Income as a Basis for Trade", in Robert Feenstra, ed., *Empirical Methods for International Trade*, MIT Press, Cambridge MA and London.

[22] KPMG (2003), "The Grey Market," Study in Cooperation with the Anti-gray Market Allience.

[23] Krishna, Kala and Cemile Yavas (2005) "When Trade Hurts: Consumption Indivisibilities and Labor Market Distrotions", *Journal of International Economics* 67, 413-427.

[24] Krugman, Paul R. (1980) "Scale Economies, Product Differentiation, and the Pattern of Trade", *American Economic Review*, Vol. 70(5), 950-959.

[25] Linder, Staffan B. (1961) *An Essay on Trade and Transformation*, Almqvist and Wiksells, Uppsala.

[26] Manova, Kalina and Zhiwei Zhang (2009), Export Prices and Heterogeneous Firm Models, *mimeo*, Stanford University.

[27] Markusen, James R. (1986) "Explaining the Volume of Trade: An Eclectic Approach", *American Economic Review*, Vol. 76(5), 1002-1011.

[28] Markusen, James R. (2010) "Putting per capita income back into trade theory," *mimeo*, University College Dublin and University of Colorado, Boulder.

[29] Maskus, Keith E. (2000) "Parallel Imports", *The World Economy*, Vol. 23(9), 1269-1284.

[30] Matsuyama, Kiminori (2000) "A Ricardian Model with a Continuum of Goods under Non-homothetic Preferences: Demand Complementarities, Income Distribution, and North-South Trade", *Journal of Political Economy*, Vol. 108(6), 1093-1120.

[31] Melitz, Marc J. (2003) "The Impact of Trade on Intra-Industry Allocation and Aggregate Industry Productivity", *Econometrica*, Vol. 71(6), 1695-1725..

[32] Mitra, Devashish and Vitor Trindade (2005) "Inequality and Trade", *Canadian Journal of Economics*, Vol. 38(4), 1253-1271.

[33] Murphy, Kevin, Andrei Shleifer, and Robert Vishny (1989) "Income Distribution, Market Size, and Industrialization," *Quarterly Journal of Economics* 104, 537-564.

[34] National Economic Research Associates (1999) *The Economic Consequences of the Choice of a Regime of Exhaustion in the Area of Trademarks*, Final Report for DGXV of the European Commission, London.

[35] Ravn Morten, Stephanie Schmitt-Grohe, and Martin Uribe (2007) "Incomplete Cost Pass-Through Under Deep Habits", NBER Working Paper No. 12961.

[36] Sauré, Philip (2009) "Productivity Growth, Bounded Marginal Utility, and Patterns of Trade", *mimeo*, Swiss National Bank.

[37] Simonovska, Ina (2009), Income Differences and Prices of Tradables, *mimeo*, University of California Davis.

[38] Trefler, Daniel (1995) "The Case of the Missing Trade and Other Mysteries", *American Economic Review*, Vol. 85(5), 1029-1046.

A Appendix 1: Three-country equilibrium

In equilibrium, the current account of each with the rest of the world has to be equalized. Denoting by N_j^i the imports of country j from country i, balanced trade implies

$$\begin{aligned} 2\tau p^P N_P^M &= p^P \left(N_P^M + N_P^R\right) \text{ for country } P \\ (1+\tau)p^P N_P^M + \tau p^M N_R^M &= p^M N_M^R + \tau p^P \left(N_M^P + N_M^R\right) \text{ for country } M \\ (1+\tau)p^P N_P^R + p^M N_M^R &= \tau p^M N_R^M + \tau p^P \left(N_R^P + N_R^M\right) \text{ for country } R \end{aligned}$$

Note that we have 5 linearly independent equations and 6 unknowns (the prices p^i, $i \in \{P, M, R\}$ are determined by the zero profit conditions). Hence, it is only possible to determine the sum of $\left(N_P^M + N_P^R\right)$. Because of this indeterminacy we are free to consider the case where all bilateral trade flows are balanced.[21] Then we have

$$\tau N_P^P = N_P^M = N_P^R \text{ and } \tau N_M^M = N_M^R.$$

Using the resource constraints we may calculate the goods produced in the three countries

$$\begin{aligned} N_P^P &= \frac{L^P \mathcal{P}}{F + \frac{1+2\tau}{a}\mathcal{P}}, \quad N_P^M = \frac{\tau L^P \mathcal{P}}{F + \frac{1+2\tau}{a}\mathcal{P}}, \quad N_P^R = \frac{\tau L^P \mathcal{P}}{F + \frac{1+2\tau}{a}\mathcal{P}} \\ N_M^M &= \frac{\left(L^M - \tau L^P\right)\mathcal{P}}{F + \frac{1+\tau}{a}\mathcal{P}} \quad N_M^R = \frac{\tau\left(L^M - \tau L^P\right)\mathcal{P}}{F + \frac{1+\tau}{a}\mathcal{P}} \\ N_R^R &= \frac{\left[L^R - \tau L^M - \tau(\tau - 1)L^P\right]\mathcal{P}}{F + \frac{1}{a}\mathcal{P}}. \end{aligned}$$

[21] Notice that countries M and R consume all goods produced in country P; and country R consumes all goods produced in country M.

B Appendix 2: Proof of proposition 3

Clearly, we have $\partial U^P(\tau)/\partial\tau > 0$. We check the sign of the derivative $\partial U^M(\tau)/\partial\tau$.

$$\begin{aligned}
\frac{\partial U^M(\tau)}{\partial \tau} &= L^P\mathcal{P}\frac{2F}{\left[F+\frac{1+2\tau}{a}\mathcal{P}\right]^2} - L^P\mathcal{P}\frac{2F}{\left[F+\frac{1+\tau}{a}\mathcal{P}\right]^2} + L^P\mathcal{P}\frac{\left[L^M/L^P+1-2\tau\right]F-\frac{(1+\tau)^2}{a}\mathcal{P}}{\left[F+\frac{1+\tau}{a}\mathcal{P}\right]^2} < 0 \\
&\text{iff} \\
\frac{\left[F+\frac{1+\tau}{a}\mathcal{P}\right]^2}{\left[F+\frac{1+2\tau}{a}\mathcal{P}\right]^2} &< \frac{-\left(L^M/L^P-\tau\right)F+(1+\tau)(F+\frac{(1+\tau)}{a}\mathcal{P})}{2F}
\end{aligned}$$

In the special case $L^M/L^P = \tau$ the inequality becomes $\frac{\left[F+\frac{1+\tau}{a}\mathcal{P}\right]^2}{\left[F+\frac{1+2\tau}{a}\mathcal{P}\right]^2} < \frac{(1+\tau)(F+\frac{(1+\tau)}{a}\mathcal{P})}{2F}$. Hence, when L^M/L^P is not too large, the middle income country will gain.

Finally, $\partial\left[U^P(\tau)\mathcal{P}+U^M(\tau)\mathcal{P}+U^R(\tau)\mathcal{P}\right]/\partial\tau < 0$. As $\partial N_R^R/\partial\tau < 0$, we must have $\partial U^R(\tau)/\partial\tau < \partial U^M(\tau)/\partial\tau$. Hence, $\partial U^R(\tau)/\partial\tau < 0$ whenever $\partial U^P(\tau)/\partial\tau > 0$.

Figure 1: Demand function

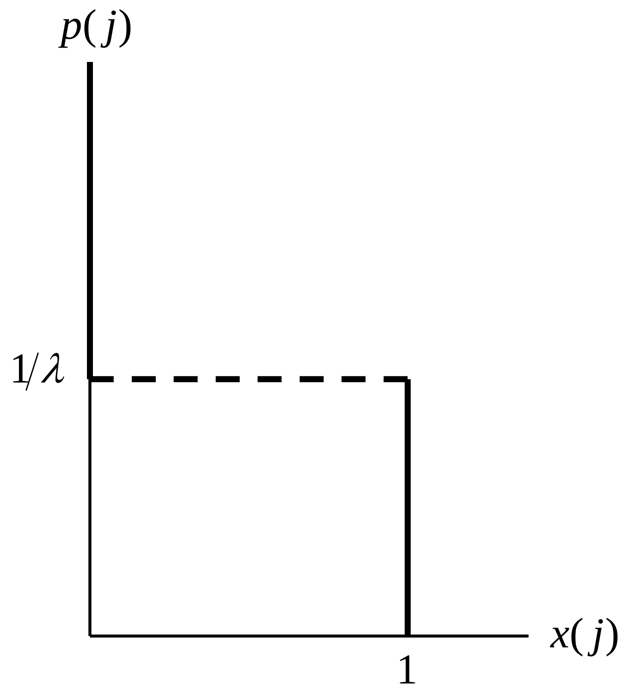

Figure 2: Partial vs. full trade equilibria

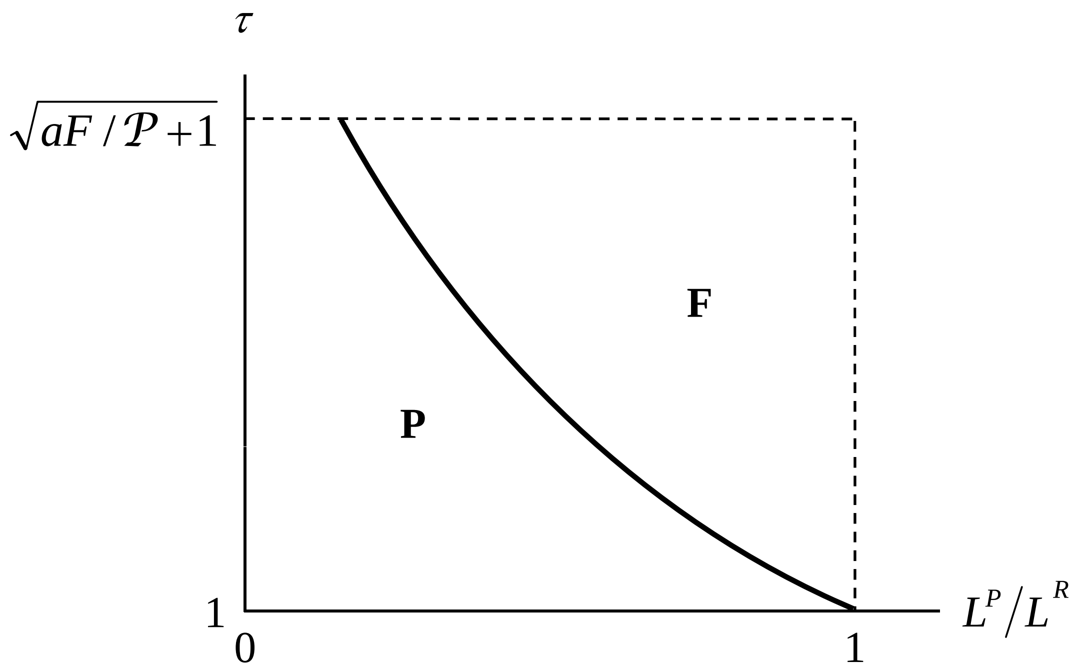

Figure 3: Trade intensity as a function of relative per-capita endowments

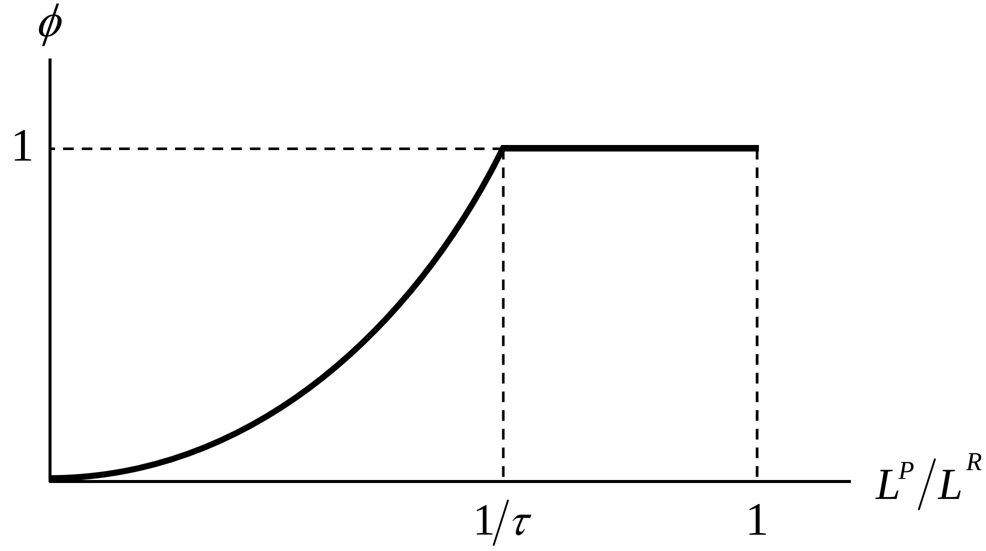

Figure 4: Welfare and trade costs

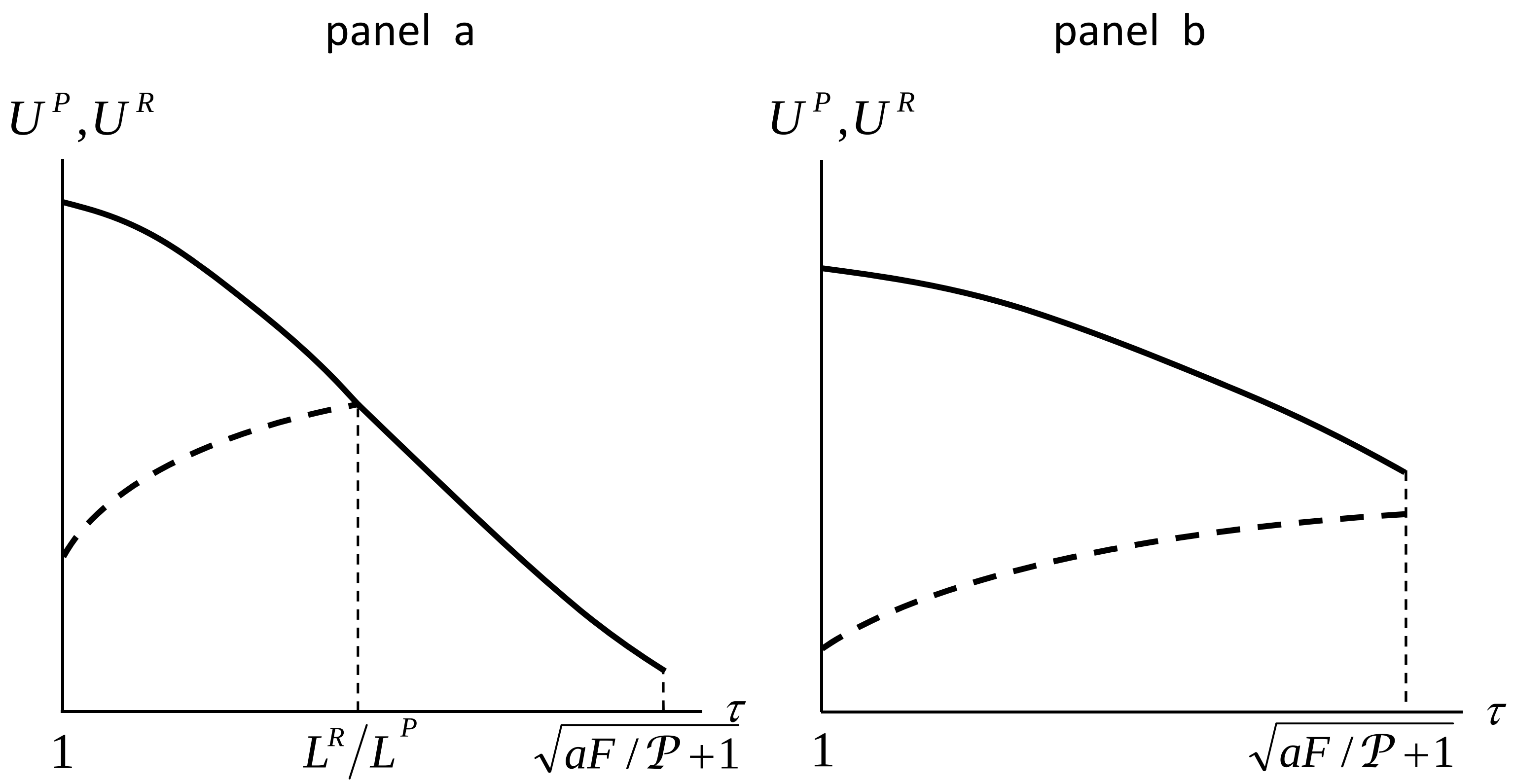

Figure 5: Relative per-capita endowments vs. relative population sizes