

Productivity Growth, Bounded Marginal Utility, and Patterns of Trade

Philip Sauré¹

European University Institute

Job Market Paper

Updated Version - Substantially Changed

Abstract

The workhorse model of the New Trade Theory fails to explain four strong and central patterns of postwar trade data. These patterns are, first, the massive increase in trade volumes, second, the small fraction of traded varieties the average country imports, third the correlation between per capita income growth and trade growth, and fourth, the correlation between trade growth and growth in the number of source country per imported good. The present paper shows that a small and reasonable change in the demand structure can reconcile the model with the data. It departs from standard theory by assuming that consumers derive bounded marginal utility from varieties. This implies that consumers purchase only the cheaper share of varieties and expensive foreign varieties that bear high transport costs are not imported. Technological progress which increases per capita consumption of the varieties in the consumption basket decreases marginal utility derived from each of them and induces consumers to extend their consumption to more expensive varieties produced at more distant locations. This additional margin along which trade can expand induces a substantial increase in the trade share as productivity grows. Productivity change is thus identified as a joint determinant and driving force of trade shares, the number of source countries per good and per capita income, explaining at the same time the trends and correlations in the data.

JEL-classification: F1, F4

Keywords: Trade Volume, Source Country

¹I would like to thank Giancarlo Corsetti, Gino Gancia, Omar Licandro, Marco Maffezzoli, Diego Puga, Morten Ravn, Karl Schlag, Jaume Ventura, and Joachim Voth for many comments and helpful discussions.

1 Introduction

The impact of Krugman's (1980) workhorse model in international economics can hardly be overestimated. As the most successful formalization of the New Trade Theory, it inspired research on economic agglomeration, trade agreements and economic development, sparked off the New Economic Geography and the literature on trade and heterogeneous firms, and hence reaches far beyond the analysis of the home market effect and intra-industry trade it originally set out to explain. The model is one of the building blocks of modern trade theory.

Yet recent literature points out four strong and important regularities of trade data that seem to mark some limits of the New Trade Theory. These regularities are, first, the massive growth of trade volumes, second, the tiny fraction of tradable varieties that countries tend to purchase, third, the parallel growth of per capita income and trade, fourth and finally, the strong correlation between trade growth and the growth in the number of source countries per imported good.

The first observation, the spectacular rise in trade shares is well-known. After Krugman (1995) observed a surprising disagreement on its causes, Baier and Bergstrand (2001) singled out tariff reductions as its most important determinant. Yi (2003) subsequently pointed out that the observed increases in world trade shares imply an excessive import elasticity in standard trade models (see also Bergoening et al (2004) on this point). Substantial progress was made to explain this "elasticity puzzle" as it was later labeled (see Yi (2003), Ruhl (2005), and Cuñat and Maffezzoli (2006)), however, the New Trade Theory remains unable to square the raise of trade volumes at reasonable import elasticities.

Haveman and Hummels (1999) observed the second empirical regularity. Defining varieties as goods differentiated by production origin², the authors report that "importers purchase a very small fraction of available varieties," while standard models predict that each country imports all varieties from all trade partners. The authors conclude that the data suggest that standard theory "considerably overstates the extend of specialization [...] or the degree to which consumers value varieties," and Anderson and van Wincoop (2003) acknowledge that their observation is an "important drawback of the existing theory."

The third observation concerns the nature of the rise in trade volumes. Figure 1 illustrates with US and the aggregate world data of the period 1972 to 2000 that the growth of trade was paralleled by an increase of per capita income. Ventura (2005) puts forward that the data strongly suggest a positive relation between trade growth and growth in per capita income. The New Trade Theory, however, predicts that trade shares and productivity parameters are entirely orthogonal, so that any correlation of per capita income and trade variables remains outside its framework.

²This is a common definition that departs from that of Krugman (1980). The latter, which identifies a variety by the firm it produces, will be adopted in the present paper.

The fourth and final observation, Broda and Weinstein (2004) report a substantial expansion of the number of source countries per imported good for all of the 20 largest importers over the period 1972 to 2000. This increase reflects a dynamic aspect of the second observation above. Countries tend to purchase each of their imports from trade an increasing number of trade partners which they had previously excluded from the set of suppliers. Figure 2 illustrates that the expansion of the average number of source countries per good closely parallels the increase of the import share in the US data. Disaggregating these data by good categories (HTS and TSUSA) confirms this pattern: Figure 3 shows a strong correlation between changes in trade volumes and changes in the average number of countries from 1972 to 2000. Finally, relying on bilateral trade data of 188 countries by 4-digit SITC good categories, Figure 4 repeats this plot adding a country dimension. The graph exhibits the same strong correlation between changes in import volumes and in the origin margin. Once again, the New Trade Theory has nothing to say about the trend nor the observed correlations.

Together this makes a list of four strong and central patterns of international trade which the New Trade Theory fails to address and one might in consequence be tempted as to downgrade the model as an analytical instrument for trade theory. The present paper shows that this is not necessary. It argues that a small and realistic deviation from the standard demand structure goes a long way in alleviating the listed drawbacks. By the paper's central assumption, consumers' marginal utility derived from each variety is bounded. This twist in the demand structure immediately implies that consumers buy only the cheaper fraction of varieties and exclude the too expensive ones from their consumption basket. Varieties from very distant destinations whose prices bear large trade costs are therefore not purchased, which explains the observation of void bilateral trade relations in some sectors. Moreover, the change of demand patterns implies that technological progress, when it is biased towards the marginal production cost, raises the number of source countries per good. The reason for this effect is slightly more subtle. A decreasing ratio of fix cost over marginal cost induces a higher per capita consumption of varieties already consumed and marginal utility derived from each of them falls. At the same time marginal utility of varieties outside the consumption basket stays trivially constant. This shift of the ratio of marginal utility means that varieties outside the consumption basket become more attractive and consumers expand their consumption basket. Consequently, the number of source countries for a given import good rises.

With the expansion of the consumption basket towards the more expensive foreign varieties, the effect just described induces also a reallocation of expenditure towards imported varieties. Productivity growth becomes a joint determinant of trade shares and the number of source countries per good, which hence explains the observed correlation between both (Figures 2-4). The remaining of the afore-

mentioned pattern of trade, the correlation of trade shares and per capita income follows, as growth of marginal productivity positively affects both of them.

In this way, the workhorse model of the New Trade Theory, slightly modified on the demand side, qualitatively explains the four empirical regularities the standard version failed to address. In order to evaluate to what extent it is able to alleviate the drawbacks quantitatively, a simple calibration exercise is performed. It shows that technological progress can account for about half of the observed increase in trade shares and the other half of the increase remains to be explained by the fall in trade cost, i.e. by reductions in tariffs and transport costs. The relatively modest reductions of US data then explain the data with a time-average of 3.6, which is still on the high side of estimates for elasticities (it peaks at an 8.8 in 1972 and is with 1.7 lowest in 2000) but constitutes substantial progress regarding the standard New Trade Model, which implies import elasticities well over 10 when explaining the trade shares with trade cost reductions only. In sum, the assumption of bounded marginal utility on the variety level thus goes a long way to alleviate the three shortcomings of the standard New Trade Model and brings it much closer to the data in these important dimensions.

The present paper is not the first one analyzing the role of non-homothetic demand for trade shares. A number of authors have formalized and tested the Linder Hypothesis according to which bilateral trade volumes correlate positively with per capita income similarity since per capita income governs the demand patterns and output structure of countries (see Thursby and Thursby (1987), Françoise and Kaplan (1996), and Hallak (2004)). Closer predecessors of the present paper are Markusen (1986) and Bergstrand (1990) who assume Stone-Geary preferences, so that consumers cover a minimum level of a homogenous, domestically produced good before demanding aggregates of imported varieties. The present model does not impose such asymmetries on demand but assumes equal valuation of varieties for consumers, while the cost of transportation creates endogenous asymmetries in demand via its effects on consumer prices. In a recent paper that stands somewhat out of this literature Hummels and Lugovskyy (2005) use a Lancaster-type utility to analyze the role of market size and per capita income on the market structure and international trade. As "richer consumers will pay more for varieties closer matched to their ideal types" they predict - and empirically confirm - that at higher per capita income markets are more segmented and own-price elasticities are lower, while the market size has opposite effects. In this respect, Hummels and Lugovskyy (2005) is very close to the present paper, which takes, however, the observed regularities of demand as a starting point, formulating them in reduced form to derive predictions concerning trade flows and the number of source countries per good.

The remainder of the paper is organized in the following way: section 2 reviews a standard multi-country New Trade Model and briefly discusses the three drawbacks. Section 3 introduces the change in the demand structure, showing the progress that

can be made there. Section 4 concludes.

2 The Standard New Trade Model

The present section aims to illustrate some implications of the Krugman (1980) model that stand in blatant contrast to the data. These implications are, first, that all countries consume all tradable goods and varieties from all countries at all times, second, that trade shares are independent of technologies and therefore of per capita income, and third, that trade shares react modestly to tariff reductions under realistic elasticities. To highlight these problems, the new trade model with a set of I different countries is briefly reviewed in the following.

Demand. Each country $i \in I$ is populated with identical individuals who have the following Cobb-Douglas preference structure over a nontradable good D_i and a composite of tradable varieties C_i

$$U_i = D_i^{1-\gamma} C_i^\gamma \quad (1)$$

The composite C_i aggregates the consumed quantities c_{ij} of the varieties $j \in J$ which the individual values with constant elasticity of substitution

$$C_i = \int_J c_{ij}^{1-1/\varepsilon} dj \quad (2)$$

The constant share of expenditure on the nontradable D will be read as the local content of the average variety. By the Cobb-Douglas structure the expenditure share on D is constant and the local content is $1 - \gamma$ throughout. When expenditure in county i equals W_i , the constant expenditure shares further allow to write the sub-budget constraints on the composite of varieties as

$$\int_J q_{ij} c_{ij} dj \leq \gamma W_i \quad (3)$$

where q_{ij} is the consumer price of the variety j in country i . Finally, consumer optimization of the variety composite implies $c_{ij} \sim q_{ij}^{-\varepsilon}$ so that the own-price demand elasticity is ε .

Supply. In each of the I countries the nontradable good D is produced competitively with a constant returns to scale technology

$$D = L_D \quad (4)$$

The tradable varieties are produced according to the increasing returns to scale technology

$$L_j = \alpha + \beta x_j \quad (5)$$

where L_j is labor and x_j is output of firm j . The positive parameter α represents a setup or entry cost in terms of units of labor while β is the marginal unit labor requirement. There is an unlimited pool of potential entrants into the market and each active firm produces one variety. Firms engage in monopolistic competition, and free entry into production ensures that operating profits exactly cover the setup cost.

Prices. In the competitive nontradable sector prices trivially equal wages within each country. In the tradable sector, international trade cost drive a wedge between producer and consumer prices. In particular, it will be assumed that shipments from country k to country i require the payment of a gross iceberg-type transport cost τ_{ki} . Consequently, in country i the consumer price of a variety j that is produced in country k is $q_{ij} = \tau_{ik}p_{ij}$ where p_{ij} is the price net of transport cost firm j sets in country i . In order to maximize profits, monopolists charge the constant markup $\varepsilon/(\varepsilon - 1)$ over marginal cost. As all firms located in a given country face the same production and pricing decisions, the prices of monopolists located in country k are all identical and can be indexed by the country of origin

$$p_k = \frac{\varepsilon}{\varepsilon - 1} \beta w_k \quad (6)$$

Note that the marginal production cost in country k is the product of marginal unit labor requirement β and the local wage w_k . With the monopolist prices (6) the relative consumer price of either pair of varieties equals the iceberg cost times the wage ratios. Thus, individual optimality condition for each country i requires

$$c_{ik} = c_{ii} (\tau_{ki} w_k / w_i)^{-\varepsilon} \quad (7)$$

This equation implies the well-known fact that all individuals consume all varieties in positive quantities³. This implies that whenever a country exports a good, it exports it to all countries worldwide, a prediction that stands in stark contrast to the trade patterns Haveman and Hummels (1999) report. In particular, the authors find that "importers purchase a very small fraction of available varieties" traded on the world market.

As a lemma of this finding, the number of source countries of the imported good equals the number of the rest-of-the-world countries and is, in particular, constant in the parameters of the model.⁴ This constitutes the second contradiction to the data since, as of Broda and Weinstein (2004) report, countries substantially increase the number of source countries per imported good over the past decades.

Equilibrium. Regardless of prices and income, the expenditure on the non-tradable good is $1 - \gamma$ so that, since prices are handed through as factor rewards,

³The cases of zero or infinite wage in one country can quickly be ruled out.

⁴This is of course true only under the implicit assumption that the number of countries in the world is constant. The increase in the number of nations that actually took place is, however, by far insufficient to explain the rise in the number of source countries per good.

the equilibrium labor allocation in the nontradable sector equals $(1 - \gamma)L_i$ in each country i . In the following description of the equilibrium, the trivial allocations in the nontradable sector will be neglected.

The equilibrium in the traded sector is determined by the sub-budget constraint, the resource constraint, and the trade balance for all countries $i \in I$. When writing n_k for the number of varieties produced in country k , and under the assumption that individuals spend all their labor income on consumption, the individual sub-budget constraint (3) in country i becomes

$$\sum_k n_k \tau_{ki} p_k c_{ik} = \gamma w_i \quad (8)$$

The resource constraint requires that within each country i the labor demand in the tradable sector equals γL_i , which is the inelastically supplied amount of labor in that sector

$$n_i \left[\alpha + \beta \sum_k L_k \tau_{ik} c_{ki} \right] = \gamma L_i \quad (9)$$

Finally balanced trade of country i implies

$$n_i p_i \sum_k L_k \tau_{ik} c_{ki} = L_i \sum_k n_k c_{ik} \tau_{ki} p_k \quad (10)$$

Combining (9) and (10) leads to

$$n_i \alpha + \beta L_i \sum_k n_k c_{ik} \tau_{ki} (p_k/p_i) = \gamma L_i$$

and determines with (6) and (8) the number of active firms in each country

$$n_i = \gamma L_i / (\alpha \varepsilon) \quad (11)$$

which does not depend on whether and how much countries trade. Using the monopolist prices (6), the consumer optimality condition (7) and the number of firms (11), one can write the individual sub-budget constraint (8) as

$$c_{ii} \sum_k L_k (\tau_{ki} w_k / w_i)^{1-\varepsilon} = (\varepsilon - 1) \alpha / \beta \quad (12)$$

and the trade balance (10)

$$\sum_k c_{kk} L_k \tau_{ik}^{1-\varepsilon} (w_k / w_i)^\varepsilon = c_{ii} \sum_k L_k (\tau_{ki} w_k / w_i)^{1-\varepsilon}$$

The last two equations combine to

$$\sum_k \frac{L_k w_k \tau_{ik}^{1-\varepsilon} w_i^{-\varepsilon}}{\sum_m L_m (\tau_{mk} w_m)^{1-\varepsilon}} = 1 \quad (13)$$

for all $i \in I$. Up to normalization, this system determines the wages w_i and together with (6), (7), (11), and (12) describes the equilibrium.⁵ Notice is that wages depend only on transport costs, the relative country-sizes and the substitution elasticity. Consequently, the trade share of country i

$$e_i = \gamma \left(1 - \frac{n_i p_i c_{ii}}{w_i} \right) = \gamma \left(1 - \frac{L_i}{\sum_k L_k (\tau_{ki} w_k / w_i)^{1-\varepsilon}} \right)$$

is independent of technology. This independence of technological progress and trade shares constitutes a third mismatch of the new trade model regarding the data. As Ventura (2005) points out, time series as well as cross section data of trade and per capita income strongly suggest a positive interconnection of productivity growth and trade shares. The model is unable to deliver such patterns.

As Yi (2003) demonstrates the observed increases in world trade volumes imply an excessive import elasticity in conventional models. Two simple calibration exercises shall show in the following that the new trade model is no exception in that respect. First, a two-country version of the model is fed with data on population size and trade costs. Population data for the US and the ROW are taken from the ERS International Macroeconomic Data Set and the US tariff and the cif/fob measure of transport cost from Feenstra et al (2002). Trade costs are assumed to be the sum trade-weighted tariffs and the cif/fob measure (both from the US data) plus 40%. The latter number is the "border cost" estimated by Anderson and van Wincoop (2004). The elasticity of substitution ε and the parameter γ are used to calibrate the trade shares to the data at the initial and final period of the time interval considered. As Figure 4a shows, this is not possible: for the values $(\varepsilon, \gamma) = (11.9, 1)$ the rise in the trade share is maximal; for higher elasticities, the initial trade share drops below the 4.5% observed in the data. The implied import elasticity⁶ of this model takes an average of 11.5 with the maximum of 11.6 and a minimum of 11.47. At these excessive elasticities, the model is able to explain little less than 70% of the increase in the trade share. As Yi (2003) states, standard theory cannot explain the rise of the trade share without assuming excessive import elasticities and, in this particular case of the new trade model, even then it matches the US import data between 1972 and 2000 only partially.

Second, a symmetric model of 20 equally-sized countries is fed with the trade cost data (the US population equals about 1/20 of the world population in the time

⁵For existence and uniqueness see the appendix. By summing up over i it is quick to check that the system does not have full rank. The missing rank reflects the fact that the trade balance of the last country holds if all other countries have balanced trade.

⁶The import elasticity does not equal ε since at positive trade shares a change of import prices leads to a partial allocation towards domestic goods. This effect tends to reduce the implied import elasticity. This reduction is larger, the larger the expenditure share of the relevant commodity (here: imports) and thus the implied import elasticity drops over time with a rising trade share.

interval considered). Each country has a set of 19 trade partners and faces a vector of bilateral trade cost. This vector of trade cost is assumed to be, up to reordering, identical across countries and is denoted by $\tau = (\tau_1, \tau_2, \dots, \tau_{M-1})$. Without loss of generality, assume that these τ_m are ordered according to ascending values. Finally, the bilateral trade costs are a composite of tariffs t and transportation cost δ and is assumed to take the linear form $\tau_m = 1.4 + t + \delta (m/(M-1))^\sigma$. Each period, the value δ is chosen such that the cif/fob measure implied by the model coincides with measure from the data. Here again, ε and γ are chosen to calibrate the trade shares to the data at the initial and final period of the time interval considered. Figure 4b shows that with $(\varepsilon, \gamma, \sigma) = (14.535, 1, 1.9)$ the multi-country model explains almost 90% of the observed increase in the trade share, yet still at an average import elasticity of 10.16 (minimum and maximum are, respectively, 10.02 and 10.21). This means a considerable improvement of the fit from the two-country to the multi-country model, which is due to the fact that the parameter δ governing the transport cost falls the observed fob/cif ratio in the data. This, in turn is the case since a reduction in transport tariffs t affects the components of the vector of bilateral trade cost uniformly so that consumption of varieties from more distant countries increases relatively more.⁷ One problem with this calibrations is clearly the excessive import elasticity needed to explain the rise in trade shares. Another, maybe less obvious drawback stems from the fact that the measured cif/fob measure implies the strong drop in the parameter δ , which then accounts for about 80% of the rise in trade share. This dominant role of the drop in transport cost contradicts the finding of Baier and Bergstrand (2001) who estimate the effect of decreases in transport cost on bilateral trade volumes to be small relative to that of tariff cuts.

This section has shown that the multi-country version of the Krugman (1980) model mismatches the data in four important dimensions: first, it fails to explain the rise in trade volumes at reasonable elasticities; second, it predicts that all countries import all tradable varieties and consequently fails, third, to account for the rise in source countries per imported good, and fourth, it predicts no response of trade shares to productivity and thus zero correlation between trade shares and per capita income. These shortcomings, however, do not imply that the model utterly misshaped to address those issues. As the following section will show, a reasonable change in the models structure suffices to set it back on track and make substantial progress in all four dimensions.

3 New Trade and Bounded Marginal Utility

This section shall show that a mild change in the above setup goes a long way in alleviating the shortcomings of Krugman's (1980) workhorse model, which the

⁷Check that by using (7) to derive $c_{ik} = c_{ij} (\tau_{ki}/\tau_{ji})^{-\varepsilon}$ under $w_i = \bar{w}$, then take derivatives and verify $d \ln(c_{ik})/dt = d \ln(c_{ij})/dt - \varepsilon(1/\tau_{ki} - 1/\tau_{ji})$ so that $d \ln(c_{ik})/dt > d \ln(c_{ij})/dt \Leftrightarrow \tau_{ki} > \tau_{ji}$.

previous section has pointed to. The change in the setup concerns the demand structure, and will be introduced next.

Demand. As in the previous section, consumers derive utility from the two final goods D and C according to (1). Within the composite of tradables consumers love variety but, in contrast to the previous section, the marginal utility derived from each of the varieties is supposed to be bounded. This latter feature constitutes a major departure from standard literature and is central to the results below. Preference structures with these characteristics have been used to study topics in trade literature and it will be convenient to follow an important precursor and assume with Young (1991) that the composite of varieties takes the form

$$C_i = \int_J \ln(c_{ij} + 1) dj \quad (14)$$

where here again c_{ij} are the consumed quantities of variety j . Writing again q_{ij} as the price of variety j in country i the demand curve of an individual located in market i is derived by maximizing (14) subject to the sub-budget constraint (3). Individual demand is then

$$c_{ij} = \max\{1/(\lambda_i q_{ij}) - 1, 0\} \quad (15)$$

where λ_i stands for the shadow price on the budget constraint.

Notice that for either two varieties consumed in positive quantities, the optimal consumption requires further

$$\frac{1 + c_{ij}}{1 + c_{il}} = \frac{q_{il}}{q_{ij}} \quad (16)$$

Supply. Production takes place as specified in the previous section. The non-tradable D is produced competitively at constant returns to scale (4) while varieties are produced with technology (5). Firms produce one variety each and engage in monopolistic competition in each market.

Prices. Here again, prices trivially equal the wages in the competitive non-tradable sector. In the tradable sector, firms can price-discriminate across different countries by assumption and consequently a firm j sets the price p_{ij} for market i in order to maximize its operating profits π_{ij} in that specific market. When firm j is located in country k and gross iceberg-type trade costs between country k and country i are $\tau_{ki} \geq 1$, the consumer price in country i of variety j is the $q_{ij} = \tau_{ki} p_{ij}$. Total profits of firm j in country i are $\pi_{ij} = L_i \tau_{ki} c_{ji} (p_{ij} - w_k \beta)$ where L_i is the number of individuals in the market i and w_k is the wage in country k , the location of production, so that $w_k \beta$ is the marginal cost of production. With (15), firm j 's profits are thus

$$\pi_{ji} = L_i \tau_{ki} (1/(\lambda_i \tau_{ki} p_{ji}) - 1) (p_{ji} - w_k \beta) \quad (17)$$

Generic profit maximization thus implies

$$p_{ji} = \sqrt{w_k \beta / (\lambda_i \tau_{ki})} \quad (18)$$

for each market i separately. Here again, within each country firms are all identical and charge the same prices in a given market so that prices can be indexed by country of origin and country of destination.

Under free entry to production, total operating profits, i.e. the sum of a firm's profits in each market, just cover the setup cost.

3.1 Closed Economy ($I = 1$)

Assume that international transport costs are too high to allow any kind of trans-border trade. Countries are in autarky and can be analyzed by looking at one representative closed economy, within which transport costs are negligible ($\tau_{ii} = 1$). Dropping country indices, firm j in the economy then sets its price according to (18). As in equilibrium all firms are identical, the shadow price on the sub-budget constraint satisfies $1/\lambda = p_j(c_j + 1) \equiv p(c + 1)$ and the monopolist prices are

$$p = (c + 1)w\beta \quad (19)$$

Note that the monopolist markup over marginal cost is $c + 1$ and the demand elasticity from (15) equals $\varepsilon = (c + 1)/c$. The specific demand structure assumed in (14) thus implies that the own-price elasticity of varieties depend on quantities consumed, is larger than unity and, in particular, unbounded. Consequently, in the context of the present paper, whose motivation partly rests on the "elasticity-puzzle," i.e. on the discrepancy between measured import elasticities and the observed rise in trade volumes, a calibration of the model is inevitable to assess the implied import elasticities at observed trade shares.

With the prices (19), the operating profit of the representative firm is

$$\pi = Lc(w\beta(c + 1) - w\beta) = w\beta Lc^2 \quad (20)$$

and the free entry condition $\pi = w\alpha$ determines the quantities consumed per individual and variety

$$c = \sqrt{\frac{\alpha}{\beta L}} \quad (21)$$

Equation (21) shows that, quite intuitively, an increase in setup over marginal costs (α/β) increases per capita consumption of the average variety. At higher values of α/β , varieties are relatively more costly to produce and less of them are produced so that the per capita consumption of each rises. Conversely, at increasing population size (L) there is more demand for varieties, which increases the number of active firms and reduces per capita consumption of each single variety.

The equilibrium number of active firms, n , is finally pinned down by combining the conditions of labor market clearing ($n[\beta x + \alpha] = \gamma L$) and goods market clearing ($x = Lc$) within the tradable sector and is calculated to be

$$n = \frac{\gamma L}{\alpha + \sqrt{\alpha\beta L}} \quad (22)$$

Observe that the number of active firms is determined by technology parameters, the population and market size. The difference between the two latter parameters is noteworthy. Market size, defined as total expenditure on varieties (γL), enters the number of firms linearly, i.e. an increase in market size induces a proportional rise in the number of active firms. Population size, however, enters expression (22) separately and, holding the market size fix, tends to reduce the number of firms. For an intuition of that finding remember that demand elasticity decreases with per capita consumption. Now compare two economies, the first with double labor force but half the parameter γ of the second. Both economies spent the same total amount of resources to produce varieties, yet per capita expenditure of the first economy is half of the second. At an equal number of firms, this implies that per capita consumption of each variety in the first economy are half of those in the second. Consequently, demand elasticities are larger in the first economy, depressing monopolist markups and profits so that finally firms in the first economy make less profits than in the second. But by free entry firms make zero total profits so that the number of firms must adjust and is in equilibrium smaller in the first economy. Per capita demand is a key determinant for the response of the number of firm to the market size. This characteristic is very much unlike the one in the classical new trade model where, due to homothetic preferences, exclusively the market size matters (not the per capita expenditure on varieties; compare (11)).

In the present model, the overall effect of an increase in population L on the number of firms is still positive, but, due to the demand effects discussed, less than proportional. Consequently, an increase in population size induces a less than proportional drop of per capita consumption of the average variety.

3.2 Two Countries ($I = 2$)

Assume now that trade costs between two countries fall to a level where agents engage in international trade. To save notation, the variables of, say, the second of the countries will be marked by an asterisk. Individuals in one country can purchase varieties produced outside of their country when incurring an iceberg-type transport cost. In other words, for every unit of a variety to arrive at the destination, $\tau > 1$ units of it have to leave the producer's country.

Firms maximize profits in each country separately, i.e. they can by assumption price discrimination across countries. Thus, the prices that firms charge in the domestic and foreign markets need to be distinguished. Let these prices be (p_d, p_f)

for firms located in country 1 and (p_d^*, p_f^*) for firms located in country 2 and denote with (c_d, c_f) and (c_d^*, c_f^*) the quantities consumed in either country. Just as in the previous section, a firm's optimal price in the domestic market is

$$p_d = (c_d + 1)w\beta \quad \text{and} \quad p_d^* = (c_d^* + 1)w^*\beta^* \quad (23)$$

With these prices, firms make the profits

$$\pi_d = Lw\beta c_d^2 \quad \text{and} \quad \pi_d^* = L^*w^*\beta^*(c_d^*)^2 \quad (24)$$

in their respective domestic market. But in addition to activities in the domestic market, firms sell in the export market. A firm that is located in one country maximizes profits in the foreign market (see (17)) by setting its export prices according to (18). As the shadow price on country 2's budget constraint satisfies $\lambda^* = p_d^*/(c_d^* + 1)$ and with (23) this leads to

$$p_f = \sqrt{\frac{w}{\tau w^*}} p_d^* \quad \text{and} \quad p_f^* = \sqrt{\frac{w^*}{\tau w}} p_d \quad (25)$$

for the export prices. Using now the optimality condition (16), the monopolist profits (17) the foreign market can be written as⁸

$$\pi_f = L^*\tau(c_f^*)^2 w\beta \quad \text{and} \quad \pi_f^* = L\tau c_f^2 w^*\beta^* \quad (26)$$

The free entry conditions in both countries require then $\pi_d + \pi_f = w\alpha$ in the first and $\pi_d^* + \pi_f^* = w^*\alpha$ in the second country.

3.2.1 Two Symmetric Countries

It will prove instructive to look at the model in under symmetry and to consider trade relations between two equally sized countries ($L = L^*$). In this simple case wages equalize and monopolist prices in the domestic and foreign markets are

$$p_d = w\beta(c_d + 1) \quad \text{and} \quad p_f = \frac{p_d}{\sqrt{\tau}} \quad (27)$$

where the asterisk can be dropped because of symmetry. Optimal consumption choice (16) then implies

$$c_d + 1 = \sqrt{\tau}(c_f + 1) \quad (28)$$

⁸Firms make less profits in foreign than in domestic markets for two reasons. First, and obviously, demand in the foreign market is lower since consumer prices, bearing additional transport cost, are higher. But second, the lower a demand induces a higher own-price elasticity and therefore depresses markups and profits in foreign markets. See also Hummels and Lugovsky (2005) for a microfoundation and empirical evidence on this point.

whenever trade volumes are positive ($c_f > 0$). The condition for trade to be positive is clearly $c_d > \sqrt{\tau} - 1$, or with (21)

$$\sqrt{\alpha/(\beta L)} > \sqrt{\tau} - 1 \quad (29)$$

In contrast to standard model, finite tariffs can, if they are high enough, make bilateral trade volumes fall to zero. This is obviously a direct implication of the bounded marginal utility: when the ratios of marginal utilities exceeds the relative price, the more expensive one drops out of the consumption bundle. While this cannot occur when marginal utilities approach infinity as quantities tend to zero it is well possible in the present setting.

Not surprisingly, higher tariffs tend to prevent trade flows between countries and may impede trade altogether. At the same time, a larger population L tends to impede trade since it increases the number of domestic firms, reduces per capita consumption of each variety and thus depresses the consumer's willingness to incur the transport costs to purchase an additional, foreign variety. On the other hand, an increase in the ratio of setup cost and marginal cost (α/β) decreases the number of local firms, increases per capita consumption of each and thus makes the consumer willing to pay for more for foreign varieties.

To solve now for the symmetric equilibrium, use condition (28) to write the operating profits from the export market (26) as

$$\pi_f = L(c_d + 1 - \sqrt{\tau})^2 w \beta \quad (30)$$

Free entry in production requires that global operating profits just cover the setup cost

$$\pi = \pi_d + \pi_f = Lw\beta (c_d^2 + (c_d + 1 - \sqrt{\tau})^2) = w\alpha$$

which leads to

$$c_d = \frac{\sqrt{\tau} - 1}{2} + \sqrt{\frac{\alpha}{2\beta L} - \left(\frac{\sqrt{\tau} - 1}{2}\right)^2}$$

Each individual consumes c_d of a domestically produced variety, this quantity increases in the ratio of setup cost and marginal productivity (α/β).

The resource constraint in the tradable sector ($n[\beta x + \alpha] = \gamma L$) and goods market clearing ($x = L(c_d + \tau c_f)$) determine the number of active firms in either of the two countries

$$n = \frac{\gamma L}{\alpha + \beta L \left((\sqrt{\tau} + 1) \sqrt{\frac{\alpha}{\beta 2L} - \left(\frac{\sqrt{\tau} - 1}{2}\right)^2} - \frac{(\sqrt{\tau} - 1)^2}{2} \right)} \quad (31)$$

This expression can be shown to be increasing in the trade cost τ . The positive relation between the level of trade cost and the number of active firms per country reflects the finding discussed in connection with equation (22). In a closed economy,

the doubling of the labor force induced a less than proportional increase in the total number of firms. Now, a doubling of the labor force in a closed economy is equivalent to a move of a given country from autarky to full trade integration with an identical country. Thus, in such a situation the number of firms *per country* must fall. Finally, when the move from autarky to trade integration takes place via gradual trade cost reduction, the above equation shows that the drop in active firms per country is in fact monotonous.

With per capita consumption and the number of firms established, one can now turn to the volume of trade. The import share, defined as nominal imports over nominal GDP, is calculated with equations (27) and (28) to be

$$e = \gamma \frac{nLp_f\tau c_f}{nLp_d c_d + nLp_f\tau c_f} = \gamma \frac{\sqrt{\frac{2\alpha}{\beta L(\sqrt{\tau}-1)^2} - 1} - 1}{2\sqrt{\frac{2\alpha}{\beta L(\sqrt{\tau}-1)^2} - 1}} \quad (32)$$

in the two countries. This expression is decreasing in the country size (L) and increasing in the ratio of setup cost and marginal productivity (α/β). For an intuition of this result note that a smaller population L and a larger ratio α/β tend to increase per capita consumption of the domestic variety. Further remember that, as discussed in connection with equation (29), demand for foreign varieties falls to zero when domestic per capita consumption is too high. Thus, starting from very low trade volumes, any parameter change that increases per capita consumption of domestic varieties must induce a more than proportional increase in per capita consumption of a foreign variety. In fact, the choice of the demand structure (14) implies that increases in per capita consumption of foreign varieties are over-proportionate along the whole range of parameters. (Take derivatives of equation (28) to see that $d\ln(c_d) < d\ln(c_f)$). As prices stay constant in the symmetric equilibrium, this means that the expenditure share on foreign varieties grows.

Notice in particular that improvements of marginal productivity lead to increases of the trade share.

This symmetric model can be brought to the data. The population of the US and the rest of the world (ROW) are supposed to be of equal size. Both labor forces evolve according to US population growth. US per capita income is taken as a proxy for the technological change, which is moreover assumed to affect marginal productivity only (i.e. US per capita income $\sim 1/\beta$). Trade costs are assumed to be the sum of tariffs and the cif/fob values plus 40%. The initial value of β and the trade share γ are jointly used to calibrate the import shares to the data at the initial and final period of the period. With $(\gamma, \beta) = (0.49, 15)$ the fit of the trade share is satisfactory. The implied import elasticity takes an average value of 6.8 with a maximum of 14.8 at the start of the period and a minimum of 3.9 at the end of it. This can be read as an improvement of the fit compared to the standard

New Trade Model. The average implied import elasticity, however, is by far larger than the target of 2 to 3 Yi (2003) argues to be reasonable. The next paragraphs shall explore to what extent the imposed symmetry can be blamed for this failure.

3.2.2 Two Asymmetric Countries

In the general case of two asymmetric countries the monopolists' prices in the domestic and foreign markets are given, respectively, by (23) and (25). The full set of prices is then

$$\begin{aligned} p_d &= w\beta(c_d + 1) & p_f &= \sqrt{\frac{w}{\tau w^*}} p_d^* \\ p_d^* &= w^*\beta(c_d^* + 1) & p_f^* &= \sqrt{\frac{w^*}{\tau w}} p_d \end{aligned} \quad (33)$$

Provided that trade volumes are positive, optimal consumer choice (16) then implies

$$c_d + 1 = \sqrt{\tau w^*/w}(c_f + 1) \quad \text{and} \quad c_d^* + 1 = \sqrt{\tau w/w^*}(c_f^* + 1) \quad (34)$$

The condition for trade volumes to be positive is then $c_f, c_f^* > 0$ or, with (21) and (34)

$$(\alpha/(\beta L) + 1)(\alpha/(\beta L^*) + 1) > \tau \quad (35)$$

Finally, the monopolists' operating profits in the respective markets are as in (24) and (26)

$$\begin{aligned} \pi_d &= w\beta L c_d^2 & \pi_f &= w\beta L^* \tau (c_f^*)^2 \\ \pi_d^* &= w^*\beta L^* (c_d^*)^2 & \pi_f^* &= w^*\beta L \tau c_f^2 \end{aligned} \quad (36)$$

so that the free entry conditions become

$$L c_d^2 + L^* \tau (c_f^*)^2 = \alpha/\beta \quad \text{and} \quad L^* (c_d^*)^2 + L \tau c_f^2 = \alpha/\beta \quad (37)$$

The resource constraints in both countries require

$$n [(L c_d + L^* \tau c_f^*) \beta + \alpha] = L \quad \text{and} \quad n^* [(L^* c_d^* + L \tau c_f) \beta + \alpha] = L^* \quad (38)$$

Finally, making use of the monopolists' prices (33) and imposing the trade balance $n L^* \tau p_f c_f^* = n^* L \tau p_f^* c_f$ leads to

$$n L^* (c_d^* + 1) c_f^* = n^* L (c_d + 1) c_f \quad (39)$$

Now, equations (38) and (39) can be combined to eliminate the number of firms n and n^*

$$\frac{(c_d^* + 1) c_f^*}{(c_d + 1) c_f} = \frac{L c_d + L^* \tau c_f^* + \alpha/\beta}{L^* c_d^* + L \tau c_f + \alpha/\beta} \quad (40)$$

while eliminating relative wages in (34) leads to

$$(c_d + 1)(c_d^* + 1) = \tau (c_f + 1)(c_f^* + 1) \quad (41)$$

The four equations (37), (40), and (41) jointly determine the quantities c_d , c_f , c_d^* , and c_f^* . This is a nonlinear system which has no analytic solution and needs to be solved numerically. One can, however, show the following

Proposition 1 *The system (37), (40), and (41) has a unique solution (c_d, c_f, c_d^*, c_f^*) .*

Proof. See appendix. ■

Given these quantities c_d, c_f, c_d^* , and c_f^* , the terms of trade are determined by (34)

$$\frac{w}{w^*} = \frac{c_f + 1}{c_d + 1} \frac{c_d^* + 1}{c_f^* + 1}$$

and equations (38) deliver the number of firms active in either country, which finally determines the trade share of the first country

$$e = \gamma \frac{n^* p_f^* \tau c_f}{n p_d c_d + n^* p_f^* \tau c_f} = \gamma \frac{n^* \sqrt{\tau w^* / w c_f}}{n c_d + n^* \sqrt{\tau w^* / w c_f}}$$

This asymmetric model can be brought to the data. Assume there are two countries with relative labor forces corresponding to those of the US and ROW. Technology ($1/\beta$) is proxied by real US per capita income and trade cost is assumed to be the sum of US tariffs and the cif/fob values plus 40%. The initial value of β and the trade share γ are jointly used to calibrate the import shares to the data at the initial and final period of the period.

3.3 A Symmetric Multicountry World ($I > 2$)

Take now an economy with I countries, which are all identical in terms of labor force. Pairs of countries differ regarding the respective transportation cost they face when engaging in bilateral trade, but to reap the virtues of symmetry assume finally that countries are symmetric in terms of potential trade partners. In particular, the vector of gross iceberg trade cost $\tau_i = (\tau_{i1}, \tau_{i2}, \dots, \tau_{iI})$ is, up to reordering, identical across countries i (i.e. for all $i, k \in I$ there is a permutation $\xi : I \rightarrow I$ such that $\xi(\tau_i) = \tau_k$). As all parameters that govern demand and supply are identical across countries, producer prices and wages equalize throughout countries.

The bounded marginal utility from varieties implies, just like in the two-country model, there is a threshold on the transportation cost $\bar{\tau}$ above which there is no bilateral trade. The defining condition for this threshold is derived from demand (15) and satisfies

$$c_{ii} = \sqrt{\bar{\tau}} - 1 \tag{42}$$

Obviously, the set of country 1's trade partners consists of those countries with whom bilateral trade is least expensive.

Consider now, say, county 1 and denote by M the set of countries it engages in bilateral trade with (M includes the country 1 itself). With monopolist prices parallel to (27), the optimal consumer choice requires

$$c_{11} + 1 = \sqrt{\tau_{1k}} (c_{1k} + 1) \quad k \in M \tag{43}$$

The profit a firm j located in country 1 makes in the market $k \in M$ is parallel to (30) and the free entry condition now becomes

$$\pi_1 = \sum_{k \in M} \pi_{1k} = \sum_{k \in M} Lw\beta (c_{11} - (\sqrt{\tau_{1k}} - 1))^2 = w\alpha \quad (44)$$

Writing the shorthand

$$T_{M,1} = \sum_{k \in M} (\sqrt{\tau_{1k}} - 1) \quad \text{and} \quad T_{M,2} = \sum_{k \in M} (\sqrt{\tau_{1k}} - 1)^2$$

condition (44) leads to an expression for per capita consumption of the domestic varieties:

$$c_{11} = \frac{T_{M,1}}{M} + \sqrt{\frac{\alpha}{\beta LM} + \left(\frac{T_{M,1}}{M}\right)^2} - \frac{T_{M,2}}{M} \quad (45)$$

Here, in slight abuse of notation, M stands for the set of trade partners and for the number of trade partners at the same time. Note that M is an endogenous variable that eventually depends on the schedule of bilateral trade costs. Formally, one has a problem of circular definition, i.e. M is defined as

$$M = \{k \in I \mid \sqrt{\tau_{1k}} < c_{11} + 1\} \quad (46)$$

while c_{11} depends on M itself. One can show, however, the following

Proposition 2 *The set M is nonempty and uniquely defined by (46).*

Proof. Assume wlog that the elements of the vector τ_1 are ordered according to ascending size. Then note that $\tau_{1l} \leq \tau_{1k}$ implies $k \in M \Rightarrow l \in M$. Consequently, any solution to (46) must be of the form $\{1, 2, \dots, n\}$ for one $n \in I$. Finally, define c_n by (44) under $M = \{1, 2, \dots, n\}$. Observe that c_n is decreasing in n , so that the sequence defined by $m_n = \max\{k \in I \mid \sqrt{\tau_{1k}} < c_n + 1\}$ is decreasing in n . By construction the set $\{1, \dots, n\}$ with n satisfying $m_n \geq n$ and $m_{n+1} \leq n + 1$ solves (46). As m_n is a monotonically falling sequence, there is maximal one such n and since $\{1\} \in M$, the set M is nonempty. ■

With the equilibrium per capita quantities of domestic varieties well defined, equations (43), (44), and the free entry condition ($n [\alpha + \beta L (\sum_{k \in M} \tau_{1k} c_{1k})] = \gamma L$) determined the number of active firms

$$n = \frac{\gamma L}{\alpha + \beta L \left[(T_{M,1} + M) \sqrt{\frac{\alpha}{\beta LM} + \left(\frac{T_{M,1}}{M}\right)^2} - \frac{T_{M,2}}{M} + \frac{(T_{M,1})^2}{M} - T_{M,2} \right]} \quad (47)$$

Finally, using again consumers' optimality condition (43) and (45), the trade share of country 1 can be expressed as

$$e = 1 - \frac{np_{11}c_{11}}{np_{11} \sum_{k \in M} \sqrt{\tau_{1k}}c_{1k}} = 1 - \frac{1}{M} \left(1 + \frac{T_{M,1}}{\sqrt{\frac{\alpha M}{\beta L} + (T_{M,1})^2} - MT_{M,2}} \right) \quad (48)$$

Equations (46) and (48) determine the representative country's number of trade partners and its trade share, the two key trade parameters which the present model aims to explain.

The set of a country's trade partners M depends on the model's parameters and may be any subset of the full set of countries I satisfying $\{1\} \subset M$. Just as in the two-country world trade costs can, if they are too high, impede bilateral trade. Intuition in the two-country model suggested that the inclination to pay for foreign varieties increases with the per capita consumption of the domestic varieties, as decreasing marginal utility from domestic varieties raised consumer's willingness to pay for the more expensive foreign ones. As population size L decreased and the ratio α/β increased per capita consumption of a domestic variety one can conjecture that these parameters drive the contraction or expansion of the set of trade partners.

The second key variable, the trade share (48) equally is expected to rise in the ratio $\alpha/(\beta L)$ not only since per capita consumption of foreign varieties increase relatively more than domestic varieties, but also since the set of trade partners expands. The dimension of trade partners - or source countries - constitutes an extra margin along which to expand trade volumes and amplifies the first effect.

The above considerations regarding the number of trade partners and the trade share prove right and the impact of population size and technology on these variables is summarizes in the following

Proposition 3 *The trade share e and the number of trade partners M increases in $\alpha/(\beta L)$.*

Proof. Note first that for any change in the set M condition (42) must be satisfied for the newcomer (dropout). Thus, the trade share as $e = 1 - \frac{c_{11}}{Mc_{11} - T_{M,1}}$ is continuous at any change in the set M . By this observation and (45) it is sufficient to prove that the equilibrium c_{11} is increasing in $\alpha/(\beta L)$. By (45) this is trivially the case whenever M is constant.

For any change in the set M consider wlog an increase in M . Define m as the index of the lowest bilateral iceberg trade cost outside the set of trade partners, i.e. $m = \operatorname{argmin}_{k \in I \setminus M} \{\tau_{1k}\}$. Whenever M increases condition $\sqrt{\tau_{1m}} = c_{11} + 1$ must hold and equation (44) is satisfied for both of the two sets M and $M \cup \{m\}$, implying that c_{11} is continuous. As the set of changes in M is countable, this proves that c_{11} is increasing in $\alpha/(\beta L)$. ■

Proposition 3 shows that the rise in trade share and the expansion of the set of trade partners are jointly driven by productivity growth at the margin. When the effect of this joint determinant is strong, it can induce the strong positive correlation between the rises in trade volumes and the number of source countries per good, which the data exhibit (see Figures 3 and 4). Moreover, it implies the common dynamics between per capita income and trade volumes that Ventura (2005) point at. Remember that the standard model of the New Trade Theory failed to address any of these patterns of the trade data, so that the results presented in Proposition 3 constitute a considerable advance towards reconciling the theory with the data. Much mileage is made by assuming that consumers value variety, but derive only bounded marginal utility from them at zero consumption levels. Up to this point, however, the results remain purely qualitative and to complete the picture it is necessary to know how the model performs quantitatively, i.e. to evaluate whether it can jointly address the level and the rise in trade shares, the number of source countries per good and per capita income at reasonable the import elasticity.

To this aim the model will be calibrated to the US trade shares. An important parameter in the calibration is the import elasticity to tariffs, which is defined as

$$\mathcal{E} = -(1+t) \frac{d}{dt} \ln(Q)$$

where Q stands for the total imported quantity. With the optimal consumer choice (43), the monopolist prices (33), and the number of firms (47) this is

$$Q = n \sum_{M \setminus \{1\}} c_{1k} = \left(\frac{(c_{11} + 1) \sum_{M \setminus \{1\}} \frac{1}{\sqrt{\tau_{1k}}} - (M - 1)}{\alpha + \beta L [c_{11}(T_{M,1} + M) - T_{M,2} - T_{M,1}]} \right)$$

Now, how does this model square the data? To evaluate its performance a parallel exercise to that of the standard New Trade Model in section 2 is performed. The symmetric model of 20 identical economies is fed with population and trade cost data of the US between 1972 and 2000 (the US population equals about 1/20 of the world population in the time interval considered). Each country has a set of 19 potential trade partners and faces a vector of bilateral trade cost, which is assumed to be, up to reordering, identical across countries and is denoted by $\tau = (\tau_1, \tau_2, \dots, \tau_{M-1})$. The τ_m are ordered according to ascending values. Finally, the bilateral trade costs are a composite of tariffs t and transportation cost δ and is assumed to take the linear form $\tau_m = 1.4 + t + \delta (m/(M - 1))^\sigma$. Each period, the value δ is chosen such that the cif/fob measure implied by the model coincides with measure from the data. The parameter γ and the initial marginal productivity β_{1972} are chosen to calibrate the trade shares to the data at the initial and final period of the time interval considered. Figure 5b shows that with $(\gamma, \beta_{1972}, \sigma) = (0.15, 14.39, 1.9)$ the multi-country model follows the time series satisfactorily with about the same success as the two-country model. A big improvement, however, concerns the implied import elasticity, which

now assumes a time-average of 3.7 with a peak of 8.8 at the start of the period and a minimum of 1.75 at the end of it. While the variance is high, the average import elasticity is close to the interval [2,3], which Yi (2003) puts forward as a realistic range.

The reason for this substantial improvement is the extension of trade along the dimension of source countries. Driven by improvements of marginal productivity ($1/\beta$), consumers purchase not only larger quantities from previous trade partners (as in the two-country model) but open new trade relations and extend the number of source countries. Figure 5b illustrates these dynamics and compares the predictions of the model to that of the data (values in 1972 are normalized). This additional margin along which trade expands amplifies the impact of technological change, which therefore explains a larger part of the increase in trade shares. The remaining part to be accounted for by the reductions in trade cost therefore shrinks, and the implied import elasticity falls.

As in the standard New Trade Model with 20 countries, the measured cif/fob values and the implied "real" transport cost δ differ considerably. The reason for this is now twofold. First, as in the standard model, an exogenous drop in δ induces consumers to increase their purchases of foreign varieties relatively more. Second, and in addition, the model with bounded marginal utility predicts that increases in marginal productivity induce consumers to buy from more distant locations. Both of these effects tend to increase the expenditure on transportation, and imply an upward bias of the conventional cif/fob measure of trade cost from the 'real' values so that the implied latter ones fall by more than the data suggest.⁹ Despite this second amplifying effect that stems from the technological change, the fraction of the rise in trade shares explained in by the drop in transportation is now 44%, which is still larger than what the estimations of Baier and Bergstrand (2001) imply (around one fourth, about a third of the impact that tariffs have), yet much below the 80% the standard New Trade Model of the previous section implied.

In sum, the model can explain two strong and important trends of trade data with reasonable parameter values: first, the massive growth of trade shares at a modest fall of trade costs and second the small fraction of trade relations that the average country engages in. It suggests technological progress as the core determinant of both variables and thereby generates two additional patterns of international trade data: first, the correlation between trade shares and per capita income and second, the correlation between the increases of trade volumes and the number of source countries.

Two clarifying remarks are in order. The first one concerns the nature of pro-

⁹Given the modeled effects are strong, conventional measures tend to understate the drop in trade costs substantially. These systematic measurement problems may be worth to analyze in detail in empirical work. In particular, an evaluation of their role in estimates like Baier and Bergstrand (1999) seems appropriate. But such an analysis is clearly outside the scope of the present paper.

ductivity growth, which was assumed to affect manufacturing efficiency only and leave entry cost unchanged. The model's direct implication that stronger increasing returns (i.e. a higher ratio of setup cost over marginal cost) imply larger trade volumes is confirmed empirically in a cross section analysis by Harrigan (1994). A time-series analysis, however, that justifies this central assumption and could lend support to the hypothesis of a strong causal effect of productivity growth on trade volumes is still to be performed. The second remark concerns previous attempts to evaluate the role of non-homothetic preferences on trade volumes. These effects are generally estimated to be small (see Bergstrand (1990) and Bergoeing and Kehoe (2003) among others). However, these quantitative studies consistently based on Stone-Geary type preferences between a homogeneous good and a composite of varieties. Consequently, they suffer from the drawback of the Krugman (1980) model in a slightly modified version: any country with positive trade share imports all tradable varieties from all other trading countries. Thus, these models rule out the existence of a non-trivial fraction of imported varieties over world tradables and thereby an endogenous expansion of the set of trade partners. The difference in the performance of the two-country and the multi-country model made clear, however, that this expansion of the set of trade partners bears a good part of the calibration's success. This qualitative difference may explain the strong discrepancy between the findings in those previous and the present paper. It finally suggests that future empirical studies should include the effect that stems from an expansion of the set of trade partners.

3.4 An Asymmetric Multicountry World

This final generalization of the model considers the multi-country model under asymmetric country size and trade costs. Denote country i 's vector of gross iceberg trade cost with $\tau_i = (\tau_{i1}, \tau_{i2}, \dots, \tau_{iI})$. Remember the monopolist prices in an asymmetric model (33), which implied under individual optimization that

$$\frac{c_{ij} + 1}{c_{ik} + 1} = \sqrt{\frac{\tau_{ji}w_j}{\tau_{ki}w_k}} \quad (49)$$

whenever c_{ij} and c_{ik} are positive. Note that consumption of domestic varieties is not necessarily positive, as relative prices can vary dramatically with wages differentials, which, in turn, are determined by relative country sizes, trade costs and technology.

Further, the set of countries country i imports from, M_i , need not in general coincide with E_i , the set of countries it exports to. To define both sets, write first

$$i_o = \operatorname{argmin}_{k \in I} \{ \sqrt{\tau_{ki}w_k} \}$$

so that varieties from i_o are the cheapest in country i . Then write

$$M_i = \left\{ k \in I \mid c_{i_o i} + 1 > \sqrt{\tau_{ki}w_k / (\tau_{i_o i}w_{i_o})} \right\} \quad (50)$$

and

$$E_i = \left\{ k \in I \mid c_{k_o k} + 1 > \sqrt{\tau_{ik} w_i / (\tau_{k_o k} w_{k_o})} \right\} \quad (51)$$

With the monopolist prices $p_{ik} = (c_{ki} + 1)\beta w_i$ of firms in country i selling in market k , generic profits from market k are still

$$\pi_{ik} = L_k \tau_{ik} \beta w_i c_{ki}^2$$

so that the free entry condition in country i remains

$$\sum_{k \in E_i} L_k \tau_{ik} c_{ki}^2 = \alpha / \beta \quad (52)$$

The resource constrain in country i requires

$$n_i \left[\alpha + \beta \sum_{k \in E_i} L_k \tau_{ik} c_{ki} \right] = L_i \quad (53)$$

and finally the trade balance is $n_i \sum_{k \in E_i} L_k \tau_{ik} p_{ik} c_{ki} = L_i \sum_{k \in M_i} n_k \tau_{ki} p_{ki} c_{ik}$ or

$$\frac{\sum_{k \in E_i} L_k \tau_{ik} (c_{ki} + 1) c_{ki}}{\alpha / \beta + \sum_{k \in E_i} L_k \tau_{ik} c_{ki}} w_i = \sum_{k \in M_i} \frac{L_k \tau_{ki} (c_{ik} + 1) w_k c_{ik}}{\alpha / \beta + \sum_{m \in E_k} L_m \tau_{km} c_{mk}}$$

With (52) this leads to

$$w_i = \sum_{k \in M_i} \frac{L_k \tau_{ki} (c_{ik} + 1) w_k c_{ik}}{\sum_{m \in E_k} L_m \tau_{km} (c_{mk} + 1) c_{mk}} \quad (54)$$

The system defined by (49) - (54) jointly determines the wages $\{w_i\}_{i \in I}$, the consumed quantities $\{c_{ik}\}_{i, k \in I}$ the number of active firms in each country $\{n_i\}_{i \in I}$ and the sets of supplied and supplying countries $\{E_i\}_{i \in I}$ and $\{M_i\}_{i \in I}$.

NOTE: This section needs to be concluded.

4 Conclusion

This paper has shown that a small and realistic twist in the demand structure goes a long way in explaining four strong and important empirical regularities that the workhorse model of the New Trade Theory fails to explain. These empirical patterns are the massive rise in trade shares, the small fraction of traded varieties countries import, the correlation between trade growth and per capita growth, and the parallel rise in the number of source countries per imported good and trade volumes, and finally the correlation between trade growth and per capita growth. By the paper's

key assumption marginal utility derived from each variety is bounded. This implies that varieties whose transport is very costly drop out of the individual's consumption basket. Technological progress that induces a higher per capita consumption of those varieties already consumed decreases marginal utility derived from them and makes consumers expand their consumption basket towards more expensive, foreign varieties. Therefore, such technological progress drives up the number of source countries per good, the trade share, and per capita income.

A Appendix

Existence of a solution to system (13). Notice first that the system

$$\sum_k \frac{L_k w_k \tau_{ik}^{1-\varepsilon} w_i^{-\varepsilon}}{\sum_m L_m (\tau_{mk} w_m)^{1-\varepsilon}} = 1 \quad (\text{A1})$$

is homogeneous of degree zero in the wages and a potential solution can be normalized by setting $w_1 = 1$.

Now confirm that each term on the LHS of is decreasing in w_i and increasing in w_k $k \neq i$ and that $\text{LHS} \rightarrow 0$ ($w_i \rightarrow \infty$) and $\text{LHS} \rightarrow \infty$ ($w_i \rightarrow 0$) and hence for every vector $(w_2, \dots, w_{i-1}, w_{i+1}, \dots, w_I) \in \mathbb{R}_+^{I-1}$ there is a unique solution

$$w_i = \varphi_i(w_2, \dots, w_{i-1}, w_{i+1}, \dots, w_I)$$

satisfying (A1). Moreover, the function $\varphi_i : \mathbb{R}_+^{I-2} \rightarrow \mathbb{R}_+$ is strictly increasing in all arguments.

Now write $\bar{\tau} = \max_{ij} \{\tau_{ij}\}$, $\bar{w} = \max_i \{w_i\}$, $\hat{w} = \min_i \{w_i\}$ and $\lambda_i = L_i / \sum_k L_k$ and check that

$$w_i^\varepsilon \leq \bar{\tau}^{\varepsilon-1} \frac{\sum_k L_k w_k}{\sum_k L_k w_k^{1-\varepsilon}} \leq \bar{\tau}^{\varepsilon-1} \frac{\sum_k L_k w_k}{L_1} \leq \bar{\tau}^{\varepsilon-1} \frac{\bar{w}}{\lambda_1}$$

and

$$w_i^\varepsilon \geq \bar{\tau}^{1-\varepsilon} \frac{\sum_k \lambda_k w_k}{\sum_k \lambda_k w_k^{1-\varepsilon}} \geq \bar{\tau}^{1-\varepsilon} \lambda_1 \hat{w}^{\varepsilon-1}$$

or

$$\varphi_i(w_2, \dots, w_{i-1}, w_{i+1}, \dots, w_I) \leq \bar{\tau}^{(\varepsilon-1)/\varepsilon} (\bar{w}/\lambda_1)^{1/\varepsilon}$$

and

$$\varphi_i(w_2, \dots, w_{i-1}, w_{i+1}, \dots, w_I) \geq \bar{\tau}^{(1-\varepsilon)/\varepsilon} (\lambda_1 \hat{w}^{\varepsilon-1})^{1/\varepsilon}$$

This implies

$$\varphi_i : \left[\lambda_1 / \bar{\tau}^{\varepsilon-1}, \bar{\tau} / \lambda_1^{1/(\varepsilon-1)} \right]^{I-2} \rightarrow \left[\lambda_1 / \bar{\tau}^{\varepsilon-1}, \bar{\tau} / \lambda_1^{1/(\varepsilon-1)} \right]$$

Therefore, when P_i is the projection on the i -th component, the function ϕ defined as

$$\phi = \begin{bmatrix} P_2 \\ \vdots \\ P_{I-1} \\ \varphi_I \end{bmatrix} \circ \begin{bmatrix} P_2 \\ \vdots \\ \varphi_{I-1} \\ P_I \end{bmatrix} \circ \dots \circ \begin{bmatrix} \varphi_2 \\ P_3 \\ \vdots \\ P_I \end{bmatrix} \begin{pmatrix} w_2 \\ w_3 \\ \vdots \\ w_I \end{pmatrix}$$

satisfies $\phi : \left[\lambda_1/\bar{\tau}^{\varepsilon-1}, \bar{\tau}/\lambda_1^{1/(\varepsilon-1)} \right]^{I-1} \rightarrow \left[\lambda_1/\bar{\tau}^{\varepsilon-1}, \bar{\tau}/\lambda_1^{1/(\varepsilon-1)} \right]^{I-1}$ and is strictly increasing in all arguments, as the functions φ_i are so. Thus,

$$\phi(\lambda_1/\bar{\tau}^{\varepsilon-1}, \dots, \lambda_1/\bar{\tau}^{\varepsilon-1}) \geq (\lambda_1/\bar{\tau}^{\varepsilon-1}, \dots, \lambda_1/\bar{\tau}^{\varepsilon-1})$$

component by component. Applying ϕ to both sides leads to $\phi^2(\lambda_1/\bar{\tau}^{\varepsilon-1}, \dots) \geq \phi(\lambda_1/\bar{\tau}^{\varepsilon-1}, \dots)$ and by induction to

$$\phi^{m+1}(\lambda_1/\bar{\tau}^{\varepsilon-1}, \dots, \lambda_1/\bar{\tau}^{\varepsilon-1}) \geq \phi^m(\lambda_1/\bar{\tau}^{\varepsilon-1}, \dots, \lambda_1/\bar{\tau}^{\varepsilon-1})$$

so that, as each component of ϕ^m is bounded above, the monotone components of $\phi^{m+1}(\lambda_1/\bar{\tau}^{\varepsilon-1}, \dots, \lambda_1/\bar{\tau}^{\varepsilon-1})$ converge to (w_2^*, \dots, w_I^*) . This proves the existence of a fixpoint of ϕ , which by construction is a solution to (A1).

Existence and Uniqueness of a solution to system (37), (40), and (41). First notice that there is nothing to show when (35) is violated. Now use (37) to write

$$c_f = \sqrt{\frac{\alpha}{\beta L \tau} - \frac{L^*}{L \tau} (c_d^*)^2} \quad \text{and} \quad c_f^* = \sqrt{\frac{\alpha}{\beta L^* \tau} - \frac{L}{L^* \tau} c_d^2}$$

with the function $c_f(c_f^*)$ decreasing in c_d^* (in c_d). Then rewrite (41) as

$$\frac{c_d + 1}{c_f^* + 1} - \tau \frac{c_f + 1}{c_d^* + 1} = 0$$

The LHS is increasing in c_d and in c_d^* . so that $c_d = g(c_d^*)$ is a decreasing function. Use (35) to check that $c_f > 0$ at $c_d \rightarrow \alpha/(\beta L)$. Finally, write (40) as

$$(c_d + 1) \left(L \frac{c_d}{c_f^*} + L^* \tau \right) + \frac{\alpha}{\beta c_f^*} - \left\{ (c_d^* + 1) \left(L^* \frac{c_d^*}{c_f} + L \tau \right) + \frac{\alpha}{\beta c_f} \right\} = 0$$

By the above observations the LHS is decreasing in c_d with $\text{LHS} \rightarrow \infty$ as $c_d \rightarrow \alpha/(\beta L)$ and $\text{LHS} \rightarrow -\infty$ as $c_d^* \rightarrow \alpha/(\beta L^*)$. Thus, there is a unique solution for c_d and, solving backward, for c_d^* and finally for c_f and c_f^* .

B References

- Acemoglu, D., and Ventura, J. 2002: "The World Income Distribution" *Quarterly Journal of Economics*, Vol. 117, No. 2, pp. 659-694
- Anderson J. and van Wincoop E. 2003: "Gravity with Gravitas: A Solution to the Border Puzzle" *American Economic Review*, Vol. 93, 1, pp. 170-192
- Anderson J. and van Wincoop E. 2004: "Trade Costs" *Journal of Economic Literature*, Volume 42, No. 3, pp. 691-751
- Baier, S. and Bergstrand, J. 2001: "The Growth of World Trade: Tariffs, Transport Costs, and Income Similarity" *Journal of International Economics*, Vol. 53, pp. 1-27
- Backus, D., Kehoe, P., and Kydland, F. 1994: "Dynamics of the Trade Balance and the Terms of Trade: The J-Curve?" *American Economic Review*, Vol. 84, 1, pp. 84-103
- Baumol, W. J. 1973: "Income and Substitution Effects in the Linder Theorem" *Quarterly Journal of Economics*, Vol. 87, No. 4, pp. 629-633
- Bergoeing R., Kehoe T. J., Strauss-Kahn V., and Yi K-M 2004: "Why Is Manufacturing Trade Rising Even as Manufacturing Output is Falling?" *American Economic Review*, 94, pp. 134-138
- Bergoeing R. and Kehoe T. J. 2003: "Trade Theory and Trade Facts" Federal Reserve Bank of Minneapolis, Research Department Staff Report 284
- Bergstrand, J. 1990: "The Heckscher-Ohlin-Samuelson Model, the Linder Hypothesis, and the Determinants of Bilateral Intra-Industry Trade" *Economic Journal*, Vol. 100, pp. 1216-1229
- Broda, Christian and Weinstein, David E. 2004: "Globalization and the Gains From Variety" NBER WP10314
- Broda, Christian and Weinstein, David E. 2004: "Variety Growth and World Welfare" *American Economic Review*, Vol. 94, 2, pp. 139-144
- Corsetti, G., Martin, P. and Pesenti, P. 2005: "Productivity Spillovers, Terms of Trade and the 'Home Market Effect'" NBER WP 11165
- Chaney, T., 2006: "Distorted Gravity: Heterogeneous Firms, Market Structure and the Geography of International Trade" mimeo.
- Cuñat, A., and Maffezzoli, M. 2003: "Can Comparative Advantage Explain the Growth of US Trade?" IGIER WP 241
- Dornbusch, R., Fischer, S., and Samuelson P. A. 1980: "Heckscher-Ohlin Trade Theory with a Continuum of Goods" *Quarterly Journal of Economics*, Vol. 95, No. 2, pp. 203-224
- Dornbusch, R., Fischer, S., and Samuelson P. A. 1977: "Comparative Advantage, Trade, and Payments in a Ricardian Model with a Continuum of Goods" *American Economic Review*, Vol. 67, No. 5, pp. 823-839

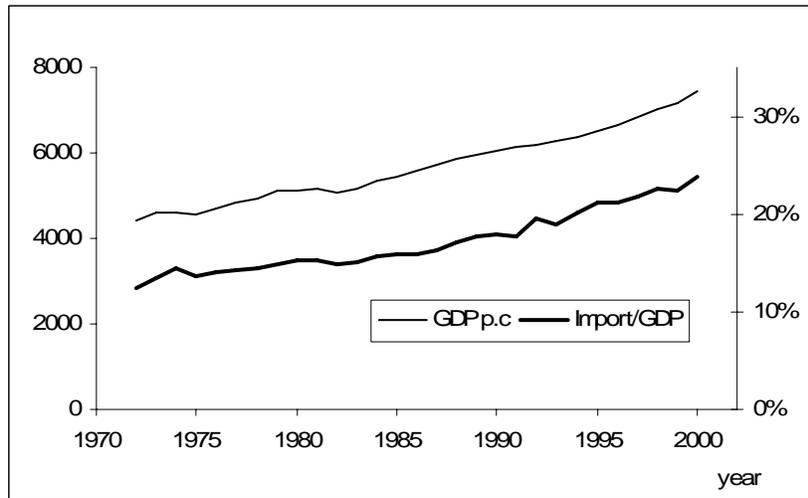
- Eaton, J. and Kortum, S. 2002: "Technology, Geography, and Trade" *Econometrica* Vol. 70, No. 5, pp. 1741-1779
- Evenett, S. J., and A. J. Venables, 2002: "Export Growth in Developing Countries: Market Entry and Bilateral Trade Flows" *mimeo*
- Feenstra, R., Romalis, J., and Schott, P. 2002. "U.S. Imports, Exports and Tariff Data 1989-2001" NBER WP 9387.
- Feenstra, R., Lipsey, R., Deng, H., Ma, A., and Mo, H. 2005: "World Trade Flows 1962-2000" NBER WP 11040.
- Harrigan, J. 1994: "Scale Economies and the Volume of Trade" *Review of Economics and Statistics*, vol. 76, pp. 321-328
- Helpman, E. and Krugman, P. 1985: Trade Policy and Market Structure, MIT press, Massachusetts
- Haveman, J. and Hummels, D., 1999: "Alternative Hypothesis and the Volume of Trade: Evidence on the Extent of Specialization", Purdue University
- Hummels, D. 1999: "Have Internatioanl Transportation Costs Declined?" University of Chicago
- Hummels, D., and Klenow, P. J., 2005: "The Variety and Quality of a Nation's Exports" *American Economic Review*, 93, 3, pp. 704-723
- Hunter, L. C., and Markusen, J. R.: "Per-Capita Income As a Determinant of Trade", in Feenstra, R. (ed.) *Empirical Methods for International Trade*, MIT Press 1988
- Kehoe, T. and Ruhl, K. 2002: "How Important is the New Goods Margin in Internatioanl Trade?" Staff Report 324, Federal Reserve Bank of Minneapolis
- Klenow, P. J., and Rodríguez-Clare, A. 1997: "Quantifying the Gains from Trade Liberalization", *mimeo* Stanford University
- Krugman, P. 1980: "Scale Economies, Product Differentiation, and the Pattern of Trade" *American Economic Review*, 70, pp. 950-959
- Krugman, P. 1995: "Growing World Trade: Causes and Consequences" *Brookings Papers on Economic Activity*, 1, pp. 327-376
- Melitz, M. 2003: "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity" *Econometrica* Vol. 71, No. 6, pp. 1695-1725
- Piketty, T. and Saez, E. 2003: "Income Inequality in the United States 1913-1998" *Quarterly Journal of Economics*
- Ruhl, K. 2005: "Solving the Elasticity Puzzle in International Ecoomics" *mimeo* University of Texas at Austin
- Schmitt, N. and Yu, Z. 2002: "Horizontal Intra-Industry Trade and the Growth of International Trade" University of Nottingham, *mimeo*

- Thursby, J. G. and Thursby, M. C. 1987: "Bilateral Trade Flows, the Linder Hypothesis, and Exchange Risk" *Review of Economics and Statistics*, Vol. 69, No. 3, pp. 488-495
- Stokey, N. 1988: "Learning by Doing and the Introduction of New Goods" *Journal of Political Economy*, Vol. 96, pp. 701-717
- Ventura, J. 2006: "A Global View of Economic Growth" *Handbook of Economic Growth*, vol. 1B, Aghion P. and Durlauf S. N. (eds.), Elsevier B. V.
- Yi, K.-M. 2003: "Can Vertical Specialization Explain the Growth of World Trade?" *Journal of Political Economy*, Vol. 111, No.1, pp. 52-102
- Young, A. 1991: "Learning by Doing and the Dynamic Effects of International Trade" *Quarterly Journal of Economics*, Vol. 106, No. 2, pp. 369-405

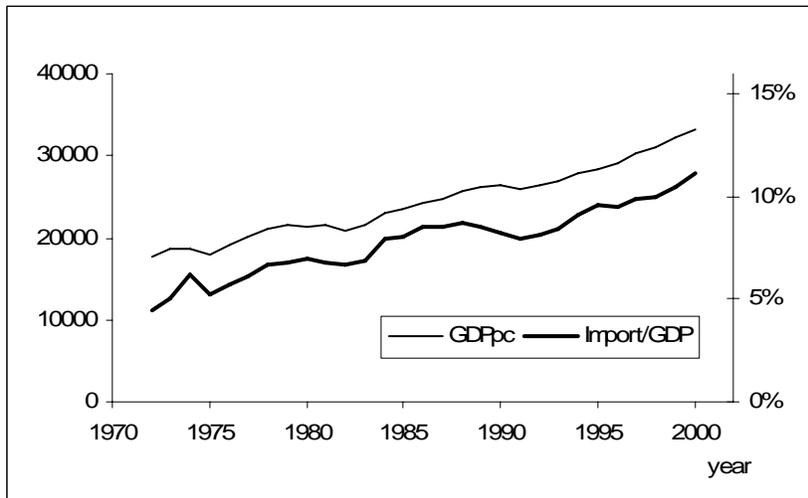
Figure 1a

Per Capita Income and Non-Oil Import Shares

World

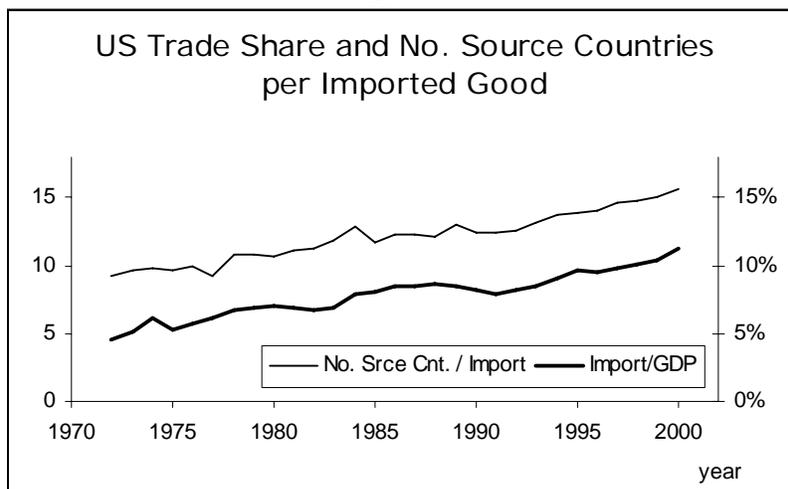


USA



Data: PWT 6.1 and Feenstra et al (2005)

Figure 2

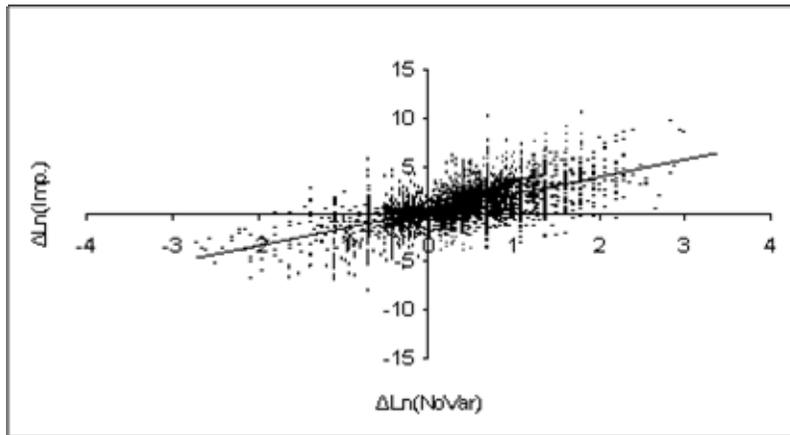


Data: Feenstra et al (2002)

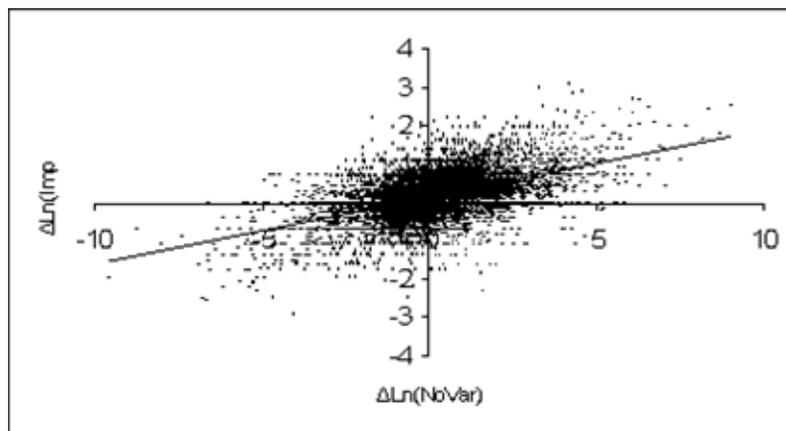
Figure 3a

Growth in Import Volume and
Growth in No. of Source Countries for US by Good
Intensive Margin

TSUSA classification (1972-1988)



HTS classification (1989-2000)

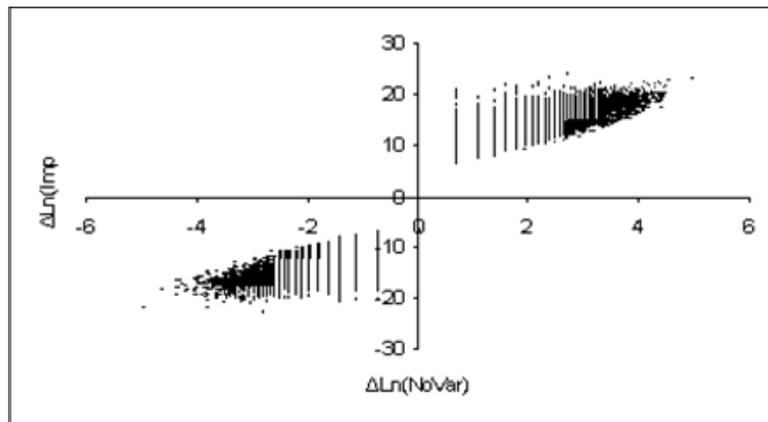


Data: Feenstra et al (2002)

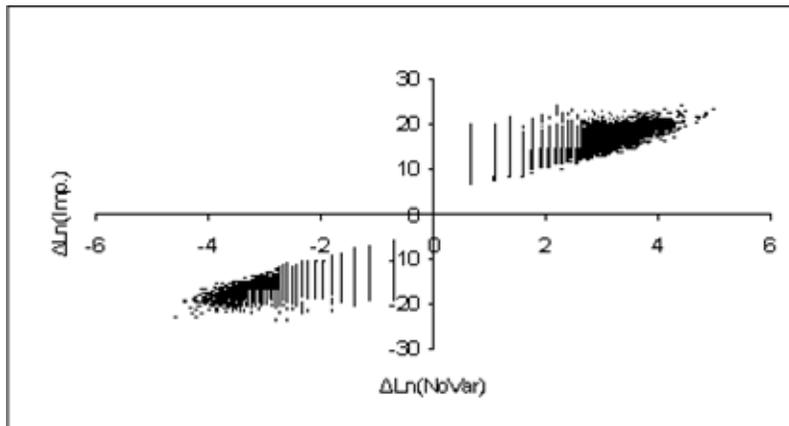
Figure 3b

Growth in Import Volume and
Growth in No. of Source Countries for US by Good
Extensive Margin

TSUSA classification (1972-1988)



HTS classification (1989-2000)

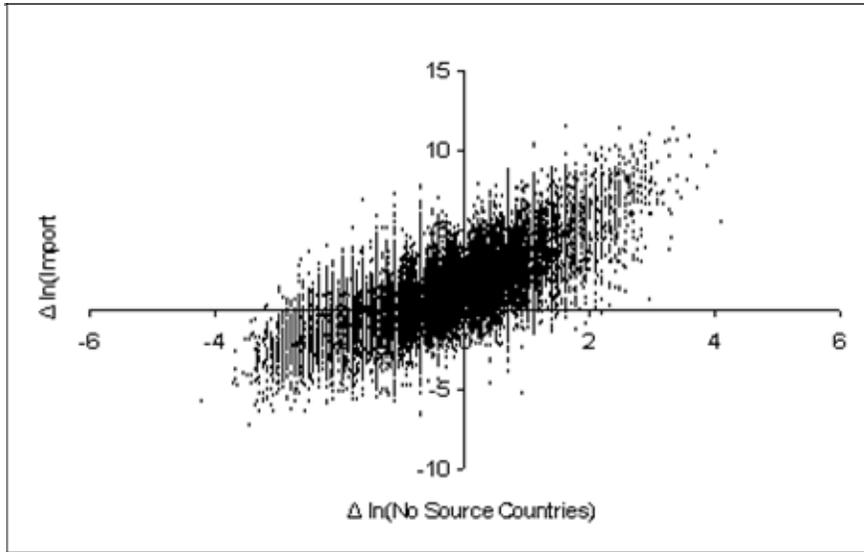


Data: Feenstra et al (2002)

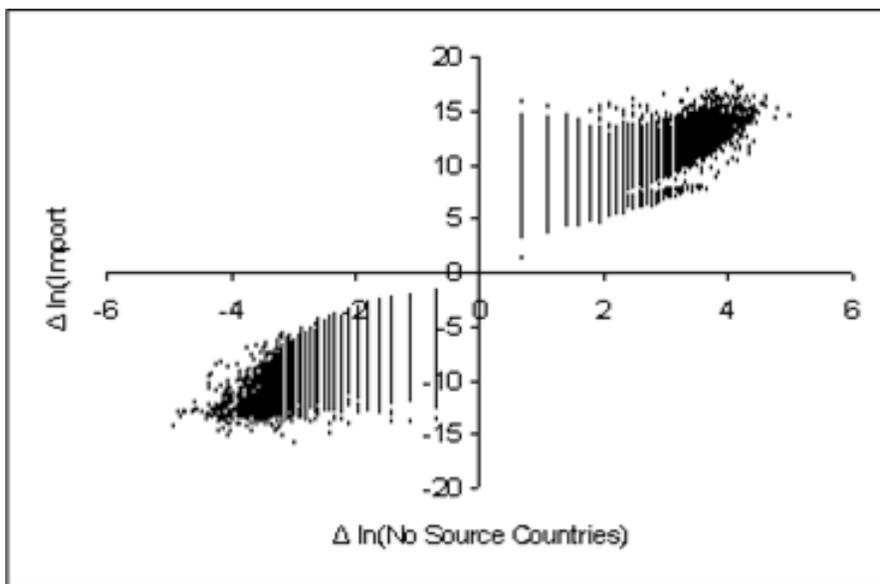
Figure 4

Change in Import Volume and Number of Source Countries
By Country and Sector (1972 - 2000)

Intensive Margin



Extensive Margin



Data: Feenstra et al (2005)

Figure 5

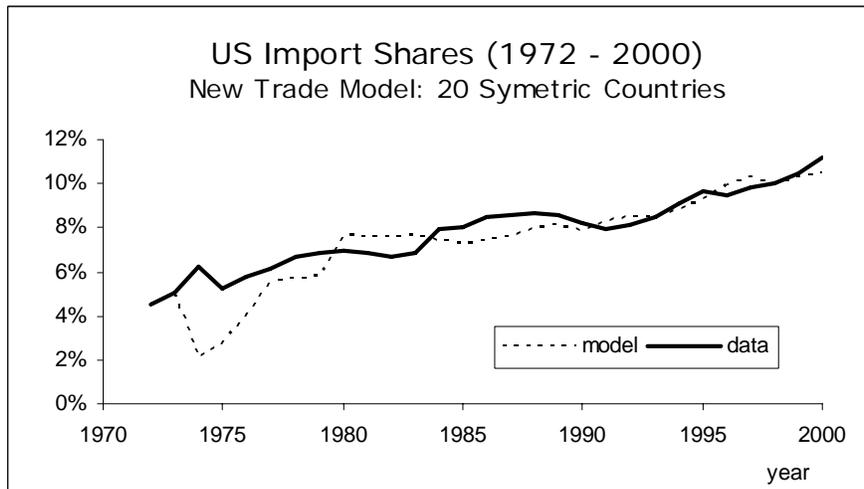
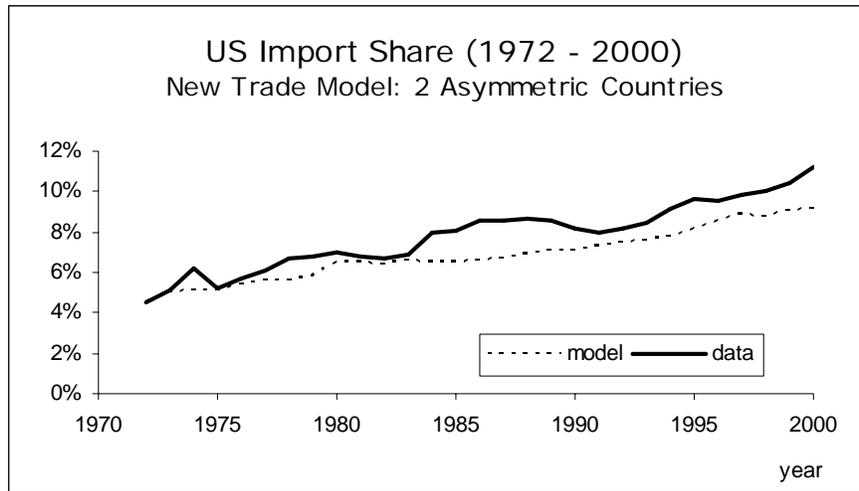


Figure 6

