Ricardian Productivity Differences and the Gains from Trade*

Andrei A. Levchenko
University of Michigan
NBER and CEPR

Jing Zhang
Federal Reserve Bank of Chicago

October 15, 2013

Abstract

This paper evaluates the role of sectoral heterogeneity in determining the gains from trade. We first show analytically that in the presence of sectoral Ricardian comparative advantage, a one-sector sufficient statistic formula that uses total trade volumes as a share of total absorption systematically understates the true gains from trade. Greater relative sectoral productivity differences lead to larger disparities between the gains implied by the one-sector formula and the true gains. Using data on overall and sectoral trade shares in a sample of 79 countries and 19 sectors we show that the multi-sector formula implies on average 30% higher gains from trade than the one-sector formula, and as much as 100% higher gains for some countries. We then set up and estimate a quantitative Ricardian-Heckscher-Ohlin model in which no version of the formula applies exactly, and compare a range of sufficient statistic formulas to the true gains in this model. Confirming the earlier results, formulas that do not take into account sectoral heterogeneity understate the true gains from trade in the model by as much as two-thirds. The one-sector formulas understate the gains by more in countries with greater dispersion in sectoral productivities.

JEL Classifications: F4

Keywords: gains from trade, comparative advantage, sufficient statistics

*We are grateful to Alan Deardorff, Andrés Rodríguez-Clare, Linda Tesar, an anonymous referee, and seminar participants at the University of Michigan, University of Nottingham, University of Warwick, and the 2013 NBER ITI Spring Meetings for helpful suggestions. The views expressed here are those of the authors and do not necessarily represent those of the Federal Reserve Bank of Chicago, the Federal Reserve System, or its Board of Governors. E-mail (URL): `alev@umich.edu` (`http://alevchenko.com`), `jzhang@chifrb.org` (`https://sites.google.com/site/jzhangzn/`).

1 Introduction

It has recently been discovered that under certain restrictions, models with different micro-structures – Armington, Eaton-Kortum, and Melitz – deliver an identical expression for gains from trade (Arkolakis, Costinot and Rodríguez-Clare 2012a). In its simplest form, the gross proportional gains from trade can be expressed as $\pi_{ii}^{-1/\theta}$, where π_{ii} is the share of spending on domestically produced goods in total spending of the economy, and θ is the elasticity of trade flows with respect to iceberg trade costs.

This result is potentially powerful for two closely related and mutually reinforcing reasons. The first is that one need not even take a precise stand on the micro-structure of the economy or motives for international exchange when evaluating the gains from trade. The second is that the gains are in fact expressed in terms of only a couple of (in principle) observable values. Thus, if the right data are available, and/or one is willing to take a stand on the value of just one or two reduced-form parameters, one could compute a country's gain from trade easily, without the heavy lifting of setting up, estimating/calibrating, and solving a large-scale quantitative trade model. This promises to dramatically lower barriers to applied policy analysis, and transform the way policy research on the benefits of trade is conducted.

Since the formula features the share of spending on domestically produced goods in total spending, an (over-)simplified message from this literature can easily be taken to be "all you need to know is the overall trade volume." By contrast, it has been understood since Ricardo (1817) that when sectors are imperfect substitutes in consumption, comparative advantage in international trade matters for the magnitude of the gains. Broadly speaking, the stronger the comparative advantage – the more different countries are in their relative technology – the larger are the gains from trade. In turn, cross-sectoral Ricardian comparative advantage will manifest itself in differences in trade shares across sectors, as countries will have less imports as a share of total spending in their comparative advantage sectors, and more imports in their comparative disadvantage sectors. Does the simple formula featuring the overall trade openness adequately capture this source of gains from trade? Arkolakis et al. (2012a) show how under certain conditions, the formula can be extended to a setting with multiple sectors, but do not discuss how the gains from trade in a multi-sector formula differ from the gains implied by the one-sector formula.

This paper shows that by ignoring the sectoral heterogeneity in productivity – and hence in trade volumes – the one-sector formula systematically understates the gains from trade. We develop this result in three ways. First, we use a simplified 2-sector, 2-country model to show analytically that the gains from trade according to the one-sector formula coincide with the true gains only when the trade volumes relative to absorption are the same in every sector. *Any* deviation from equal trade volumes across sectors – holding aggregate trade volumes relative to absorption constant – leads to larger gains than what is implied by the one-sector formula. Greater differences in relative sectoral productivities, implying greater dispersion in sectoral trade shares, increase both the gains from trade and the disparity between the true gains and those implied by the one-sector formula.

Second, the differences between one- and multi-sector formula gains are large in magnitude when computed on observed trade flows. We use actual data on output and cross-border trade at sector level in a large sample of countries to compare the gains implied by the one-sector formula to those implied by the formula that uses information on sectoral trade shares. Almost without exception, the gains according to the one-sector formula are lower than those according to the multi-sector formula, confirming the analytical results. The multi-sector formula produces 30% higher gains from trade on average, and as much as 100% higher in countries with a large dispersion in sectoral trade shares. The disparity is also highly non-uniform across countries. There are many examples in which total trade volumes – and thus the gains according to the one-sector formula – are quite similar for a pair of countries, but the multi-sector formula gains differ substantially because of very different degrees of dispersion in sectoral trade shares.

The third part of the paper relates the dispersion in the sectoral trade shares and the resulting bias in the one-sector formulas explicitly to Ricardian comparative advantage. We use the sectoral productivity estimates and the large-scale 79-country 20-sector Ricardian-Heckscher-Ohlin model built by Levchenko and Zhang (2011, 2012). In addition to Ricardian productivity differences, the framework incorporates multiple factors of production, a non-tradeable sector, and intermediate input linkages between and within the sectors. The parameters are calibrated using observed production technologies and input-output matrices, and estimated using bilateral sector-level trade. The model matches relative factor prices and both overall and bilateral trade quite well. In this framework, no version of the formula applies. Because we fully solved it, we can compute the true gains from trade inside this model. We then compare these gains to a series of sufficient statistic formulas, to evaluate which types of formulas work best, and thus which features of the real world are a must to include in any formula-based analysis. It is important to emphasize that we do not claim that gains from trade computed in this model coincide with the true gains from trade in reality. No matter how rich the model, the real world will inevitably be more complex. In our view, the exercise is nonetheless informative because the model does capture, in a quantitative way, several distinct sources of the gains from trade, which can then be compared to the sufficient statistic formulas' attempts to also capture (some of) those gains.

Our main finding is that the simpler formulas that do not use information on sectoral trade volumes understate the true gains from trade dramatically, often by more than two-thirds. The error in the formulas across countries is strongly negatively correlated to the strength of Ricardian comparative advantage: the one-sector formula-implied gains understate the true gains from trade by more in countries with greater dispersion in sectoral productivity. The model-based exercise thus reinforces the main result of the paper that accounting for sectoral heterogeneity in productivity is essential for a reliable assessment of the gains from trade.

We conclude from the analytical results, the data-based evaluation, and the model-based evaluation that policy analysis must take sectoral variation in trade volumes into account when computing the gains from trade using the sufficient statistic approach.

This paper contributes to the recent literature on the welfare gains from trade that arose in parallel with, and in response to, Arkolakis et al. (2012a). A number of papers explore the impact of market structure (the extensive margin, variable markups) on the gains from trade (see, among others, Arkolakis, Demidova, Klenow and Rodríguez-Clare 2008, Atkeson and Burstein 2010, Feenstra 2010, Corcos, Del Gatto, Mion and Ottaviano 2012, Arkolakis, Costinot, Donaldson and Rodríguez-Clare 2012b, de Blas and Russ 2012, di Giovanni and Levchenko 2013).

This paper addresses sectoral heterogeneity instead. In that respect, our analysis is most closely related to Imbs and Méjean (2011) and Ossa (2012), who explore the biases in the one-sector gains from trade formula induced by sectoral heterogeneity, and Costinot and Rodríguez-Clare (2012), who provide a series of gains from trade results under different assumptions, including one-sector and multi-sector formulas. Relative to these papers, our substantive point is complementary and distinct. We explore variation in sectoral productivity and trade volumes, whereas Imbs and Méjean (2011), Ossa (2012), and Costinot and Rodríguez-Clare (2012) focus on the variation in the trade elasticities across sectors. To make this distinction as transparent as possible, our analysis is carried out under the same trade elasticity in all sectors. While Imbs and Méjean (2011) and Costinot and Rodríguez-Clare (2012) also allude to the dispersion in trade shares across sectors, we relate it explicitly to Ricardian comparative advantage and present a complete elaboration of this effect both analytically and quantitatively. In addition, we provide a model-based evaluation of the performance of the different sufficient statistic formulas. Our argument that sectoral productivity heterogeneity increases the gains from trade is related to the arguments in Melitz and Redding (2013) regarding the impact of firm heterogeneity.

The last part of the paper is based on the quantitative Ricardian-Heckscher-Ohlin model developed and estimated by Levchenko and Zhang (2011, 2012) and di Giovanni, Levchenko and Zhang (2013). This paper shares with our earlier work the emphasis on the quantitative welfare impact of sectoral technology differences. While our earlier papers focus on technological change or trade integration of individual regions (China, Eastern Europe), this paper investigates the performance of the various sufficient statistic formulas for the gains from trade. The model-based exercise in the last part of the paper is methodologically related to studies that evaluate empirical estimation approaches by applying them to model-simulated data in which the true data-generating process is by construction known. See, for instance, Chari, Kehoe and McGrattan

(2008) for an application of this strategy in the business cycle context, and Hauk and Wacziarg (2009) for an application in the economic growth context. Our approach of looking for observables that accurately capture the gains from trade bears an affinity to Burstein and Cravino (2012), who explore the conditions under which real GDP data will reflect the welfare gains from reductions in trade costs.

The rest of the paper is organized as follows. Section 2 derives the analytical results on the impact of sectoral dispersion in trade flows on the gains from trade. Section 3 compares one-sector and multi-sector sufficient statistic formulas for the gains from trade using data on the trade volumes and production. Section 4 sets up the quantitative framework and discusses the results of comparing a range of sufficient statistic formulas to the true gains from trade. Section 5 concludes.

2 Analytical Results

This section presents a simple model to illustrate the role of sectoral heterogeneity in determining the gains from trade. The exercise makes two main points. First, the one-sector formula for the gains from trade systematically understates the true gains from trade when there is dispersion in trade volumes across sectors. Holding the overall trade volume – and thus the gains implied by the one-sector formula – constant, the true gains from trade increase in the cross-sectoral dispersion in trade shares. And second, the degree of dispersion in trade shares across sectors is intimately related to the strength of Ricardian comparative advantage.

Consider a two-country (1 and 2) two-sector (a and b) Eaton and Kortum (2002, henceforth EK) model. Labor is the only input in production, and both countries are endowed with one unit: $L_1 = L_2 = 1$. International trade is costless. Each sector is a CES composite of a continuum of varieties $[0, 1]$ that do not overlap across sectors, and country i can produce each infinitesimal variety with productivity drawn randomly from the Fréchet distribution with dispersion parameter θ common to all countries and sectors, and central tendency parameter T_i^j in country i, sector j. The parameter T_i^j determines the average productivity draw in country i sector j. Thus, as a shorthand we will refer to T_i^j as country i's "productivity" in sector j. As the structure of the EK model is standard, we do not reproduce the functional forms and derivations here (the more full-fledged multi-sector EK model is described in Section 4), and focus instead on the outcomes.

Preferences are Cobb-Douglas over the broad sectors of the economy, an assumption adopted throughout the paper. Costinot and Rodríguez-Clare (2012) show that the overall gains from trade depend heavily on the "upper level" elasticity of substitution, and point out that we currently have no reliable estimates of what that elasticity would be across broad sectors of the economy. The level of sectoral disaggregation we employ throughout the paper is quite coarse

– 19 tradeable sectors – making the Cobb-Douglas assumption relatively more plausible than it would be with finely disaggregated sectors.[1] The result that sectoral dispersion in trade shares will increase the gains from trade conditional on the overall trade volume is more general, however, and would obtain whenever the elasticity of substitution between sectors is lower than the elasticity of substitution within sectors – a reasonable assumption.

To obtain analytical results under endogenous factor price determination, we assume (i) equal expenditure shares in the two sectors and (ii) a mirror image of productivities across sectors and countries: $T_1^a = T_2^b$ and $T_1^b = T_2^a$. Of course, relative productivities are not generically the same in the two countries, $T_1^a/T_1^b \neq T_2^a/T_2^b$, and thus the strength of comparative advantage can vary. The mirror image assumption on sectoral productivity and symmetry in the utility function ensure that the wages are equal in the two countries, $w_1 = w_2$, which we set as the numeraire. Together with the normalization of labor endowments to 1 this implies that the total income/expenditure in each country is equal to 1, and trade flows expressed as shares of total expenditure also equal absolute trade flows.

Denote by π_{ni}^j the share of total expenditure in country n on goods coming from country i in sector j. The import shares in country 1 from country 2 are given by

$$\pi_{12}^a = \frac{T_2^a w_2^{-\theta}}{T_1^a w_1^{-\theta} + T_2^a w_2^{-\theta}} = \frac{T_2^a}{T_1^a + T_2^a}, \tag{1}$$

$$\pi_{12}^b = \frac{T_2^b w_2^{-\theta}}{T_1^b w_1^{-\theta} + T_2^b w_2^{-\theta}} = \frac{T_2^b}{T_1^b + T_2^b} = \frac{T_1^a}{T_1^a + T_2^a}, \tag{2}$$

where the last equality uses the mirror image assumption we put on the relative productivities ($T_1^a = T_2^b$ and $T_1^b = T_2^a$).

Importantly, regardless of the sectoral import shares, the assumptions on preferences and technology imply that the overall imports (and thus domestically produced goods) as a share of total absorption is always one half: $\pi_{11} = \pi_{22} = \frac{1}{2}$, regardless of the strength of comparative advantage.[2] An "econometrician" that assumes this economy is well-characterized by a one-sector

[1] As a related point, our results cannot be "extrapolated" to argue that adopting an ever finer level of sectoral disaggregation will lead to ever higher implied gains from trade, for two reasons. First, the multi-sector gains from trade formulas explored in the paper only apply if the utility over these sectors is Cobb-Douglas. As we disaggregate sectors further and further, the Cobb-Douglas assumption will become implausible. While the assumption of unitary elasticity of substitution is palatable at the 19-sector level of disaggregation (Food Products, Textiles, Apparel, etc), it will be implausible for a 190-sector level of disaggregation, since by construction more finely disaggregated products will be more similar to each other, and thus more substitutable. Second, the model relies on the EK structure within each sector – that is, there is a continuum of varieties within each sector, and at the sector level countries have interior trade shares. Under much more disaggregated sectors, the EK assumption within each sector will be strained.

[2] To see this, note that the total imports as a share of spending in country 1 are given by:

$$\pi_{12} = \frac{w_1 L_1 \frac{1}{2}\pi_{12}^a + w_1 L_1 \frac{1}{2}\pi_{12}^b}{w_1 L_1} = \frac{1}{2},$$

EK model with labor as the only input in production will compute the (gross) gains from trade as[3]

$$\pi_{ii}^{-\frac{1}{\theta}}. \tag{3}$$

This implies that gains from trade computed using only aggregate trade volumes are always constant in this model.

However, as comparative advantage and thus sectoral trade shares change, the true gains from trade will change as well. Welfare in country i is given by the indirect utility function, and corresponds to real income

$$\frac{w_i}{(p^a p^b)^{\frac{1}{2}}}, \tag{4}$$

where p^j is the price of sector $j = a, b$ composite, which does not vary by country since trade is costless. Standard steps lead to the expression for welfare as a function of technology and trade shares by sector:[4]

$$\frac{w_i}{(p^a p^b)^{\frac{1}{2}}} = \Gamma^{-1}\left(T_i^a T_i^b\right)^{\frac{1}{2\theta}}\left(\pi_{ii}^a \pi_{ii}^b\right)^{-\frac{1}{2\theta}}, \tag{5}$$

where Γ is a constant. Thus, the true gains from trade in this model are expressed as:

$$\left(\pi_{ii}^a \pi_{ii}^b\right)^{-\frac{1}{2\theta}} = \left(\pi_{ii}^a(1-\pi_{ii}^a)\right)^{-\frac{1}{2\theta}}, \tag{6}$$

where the second equality is due to the fact that $\pi_{11}^b = \pi_{22}^a = 1 - \pi_{21}^a = 1 - \pi_{12}^a$. The following result is immediate, coming from differentiation of (6) with respect to π_{ii}^a:

Proposition 1. *When the share of absorption spent on domestically produced goods is the same across sectors, $\pi_{ii}^a = \pi_{ii}^b$, (i) the true gains from trade attain their minimum, and (ii) the gains from trade implied by the one-sector formula* (3) *and computed based on aggregate imports and absorption coincide with the true gains from trade. Therefore, the one-sector formula understates the true gains from trade as long as $\pi_{ii}^a \neq \pi_{ii}^b$.*

The result stated in the proposition is illustrated in Figure 1(a). It plots the true gains from trade and the gains from trade implied by the one-sector formula against the share of spending on domestic goods in sector a, π_{ii}^a. Because of the symmetry assumptions we imposed, the total trade volume as a share of absorption is fixed throughout, and so the gains implied by the one-sector model are constant as π_{ii}^a varies. The true gains from trade, computed using information on sectoral trade volumes, are always at least as great as the gains implied by the one-sector formula,

where the last equality comes from the fact, immediate from (1)-(2), that $\pi_{12}^b = 1 - \pi_{12}^a$. Since the share of total spending on domestically produced goods and on imports sum to 1, $\pi_{11} = 1 - \pi_{12} = \frac{1}{2}$.

[3]See Eaton and Kortum (2002), Arkolakis et al. (2012a).

[4]The price index in sector j is given by $p^j = \Gamma(T_1^j + T_2^j)^{1/\theta}$ (Eaton and Kortum 2002). Combining this expression with the trade shares (1)-(2) to write p^j as a function of π_{ii}^j and plugging into (4) yields (5).

and the deviation gets larger the larger is the asymmetry between the sectoral trade shares π^j_{ii}. The one-sector formula yields the true gains from trade only when the trade shares are identical in the two sectors: $\pi^a_{ii} = \pi^b_{ii} = \frac{1}{2}$.

Finally we draw the connection between the magnitude of the gains from trade and the strength of comparative advantage.

Corollary 1. *The deviation between the true gains and the gains implied by the one-sector formula increases as the comparative advantage becomes stronger – that is, as the differences in relative* T^j_i*'s grow.*

The corollary is immediate from equations (1)–(2). The point at which the welfare gain implied by the one-sector formula coincides with the true gains corresponds to identical technology in the two countries: $T^a_1 = T^b_1 = T^a_2 = T^b_2$. The sectoral absorption shares in the gains from trade formula (6) are more different the more different are T^a_1 and T^b_1. Figure 1(b) depicts this relationship. It plots the percentage difference between the gains from trade implied by the one-sector formula and the true gains against the dispersion in sectoral T^j_i's, measured by the standard deviation between T^a_1 and T^b_1. As expected, greater dispersion in T^j_i's results in the one-sector formula missing the true gains by more.

3 Dispersion in Sectoral Trade Shares and Gains from Trade

It is possible that while analytically we can show that the one-sector formula systematically understates the gains from trade in a multi-sector environment, real-world sectoral trade shares are such that this understatement is not large in magnitude. In this section we use actual data on manufacturing production and trade for a sample of 79 countries to assess how large are the disparities between the gains from trade implied by the one-sector and the multi-sector formulas under the observed sectoral trade shares.

For 52 countries in the sample, information on output comes from the 2009 UNIDO Industrial Statistics Database. For the 27 European Union countries plus FYR Macedonia, the EUROSTAT database contains data of superior quality, and thus for those countries we used EUROSTAT production data. The two output data sources are merged at the roughly 2-digit ISIC Revision 3 level of disaggregation, yielding 19 manufacturing sectors. Bilateral trade data come from the United Nations' COMTRADE database, concorded to the same sectoral classification. Sectoral absorption shares are averaged for the period 2000-2007, which is the time period on which we carry out the analysis. Appendix Table A1 lists the countries used in the analysis, while Appendix Table A2 lists the sectors.

We compare the gains from trade implied by two formulas: a one-sector formula in which the manufacturing sector is treated as one, and a multi-sector formula. To scale the aggregate gains from trade appropriately, we augment these formulas with a non-tradeable sector. The one-sector formula is

$$\pi_{ii}^{-\frac{\xi}{\theta}} - 1. \tag{7}$$

The multi-sector formula (see Arkolakis et al. 2012a, section IV.A) is:

$$\prod_{j=1}^{J} \left(\pi_{ii}^{j}\right)^{-\frac{\xi\omega_j}{\theta}} - 1, \tag{8}$$

where j indexes sectors.

To be precise, (7) is the gains from trade in a model with labor as the only input in production, one tradeable and one non-tradeable sector, with utility Cobb-Douglas in the two sectors and ξ as the expenditure share of tradeables. The tradeable sector can be an EK-type sector as in Sections 2 and 4, or an Armington-type sector that aggregates goods from different origin countries. The non-tradeable sector can be either an EK-type sector or a sector producing a homogeneous good. The model that yields (8) as the gains from trade is the same except tradeables are composed of j sectors, over which utility is Cobb-Douglas with expenditure shares ω_j. Once again, each tradeable sector can be either EK or Armington.

For a fully-fledged description of one particular model that can lead to these gains from trade formulas, we can refer to Section 4 below. Formula (7) captures the gains from trade in the model set out in Section 4 when (i) there is only 1 traded sector ($J = 1$), (ii) there is no capital ($K_n = 0$ $\forall n$ and $\alpha_j = 1$ $\forall j$) and (iii) there are no input-output linkages ($\beta_j = 1$ $\forall j$). Formula (8) captures the gains from trade in the model in Section 4 under conditions (ii)-(iii), but with $J > 1$.

In formulas (7) and (8), the absorption shares in manufacturing, both in aggregate (π_{ii}) and at the sector level (π_{ii}^{j}) come directly from the data on production and trade. To implement these formulas, we must take a stand on a number of parameters. We adopt two alternative assumptions on the Cobb-Douglas shares of sectors in consumption ω_j. First, to make the mechanics behind sectoral heterogeneity's effect as transparent as possible, we set those weights to be equal across sectors. Second, since ω_j's are equivalent to consumption shares, we set them to the actual absorption shares of each manufacturing sector in each country. Since those shares differ across countries, and since a great deal of gross absorption of manufacturing output goes to intermediate input usage, this approach may be less transparent, but it will account for inherent differences in sector sizes. The other two parameters (θ and ξ) do not matter for the qualitative conclusions about the direction of the effect, since they both exponentiate the whole formula. Anticipating the quantitative exercise below, in which this elasticity has a concrete interpretation in the context of the Eaton-Kortum model, we set θ to 8.28, which is the preferred value in Eaton and Kortum (2002). We set the Cobb-Douglas utility weight of the manufacturing sector to 0.35 (Alvarez and

Lucas 2007).

Table 1 presents the summary statistics for the gains from trade implied by the one-sector formula (7) and the multi-sector formula (8), under two assumptions on the utility weights ω_j. The bottom two rows report the summary statistics for the proportional difference between the multi-sector gains from trade and the one-sector gains from trade. The clear result is that the one-sector formula systematically understates the gains from trade relative to the multi-sector formula. The difference is substantial in proportional terms: the gains implied by the multi-sector formulas are 32% and 56% larger, on average, than the one-sector gains. In the more extreme cases, the gains implied by the multi-sector formulas are twice as large as the gains implied by the one-sector formula. The direction of the bias is also (nearly) universal: in every single one of the 79 countries, the expenditure-weighted formula implies larger gains than the one-sector formula, and in 78 out of 79 countries, the equal-weighted formula implies larger gains.

Figure 2 presents the results graphically, by plotting the multi-sector formula gains on the y-axis against the one-sector formula gains on the x-axis. The solid dots denote the gains implied by the equal-weighted sectors, while the hollow dots denote the gains under the expenditure weights. For convenience, a 45-degree line is added to the plot. It is immediate that nearly all the dots are above the 45-degree line – the multi-sector gains are larger than the one-sector gains. It is also clear that the larger are the total gains, the greater is the deviation between the one-sector and the multi-sector formulas.

To provide intuition for how sectoral heterogeneity conditions the gains from trade, consider the comparison between Bolivia and the Czech Republic. In the data, these two countries have similar overall openness in the manufacturing sector, and thus similar gains from trade according to the one-sector formula. In fact, Czech Republic's one-sector formula gains, at 2.39%, are slightly larger than the one-sector formula gains for Bolivia, which are 2.06%. However, at a similar – in fact slightly lower – level of overall openness, the dispersion in trade volumes across sectors is much greater in Bolivia. Figure 3 depicts the shares of domestically produced goods in sectoral spending, π_{ii}^j, for Bolivia and Czech Republic, ranking sectors from the most open (lowest π_{ii}^j) to the least open for each country. It is immediate that the variation in openness across sectors in Bolivia is higher. In its most open sectors, a far greater share of total expenditure goes to imported goods, and in its least open sector, a lower share of expenditure goes to imports, compared to the Czech Republic. This greater heterogeneity implies that Bolivia's gains from trade are actually higher than Czech Republic's, not lower as implied by the one-sector formula. The differences in sectoral heterogeneity also mean that the deviation between the one- and the multi-sector gains is much larger in Bolivia. The expenditure-weighted multi-sector gains in Bolivia are 3.93%, nearly double the gains implied by the one-sector formula, and the equal-weighted gains are 5.57%, or 2.7 times larger. By contrast, for the Czech Republic, the multi-sector gains are 2.67% and 2.68%, or only 11−12% larger than the one-sector gains.

The pattern that greater dispersion in sectoral π_{ii}^{j}'s implies greater deviations between multi- and one-sector gains is a general one. Figure 4 plots the proportional difference between the one-sector gains and the expenditure-weighted multi-sector gains against the standard deviation of π_{ii}^{j}'s across sectors in each country, which we use as a measure of dispersion in trade shares across sectors. The correlation between the two is -0.82. In countries with larger dispersion in trade shares across sectors, the one-sector formula understates the gains by more, relative to the multi-sector formula.

4 Quantitative Assessment

The previous section performed a data-driven exercise: it used actual trade shares by sector to compare the gains implied by the different formulas. However, variation in sectoral trade shares alone does not itself establish Ricardian comparative advantage, since trade shares can differ for a variety of reasons. In addition, the world economy is much more complex than the models that lead to either formula (7) or formula (8), what we cannot do purely with the data is compare the formula-implied gains to the true gains from trade.

While we cannot ever know the true gains from trade in the real world, in this section we implement a large-scale quantitative trade model, and conduct a model-based exercise in which we compare the true gains from trade in that model to a range of possible sufficient statistic formulas calculated based on model quantities that can (in principle) be measured in actual data.

The exercise has two goals. First, we relate the performance of the different sufficient statistic formulas explicitly to Ricardian comparative advantage. And second, we evaluate which, if any, sufficient statistic formulas that can be computed using real-world data are a good approximation to the true gains from trade in a world that is more complex. Aside from the sufficient statistic formula extensions that account for sectoral heterogeneity such as (8), a number of other extensions are known. For instance, sufficient statistic formulas can be enriched to take into account input linkages, which are well-known to increase the gains from trade. In the world characterized by multiple sectors, input linkages, as well as many additional features at the same time, it is an open quantitative question which extensions to the sufficient statistic formula are essential to get closer to the true gains. This exercise can thus assess the relative importance of the main mechanism in this paper – cross-sectoral dispersion in trade shares – compared to other sources of error in the simple gains from trade formulas.

Different sufficient statistic formulas also have very different data requirements. Clearly, a one-sector formula that uses only total trade flows and gross output requires less data than a multi-sector formula. Thus, from the perspective of applied trade policy evaluation using the sufficient statistic approach, it is important to know what are the minimum data requirements for a reliable assessment of the gains from trade.

4.1 The Environment

The world is comprised of N countries, indexed by n and i. There are J tradeable sectors, plus one nontradeable sector $J+1$. Utility over these sectors in country n is given by

$$U_n = \left(\prod_{j=1}^{J} \left(Y_n^j\right)^{\omega_j} \right)^{\xi_n} \left(Y_n^{J+1}\right)^{1-\xi_n}, \tag{9}$$

where ξ_n denotes the Cobb-Douglas weight of the tradeable sector composite good, ω_j is the share of tradeable sector j in total tradeable expenditure (and $\sum_{j=1}^{J} \omega_j = 1$), Y_n^{J+1} is the nontradeable-sector composite good, and Y_n^j is the composite good in tradeable sector j.[5]

Each sector j aggregates a continuum of varieties $q \in [0,1]$ unique to each sector using a CES production function:

$$Q_n^j = \left[\int_0^1 Q_n^j(q)^{\frac{\varepsilon-1}{\varepsilon}} dq \right]^{\frac{\varepsilon}{\varepsilon-1}},$$

where ε denotes the elasticity of substitution across varieties q, Q_n^j is the total output of sector j in country n, and $Q_n^j(q)$ is the amount of variety q that is used in production in sector j and country n. Producing one unit of good q in sector j in country n requires $\frac{1}{z_n^j(q)}$ input bundles.

Production uses labor (L), capital (K), and intermediate inputs from other sectors. The cost of an input bundle is:

$$c_n^j = \left(w_n^{\alpha_j} r_n^{1-\alpha_j} \right)^{\beta_j} \left(\prod_{k=1}^{J+1} \left(p_n^k\right)^{\gamma_{k,j}} \right)^{1-\beta_j},$$

where w_n is the wage, r_n is the return to capital, and p_n^k is the price of intermediate input from sector k. The value-added based labor intensity is given by α_j, and the share of value added in total output by β_j. Both vary by sector. The shares of inputs from other sectors, $\gamma_{k,j}$ vary by output industry j as well as input industry k.

As standard in the EK model, productivity $z_n^j(q)$ for each $q \in [0,1]$ in each sector j is random, and drawn from the Fréchet distribution with cdf:

$$F_n^j(z) = e^{-T_n^j z^{-\theta}}.$$

In this distribution, the absolute advantage term T_n^j varies by both country and sector, with higher values of T_n^j implying higher average productivity draws in sector j in country n. The

[5]Note that unlike in Sections 2 and 3, the expenditure share of tradeables ξ_n will differ across countries and calibrated using data as described below.

parameter θ captures dispersion, with larger values of θ implying smaller dispersion in draws.

The production cost of one unit of good q in sector j and country n is thus equal to $c_n^j/z_n^j(q)$. Each country can produce each good in each sector, and international trade is subject to iceberg costs: $d_{ni}^j > 1$ units of good q produced in sector j in country i must be shipped to country n in order for one unit to be available for consumption there. The trade costs need not be symmetric – d_{ni}^j need not equal d_{in}^j – and will vary by sector. We normalize $d_{nn}^j = 1$ for any n and j.

All the product and factor markets are perfectly competitive, and thus the price at which country i supplies tradeable good q in sector j to country n is:

$$p_{ni}^j(q) = \left(\frac{c_i^j}{z_i^j(q)}\right) d_{ni}^j.$$

Buyers of each good q in tradeable sector j in country n will only buy from the cheapest source country, and thus the price actually paid for this good in country n will be:

$$p_n^j(q) = \min_{i=1,\ldots,N} \left\{p_{ni}^j(q)\right\}.$$

The model thus contains two features that make a closed-form expression for gains from trade impossible (or at least currently unknown): multiple factors of production and cross-sectoral input-output linkages, both within the tradeable sector, and between tradeables and non-tradeables.

4.2 Characterization of Equilibrium

The **competitive equilibrium** of this model world economy consists of a set of prices, allocation rules, and trade shares such that (i) given the prices, all firms' inputs satisfy the first-order conditions, and their output is given by the production function; (ii) given the prices, the consumers' demand satisfies the first-order conditions; (iii) the prices ensure the market clearing conditions for labor, capital, tradeable goods and nontradeable goods; (iv) trade shares ensure balanced trade for each country.

The set of prices includes the wage rate w_n, the rental rate r_n, the sectoral prices $\{p_n^j\}_{j=1}^{J+1}$, and the aggregate price P_n in each country n. The allocation rules include the capital and labor allocation across sectors $\{K_n^j, L_n^j\}_{j=1}^{J+1}$, final consumption demand $\{Y_n^j\}_{j=1}^{J+1}$, and total demand $\{Q_n^j\}_{j=1}^{J+1}$ (both final and intermediate goods) for each sector. The trade shares include the expenditure share π_{ni}^j in country n on goods coming from country i in sector j.

4.2.1 Demand and Prices

It can be easily shown that the price of sector j's output will be given by:

$$p_n^j = \left[\int_0^1 p_n^j(q)^{1-\varepsilon} dq \right]^{\frac{1}{1-\varepsilon}}.$$

Following the standard EK approach, it is helpful to define

$$\Phi_n^j = \sum_{i=1}^{N} T_i^j \left(c_i^j d_{ni}^j \right)^{-\theta}.$$

This value summarizes, for country n, the access to production technologies in sector j. Its value will be higher if in sector j, country n's trading partners have high productivity (T_i^j) or low cost (c_i^j). It will also be higher if the trade costs that country n faces in this sector are low. Standard steps lead to the familiar result that the price of good j in country n is simply

$$p_n^j = \Gamma \left(\Phi_n^j \right)^{-\frac{1}{\theta}}, \tag{10}$$

where $\Gamma = \left[\mathbf{\Gamma} \left(\frac{\theta+1-\varepsilon}{\theta} \right) \right]^{\frac{1}{1-\varepsilon}}$, with $\mathbf{\Gamma}$ the Gamma function. The consumption price index in country n is then:

$$P_n = B_n \left(\prod_{j=1}^{J} (p_n^j)^{\omega_j} \right)^{\xi_n} (p_n^{J+1})^{1-\xi_n}, \tag{11}$$

where $B_n = \xi_n^{-\xi_n} (1 - \xi_n)^{-(1-\xi_n)}$.

Both capital and labor are mobile across sectors and immobile across countries, and trade is balanced. The budget constraint (or the resource constraint) of the consumer is thus given by

$$\sum_{j=1}^{J+1} p_n^j Y_n^j = w_n L_n + r_n K_n, \tag{12}$$

where K_n and L_n are the endowments of capital and labor in country n.

Given the set of prices $\{w_n, r_n, P_n, \{p_n^j\}_{j=1}^{J+1}\}_{n=1}^{N}$, we first characterize the optimal allocations from final demand. Consumers maximize utility (9) subject to the budget constraint (12). The first order conditions associated with this optimization problem imply the following final demand:

$$p_n^j Y_n^j = \xi_n \omega_j (w_n L_n + r_n K_n), \text{ for all } j = \{1, .., J\}$$

and

$$p_n^{J+1} Y_n^{J+1} = (1 - \xi_n)(w_n L_n + r_n K_n).$$

4.2.2 Production Allocation and Market Clearing

Let Q_n^j denote the total sectoral demand in country n and sector j. Q_n^j is used for both final consumption and as intermediate inputs in domestic production of all sectors. Denote by $X_n^j = p_n^j Q_n^j$ the total *spending* on the sector j goods in country n, and by X_{ni}^j country n's total spending on sector j goods coming from country i, i.e. n's imports of j from country i. The EK structure in each sector j delivers the standard result that the probability of importing good q from country i, π_{ni}^j is equal to the share of total spending on goods coming from country i, X_{ni}^j / X_n^j, and is given by:

$$\frac{X_{ni}^j}{X_n^j} = \pi_{ni}^j = \frac{T_i^j \left(c_i^j d_{ni}^j\right)^{-\theta}}{\Phi_n^j}. \tag{13}$$

The market clearing condition for expenditures on sector j in country n is:

$$p_n^j Q_n^j = p_n^j Y_n^j + \sum_{k=1}^{J} (1-\beta_k)\gamma_{j,k} \left(\sum_{i=1}^{N} \pi_{in}^k p_i^k Q_i^k \right) + (1-\beta_{J+1})\gamma_{j,J+1} p_n^{J+1} Q_n^{J+1}.$$

Total expenditure in sector $j = 1, ..., J+1$ in country n, $p_n^j Q_n^j$, is the sum of (i) domestic final consumption expenditure $p_n^j Y_n^j$; (ii) expenditure on sector j goods as intermediate inputs in all the traded sectors $\sum_{k=1}^{J} (1-\beta_k)\gamma_{j,k} (\sum_{i=1}^{N} \pi_{in}^k p_i^k Q_i^k)$, and (iii) expenditure on the j's sector intermediate inputs in the domestic non-traded sector $(1-\beta_{J+1})\gamma_{j,J+1} p_n^{J+1} Q_n^{J+1}$. These market clearing conditions summarize two important features of the world economy captured by our model: complex international production linkages, as much of world trade is in intermediate inputs, and a good crosses borders multiple times before being consumed (Hummels, Ishii and Yi 2001); and two-way input linkages between the tradeable and the nontradeable sectors.

In each tradeable sector j, some goods q are imported from abroad and some goods q are exported to the rest of the world. Country n's exports in sector j are given by $EX_n^j = \sum_{i=1}^{N} \mathbb{1}_{i \neq n} \pi_{in}^j p_i^j Q_i^j$, and its imports in sector j are given by $IM_n^j = \sum_{i=1}^{N} \mathbb{1}_{i \neq n} \pi_{ni}^j p_n^j Q_n^j$, where $\mathbb{1}_{i \neq n}$ is the indicator function. The total exports of country n are then $EX_n = \sum_{j=1}^{J} EX_n^j$, and total imports are $IM_n = \sum_{j=1}^{J} IM_n^j$. Trade balance requires that for any country n, $EX_n - IM_n = 0$.

Given the total production revenue in tradeable sector j in country n, $\sum_{i=1}^{N} \pi_{in}^j p_i^j Q_i^j$, the optimal sectoral factor allocations must satisfy

$$\sum_{i=1}^{N} \pi_{in}^j p_i^j Q_i^j = \frac{w_n L_n^j}{\alpha_j \beta_j} = \frac{r_n K_n^j}{(1-\alpha_j)\beta_j}.$$

For the nontradeable sector $J+1$, the optimal factor allocations in country n are simply given by

$$p_n^{J+1} Q_n^{J+1} = \frac{w_n L_n^{J+1}}{\alpha_{J+1}\beta_{J+1}} = \frac{r_n K_n^{J+1}}{(1-\alpha_{J+1})\beta_{J+1}}.$$

Finally, the feasibility conditions for factors are given by, for any n,

$$\sum_{j=1}^{J+1} L_n^j = L_n \text{ and } \sum_{j=1}^{J+1} K_n^j = K_n.$$

4.3 Estimation, Calibration, and Solution

The core implementation step of this model is the estimation of the sector-level technology parameters T_n^j for a large set of countries. The technology parameters in the tradeable sectors relative to a reference country (the U.S.) are estimated using data on sectoral output and bilateral trade. The procedure relies on fitting a structural gravity equation implied by the model. Intuitively, if controlling for the typical gravity determinants of trade, a country spends relatively more on domestically produced goods in a particular sector, it is revealed to have either a high relative productivity or a low relative unit cost in that sector. The procedure then uses data on factor and intermediate input prices to net out the role of factor costs, yielding an estimate of relative productivity. This step also produces estimates of bilateral, sector-level trade costs d_{ni}^j. The next step is to estimate the technology parameters in the tradeable sectors for the U.S.. This procedure requires directly measuring TFP at the sectoral level using data on real output and inputs, and then correcting measured TFP for selection due to trade. Third, the nontradeable technology for all countries is calibrated using the first-order condition of the model and the relative prices of nontradeables observed in the data. The detailed procedures for all three steps are described in Levchenko and Zhang (2011) and reproduced in Appendix A.

Estimation of sectoral productivity parameters T_n^j and trade costs d_{ni}^j requires data on total output by sector, as well as sectoral data on bilateral trade, that are described at the beginning of Section 3. Productivity and trade cost estimation requires an assumption on the dispersion parameter θ. We pick the value of $\theta = 8.28$, which is the preferred estimate of EK, and in addition assume that it does not vary across sectors.[6]

[6] There are no reliable estimates of how θ varies across sectors, and thus we do not model this variation. Shikher (2004, 2005, 2011), Eaton, Kortum, Neiman and Romalis (2011), and Burstein and Vogel (2012), among others, follow the same approach of assuming the same θ across sectors. Caliendo and Parro (2010) use tariff data and triple differencing to estimate sector-level θ. However, their approach may impose too much structure and/or be dominated by measurement error: at times the values of θ they estimate are negative. In addition, in each sector the restriction that $\theta > \varepsilon - 1$ must be satisfied, and it is not clear whether Caliendo and Parro (2010)'s estimated sectoral θ's meet this restriction in every case. Our approach is thus conservative by being agnostic on this variation across sectors. It is also important to assess how the results below are affected by the value of this parameter. One may be especially concerned about how the results change under lower values of θ. Lower θ implies greater within-sector heterogeneity in the random productivity draws. Thus, trade flows become less sensitive to the costs of the input bundles (c_i^j), and the gains from intra-sectoral trade become larger relative to the gains from inter-sectoral trade. In Levchenko and Zhang (2011), we estimated the sectoral productivities for a sample of 75 countries assuming instead a value of $\theta = 4$, which has been advocated by Simonovska and Waugh (2011) and is at or near the bottom of the range that has been used in the literature. Overall, the results are remarkably similar. The correlation between estimated T_i^j's under $\theta = 4$ and under $\theta = 8.28$ is above 0.95, and there is actually somewhat greater variability in T_i^j's under $\theta = 4$.

In order to implement the model numerically, we must in addition calibrate the following sets of parameters: (i) preference parameters ω_j and ξ_n; (ii) production function parameters ε, α_j, β_j, $\gamma_{k,j}$ for all sectors j and k; (iii) country factor endowments L_n and K_n.

The share of expenditure on traded goods, ξ_n in each country is sourced from Yi and Zhang (2010), who compile this information for 30 developed and developing countries. For countries unavailable in the Yi and Zhang data, values of ξ_n are imputed based on fitting a simple linear relationship to log PPP-adjusted per capita GDP from the Penn World Tables. The fit of this simple bivariate linear relationship is quite good, with the R^2 of 0.55. The taste parameters for tradeable sectors ω_j were set equal to final consumption expenditure shares in the U.S. sourced from the U.S. Input-Output matrix.

The production function parameters α_j and β_j are estimated using the output, value added, and wage bill data from EUROSTAT and UNIDO. To compute α_j for each sector, we calculate the share of the total wage bill in value added, and take a simple median across countries (taking the mean yields essentially the same results). To compute β_j, we calculate the ratio of value added to total output for each country and sector, and take the median across countries.

The intermediate input coefficients $\gamma_{k,j}$ are obtained from the Direct Requirements Table for the United States. We use the 1997 Benchmark Detailed Make and Use Tables (covering approximately 500 distinct sectors), as well as a concordance to the ISIC Revision 3 classification to build a Direct Requirements Table at the 2-digit ISIC level. The Direct Requirements Table gives the value of the intermediate input in row k required to produce one dollar of final output in column j. Thus, it is the direct counterpart of the input coefficients $\gamma_{k,j}$. Note that we assume these to be the same in all countries.[7] In addition, we use the U.S. I-O matrix to obtain α_{J+1} and β_{J+1} in the nontradeable sector.[8] The elasticity of substitution between varieties within each tradeable sector, ε, is set to 4 (of course, as is well known, this value plays no role in this model, beyond affecting the value of the constant Γ). Appendix Table A2 lists the key parameter values for each sector: α_j, β_j, the share of nontradeable inputs in total inputs $\gamma_{J+1,j}$, and the taste parameter ω_j.

The total labor force in each country, L_n, and the total capital stock, K_n, are computed based on the Penn World Tables 6.3. Following the standard approach in the literature (see, e.g. Hall and Jones 1999, Bernanke and Gürkaynak 2001, Caselli 2005), the total labor force is calculated from

[7] di Giovanni and Levchenko (2010) provide suggestive evidence that at such a coarse level of aggregation, Input-Output matrices are indeed similar across countries. To check robustness of the results, Levchenko and Zhang (2011) collected country-specific I-O matrices from the GTAP database. Productivities computed based on country-specific I-O matrices were very similar to the baseline values, with the median correlation of 0.98, and all but 3 out of 75 countries with a correlation of 0.93 or above, and the minimum correlation of 0.65.

[8] The U.S. I-O matrix provides an alternative way of computing α_j and β_j. These parameters calculated based on the U.S. I-O table are very similar to those obtained from UNIDO, with the correlation coefficients between them above 0.85 in each case. The U.S. I-O table implies greater variability in α_j's and β_j's across sectors than does UNIDO.

data on the total GDP per capita and per worker.[9] The total capital stock is calculated using the perpetual inventory method that assumes a depreciation rate of 6%: $K_{n,t} = (1-0.06)K_{n,t-1} + I_{n,t}$, where $I_{n,t}$ is total investment in country n in period t. For most countries, investment data start in 1950, and the initial value of K_n is set equal to $I_{n,0}/(\gamma + 0.06)$, where γ is the average growth rate of investment in the first 10 years for which data are available.

Given the estimated sectoral productivities, factor endowments, trade costs, and model parameters, we solve the system of equations defining the equilibrium under the baseline values. The algorithm for solving the model is described in Levchenko and Zhang (2011). Then, to compute the true gains from trade in the model we re-solve it under the assumption that each country is in autarky, and compare the autarky welfare to the welfare in the baseline. Finally, we compute in the baseline model a range of values that are (potentially) observable in the real-world data, and that are then used to compute gains from trade according to a range of formulas.

4.4 Model Fit

Table 2 compares the wages, returns to capital, and the trade shares in the baseline model solution and in the data. The top panel shows that mean and median wages implied by the model are very close to the data. The correlation coefficient between model-implied wages and those in the data is above 0.99. The second panel performs the same comparison for the return to capital. Since it is difficult to observe the return to capital in the data, we follow the approach adopted in the estimation of T_n^j's and impute r_n from an aggregate factor market clearing condition: $r_n/w_n = (1-\alpha)L_n/(\alpha K_n)$, where α is the aggregate share of labor in GDP, assumed to be 2/3. Once again, the average levels of r_n are very similar in the model and the data, and the correlation between the two is about 0.95.

Next, we compare the trade shares implied by the model to those in the data. The third panel of Table 2 reports the spending on domestically produced goods as a share of overall spending, π_{nn}^j. These values reflect the overall trade openness, with lower values implying higher international trade as a share of absorption. Though we under-predict overall trade slightly (model π_{nn}^j's tend to be higher), the averages are quite similar, and the correlation between the model and data values is 0.92. Finally, the bottom panel compares the international trade flows in the model and the data. The averages are very close, and the correlation between model and data is about 0.91.

We conclude from this exercise that our model matches quite closely the relative incomes of countries as well as bilateral and overall trade flows observed in the data.

[9] Using the variable name conventions in the Penn World Tables, $L_n = 1000 * pop * rgdpch/rgdpwok$.

4.5 Comparison of Candidate Gains from Trade Formulas

All throughout, welfare is defined as the indirect utility function. Straightforward steps can be used to show that indirect utility in each country i is equal to total income divided by the price level. Since the model is competitive, total income equals the total returns to factors of production. Expressed in per capita terms welfare is thus:

$$\frac{w_i + r_i k_i}{P_i},$$

where $k_i = K_i/L_i$ is capital per worker, and P_i comes from equation (11).

The true gains from trade in this model are computed by solving the baseline model, calculating welfare, and comparing this welfare to a counterfactual scenario in which all countries are assumed to be in autarky. The question is, from a pragmatic perspective, whether there is a formula based on quantities potentially observable in the data that can approximate the true gains from trade well. We go through a sequence of candidate formulas, with different data requirements, to see which data are essential to reliably compute the gains from trade using a formula.

To facilitate comparisons, Table 3 summarizes the formulas, the underlying models that deliver those formulas as the exact gains from trade, and the data requirements for implementing each formula. We proceed in the increasing order of data requirements. The simplest is a one-sector formula that does not distinguish explicitly between the traded and the non-traded sector, and relies only on observing the aggregate trade volume relative to the total gross output. To compute it, one would only need to collect data on aggregate imports and exports, as well as the gross output in the economy. The latter may not be as readily available as total GDP, but could be approximated, for instance, by "grossing up" total GDP by a factor that corresponds to one minus the share of intermediates in total output. One would then take these data and apply the formula $\pi_{ii}^{-1/\theta} - 1$ where $\pi_{ii} = (\text{OUTPUT} - \text{EXPORTS})/(\text{OUTPUT} - \text{EXPORTS} + \text{IMPORTS})$.

This formula can be augmented to account for intermediate input linkages. Since it does not distinguish between the tradeable and the non-tradeable sectors, the strength of the intermediate input linkages simply becomes the share of value added to total gross output, and thus to augment this formula to include input linkages, no extra data are required beyond aggregate value added, which is just total GDP. The augmented formula becomes $\pi_{ii}^{-1/\beta\theta} - 1$, where $\beta = $ VALUE ADDED/OUTPUT in the aggregate economy.

To take explicit account of the fact that much of the domestic GDP is in the non-tradeable sector, one could use some information on the output of tradeables, and the share of tradeables in value added, to refine this formula. The data requirements for implementing this formula call only for aggregate imports and exports, the gross output of the tradeable sector, and the share of tradeables in total value added. The formula for the gains from trade is then given by (7) (where now ξ_i can vary by country).

This formula can also be augmented with intermediate input linkages, under the assumption that these input linkages are strictly within-sector. That is, the tradeable sector only uses tradeable intermediates, and the non-tradeable only non-tradeable intermediates. The formula then becomes:

$$\pi_{ii}^{-\frac{\xi_i}{\beta\theta}} - 1. \tag{14}$$

In terms of measuring intermediate input linkages, there are now a couple of ways to proceed. The first, less data-demanding way is to measure the linkages by the share of value added to output in the tradeable and the non-tradeable sectors, exactly as above. That approach would give the input linkages the maximum strength, but may potentially overstate the true gains from trade. This is because a significant share of input usage in the tradeable sector goes to non-tradeables (see Appendix Table A2), and those are not subject to comparative-advantage driven gains from trade. Thus, measuring the strength of input linkages simply by the share of value added in total output overstates international trade's benefits acting through this channel.

The second approach to calibrating the linkages is to isolate the share of intermediate input usage in the tradeable sector that goes to tradeables. This will quantify the intermediate-input driven gains more precisely, but raises the data requirements. In particular, now we need to know not just the total value added and total output, but the breakdown of the input usage into tradeables and non-tradeables. In other words, we now need the (2-sector) input-output table, and the β in (14) is now one minus the share of spending on tradeable intermediate inputs in the tradeable sector gross output.

Finally, one can implement explicitly multi-sector formulas. The data requirements are much higher in this case, as we now need imports, exports, and gross output at the sector level for tradeables. Without input-output linkages, the formula for the gains from trade in an explicitly multi-sector context is given by (8) (where again ξ_i is now country-specific).

This formula can be extended to incorporate within-sector intermediate linkages, in the two ways that parallel the two-sector tradeable-non-tradeable formula. The first way, that does not require any additional data, is to proxy for the strength of intermediate input linkages by the share of value added to total output in each sector. In that case, the gains from trade formula becomes

$$\prod_{j=1}^{J} \left(\pi_{ii}^{j}\right)^{-\frac{\xi_i \omega_j}{\beta_j \theta}} - 1, \tag{15}$$

where β_j is simply the value added over output in sector j.

Once again, this approach will overstate the contribution of international trade through the input channel because some intermediates used in the tradeable sector are non-tradeable, and thus not subject to the productivity gains from trade. To confine the impact of intermediate inputs to the tradeable inputs only, one must use the full multi-sector input-output table, and for each sector j set β_j to one minus the share of tradeable inputs in total output in each sector. Note that this last option is the most data-intensive, requiring the complete input-output table in addition to data on sectoral output and cross-border trade. While formula (15) does capture some important intermediate input linkages, it still mismeasures the nature of the real-world input linkages because it assumes that all of the linkages are within-sector. That is, the Textile sector uses only Textiles as intermediate inputs, the Apparel sector uses only Apparel, and so on. Thus, to the extent that cross-sectoral input linkages are important – for instance, Apparel uses a great deal of Textiles as intermediates – this formula will miss those. It is ultimately a quantitative question, given the observed input-output matrices, how much those linkages matter.

Before moving on to the results, we pause to clarify the nature of the exercise. Each of the formulas offered above can be calculated based on observable data, and thus would not require the researcher to implement, estimate/calibrate, and perform counterfactuals in, a complete quantitative model. Thus, the potential promise of these formulas is to enable researchers to estimate gains from trade quickly and easily.

For each of the formulas offered above, there is a model under which that formula represents the exact gains from trade. However, all of these behind-the-scenes models are simplifications, both with respect to reality, and with respect to the full quantitative model. The question we answer below is whether, for the purposes of computing the gains from trade, any of these simplified models represent reasonable approximations to (i) the full quantitative model and (ii) the real world. Our exercise provides the complete answer to (i). With respect to (ii), our results are of course more suggestive, but here the discipline of the estimation and calibration of the quantitative model to real world data, and the match to observed trade flows is helpful.

Table 4 reports the summary statistics for the gains from trade in our 79-country model. The top row reports the true gains from trade implied by the model. On average, the gains are 7.22%, ranging from 1% to 22% in this sample of countries. The rest of the table presents the results of using the formulas to compute the gains from trade. By and large, the message from this table is that the formulas tend to under-predict the gains from trade, often by a significant margin. At the extreme, the one-sector, no-intermediate formula delivers the gains from trade of 1.88% at the mean, or less than a third of the true model gains. The range of gains across countries predicted by these formulas is also much narrower than the range of true gains. Thus, the formulas tend to understate both the mean, and the variation in the gains from trade.

The one exception to this regularity is the multi-sector formula that assumes all of the intermediate good usage to come from the sector itself (second-to-last row). That formula overstates both the average gains, and the dispersion. As we mention above, this overstatement is due to the fact that in the real world and in the quantitative model, a large share (about 40%) of intermediates are actually non-tradeables. These non-tradeables do not benefit from imported, more productive varieties when the country opens to trade, and thus the formula featuring the total intermediate input usage in effect overstates the productivity-enhancing impact of imported intermediates.

Table 5 summarizes the performance of the different formulas relative to the true gains. The first 4 columns present the summary statistics for the proportional deviation of the formula-implied gains from the true ones. We see that most of the formulas understate the true gains by 40 to 70%, with the exception of the second-to-last formula, which overstates them by nearly 50%. For most formulas, deviations can be quite large for individual countries. At the extreme, some formulas miss 80 to 90% of the true gains from trade, while the multi-sector, total linkage formula overstates the gains for one of the countries by 100%. We can also see that the bias is quite systematic, with most of the formulas understating the gains for every single country, while one of the formulas overstates the gains for every country.

The clear winner is in the last row. The formula that features multiple sectors, intermediate inputs, and that only uses tradeable intermediate inputs to reflect the intermediate input linkages gets closest to the true gains. This is not surprising, given that it is the most sophisticated and data-intensive approach. However, what is striking is how much better the results are compared to the simpler alternatives. On average, the last formula understates the gains by a relatively modest 11%. The range of its deviations from the true model gains is narrow, from -0.38 to 0.23, and nearly symmetric around zero. Columns (5) and (6) of Table 5 report the mean squared error and the mean absolute error for each formula relative to the true gains. The last row emerges as the clear winner, with both the mean squared error and the mean absolute errors nearly an order of magnitude lower than every other formula, and with the mean absolute error being nearly 2 times smaller than the next best formula.

We conclude from this exercise that, not surprisingly, in order to obtain reliable results for the gains from trade, it is essential to both (i) use sector-level data, and (ii) reflect input-output linkages carefully using input-output tables. Point (i) thus confirms analytical results in Section 2 and the data-driven exercise in Section 3.

Finally, it turns out that the second-best formula is not also the second-most complicated or data-intensive. Instead, the second-best formula is a two sector (tradeable/non-tradeable) formula with input linkages, that features the ratio of tradeable value added to tradeable output as the measure of linkages. The average deviation for that formula from the true gains is also -11%, just as for the best one. It is also the only formula aside from the winner that produces both positive and negative deviations from true gains. Finally, its mean squared and absolute errors are substantially lower than the rest of the field, though still much higher than the winner's.

This potentially suggests that one could get fairly close to the true gains with much lower data requirements. To compute that formula, one needs total imports and exports, and total tradeable and non-tradeable output and value added. No sectoral output and trade flows, and no input-output matrices are required to implement that formula. However, it should be clear from the discussion above that this a case of "two wrongs make a right." As shown throughout the paper, ignoring the dispersion among tradeable sectors leads to a systematic understatement of the true gains. On the other hand, attributing all of the intermediate usage in the tradeable sector to tradeables overstates the gains from trade coming from input linkages. Those two opposing forces apparently largely cancel out. Since this cancelling out is clearly not a general feature of all models and calibrations, we would be cautious to conclude that one could use this simpler and easier to calculate formula as a good surrogate when sectoral data and IO matrices are not available.

4.6 The Role of Ricardian Comparative Advantage

The simple analytical framework in Section 2 illustrates that the one-sector formula will understate the gains by more the stronger is Ricardian comparative advantage – that is, the greater is the dispersion in T_i^j's. Column (7) of Table 5 reports the correlation between the error of the formula relative to the true gains and a simple heuristic measure of how strong is a country's comparative advantage, namely the coefficient of variation in sectoral productivities relative to the world frontier. Countries with a high coefficient of variation are considered to have "strong comparative advantage," in the sense that their technology has high relative dispersion across sectors. The intuition we built using analytical results suggests that in countries with a high coefficient of variation in technology the one-sector formulas will understate the gains by more: a negative correlation. This is confirmed in the quantitative exercise: the correlations between the strength of comparative advantage and the errors in the one-sector formulas are all negative and highly significant. These negative correlations disappear when we go to the multi-sector formulas (bottom 3 rows). Since the multi-sector formulas take proper account of sectoral heterogeneity, in those formulas there is no longer a systematically greater understatement of the gains for countries with stronger comparative advantage.

The negative relationship is depicted graphically in Figure 5. In countries with greater dispersion in sectoral productivity, the one-sector formula understates the true gains by more. This is the quantitative, productivity estimates-based counterpart of Figure 1(b) in the analytical section.

5 Conclusion

The discovery that several models with very different micro-structures deliver closed-form expressions for the gains from trade that are both (i) identical to each other and (ii) computable based on a small number of observables is a potentially transformative one in the realm of policy research. If made operational, this approach can at the same time dramatically lower barriers to quantitative evaluation of trade policy, and render the results much more general.

This paper takes a step toward assessing the applicability of this approach to real-world policy analysis. It starts from the observation that when there are multiple sectors, a one-sector formula that only incorporates information on the total trade volume relative to absorption systematically understates the true gains from trade. The basic economic intuition for this result can be gleaned from the Ricardian motive for trade: larger sectoral productivity differences will raise the gains from trade even holding constant the overall trade volume. We then use actual data on sectoral and aggregate absorption shares in the manufacturing sector in a sample of 79 countries to show that this effect is large quantitatively: the multi-sector formula implies gains from trade that are 30% higher on average than the gains according to the one-sector formula, and as much as 100% higher in countries with large dispersion in sectoral trade shares.

Finally, we set up a quantitative model with many features relevant in the real world, such as multiple factors of production, a non-tradeable sector, and the full set of input-output linkages between sectors. The model is implemented using actual data on production functions, input-output matrices, and trade flows. It matches relative incomes and bilateral and overall trade flows quite closely. We evaluate whether there are sufficient statistic formulas that, when computed inside this model, can approximate well the true gains from trade in the model. It turns out that augmented formulas that take into account input-output linkages but not sectoral trade shares, or sectoral trade shares but not input-output linkages understate the true model gains substantially. However, an appropriate combination of information on sectoral trade flows and input linkages performs quite well relative to others, understating the true model gains by only 11% on average.

We conclude that sectoral heterogeneity in productivity and trade volumes has a first-order impact on the gains from trade, and that it should be taken into account in exercises that use the sufficient statistic approach to compute the gains from trade.

Appendix A Procedure for Estimating T_n^j, d_{ni}^j, and ω_j

This appendix reproduces from Levchenko and Zhang (2011) the details of the procedure for estimating technology, trade costs, and taste parameters required to implement the model. Interested readers should consult that paper for further details on estimation steps and data sources.

A.1 Tradeable Sector Relative Technology

We now focus on the tradeable sectors. Following the standard EK approach, first divide trade shares by their domestic counterpart:

$$\frac{\pi_{ni}^j}{\pi_{nn}^j} = \frac{X_{ni}^j}{X_{nn}^j} = \frac{T_i^j \left(c_i^j d_{ni}^j\right)^{-\theta}}{T_n^j \left(c_n^j\right)^{-\theta}},$$

which in logs becomes:

$$\ln\left(\frac{X_{ni}^j}{X_{nn}^j}\right) = \ln\left(T_i^j (c_i^j)^{-\theta}\right) - \ln\left(T_n^j (c_n^j)^{-\theta}\right) - \theta \ln d_{ni}^j.$$

Let the (log) iceberg costs be given by the following expression:

$$\ln d_{ni}^j = d_k^j + b_{ni}^j + CU_{ni}^j + RTA_{ni}^j + ex_i^j + \nu_{ni}^j,$$

where d_k^j is an indicator variable for a distance interval. Following EK, we set the distance intervals, in miles, to [0, 350], [350, 750], [750, 1500], [1500, 3000], [3000, 6000], [6000, maximum). Additional variables are whether the two countries share a common border (b_{ni}^j), belong to a currency union (CU_{ni}^j), or to a regional trade agreement (RTA_{ni}^j). Following the arguments in Waugh (2010), we include an exporter fixed effect ex_i^j. Finally, there is an error term ν_{ni}^j. Note that all the variables have a sector superscript j: we allow all the trade cost proxy variables to affect true iceberg trade costs d_{ni}^j differentially across sectors. There is a range of evidence that trade volumes at sector level vary in their sensitivity to distance or common border (see, among many others, Do and Levchenko 2007, Berthelon and Freund 2008).

This leads to the following final estimating equation:

$$\ln\left(\frac{X_{ni}^j}{X_{nn}^j}\right) = \underbrace{\ln\left(T_i^j (c_i^j)^{-\theta}\right) - \theta ex_i^j}_{\text{Exporter Fixed Effect}} \underbrace{- \ln\left(T_n^j \left(c_n^j\right)^{-\theta}\right)}_{\text{Importer Fixed Effect}}$$
$$\underbrace{-\theta d_k^j - \theta b_{ni}^j - \theta CU_{ni}^j - \theta RTA_{ni}^j}_{\text{Bilateral Observables}} \underbrace{-\theta \nu_{ni}^j}_{\text{Error Term}}.$$

This equation is estimated for each tradeable sector $j = 1, ...J$. Estimating this relationship will thus yield, for each country, an estimate of its technology-cum-unit-cost term in each sector j, $T_n^j (c_n^j)^{-\theta}$, which is obtained by exponentiating the importer fixed effect. The available degrees of freedom imply that these estimates are of each country's $T_n^j (c_n^j)^{-\theta}$ relative to a reference country, which in our estimation is the United States. We denote this estimated value by S_n^j:

$$S_n^j = \frac{T_n^j}{T_{us}^j} \left(\frac{c_n^j}{c_{us}^j} \right)^{-\theta},$$

where the subscript us denotes the United States. It is immediate from this expression that estimation delivers a convolution of technology parameters T_n^j and cost parameters c_n^j. Both will of course affect trade volumes, but we would like to extract technology T_n^j from these estimates. In order to do that, we follow the approach of Shikher (2004). In particular, for each country n, the share of total spending going to home-produced goods is given by

$$\frac{X_{nn}^j}{X_n^j} = T_n^j \left(\frac{\Gamma c_n^j}{p_n^j} \right)^{-\theta}.$$

Dividing by its U.S. counterpart yields:

$$\frac{X_{nn}^j / X_n^j}{X_{us,us}^j / X_{us}^j} = \frac{T_n^j}{T_{us}^j} \left(\frac{c_n^j}{c_{us}^j} \frac{p_{us}^j}{p_n^j} \right)^{-\theta} = S_n^j \left(\frac{p_{us}^j}{p_n^j} \right)^{-\theta},$$

and thus the ratio of price levels in sector j relative to the U.S. becomes:

$$\frac{p_n^j}{p_{us}^j} = \left(\frac{X_{nn}^j / X_n^j}{X_{us,us}^j / X_{us}^j} \frac{1}{S_n^j} \right)^{\frac{1}{\theta}}. \tag{A.1}$$

The entire right-hand side of this expression is either observable or estimated. Thus, we can impute the price levels relative to the U.S. in each country and each tradeable sector.

The cost of the input bundles relative to the U.S. can be written as:

$$\frac{c_n^j}{c_{us}^j} = \left(\frac{w_n}{w_{us}} \right)^{\alpha_j \beta_j} \left(\frac{r_n}{r_{us}} \right)^{(1-\alpha_j)\beta_j} \left(\prod_{k=1}^{J} \left(\frac{p_n^k}{p_{us}^k} \right)^{\gamma_{k,j}} \right)^{1-\beta_j} \left(\frac{p_n^{J+1}}{p_{us}^{J+1}} \right)^{\gamma_{J+1,j}(1-\beta_j)}.$$

Using information on relative wages, returns to capital, price in each tradeable sector from (A.1), and the nontradeable sector price relative to the U.S., we can thus impute the costs of the input bundles relative to the U.S. in each country and each sector. Armed with those values, it is straightforward to back out the relative technology parameters:

$$\frac{T_n^j}{T_{us}^j} = S_n^j \left(\frac{c_n^j}{c_{us}^j} \right)^{\theta}.$$

A.2 Trade Costs

The bilateral, directional, sector-level trade costs of shipping from country i to country n in sector j are then computed based on the estimated coefficients as:

$$\ln \hat{d}_{ni}^j = \theta \hat{d}_k^j + \theta \hat{b}_{ni}^j + \theta \widehat{CU}_{ni}^j + \theta \widehat{RTA}_{ni}^j + \theta \widehat{ex}_i^j + \theta \widehat{\nu}_{ni}^j,$$

for an assumed value of θ. Note that the estimate of the trade costs includes the residual from the gravity regression $\theta \widehat{\nu}_{ni}^j$. Thus, the trade costs computed as above will fit bilateral sectoral trade flows exactly, given the estimated fixed effects. Note also that the exporter component of the trade costs $\widehat{ex}_i^j$ is part of the exporter fixed effect. Since each country in the sample appears as both an exporter and an importer, the exporter and importer estimated fixed effects are combined to extract an estimate of $\theta \widehat{ex}_i^j$.

A.3 Complete Estimation

So far we have estimated the levels of technology of the tradeable sectors relative to the United States. To complete our estimation, we still need to find (i) the levels of T for the tradeable sectors in the United States; (ii) the taste parameters ω_j, and (iii) the nontradeable technology levels for all countries.

To obtain (i), we use the NBER-CES Manufacturing Industry Database for the U.S. (Bartelsman and Gray 1996). We start by measuring the observed TFP levels for the tradeable sectors in the U.S.. The form of the production function gives

$$\ln Z_{us}^j = \ln \Lambda_{us}^j + \beta_j \alpha_j \ln L_{us}^j + \beta_j (1 - \alpha_j) \ln K_{us}^j + (1 - \beta_j) \sum_{k=1}^{J+1} \gamma_{k,j} \ln M_{us}^{k,j}, \tag{A.2}$$

where Λ^j denotes the measured TFP in sector j, Z^j denotes the output, L^j denotes the labor input, K^j denotes the capital input, and $M^{k,j}$ denotes the intermediate input from sector k. The NBER-CES Manufacturing Industry Database offers information on output, and inputs of labor, capital, and intermediates, along with deflators for each. Thus, we can estimate the observed TFP level for each manufacturing tradeable sector using the above equation.

If the United States were a closed economy, the observed TFP level for sector j would be given by $\Lambda_{us}^j = (T_{us}^j)^{\frac{1}{\theta}}$. In the open economies, the goods with inefficient domestic productivity draws will not be produced and will be imported instead. Thus, international trade and competition introduce selection in the observed TFP level, as demonstrated by Finicelli, Pagano and Sbracia (2013). We thus use the model to back out the true level of T^j_{us} of each tradeable sector in the United States. Here we follow Finicelli et al. (2013) and use the following relationship:

$$(\Lambda^j_{us})^\theta = T^j_{us} + \sum_{i \neq us} T^j_i \left(\frac{c^j_i d^j_{us,i}}{c^j_{us}} \right)^{-\theta}.$$

Thus, we have

$$(\Lambda^j_{us})^\theta = T^j_{us} \left[1 + \sum_{i \neq us} \frac{T^j_i}{T^j_{us}} \left(\frac{c^j_i d^j_{us,i}}{c^j_{us}} \right)^{-\theta} \right] = T^j_{us} \left[1 + \sum_{i \neq us} S^j_i \left(d^j_{us,i} \right)^{-\theta} \right]. \tag{A.3}$$

This equation can be solved for underlying technology parameters T^j_{us} in the U.S., given estimated observed TFP Λ^j_{us}, and all the S^j_i's and $d^j_{us,i}$'s estimated in the previous subsection.

Finally, we estimate the nontradeable sector TFP using the relative prices. In the model, the nontradeable sector price is given by

$$p^{J+1}_n = \Gamma (T^{J+1}_n)^{-\frac{1}{\theta}} c^{J+1}_n.$$

Since we know the aggregate price level in the tradeable sector p^T_n, c^{J+1}_n, and the relative price of nontradeables (which we take from the data), we can back out T^{J+1}_n from the equation above for all countries.

References

Alvarez, Fernando and Robert E., Jr. Lucas, "General Equilibrium Analysis of the Eaton-Kortum Model of International Trade," *Journal of Monetary Economics*, September 2007, *54* (6), 1726–1768.

Arkolakis, Costas, Arnaud Costinot, and Andrés Rodríguez-Clare, "New Trade Models, Same Old Gains?," *American Economic Review*, February 2012, *102* (1), 94–130.

____ , ____ , Dave Donaldson, and Andrés Rodríguez-Clare, "The Elusive Pro-Competitive Effects of Trade," March 2012. mimeo, Yale University, MIT, and UC Berkeley.

____ , Svetlana Demidova, Peter J. Klenow, and Andrés Rodríguez-Clare, "Endogenous Variety and the Gains from Trade," *American Economic Review Papers and Proceedings*, May 2008, *98* (2), 444–450.

Atkeson, Andrew and Ariel Burstein, "Innovation, Firm Dynamics, and International Trade," *Journal of Political Economy*, June 2010, *118* (3), 433–484.

Bartelsman, Eric J. and Wayne Gray, "The NBER Manufacturing Productivity Database," October 1996. NBER Technical Working Paper 205.

Bernanke, Ben and Refet Gürkaynak, "Is Growth Exogenous? Taking Mankiw, Romer, and Weil Seriously," *NBER Macroeconomics Annual*, 2001, *16*, 11–57.

Berthelon, Matias and Caroline Freund, "On the Conservation of Distance in International Trade," *Journal of International Economics*, July 2008, *75* (2), 310–320.

Burstein, Ariel and Javier Cravino, "Measured Aggregate Gains from International Trade," January 2012. mimeo, UCLA.

____ and Jonathan Vogel, "International Trade, Technology, and the Skill Premium," April 2012. mimeo, UCLA and Columbia University.

Caliendo, Lorenzo and Fernando Parro, "Estimates of the Trade and Welfare Effects of NAFTA," January 2010. mimeo, University of Chicago.

Caselli, Francesco, "Accounting for Cross-Country Income Differences," in Steven Durlauf Philippe Aghion, ed., *Handbook of Economic Growth*, Vol. 1, Elsevier-North Holland, 2005, chapter 9, pp. 679–741.

Chari, V.V., Patrick J. Kehoe, and Ellen R. McGrattan, "Are structural VARs with long-run restrictions useful in developing business cycle theory?," *Journal of Monetary Economics*, November 2008, *55* (8), 1337–1352.

Corcos, Gregory, Massimo Del Gatto, Giordano Mion, and Gianmarco I.P. Ottaviano, "Productivity and Firm Selection: Quantifying the 'New' Gains from Trade," *Economic Journal*, June 2012, *122* (561), 754–798.

Costinot, Arnaud and Andrés Rodríguez-Clare, "Trade Theory with Numbers: Quantifying the Consequences of Globalization," September 2012. Forthcoming, Handbook of International Economics.

de Blas, Beatriz and Katheryn Russ, "Understanding Markups in the Open Economy," July 2012. mimeo, Universidad Autónoma de Madrid and UC Davis.

di Giovanni, Julian and Andrei A. Levchenko, "Putting the Parts Together: Trade, Vertical Linkages, and Business Cycle Comovement," *American Economic Journal: Macroeconomics*, April 2010, *2* (2), 95–124.

____ and ____ , "Firm Entry, Trade, and Welfare in Zipf's World," *Journal of International Economics*, March 2013, *89* (2), 283–296.

di Giovanni, Julian, Andrei A. Levchenko, and Jing Zhang, "The Global Welfare Impact of China: Trade Integration and Technological Change," May 2013. Forthcoming, *American Economic Journal: Macroeconomics.*

Do, Quy-Toan and Andrei A. Levchenko, "Comparative Advantage, Demand for External Finance, and Financial Development," *Journal of Financial Economics*, December 2007, *86* (3).

Eaton, Jonathan and Samuel Kortum, "Technology, Geography, and Trade," *Econometrica*, September 2002, *70* (5), 1741–1779.

____ , ____ , Brent Neiman, and John Romalis, "Trade and the Global Recession," January 2011. NBER Working Paper No. 16666.

Feenstra, Robert C., "Measuring the Gains From Trade Under Monopolistic Competition," *Canadian Journal of Economics*, February 2010, *43* (1), 1–28.

Finicelli, Andrea, Patrizio Pagano, and Massimo Sbracia, "Ricardian Selection," *Journal of International Economics*, January 2013, *89* (1), 96–109.

Hall, Robert and Charles Jones, "Why Do Some Countries Produce So Much More Output per Worker then Others," *Quarterly Journal of Economics*, 1999, *114*, 83–116.

Hauk, William and Romain Wacziarg, "A Monte Carlo study of growth regressions," *Journal of Economic Growth*, June 2009, *14* (2), 103–147.

Hummels, David, Jun Ishii, and Kei-Mu Yi, "The Nature and Growth of Vertical Specialization in World Trade," *Journal of International Economics*, June 2001, *54*, 75–96.

Imbs, Jean and Isabelle Méjean, "Trade Elasticities," August 2011. Mimeo, Paris School of Economics and Ecole Polytechnique.

Levchenko, Andrei A. and Jing Zhang, "The Evolution of Comparative Advantage: Measurement and Welfare Implications," February 2011. NBER Working Paper No. 16806.

____ and ____ , "Comparative Advantage and the Welfare Impact of European Integration," *Economic Policy*, October 2012, *27* (72), 567–602.

Melitz, Marc J. and Stephen J. Redding, "Firm Heterogeneity and Aggregate Welfare," March 2013. mimeo, Harvard University and Princeton University.

Ossa, Ralph, "Why Trade Matters After All," May 2012. NBER Working Paper No. 18113.

Ricardo, David, *On the Principles of Political Economy and Taxation*, Vol. 1 of *Works and Correspondence of David Ricardo*, Cambridge University Press, 1817.

Shikher, Serge, "Putting industries into the Eaton-Kortum model," July 2004. Forthcoming, *Journal of International Trade and Economic Development.*

____ , "Accounting for International Trade," August 2005. mimeo, Suffolk University.

____ , "Capital, technology, and specialization in the neoclassical model," *Journal of International Economics*, March 2011, *83* (2), 229–242.

Simonovska, Ina and Michael E. Waugh, "The Elasticity of Trade: Estimates and Evidence," February 2011. UC Davis Working Paper No. 11-2.

Waugh, Michael, "International Trade and Income Differences," *American Economic Review*, December 2010, *100* (5), 2093–2124.

Yi, Kei-Mu and Jing Zhang, "Structural Change in an Open Economy," April 2010. Mimeo, Federal Reserve Bank of Philadelphia and University of Michigan.

Table 1. Welfare Gains from Trade Implied by the One- and Multi-Sector Formulas

	Mean	Std. Dev.	Min	Max
$\pi_{ii}^{-\frac{\xi}{\theta}} - 1$	0.022	0.010	0.005	0.051
$\prod_{j=1}^{J}\left(\pi_{ii}^{j}\right)^{-\frac{\xi\omega_j}{\theta}} - 1$, Equal-weighted	0.035	0.017	0.006	0.067
$\prod_{j=1}^{J}\left(\pi_{ii}^{j}\right)^{-\frac{\xi\omega_j}{\theta}} - 1$, Expenditure-weighted	0.030	0.016	0.005	0.080
Pct difference, equal-weighted to one-sector	0.567	0.421	-0.069	1.706
Pct difference, exp.-weighted to one-sector	0.321	0.222	0.032	0.990

Notes: This table presents the summary statistics for the gains from trade implied by the one-sector formula, the equal-weighted multi-sector formula, and the expenditure-weighted multi-sector formula in the sample of 79 countries. The bottom two rows present the summary statistics for the proportional deviations of the multi-sector formulas relative to the one-sector one.

Table 2. The Fit of the Baseline Model with the Data

		model	data
Wages:			
	mean	0.387	0.351
	median	0.131	0.150
	corr(model, data)	*0.994*	
Return to capital:			
	mean	0.881	0.939
	median	0.664	0.698
	corr(model, data)	*0.946*	
π_{nn}^{j}			
	mean	0.621	0.569
	median	0.682	0.609
	corr(model, data)	*0.918*	
π_{ni}^{j}, $i \neq n$			
	mean	0.0052	0.0056
	median	0.0001	0.0001
	corr(model, data)	*0.908*	

Notes: This table reports the means and medians of wages relative to the U.S. (top panel); return to capital relative to the U.S. (second panel), share of domestically produced goods in overall spending (third panel), and share of goods from country i in overall spending (bottom panel) in the model and in the data. Wages and return to capital in the data are calculated as described in Appendix A.

Table 3. Formulas, Underlying Models, and Data Requirements

Formula	Model	Data requirements
$\pi_{ii}^{-\frac{1}{\theta}} - 1$	One-sector, no linkages	(1): Aggregate imports, exports, gross output
$\pi_{ii}^{-\frac{1}{\beta\theta}} - 1$	One-sector, with linkages	(2): (1) + aggregate value added
$\pi_{ii}^{-\frac{\xi_i}{\theta}} - 1$	Tradeable and non-tradeable sectors, no linkages	(3): (2) + breakdown of gross output and value added into tradeables and non-tradeables
$\pi_{ii}^{-\frac{\xi_i}{\beta\theta}} - 1$:	Tradeable and non-tradeable sectors, within-sector linkages	
β = total interm. use		(3)
β = within-sector interm. use		(4): (3) + T/NT IO matrix
$\prod_{j=1}^{J}\left(\pi_{ii}^{j}\right)^{\frac{-\xi_i\omega_j}{\theta}} - 1$	Multi-sector, no linkages	(5): sector-level output, imports, and exports, share of non-tradeables in aggregate value added
$\prod_{j=1}^{J}\left(\pi_{ii}^{j}\right)^{\frac{-\xi_i\omega_j}{\beta_j\theta}} - 1$:	Multi-sector, within-sector linkages	
β_j = total linkages		(6): (5) + sectoral value added
β_j = tradeable only sectoral linkages		(7): (6) + full IO matrix

Notes: This table presents the formulas for the gains from trade, underlying models that yield those formulas, and the data requirements necessary to compute them.

Table 4. Welfare Gains from Trade

Formula	Mean	Std. Dev.	Min	Max	No. Countries
True Gains	7.22	4.92	1.05	22.12	79
$\pi_{ii}^{-\frac{1}{\theta}} - 1$	1.88	0.83	0.35	4.11	79
$\pi_{ii}^{-\frac{1}{\beta\theta}} - 1$	3.62	1.67	0.70	8.13	79
$\pi_{ii}^{-\frac{\xi_i}{\theta}} - 1$	1.97	1.08	0.37	5.11	79
$\pi_{ii}^{-\frac{\xi_i}{\beta\theta}} - 1$, $\beta =$ total linkages	5.81	3.16	1.05	14.83	79
$\pi_{ii}^{-\frac{\xi_i}{\beta\theta}} - 1$, $\beta =$ within-sector linkages	3.46	1.90	0.63	9.09	79
$\prod_{j=1}^{J} \left(\pi_{ii}^{j}\right)^{\frac{-\xi_i\omega_j}{\theta}} - 1$	3.65	2.46	0.52	10.81	79
$\prod_{j=1}^{J} \left(\pi_{ii}^{j}\right)^{\frac{-\xi_i\omega_j}{\beta_j\theta}} - 1$, $\beta_j =$ total linkages	10.87	7.85	1.45	35.30	79
$\prod_{j=1}^{J} \left(\pi_{ii}^{j}\right)^{\frac{-\xi_i\omega_j}{\beta_j\theta}} - 1$, $\beta_j =$ tradeable only sectoral linkages	6.49	4.54	0.89	20.21	79

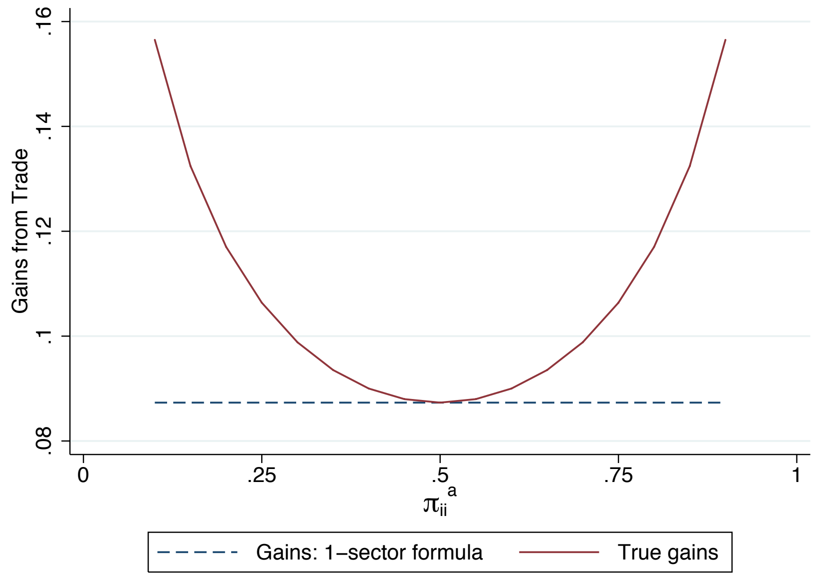
Notes: This table presents the summary statistics for the gains from trade. In the first row are the true gains computed using the full quantitative model. The rest of the rows summarize the gains from trade in this sample of countries implied by various formulas.

Table 5. Comparison of Formula-Implied Gains to True Model-Implied Gains

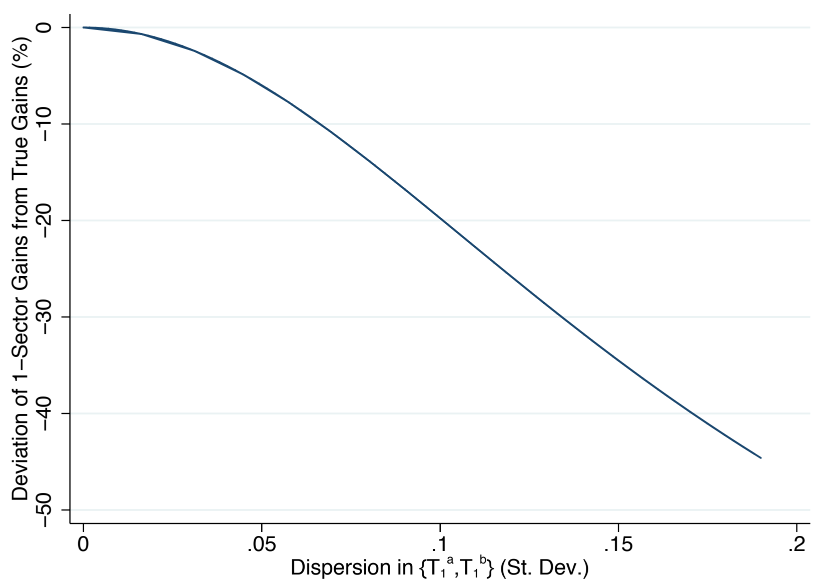
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Prop. Deviation from True Gains				Mean	Mean	Corr
Formula	Mean	Std. Dev.	Min	Max	sq. error	abs. error	(Error, $\text{cv}((T_i^j)^{1/\theta})$)
$\pi_{ii}^{-\frac{1}{\theta}} - 1$	-0.69	0.09	-0.90	-0.50	46.15	5.34	-0.531***
$\pi_{ii}^{-\frac{1}{\beta\theta}} - 1$	-0.41	0.17	-0.78	-0.10	25.50	3.60	-0.438***
$\pi_{ii}^{-\frac{\xi_i}{\theta}} - 1$	-0.70	0.08	-0.89	-0.56	43.60	5.25	-0.300***
$\pi_{ii}^{-\frac{\xi_i}{\beta\theta}} - 1$, $\beta =$ total linkages	-0.11	0.22	-0.64	0.28	9.02	1.77	-0.287**
$\pi_{ii}^{-\frac{\xi_i}{\beta\theta}} - 1$, $\beta =$ within-sector linkages	-0.47	0.13	-0.78	-0.23	25.49	3.76	-0.273**
$\prod_{j=1}^{J} \left(\pi_{ii}^j\right)^{-\frac{\xi_i \omega_j}{\theta}} - 1$	-0.49	0.07	-0.65	-0.27	19.50	3.58	0.030
$\prod_{j=1}^{J} \left(\pi_{ii}^j\right)^{-\frac{\xi_i \omega_j}{\beta_j\theta}} - 1$, $\beta_j =$ total linkages	0.48	0.18	0.04	1.02	23.45	3.65	0.275**
$\prod_{j=1}^{J} \left(\pi_{ii}^j\right)^{-\frac{\xi_i \omega_j}{\beta_j\theta}} - 1$, $\beta_j =$ tradeable only sectoral linkages	-0.11	0.11	-0.38	0.23	1.73	0.94	0.162

Notes: This table presents the comparison of the gains from trade implied by each formula and the true gains from trade. The first 4 columns report the summary statistics for the proportional deviations in the gains from trade according to the formula compared to the true gains. Columns (5) and (6) present the mean squared error and the mean absolute error for the formula-based gains from trade relative to the true ones. Column (7) reports the correlation of the error of the formula with respect to the true gains from trade with the coefficient of variation of a country's average sectoral productivities $(T_i^j)^{1/\theta}$'s relative to the world frontier, along with the level of significance. ***: significant at 1%; **: significant at 5%.

Figure 1. Illustration of Analytical Results



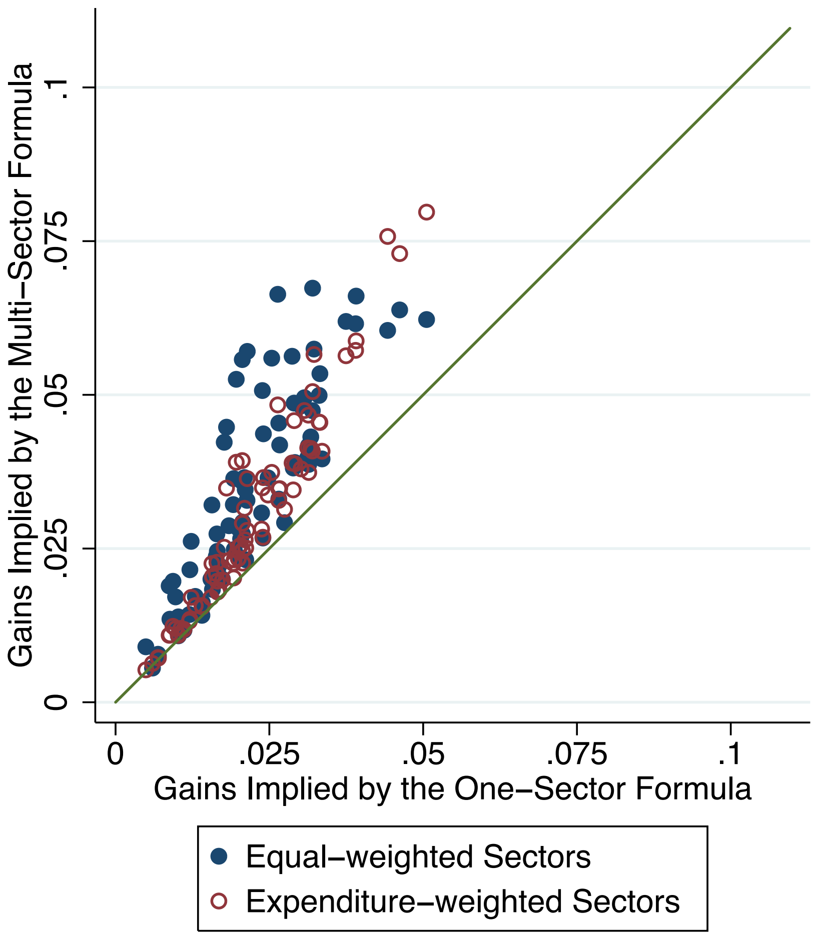
(a) Sectoral Trade Shares and the Gains from Trade



(b) Comparative Advantage and the Gains from Trade

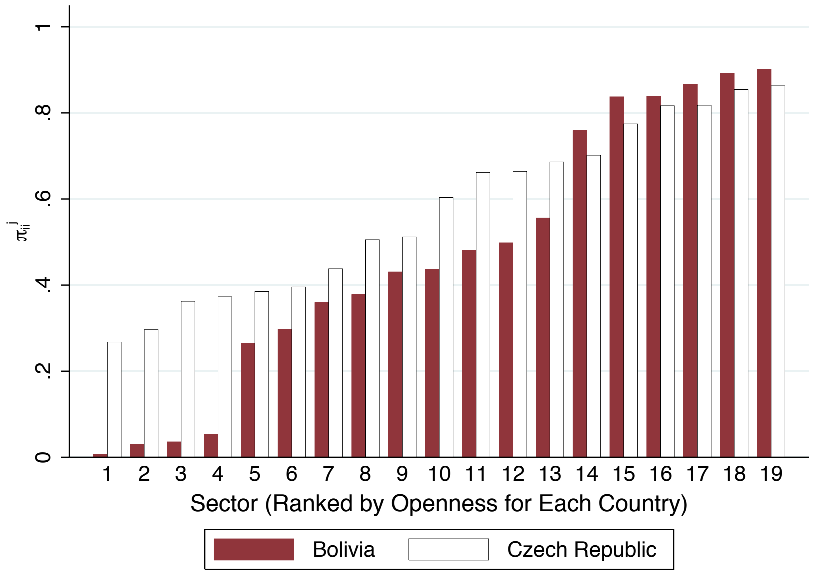
Notes: The top panel depicts the true gains from trade in the 2-sector analytical model as well as the gains from trade implied by the one-sector formula that only uses aggregate trade as a share of absorption. The bottom panel displays the percentage difference between the gains implied by the one-sector formula and true gains as a function of the dispersion in T_1^a and T_1^b, with dispersion measured by the standard deviation among them.

Figure 2. Estimated Gains from Trade According to One-Sector and Multi-Sector Formulas



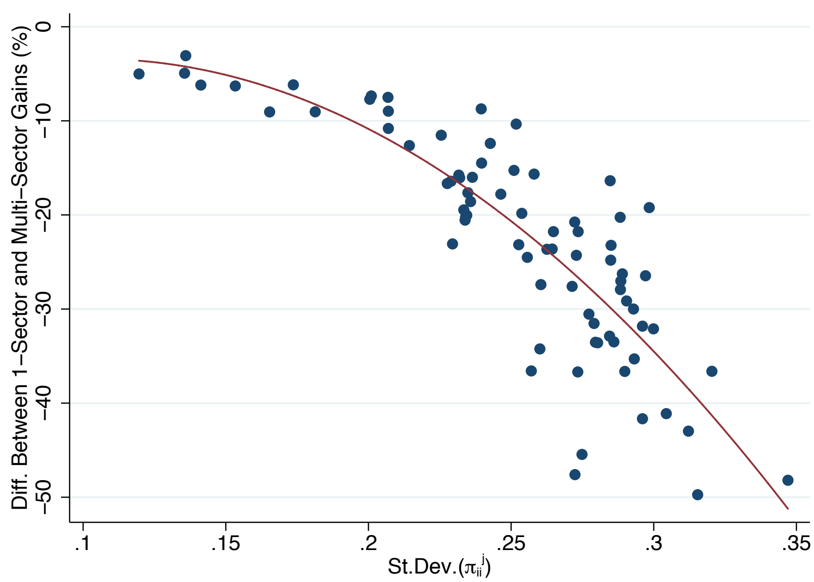
Notes: This figure plots the gains from trade implied by the one-sector formula on the x-axis against the gains from trade for the same country implied by the equal-weighted (solid dots) and the expenditure-share-weighted (hollow dots) multi-sector formulas, along with the 45-degree line.

Figure 3. Dispersion in Sectoral Trade Shares: Bolivia and the Czech Republic



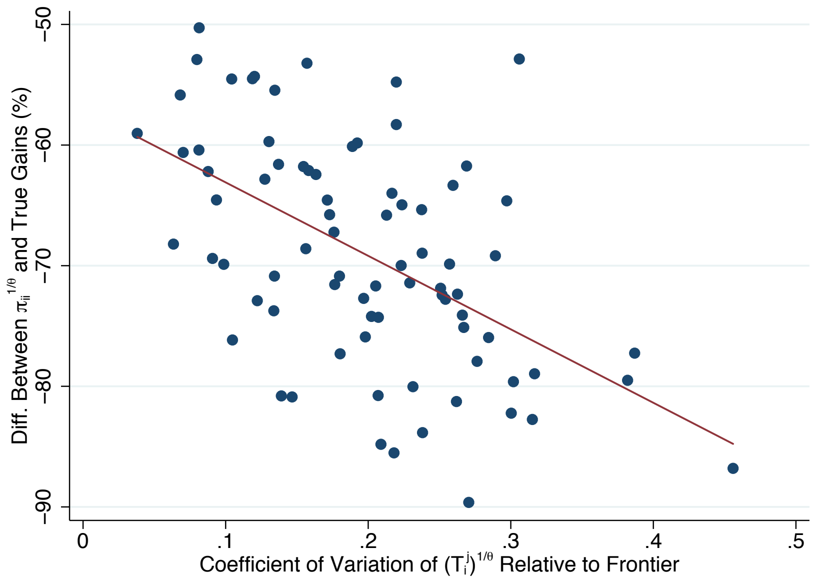
Notes: This figure displays the π_{ii}^{j} in the 19 manufacturing sectors for Bolivia and the Czech Republic, ranked for each country by sectoral trade openness (sectors with lower π_{ii}^{j} correspond to sectors with lower shares of spending on domestic goods in the total sectoral spending.

Figure 4. Dispersion in Sectoral Trade Shares and the Gains from Trade



Notes: This figure plots the percentage difference between the gains from trade implied by the one-sector formula relative to the multi-sector formula against the standard deviation of π_{ii}^{j} for that country. The curve through the data is the quadratic fit.

Figure 5. Comparative Advantage and the Gains from Trade



Notes: This figure plots the percentage difference in the gains from trade implied by the one-sector formula $\pi_{ii}^{1/\theta}$ relative to the true gains against the coefficient of variation in the sectoral productivities $(T_i^j)^{1/\theta}$ relative to the world frontier. The line through the data is the OLS fit.

Table A1. Country Coverage

Argentina	Korea, Rep.
Australia	Kuwait
Austria	Latvia
Bangladesh	Lithuania
Belgium-Luxembourg	Macedonia, FYR
Bolivia	Malaysia
Brazil	Mauritius
Bulgaria	Mexico
Canada	Netherlands
Chile	New Zealand
China	Nigeria
Colombia	Norway
Costa Rica	Pakistan
Czech Republic	Peru
Denmark	Philippines
Ecuador	Poland
Egypt, Arab Rep.	Portugal
El Salvador	Romania
Estonia	Russian Federation
Ethiopia	Saudi Arabia
Fiji	Senegal
Finland	Slovak Republic
France	Slovenia
Germany	South Africa
Ghana	Spain
Greece	Sri Lanka
Guatemala	Sweden
Honduras	Switzerland
Hungary	Taiwan Province of China
Iceland	Tanzania
India	Thailand
Indonesia	Trinidad and Tobago
Iran, Islamic Rep.	Turkey
Ireland	Ukraine
Israel	United Kingdom
Italy	United States
Japan	Uruguay
Jordan	Venezuela, RB
Kazakhstan	Vietnam
Kenya	

Notes: This table reports the countries in the sample.

Table A2. Sectors

ISIC code	Sector Name	α_j	β_j	$\gamma_{J+1,j}$	ω_j
15	Food and Beverages	0.290	0.290	0.303	0.166
16	Tobacco Products	0.272	0.490	0.527	0.019
17	Textiles	0.444	0.368	0.295	0.015
18	Wearing Apparel, Fur	0.468	0.369	0.320	0.059
19	Leather, Leather Products, Footwear	0.469	0.350	0.330	0.014
20	Wood Products (Excl. Furniture)	0.455	0.368	0.288	0.008
21	Paper and Paper Products	0.351	0.341	0.407	0.012
22	Printing and Publishing	0.484	0.453	0.407	0.005
23	Coke, Refined Petroleum Products, Nuclear Fuel	0.248	0.246	0.246	0.043
24	Chemical and Chemical Products	0.297	0.368	0.479	0.006
25	Rubber and Plastics Products	0.366	0.375	0.350	0.012
26	Non-Metallic Mineral Products	0.350	0.448	0.499	0.072
27	Basic Metals	0.345	0.298	0.451	0.002
28	Fabricated Metal Products	0.424	0.387	0.364	0.014
29C	Office, Accounting, Computing, and Other Machinery	0.481	0.381	0.388	0.134
31A	Electrical Machinery, Communication Equipment	0.369	0.368	0.416	0.074
33	Medical, Precision, and Optical Instruments	0.451	0.428	0.441	0.060
34A	Transport Equipment	0.437	0.329	0.286	0.219
36	Furniture and Other Manufacturing	0.447	0.396	0.397	0.067
4A	Nontradeables	0.561	0.651	0.788	
	Mean	0.414	0.393	0.399	0.053
	Min	0.244	0.243	0.246	0.002
	Max	0.561	0.651	0.788	0.219

Notes: This table reports the sectors used in the analysis. The classification corresponds to the ISIC Revision 3 2-digit, aggregated further due to data availability. α_j is the value-added based labor intensity; β_j is the share of value added in total output; $\gamma_{J+1,j}$ is the share of nontradeable inputs in total intermediate inputs; ω_j is the tradeable sector j's share of total tradeable expenditure. Variable definitions and sources are described in detail in the text.