Essays in International Trade

Amanda Rae Geary Kurzendoerfer
Florence, Kentucky

Master of Arts, University of Virginia, 2010
Bachelor of Arts, University of Kentucky, 2004

A Dissertation presented to the Graduate Faculty
of the University of Virginia in Candidacy for the Degree of
Doctor of Philosophy

Department of Economics

University of Virginia
May, 2015

Abstract

The first chapter in this dissertation studies the importance of distinguishing between intermediate and final use for the gains from trade. The domestic expenditure share, which is central to calculating the gains from trade across a wide class of models, varies by end use. Failure to account for this heterogeneity in a one-sector, one-factor model systematically understates the gains from trade and conceals differences in the returns to openness across end use. I construct a multi-sector, multi-factor model with input-output linkages that incorporates variation in end use to investigate the full extent of these discrepancies, and to explore the relationship between relative income and intermediate relative to final use estimates of Ricardian comparative advantage, trade costs, and prices. I estimate the parameters of the model for 38 countries and 32 manufacturing and service industries using the World Input-Output Database. Lower income countries have a comparative disadvantage in producing and importing intermediate relative to final goods, which results in a higher relative price of intermediates for these countries. Including end-use variation raises the gains from trade by 14.4 percent on average.

The second chapter examines the relationship between distance and bilateral trade at the industry level, where zero trade flows are prevalent. I expand upon the work of Berthelon and Freund (2008) by incorporating zero trade flows into the estimation using Tobit and Poisson pseudo-maximum-likelihood methods. Methods that incorporate zeros reveal that distance sensitivity is either decreasing (Tobit) or increasing only slightly (PPML) over time. This stands in contrast to the OLS approach, which shows a significant increase in distance sensitivity. I find that more substitutable goods and those with higher trade costs are likely to exhibit higher sensitivity to distance, but that these qualities cannot explain the small and often insignificant changes in Poisson-estimated distance elasticity over time.

The last chapter decomposes U.S. water use, which has followed a remarkable pattern since 1950, not mimicking the almost uninterrupted 110 percent increase in the size of the U.S. population, the relatively steady 570 percent growth in real GDP, and the 220 percent improvement in per capita GDP. After doubling between 1950 and 1980, the total volume of water withdrawn has stabilized and even decreased in recent years. Our decomposition shows that between 35 and 50 percent of the productivity gains that allowed the U.S. to produce each dollar of its GDP with increasingly less water stems from long-term structural changes of the U.S. economy since 1950 (growing service economy, declining manufacturing and agricultural sectors). The remaining 50 to 65 percent is due to improved production techniques, and in particular due to water productivity improvements in the electricity-generating sector, especially since the mid to late 1970s. We argue that while globalization has helped reduce U.S. water use particularly since 1980, the U.S. ability to import more water-intensive goods is not the main reason U.S. water use plateaued.

Acknowledgements

I would like to thank my advisors James Harrigan, John McLaren, and Ariell Reshef for sharing their expertise and offering many fine suggestions that influenced and improved this research. I am also abundantly grateful to Peter Debaere for his guidance and ready encouragement.

This research was supported by the Albert Gallatin Graduate Research Fellowship and the Bankard Fund for Political Economy, both at the University of Virginia. I am thankful for the generous support of these programs.

Above all, this work would not have been possible without the devotion of Kevin, my staunchest supporter.

For Harry and Jack

Contents

Chapter 1

Comparative Advantage, End Use, and the Gains from Trade

1.1 Introduction

In the Ricardian model of international trade, countries benefit from trade by specializing in the activities in which they are relatively more productive. A trade liberalization allows countries to produce and export more of their comparative advantage sectors and import more of their comparative disadvantage sectors. The larger the productivity differences, the larger the reallocations, and the larger the gains from trade. Productivity differences are therefore central to determining the gains from trade. International trade data suggests that productivity varies by intermediate and final end use; that is, some countries are relatively better at producing goods intended for intermediate use and others are relatively better at producing goods intended for final consumption. Despite apparent productivity differences, comparative advantage by end use has not been explored as an avenue for the gains from trade. In this paper, I construct a general equilibrium Ricardian trade model that features productivity differences by intermediate and final use to determine their contribution to the gains from trade.

Distinguishing productivity by end use highlights the different roles of intermediate and final goods in an economy, and their different contributions to the gains from trade. Intermediates are used in the production of intermediates, which are used in the production of other intermediates and so on (and ultimately final goods), so the gains from trade are magnified when intermediate productivity improves or barriers to trading intermediates are reduced. In contrast, final goods are consumed once, so the benefit of a productivity improvement or trade liberalization in final goods passes directly to the consumer, but does not accumulate through the production process. The existing literature does not incorporate productivity differences that arise by virtue of a good's end-use classification, potentially masking an asymmetric response of the gains from trade to adjustments in the characteristics of intermediate and final use trade.

The evidence that productivity varies by end use comes from a single statistic, the domestic expenditure share. In a Ricardian framework with costly trade, a country's share of total expenditure on domestically produced goods, or its domestic expenditure share, contains information about its comparative advantage and access to imports. A high share implies that a country is either very productive at producing a particular good or that it faces significant barriers to importing the good from low cost locations. Figure 1.1 plots the domestic expenditure share for intermediates against the domestic expenditure share for final goods for 40 countries.[1] If productivity and trade barriers did not vary by end use, the shares would not vary by end use, and the points in Figure 1.1 would lie on the 45°-line. As the figure shows, intermediate and final domestic expenditure shares are correlated—countries that purchase a large share of intermediates from home tend also to purchase a large share of final goods

[1] All calculations are based on data from the World Input-Output Database (WIOD), http://www.wiod.org/new_site/home.htm, which I describe in Section 1.6.

from home—but the difference is often large and varies by country. The difference between shares ranges from as much as 32 percent (Luxembourg), to as little as minus three percent (Russia, the only country for which the intermediate domestic share is higher). Within-country differences in intermediate and final domestic expenditure shares indicate that productivity, trade costs, or both vary by end use.

Because the domestic expenditure share captures information about a country's comparative advantage and access to imports, it is central to determining the country's gains from trade. Arkolakis, Costinot, and Rodríguez-Clare (2012) show that the domestic expenditure share and the trade cost elasticity are the only variables needed to compute the gains from trade relative to autarky across a wide class of models. I show that the expression for the gains from trade in a simple one-sector, one-factor version of the full model with end-use variation is a function of *both* the intermediate and final domestic expenditure shares, and the trade cost elasticity. This is in contrast to the same model without end-use variation (Eaton and Kortum, 2002), in which the *overall* domestic expenditure share and the trade cost elasticity determine the gains from trade. I show that the model without end-use variation will always understate the gains from trade when trade is balanced (and the intermediate and final domestic expenditure shares are not the same). Further, I demonstrate the asymmetry of the elasticity of the gains from trade with respect to intermediate and final domestic expenditure shares.

Differences in intermediate and final domestic expenditure shares generate gains from trade, and the shares contribute asymmetrically to the gains from trade. Determining the underlying productivity differences that generate differences in the shares is therefore an informative exercise. The simple model provides an expression that relates intermediate relative to final domestic expenditure shares to relative technology and relative prices. Relative prices reflect a country's ability to access intermediates vis à vis final goods at low cost. In a first look at comparative advantage by use, I use data on the relative price of intermediates—which is sharply decreasing in income—to extract relative productivities from the domestic expenditure shares. I find that low income countries have a comparative disadvantage in the production of intermediates.

The simple model provides an analytical expression for the gains from trade and the relationship between domestic expenditure shares and comparative advantage, but it does not incorporate the full extent of productivity differences by end use. The data also show that domestic expenditure shares vary by end use within industries. Figure 1.2 plots the intermediate share against the final share for 32 goods and service industries in 38 countries.[2] The point *Japan, Leather Goods*, for example, demonstrates that Japan turns to domestic producers for 92 percent of its intermediate leather requirements, but is considerably more open in its purchases of leather final goods—the domestic expenditure share is just 20 percent. To capture this variation, and to incorporate the fact that an industry's output is used in varying intensities by other industries, I construct a multi-industry Eaton and Kortum (2002) model with input-output linkages and end-use variation within industries. The model features Ricardian motives for trade at the industry-by-end-use level, and also incorporates multiple factors (labor and capital). The full model does not provide an analytical expression for the gains from trade, so I estimate the parameters using three different regression techniques and solve the model numerically to determine the contribution of end-use variation to the gains from trade; I find that the gains from trade are 14.4 percent higher in a model with end-use variation than in a model without.

I also use the parameter estimates from the full model to provide a closer look at comparative advantage. The estimates support the aggregate result that lower

[2] I combine the three small, open economies Cyprus, Luxembourg, and Malta, and some industries to avoid observations of zero gross output, as I describe in Section 1.6 (See Tables 1.5 and 1.6 for the country and industry aggregation schemes.)

income countries have a comparative disadvantage in producing intermediates relative to final goods. I find that this result is driven by a comparative disadvantage in intermediate agriculture and manufacturing industries (and to some extent service industries) in these countries. Further, low income countries pay relatively more to import intermediates. A comparative disadvantage in intermediate production and a high cost to import are consistent with low income countries paying a higher relative price for intermediates, which the aggregate data show and my parameter estimates support. I also show that intermediates are more tradable than final goods, and that the estimates imply a Balassa-Samuelson effect: countries that have a comparative advantage in the production of more tradable goods (intermediates) pay a relatively higher price for less-tradable goods (final goods).

Recent literature has sought to quantify the gains from trade under the different sources of heterogeneity that Arkolakis et al. present in their theoretical paper. Examples include Costinot and Rodríguez-Clare (2013), Levchenko and Zhang (2014), and Caliendo and Parro (2012). Costinot and Rodríguez-Clare find that multiple sectors and tradable intermediate goods have larger effects on the gains from trade than market structure and firm-level heterogeneity. Levchenko and Zhang find that sectoral heterogeneity increases the gains from trade by 30 percent relative to a one-sector model, and show analytically that the one-sector model will always understate the gains from trade. Caliendo and Parro estimate the welfare effects of NAFTA, and find that welfare is reduced by more than 40 percent when intermediate goods and country-varying input-output linkages are not considered. This paper is the first to quantify the contribution of end-use variation to the gains from trade. I do this using a model that includes end-use variation, as well as the sectoral heterogeneity, tradable intermediate inputs, and input-output linkages that the literature described above has shown are important channels for the gains from trade. This paper and those above rely on the multi-sector Eaton and Kortum framework that was introduced by Shikher (2011, 2012, and 2013) and Chor (2010). The model is also related to Melitz and Redding (2014), who show that the gains from trade in a model with sequential production become arbitrarily large as the number of production stages increases. Distinguishing end-use, I construct a model with two stages of production: intermediates (stage one) are required to produce final goods (stage two).[3] This empirics in this paper are related to Levchenko and Zhang (2013), who use a multi-sector Eaton and Kortum model to estimate technology parameters and find that comparative advantage has weakened over time. I also use the parameter estimates to assess comparative advantage, but by end use and across countries rather than by industry and over time.

Literature that features end-use variation centrally is outside the context of the literature on the gains from trade, and typically focuses on the importance of low trade barriers and productivity in intermediates vis à vis final goods. Amiti and Konings (2007) find that, in the context of an Indonesian trade liberalization, a decline in tariffs on intermediate inputs leads to a productivity gain for firms that import their inputs that is at least twice as high as the gain from reducing tariffs on final goods. Jones (2011) shows that linkages through intermediate goods generate a productivity multiplier that helps to explain large income differences across countries. A United Nations Conference on Trade and Development (2013) report discusses the importance of participation in global value chains—which is determined by the proportion of a country's exports that are part of a multi-stage production process, and is therefore an indication of participation in intermediate goods trade—for generating employment and increasing GDP and income growth. These papers demonstrate that there are

[3]The structure of intermediate production itself is "roundabout" rather than sequential, in that any intermediate input can be used in the production of another intermediate.

important benefits to improved competitiveness in intermediates. I explore this idea further—first by showing analytically that the gains from trade are more responsive to changes in intermediate trade, and second by showing that technology, trade costs, and prices vary by end use in a way that is related to income.

The rest of this paper is organized as follows. In Section 1.2 I set up a one-sector, one-factor model with end-use variation. I show that a model that fails to account for this variation will weakly understate the gains from trade, and that the size of the discrepancy depends on the ratio of final to intermediate domestic expenditure shares and the labor share. I show the circumstances under which the gains from trade are more responsive to changes in the intermediate domestic expenditure than to changes in the final domestic expenditure share, and use aggregate data to demonstrate the magnitude of the discrepancy in the gains from trade and to quantify the elasticities. In Section 1.3 I take a first look at comparative advantage, showing the implications for relative technology levels given data on intermediate and final prices and domestic expenditure shares. In Section 1.4 I set up the full general equilibrium model, incorporating variation in end use at the industry level, input-output linkages, and capital. In Section 1.5 I describe the estimation procedure and the data. Section 1.6 describes the data and implementation for estimation. Section 1.7 presents the results of the estimation, shows that the relative parameter estimates are related to income, and demonstrates evidence of the Balassa-Samuelson effect. In Section 1.8 I use the estimated parameters to solve the full general equilibrium model to show the effect of incorporating end-use heterogeneity on the gains from trade. Section 1.9 concludes.

1.2 Simple model and some magnitude

I first describe an extension of the Eaton and Kortum (2002) model that incorporates variation in end use. I demonstrate that the standard model understates the gains from trade, and that the discrepancy depends on two variables: the ratio of the intermediate and final domestic expenditure shares and the labor share. I also show that the gains from trade are more responsive to changes in the intermediate domestic expenditure share when the intermediate share in total output is greater than 50 percent, and more responsive than the standard model implies when the intermediate domestic expenditure share is less than the final domestic expenditure share. I then turn to the data to demonstrate the magnitude of the discrepancy and the elasticities of the gains from trade with respect to intermediate and final domestic expenditure shares for the 38 countries in my sample.

1.2.1 Simple model with end-use variation

There are N countries. Production is Cobb-Douglas over labor and intermediates, with unit costs in country i given by $c_i = w_i^{\beta_i}(p_i^I)^{1-\beta_i}$, where w_i is the wage, p_i^I is the price of a bundle of intermediates, and β_i is the labor share in total output ($0 < \beta_i < 1$).[4] Countries produce varieties of intermediate and final goods, and varieties are produced with productivities that vary by end use. End use is distinguished by $u = \{I, F\}$, varieties are indexed by l on $[0,1]$, and productivity is given by $z_i^u(l)$. Productivity is drawn from a Fréchet distribution with location parameter T_i^u and dispersion parameter θ. T_i^u is the absolute productivity level for country i, end use

[4] In a minor departure from Eaton and Kortum, I allow the labor share to vary by country. This has implications for the full general equilibrium solution, but, other than allowing the elasticity of the gains from trade with respect to the overall domestic expenditure share to vary by country, it does not change the standard gains from trade formula.

u, and the ratio of intermediate to final technology levels determines comparative advantage in producing goods suited for each end use. That is, $T_i^I/T_i^F > T_{i'}^I/T_{i'}^F$ means that country i has a comparative advantage in the production of intermediate relative to final goods compared to country i'. Trade costs vary by end use and take the iceberg form: τ_{ni}^u units of the good destined for end use u in country n must be shipped from i for one unit to arrive (within-country trade costs are normalized to one, $\tau_{ii}^u = 1$). Perfect competition implies that a buyer in country n would pay $p_{ni}^u(l) = c_i \tau_{ni}^u / z_i^u(l)$, the productivity-adjusted unit cost times the iceberg trade cost, if the variety were bought from country i. Buyers, who can be producers shopping for intermediates or consumers shopping for final goods, purchase the variety from the lowest-cost source and combine varieties in CES fashion. The technology distribution and CES price index yield a closed form expression for prices paid for intermediate and final goods in the destination country: $p_n^u = \gamma \left[\sum_{i=1}^N T_i^u (c_i \tau_{ni}^u)^{-\theta}\right]^{-1/\theta}$. The probability that country i is the lowest cost provider of variety l to country n, which is also the fraction of expenditure by country n on goods from country i is $\pi_{ni}^u = \frac{X_{ni}^u}{X_n^u} = T_i^u \left(\frac{\gamma c_i \tau_{ni}^u}{p_n^u}\right)^{-\theta}$, where X_{ni}^u is expenditure by country n on goods of end use u from country i and X_n^u is expenditure by country n on goods of end use u from all countries. The fraction of expenditure by country i on goods from itself, the domestic expenditure share, is $\pi_{ii}^u = T_i^u \left(\frac{\gamma c_i}{p_i^u}\right)^{-\theta}$.

To solve for the gains from trade, I first find real wages by substituting the unit cost function into the final domestic expenditure share equation ($u = F$) and rearranging:

$$\frac{w_i}{p_i^F} = \gamma^{-1/\beta_i} \left(\frac{T_i^F}{\pi_{ii}^F}\right)^{1/\beta_i\theta} \left(\frac{p_i^F}{p_i^I}\right)^{1/\beta_i - 1}. \tag{1.1}$$

Welfare is measured by the purchasing power of wages in terms of the final good, the price of which may differ from the price of the intermediate good. The price of the intermediate good affects real wages indirectly through the use of intermediates in production of the final good (and through general equilibrium effects on the wage). I next use the domestic expenditure share equation for both final and intermediate goods to write relative prices as a function of relative domestic expenditure shares and technology levels:

$$\frac{p_i^F}{p_i^I} = \left[\left(\frac{\pi_{ii}^I}{\pi_{ii}^F}\right)\left(\frac{T_i^F}{T_i^I}\right)\right]^{-1/\theta}, \tag{1.2}$$

and substitute (1.2) into (1.1):

$$\frac{w_i}{p_i^F} = \gamma^{-1/\beta_i}\left[\left(\frac{T_i^F}{\pi_{ii}^F}\right)^{\beta_i}\left(\frac{T_i^I}{\pi_{ii}^I}\right)^{1-\beta_i}\right]^{1/\beta_i\theta}. \tag{1.3}$$

The change in real wages, $\widehat{W} \equiv \left(w_i/p_i^F\right)'/\left(w_i/p_i^F\right)$, associated with a move from autarky ($\pi_{ii}^u = 1$) to trade is then:

$$\widehat{W} = \left[\left(\pi_{ii}^F\right)^{\beta_i}\left(\pi_{ii}^I\right)^{1-\beta_i}\right]^{-1/\beta_i\theta}. \tag{1.4}$$

This expression has the counterpart $\pi_{ii}^{-1/\beta_i\theta}$ in the standard model. The expressions differ in terms of the interior component: the model with end-use variation relies on a geometric weighted average of the intermediate and final domestic expenditure shares, while the standard formulation depends on the overall domestic expenditure share (π_{ii}). Before formalizing the conditions under which the two gains from trade expressions diverge, I discuss the intuition behind the expression that incorporates end-use variation.

Rearranging the exponents, we can rewrite (1.4) as $\widehat{W} = \left(\pi_{ii}^F\right)^{-1/\theta}\left(\pi_{ii}^I\right)^{-(1-\beta_i)/\beta_i\theta}$. The elasticity of the gains from trade with respect to the final domestic expenditure share is $-1/\theta$, as it is in the gains from trade expression with no intermediates (see

Arkolakis et al. equation (1)), where θ is the trade cost elasticity. Final goods are not used in the production of other goods, so openness in final goods is not subject to the amplification in the gains from trade that arises when a good is part of an input-output loop. In contrast, intermediates are used in the production of other intermediates, so the gains from trade are amplified by the share of intermediates in total expenditure (precisely by the labor share, which is one minus the intermediate share), hence the presence of β_i in the denominator of the exponent on the intermediate domestic expenditure share (see Arkolakis et al. Section IV.B). Because welfare is measured by the purchasing power of wages in terms of final goods, this magnification effect is only directly relevant to the gains from trade through the extent to which final goods rely on intermediates, hence the presence of $1-\beta_i$ in the numerator of the exponent. Thus, we can think of the gains from trade as being determined directly by openness in final goods, and indirectly by openness in intermediates through two channels: the effect on other intermediates, and on final goods.

I now show that the gains from trade in a model without end-use variation will systematically understate the true gains from trade, and that the size of the discrepancy depends on the ratio of intermediate to final domestic expenditure shares and the labor share. First, rewrite the overall domestic trade share π_{ii} as a linear combination of the final and intermediate domestic expenditure shares, where β_i and $1-\beta_i$ are the weights when trade is balanced: $\pi_{ii} = \beta_i \pi_{ii}^F + (1-\beta_i)\pi_{ii}^I$.[5] Now we can easily compare the gains from trade formula with end use variation to the standard

[5] To see that the labor and intermediate shares in total output are the correct weights when trade is balanced, first recall notation—that X_i^u is expenditure by country i on goods of end use u. In equilibrium, payments to labor (the only factor of production) equal total expenditure on final goods X_i^F, and total output equals total expenditure, X_i. Thus, β_i is also the share of expenditure on final goods in total expenditure, X_i^F/X_i, and $1-\beta_i$ is the share of expenditure on intermediate goods in total expenditure, X_i^I/X_i. We can write the overall domestic expenditure share as $\pi_{ii} = (X_{ii}^F + X_{ii}^I)/X_i$, which is the same as $(X_i^F/X_i)(X_{ii}^F/X_i^F) + (X_i^I/X_i)(X_{ii}^I/X_i^I)$. It follows then that $\pi_{ii} = \beta_i \pi_{ii}^F + (1-\beta_i)\pi_{ii}^I$.

formulation: in the former the interior component is a geometric weighted average of the final and intermediate domestic trade shares, and in the latter it is a linear weighted average of the final and intermediate domestic trade shares. That is,

$$\widehat{W}_{End\text{-}Use} = \left[\left(\pi_{ii}^F\right)^{\beta_i}\left(\pi_{ii}^I\right)^{1-\beta_i}\right]^{-1/\beta_i\theta}, \tag{1.5}$$

$$\widehat{W}_{Standard} = \left[\beta_i \pi_{ii}^F + (1-\beta_i)\pi_{ii}^I\right]^{-1/\beta_i\theta}. \tag{1.6}$$

Taking the logarithm of each interior component we see that, by Jensen's Inequality, the geometric expression will always be less than or equal to the linear expression:

$$\beta_i \ln \pi_{ii}^F + (1-\beta_i)\ln \pi_{ii}^I \leq \ln(\beta_i \pi_{ii}^F + (1-\beta_i)\pi_{ii}^I), \tag{1.7}$$

and strictly less when $\pi_{ii}^F \neq \pi_{ii}^I$. Because the gains from trade formulas are decreasing in their interior components, the standard formulation will always understate the end-use formulation when the intermediate and final domestic expenditure shares are not the same. This is Proposition 1.

Proposition 1 *When trade is balanced, the gains from trade in the standard one-sector model weakly understate the gains from trade in the one-sector model with end-use variation. That is:*

$$\widehat{W}_{Standard} = \left[\beta_i \pi_{ii}^F + (1-\beta_i)\pi_{ii}^I\right]^{-1/\beta_i\theta} \leq \left[\left(\pi_{ii}^F\right)^{\beta_i}\left(\pi_{ii}^I\right)^{1-\beta_i}\right]^{-1/\beta_i\theta} = \widehat{W}_{End\text{-}Use}.$$

The inequality is strict when $\pi_{ii}^F \neq \pi_{ii}^I$.

A corollary to the proposition is related to the size of the discrepancy, which we can determine analytically by taking the ratio of the end-use (1.5) and standard (1.6) versions and rearranging.

Corollary 1 *For a given θ, the discrepancy in the gains from trade between the end-use and standard models depends on the ratio of domestic trade shares (π_{ii}^F/π_{ii}^I) and the labor share (β_i):*

$$\widehat{W}_{End-Use}/\widehat{W}_{Standard} = \left(\frac{\pi_{ii}^F}{\pi_{ii}^I}\right)^{-1/\theta}\left[\beta_i\left(\frac{\pi_{ii}^F}{\pi_{ii}^I}\right) + (1-\beta_i)\right]^{1/\beta_i\theta}. \tag{1.8}$$

The further apart are the final and intermediate domestic trade shares and the lower β_i, the larger the discrepancy. The size of the overall trade share π_{ii} does not matter; it is the extent to which the domestic trade shares are different that affects the discrepancy in the gains from trade. Figure 1.3 plots the discrepancy in the gross gains from trade against a potential range of the ratio of final to intermediate domestic expenditure shares and the range of possible labor shares for $\theta = 4$.[6] As the figure shows, the discrepancy is largest when π_{ii}^F and π_{ii}^I are most different and β_i is low.

Turning now to the elasticity of the gains from trade with respect to each domestic expenditure share, equation (1.5) shows that the elasticities with respect to the final and intermediate domestic expenditure shares are $-1/\theta$ and $-(1-\beta_i)/\beta_i\theta$, respectively. Thus the elasticity with respect to the intermediate share will be larger than the elasticity with respect to the final share when $\beta_i < 0.5$, and it will be larger by a factor of $(1-\beta_i)/\beta_i$. The lower the labor share, the more responsive are the gains from trade to the intermediate domestic expenditure share than to the final share. As discussed previously, this is because intermediates are used more intensively in the production of other intermediates and in the production of final goods when the labor share is low. We can also compute the elasticity of the gains from trade with respect to each domestic trade share for the standard formulation. From equation (1.6), we

[6] I use $\theta = 4$ here and throughout the paper following Simonovska and Waugh (2014), who show that the Eaton and Kortum (2002) estimator is biased and will overestimate the elasticity of trade in finite sample sizes. Simonovska and Waugh develop a new estimator that reduces the bias and yields an estimate of θ that is roughly equal to 4.

can show that the elasticity with respect to the final domestic expenditure share is $(-1/\theta)(\pi_{ii}^F/\pi_{ii})$, and with respect to the intermediate domestic expenditure share is $(-(1-\beta_i)/\beta_i\theta)(\pi_{ii}^I/\pi_{ii})$. Thus the gains from trade are $((1-\beta_i)/\beta_i)(\pi_{ii}^I/\pi_{ii}^F)$ times more responsive to changes in the intermediate domestic expenditure share than to changes in the final domestic expenditure share, and the standard model will understate the importance of changes in the intermediate domestic expenditure share when $\pi_{ii}^I < \pi_{ii}^F$.

The analytical expressions for the discrepancies between the end-use and standard models discussed above rely on the assumption of balanced trade—that the overall domestic expenditure share can be written as a linear combination of the intermediate and final domestic expenditure shares with respective weights β_i and $1-\beta_i$. Trade is not balanced in practice, however, and a researcher following the standard procedure observes only the overall domestic expenditure share. It is therefore important to quantify the size of the actual discrepancy, or the discrepancy that would result from using the observed overall domestic expenditure share (not the labor share weighted average) to compute the gains from trade.

1.2.2 Size of the discrepancy

Table 1.1 reports the average overall, final, and intermediate domestic expenditure shares, the average ratio of final to average intermediate shares, and the average labor share for countries in three income classifications for the year 2007.[7] Income classifications are determined by the World Bank (see Table 1.5); high income countries are the developed economies in North America, Europe, and the Asia-Pacific region,

[7]The simple model does not include capital as a factor of production, so I net capital compensation out of the labor share calculation. That is, *labor share* = *labor compensation*/(*gross output* − *capital compensation*).

upper middle income economies include the transition economies in Central and Eastern Europe, as well as Brazil and Mexico, and the lower middle income countries are China, India, and Indonesia. Higher income countries are more open overall, and for the intermediate and final classifications individually. All groups purchase a larger share of final goods and services from home than intermediate goods and services ($\pi_{ii}^F/\pi_{ii}^I > 1$), and this fact tends to be more pronounced for the richer countries: the average ratio of final to intermediate domestic expenditure shares is 1.17 for the high income group and 1.06 for the lower middle income group. Labor shares are on average lower in lower income countries owing to relatively less service-sector output in these countries (34 percent for high income and 27 percent for lower middle income). Recalling equation (1.8) and Figure 1.3, the discrepancy in the gains from trade across the two models is increasing in the domestic expenditure share ratio (when it is greater than one) and decreasing in the labor share, so it is not obvious a priori which group will experience the largest discrepancy.

Table 1.2 maps the domestic expenditure shares and labor shares into the gains from trade under the standard and end-use models, reports the discrepancy between the two, and also shows the relative elasticity of the gains from trade with respect to the intermediate and final domestic expenditure shares. As the table shows, the average discrepancy in the gains from trade across the two models is largest for the lower income countries (15.1 percent for upper middle and 16.2 percent for lower middle compared to 11.1 percent for high income). In this instance, the lower average labor shares (which increase the size of the discrepancy) in the upper middle and lower middle income classifications offset the effect of their lower average domestic expenditure share ratios (which reduce the size of the discrepancy) relative to the high income classification. The last column in Table 1.2 reports the relative elasticity of the gains from trade with respect to the intermediate and final domestic expenditure shares, which is $(1-\beta_i)/\beta_i$. The gains from trade are three times as responsive to the intermediate domestic expenditure share as to the final domestic expenditure share in the lower middle income group. This distinction between the responsiveness between intermediate and final domestic expenditure share is altogether missed in the standard model, and the relative elasticities are equal to one.

Underlying the averages is a considerable amount of variation across countries. Table 1.3 shows the country-level domestic expenditure shares, ratios, and labor shares, and Table 1.4 shows the gains from trade discrepancies and the relative elasticities. The size of the discrepancy ranges from -5 percent for Russia to 44 percent for Mexico.[8] The relative elasticity is as low as 1.3 in Greece, and the gains from trade are nearly five times as responsive to the intermediate domestic expenditure share as to the final domestic expenditure share in China. This is directly a consequence of China's low labor share and demonstrates the disproportionate importance of intermediates given their large share in China's production.

1.3 Comparative advantage, a first look

Domestic expenditure shares vary by end use and have different effects on the gains from trade, so as a next step I look at the factors that contribute to differences in relative domestic expenditure shares: prices and productivity. In this section I combine country-level data on prices of intermediate and final goods with the intermediate and final domestic expenditure shares to make an inference about the nature of comparative advantage across countries and end use. I continue to use the simple model in

[8] The analytical discrepancy is always weakly positive when trade is balanced. The gains from trade under the standard model are calculated using the overall domestic expenditure share and trade is not necessarily balanced, so it is possible that the discrepancy is negative—as is the case for Russia.

this section, showing in Section 1.4 the implications for comparative advantage using the full model.

A country that sources a relatively larger share of intermediates than final goods domestically will have a higher relative technology in producing intermediates or a higher relative price of intermediates. We can see this by rearranging equation (1.2):

$$\frac{\pi_{ii}^I}{\pi_{ii}^F} = \left(\frac{T_i^I}{T_i^F}\right)\left(\frac{p_i^I}{p_i^F}\right)^{\theta}. \tag{1.9}$$

If trade were completely costless and consequently the law of one price held, the relative price would be the same across countries, and differences in the relative domestic expenditure share would be governed only by relative technology levels. We would then conclude that comparative advantage in the production of intermediates is decreasing in income, as $(\pi_{ii}^I/\pi_{ii}^F)_{LowerMiddle} > (\pi_{ii}^I/\pi_{ii}^F)_{UpperMiddle} > \pi_{ii}^I/\pi_{ii}^F)_{High}$, see Table 1.1. Trade is far from costless, however, so we cannot make a statement about the relationship between comparative advantage and domestic expenditure shares without some knowledge of relative prices. Prices are in principle observable, so together with relative domestic expenditure shares and an estimate of θ we can extract relative technology levels using the expression above.

I obtain the price of intermediates from the GGDC Productivity Level Database for the benchmark year 1997 and the price of final goods from the OECD, also for the year 1997. The intermediate prices are constructed from sectoral intermediate input PPPs, which reflect each sector's cost of acquiring intermediate deliveries.[9] The final prices are the PPPs for GDP, which cover both final consumption expenditure

[9]The price of intermediate inputs in a country is computed as the geometric average of the PPP for sectoral intermediate inputs (PPP_II), with the share of sectoral intermediate expenditure (II) in total intermediate expenditure as the weights. The data are available at http://www.rug.nl/research/ggdc/data/ggdc-productivity-level-database. See Inklaar and Timmer (2008) and Timmer, Ypma, and van Ark (2007) for a detailed discussion of the construction of the PPPs.

(household and government) and gross capital formation.[10] I take the ratio of the intermediate to the final price and normalize it to one in the US. The price data are available for 26 countries, and unfortunately exclude the lowest income countries in my initial data set. Nonetheless, there is a strong inverse relationship between the price of intermediates relative to final goods and per capita income.

Figure 1.4, Panel (a) plots the relationship between relative domestic expenditure shares and income, and Panel (b) plots the relationship between relative prices and income.[11] Given that the ratio of domestic expenditure shares is flat to decreasing with respect to income (the inverse relationship is weaker here where the lowest income countries are excluded), and the price ratio is sharply decreasing (and also raised to a power $\theta > 1$), we can infer from equation (1.9) that the relative technology to produce intermediate goods will be increasing in income. Panel (c) of Figure 1.4 plots the precise relationship between relative technology and income, calculated under the assumption that $\theta = 4$. This calculation shows not only that lower income countries tend to have a comparative disadvantage in producing intermediates, but also that there is considerable variation in comparative advantage across countries—the relative technology level for the country with the largest comparative advantage in producing intermediates, Denmark, is eight times that of the country with the largest comparative disadvantage in producing intermediates, the Czech Republic. The large amount of variation suggests that productivity differences at the end-use level provide an important channel for the gains from trade. The calculation is only suggestive, however, as it relies on the assumptions of a very basic model, uses highly aggregate data, and price data that may be measured with error. In the next section,

[10] The PPPs for GDP are available at http://stats.oecd.org/#.

[11] Per capita income is given by output-side real GDP at chained PPPs in 2005 US dollars (*rgdpo*) per person (*pop*) for the year 1997 from the Penn World Tables Version 8.0, available at http://www.rug.nl/research/ggdc/data/pwt/pwt-8.0.

I describe the full model, which incorporates many industries, labor and capital, and input-output linkages, and generates prices that vary by industry and end use. I use the full model to further assess the relationship between comparative advantage and income, and to evaluate the relationship between relative trade costs and relative prices and income. In Section 1.8, I use the full model to quantify the contribution of end-use variation to the gains from trade.

1.4 Full model

In this section I construct the full model, which incorporates many sectors, input-output linkages, and end use variation. I allow the technology and trade cost parameters to vary by industry and end use, which generates prices and trade shares that also vary by industry and end use. I do this to capture the variation in domestic expenditure shares at this level (Figure 1.2), and to incorporate end use variation as a channel for the gains from trade. The model is most closely related to the model described in Caliendo and Parro (2012). In the Caliendo and Parro model (and other multi-sector Eaton and Kortum models), an industry's output can be used both as an intermediate and as a final good, and the productivity and trade cost estimates are a composite of the productivity levels and trade costs associated with each type of end use. Assessing comparative advantage by end use and determining its effect on the gains from trade, however, requires a clear delineation between intermediate and final goods. I ensure that the sectoral productivity and trade cost measures do not confound differences across end use by completely separating intermediate and final goods within a sector; that is, an intermediate good is never used as a final good, and a final good is never used as an intermediate.[12] This characterization is consistent

[12]The Caliendo and Parro model is flexible enough to handle this adjustment (by setting consumption shares to zero for intermediates and input shares to zero for final goods). Solving the

with the data, which classifies all sectoral trade flows and domestic production as destined for either intermediate or final use.

1.4.1 Production

Countries are denoted n and i, and industries are denoted k. End use is given by $u = [I, F]$. The cost of production in country i, industry k is a Cobb-Douglas function of labor, capital, and intermediate inputs:

$$c_i^k = \left(w_i^{\alpha_i^k} r_i^{\iota_i^k} \right)^{\beta_i^k} \left(\rho_i^k \right)^{1-\beta_i^k}, \tag{1.10}$$

where w_i is the wage, r_i is the rental rate, and ρ_i^k is the price of a bundle of intermediates. Labor and capital are mobile across industries within a country, and their shares in value added are α_i^k and ι_i^k. The share of value added in gross output is β_i^k and the share of intermediates in gross output is $1 - \beta_i^k$. The price of the bundle of intermediates used to produce an industry k good in country i is a Cobb-Douglas function of the prices of intermediate inputs from each industry k':

$$\rho_i^k = \prod_{k'} \left(p_i^{I,k'} \right)^{\eta_i^{k,k'}}, \tag{1.11}$$

where $\eta_i^{k,k'}$ is industry k's share of total expenditure spent on intermediates from industry k'. The input shares vary by country, and $\sum_{k'} \eta_i^{k,k'} = 1$. The literature commonly assumes constant industry-level factor and input shares across countries. I exploit the World Input-Output Database to calculate country-specific industry-level shares and find that the shares are not particularly similar across countries.[13]

model, however, requires knowledge of trade and domestic production by end use, which is not available in the widely used trade data.

[13]The coefficient of variation across countries within an industry (taking the average coefficient of variation across all industries) is 0.37 for labor shares and 0.55 for capital shares. For input

I allow input costs to vary by industry and not use, implying that Heckscher-Ohlin motives for trade exist only across industries. This decision is driven primarily by data availability. Use-varying costs would require labor, capital, and input shares that vary by use, and to my knowledge this data does not exist.

Ricardian comparative advantage at the end-use level enters through the productivity parameter $z_i^{k(u)}(l)$. Each industry k in country n produces a continuum of goods indexed by l on $[0,1]$ for intermediate use and for final use. In country i, industry k's efficiency in producing a good for end use u is given by $z_i^{k(u)}(l)$. Iceberg trade costs are given by $\tau_{ni}^{k(u)}$. The unit cost of a good l produced by industry k in country i for end use u in country n is then $p_{ni}^{k(u)}(l) = c_i^k \tau_{ni}^{k(u)} / z_i^{k(u)}(l)$. Markets are perfectly competitive, so $p_{ni}^{k(u)}(l)$ is the price that buyers in country n would pay if the good were bought from country i. Instead, buyers shop around the world and purchase the good from the country with the lowest price. The price actually paid is then $p_n^{k(u)}(l) = \min \left\{ p_{ni}^{k(u)}(l); i = 1, \ldots, N \right\}$. Facing these prices, buyers of end-use u goods in country n purchase amounts of industry k goods to maximize a CES objective function. The price index for the CES objective function is $p_n^{k(u)} = \left[\int_0^1 p_n^{k(u)}(l)^{1-\sigma} \, \mathrm{d}l \right]^{1/(1-\sigma)}$, where σ is the elasticity of substitution between goods.

1.4.2 Technology

The efficiency parameter $z_i^{k(u)}(l)$ is the realization of a random variable drawn from a Fréchet distribution $F_i^{k(u)}(z) = e^{-T_i^{k(u)} z^{-\theta}}$. The parameter $T_i^{k(u)}$ governs the average efficiency with which goods are produced, and a higher value of $T_i^{k(u)}$ implies a higher level of technology. Variation in end use within country and industry implies

shares—looking only at the diagonal entries to get a sense of variation in the shares of the most important input—this measure is 0.66.

that, though a country may have an advantage in producing an industry k good for intermediate use, it may not be well suited to producing the industry k good for final consumption. We can therefore think of production of the industry k good as being tailored to suit the needs of a particular end use. The parameter θ governs the spread of the distribution; lower values imply more variation. More variation in efficiency draws (lower θ) increases the likelihood that technological advantage will overcome high production or transport costs, implying that trade flows will be more influenced by Ricardian comparative advantage.[14]

1.4.3 Consumption

Consumers have CES preferences over final goods produced by each industry k with elasticity of substitution σ, and Cobb-Douglas preferences over industries. The share of final consumption expenditure on each industry is $\eta_i^{F,k}$, with $\sum_k \eta_i^{F,k} = 1$.

1.4.4 Prices

The technology distribution and the CES price index (for consumers and buyers of intermediates) yield a closed form expression for prices in each destination country n that vary by industry k and end use u:

$$p_n^{k(u)} = \gamma \left[\sum_{i=1}^{N} T_i^{k(u)} (c_i^k \tau_{ni}^{k(u)})^{-\theta} \right]^{-1/\theta}, \tag{1.12}$$

[14]It is possible to embed correlation across end use within industries, resulting in the joint distribution $F_i^k(\mathbf{z}) = \exp\left\{ - \left[\sum_u \left(T_i^{k(u)} z^{-\theta} \right)^{1/\rho} \right]^{\rho} \right\}$, where $\mathbf{z}$ is the vector $[z_i^{I,k}, z_i^{F,k}]$ and ρ is a measure of correlation that rises as correlation decreases, with $0 < \rho \leq 1$. The parameter ρ is not separately identifiable from θ, and introducing correlation (low ρ) reduces the strength of comparative advantage in the same way that higher θ reduces the strength of comparative advantage.

where $\gamma = \left[\Gamma\left(1+\frac{1-\sigma}{\theta}\right)\right]^{1/(1-\sigma)}$ and Γ is the gamma function.[15] Prices in country n are a function of its access ($\tau_{ni}^{k(u)}$) to the technology and costs of all countries i.

1.4.5 Trade

The probability that industry k in country i is the lowest-cost provider of good l for end use u in country n is $\pi_{ni}^{k(u)} = T_i^{k(u)}\left(\frac{\gamma c_i^k \tau_{ni}^{k(u)}}{p_n^{k(u)}}\right)^{-\theta}$.[16] Because there is a continuum of goods, $\pi_{ni}^{k(u)}$ is also the fraction of goods that end use u in country n buys from industry k in country i. Further, the distribution of minimum prices is invariant to the source country, so the average price per good is also invariant to the source. This means that $\pi_{ni}^{k(u)}$ is the fraction of country n, end use u expenditure on industry k goods that come from country i:

$$\pi_{ni}^{k(u)} = \frac{X_{ni}^{k(u)}}{X_n^{k(u)}} = T_i^{k(u)}\left(\frac{\gamma c_i^k \tau_{ni}^{k(u)}}{p_n^{k(u)}}\right)^{-\theta}, \tag{1.13}$$

where $X_n^{k(u)}$ is total spending on industry k goods by end use u in country n, and $X_{ni}^{k(u)}$ is spending on the goods that come from country i. A destination country will purchase a larger share of its industry k, end-use u requirements from a country with a higher technology level, lower costs, or with which it has lower bilateral trade costs. A high price in the destination country increases the share that the country

[15] The efficiency parameter $z_i^{k(u)}(l)$ is the realization of the random variable $Z_i^{k(u)}$, so the delivered price of a good $p_{ni}^{k(u)}(l)$ is a realization of the random variable $P_{ni}^{k(u)} = c_i^k \tau_{ni}^{k(u)}/Z_i^{k(u)}$, and the lowest price is a realization of $P_n^{k(u)} = \min\left\{P_{ni}^{k(u)}; i = 1, \ldots, N\right\}$. Substituting the expression for $P_{ni}^{k(u)}$ into the technology distribution yields a distribution of prices $G_{ni}^{k(u)}(p) = 1 - F_i^{k(u)}(c_i^k \tau_{ni}^{k(u)}/p) = 1 - e^{-T_i^{k(u)}(c_i^k \tau_{ni}^{k(u)})^{-\theta} p^\theta}$. Buyers purchase the good from the country with the lowest price, so the price distribution is the distribution of minimum prices: $G_n^{k(u)}(p) = 1 - \prod_{i=1}^N [1 - G_{ni}^{k(u)}(p)] = 1 - e^{-\Phi_n^{k(u)} p^\theta}$, where $\Phi_n^{k(u)} = \sum_{i=1}^N T_i^{k(u)}(c_i^k \tau_{ni}^{k(u)})^{-\theta}$. Substituting this distribution into the CES price index yields the expression for $p_n^{k(u)}$.

[16] This probability is $\Pr\left[p_{ni}^{k(u)}(l) \leq \min\left\{p_{ni'}^{k(u)}(l); i' \neq i\right\}\right] = \int_0^\infty \prod_{i' \neq i}[1 - G_{ni'}^{k(u)}(p)]\, \mathrm{d}G_{ni}^{k(u)}(p)$. Substituting the distribution of prices $G_{ni}^{k(u)}$ yields the expression shown in the equation.

will purchase from a given origin country relative to a destination country with a lower price.

1.4.6 Market clearing

Total expenditure by country n on industry k goods X_n^k can be divided into expenditure on intermediates $X_n^{I,k}$ and expenditure on final goods $X_n^{F,k}$: $X_n^k = X_n^{I,k} + X_n^{F,k}$, and we can allocate intermediate and final expenditure to each origin country i using the trade shares $\pi_{ni}^{k(u)}$:

$$X_{ni}^k = \pi_{ni}^{I,k} X_n^{I,k} + \pi_{ni}^{F,k} X_n^{F,k}. \tag{1.14}$$

Goods markets clear, so the value of industry output Q_i^k equals the sum of expenditure by all countries n on industry k goods from country i: $Q_i^k = \sum_{n=1}^N X_{ni}^k$. Substituting (1.14) into the goods market clearing equation, we have:

$$Q_i^k = \sum_{n=1}^{N} \left(\pi_{ni}^{I,k} X_n^{I,k} + \pi_{ni}^{F,k} X_n^{F,k} \right). \tag{1.15}$$

Recalling the Cobb-Douglas production structure, equilibrium industry expenditures on labor and capital are a constant share of industry output:

$$w_n L_n^k = \alpha_n^k \beta_n^k Q_n^k \text{ and } r_n K_n^k = \iota_n^k \beta_n^k Q_n^k, \tag{1.16}$$

where L_n^k and K_n^k are the industry demands for labor and capital. Factor markets clear, so $\sum_k L_n^k = L_n$ and $\sum_k K_n^k = K_n$. Industry expenditure on intermediates is a fraction $1 - \beta_n^k$ of industry output, so we can write expenditure on industry k

intermediates as a function of output in all industries k' using the input shares $\eta_n^{k',k}$:

$$X_n^{I,k} = \sum_{k'} \eta_n^{k',k}(1-\beta_n^{k'})Q_n^{k'}. \tag{1.17}$$

I do not require that trade is balanced. Denote S_n as the exogenous trade surplus of country n, with $\sum_n S_n = 0$ and $S_n = \sum_k S_n^k$. The industry-level trade surplus S_n^k is output minus expenditure, $S_n^k = Q_n^k - X_n^k$, so we can write equation (1.17) as:

$$X_n^{I,k} = \sum_{k'} \eta_n^{k',k}(1-\beta_n^{k'})(X_n^{k'} + S_n^{k'}). \tag{1.18}$$

Final consumption expenditure X_n^F equals national income Y_n, the sum of payments to labor and capital across all industries, minus the trade surplus S_n:

$$X_n^F = Y_n - S_n = \sum_k (w_n L_n^k + r_n K_n^k) - S_n. \tag{1.19}$$

Final consumption expenditure is allocated to each industry k by consumption shares $\eta_n^{F,k}$, so $X_n^{F,k} = \eta_n^{F,k} X_n^F$. This equation, and equations (1.16), (1.17), and (1.19) imply that we can write expenditure on industry k intermediates $X_n^{I,k}$ and expenditure on industry k final goods $X_n^{F,k}$ as functions of payments to the factors of production. That is,

$$X_n^{I,k} = \sum_{k'} \frac{\eta_n^{k',k}(1-\beta_n^{k'})}{\alpha_n^{k'}\beta_n^{k'}} w_n L_n^{k'} \tag{1.20}$$

and

$$X_n^{F,k} = \eta_n^{F,k} \sum_k (w_n L_n^k + r_n K_n^k - S_n^k). \tag{1.21}$$

Substituting equations (1.20) and (1.21) into (1.15), we can write:

$$Q_i^k = \sum_{n=1}^{N} \left[\pi_{ni}^{I,k} \left(\sum_{k'} \frac{\eta_n^{k',k}(1-\beta_n^{k'})}{\alpha_n^{k'}\beta_n^{k'}} w_n L_n^{k'} \right) + \pi_{ni}^{F,k} \eta_n^{F,k} \sum_k (w_n L_n^k + r_n K_n^k - S_n^k) \right]. \tag{1.22}$$

This equation, along with the cost and price equations (1.10)-(1.12), the trade share equation (1.13), and the factor market clearing and trade balance conditions, characterizes the solution. The parameters are α_n^k, ι_n^k, β_n^k, $\eta_n^{k,k'}$, $T_n^{k(u)}$, $\tau_{ni}^{k(u)}$, L_n, K_n, S_n, and θ. The model solves for costs c_n^k, wages w_n, rental rates r_n, prices $p_n^{k(u)}$, trade shares $\pi_{ni}^{k(u)}$, industry demands for labor and capital, L_n^k and K_n^k, and each industry-level trade surplus S_n^k.

1.5 Estimation

In this section I describe the procedure that I use to estimate and recover the parameters of the model. I use the estimated parameters to solve the model and to quantify the contribution of end-use variation to the gains from trade (Section 1.8). I also use the parameter estimates to understand the extent to which intermediate relative to final technology, trade costs, and prices are related to a country's income level (Section 1.7).

1.5.1 Deriving the estimating equation

The trade share equation (1.13) forms the basis of the estimation procedure. I follow Levchenko and Zhang (2013) to estimate the technology and trade cost parameters. I begin by normalizing the trade share equation by its domestic counterpart $\pi_{nn}^{k(u)}$. Dividing by the domestic trade share eliminates prices $p_n^{k(u)}$ and clearly illustrates comparative advantage: a country will import a larger share than it purchases domestically if the exporting country has an overall productivity and cost advantage, inclusive of trade costs (which are normalized to one in the domestic country):

$$\frac{\pi_{ni}^{k(u)}}{\pi_{nn}^{k(u)}} = \frac{T_i^{k(u)}}{T_n^{k(u)}} \left(\frac{c_i^k \tau_{ni}^{k(u)}}{c_n^k} \right)^{-\theta}. \tag{1.23}$$

Log-linearizing, this equation becomes

$$\ln \left(\frac{\pi_{ni}^{k(u)}}{\pi_{nn}^{k(u)}} \right) = \ln \left(T_i^{k(u)} \left(c_i^k \right)^{-\theta} \right) - \ln \left(T_n^{k(u)} \left(c_n^k \right)^{-\theta} \right) - \theta \ln \tau_{ni}^{k(u)}.[17] \tag{1.24}$$

The first two terms on the right hand side of the equation measure the origin and destination country's technology and cost advantage for producing industry k goods for end use u. I estimate the size of this advantage using fixed effects $S_i^{k(u)}$ and $S_n^{k(u)}$. Next, I specify a functional form for the trade cost parameter $\tau_{ni}^{k(u)}$ using trade cost proxies that are standard in the gravity literature: distance, presence of a shared border, and common language. Log trade costs are given by

$$\ln \tau_{ni}^{k(u)} = \left(d^{k(u)} \right)_m + b^{k(u)} + l^{k(u)} + ex_i^{k(u)}. \tag{1.25}$$

where $\left(d^{k(u)} \right)_m$ is the effect of lying in distance interval m, $b^{k(u)}$ is the effect of having a shared border, and $l^{k(u)}$ is the effect of sharing a language. The dummy variable associated with each effect is suppressed to simplify notation. The distance intervals in miles, following Eaton and Kortum, are: [0,375), [375, 750), [750, 1500), [1500,3000), [3000,6000), and [6000, max]. I also include an exporter fixed effect $ex_i^{k(u)}$; Waugh (2010) shows that exporter fixed effects, as opposed to importer fixed effects, produce estimates that are more consistent with the observed pattern of prices and country incomes. Substituting the trade cost specification (1.25) into equation (1.24), replac-

[17] Taking logs drops zeros from the estimation. I discuss dropped observations in Section 1.6.

ing the technology and cost advantage terms with fixed effects, and incorporating an error term $\varepsilon_{ni}^{k(u)}$, we arrive at the estimating equation:

$$\ln\left(\frac{\pi_{ni}^{k(u)}}{\pi_{nn}^{k(u)}}\right) = S_i^{k(u)} - S_n^{k(u)} - \theta\left(d^{k(u)}\right)_m - \theta b^{k(u)} - \theta l^{k(u)} - \theta e x_i^{k(u)} + \varepsilon_{ni}^{k(u)}. \quad (1.26)$$

The fixed effects $S_i^{k(u)}$ and $S_n^{k(u)}$ measure the same object—the technology-adjusted unit cost—so I restrict them to be symmetric. That is, $S_i^{k(u)} = S_n^{k(u)}$ for all $i = n$. Further, the estimating equation reduces to an identity for observations in which $i = n$, so domestic flows are omitted. I estimate the equation using OLS and Poisson and Gamma pseudo-maximum likelihood (PML) methods. I perform the Poisson and Gamma PML estimation methods to incorporate zeros—estimating the equation in logs drops zero trade flows—and to address the problem posed by heteroskedasticity that arises in log-transformed regressions as discussed in Santos-Silva and Tenreyro (2006).[18]

1.5.2 Recovering the parameters

In this subsection I describe the method that I use to recover the values $T_i^{k(u)}$, $\tau_{ni}^{k(u)}$, and $p_i^{k(u)}$. These estimates are used to investigate the relationship between aspects of comparative advantage and a country's income level, and the technology and trade cost parameters are used to solve the model. Each step requires an estimate of θ, which I again take to be four.

Recall that the estimated fixed effect $S_i^{k(u)}$ measures the technology-adjusted unit

[18]Santos-Silva and Tenreyro (2006) show that when the variance of the error term in a multiplicative model depends on the regressors, the expected value of the error term in the log-linearized model will also depend on the regressors.

cost:

$$S_i^{k(u)} = \ln\left(T_i^{k(u)}\left(c_i^k\right)^{-\theta}\right).^{19} \tag{1.27}$$

To find prices I follow the method of Shikher (2012) by substituting the exponentiated fixed effect $\exp\left(S_i^{k(u)}\right)$ into the domestic expenditure share equation and rearranging:

$$p_i^{k(u)} = \left(\frac{\pi_{ii}^{k(u)}}{\exp\left(S_i^{k(u)}\right)}\right)^{1/\theta}. \tag{1.28}$$

To recover the technology parameter $T_i^{k(u)}$, first construct unit costs c_i^k using the Cobb-Douglas functional form: $c_i^k = \left(w_i^{\alpha_i^k} r_i^{\iota_i^k}\right)^{\beta_i^k}\left(\rho_i^k\right)^{1-\beta_i^k}$. Wages, rental rates, and labor and capital shares are data from the World Input-Output Database, and the price of a bundle of intermediates $\rho_i^k = \prod_{k'}\left(p_i^{I,k'}\right)^{\eta_i^{k,k'}}$ is constructed using prices derived as described above. Extract $T_i^{k(u)}$ from the fixed effect $S_i^{k(u)}$ using this value of c_i^k and equation (1.27). The trade cost parameters $\tau_{ni}^{k(u)}$ are constructed by exponentiating equation (1.25): $\ln\tau_{ni}^{k(u)} = \left(d^{k(u)}\right)_m + b^{k(u)} + l^{k(u)} + ex_i^{k(u)}$.[20]

1.6 Data and implementation

I estimate the parameters of the model using the World Input-Output Database (WIOD), a global input-output table that reports trade flows between 35 industries (both manufacturing and service classifications) and 40 countries (and a rest of world aggregate) for the years 1995 through 2009. The 40 countries comprise 85 percent of world trade and include 29 countries classified as high income and 11 classified

[19] The fixed effects are estimated relative to a reference country, which I take to be the US, so all variables used in the recovery of the parameters are also transformed to be relative to the US.

[20] The US is also the reference country for the exporter fixed effect, so the trade cost estimates are, net of all bilateral components, relative to the cost to export from the US for each industry-end-use pair.

as upper middle or lower middle income by the World Bank in 2007. The data set distinguishes the exporting country and industry and the importing country and industry.[21] Because I am interested in the distinction between intermediate and final use, I aggregate all industry-use categories to create the intermediate classification, and all final consumption, investment, and inventory categories to create the final end-use classification.[22]

In order to minimize the number of trade zeros while keeping the data as disaggregate as possible, I combine countries or industries that have zero industry output. This aggregation scheme eliminates all country-by-industry output zeros, and reduces the number of countries from 40 to 38 and the number of industries from 35 to 32. See Tables 1.5 and 1.6 for WIOD countries and industries and the aggregation scheme. I estimate the model for the year 2007 because it is the most recent year that fully predates the trade collapse, and because capital stocks are provided only for a limited set of countries in 2008 and 2009. I exclude the rest of world aggregate because of the difficulty to create distance, border, and shared language variables for this region. The dimensions of the final data set are 38 origin by 38 destination countries by 32 industries by 2 types of end use.

[21] WIOD distinguishes use by allocating HS 6-digit import flows from the UN COMTRADE database to end-use categories (intermediate, final consumption, and investment) using a correspondence based on the Broad Economic Categories (BEC) from the United Nations Statistics Division. When a product can reasonably be classified into more than one end-use category, weights are applied to divide the trade flow into the relevant categories. Services trade is taken from various sources (UN, Eurostat, and OECD), and is split into end-use categories using average use shares from import input-output tables from Eurostat. Within the intermediate, final consumption, and investment categories trade flows are allocated by proportionality assumption. See Timmer (2012) for a detailed discussion of the construction of the World Input-Output Database.

[22] In some country-by-industry observations, the change in inventories is negative, reflecting a decline in inventories, and large enough that the total final use value is negative. I handle negative inventories using the method of Costinot and Rodríguez-Clare (2013) (see the online appendix to their paper), which is to set negative inventories to zero, and recalculate the total output vector and matrix of intermediate flows using the identity $X = (I - A)^{-1}F$, where X is the total output vector, A is the matrix of direct input coefficients, and F is the final demand vector, with negative inventories set to zero and positive inventories left unchanged.

I use the Socio-Economic Accounts (SEA) that accompany the WIOD to construct wages, rental rates, and labor and capital shares. Wages are calculated as total labor compensation in a country (LAB) divided by the total number of hours worked by persons engaged (H_EMP). The rental rate is constructed by dividing total capital compensation (CAP) by the value of the capital stock (K_GFCF), which is converted from real to nominal values using the price index for gross fixed capital formation ($GFCF\_P$). Labor and capital compensation and the value of the capital stock are converted to US dollars using exchange rates provided by WIOD. Labor and capital shares are computed by dividing labor compensation (LAB) and capital compensation (CAP) by gross output (GO). Input shares are constructed directly from WIOD by dividing country-by-industry total expenditure on intermediates by country-by-industry expenditure on intermediates from a particular industry. I compute each country's trade surplus using WIOD, excluding trade with the rest of the world aggregate. I do this to achieve balanced "world" trade in the sample of countries that I use in the simulation. Per capita income, which is used to investigate the relationship between comparative advantage and a country's level of development in Section 1.7.2 is given by output-side real GDP at chained PPPs in 2005 US dollars (*rgdpo*) per person (*pop*) for the year 2007 from the Penn World Tables Version 8.0.

The estimation strategy requires taking the log of relative trade shares, so zeros are not included. In total, 5.9 percent of the relative trade share observations are zeros. The prevalence of zeros varies by industry, and is higher in service industries—10 percent in service industries and 1.7 percent in goods industries. Within industries, across end use, the proportions of zeros are very similar. This means that, to the extent that missing observations introduce bias in the OLS estimates, concerns should be less pronounced for within-industry, across-end use comparisons, which are the focus of this paper. Even if zeros do not pose a significant problem, estimating log-transformed regression equations will produce inconsistent estimates when heteroskedasticity is present. To account for zeros and this problem posed by heteroskedasticity, I estimate the model using Poisson and Gamma pseudo-maximum likelihood (PML) methods in addition to OLS. I follow the procedure outlined in Head and Mayer (2014) to determine which of the three sets of estimates are most reliable.

1.7 Results

In this section I discuss the choice of estimation method and use the parameter estimates to take a closer look at technology, trade costs, and prices by end use as they relate to income. I also describe the trade cost estimates by end use, and show that the estimates imply a Balassa-Samuelson effect.

1.7.1 Evaluating the estimation methods

Determining whether to use the OLS, Poisson PML, or Gamma PML estimates requires assessing the similarity of the estimates across models. Head and Mayer (2014) provide recommendations for three scenarios: (1) the parameter estimates across the three methods are similar, (2) Poisson and Gamma PML estimates are similar but distinct from the OLS estimates, (3) the Gamma and OLS coefficients are similar and the Poisson are smaller in absolute magnitude. To assess the similarity of the high number of estimates, I regress the set of estimates from one method on the set of estimates for the other methods and force the coefficient on the regressor to equal one, ensuring that a good fit signifies that the estimates are not just correlated, but also similar in magnitude.[23] The R-squared from each regression is reported in Table 1.7. I report the R-squared for the trade cost coefficients (distance, border, and language), the fixed effects (competitiveness and exporter), and for all coefficients. The trade cost estimates are similar across models, and the OLS estimates are particularly close to both the Poisson and Gamma estimates: R-squared 0.84 and 0.81, respectively. The fixed effects are less similar and reduce the strength of the overall fit, but the R-squared remains close to or above 0.5 in each case; this points toward scenario (1) from Head and Mayer. Further, the Poisson and Gamma estimates are less similar to each other than the OLS estimates are to each of these methods (R-squared 0.48 versus 0.54 and 0.57), which does not favor scenario (2). Scatter plots that correspond to the R-squared calculations, provided in Figure 1.5, depict the relationship between coefficients across models. Regarding scenario (3), the Gamma and OLS estimates are similar, but the Poisson estimates are not smaller in absolute magnitude. Table 1.8 shows the average absolute value of the estimate for each set of parameters, and the Poisson estimates are not systematically lower than the others. This points toward scenario (1), in which case the log-linear model is well specified and consistency of the estimates is not a concern. I proceed here with the OLS estimates, and all exercises performed using the Poisson and Gamma estimates are available upon request.

1.7.2 A closer look at comparative advantage

The exercise in Section 1.3 indicated that low income countries have a comparative disadvantage in intermediate relative to final goods, and that these countries pay relatively higher prices for intermediates. In this section I use the parameter estimates

[23]The estimating equation produces 5,248 parameter estimates: there are $(i-1)*u*k$ competitiveness fixed effects $S_i^{k(u)}$, $(i-1)*u*k$ exporter fixed effects $ex_i^{k(u)}$, $m*u*k$ distance coefficients $\left(d^{k(u)}\right)_m$, $u*k$ border coefficients $b^{k(u)}$, and $u*k$ common language coefficients $l^{k(u)}$.

to investigate these relationships further.

Before assessing the relationship between comparative advantage at the end-use level and income, I first evaluate the relationship between the individual intermediate and final estimates and income. I separately regress the intermediate and final technology, trade cost, and price estimates on log per capita GDP and industry fixed effects. The estimating equation for the technology and price estimates is:

$$\ln \Upsilon_i^{k(u)} = \beta_0 + \beta_1 \ln GDP_i + \alpha^{k(u)} + \varepsilon_i^{k(u)}, \text{ for } u = \{I, F\}, \tag{1.29}$$

where $\Upsilon_i^{k(u)}$ represents technology $(T_i^{k(u)})^{1/\theta}$ (the mean of each Fréchet distribution) or price $p_i^{k(u)}$, $\alpha^{k(u)}$ are the fixed effects, and $\varepsilon_i^{k(u)}$ is the error term. I expect that β_1 will be positive in the technology and price regressions because higher income countries are on average more productive and pay higher wages, which imply higher input costs. For trade costs, which vary by origin and destination country, the estimating equation is:

$$\ln \tau_{ni}^{k(u)} = \beta_0 + \beta_1 \ln GDP_c + \alpha^{k(u)} + \varepsilon_{ni}^{k(u)}, \text{ for } u = \{I, F\} \text{ and } c = \{n, i\}, \tag{1.30}$$

where the subscript c on the variable $\ln GDP_c$ indicates whether trade costs are regressed on exporter or importer income. I expect β_1 to be negative in the trade cost regressions, reflecting better transport infrastructure and more open trade policies in higher income countries. I run the regressions for all industries together and for four broad industry classifications: Agriculture, Mining, Manufacturing, and Services. Table 1.9 present the results.[24] High income countries have higher average technol-

[24]The dependent variable is a function of estimates, so I have also followed the Lewis and Linzer (2005) FGLS method to account for sampling error in the estimation of the dependent variable, using bootstrapped standard errors of the technology, trade cost, and price estimates to construct the weights that are applied in the second-stage WLS regression. The Stata routine for the procedure

ogy levels for both intermediates and final goods than low income countries in all categories except Mining. The coefficient on income in the price regressions is also positive in the majority of the regressions. It is notably not statistically different from zero in the intermediate Mining and Manufacturing categories, likely due to the very tradable nature of these goods and, in the case of Mining, the lack of a relationship between technology and income. The export trade cost regressions show that the cost to export is decreasing in income for all categories except intermediate Agriculture and Mining, perhaps reflecting trade policies in lower income countries that promote commodity exports. The estimates from the import trade cost regressions show that higher income countries also pay less to import than lower income countries. The relationship is less pronounced than it is for export trade costs, but it exists for all industry categories. The signs of the coefficients are as expected—positive for the technology and price regressions and negative for the trade cost regressions—in every specification that includes all industries and in the majority of the industry category specifications.

To evaluate comparative advantage at the end-use level, I next regress relative values of the estimates on income. The specifications are the same as above, except the left hand side is now the log of the ratio of the intermediate estimate to the final estimate. The estimating equation for technology and prices is:

$$\ln\left(\frac{\Upsilon_i^{I,k}}{\Upsilon_i^{F,k}}\right) = \gamma_0 + \gamma_1 \ln GDP_i + \alpha^k + \mu_i^k, \tag{1.31}$$

edvreg does not allow clustered standard errors, and the standard errors are more conservative when they are clustered and the Lewis and Linzer approach is not used. For this reason I present the clustered standard error estimates rather than the Lewis and Linzer estimates.

and the estimating equation for trade costs is:

$$\ln\left(\frac{\tau_{ni}^{I,k}}{\tau_{ni}^{F,k}}\right) = \gamma_0 + \gamma_1 \ln GDP_c + \alpha^k + \mu_{ni}^k, \text{ for } c = \{n, i\}. \tag{1.32}$$

The exercise in Section 1.3 that related relative domestic expenditure shares to relative technology and relative prices showed that low income countries have a comparative disadvantage in the production of intermediates, and the data showed that the relative price of intermediates is higher in these countries. It is likely that these findings do not hold for every industry category, but I do expect broadly similar results—that γ_1 is positive in the technology regressions and negative in the price regressions. Given that relative prices are decreasing in income, it is reasonable to expect that it is more difficult for lower income countries to import intermediates relative to final goods, implying that γ_1 is negative in the import trade cost regression; the price data do not have implications for the export trade cost regressions, however. The results are shown in Table 1.10. High income countries have an overall comparative advantage in intermediates that is driven by comparative advantage in the Agriculture and Manufacturing sectors. The export trade cost regression coefficients are significant and positive for Agriculture and Manufacturing, indicating that lower income countries are able to export intermediate goods in these industries at a relatively lower cost than final goods compared to high income countries. The coefficients from the import trade cost regressions are mostly negative, indicating that low income countries have relatively more difficulty importing intermediates than final goods relative to high income countries. Relative prices are negatively related to income in all categories. This is consistent with the aggregate data, and with the fact that low income countries have a comparative disadvantage in intermediates and that it costs these countries relatively more to import intermediates. Recalling equation (1.12), prices are a function of the states of technology around the world and the importing country's access to these technologies via trade costs. If low income countries are not productive in intermediates and pay more to import them, they will pay a higher overall price.

1.7.3 Trade costs

Table 1.11 takes a closer look at trade costs. Each column shows the average coefficient across industries for intermediates and final use—for all industries, goods, and services. The familiar gravity result that trade decreases with distance and increases with the presence of a shared border and common language holds up by end use, and for both goods and services classifications. Across all industries, final goods and services are less tradable than intermediates, and the size of the barriers are large. The average implied effect on cost at a distance of [1500,3000) miles is 293 percent for final goods and services and 209 percent for intermediates with $\theta = 4$.[25] Not surprisingly, services are much less tradable than goods, and the result that final use goods or services tend to be less tradable than their intermediate counterparts holds up within these classifications (with the exception of the two furthest distance intervals for goods), and particularly so for services. This reflects the fact that final services (restaurant services or haircuts, for example) must often be consumed at the location of production, but intermediate services (financial services or information technology) need not.

1.7.4 Balassa-Samuelson effect

The results in Tables 1.10, 1.11, and 1.12 provide evidence of a Balassa-Samuelson effect, which says that countries with a higher productivity in the tradables sector

[25]The implied percentage effect on cost is $100(e^{-\hat{d}/\theta} - 1)$ for an estimated coefficient $\hat{d}$.

will have a higher relative price of nontradables. I treat all goods as tradable, but final goods are comparatively less tradable than intermediates, as Table 1.11 shows. Greater tradability in intermediates means that the prices of intermediates should be less variable across locations than the prices of final goods and services. Table 1.12 demonstrates this by reporting the standard deviation by end-use classification for industry-demeaned prices. Intermediate prices are less variable than final prices, and the same holds within goods and services classifications. Higher income countries have a higher technology level in intermediates, overall and for goods and services, so it follows that these countries will have a higher relative price of the less tradable good—or, equivalently, that lower income countries will have a higher relative price of the more tradable good, the intermediate (Table 1.10).

1.8 Simulation

In this section I solve the full general equilibrium model to determine the effect that incorporating end-use variation has on the gains from trade relative to a model without end-use variation. The labor, capital, and input shares (α_n^k, β_n^k, and $\eta_n^{k,k'}$), size of the labor force (L_n), and capital stock (K_n) are constructed from WIOD as described in Section 1.6. The technology and trade cost parameters ($T_i^{k(u)}$ and $\tau_{ni}^{k(u)}$) are estimated according to the procedure described in Section 1.5, and θ is taken to be 4. The model solves for costs c_n^k, wages w_n, rental rates r_n, prices $p_n^{k(u)}$, trade shares $\pi_{ni}^{k(u)}$, industry demand for labor and capital, L_n^k and K_n^k, and each industry-level trade surplus S_n^k. I solve the model with and without end-use variation and compare the gains from trade. In the version without end-use variation, I re-estimate the parameters using trade data that is not distinguished by end use—that is, the left side of the estimating equation (1.26) is $\ln\left(\frac{\pi_{ni}^k}{\pi_{nn}^k}\right)$. The gains from trade relative to autarky for the models with and without end-use variation, and the discrepancy between the two, are shown in Tables 1.13 and 1.14. In line with the literature, sectoral heterogeneity, input-output linkages, and multiple factors tend to raise the gains from trade: the gains from trade are larger under the full model than under the simple model (recall Tables 1.2 and 1.4) on average and for the majority of countries. As Tables 1.13 and 1.14 show, end-use variation also raises the gains from trade. The gains from trade are 14.4 percent higher on average under the model with end-use variation than under the model without, and are also higher for each income classification and for each country. Relative to the addition of other forms of heterogeneity, the contribution of end-use variation is sizeable. The gains from trade contributed by end-use variation alone are more than one-third the size of the gains from trade contributed by sectoral heterogeneity, input-output linkages, and multiple factors of production.[26] As in the analytical exercise, the contribution of end-use variation to the gains from trade is the largest for the lower income countries—26.3 percent on average for the upper middle income countries and 25.6 percent for the lower middle income countries.

1.9 Conclusion

In this paper I show that a proper calculation of the gains from trade requires allowing for differences in the characteristics of intermediate and final goods trade. Domestic expenditure shares and prices vary by intermediate and final use, indicating the presence of productivity differences that generate gains from trade. This source of productivity differences has not previously been identified nor has it been explored as

[26]The average gains from trade under the full model with end-use variation are 33.3 percent, compared to 29.7 percent for the full model without end-use variation, and 19.4 percent for the standard one-sector, one-factor model.

an avenue for the gains from trade. Distinguishing intermediate and final goods trade is of added importance because intermediates are used in the production of other goods and final goods are not—meaning that the gains from trade in intermediates, but not final goods, accumulate through the production process. I construct a simple model that allows for productivity differences in the production of intermediate and final goods, and show analytically that the gains from trade are always weakly understated in a model that does not include this variation. Using a novel data set, I show that the average discrepancies using the simple model and country-level data are 11.1 percent for high income countries, 15.1 percent for upper middle income countries, and 16.2 percent for lower middle income countries, and that the gains from trade are two to three times more responsive to changes in intermediates trade than to final goods trade. To more fully assess the size of the discrepancy, I construct a model that features variation in intermediate and final use at the industry level, linkages between industries, and multiple factors of production. Solving the model numerically, I find that the average gains from trade are 9.0 percent, 26.3 percent, and 25.6 percent higher for high, upper middle, and lower middle income countries, respectively, under the end-use model. Low income countries benefit more from trade across intermediate and final use, and this appears to be related to the nature of comparative advantage. The parameter estimates show that low income countries have a comparative disadvantage in the production of intermediates; thus, opening to trade allows these countries to import intermediates—which generate cumulative gains from trade—from the more productive high income countries. Given their comparative disadvantage in intermediates, access to imported intermediates is particularly central to welfare in lower income countries. Despite this, the parameter estimates reveal that trade costs pose a disproportionate burden for trade in intermediates in low income countries: lower income countries pay relatively more to import intermediate goods than final goods compared to higher income countries. The combination of a comparative disadvantage in intermediates and a relatively higher cost to import intermediates results in a higher relative price of intermediates in low income countries. Higher prices of intermediates present an important policy challenge, as they limit the competitiveness of countries seeking greater access to international production networks. This study suggests that policies that target productivity improvements in intermediates and the lowering of barriers to trading intermediates may generate important welfare gains in low income countries.

1.10 Tables

Table 1.1: Determinants of the Gains from Trade Discrepancy: Domestic Expenditure Shares and Labor Shares

Income Classification	Overall	Final	Int.	Final/Int.	Labor Share
High	0.80	0.86	0.75	1.17	0.34
Upper Middle	0.83	0.87	0.79	1.11	0.31
Lower Middle	0.93	0.96	0.91	1.06	0.27

Notes: Income classifications are for the year 2007 and are defined by the World Bank, see Table 1.5. Average shares, the ratios of final to intermediate domestic expenditure shares, and labor shares for the groups are simple averages across countries. Labor shares are computed as $labor\ compensation/(gross\ output - capital\ compensation)$ for each country.

Table 1.2: Gains from Trade: Comparison of End-Use and Standard Models

Income Classification	Standard	End-Use	Discrepancy	Relative Elasticity
High	0.214	0.245	0.111	2.06
Upper Middle	0.178	0.204	0.151	2.27
Lower Middle	0.074	0.085	0.162	3.05

Notes: Income classifications are for the year 2007 and are defined by the World Bank, see Table 1.5. Gains from trade are computed with $\theta = 4$, and are net gains from trade ($\widehat{W} - 1$). Gains from trade, the discrepancy between the two models (reported in percent), and the relative elasticities are averages across countries within each income classification.

Table 1.3: Determinants of the Gains from Trade Discrepancy: Domestic Expenditure Shares and Labor Shares

	Overall	Final	Int.	Final/Int.	Labor Share
Australia	0.92	0.93	0.92	1.01	0.36
Austria	0.77	0.83	0.71	1.17	0.37
Belgium	0.71	0.80	0.64	1.25	0.32
Bulgaria	0.72	0.77	0.68	1.15	0.22
Brazil	0.96	0.97	0.94	1.04	0.37
Canada	0.84	0.87	0.81	1.08	0.39
China	0.93	0.96	0.92	1.04	0.17
Cyprus, Luxembourg, & Malta	0.58	0.77	0.46	1.65	0.25
Czech Republic	0.74	0.81	0.70	1.15	0.24
Germany	0.82	0.86	0.79	1.10	0.37
Denmark	0.78	0.85	0.70	1.22	0.38
Spain	0.87	0.90	0.85	1.06	0.34
Estonia	0.74	0.80	0.69	1.16	0.31
Finland	0.82	0.89	0.77	1.15	0.33
France	0.88	0.91	0.85	1.07	0.38
United Kingdom	0.88	0.90	0.86	1.04	0.40
Greece	0.84	0.89	0.75	1.19	0.44
Hungary	0.67	0.80	0.57	1.41	0.29
Indonesia	0.93	0.96	0.89	1.07	0.33
India	0.94	0.96	0.91	1.06	0.32
Ireland	0.68	0.80	0.60	1.33	0.29
Italy	0.89	0.91	0.87	1.05	0.35
Japan	0.95	0.96	0.94	1.03	0.37
Korea	0.88	0.92	0.85	1.08	0.31
Lithuania	0.74	0.80	0.67	1.19	0.36
Latvia	0.78	0.80	0.75	1.06	0.31
Mexico	0.86	0.92	0.76	1.22	0.30
Netherlands	0.78	0.86	0.70	1.22	0.36
Poland	0.81	0.86	0.77	1.12	0.29
Portugal	0.83	0.85	0.80	1.06	0.36
Romania	0.81	0.85	0.76	1.12	0.36
Russia	0.91	0.89	0.92	0.97	0.36
Slovak Republic	0.70	0.80	0.62	1.28	0.19
Slovenia	0.72	0.79	0.66	1.19	0.34
Sweden	0.80	0.86	0.75	1.16	0.36
Turkey	0.90	0.93	0.86	1.09	0.25
Taiwan	0.79	0.86	0.74	1.17	0.29
United States	0.94	0.95	0.92	1.03	0.42

Notes: Labor shares are computed as $labor\ compensation/(gross\ output - capital\ compensation)$.

Table 1.4: Gains from Trade: Comparison of End-Use and Standard Models

Country	Standard	End-Use	Discrepancy	Relative Elasticity
Australia	0.059	0.061	0.019	1.80
Austria	0.195	0.215	0.103	1.71
Belgium	0.303	0.337	0.113	2.12
Bulgaria	0.448	0.505	0.126	3.52
Brazil	0.030	0.035	0.156	1.73
Canada	0.117	0.125	0.073	1.60
China	0.108	0.119	0.094	4.93
Cyprus, Luxembourg, and Malta	0.730	0.898	0.230	2.99
Czech Republic	0.376	0.405	0.077	3.20
Germany	0.139	0.148	0.069	1.69
Denmark	0.182	0.206	0.128	1.63
Spain	0.106	0.114	0.077	1.95
Estonia	0.275	0.306	0.112	2.26
Finland	0.157	0.176	0.118	2.02
France	0.087	0.095	0.089	1.66
United Kingdom	0.083	0.087	0.044	1.47
Greece	0.106	0.132	0.248	1.30
Hungary	0.410	0.495	0.207	2.45
Indonesia	0.059	0.071	0.197	2.06
India	0.054	0.064	0.195	2.16
Ireland	0.396	0.455	0.149	2.48
Italy	0.087	0.093	0.069	1.90
Japan	0.035	0.038	0.077	1.71
Korea	0.105	0.114	0.078	2.20
Lithuania	0.232	0.266	0.149	1.77
Latvia	0.222	0.235	0.059	2.19
Mexico	0.138	0.198	0.436	2.32
Netherlands	0.186	0.211	0.134	1.76
Poland	0.193	0.215	0.115	2.44
Portugal	0.137	0.147	0.067	1.78
Romania	0.162	0.181	0.114	1.81
Russia	0.068	0.065	-0.048	1.74
Slovak Republic	0.593	0.731	0.232	4.14
Slovenia	0.270	0.295	0.093	1.94
Sweden	0.166	0.183	0.098	1.79
Turkey	0.108	0.136	0.253	2.92
Taiwan	0.224	0.251	0.119	2.50
United States	0.039	0.042	0.069	1.41

Notes: Gains from trade are computed using the domestic expenditure and labor shares reported in Table 1.3 with $\theta = 4$, and are net gains from trade ($\widehat{W} - 1$). The discrepancy is the percent difference across the two models.

Table 1.5: List of Countries

Country	Abbreviation	Income Classification, 2007
Australia	AUS	High
Austria	AUT	High
Belgium	BEL	High
Bulgaria	BGR	Upper Middle
Brazil	BRA	Upper Middle
Canada	CAN	High
China	CHN	Lower Middle
Cyprus, Luxembourg, and Malta	CYP-LUX-MLT	High
Czech Republic	CZE	High
Germany	DEU	High
Denmark	DNK	High
Spain	ESP	High
Estonia	EST	High
Finland	FIN	High
France	FRA	High
United Kingdom	GBR	High
Greece	GRC	High
Hungary	HUN	High
Indonesia	IDN	Lower Middle
India	IND	Lower Middle
Ireland	IRL	High
Italy	ITA	High
Japan	JPN	High
Korea	KOR	High
Lithuania	LTU	Upper Middle
Latvia	LVA	Upper Middle
Mexico	MEX	Upper Middle
Netherlands	NLD	High
Poland	POL	Upper Middle
Portugal	PRT	High
Romania	ROM	Upper Middle
Russia	RUS	Upper Middle
Slovak Republic	SVK	High
Slovenia	SVN	High
Sweden	SWE	High
Turkey	TUR	Upper Middle
Taiwan	TWN	High
United States	USA	High

Notes: This table shows the list of countries, and their abbreviations and 2007 income classifications, included in the World Input-Output Database. Income classifications are determined by GNI per capita thresholds set by the World Bank. The thresholds, in US dollars, for Lower Middle, Upper Middle, and High income countries, respectively, are: \$936-\$3,705; \$3,706-\$11,455, and $>$ \$11,455.

Table 1.6: List of Industries

NACE Code	Description	Classification	Aggregation
AtB	Agriculture, Hunting, Forestry and Fishing	Agriculture	AtB
C	Mining and Quarrying	Mining	C
15t16	Food, Beverages and Tobacco	Manufacturing	15t16
17t18	Textiles and Textile Products		17t18
19	Leather, Leather and Footwear		19
20	Wood and Products of Wood and Cork		20
21t22	Pulp, Paper, Paper, Printing and Publishing		21t22
23	Coke, Refined Petroleum and Nuclear Fuel		23
24	Chemicals and Chemical Products		24
25	Rubber and Plastics		25
26	Other Non-Metallic Mineral		26
27t28	Basic Metals and Fabricated Metal		27t28
29	Machinery, Nec		29
30t33	Electrical and Optical Equipment		30t33
34t35	Transport Equipment		34t35
36t37	Manufacturing, Nec; Recycling		36t37
E	Electricity, Gas and Water Supply	Services	E
F	Construction		F
50	Sale, Maintenance and Repair of Motor Vehicles and Motorcycles		50, 51
51	Wholesale Trade and Commission Trade		
52	Retail Trade, Except of Motor Vehicles and Motorcycles		52
H	Hotels and Restaurants		H
60	Inland Transport		60
61	Water Transport		61
62	Air Transport		62
63	Other Supporting and Auxiliary Transport Activities		63
64	Post and Telecommunications		64
J	Financial Intermediation		J
70	Real Estate Activities		70
71t74	Renting of M&Eq and Other Business Activities		71t74
M	Education		M
N	Health and Social Work		N
L	Public Admin and Defence; Compulsory Social Security		L, O, P
O	Other Community, Social and Personal Services		
P	Private Households with Employed Persons		

Notes: This table shows the NACE code, description, classification, and aggregation scheme for industries in the World Input-Output Database.

Table 1.7: OLS, PPML, and GPML R-squared

	PPML v. OLS	GPML v. OLS	PPML v. GPML
Trade Cost Estimates	0.84	0.81	0.66
Fixed Effects	0.43	0.50	0.40
Entire Regression	0.54	0.57	0.48

Notes: This table shows the R-squared from a regression of the coefficients from one estimation method against the coefficients from another estimation method with the coefficient on the independent variable constrained to be one.

Table 1.8: Average Absolute Value of Estimate

	OLS	PPML	GPML
Distance [0,375)	3.78	4.38	4.23
Distance [375,750)	4.02	4.87	4.64
Distance [750,1500)	4.40	5.40	5.18
Distance [1500,3000)	4.99	6.10	6.15
Distance [3000,6000)	6.30	6.22	6.92
Distance [6000,max]	7.02	6.84	7.53
Shared border	0.74	0.82	0.95
Shared language	0.32	0.44	0.57
Competitiveness Fixed Effect	1.51	1.67	2.01
Exporter Fixed Effect	3.72	3.30	3.75

Notes: This table shows the average value of the absolute value of the estimated coefficients across estimation methods.

Table 1.9: Regression of Log Parameters on Log GDP Per Capita, OLS Estimates

	Intermediate				Final			
	Coef.	Std. Err.	R-sq.	N	Coef.	Std. Err.	R-sq.	N
Technology								
All Industries	0.43***	(0.04)	0.53	1,184	0.37***	(0.04)	0.56	1,184
Agriculture	0.53***	(0.11)	0.26	37	0.34***	(0.12)	0.10	37
Mining	0.02	(0.16)	0.00	37	-0.10	(0.15)	0.01	37
Manufacturing	0.31***	(0.05)	0.40	518	0.22***	(0.05)	0.49	518
Services	0.55***	(0.06)	0.56	592	0.52***	(0.06)	0.48	592
Trade Costs (Exporter)								
All Industries	-0.33***	(0.09)	0.31	45,394	-0.37***	(0.10)	0.43	45,357
Agriculture	-0.14	(0.13)	0.02	1,444	-0.39***	(0.14)	0.09	1,444
Mining	-0.17	(0.15)	0.02	1,333	-0.36***	(0.13)	0.11	1,333
Manufacturing	-0.26***	(0.09)	0.18	20,068	-0.35***	(0.11)	0.24	20,105
Services	-0.41***	(0.11)	0.24	22,549	-0.38***	(0.11)	0.30	22,475
Trade Costs (Importer)								
All Industries	-0.09*	(0.05)	0.26	45,394	-0.09*	(0.05)	0.39	45,357
Agriculture	-0.13*	(0.07)	0.01	1,444	-0.15*	(0.08)	0.01	1,444
Mining	-0.15**	(0.07)	0.01	1,333	-0.07*	(0.04)	0.00	1,333
Manufacturing	-0.11*	(0.06)	0.13	20,068	-0.09*	(0.05)	0.15	20,105
Services	-0.07*	(0.04)	0.17	22,549	-0.08*	(0.04)	0.25	22,475
Prices								
All Industries	0.07**	(0.03)	0.65	1,184	0.12***	(0.03)	0.66	1,184
Agriculture	0.14***	(0.05)	0.23	37	0.31***	(0.04)	0.61	37
Mining	0.08	(0.06)	0.04	37	0.26***	(0.07)	0.23	37
Manufacturing	0.04	(0.03)	0.46	518	0.09***	(0.03)	0.51	518
Services	0.09**	(0.04)	0.61	592	0.13***	(0.04)	0.45	592

Notes: All regressions include industry fixed effects. The export trade cost estimates are obtained by regressing bilateral trade costs on exporter GDP and the import trade cost estimates are obtained by regressing bilateral trade costs on importer GDP. Standard errors are clustered at the country level. Significance at the one percent level is represented by ***, at the five percent level by **, and at the ten percent level by *.

Table 1.10: Regression of Log Relative Parameters on Log GDP Per Capita, OLS Estimates

	Intermediate/Final			
	Coef.	Std. Err.	R-sq.	N
Technology				
All Industries	0.06***	(0.02)	0.36	1,184
Agriculture	0.19***	(0.04)	0.26	37
Mining	0.12	(0.12)	0.04	37
Manufacturing	0.09***	(0.03)	0.15	518
Services	0.03*	(0.02)	0.46	592
Trade Costs (Exporter)				
All Industries	0.04*	(0.02)	0.41	45,283
Agriculture	0.25***	(0.07)	0.23	1,444
Mining	0.18	(0.13)	0.06	1,333
Manufacturing	0.09**	(0.04)	0.18	20,031
Services	-0.04*	(0.02)	0.48	22,475
Trade Costs (Importer)				
All Industries	-0.01*	(0.00)	0.41	45,283
Agriculture	0.01	(0.01)	0.00	1,444
Mining	-0.08**	(0.04)	0.01	1,333
Manufacturing	-0.02**	(0.01)	0.15	20,031
Services	0.00***	(0.00)	0.47	22,475
Prices				
All Industries	-0.05***	(0.02)	0.58	1,184
Agriculture	-0.17***	(0.04)	0.27	37
Mining	-0.18*	(0.10)	0.12	37
Manufacturing	-0.05**	(0.02)	0.30	518
Services	-0.04**	(0.01)	0.56	592

Notes: All regressions include industry fixed effects. The export trade cost estimates are obtained by regressing bilateral trade costs on exporter GDP and the import trade cost estimates are obtained by regressing bilateral trade costs on importer GDP. Standard errors are clustered at the country level. Significance at the one percent level is represented by ***, at the five percent level by **, and at the ten percent level by *.

Table 1.11: Trade Cost Components, OLS Estimates

	Coefficients					
	All Industries		Goods		Services	
Variable	Int.	Final	Int.	Final	Int.	Final
Distance [0,375)	-2.13	-3.15	0.93	0.64	-5.19	-6.93
Distance [375,750)	-3.22	-4.21	-0.36	-0.63	-6.07	-7.80
Distance [750,1500)	-3.87	-4.79	-1.33	-1.41	-6.41	-8.16
Distance [1500,3000)	-4.51	-5.47	-2.10	-2.26	-6.93	-8.68
Distance [3000,6000)	-5.90	-6.70	-3.80	-3.64	-8.01	-9.77
Distance [6000,max]	-6.66	-7.37	-4.48	-4.10	-8.84	-10.64
Shared border	0.76	0.73	0.78	0.69	0.73	0.76
Shared language	0.28	0.33	0.25	0.31	0.32	0.35

Notes: The trade cost components are the average distance, border, and language coefficients across industries from equation (1.26).

Table 1.12: Standard Deviation of Prices by End Use, OLS Estimates

	All Industries	Goods	Services
Intermediate	0.27	0.15	0.35
Final	0.34	0.24	0.41

Notes: Prices are demeaned by industry before computing the standard deviation.

Table 1.13: Gains from Trade Simulation: Comparison of End-Use and Standard Models

Income Classification	No End-Use	End-Use	Discrepancy
High	0.343	0.372	0.090
Upper Middle	0.245	0.312	0.263
Lower Middle	0.060	0.076	0.256

Notes: Income classifications are for the year 2007 and are defined by the World Bank, see Table 1.5. Gains from trade are computed solving the full model (with and without end-use variation) with parameter values obtained as described in Sections 1.5 and 1.6, and are net gains from trade ($\widehat{W}-1$). Gains from trade and the discrepancy between the two models (reported in percent), are averages across countries within each income classification.

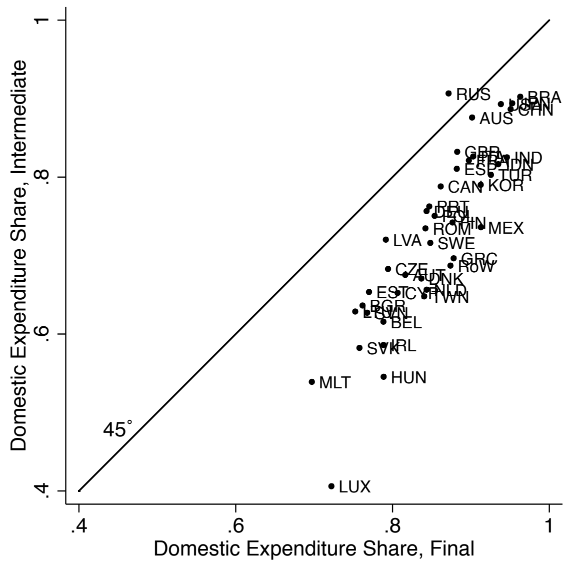
Table 1.14: Gains from Trade Simulation: Comparison of End-Use and Standard Models

Country	No End-Use	End-Use	Discrepancy
Australia	0.063	0.069	0.084
Austria	0.411	0.444	0.082
Belgium	0.988	1.052	0.064
Bulgaria	0.639	0.675	0.055
Brazil	0.020	0.029	0.434
Canada	0.283	0.299	0.058
China	0.079	0.104	0.321
Cyprus, Luxembourg, and Malta	1.172	1.315	0.122
Czech Republic	0.674	0.707	0.050
Germany	0.234	0.250	0.068
Denmark	0.421	0.439	0.043
Spain	0.160	0.172	0.073
Estonia	0.465	0.516	0.109
Finland	0.149	0.159	0.064
France	0.205	0.209	0.019
United Kingdom	0.135	0.145	0.075
Greece	0.213	0.242	0.133
Hungary	0.621	0.667	0.075
Indonesia	0.036	0.045	0.234
India	0.065	0.078	0.212
Ireland	0.477	0.501	0.051
Italy	0.118	0.153	0.295
Japan	0.039	0.045	0.135
Korea	0.185	0.201	0.087
Lithuania	0.263	0.332	0.264
Latvia	0.536	0.918	0.713
Mexico	0.108	0.121	0.119
Netherlands	0.377	0.399	0.059
Poland	0.255	0.274	0.074
Portugal	0.193	0.211	0.093
Romania	0.202	0.228	0.126
Russia	0.069	0.097	0.402
Slovak Republic	0.343	0.375	0.092
Slovenia	0.548	0.599	0.093
Sweden	0.198	0.203	0.026
Turkey	0.099	0.117	0.177
Taiwan	0.166	0.198	0.189
United States	0.059	0.064	0.096

Notes: Gains from trade are computed solving the full model (with and without end-use variation) with parameter values obtained as described in Sections 1.5 and 1.6, and are net gains from trade ($\widehat{W} - 1$). The discrepancy is the percent difference across the two models.

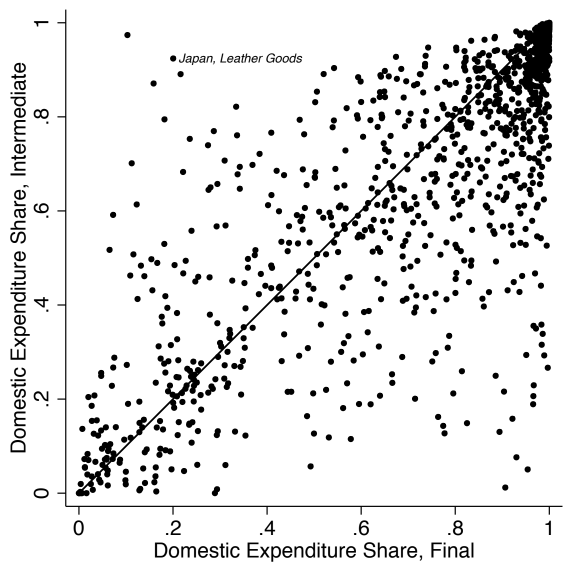
1.11 Figures

Figure 1.1: Country-Level Domestic Expenditure Share, Intermediate vs. Final



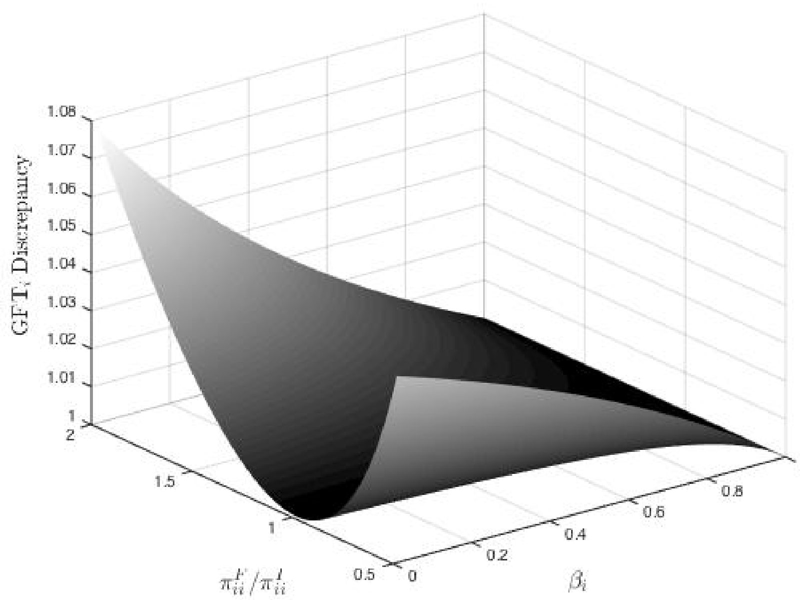
Notes: This figure plots the intermediate domestic expenditure share π_{ii}^{I} against the final domestic expenditure share π_{ii}^{F} for the 40 countries in the sample. The 45°-line is included for reference.

Figure 1.2: Country-by-Industry-Level Domestic Expenditure Share, Intermediate vs. Final



Notes: This figure plots the intermediate domestic expenditure share $\pi_{ii}^{I,k}$ against the final domestic expenditure share $\pi_{ii}^{F,k}$ for the 38x32 country-industry pairs in the sample. The 45°-line is included for reference.

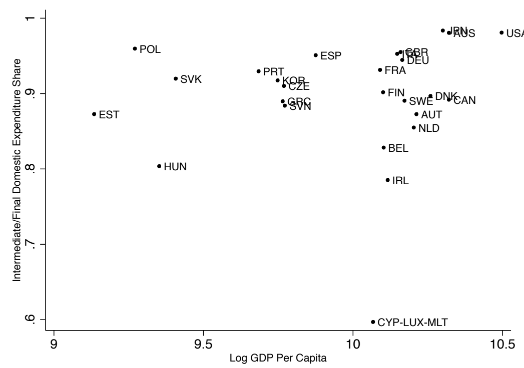
Figure 1.3: Gains from Trade Discrepancy, End-Use vs. Standard Model



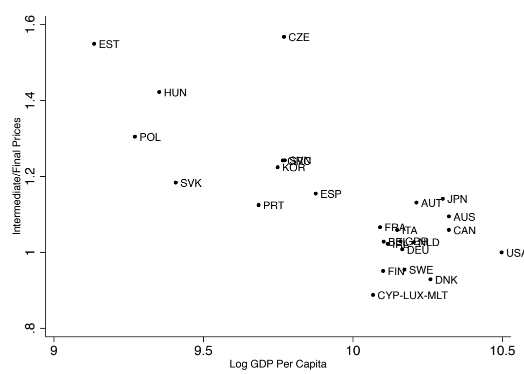
Notes: This figure plots the discrepancy between the end-use and standard gains from trade formulas given by equation (1.8) for $\theta = 4$.

Figure 1.4: Domestic Expenditure Share, Prices, and Comparative Advantage

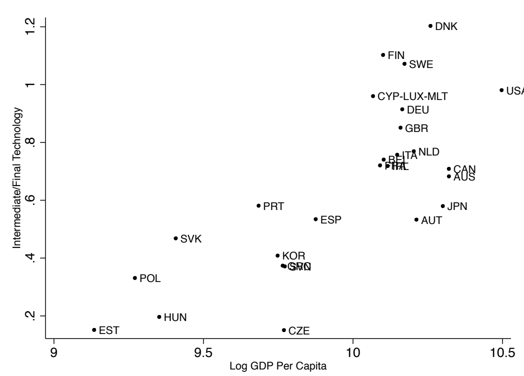
(a) Relative Domestic Expenditure Shares and Income



(b) Relative Prices and Income



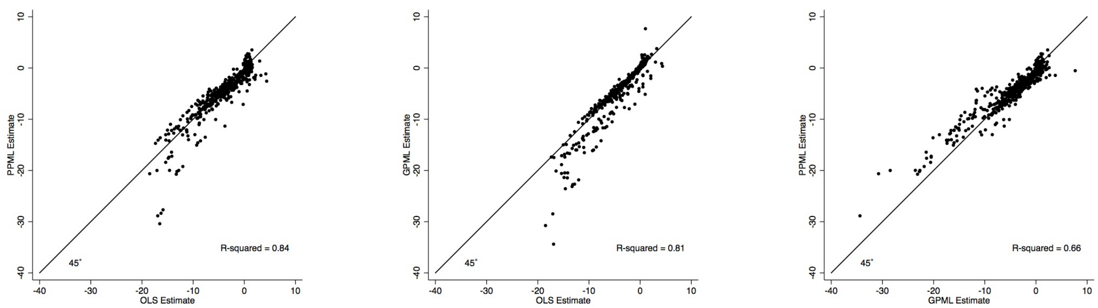
(c) Relative Technology and Income



Notes: This figure plots the ratio of country-level intermediate to final domestic expenditure shares (a), prices (b), and technology levels T_i^I/T_i^F implied by equation (1.9) (c) against log GDP per capita for the 27 countries with relative price data.

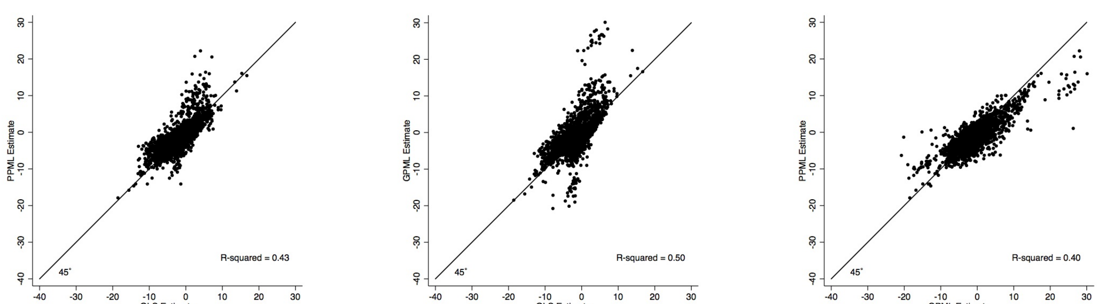
Figure 1.5: OLS v. Poisson PML v. Gamma PML

(a) Trade Cost Coefficients



Notes: This figure plots the estimated trade cost coefficients (distance, border, and language effects) for one estimation method (OLS, PPML, or GPML) against the same coefficients for another estimation method (OLS, PPML, or GPML). The 45°-line is provided for reference.

(b) Fixed Effects



Notes: This figure plots the estimated fixed effects (competitiveness and exporter) for one estimation method (OLS, PPML, or GPML) against the same coefficients for another estimation method (OLS, PPML, or GPML). The 45°-line is provided for reference.

Chapter 2

Decomposing the Distance Effect when Zero Trade Flows are Prevalent

2.1 Introduction

It is well documented that distance affects international trade. Disdier and Head (2008) consider 103 papers and 1,467 distance elasticity estimates and find that a 10 percent increase in distance reduces bilateral trade by 9 percent on average. Distance elasticity estimates capture the impact of trade costs that are either a direct consequence of, or more peripherally related to, geographical proximity. Direct distance-related trade costs include freight and time-in-transit costs, and communication costs and cultural differences may act as barriers to trade that are more indirectly related to the physical distance goods travel. Evidence suggests that trade costs of each type are declining. Hummels (2007) shows that prices of ocean and air freight have declined steadily in the 20 years since the mid-1980s. Harrigan (2010) demonstrates that the decline in the cost of air transport, a much faster shipping method, has been sharper since 1990, and that transport has shifted accordingly toward air in the US. Rauch and Trinidade (2003) note the importance of the internet in allowing firms searching for trading partners to make better "first cuts," which reduces the cost of finding a suitable match. It is reasonable then to suspect that the importance of distance to international trade is waning. In fact, evidence points to the contrary. Despite the apparent decline in distance-related trade costs, distance seems to inhibit trade to an increasing extent. Disdier and Head (2008) find that estimated distance effects have not fallen, and in most specifications are rising. This presents an important puzzle that many authors have attempted to resolve, but the literature remains inconclusive.

Berthelon and Freund (2008) identify aggregation as a potential problem, noting that rising distance sensitivity might be the result of increased trade in distance-sensitive products (a compositional effect), rather than individual industries necessarily becoming more sensitive to distance. They find, however, that adjustments in the composition of trade over time (1985-2005) have had little effect on the distance elasticity and that increased distance sensitivity is the result of increases in 40 percent of industries. With this knowledge they investigate the reasons that industry-level distance sensitivity has risen. Rather than changes in trade costs affecting distance sensitivity, they find that substitutability—whether a good is considered homogenous or its elasticity of substitution has risen—is positively related to an increase in the distance effect. Berthelon and Freund use ordinary least squares and a log-linearized gravity equation to estimate the distance elasticities that underlie these conclusions. This is the standard approach, but it is limited in that it cannot handle observations of zero trade. Zeros in the trade data are prolific, and to exclude them is to overlook relevant information about trade patterns and, importantly, may bias the elasticity estimates. Zero trade often occurs when trading partners are distant, meaning that estimates that exclude zeros are likely to understate the importance of distance.

The standard log-linearized gravity equation presents an additional problem. Santos Silva and Tenreyro (2006) show that, in the presence of heteroskedasticity, elasticity estimates derived from a log-linearized specification are biased. They propose a Poisson pseudo-maximum-likelihood (PPML) method that is consistent in the presence of heteroskedasticity and is estimated in levels rather than logs, providing a natural way to handle zeros.

In this paper, I expand upon the work of Berthelon and Freund by incorporating zero trade flows into the estimation of disaggregate gravity equations and acknowledging the potential threat that heteroskedasticity poses to the consistency of the estimates. I do this using Tobit and PPML methodologies, in addition to a baseline OLS specification. Tobit estimation is useful in that it acknowledges the censoring of observations below zero, and thus the large mass of observations at or near that level. However, zero-valued observations still cannot be handled directly in a Tobit specification of a log-linear gravity model. The PPML procedure is thus an improvement in that trade values enter in levels rather than logs, eliminating the problem of incorporating zero-valued observations. The Tobit and PPML methodologies are discussed in greater depth in Section 2.5. I estimate the models using bilateral trade data for 224 countries and 764 industries over the periods 1997-2000, 2001-2005, and 2006-2009.

I find that the incorporation of zeros and the methodology used to estimate the distance coefficients matters. Like Berthelon and Freund, I find that distance sensitivity increases over time in the OLS specification. Incorporating zero trade flows and adjusting the methodology paints a different picture. Whereas distance sensitivity increases by 9.6 percent on average from the first time period to the second and 7.3 percent from the second to the third using the OLS specification, the corresponding figures for the PPML technique are 0.4 percent and 1.3 percent. Tobit estimation shows that distance sensitivity has declined, by 3.2 percent and 14.3 percent. Further, measures of substitutability and trade costs affect the level of the distance coefficient, but not the change in the distance coefficient when the PPML method is used to estimate distance sensitivity.

In Section 2.2, I discuss literature related to the inclusion of zeros and disaggregation in gravity equation estimation, and how these adjustments affect the direction of the estimated distance coefficients over time. Section 2.3 presents the gravity model of international trade and describes the Berthelon and Freund aggregation method in detail. I then describe the data and discuss what constitutes a zero for the purposes of this paper in Section 2.4. The three estimation methods are presented in Section 2.5. Sections 2.6 and 2.7 present the results of the gravity equation estimation and disaggregation. Section 2.8 examines the reasons distance sensitivity varies by industry and Section 2.9 concludes.

2.2 Related literature

Many recent papers have acknowledged the problem of discarding zero trade flow observations in a gravity model and present ways to handle them. Common approaches are Tobit estimation, which acknowledges the positive mass of observations near zero but does not explicitly allow for them in a log-linearized model, and Poisson pseudo-maximum-likelihood, which is estimated in levels, and can therefore incorporate zero observations directly. Disdier and Head (2008) find that methods that incorporate zeros tend to produce larger distance coefficients (PPML notwithstanding), which is not surprising given that countries that do not trade are farther apart on average than those that do. Figure 2.1, discussed in Section 2.4, demonstrates this fact. Though the coefficients tend to be larger, they are also more likely to decline over time, suggesting that methods that allow for zeros provide a potential solution to the distance puzzle.

Felbermayr and Kohler (2006) estimate a Tobit model with panel data for the years 1950-1997, adding one to all trade values to incorporate zeros. Whereas an OLS specification shows that distance coefficients have increased over time, they find that distance sensitivity has actually declined as time has progressed when the extensive margin of trade is considered. However, procedures that incorporate zeros into a log-linearized specification by increasing all trade values by a constant have been criticized. King (1998) demonstrates that the choice of constant can substantially affect the estimated coefficient. Coe, Subramanian, and Tamirisa (2007) estimate a panel gravity model for the periods 1980-1989 and 1990-2000 using both Tobit and Poisson pseudo-maximum-likelihood, and find that the estimated distance coefficients decline in both periods.

Liu (2009) also uses Tobit and Poisson pseudo-maximum-likelihood approaches, but considers a much longer time horizon, 1948-2003. His main objective is to examine the effect of GATT/WTO on trade over time, but in the process finds that many of the coefficients estimated using a random effects Tobit model are unrealistically large. Liu cites probable violation of the homoskedasticity and normality assumptions as likely causes, and considers the Poisson estimates to be more reliable. The Poisson method also tends to produce distance elasticities that are smaller than the OLS estimates. Using a cross section of 136 countries, Santos Silva and Tenreyro (2006) find that the PPML method produces distance elasticity estimates that are close to half the size of the OLS estimates. This has been a consistent finding in more recent work using the Poisson method. Liu (2009), An and Puttitanun (2009), and Boulhol and de Serres (2010) all find estimated PPML distance elasticities to be smaller than OLS distance elasticities.

Each of the papers mentioned above estimates distance sensitivity using country-level data. Papers that estimate gravity equations at a disaggregate level are less common. In addition to Berthelon and Freund (2008), Siliverstovs and Schumacher (2008) use disaggregate data to address the distance puzzle. They estimate gravity equations using OLS for 25 manufacturing sectors and show that distance sensitivity increases for only five of these industries. This stands in contrast to Berthelon and Freund's finding that distance sensitivity increases in 40 percent of industries. The level of aggregation is the likely reason for the discrepancy, as Berthelon and Freund consider 776 industries, a significantly finer degree of disaggregation. Despite the extensive distance literature and the renewed focus on zero trade flows, to my knowledge there are no papers that combine a handling of zeros in gravity equation estimation with the extremely disaggregate level of data that I use.

2.3 Gravity model and aggregation

In this section I follow Berthelon and Freund in detailing the problems that arise from aggregation in estimating distance elasticities using a gravity model. Aggregating across industries introduces two potential problems. First, distance does not enter log-linearly in an aggregate gravity equation that is constructed from industry-level gravity equations. This means that the distance coefficient estimated on an aggregate sample will be biased. Second, changes in aggregate elasticity may be the result of changes in the composition of trade rather than actual adjustments in distance sensitivity at the industry level.

The standard gravity model relates bilateral trade flows to the product of the income of each trading partner and the distance between them. At the industry-level, the gravity equation takes the form

$$x_{ijk} = A_k \frac{Y_i^{\alpha} Y_j^{\beta}}{D_{ij}^{\gamma_k}}, \tag{2.1}$$

where x_{ijk} is exports from country i to country j in industry k, $Y_{(i)j}$ is GDP in country $i(j)$, and D_{ij} is the distance between i and j. A_k is a constant term that reflects the share of industry k exports among total exports, α and β are income elasticities, and γ_k is the distance elasticity in industry k.

Summing across industries, the gravity equation becomes

$$X_{ij} = Y_i^{\alpha} Y_j^{\beta} \sum_k \frac{A_k}{D_{ij}^{\gamma_k}}, \tag{2.2}$$

where $X_{ij} = \sum_k x_{ij}$ is total exports from country i to country j.

Taking the log of both sides of equation (2.2), we can write

$$\ln X_{ij} = \alpha \ln Y_i + \beta \ln Y_j + \ln \left(\sum_k \frac{A_k}{D_{ij}^{\gamma_k}} \right), \tag{2.3}$$

noticing that distance does not enter the equation log-linearly.

Estimating the standard gravity equation in the aggregate X_{ij} rather than from the individual x_{ijk}'s yields the following:

$$\ln X_{ij} = \alpha \ln Y_i + \beta \ln Y_j + \gamma \ln D_{ij}. \tag{2.4}$$

Equations (2.3) and (2.4) are clearly not the same. Estimating the latter does not allow for the possibility that distance elasticities vary across industries, and will therefore produce biased estimates of the effect of distance on trade.

To illustrate the second problem, note that the aggregate distance elasticity γ is the weighted average of industry-level distance elasticities, γ_k:

$$\gamma = \sum_k s_k \gamma_k, \tag{2.5}$$

where s_k is industry k's share of total exports.

Thus, an adjustment in the aggregate distance elasticity γ may be the result of a change in the composition of trade, a change in the distance sensitivity of individual industries, or a combination of both. We can therefore decompose a change in the aggregate distance elasticity over time in the following way:

$$\Delta\gamma_t \cong \sum_k \Delta s_k \gamma_k + \sum_k s_k \Delta\gamma_k. \tag{2.6}$$

The first term on the right is the compositional effect, which arises from a change in the trade share of a particular industry, and the second term is the distance sensitivity effect, which is the result of industry-level changes in distance elasticity.

For each industry k and time period t, I estimate the distance elasticity and calculate trade shares to determine the overall change in distance sensitivity. I take three approaches to estimating distance elasticity: OLS following Berthelon and Freund, and Tobit and PPML to incorporate zeros. Before discussing these approaches in Section 2.5, I describe the data used to estimate the industry-level gravity equations.

2.4 Data

2.4.1 Trade and distance data

I use 6-digit Harmonized System (HS) 1988/92 import data from UN COMTRADE concorded to the 4-digit SITC Revision 2 classification to most closely match the data used in Berthelon and Freund. Import data is believed to be tracked more carefully because import duties constitute a larger source of revenue than export duties, which are less common. The data cover 224 reporting and partner countries and 764 industries. Thus, a matrix of all possible reporter-partner-industry observations in a given timeframe contains over 38 million entries. In order to smooth over potential noise in year-to-year trade, I use average industry-level bilateral trade in three time periods: 1997-2000, 2001-2005, and 2006-2009.

Despite efforts to replicate Berthelon and Freund as closely as possible, there are a few differences in the data. First, the 6-digit HS data concords to 764 SITC 4-digit industries, just fewer than Berthelon and Freund's 776 industries. Second, Berthelon and Freund use trade data for 100 reporting and 179 partner countries. Because I cannot be certain which countries they include, and to incorporate as many relevant trade flows (or trade zeros) as possible, I use the entire set of 224 reporting and partner countries. Last, I evaluate different time periods. Berthelon and Freund use average trade in two periods separated by a number of years (1985-1989 and 2001-2005), whereas I use average trade in three consecutive and more recent time periods.

Bilateral distances are Great Circle distances between the main city, typically the capital, of each trading partner. CEPII provides the data. Portworld (www.portworld.com) provides actual sailing distances between ports, but only for 58 countries, a small subset of the 224 countries available in the trade data. I find that port-to-port distances between these 58 countries are 85 percent correlated with Great Circle distances between the same countries. Due to the limited availability of actual distance data from this source, and its strong correlation with Great Circle distances, I use only Great Circle distances in my analysis.

Following Berthelon and Freund, I summarize the trade and distance data by calculating the average distance that world trade travels. The average distance, or ADIS, is calculated as

$$ADIS = \sum_{ij} \frac{x_{ij}}{X_W} D_{ij}, \tag{2.7}$$

where x_{ij} is the value of exports from country i to country j, X_W is total world trade, and D_{ij} is the distance between i and j. Berthelon and Freund show that the average distance trade travels has been roughly constant from 1985 to 2005. Figure 2.1 shows this calculation for the period 1997 to 2009. I find that average distance has increased steadily—albeit slowly—from 2005, rising 8 percent from the series low in 2006 to the most recent year, 2009. Despite creeping higher, trade still travels a shorter distance on average than it did in 1997. I also plot the average distance between countries that do not trade. Unsurprisingly, countries that do not trade are farther apart, by 77 percent on average in 2009. Interestingly, as the average distance that trade travels has risen over recent years, the average distance of zero trade has fallen. This indicates that new trade relationships have not necessarily formed between the closest of the previously non-trading partners.

2.4.2 Zeros

Log-linearized specifications of the gravity equation do not allow a complete characterization of the relationship between bilateral trade and distance because they cannot handle zero trade values. The inability to incorporate zeros not only reduces the amount of information available, but it may also bias the distance elasticity estimates if the presence of zero trade values is not independent of distance. As Figure 2.1 demonstrated, there is indeed a relationship. Countries that do not trade are farther apart than those that do. This implies that distance estimates that do not incorporate zeros will be biased downward, underestimating the sensitivity of trade to distance.

Before proceeding further, it is important to clarify what constitutes a zero, as UN COMTRADE data does not explicitly contain observations recorded as zero. The observation is either missing, very small, or the record does not exist. In all cases that the trade value is missing, the industry is also missing. These observations, which comprise fewer than one percent of the total number of observations, cannot be classified by industry so they are dropped from the analysis. In other cases, the trade value is very small, sometimes as low as one dollar for an observation. In the Tobit specifications, I define small observations (less than $100) as effective zeros, and consider them censored observations. Last, and most prevalent, are reporter-partner-industry combinations for which there is simply no record. Because it is possible for a trading relationship to exist for all country-pairs across all industries (with a caveat discussed momentarily), I also classify these occurrences as zeros. Table 2.1 summarizes the incidence of zeros. Roughly half of all country-pairs do not trade with each other, and the occurrence of industry-level zeros is even more prevalent. Over 90 percent of all possible trade flows (country-pair by industry) do not exist.

Following Baldwin and Harrigan (2011), I further refine the definition of a zero by introducing the concept of export and import zeros. I classify an export zero as a product that is exported by a particular country to at least one other country. Trade might fail to exist because the exporting country does not produce the product in question. If we assume that a product never exported is a product never produced, it is impossible, at any distance, for this product to be traded. Consequently, I do not consider these non-traded, potentially non-produced products as zeros. This significantly reduces the number of zeros, but at more than 60 percent of all observations, export zeros still constitute a large portion of the sample. Import zeros arise when a country imports a product from at least one other country, but not others. When import zeros are considered, any product that a particular country never imports is thus excluded from the sample. As Table 2.1 shows, the addition of import zeros further reduces the sample size, but does not change the fact that zeros are prolific. Products classified as both export and import zeros account for over 40 percent of all observations.

2.5 Methodology

In this section I describe the three approaches I take to estimating industry-level distance elasticities. Each method—OLS, Tobit, and PPML—uses the same underlying specification. That is, I regress a measure of trade value (in logs or levels) on log distance and importer and exporter fixed effects. I follow Anderson and van Wincoop (2003) in using importer and exporter fixed effects to account for multilateral resistance, and to control for country-specific characteristics such as income, technology level, and openness to trade. In using the three approaches, my interests are two-fold. I plan first to evaluate how the specification change alone affects the estimates, and second to understand how the addition of zeros changes the estimates.

2.5.1 OLS

Replicating the methodology used by Berthelon and Freund, I estimate the log-linearized gravity equation with importer and exporter fixed effects using ordinary least squares as a baseline specification:

$$\ln x_{ij} = \alpha_i + \alpha_j + \gamma \ln D_{ij} + \varepsilon_{ij}, \tag{2.8}$$

where x_{ij} is the value of exports from i to country j, α_i and α_j are importer and exporter fixed effects, D_{ij} is the distance between i and j, and ε_{ij} is the error term. The parameter of interest is γ, the distance elasticity. I estimate this equation for 764 industries for each of the three time periods. Zero observations are not included.

2.5.2 Tobit

Next, I use a Tobit specification to account for the fact that trade values cannot be negative, and are thus censored at zero. In the data, however, the censoring point is just higher than zero, at one dollar. I have chosen to increase the censoring point in the model further to \$100 because there is a high concentration of trade values less than \$100, after which the frequency declines more steadily as trade values increase. As in the OLS specification, I regress log trade values on log distance, the only procedural difference being the use of a Tobit model. I run the Tobit specification using two different samples. In the first, I include only observations used in the OLS specification (no zeros). I present results for this specification to gauge how the change in specification (OLS to Tobit) affects the distance elasticity. In the second, I expand the data set to include export and import zeros, coding each unrecorded observation as one dollar, the lowest value in the original data set. This is the primary Tobit specification, and it can be compared against the baseline Tobit specification to gauge the effect of including zeros on the distance elasticity. Because the effect of distance varies with each observation, the results that I present for the Tobit specifications are average partial effects. Specifically, the effect of a change in an independent variable x_j on the dependent variable y in a Tobit model (when the censoring point is zero) is given by

$$\frac{\partial E(y \mid \mathbf{x})}{\partial x_j} = \Phi(\mathbf{x}\boldsymbol{\beta}/\sigma)\beta_j \tag{2.9}$$

which varies with each value of the vector $\mathbf{x}$. In order to obtain one elasticity estimate for a regression equation, I calculate the average partial effect, which is the following:

$$\left[N^{-1} \sum_{i=1}^{N} \Phi(\mathbf{x}_i\hat{\boldsymbol{\beta}}/\hat{\sigma}) \right] \hat{\beta}_j. \tag{2.10}$$

The average partial effect is then the average probability of observing an uncensored observation (note: $P(y > 0 \mid \mathbf{x}) = \Phi(\mathbf{x}\boldsymbol{\beta}/\sigma)$) multiplied by the coefficient on the variable in question.

2.5.3 Poisson pseudo-maximum-likelihood

While the Tobit specification recognizes the positive probability mass at very low trade values, it does not explicitly incorporate missing values, as the dependent variable remains logged trade value. The PPML method, proposed by Santos Silva and Tenreyro (2006), is estimated in levels, eliminating the problem of incorporating zero trade values. Further, PPML allows for consistent estimation in the presence of heteroskedasticity, a feature not shared by OLS estimation of the log-linearized gravity equation. Santos Silva and Tenreyro demonstrate the inconsistency of OLS using a log-linearized gravity equation of the form

$$\ln X_{ij} = \ln \alpha_0 + \alpha_1 \ln Y_i + \alpha_2 \ln Y_j + \alpha_3 \ln D_{ij} + \ln \eta_{ij}, \tag{2.11}$$

where X_{ij} is the value of exports from country i to j, $Y_{i(j)}$ is the income of country $i(j)$, D_{ij} is the distance between i and j, and η_{ij} is an error term with $E\left(\ln \eta_{ij} \mid Y_i, Y_j, D_{ij}\right) = 0$.

The consistency of OLS relies on the assumption that $\ln \eta_{ij}$ is statistically independent of the regressors. When the error term is specified as a logarithm, this assumption cannot be made, as the expected value of the logarithm of a random variable is a function of higher order moments of the random variable. Therefore, if the variance of η_{ij} depends on the regressors, the expected value of $\ln \eta_{ij}$ depends on the regressors, and OLS is neither consistent nor unbiased. It is not difficult to imagine that the variance of the error term varies with a country's income. Trade values vary immensely, with values ranging from zero to multiple billions of dollars within a single product category. Large countries with large trade values are likely to deviate more from the gravity relationship than small countries with small trade values would be expected to deviate. Santos Silva and Tenreyro (2006) find evidence to suggest that heteroskedasticity is in fact present in the data, and propose the Poisson pseudo-maximum-likelihood method, an alternative to OLS that allows for consistent estimation in the presence of heteroskedasticity.

Under the assumption that the conditional variance of the dependent variable is proportional to the conditional expectation of the dependent variable, Santos Silva and Tenreyro show that the nonlinear least squares (NLS) estimator of β from the model $y_i = \exp(x_i\beta) + \varepsilon_i$, which is analogous to the gravity model when $x_i = \ln D_{ij}$, is numerically equivalent to the Poisson pseudo-maximum-likelihood estimator. The NLS estimator of β is defined as the following

$$\hat{\beta} = \underset{b}{argmin} \sum_{i=1}^{n} [y_i - \exp(x_i b)]^2, \tag{2.12}$$

which implies the first order conditions

$$\sum_{i=1}^{n} [y_i - \exp(x_i\hat{\beta})] \exp(x_i\hat{\beta}) x_i = 0. \tag{2.13}$$

Santos Silva and Tenreyro note that this set of first order conditions gives more weight to larger and potentially noisier observations, implying the estimator may be inefficient. Because the form of the conditional variance of y_i is unknown, they propose making the assumption that it is proportional to the conditional expectation of y_i, and use this information to create a weighted NLS estimator. Dividing by

$E(Y_i \mid X_i) = \exp(x_i\hat{\beta})$, the set of first order conditions becomes

$$\sum_{i=1}^{n}[y_i - \exp(x_i\hat{\beta})]x_i = 0. \tag{2.14}$$

Santos Silva and Tenreyro note that the estimator defined by equation (2.14) is numerically equivalent to the Poisson pseudo-maximum-likelihood estimator. They propose the PPML method over OLS when estimating a log-linearized equation for its ability to consistently estimate the coefficients in the presence of heteroskedasticity and its ability handle zeros.

As with the Tobit specification I estimate the PPML model using two different samples. I regress trade value on the log of distance and importer and exporter fixed effects, first on a sample that does not include zeros, and second on the full trade matrix inclusive of export and import zeros. Estimates obtained from the sample of positive trade values will be used to compare the effect of the methodology (PPML vs. OLS and Tobit) on the distance coefficient. Those obtained from the full sample are used to evaluate the effect on distance sensitivity of adding zeros.

2.6 Gravity equation estimation and distance coefficients

In this section I present results from the gravity equation estimation procedures described in the previous section. With bilateral import data for 764 industries, I estimate the effect of distance on average trade in three periods (1997-2000, 2001-2005, and 2006-2009) using OLS, Tobit, and PPML specifications. Summary results are presented in Table 2.2.

The OLS distance coefficients that I estimate are centered around -1.1, implying a one percent increase in distance reduces trade by 1.1 percent on average. This is slightly larger than the -0.9 average estimate found in Disdier and Head's meta-analysis, and well within their range of -2.33 to 0.4. Berthelon and Freund estimate OLS coefficients that are slightly larger, centered around -1.4, likely owing to one or more of the data discrepancies discussed earlier. Figure 2.2 provides a histogram of the estimates.

Results from the Tobit estimation using the baseline data set are similar in magnitude to the OLS results. This is not surprising given that I estimate the model on a positive sample and consider only the few observations less than $100 to be censored. Interestingly, the Tobit estimates on the zero-inclusive sample are smaller. This seems to be the result of the very low probability of observing trade values greater than the censoring point (roughly 15 percent), because the coefficients on the distance variables themselves are quite large (average in excess of -3.0).

The PPML estimates of distance sensitivity using the baseline data set are similar, but just smaller in magnitude than the OLS and Tobit baseline estimates, an indication that the choice of methodology has little impact when zeros are not considered. The PPML estimates for the zero-inclusive sample are also similar in magnitude to the OLS estimates, and smaller in the later two later periods. The findings in these time periods are consistent with the literature that tends to show smaller PPML than OLS estimates.

With the exception of the Tobit model estimated on a zero-inclusive sample, I find that the distance coefficients are similar across methodologies. Further, in each of the OLS and PPML specifications I find that distance sensitivity is increasing over time—a finding that seems to corroborate the distance puzzle. However, the extent to which distance sensitivity increases varies by methodology. As Figure 2.3 shows, the OLS procedure yields an increase in the distance coefficient of more than 6 percent from each time period to the next, whereas the PPML-zeros specification shows that the changes have been much smaller—at or below 2 percent in each case. A comparison of means that weights each coefficient by the inverse of its standard error reveals that the PPML coefficients are not statistically different across time periods, but that the OLS estimates show a significant increase from each time period to the next.

A primary difference between the PPML procedure and the OLS procedure is its ability to consistently estimate the coefficients in the presence of heteroskedasticity. That the difference between the OLS and PPML distance coefficients varies over time indicates that the bias induced by heteroskedasticity may be stronger in some time periods than others. If we assume that samples with high variation in trade values are more likely to exhibit heteroskedasticity—in that the error term is likely to be larger for larger values of trade, we might conclude that the size of the bias in the OLS estimates is larger during time periods with more highly varied trade values. In fact, the variance increased from 1997-2000 to 2001-2005 in 77 percent of industries and from 2001-2005 to 2006-2009 in 91 percent of industries. This implies that the size of the bias in the OLS estimates may be increasing over time, and is consistent with the fact that the PPML and OLS estimates diverge over time. In contrast to the increase in distance sensitivity revealed by the OLS and PPML approaches, the Tobit approach indicates that distance sensitivity is falling. This finding, along with the fact that the PPML estimates are not statistically different across time, demonstrates that the distance puzzle is challenged when zero trade flows are considered.

2.7 Distance-sensitivity and compositional effect

Following Berthelon and Freund, I now compute the compositional and distance-sensitivity effects as shown in equation (2.6). Calculating the trade shares s_k is a straightforward computation of each industry's share of total trade in a given time period, and Δs_k is the difference in the calculated shares for two adjacent time periods. The distance elasticities γ_k are those described in the previous section. Obtaining $\Delta\gamma_k$ requires only a simple difference of estimated distance elasticities, but I am interested in whether the difference is significant, so I estimate $\Delta\gamma_k$ directly using each methodology.

For OLS, I estimate a differenced version of equation (2.8):

$$\ln x_{ijt} - \ln x_{ijt-1} = (\alpha_{it} - \alpha_{it-1}) + (\alpha_{jt} - \alpha_{jt-1}) + (\gamma_t - \gamma_{t-1}) \ln D_{ij} + (\varepsilon_{ijt} - \varepsilon_{ijt-1}), \quad (2.15)$$

where $\gamma_t - \gamma_{t-1} = \Delta\gamma_k$ is the parameter of interest.

The nonlinearity in the Tobit and PPML models require that I estimate the change in elasticity on a combined sample of two time periods (1997-2000 and 2001-2005, for example), using interactions with a time period indicator to parse out the effect of distance on the change in trade. A linearized representation of the specification is the following:

$$\ln x_{ijt} = \alpha_i + \alpha_i * t + \alpha_j + \alpha_j * t + \gamma \ln D_{ij} + \Delta\gamma(\ln D_{ij} * t) + \nu_{ij}, \quad (2.16)$$

where $t = 0$ in the earlier time period and $t = 1$ in the later time period.

Decomposition results are presented in Table 2.3. The distance-sensitivity effect is much stronger than the compositional effect across most specifications and time periods. Similar to Berthelon and Freund, I find that trade shares have changed little over time, so it is not surprising that the compositional effect is small. From 1997-2000 to 2001-2005, the change in trade share is less than 0.1 percent for 98 percent of industries. The corresponding figure for the 2001-2005 to 2006-2009 time period is 96 percent. Industry-level changes in distance-sensitivity are thus largely responsible for increased sensitivity to distance over time. In line with the finding that distance sensitivity increases very little using the PPML method, the PPML distance sensitivity effects are also small, and much smaller than the OLS distance sensitivity effects.

Table 2.4 shows the sign and significance of the change in the distance coefficient over time for the baseline OLS and PPML-zeros specifications.[1] Using the OLS specification, distance sensitivity increased (negative and significant category) for 42 percent of industries from 1997-2000 to 2001-2005 and 27 percent of industries from 2001-2005 to 2006-2009.[2] This indicates that distance is becoming more important for a sizable number of industries, and is similar to Berthelon and Freund's finding that distance sensitivity increased in 40 percent of industries from 1985 to 2005. However, the result that distance is becoming significantly more important for a large number of industries does not hold up in the PPML specification. The PPML-zeros procedure produces a negative and significant change in just 4 percent of industries from 1997-2000 to 2001-2005 and 2 percent of industries from 2001-2005 to 2006-2009. If not surprising—given the earlier result that the average change in distance coefficient was much smaller using the PPML specification—this finding is nonetheless

[1]Calculating standard errors for the average partial effects in the Tobit regressions requires bootstrapping or the delta method, and is extremely computationally intensive regardless of which method is used. Generating a sufficient number of replications for the bootstrap is not feasible given the already extensive computational requirements needed to estimate the Tobit regression with the large number of fixed effects. The delta method requires evaluating the matrix of partial derivatives for each observation (Greene (2012)), which is also exceedingly time consuming given the large number of observations and number of regressions. For these reasons, I do not present standard errors for the Tobit average partial effects.

[2]The total number of industries considered here is less than 764 because standard errors could not be calculated for all PPML change regressions (the variance matrix was highly singular). Elimination of some fixed effects solves the problem, but any method to remove some while leaving others would deviate from the underlying gravity model and be fairly arbitrary. In order to make an accurate comparison across specifications, I have used only the subset of industries for which calculation of the PPML standard errors was feasible.

important. Whereas the OLS findings suggest that the distance puzzle remains intact and influential for a subset of industries, the PPML results suggest that changes in distance sensitivity are more subtle and also more evenly split between industries experiencing positive and negative adjustments in distance sensitivity.

Nevertheless, the magnitude, direction, and significance of the change in distance coefficients clearly vary across industries. The question then remains—why has distance become more important for some industries? To answer this question, I focus attention to the preferred specifications: OLS as the baseline, and PPML-zeros as the specification that best handles zeros and heteroskedasticity.

2.8 Determinants of distance sensitivity

In this section I follow Berthelon and Freund in examining industry characteristics that might account for the variation in distance sensitivity across industries. These characteristics include measures of substitutability and trade costs. Considering the first, products that are more easily substituted are expected to be more sensitive to distance. To the extent that trade costs increase with distance, easily substitutable products should be more readily imported from a nearby trading partner. Conversely, there may be few alternatives but to import differentiated products with no close substitutes than from countries that are far away. The explanatory variables that I use to capture a product's substitutability are, directly, its elasticity of substitution, and whether it can be classified as a homogeneous or differentiated good. The elasticities of substitution by industry are given by Broda and Weinstein (2006), and are constructed from US import data for the periods 1972-1988 and 1990-2001. I use the 4-digit SITC Revision 3 estimates for the period 1990-2001, which most closely matches my timeframe, keeping estimates for industries that were not reclassified from Revision 2 to Revision 3. Like Berthelon and Freund, I have also dropped observations for which the increase in the elasticity of substitution was more than 2000 percent (6 observations). The classification of a good as homogeneous or differentiated comes from Rauch (1999). Rauch classifies goods into three categories: those that are traded on an exchange, those that are not but have 'reference' prices, and all other goods. Goods that fall into the first two categories are considered homogeneous—and thus more substitutable—while those in the third are considered differentiated. The Rauch classifications are at the 4-digit SITC Revision 2 level.

To examine the effect of trade costs on the distance elasticity, I use the simple average tariff and the cost of insurance and freight as a share of export value, again following Berthelon and Freund. I use the average ad valorem equivalent applied rate MFN tariff from the World Integrated Trade System (WITS) by industry for each of the three time periods. The tariffs are obtained at the 4-digit SITC Revision 2 level. Transport costs are CIF/FOB rates from US customs data. I use exports from the EU to the US to obtain a measure of industry-level transport costs that has good coverage (698 industries), but that does not vary along dimensions related to distance. The distance coefficient captures transport costs associated with geographical proximity, so I am interested in aspects of transport costs related to features of the industry—principally physical characteristics like size and weight. Summary statistics for the explanatory variables are provided in Table 2.5.

I use weighted least squares to estimate the effect of these industry characteristics on the distance coefficients obtained from the OLS and PPML-zeros specifications. The distance coefficients are estimated with varying degrees of precision, so I use weighted least squares (with the inverse of the standard errors of the estimated distance coefficients as weights) to downweight the residuals for observations with large variances. I also estimate the model with 1-digit industry fixed effects to remove any variation associated solely with broader characteristics of the good. Table 2.6 shows the results. Like Berthelon and Freund, I find that the Broda and Weinstein elasticities of substitution are generally not significantly correlated with the distance coefficients from the OLS specifications. The PPML specifications do provide instances, however, in which the elasticity of substitution is significant and has the expected sign, but the size of the effect is small. For example, the PPML-zeros specification for 1997-2000 (Column 1) indicates that a one-unit change in the elasticity of substitution (7.8 percent of one standard deviation) increases distance sensitivity by 0.005 (1.1 percent of one standard deviation). Being a reference priced good is associated with increased distance sensitivity in most specifications, particularly in the PPML-zeros specifications. The measures of trade costs are significantly correlated with the distance coefficients, have the expected sign in almost all cases, and are of similar magnitude to the Berthelon and Freund estimates. Industries with higher tariffs and freight costs are more sensitive to distance. Using the PPML-zeros 2006-2009 specification as an example (Column 6), we see that a one percentage point increase in average tariffs is associated with an increase of 0.02 in the distance coefficient, which is 2.1 percent of the mean distance coefficient for this specification and time period. The distance elasticity is particularly sensitive to freight costs, as a one percent increase in freight costs corresponds to an increase of more than two in the distance coefficient.

Next, I examine the effect of industry characteristics on the change in distance coefficients, again using the distance coefficients estimated with the OLS and PPML-zeros specifications and weighted least squares. Results are reported in Table 2.7. While I was not able to estimate the effect of a change in the elasticity of substitution on the change in distance sensitivity (my more recent timeframe corresponds with only the later timeframe of Broda and Weinstein's elasticity estimates), I do find that the elasticity of substitution itself is positively related to an increase in distance sensitivity, and that homogeneous goods (exchange traded and reference priced) are associated with increased distance sensitivity using the OLS-estimated distance coefficients. This echoes Berthelon and Freund's finding that the substitutability of a good seems to play a more prominent role as time progresses. However, there is very little evidence from the PPML-estimated coefficients that measures of substitutability affect distance sensitivity. Tariffs and transport costs are also shown not to significantly affect the change in distance elasticity in a systematic way in either the OLS or PPML specifications. The most prominent conclusion to be drawn from the preferred PPML-zeros specification is that measures of substitutability or trade costs cannot be said definitively to affect changes in distance sensitivity over time.

2.9 Conclusion

Using industry-level data, I find that the effect of distance on trade is sensitive to the incorporation of zeros and to the estimation method. The OLS specification of the gravity model, which neither incorporates zero trade flows nor effectively acknowledges the presence of heteroskedasticity, shows a pronounced increase in distance sensitivity on average, and for many industries individually. In contrast, estimation methods that incorporate zeros sufficiently challenge the distance puzzle. Tobit estimation yields coefficients that decline over time on average, and the PPML method produces coefficients that increase on average, but without statistical significance. In further contrast to the OLS results, very few industries experience a statistically significant increase in distance sensitivity when the PPML method is employed. Using the PPML-estimated coefficients I find that exchange traded and reference priced goods, as well as those with higher tariffs and transport costs, tend to be more sensitive to distance. However, these measures of substitutability and trade costs cannot explain changes in the distance coefficient over time. This is likely a consequence of the fact that PPML-estimated coefficients change little and not significantly over time, and provides further indication—along with the Tobit results—that the distance puzzle loses potency when the estimation procedure is better-suited to the characteristics of the trade data.

2.10 Tables

Table 2.1: Incidence of Zeros, Number (Thousands) and Percent of Possible Observations

	97-00		01-05		06-09	
Country-level						
Potential observations	50.0		50.0		50.0	
Zeros	26.0	52%	21.5	43%	22.3	45%
Industry-level						
Potential observations	38,163		38,163		38,163	
Zeros	35,762	94%	34,886	91%	34,917	91%
Export zeros	23,616	62%	25,325	66%	24,827	65%
Export zeros (and $\leq$ \$100)	23,646	62%	25,482	67%	25,001	66%
Export and import zeros (and $\leq$ \$100)	15,175	40%	17,821	47%	16,268	43%

Notes: This table shows the number (in thousands) of zero trade observations and percent of possible observations for country pairs overall and at the industry level for each time period.

Table 2.2: Distance Coefficients, Summary Statistics

	97-00		01-05		06-09	
Method	Avg. Coef.	Std. Dev.	Avg. Coef.	Std. Dev.	Avg. Coef.	Std. Dev.
OLS	-0.989	0.325	-1.084	0.349	-1.163	0.374
Tobit	-0.972	0.319	-1.035	0.335	-1.107	0.359
Tobit-zeros	-0.507	0.176	-0.491	0.173	-0.421	0.164
PPML	-0.894	0.382	-0.912	0.363	-0.915	0.402
PPML-zeros	-1.036	0.444	-1.041	0.415	-1.054	0.501

Notes: This table reports the average distance coefficient and the standard deviation of the coefficients for each estimation method and time period. Regressions are robust to heteroskedasticity and include importer and exporter fixed effects.

Table 2.3: Decomposition of the Distance Coefficient

	Distance-Sensitivity Effect		Compositional Effect		Total	
Method	97-00 to 01-05	01-05 to 06-09	97-00 to 01-05	01-05 to 06-09	97-00 to 01-05	01-05 to 06-09
OLS	-0.092	-0.088	-0.007	-0.016	-0.099	-0.104
Tobit	-0.062	-0.081	-0.008	-0.016	-0.070	-0.097
Tobit-zeros	0.067	0.039	-0.011	-0.029	0.056	0.010
PPML	-0.009	0.006	-0.012	-0.030	-0.021	-0.023
PPML-zeros	-0.007	0.011	-0.020	-0.038	-0.027	-0.027

Notes: This table shows the decomposition of the change in the aggregate distance coefficient from one time period to the next into distance sensitivity and compositional effects for each estimation method.

Table 2.4: Sign and Significance of Change in Industry-Level Distance Coefficients, 95 Percent Confidence Level

	97-00 to 01-05		01-05 to 06-09	
	Number	Percent	Number	Percent
OLS Specification				
Positive and Significant	4	0.6	3	0.4
Positive and Insignificant	83	13.3	138	19.9
Negative and Significant	263	42.1	186	26.9
Negative and Insignificant	274	43.9	365	52.7
Total	624	100.0	692	100.0
PPML-zeros Specification				
Positive and Significant	5	0.8	23	3.3
Positive and Insignificant	254	40.7	393	56.8
Negative and Significant	27	4.3	11	1.6
Negative and Insignificant	338	54.2	265	38.3
Total	624	100.0	692	100.0

Notes: This table shows the number and percent of the estimated change distance coefficients (equations (2.15) and (2.16)) that are positive, negative, and significant at the 95 percent confidence level for the OLS and PPML-zeros specifications.

Table 2.5: Summary Statistics, Explanatory Variables

	Period	Mean	Std. Dev.	Min.	Med.	Max.	N
Average tariff	97-00	12.15	6.87	2.25	10.53	59.44	763
	01-05	10.22	6.59	1.83	8.44	53.18	763
	06-09	8.98	5.73	0.68	7.42	45.96	763
Change 1		-1.93	2.04	-32.44	-1.89	11.13	763
Change 2		-1.25	1.39	-12.84	-1.05	6.03	763
Cost of insurance and freight	97-00	0.07	0.08	0.00	0.05	0.76	698
	01-05	0.07	0.07	0.00	0.05	0.57	697
	06-09	0.06	0.07	0.00	0.04	0.67	698
Change 1		0.00	0.04	-0.28	0.00	0.21	697
Change 2		-0.01	0.04	-0.34	0.00	0.46	697
Elasticity of substitution	72-88	7.17	16.76	1.10	2.80	131.50	670
	90-01	5.39	12.85	1.00	2.40	131.50	514
Change		-0.64	18.66	-129.10	-0.20	128.20	513
Differentiated		0.59	0.49	0.00	1.00	1.00	686
Reference		0.30	0.46	0.00	0.00	1.00	686
Homogeneous		0.11	0.31	0.00	0.00	1.00	686

Notes: This table shows summary statistics for potential determinants of the level and change in distance sensitivity. Average tariffs are the average ad valorem equivalent applied rate MFN tariff from the World Integrated Trade System (WITS), cost of insurance and freight is the CIF/FOB rate as a share of export value for exports from the EU to the US, elasticities of substitution are from Broda and Weinstein (2006), and binary goods classifications are from Rauch (1999).

Table 2.6: Determinants of Distance Elasticities

	Distance coefficient 1997-2000				Distance coefficient 2001-2005				Distance coefficient 2006-2009			
	(1)		(2)		(3)		(4)		(5)		(6)	
Specification: OLS												
Elasticity of substitution	0.00	(0.00)	-0.00	(0.00)	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)
Rauch ex. traded goods	0.16***	(0.05)	0.08	(0.05)	0.12**	(0.06)	0.08	(0.06)	0.06	(0.07)	-0.02	(0.07)
Rauch ref. priced goods	0.04	(0.03)	-0.02	(0.04)	-0.04	(0.03)	-0.08*	(0.04)	-0.06	(0.04)	-0.17*	(0.10)
Average tariff	-0.01**	(0.00)	-0.01***	(0.00)	-0.01***	(0.00)	-0.01***	(0.00)	-0.01***	(0.00)	-0.01**	(0.01)
Insurance and freight	-0.75***	(0.17)	-1.21***	(0.19)	-0.22	(0.15)	-0.36*	(0.20)	-0.85***	(0.22)	-1.29***	(0.26)
Observations	442		442		442		442		442		442	
R-squared	0.09		0.27		0.07		0.26		0.09		0.29	
Industry FE (1-digit)	No		Yes		No		Yes		No		Yes	
Specification: PPML-zeros												
Elasticity of substitution	-0.01***	(0.00)	-0.00**	(0.00)	-0.00	(0.00)	-0.00	(0.00)	-0.00	(0.00)	-0.00	(0.00)
Rauch ex. traded goods	0.03	(0.07)	0.00	(0.07)	-0.14	(0.11)	-0.10	(0.12)	-0.26**	(0.11)	-0.22**	(0.11)
Rauch ref. priced goods	-0.06	(0.05)	-0.02	(0.06)	-0.22***	(0.06)	-0.16**	(0.07)	-0.11*	(0.06)	-0.09	(0.06)
Average tariff	-0.01***	(0.00)	-0.02***	(0.01)	-0.02***	(0.00)	-0.02***	(0.00)	-0.02***	(0.00)	-0.02***	(0.00)
Insurance and freight	-1.63***	(0.41)	-1.57***	(0.42)	-0.18	(0.15)	-0.10	(0.13)	-2.84***	(0.46)	-2.44***	(0.40)
Observations	442		442		442		442		442		442	
R-squared	0.30		0.39		0.16		0.27		0.32		0.40	
Industry FE (1-digit)	No		Yes		No		Yes		No		Yes	

Notes: This table shows the results from regressions of the distance coefficients for each time period and for the OLS and PPML-zeros specifications on industry characteristics. Coefficients are estimated using weighted least squares with the standard errors of estimated distance coefficients as weights. Robust standard errors reported in parentheses. Significance at the one percent level is represented by ***, at the five percent level by **, and at the ten percent level by *.

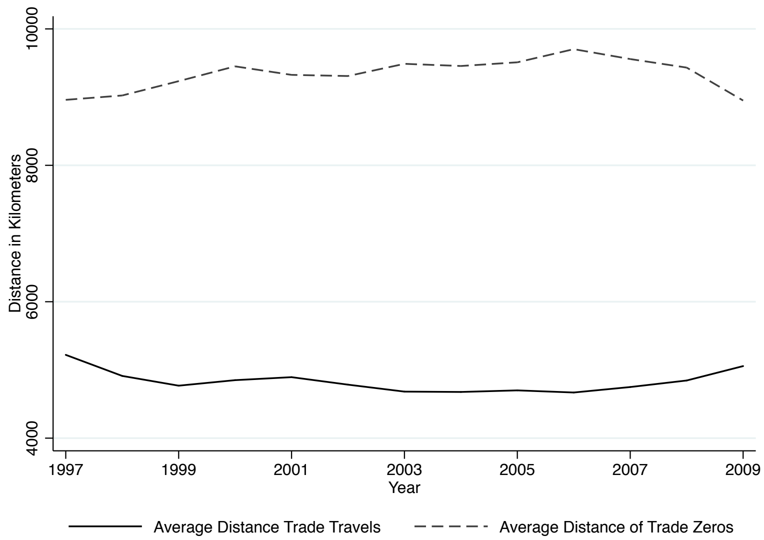
Table 2.7: Determinants of Changes in Distance Elasticities

	Change coefficient, 1997-2000 to 2001-2005				Change coefficient, 2001-2005 to 2006-2009			
	(1)		(2)		(3)		(4)	
Specification: OLS								
Elasticity of substitution	0.00	(0.00)	0.00	(0.00)	-0.00***	(0.00)	-0.00***	(0.00)
Rauch ex. traded goods	-0.03**	(0.02)	-0.00	(0.02)	-0.06***	(0.02)	-0.03*	(0.02)
Rauch ref. priced goods	-0.04***	(0.01)	-0.02**	(0.01)	-0.03***	(0.01)	-0.00	(0.01)
Change in average tariff	-0.00	(0.00)	-0.00	(0.00)	0.00	(0.00)	0.00	(0.00)
Initial average tariff	-0.00	(0.00)	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)
Change in insurance and freight	-0.03	(0.04)	-0.04	(0.04)	-0.04	(0.07)	-0.01	(0.07)
Initial insurance and freight	-0.07	(0.06)	-0.09	(0.06)	-0.05	(0.05)	-0.02	(0.06)
Initial distance elasticity	0.04***	(0.01)	0.02	(0.01)	0.06***	(0.01)	0.05***	(0.01)
Observations	442		442		442		442	
R-squared	0.11		0.20		0.16		0.23	
Industry FE (1-digit)	No		Yes		No		Yes	
Specification: PPML-zeros								
Elasticity of substitution	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)
Rauch ex. traded goods	-0.05*	(0.03)	-0.02	(0.03)	0.00	(0.01)	-0.01	(0.01)
Rauch ref. priced goods	-0.01	(0.01)	-0.01	(0.01)	0.00	(0.00)	-0.00	(0.01)
Change in average tariffs	-0.00	(0.00)	-0.00	(0.00)	0.00***	(0.00)	0.00	(0.00)
Initial average tariff	0.00	(0.00)	0.00	(0.00)	0.00***	(0.00)	0.00***	(0.00)
Change in insurance and freight	-0.03	(0.03)	-0.04	(0.03)	0.05*	(0.03)	-0.004	(0.03)
Initial insurance and freight	-0.059	(0.09)	-0.122	(0.08)	0.074***	(0.03)	0.02	(0.03)
Initial distance elasticity	-0.03*	(0.02)	-0.03**	(0.02)	-0.01	(0.01)	-0.00	(0.01)
Observations	442		442		442		442	
R-squared	0.04		0.12		0.11		0.25	
Industry FE (1-digit)	No		Yes		No		Yes	

Notes: This table shows the results from regressions of the changes in distance coefficients between each time period and for the OLS and PPML-zeros specifications on industry characteristics. Coefficients are estimated using weighted least squares with the standard errors of estimated distance coefficients as weights. Robust standard errors reported in parentheses. Significance at the one percent level is represented by ***, at the five percent level by **, and at the ten percent level by *.

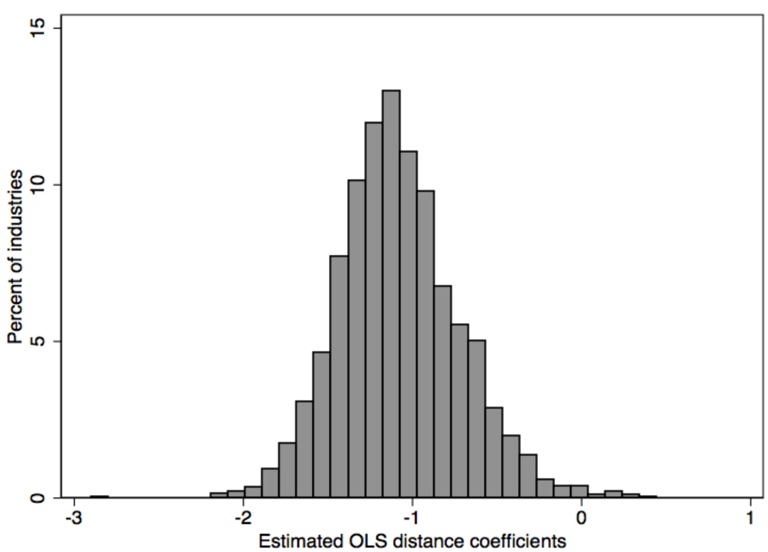
2.11 Figures

Figure 2.1: Average Distance of Trade and Trade Zeros, 1997-2009



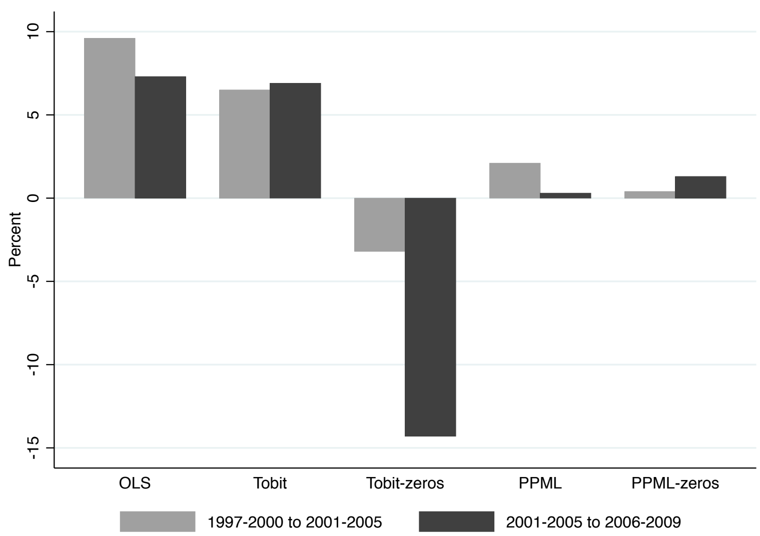
Notes: This figure shows the trade-weighted average distance of trade and trade zeros calculated using equation (2.7).

Figure 2.2: OLS Distance Coefficients: 1997-2000, 2001-2005, and 2006-2009



Notes: This figure shows a histogram of distance coefficients (γ) estimated with OLS using equation (2.8) for the 764 industries and three time periods.

Figure 2.3: Change in Estimated Distance Coefficients



Notes: This figure shows the percent change in the average distance coefficient reported in Table 2.2.

Chapter 3

Decomposing U.S. Water Use Since 1950: Is the U.S. Experience Replicable?[1]

Co-authored with Peter Debaere, Darden School of Business

3.1 Introduction

Water use in the U.S. has followed a remarkable pattern since 1950. After doubling between 1950 and 1980, the total volume of water withdrawn has virtually remained unchanged. Moreover, the newly released U.S. Geological Survey (USGS) data for 2010 reveal that water use has even slightly decreased in the last few years.[2] Against the background of the California water crisis and the mounting global fears of freshwater scarcity, the leveling off and slight decrease in U.S. water use is a fascinating

[1]We received research funding from the Darden Foundation. Chris Hendrickson graciously provided additional access to his water data. Jorge Miranda provided excellent research assistance. We benefited from presenting the paper in the University of Virginia Department of Economics, and the Annual Geophysicist Union in San Francisco, as well as from comments by Paolo D'Odorico, James Harrigan, Arik Levinson, John Mclaren, Ariell Reshef, and Brian Richter. All remaining errors are ours.

[2]Maupin et al. (2014) Estimated Use of Water in the United States in 2010, USGS Circular 1405.

development.[3] The decreasing water use is especially striking since the water data do not track the relatively steady increase in U.S. population, GDP, and per capita GDP over the last sixty years. As a matter of fact, population has more than doubled, GDP has increased more than sixfold and income per capita has tripled since 1950. Today, half of the world's cities lie in water-stressed river basins and one-fifth of the world population suffers from water scarcity.[4] Population growth, economic growth, and rising standards of living and the lifestyle changes they entail are very often expected to increase the demand for water and to further strain the available water resources.[5] To face the challenge of managing water effectively in the 21st century, however, will require a solid understanding of the exact drivers of water use and of the determinants of countries' overall water efficiency.[6] While the focus in the literature is often on technology improvements, in this paper we show how long-term structural changes of the U.S. economy *next to* technological progress have allowed the United States to produce each dollar of its GDP with increasingly less water.[7] The latter has enabled the United States to call to a halt the increase of its overall water use in spite of continued population and GDP growth. In addition, our analysis should prove relevant beyond the U.S. context. In the absence of long-run, high-quality data of water use on a global scale, our analysis reveals, to some extent, whether the U.S. experience is replicable in other parts of the world, and also what such replication

---

[3] Pilita Clark, A world without water, Financial Times, July 14, 2014; The Economist, For Want of Drink, Special Report on Water, May 22nd 2010. National Geographic, Water Crisis News, http://news.nationalgeographic.com/news/archives/water-crisis/, Fang, Kenny, The Global Water Crisis: The Innovations to Watch For, Wall Street Journal, 2007, http://www.wsj.com/articles/SB119042333799235895.

[4] Richter et al. (2013) and UN Word Water Assessment Programme (2009).

[5] UN (2012), Rosegrant et al. (2002), Alcamo et al. (2003), and Vörösmarty et al. (2000) also consider the role of climate change for future water stress.

[6] See Gleick (2003a, 2003b).

[7] See Gleick (2003a, 2003b) for discussion of technology improvements.

would require.[8]

In this paper we explain U.S. water use in economic terms.[9] We start by tying water use to the dramatic long-term structural changes of the U.S. economy. At the heart of this structural change is the rise of the United States as a service economy, and the accelerated demand for services since the mid-1970s to early 1980s. As a matter of fact, at the beginning of the 21st century, about 75 percent of GDP is spent by consumers, local and federal governments and investors on services. This is more than 25 percentage points higher than in 1950; see Figure 3.1.[10] Related to the shift towards services, we also consider the drastic decrease of the U.S. manufacturing sector, the secular decline of agriculture, and the role of globalization that has made the U.S. an increasingly more open economy.[11] Note that the rise of the U.S. service sector and its implications for water use is directly relevant for assessing water use beyond the United States since the *World Development Indicators* reveal that a steadily

[8] A major challenge is the international comparability of water data, see also Gleick (2003a, 2003b). Shiklomanov (1998), for example, compiled aggregate, international water use data from country-specific statistics with varying methodologies. Flörke et al. (2013) reconstructed historical international water use relying on assumptions about human behavior and extrapolating trends in water productivity growth, some of which we investigate.

[9] In this paper, water use refers to surface and ground water withdrawal and not to water consumption. Water withdrawal equals both consumptive use and non-consumptive use that flows back to the environment. Consistent with USGS, we do not include hydropower withdrawals since this water is returned virtually directly to the environment. Our focus on blue water withdrawal is informed by data limitations (no consumptive data are available for 1950-1955 or 2000-2010, nor are disaggregate data available) and by our economic perspective: you pay, if at all, for water withdrawal, not consumptive use. Note, however, that the limited USGS data on water consumption between 1960 and 1995 mimic the longer pattern of water withdrawal.

[10] The reported shares are in current prices. After correcting for changing relative prices, growth in the share of services persists, and so does the slight acceleration since 1980. Final services demand is the relevant measure for our approach, see below. Value added or employment numbers also reveal a shift towards services, see Buera and Kaboski (2012).

[11] The extensive literature on the environmental Kuznets curve documents the inverted-U-shaped relationship between environmental degradation and countries' increasing per capita GDP. Jia et al. (2006) confirms a Kuznets curve for water use among industrialized countries. We explain overall (not per capita) water use in the United States (with positive population growth a Kuznets curve can imply either more or less water use) allowing for scale, changes in composition (of inputs and final demand) as well as changing technology, and globalization.

increasing share of services in world GDP in the last couple decades.[12] The same is true for the U.S.' worsening trade balance since the late 1970s. It raises the question as to whether the United States was able to reduce its water use by importing more water-intensive goods, which would question the ability of the rest of the world to follow the U.S. example.

Our study builds on Leontief's seminal (1970) input-output analysis that has deep roots in economics.[13] We consider *total* water use of all final goods that are produced in the United States and sold to U.S. consumers, investors, governments, and foreigners—these final goods by definition make up all of U.S. GDP. We study both the water that is *directly* withdrawn during the production process of those goods as well as the water that is *indirectly* contained in the intermediates that are employed. While increasingly popular in environmental studies, input-output analyses are not prominent in water studies and have not been systematically applied to explaining the dynamics of water use.[14]

An input-output analysis of direct and indirect water use holds great promise for investigating water use. So far, the standard presentation of water use data by the USGS shows *direct* water use for a few key sectors, with direct water use in the large services sector barely registering at five percent. The singular focus on direct water use limits our economic understanding because water is a very important resource

[12]World Development Indicators, http://databank.worldbank.org/data/. See also Timmer et al. (2014), and Uy,Yi, and Zhang (2013).

[13]See especially Miller and Blair (2009), the standard reference on input-output analysis.

[14]Our study complements innovative research on virtual water by Hoekstra and co-authors. Chapagain and Hoekstra (2008), Hoekstra and Chapagain (2008), Hoekstra and Mekonnen (2012), and Mekonnen and Hoekstra (2011) calculate virtual water use in agriculture in great detail. Without input-output tables, the interaction between agriculture and other sectors of the economy lacks detail. Hoekstra and co-authors study water consumption (not withdrawal), which explains their singular focus on agriculture. Blackhurst et al. (2010) is one of few who use US input-output tables to calculate the direct and indirect water content of the sectors of the U.S. economy for 2002, as do Di Cosmo et al. (2012) who study direct and indirect water content of EU countries for 2005. Like us, Blackhurst et al. (2010) and Di Cosmo et al. (2012) focus on water withdrawal (not consumption).

for intermediate goods that are also used in services. Electricity generation and agriculture, for example, are responsible for over 70 percent of direct water withdrawals. While agriculture and electricity generation together account for a mere three percent of U.S. GDP in recent years, their output is widely used as an intermediate in other sectors. An input-output framework helps tie the majority of direct water use to the rest of the economy, and reveals that the service sector is in fact the largest total (direct and indirect) water user of the economy. Linking intermediate and final goods is all the more important since the open economy that the United States has become remains very dependent on domestic electricity generation and agriculture production.[15] Indeed, electricity generation is to a very large extent for the United States a non-traded good that could only be sourced from abroad at considerable cost, which is why increases in final demand will put additional stress on U.S. water resources. Similarly, the United States has a revealed comparative advantage in agricultural production, and substituting foreign agricultural inputs for domestic ones is prone to drive up production costs significantly.

Our empirical analysis decomposes U.S. water use in terms of its key drivers, and links the improving water productivity of the U.S. economy to the structural change and technological improvements. We modify in two ways the conventional decomposition that is perhaps best exemplified by Levinson (2009).[16] Following Levinson (2009), we relate water use to 1) the changing 'scale' of the economy, 2) the changing water productivity of water use at the sectoral level or the changing 'technique or technology improvements' and 3) the changing composition of the economy. Our focus on final demand and GDP lets us identify changes in scale with GDP growth, as well as break down the composition effect into a) the changing domestic and interna-

[15]If electricity were traded internationally at comparable cost, for example, there would be less of a need to tie final goods demand to electricity use in an analysis to understand U.S. water use.

[16]See also Brock and Taylor (2005).

tional demand for U.S. final products (‘demand composition’), and b) the changing composition of the inputs that are used in the production process (‘input composition’).[17] Our second innovation follows from combining an aggregate and disaggregate analysis. Initially, we consciously work with aggregate sectors such as services, manufacturing, and agriculture that cover all sectors of the economy. Because of this fairly aggregate level, our decomposition displays the *between*-sector shifts of demand and inputs used since 1950. In other words, we can directly link our decomposition to the long-run structural changes of the U.S. economy whose drivers (human capital accumulation, the skill premium, and globalization) are relatively well understood.[18] The latter, at least to some extent, addresses a perceived shortcoming of decomposition analyses as failing to establish a causal link between the phenomenon that is studied (e.g. pollution or in our case water use) and the evolution of GDP, see Levinson (2008) and Levinson and O’Brien (2015).[19] Indeed, few will argue that the larger structural shift that has fueled the emergence of the service sector is driven by water scarcity, or by changing water prices for that matter. Finally, tying total water use to a growing service sector and a declining manufacturing and agricultural sector makes intuitive sense, since services is by far the least water intensive sector.

We find that the changing composition of the U.S. economy is responsible for

---

[17] See Section 3.2 for details. Levinson’s (2009) study of air pollution by manufacturing since 1972 illustrates the standard decomposition well. Our study is different in a number of ways: 1) Unlike pollution, most water is used in a few intermediate goods producers, which warrants our focus on the total water content of final demand. Levinson (except when considering pollution content of international trade) studies the direct pollution content of sectors’ gross output. 2) Focusing on final demand lets us also break down the composition effect into a demand and input component, as well as link the scale effect to changing GDP—after all, final demand across the economy sums to total value added or GDP. Levinson and others investigate changing gross output (not value added) of individual sectors or the entire economy.

[18] See Section 3.3, and Buena and Kaboski (2012).

[19] In Levinson (2015)’s words, extending his critique to Environmental Kuznets curves (EKCs): “But EKCs are simply conditional correlations, without meaningful interpretations other than that pollution does not necessarily increase with economic growth.” Because of this critique, Levinson (2015) turns to examining Environmental Engel Curves.

between 35 and 50 percent of the increased water productivity in the United States. The larger part of the changing composition comes from the shifting final demand by consumers, investors, governments, and foreign customers away from manufacturing and agricultural products towards services. By default, the overall water productivity gain that is not explained by the shifts between the key sectors, between 50 and 65 percent, is booked as technique improvement in the initial analysis. This sizable share is good news if one considers technique improvements opportunities for replication abroad. Transferring technology can be a more actionable way of bringing about less water use, especially when compared to the slow-moving process of structural shifts towards a less water-intensive service economy. In addition, we document that more than 60 percent of these technical water savings are driven by lower water needs per kilowatt-hour in the electricity-generating sector. This finding underscores the role that public infrastructure and regulation can potentially play in constraining water use.

To make sure the fairly aggregate analysis does not bias our findings, we complement our *between*-sector calculations with a more common, granular decomposition that includes the shifts in demand *within* the key sectors. For this exercise we rely on 81 disaggregate sectors based on Blackhurst et al. (2010) who provide very disaggregate sectoral water use data for 2000. Our analysis confirms the shift over time towards less water-intensive products also *within* services and *within* manufacturing. What stands out, however, is that the shift toward less water-intensive products is only slightly more pronounced with disaggregate data than with the aggregate analysis. The aggregate between-sector shifts capture 75 percent of the shift toward less water-intensive (disaggregate) products since 1950, which underscores the explanatory power of the broad structural changes to understand water use in the United States. Note also that the slightly more pronounced shift at the disaggregate level suggests that our estimate of the contribution of technology should be interpreted as an upper bound. At the same time, our more disaggregate analysis documents the uneven pattern of technological progress that took off especially since the 1970s.

Our analysis finally considers the role of globalization and whether the recent stabilization and decrease in overall water use is due to imports of more water-intensive products. We study the hypothetical scenario where the United States would have to produce all the goods it consumes itself (with its own technology). In that case, we find that water use would have peaked in 2005, instead of in 1980. This result is driven by the increase in (U.S.) water content of net imports since the 1980s and especially by the worsening international trade deficit of the United States. It is important to note, however, that the magnitude of these water savings is a relatively limited: 17 percent of overall savings that can be attributed to the changing composition (both demand and input composition) and a mere one percent of overall water use. This finding is important for the international replicability of the U.S. experience. Running a trade deficit or shifting imports towards more water-intensive imports could not be a recipe for increasing worldwide water conservation.

The article is structured as follows. First, we lay out the analytical framework that guides the analysis, and which will be the basis for our description and decomposition of total U.S. water use. In the next section we summarize direct water use data, before we specify sectors' total water use and its link to international trade. The third-to-last section then presents the results of our decomposition exercise. We finally conclude after we have corroborated and interpreted our findings in light of more detailed disaggregation.

3.2 The analytical framework

To gain a deeper understanding of the drivers of water use in the United States, we break down water use by its key sources. We propose a modification of the conventional decomposition into scale, composition, and technology effects that accommodates the specifics of water use and how water is reported in the water use statistics. At the same time, however, our approach should be applicable more broadly: 1) We look at the 'scale' of the U.S. economy and in particular how the changing overall size of the U.S. economy as measured by its GDP affects its water use. 2) We investigate how the 'composition' or how the changing sectoral structure of the economy affects water use. We propose to break up the traditional composition effect into two segments. One part reflects how output of final goods or, alternatively, the demand for final U.S. products changes. Final products are the products that are produced in the United States and bought by its consumers, investors, governments, and also by foreigners for their own use and not to be employed as intermediates in further production. This part of the composition effect reflects most clearly the changing demand that is driven by changing incomes as well as shifting preferences for U.S. products domestically and abroad. The second element of the composition effect is determined by the changing links between the sectors in the economy as mapped by the input-output table. The input-output table lays out how intermediate goods from one sector are used in another. This second composition effect captures changes in how intermediate goods are being combined into final goods. It is not so much associated with changing demand, but rather with the changing production process of final goods as such. 3) We finally investigate how 'technique' or the changing technologies yield water efficiency gains.[20]

[20] This breakdown into scale, composition, and technique follows an emerging convention in environmental economics; see Grossman and Krueger (1993), Copeland and Taylor (2005), and Levinson

Equation (3.1) is a good starting point to introduce the equation that guides our decomposition of total water use in the United States, or of W. We define W as the following multiplication of vectors and a scalar,

$$W = \underline{\mathbf{w}}'\boldsymbol{\theta}Y \tag{3.1}$$

The $n \times 1$ vector $\underline{\mathbf{w}}$ captures for each sector i the total domestic water that is needed to produce one dollar of its final output in the United States. The vector $\underline{\mathbf{w}}$ encompasses both direct and indirect water use. This includes both the water used directly in a sector's output as well as the water contained in the intermediate products that are employed in the sector. Since the sum of domestic and foreign demand for sectors' final products totals a country's GDP, the U.S. total water use, W, is obtained by simply multiplying $\underline{\mathbf{w}}$ by the value of final demand in each sector, which is identical to the product of $\boldsymbol{\theta}$, an $n \times 1$ vector of the shares of sectors' final output/demand in U.S. GDP, and by Y, a scalar that measures U.S. GDP.

We borrow from Leontief's (1970) input-output analysis to calculate the total (direct and indirect) water use vector $\underline{\mathbf{w}}$. Equation (3.2) characterizes the well-known relationship between sectors' gross output (or, the total value of shipments) as the sum of the intermediate and the final products that sectors sell.

$$\mathbf{x} = \mathbf{A}\mathbf{x} + \mathbf{y} \tag{3.2}$$

The $n \times 1$ vector $\mathbf{x}$ contains sectors' gross output. $\mathbf{Ax}$ is the product of the gross output vector and an $n \times n$ matrix $\mathbf{A}$ of input coefficients that characterizes sectors' intermediate goods use. To be precise, the elements a_{ij} of $\mathbf{A}$ indicate how much of sector i's intermediates are used in another sector j (as a fraction of gross output

(2009).

in sector j). The input coefficients are directly derived from the U.S. input-output table. The $n \times 1$ vector $\mathbf{y}$ reports sectors' final output. Sectors produce these final products for domestic and foreign customers. With some matrix manipulation we can rewrite equation (3.2) and directly relate sectors' gross output to their final demand as in equation (3.3).

$$\mathbf{x} = (\mathbf{I} - \mathbf{A})^{-1}\mathbf{y} = \mathbf{L}\mathbf{y} \tag{3.3}$$

To be clear, $(\mathbf{I} - \mathbf{A})^{-1}$ is the famous Leontief matrix $\mathbf{L}$; $\mathbf{I}$ is the $n \times n$ identity matrix. The elements of the Leontief matrix l_{ij} report the total amount of sector i's intermediate output required to generate one dollar of final output in industry j, which includes whatever amount of sector i is used in all other industries whose intermediates are employed in j, as well as the amount of i used in the inputs to those industries.

We obtain total U.S. water use or W in equation (3.4a) when we pre-multiply equation (3.3) with the $n \times 1$ vector $\mathbf{w}$ of sectors' direct water use per unit of gross output in the United States. Using the notation of equation (3.1) we can rewrite expression (3.4a) as equation (3.4b). To be clear, the vector of total water use $\underline{\mathbf{w}}$ equals $\mathbf{w}'\mathbf{L}$.

$$W = \mathbf{w}'\mathbf{x} = \mathbf{w}'\mathbf{L}\mathbf{y} \tag{3.4a}$$

$$W = \mathbf{w}'\mathbf{L}\boldsymbol{\theta}Y \tag{3.4b}$$

The last expression (3.4b) is the equation that we will use in our decomposition to study the water use over time. To isolate the scale effect, we want to study what water use W would be if only the scale Y changed, all else equal. To be clear: as Y changes, we will assume there is no change in the technology $\mathbf{w}$, nor in the distribution of input use as found in $\mathbf{L}$ or in the distribution of final demand that is characterized by $\boldsymbol{\theta}$. Similarly, to isolate the two composition effects and the effect of technology, we will respectively let $\mathbf{L}$, $\boldsymbol{\theta}$ or $\mathbf{w}$ change over time while keeping all other factors the same.

Note that expressions (3.4a) and (3.4b) are particularly well fit for a decomposition of U.S. water use, in light of how the water data are reported and of how water use is distributed. As will be especially clear when we describe U.S. water data in the next section, the heaviest reported water users that are responsible for over 60 percent of overall water use are electricity generation, agriculture, and water utilities. Together these sectors are relatively small, accounting for less than three percent of GDP. Moreover, they are largely providing intermediate inputs to other sectors that comprise 97 percent of GDP. To understand changes in water use, therefore, it is instrumental to figure out what is driving this intermediate good demand, which warrants our focus on changing final demand, the engine for demand for intermediates. To be explicit, understanding that water use increases because water utilities deliver more water, or because more electricity is generated is one thing. It is especially informative to understand why more water is used, and which sectors use the electricity and for what. Therefore, once we focus on sectors' final demand and consider their total water use, we are explicitly accounting for the water contained in their intermediates and we link final and intermediate demand. Note also that our focus on final demand allows us to explicitly link water use to GDP, since the sum of sectoral final demand is total GDP.[21]

[21]If one focuses on the direct water content, one would have to interpret scale as total gross output (which is different from GDP). For reference, Levinson (2009) calculates the direct pollution content of gross output for manufacturing. Only when considering exports and imports does he

A second reason that our decomposition based on (3.4a) and (3.4b) will be informative relates to the non-tradable nature of a major water user such as electricity generation. For a large country such as the United States, most electricity is generated domestically. Because of this, any increase in production of all sectors that use electricity will put additional pressure on U.S. water resources. Our total (direct and indirect) water use measures reflect this. A similar argument could be made with respect to agriculture. Since the U.S. has a comparative advantage in agricultural production, increases in domestic manufacturing through its use of domestic agricultural inputs will pose additional pressure on domestic water resources.[22]

We will proceed with the implementation of the input-output analysis in two steps. First we will conduct the investigation at a relatively aggregate level, allowing us to directly link water use to the broad structural shifts *between* the major sectors in the economy. Next, we will rely on a more granular approach with more disaggregate data that allows us to study water use *within* the major sectors. Comparing between and within results will, on the one hand, reveal any biases involved and, more importantly, underscore the significant contribution of the large structural changes for understanding U.S. water use.

3.3 Data

Figure 3.2 shows how water withdrawals have evolved in the United States between 1950 and 2010 based on data from the USGS. The graph shows the cumulative water use of eight sectors of the economy—the data representation closely relates to how the USGS presents the water data. The USGS provides blue water withdrawal (not

---

focus on the total (direct and indirect) pollution.

[22]See Debaere (2014).

consumption) data since 1950 at five year intervals for the following eight aggregate sectors: industrial (manufacturing), mining, water utilities, electric utilities, livestock, agriculture, commercial use (services), and residential use. From an economic point of view, withdrawal data (rather than consumption data of water that is not returned to the environment) are more relevant, since you tend to pay for withdrawal, not consumption. In addition, water withdrawal data are preferred for our analysis since the consumption data are of lower quality, and only available for a limited period of time (1960-1995).[23] We also do not have for water consumption a comparable source to Blackhurst et al. (2010) that provides very disaggregate water withdrawal data (see discussion in Section refsec:water˙disaggr). Moreover, our choice of water withdrawal as the focus of the analysis relates well to the point made by Gleick (2003) that over-emphasizing water consumption rather than withdrawal sometimes tends to underestimate the real *local* water savings associated with water withdrawal reductions, even when such withdrawal reductions were to have no impact on the water availability downstream.[24] Note also that the USGS data for electricity generation traditionally do not include hydropower, and we follow that convention.

In some years, the USGS does not break down industrial water use into water use by mining and manufacturing, nor does it identify residential demand. We rely on secondary data sources (often the input-output tables) to attribute water use to these sectors.[25] Note that sectoral water use as reported by USGS typically refers to self-supplied water, i.e., water supplied to wells—a notable exception is residential demand. For our decomposition of water use across the major, aggregate sectors of the economy, as well as in Figure 3.2, we have supplemented the provided self-supplied

[23]The consumption data between 1960 and 1995 do show the same pattern as the withdrawal data.

[24]Indeed, local reductions in withdrawal, for example, make it not necessary to tap into additional resources that could have downstream consequences.

[25]See Appendix.

water data of the major sectors with the water that these sectors draw from water utilities.[26] To distribute water from utilities across sectors we rely on estimates by the USGS, complemented by information from the input-output tables that specifies the payments of other sectors to the water utilities. Note that we have to interpolate the data from the input-output tables to match the five-year intervals of the USGS water data.[27]

After we have redistributed the water use by the water utilities to all its customers, the water that remains is that which water utilities use in their process. As should be clear from Figure 3.2, agriculture and electric utilities are by far the heaviest (direct) water users, which over the entire period are responsible for 70 percent or more of all water withdrawn in the United States. Direct water use by services, on the other hand, is relatively minor at less than five percent in 2010.

As noted, to implement our decomposition of water use across the eight aggregate sectors, we match the USGS data with the sectoral information of the input-output data. We take the input-output data from the *Bureau of Economic Analysis*. Given the level of aggregation of our structural analysis, it is relatively straightforward to match gross output data and water use data in order to construct the direct water use vectors $\mathbf{w}$, and to link the gross output $\mathbf{x}$ and final output data $\mathbf{y}$. We also draw on the export and import data as provided in the input-output tables.[28,29]

[26]See Appendix.

[27]The available data from the input-output tables are 1947, 1958, 1963, 1967, 1972, 1977, 1982, 1987, 1992, 1997, 2002, and 2007. We interpolate to match the water data that are available every five years from 1950 to 2010

[28]In some years, the utilities sector is not divided into electric and water utilities, and net exports are not divided into exports and imports. See Appendix for the methods we use to impute these values.

[29]In some sectors and years, the gross output numbers going across the rows of the input-output table do not match the gross output going down the columns. Even though the maximum difference is less than one percent, we need the gross output numbers to match for the decompositions, so that total water use matches the sum of water use across sectors. To maintain the equality of gross output by rows and columns, we adjust final demand.

The input-output data are reported in nominal values. To deflate the nominal values we take two different approaches, with our preference going to the double deflation method because it allows deflation by sector-specific price indexes. In the double deflation method, the output of a particular industry is first deflated by that industry's price index. The value added price index is then derived so that the fundamental identity that the total value of output equals the total value of input holds. See Miller and Blair (2009) for a detailed description of the double deflation procedure. The price indexes for our aggregate categories are chain-type price indexes for gross output by sector from the *Bureau of Economic Analysis*.[30,31] In the second deflation method, we simply deflate all values in the input-output table by the price index for value added, also from the BEA.

3.4 Total water use across sectors and international trade

Figure 3.3 presents U.S. water data in a different way. We depict the total (direct and indirect) water that is contained in the final demand for the eight sectors mentioned above. When compared to the standard categorization of water use in Figure 3.2, Figure 3.3 tells a very different story that more clearly reflects the dramatic changes in the U.S. economy since 1950. As the United States became a service economy, the relative importance of its manufacturing sector diminished and agriculture expe-

---

[30] Prior to 1977, real gross output is not available to construct the price indexes, so we use appropriately scaled price indexes for corresponding sectors from the U.S. Bureau of Labor Statistics.

[31] We use the BEA price indexes for all sectors except the electricity-generating sector, for which we use the real total price of electricity (supplied to residential and industrial consumers) from the Energy Information Administration. This price index better captures price movement in the electricity-generating sector than the price index for utilities, which is the most closely related price index provided by the BEA.

rienced a secular decline as a fraction of GDP. As noted, since no sector in the U.S. economy can produce without electricity, the heaviest direct user of water, the total (direct and indirect) water content calculations of final demand show how much the final demand for products in various sectors puts pressure on the U.S. water resources. The same is true to some extent for the intermediate use of agricultural products. Since the United States has a comparative advantage in agriculture, increased input demand due to final demand for manufacturing goods will also strain U.S. water resources. As Figure 3.3 illustrates, manufacturing and the service sector's total (direct and indirect) water use comprise 60 percent or more of water use, and the role of the service sector as far as total water demand goes is ever increasing.[32] As a matter of fact, in 1950 services total water use was a mere 18 percent of all water use in the U.S, and grew by just three percent over the next three decades. After 1980, water use accelerated markedly, and by 2010 it stood at 35 percent of total water use. Manufacturing, on the other hand, initially accounted for 54 percent of U.S. water use, a number that dropped to 39 percent in 1980 and finally 23 percent in 2010.

To fully assess U.S. water, we also want to study how globalization has altered U.S. water use. We ask the question whether the United States would use more or less water in the hypothetical situation that it had to produce the goods it consumes (as in a closed economy).[33] To determine whether U.S. water saving is due to its exchange with the rest of the world economy, we calculate the total water content of net imports in Figure 3.4 *using U.S. technology*.[34] The exercise is particularly relevant if one is interested in assessing to what extent the U.S. experience can be replicated

[32]The line for electricity utilities reflects the direct electricity used by residents and the water use it implies.

[33]Needless to say, this exercise is just a thought experiment—we assume that prices, production and consumption patterns are not altered by becoming a closed economy.

[34]There is a long tradition in the international trade literature to calculate the total factor content of net trade, see Baldwin (2008), as opposed to just comparing the direct factor content of exports and imports which ignores the factors used in the production of the intermediate goods used.

abroad.

Before assessing our findings, we should clarify two complications that come with calculating the water content of net imports. First, in order to produce goods, the U.S. economy uses imported intermediate inputs. There is no perfect way to account for the imported intermediates, which are typically not fully specified in an input-output table. The most common way to "scrub" the imports from intermediate inputs is to use a proportionality assumption. In this case one assumes that each sector uses imported intermediates to the same extent (i.e., the use of imported intermediates does not vary across the various sectors that one sector produces for). To that effect we multiply the input-output coefficients a_{ij} in the input coefficient matrix $\mathbf{A}$ by the adjustment factor for sector i. We follow Levinson (2009) and Miller and Blair (2009) and use for each sector the ratio of $imports/(domestic\ production + imports - exports)$ as the adjustment factor.[35] Note that because we are interested in evaluating what it would cost the United States in terms of water to produce the imported products itself, we evaluate imports by U.S. technology.[36]

We also want to address a second concern that relates to the level of aggregation as we assess the water content of net imports—aggregation will also play a role in how we interpret our decomposition results below. We can easily calculate the water content of net imports with our relatively aggregate data as $W_T = \underline{\mathbf{w}}'\mathbf{T} = \mathbf{w}'\mathbf{L}\mathbf{T}$, where $\mathbf{T}$ is the vector of imports minus exports and $\underline{\mathbf{w}}$ and $\mathbf{w}$ respectively the vector of total and direct water use for our major sectors in the United States. One might be concerned, however, about systematic differences in the mix of water-intensive

[35]This adjustment factor is implied by the proportionality assumption, as described in Antràs et al. (2012), who use the assumption to construct an open-economy adjustment to their measure of the upstreamness of production.

[36]We are able to relax the proportionality assumption with import shares computed from the Asian Input-Output tables, which distinguish U.S. imports by use. We apply the distribution of import shares across use within an industry for the year 2000 to the import shares computed using the BEA input-output tables. The decomposition results are unchanged.

products *within* the major sectors that we consider between imports, exports, and what the U.S. produces. In such a case, applying the aggregate $\underline{\mathbf{w}}$ and in particular the $\mathbf{w}$ measures that are based on the water use of the mix of goods of U.S. production sectors to the aggregate trade data $\mathbf{T}$ should bias the water content calculations.[37] We therefore propose to adjust our water content of net trade measures W_T by using more disaggregate data, while at the same time working around the constraint that we only have disaggregate water use data for one year—Blackhurst et al. (2010) disaggregate the USGS water data for 2000 and assign them to the 428 NAICS categories of the input-output table.

Here is how we proceed. For each year t in our sample, 1) we calculate the water content of exports and imports with our aggregate data that are readily available, $W_t^E = \mathbf{w}_t'\mathbf{L}_t\mathbf{E}_t$ and $W_t^{IM} = \mathbf{w}_t'\mathbf{L}_t\mathbf{IM}_t$, where $\mathbf{E}_t$ and $\mathbf{IM}_t$ are respectively the export and import vector; 2) we construct the water content of exports and imports for that year with the aggregate water use vector for the year 2000, $\mathbf{w}_{(2000)}$, and obtain $W_{t(2000)}^E = \mathbf{w}_{(2000)}'\mathbf{L}_t\mathbf{E}_t$ and $W_{t(2000)}^{IM} = \mathbf{w}_{(2000)}'\mathbf{L}_t\mathbf{IM}_t$ and, 3) we calculate the water content of the 81 disaggregate export and import sectors using the vectors $\mathbf{E}_{dt}$ and $\mathbf{IM}_{dt}$, the disaggregate water use vector for 2000, $\mathbf{w}_{d(2000)}$, and the disaggregate Leontief matrix, $\mathbf{L}_d$, to obtain $W_{dt(2000)}^E = \mathbf{w}_{d(2000)}'\mathbf{L}_{dt}\mathbf{E}_{dt}$ and $W_{dt(2000)}^{IM} = \mathbf{w}_{d(2000)}'\mathbf{L}_{dt}\mathbf{IM}_{dt}$. Because of differences in product mix between exports, imports, and production within more aggregate sectors, it is possible that the total water content of, say, exports, $W_{t(2000)}^E$ differs from the disaggregate calculation $W_{dt(2000)}^E$. If we find that $W_{dt(2000)}^E$ differs from $W_{t(2000)}^E$ by a factor α_t in a particular year, we propose to pre-multiply our aggregate W_t^E measures by α_t to correct for a potential bias. Needless to say, α_t will vary over time. We do the same for the factor content of imports. For reference,

[37] This concern does not arise when dealing with domestic final demand, as by construction $\underline{\mathbf{w}}'\boldsymbol{\theta}Y = \underline{\mathbf{w}}_d'\boldsymbol{\theta}_dY$.

we find that both the product mix of exports and imports tends to be somewhat less water-intensive than that of production, and on balance, especially for the later years, the water content of net trade that takes into account the variation in product mix is less water-intensive than the more aggregate net water content of trade, see Appendix.

Note that to make the disaggregation work, we have to address the changing classifications of the input-output tables, which is a challenge. In particular, before 1997 the sectors in the input-output tables were classified using SIC codes within 81 broad categories that we can easily follow through time. In later years, however, NAICS classification codes are used. We have to reconcile the 428 input-output sectors in NAICS codes with the 81 categories of the input-output table that we were following before. We build a concordance between 1997 and later years that largely follows Cicas et al. (2006).[38]

The lowest line in Figure 3.4 that captures the water content of U.S. net imports documents that there has been a significant change in the U.S. water exchange with the rest of the world over time. From 1950 to the 1980s the water content of net imports was negative as the water content of its exports was higher than the water it would take the United States to produce the imports itself. Since the 1990s, the total water use of net imports has turned positive, however. By 2010 the United States imported on net 3.8 billion gallons of water per day through its trade. This evidence suggests that some of the water savings achieved in the United States are due to its changing exchange with the rest of the world. Of particular interest here is the top curve in Figure 3.4. We have added the water contained in net imports (bottom line) to U.S. domestic water use (middle line), which is tantamount to assuming hypothetically that the United States would be producing all the goods it consumes

[38] See Appendix for further details about the concordance.

(just like a closed economy).[39] As Figure 3.4 makes clear, under such hypothetical scenario the peak of U.S. water use would be in 2005 instead of 1980, suggesting indeed that overall water use has not leveled off but increased virtually continuously. While Figure 3.4 is qualitatively of interest, it should be noted that on balance the water content of net imports is only one percent of total water use, and as we will show unlikely an impediment for the replicability of water saving abroad.

3.5 Decomposing U.S. water use

Figure 3.5 presents our key findings for the decomposition of U.S. water use since 1950 that uses double deflation with industry-specific price deflators (see Table 3.1 for the quantified effects).[40] The decomposition compares U.S. water use under four hypothetical scenarios relative to what it was in 1950. Before emphasizing some of our key findings, let's make sure we understand the meaning of the various curves in Figure 3.5. For ease of interpretation we have scaled all curves by total 1950 water use. or by 180 billion gallons a day. In this way, it is fairly straightforward to assess changes in water use.

The lowest curve shows the increase of actual total water use as observed in the USGS data since 1950. Following equation (3.4b), observed water use, W_t, equals at every moment in time $\mathbf{w}_t'\mathbf{L}_t\boldsymbol{\theta}_t Y_t$, which involves changes in all of its components (sectoral water productivity, the changing input-output matrix, the share of sectoral

[39] It should be emphasized that this is nothing but a hypothetical scenario since prices are assumed not to change.

[40] We have also performed the decomposition uniformly applying the GDP deflator across all sectors, see Figure A.1 and Table A.2. Such an analysis increases the contribution of the changing composition to the water efficiency gains relative to the one we obtain with double deflation, and decreases the contribution of technological progress. Closer analysis reveals that deflating especially electricity generation with the GDP deflator fails to correct for the particular pattern of electricity pricing, understating technological progress.

final demand in GDP, and GDP). The top line singles out what happens to water use (relative to water use in 1950) once the scale of the economy changes, all else equal. In other words, it depicts what water use would have been if the economy had grown at its actual rate, while the water productivity (technology), the input-output relationships, and the distribution of final demand did not change from their 1950 levels. We calculate the water use that is implied by the changing scale of the economy as $\mathbf{w}'_{1950}\mathbf{L}_{1950}\boldsymbol{\theta}_1 950Y_t$ and divide it by 1950 water use. For ease of interpretation we introduce the subsequent changes (composition and technology) in a cumulative fashion. The second and third curves from the top are labeled respectively scale plus demand composition, and scale plus demand and input composition. They are calculated as $\mathbf{w}'_{1950}\mathbf{L}_t\boldsymbol{\theta}_{1950}Y_t$ and $\mathbf{w}'_{1950}\mathbf{L}_t\boldsymbol{\theta}_t Y_t$ and compared to 1950 water use. The second and third curves allow the distribution of final demand, $\boldsymbol{\theta}_t$, and the input-output structure of production, $\mathbf{L}_t$, to change over time in addition to the change in GDP.

With these definitions in mind, it is relatively straightforward to interpret the curves and the vertical differences between them.

- Since all curves are normalized by 1950 water use, they indicate how strong the changing scale or the changing scale *plus* (input and demand) composition would have pushed water use up compared to 1950, as well as how much actual water use did rise since 1950.

- Of even greater interest is the vertical difference between the lowest curve (the actual water use through time) and the highest curve (the scale effect) at every moment in time, which measures all realized water savings relative to 1950. As a matter of fact, at every moment in time the ratio of scale to actual water use is a measure of how water productivity has evolved since 1950, and how much less water it takes to produce one dollar's worth of GDP since 1950. The vertical distance between the first and second curve, the second and third curve, etc.—all the way down to the lowest (actual water use) curve—shows how the overall water productivity gains for the United States as a whole can be broken down.

- The difference between the first and the second curve (relative to the difference between the first and the lowest curve) reveals how important the changing composition due to the changing structure of final demand is for the improving water productivity in the United States.

- The difference between the second and the third curve informs us about water-productivity gains due to the evolving composition associated with the changing input-output matrix. This difference gets at the varying ways in which various intermediates are being combined to produce a final good. This reflects a changing production process.

- The difference between the third curve and the lowest curve that marks the actual water use attributes all of the remaining water productivity gains to improvements at the sectoral level that are summarized under the label technique or technology.

When looking more closely at the data, a few observations stand out. First and foremost, the United States has experienced substantial gains in overall water productivity between 1950 and 2010. While water use in 2010 was 1.95 times what it was in 1950, the upper, scale curve indicates that water use would have been 6.71 times the 1950 level if technology as well as the structure of demand and of input use had not changed since then. What can account for this overall increase in water productivity of almost 250 percent ($6.71/1.95 = 3.44$) since 1950? The most drastic water savings have occurred since the mid 1970s/early 1980s around which time water use was increasingly disconnected from GDP growth. Before water use was increasing virtually in step with the growing economy: real GDP grew 2.8 times between 1950 and 1980, and actual water use rose 2.4 times. Since 1980 the picture has been very different. Actual water use has decreased slightly (by 19 percent) since 1980, whereas GDP for 2010 has increased 2.4 times its 1980 level. In sum, while overall water productivity has increased a mere 1.2 times ($1.2 = 2.8/2.4$) between 1950 and 1980, it has grown 2.7 times ($2.7 = 2.4/0.89$) since 1980.

Tracking for 2010 the vertical difference between the top curve and the second curve from the top, we notice that the changing composition associated with the changing final demand can account for 35 percent of the water productivity gains since 1950. The difference between the second and the third curve from the top assigns another 15 percent of the productivity gains to the changing input-output structure. We thus find a total composition effect of 50 percent, which underscores how crucial the structural shifts in the economy have been for slowing down water use in the United States.

In our analysis we study the impact of the structural changes of the U.S. economy on water use. We can draw on an important literature to explain what is behind the dramatic increase in the size of the service sector and the associated decrease of manufacturing and agriculture. A recent contribution, Buera and Kaboski (2012) provides a succinct summary of the key papers in the literature, from the early observers of the growth of the employment share of the service sector to the more recent theoretical contributions that also address increased final demand in services. A key factor in the emergence of the service sector is the very distinct human capital accumulation in the United States that is associated with the higher returns of skill acquisition in spite of an increase in supply of high-skilled labor.[41] According to Buera and Kaboski, increased specialization in skills gives way to an increasingly important role for the market to provide services that used to be provided in-house or in-family. Buera and Kaboski's data show that for much of U.S. history, the size of the service sector stayed relatively stable. By 1950 it began to increase, and since 1980 we even saw an acceleration of that share. Being able to link our decomposition to the structural change literature is of particular importance since it brings in an element of causation to one of the key drivers of water use: few will argue that the larger structural shift that has driven the emergence of the service sector is driven by water scarcity.[42]

The increasing relative size of the service sector, however, implies a decline in the relative size of the other sectors. Over the period that we study the decline of manufacturing is most pronounced since the 1980s. There is also a significant literature on the impact of globalization on the size of the manufacturing sector. A most recent article by Autor et al. (2014) in particular shows how increased competition due to the emerging Chinese economy as a major exporter of manufacturing products has hastened the decline in U.S. manufacturing.[43] Note that the structural change that we study makes sense in the context of our attempt to explain the stabilization of water use in the United States. Indeed, the rising service sector is one of the least water-intensive sectors, whereas the declining sectors, manufacturing and especially agriculture, are the more water-intensive ones. For reference, Table 3.1 reveals significant differences in water productivity at the sectoral level by multiple measures. For comparison, we also include total water use per final demand in a sector. Note that

---

[41]The 20 percent increase in the service sector as a share of value added is entirely explained by the rise of high-skill services, see Buera and Kaboski (2012).

[42]This, at least for one of the key drivers that mitigated the fast increase in water use, addresses a criticism that is sometimes leveled at such decompositions, see Levinson (2008) and Levinson and O'Brien (2015).

[43]For related literature see Autor et al. (2014).

while the very stark differences in water productivity remain, counting total water use relative to final demand does reduce the extent of the sectoral water productivity differences by a factor of ten.

The remaining difference between the third curve and the lowest one in Figure 3.5 indicates that about 50 percent of the water productivity gains can be attributed to productivity gains in the water use within the various sectors. For those eager to replicate the successful reduction in U.S. water use, this is good news, as technological improvements (especially when compared to the slow-moving shift in a country's sectoral structure) are more likely to be influenced by policy and are also potentially faster to implement. We will come back to this finding as we compare our results with a more disaggregate analysis. Before doing so, however, we intend to refine the technology result, assess the impact of globalization, and vary the ranking of the decomposition.

3.5.1 A closer look at technology

In this section, we take a closer look at the technology improvements. We impute the actual water savings (improvements in water per kilowatt-hour (kWh)) within the thermic electricity-generating sector, which is a good proxy for technical/technological advancements.[44] The imputation yields the fourth curve (the one right above the actual water use curve). What is striking is that the fourth curve lies not too far above the actual water use curve. This underscores the key nexus between energy generation and water use, and between technological improvements and water-

[44] We formally do this by calculating (and drawing) $W_t = \mathbf{w}^{*'} \mathbf{L}_t \boldsymbol{\theta}_t Y_t$, where the sectoral water productivity measures in $\mathbf{w}^*$ are identical to those of $\mathbf{w}_{1950}$, except for water use/gross output in electricity generation. Water productivity in electricity generation is allowed to change with the improvements in water/kWh in the data (U.S. Energy Information Administration), while kWh/gross output is held constant at its 1950 level. In particular $\mathbf{w}^*{}_t$ for electricity generation $= (w/kWh)_t * (kWh/gross\ output)_t$.

productivity gains. There are non-negligible returns to water saving technology in the electricity-generating sector. As a matter of fact, the technological improvements by the electricity-generating sector are responsible for the vast majority (64 percent) of the water-productivity gains due to technique or technology.

While we do not formally investigate what is driving the move towards more water saving technology in electricity generation, a few facts have to be brought in. As Kenny et al. (2009) points out the Clean Water Act that amended the 1972 Federal Water Pollution Control Act most likely played a key role. The Clean Water Act regulated not only the technology of cooling water intake that should minimize the environmental effect, but also the cooling system thermal discharges. Increasingly since the 1970s power plants reduced their water use significantly by recycling water, or by using air-cooled systems instead of once-through cooling systems. This phenomenon has had a significant hand in accounting for the improvement in the water productivity. Corroborating this analysis is the fact that water use/gross output for electricity generation (relative to its 1980s value) shows the strongest decline of all sectors that we consider, see Figure 3.10. We found that finding the proper deflator (we use the nominal electricity price for residential and industrial use) is also important.[45]

3.5.2 International trade and the decomposition

In this section, we get back to the impact of international trade. As noted, the changing composition of the U.S. economy plays an important role in accounting for its water savings. The question is whether and to what extent this aspect of

[45]Using the price index for gross output of the utilities sector overstates the technological improvements after 1980 and understates the improvements before 1980. Using the GDP deflator does not capture the industry-specific price movements and understates the role of technological progress, see Figure A.1.

water saving is to be attributed to the import of more water-intensive goods versus the export of less water-intensive products. Figure 3.6 is similar to Figure 3.5 (see Table 3.3 for the quantified effects). For simplicity, we have lumped both composition effects together. In addition, we include two new curves. One curve is similar to the top curve of Figure 3.4 and adds the water content of net imports to domestic water use (while not allowing technique to change) in order to assess how much water saving that is associated with the changing composition of the U.S. economy comes from net imports. The other curve, then, corrects for the size of the U.S. current account. Notably, there has been a relatively dramatic change in the external position of the United States. In 1950 U.S. exports were larger than imports with a trade surplus was about 3.2 percent of GDP. In 2010, on the other hand, U.S. imports far outstripped exports, and the trade deficit was 3.3 percent of GDP.

As one can see from Figure 3.6, trade contributes a relatively small portion of the overall composition effect. If the United States had to produce all of its imports (the no-trade scenario) it would be saving 17 percent less water. If the United States were forced, on the other hand to run a level trade balance compared to 1950, it would reduce its savings by 16 percent. In other words, the trade deficit accounts for over 94 percent (16/17) of the water saving, and less than 6 percent (1/17) is to be attributed to a shift towards more water-intensive products. The relatively moderate role of imports in water savings is good news for the international replicability of U.S. water savings: running a current account deficit or shifting imports towards more water-intensive products is not a recipe for water saving that can be implemented by all countries of the world.

3.5.3 Varying the order of the decomposition

An attractive feature of our decomposition is that it explains the total change in water use and water productivity and breaks it down into the salient components, considering the various drivers of water use in the entire economy (scale, composition, and technology) cumulatively. A disadvantage is that the decomposition ignores any interactions between the various components, which is why we investigate the robustness of the decomposition with variations in the sequencing of the changes.[46] While we did investigate all possible orderings, only a few of them are meaningful. The decomposition that we presented so far is the most intuitive one that is directly in line with Leontief's input-output analysis in which change is driven by the changing final demand. There is one change of sequencing that we did want to report. We reversed the order of the technological component versus the input and demand composition component. For the entire period we find a stronger role for technology (65 percent instead of 50 percent) and a smaller role for the sectoral composition (35 percent instead of 50 percent). Reporting this decomposition lets us describe the range of the technology and composition effects. Note that other orderings are really not meaningful. Since our focus is on explaining changing water productivity (GDP/actual water use), the scale effect has to remain the first change to consider in the decomposition, and the actual water use by default the last one. We break down the overall composition into a demand and an input component. It is hard to rationalize inserting the technology component in between both composition effects.[47]

[46]Note that many decomposition studies cannot investigate the robustness as they do not have proper measures for all components, often attributing the residual of scale and composition to technology.

[47]If one would, one would not find any meaningful difference in the overall decomposition.

3.5.4 Aggregating up

In this section and in Figure 3.7, we aggregate all the USGS sectors that we have been using in the decomposition up to two, services and the rest of the economy. The objective is to illustrate the robustness of the key role that services plays in the increased water productivity. We want to convince ourselves that what we have done so far, classifying the (in terms of water use) quite sizeable residential water and electricity demand as separate entities, did not distort the role of services in the decomposition of U.S. water saving. Lumping both water and electricity together with other goods is in line with much of the literature that classifies water and electricity use as a good, not a service, see Reshef (2013). Figure 3.7 is quite similar to Figure 3.5. The latter should not surprise. For one, residential water and electricity use are very small in terms of GDP and have not been growing at the rate that services has. In addition, both residential water and electricity use are far more water intensive than services. As such, they add to the water intensity of the slower-growing rest of the economy, which can only emphasize the water saving through the emergence and fast expansion of less water-intensive services.

3.6 Disaggregation

So far, we have shown how switches between the major sectors of the economy in terms of final demand and inputs used have played a non-negligible role next to technological progress in driving the pattern of water use in the United States. Using the more disaggregate water data by Blackhurst et al. (2010) we want to double check this observation and investigate whether and to what extent our finding is a consequence of ignoring action within our larger sectors. The between-sector analysis attributes improvements that cannot be explained by composition changes in final demand or input use between our major sectors to changes in technique/technology. Moreover, it treats the major sectors as relatively homogenous units. If it were the case that there were within agriculture, within manufacturing, or within services a shift towards the production of more water-intensive goods, our attribution of 50 (65) percent of the water productivity gains to technology would underestimate the total contribution of technology. The technological improvements within the respective sectors would simply have been offset by the consumption (and production) of more water-intensive final goods in those sectors. Alternatively, with a shift towards more consumption and production of less water-intensive goods within sectors, we would have overstated the technique/technology contribution.

In Figure 3.8a, we first confirm the important role of an overall composition shift over time towards less water-intensive goods also at the more disaggregate level—to simplify we lump together demand and input composition. Note that since disaggregate sectoral water use data are only for 2000, $\mathbf{w}_{2000}$, our analysis is more constrained than the between-sector analysis that drew on aggregate sectoral water use data for all the years of the period that we study. To tease out the composition effect, we compare total water use in 2000, which equals $\mathbf{w}'_{d2000}\mathbf{L}_{d2000}\boldsymbol{\theta}_{d2000}Y_{2000}$, with what water use would have been if the United States had to generate its 2000 GDP with its 2000 water technology, but with the composition (final demand and input combinations) of the other periods, or with $\mathbf{w}'_{d2000}\mathbf{L}_{dt}\boldsymbol{\theta}_{dt}Y_{2000}$. We find indeed that the composition of the earlier years would have given way to significantly more water use—on the order of 1.8 times as much water as in 2000 if the 1950 composition were used. Because of this shift towards less water-intensive goods over time within our broader sectors, the more aggregate between-analysis that we presented before by construction underestimates the extent of the composition effect, and hence overestimates the role of technology. What is striking, however, is that the big structural between-sector shifts (in particular the emergence of the service sector) that we have focused on and that are hard to be rationalize in terms of water scarcity or rising water prices in recent years account for a very important fraction of the composition effect. Calculating $\mathbf{w}'_{2000}\mathbf{L}_t\boldsymbol{\theta}_t Y_{2000}$ with aggregate data for 1950, we capture 75 percent of the composition effect as calculated by disaggregate data $\mathbf{w}'_{d2000}\mathbf{L}_{dt}\boldsymbol{\theta}_{dt} Y_{2000}$: the 1950 hypothetical water use is 1.4 times as high as that of 2000.[48]

To tease out the impact of technology, in Figure 3.8b we compare with disaggregate data, total actual water use in United States or, $\mathbf{w}'_{dt}\mathbf{L}_{dt}\boldsymbol{\theta}_{dt} Y_t$, with what it would have been had 2000 technology been used, or with $\mathbf{w}'_{d2000}\mathbf{L}_{dt}\boldsymbol{\theta}_{dt} Y_t$. We also draw the same hypothetical water use with aggregate data, or $\mathbf{w}'_{2000}\mathbf{L}_t\boldsymbol{\theta}_t Y_t$. As one can see, there is less of an improvement in technology with disaggregate data than what our aggregate analysis suggested. As noticed before, technological improvement is unevenly distributed over the time frame, and improvement picks up after 1975. Since we have only disaggregate water use data for one year to work with, we are constrained in how we can measure technological progress for individual sectors. We, for example, cannot compare $\mathbf{w}'_{dt}\mathbf{L}_{dt}\boldsymbol{\theta}_{idt} Y_{it}$ with $\mathbf{w}'_{d2000}\mathbf{L}_{dt}\boldsymbol{\theta}_{idt} Y_{it}$, for agriculture, manufacturing and services, which would let us directly tease out the role of technology.[49] In Figure 3.10 we plot the aggregate $\mathbf{w}$'s for the various sectors—to make the ratios comparable we divide each water intensity by its 1980 value. What stands out in the aggregate data is the change in the water to gross output ratio comes from

[48]For reference, in Figure 3.8a, we only let the composition change over time. If one wanted to get a sense of what the comparable impact of technology as the only changing factor would be, one could compare water use of 2000 ($\mathbf{w}'_{2000}\mathbf{L}_{2000}\boldsymbol{\theta}_{2000} Y_{2000}$) with what water use would have been if one were to produce the 2000 GDP with the composition of 2000, yet with technology that evolves over time (i.e., $\mathbf{w}'_t\mathbf{L}_{2000}\boldsymbol{\theta}_{2000} Y_{2000}$). One would find that water use in 1950 would have been 1.9 times that of 2000. Note that this comparison should be with aggregate data since we have no disaggregate $\mathbf{w}$ vector that changes over time.

[49]To be explicit, $\mathbf{w}'_{dt}\mathbf{L}_{dt}\boldsymbol{\theta}_{idt} Y_{it}$ does not correspond to the water use within agriculture, manufacturing or services as reported by the USGS, as it refers to the water contained in final goods demand/production, not water contained in gross output.

agriculture and especially from the electricity-generating sector. Since 1960 there is a continuous improvement in electricity generation, whereas for agriculture we have to wait till 1980. Combined with the fact that there is initially a slight shift towards more water-intensive products within agriculture, we know that the technological improvement since 1980 may have been stronger in agriculture than what we observe from Figure 3.10.

In Figure 3.9 we look inside manufacturing and services and compare for those sectors the water that is contained in final demand in 2000, $\mathbf{w}'_{d2000}\mathbf{L}_{d2000}\boldsymbol{\theta}_{id2000}Y_{i2000}$, with what water use would have been if one were to produce the final demand of 2000 in each of the sectors with the 2000 technology but with the changing internal composition, or with $\mathbf{w}'_{d2000}\mathbf{L}_{dt}\boldsymbol{\theta}_{idt}Y_{i2000}$—$i$ stands for manufacturing or services. There is a clear shift toward less water-intensive sectors within manufacturing and services over time.

3.7 Conclusion

In recent years water stress and water scarcity around the globe have received increasingly more attention in the public domain. To manage water stress adequately in the United States and in the global economy, it is essential that we understand the long-term drivers behind water use. In this paper we have decomposed the long-term, blue water use for the United States that in spite of significant GDP growth has stabilized and even decreased since 1980. To shed light on the significant overall water productivity growth that made this stabilization in water use possible, we have explicitly linked the main sectors of the U.S. economy to water through their direct and indirect water use. The literature tends to favor technological/behavioral explanations of water productivity improvements; we have documented that the changing structure of final demand and production of the U.S. economy (the evolving service economy, the decline in manufacturing, and the secular decline of agriculture) has played a critical role. It is not the case that water savings are solely driven by improvements in technology. Thirty-five to 50 percent comes from the changing composition of the U.S. economy. Moreover, as far as technological improvements go, the lion's share comes from efficiency gains in the electricity-generating sector.

Our conclusions for the United States are directly relevant for the global economy, especially since long-term, detailed and internationally comparable water data are not available on a global scale. We do not find that the majority of the productivity gains in the United States are at the expense of the rest of the world. The U.S. current account deficit and imports of water-intensive goods have an only limited impact on the overall outcome. More importantly, our finding that structural change that moves an economy towards services slows down water withdrawals is relevant for a world economy that is increasingly oriented towards services. Our analysis also suggests that water productivity gains can emanate from efficiency gains in the electricity-generating sector.

Read against the background of increased water stress in the United States our analysis at the same time raises questions for further research. While our analysis of water demand shows growing restraint and even a slight decrease in demand, it does not suggest whether such restraint is sufficient. In other words, the demand-side analysis presented here needs to be complemented with an analysis of water supply, to see whether, in spite of the slight decrease in water withdrawal, a sustainable level of water use has been reached. Merging water supply with demand might moreover shed light on some of the underlying reasons water productivity gains are achieved. In particular, our analysis takes within-sector shifts towards less water-intensive products as well as technological improvements that increase water productivity as given. It is an open, empirical question whether a causal connection can be established between the extent of the water scarcity in a region, or some other environmental stress in a particular time period, and technological progress. Whether and how scarcity and drought trigger innovation and efficiency gains is an important question that that needs more empirical research with micro-level data.

Finally, our findings should also matter for ongoing discussions about societal water redistribution in the wake of water crises such as the decade-long Big Dry in Australia and the current water stress in California. In those discussions, mechanisms—such as water markets—tend to be favored that channel water from less to more productive water users. Our analysis confirms that water is on average more productively used in services and manufacturing compared to agriculture. However, what our analysis also emphasizes is the need to consider the indirect water use of the key sectors caught in the debate (services, manufacturing and agriculture) especially since much of the intermediate water use (such as the water linked to electricity generation) is non-tradable. Including the indirect water use of sectors indicates that the average water productivity differences across sectors are still significant but not as stark as initially assumed.

3.8 Tables

Table 3.1: Decomposition with Double Deflation

Increase in Water Use Since 1950, Allowing the following to change over time:	
Scale	6.71
Scale and Final Demand Composition	5.05
Scale, Final Demand and Input Composition	4.32
Scale, Final Demand and Input Composition, and Water Efficiency of Power Generation Sector	2.80
Scale, Final Demand and Input Composition, Water Efficiency of Power Generation Sector, and Technique (Actual Water Use)	1.95
Fraction of Water Productivity Improvement Explained by:	
Final Demand Composition	0.35
Input Composition	0.15
Water Efficiency of Power Generation Sector	0.32
Technique	0.18

Notes: The table reports the size of the effects and the fraction of the overall water productivity improvement explained by each effect for the decomposition shown in Figure 3.5.

Table 3.2: Measures of Sectoral Water Efficiency Relative to Services (2005)

	Direct Water/ Gross Output	Direct Water/ Value Added	Total Water/Final Demand
Agriculture	712.9	1,106.4	161.3
Manufacturing	3.5	5.8	4.5
Services	1.0	1.0	1.0

Notes: This table shows measures of water use per dollar of output measure relative to the services sector.

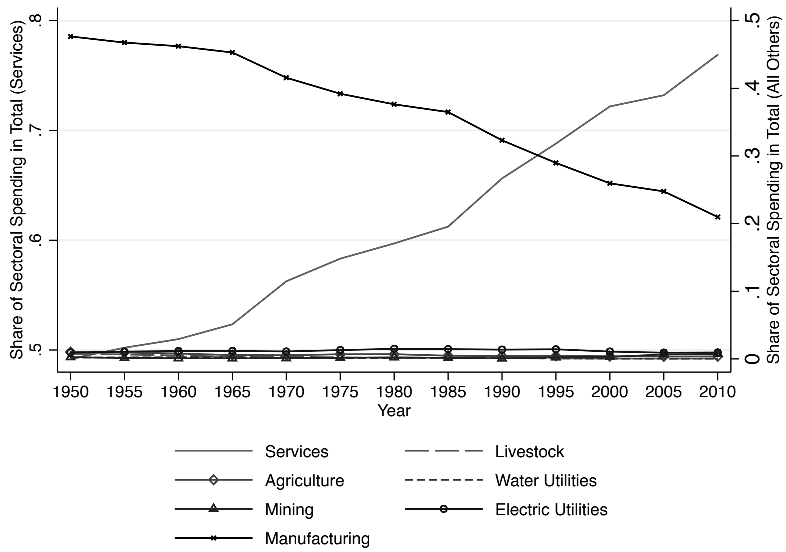
Table 3.3: Decomposition: Hypothetical Closed Economy

Increase in Water Use Since 1950, Allowing the following to change over time:	
Scale	6.71
Scale, Final Demand and Input Composition; No Trade	4.74
Scale, Final Demand and Input Composition; Current Account = 0	4.71
Scale, Final Demand and Input Composition	4.32
Scale, Final Demand and Input Composition, and Technique (Actual Water Use)	1.95
Fraction of Composition Effect Explained by:	
Trade	0.17
Current Account	0.16

Notes: The table reports the size of the effects and the fraction of the final demand and input composition effect explained by each trade scenario for the decomposition shown in Figure 3.6.

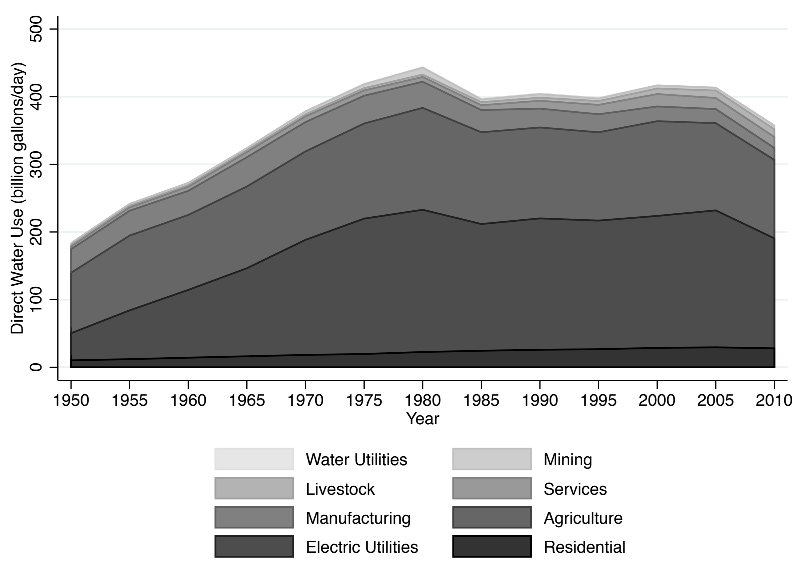
3.9 Figures

Figure 3.1: The Changing Structure of Final Demand Spending (as Fractions of GDP)



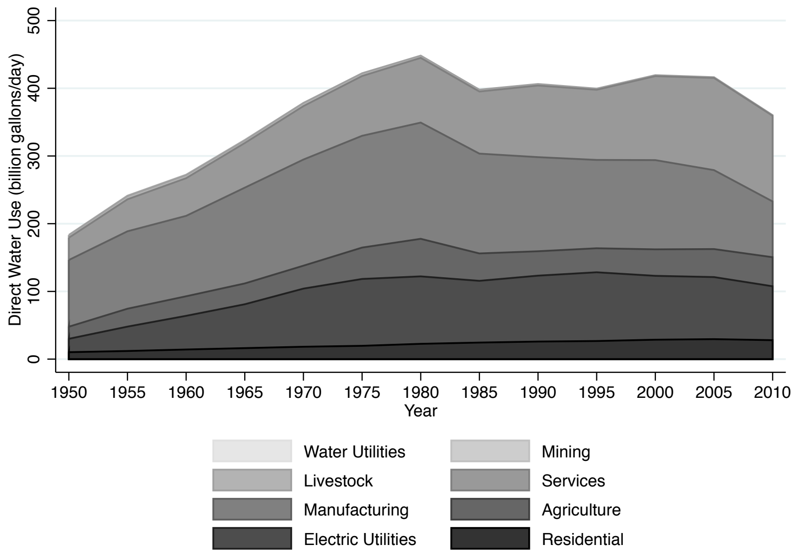
Notes: This figure shows each sector's share of total final demand spending.

Figure 3.2: Direct Water Use



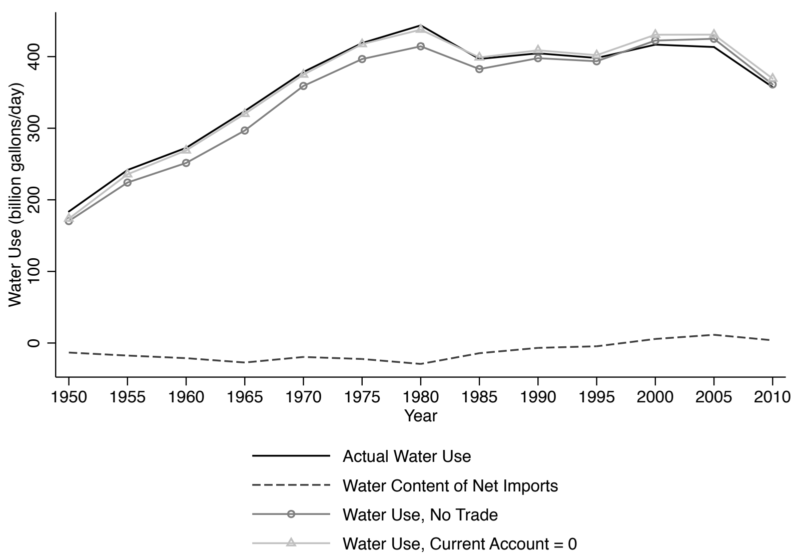
Notes: Data USGS, calculations authors.

Figure 3.3: Total (Direct and Indirect) Water Use



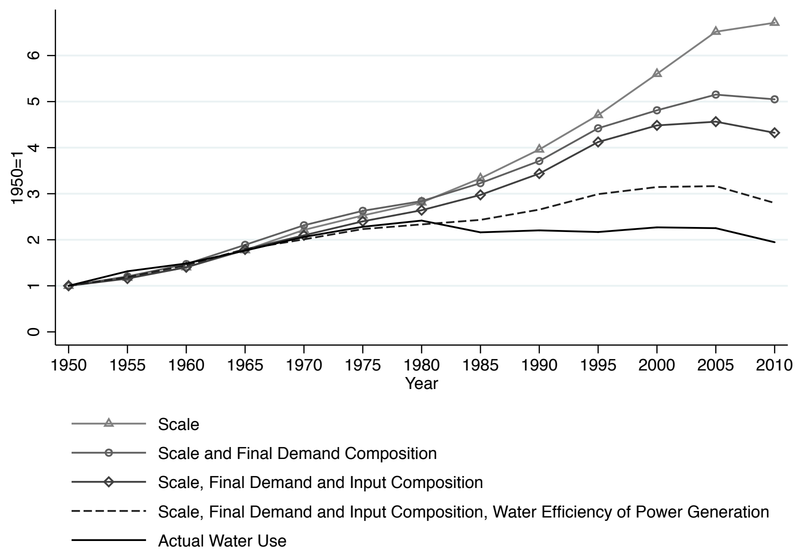
Notes: Data USGS, calculations authors.

Figure 3.4: Water Content of Net Imports and the Hypothetical Closed Economy



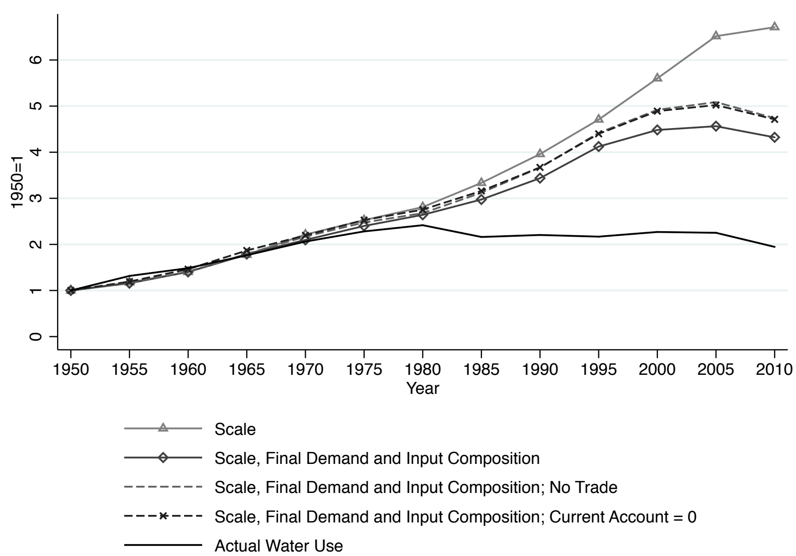
Notes: This figure shows the water content of net imports, actual water use, and hypothetical water use under the two trade scenarios ($Water\ use\ of\ Closed\ Economy\ =\ Actual\ water\ use\ +\ Water\ use\ of\ net\ imports$).

Figure 3.5: Decomposition with Double Deflation



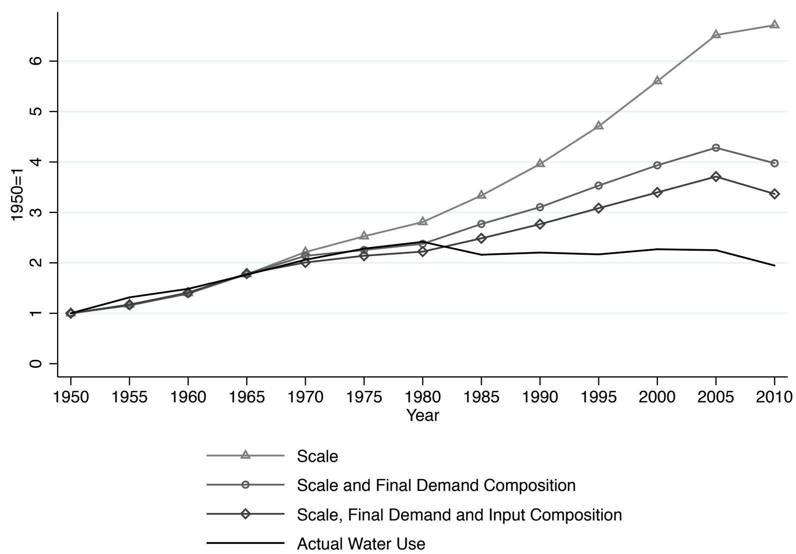
Notes: This figure shows the decomposition using the double deflation method. Actual water use allows all components—scale, final demand and input composition, power generation water efficiency, and technique—to change over time. The size of each effect is reported in Table 3.1.

Figure 3.6: Decomposition: Hypothetical Closed Economy



Notes: This figure shows the decomposition using the double deflation method, adding curves for the hypothetical trade scenarios. Actual water use allows all components—scale, final demand and input composition, power generation water efficiency, and technique—to change over time. The size of each effect is reported in Table 3.3.

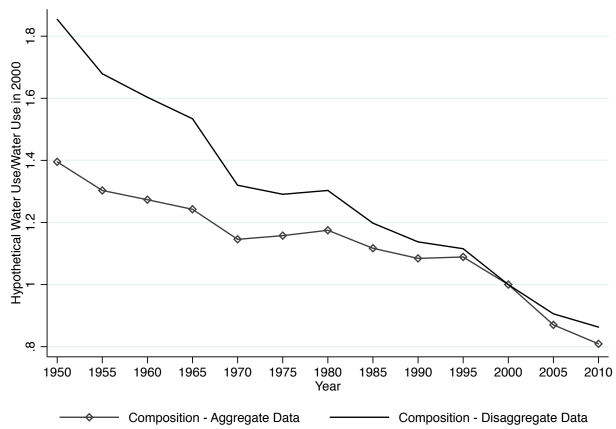
Figure 3.7: Decomposition with Services and Rest of the Economy



Notes: This figure shows the decomposition using the double deflation method, classifying all non-service sectors as the rest of the economy. Actual water use allows all components—scale, final demand and input composition, power generation water efficiency, and technique—to change over time.

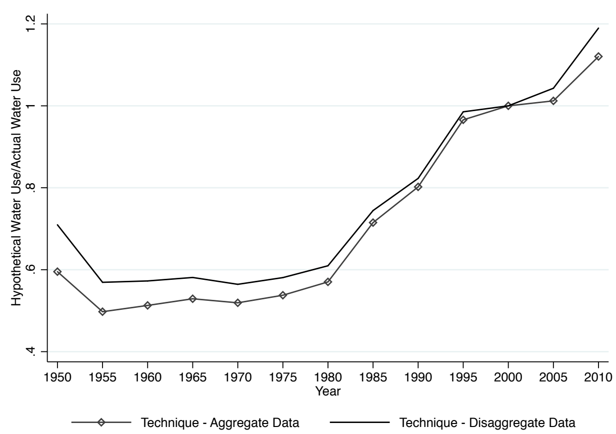
Figure 3.8: Disaggregation Exercise

(a) Changing composition toward less-water-intensive products over time: Aggregate vs. Disaggregate Data



Notes: Own calculations, comparing what water use would have been producing 2000 GDP with 2000 technology but letting the composition (inputs and demand) change over time to actual water use in 2000.

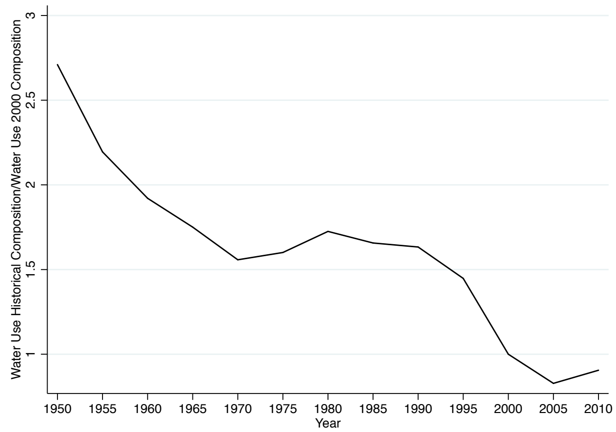
(b) Uneven technological progress



Notes: Own calculations, comparing what water use would have been producing actual GDP with 2000 technology to actual water use for disaggregate and aggregate data.

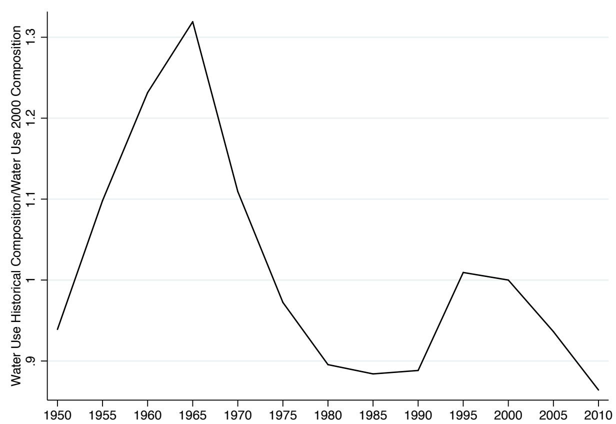
Figure 3.9: Changing Product Compositions by Sector

(a) Changing product composition inside Manufacturing



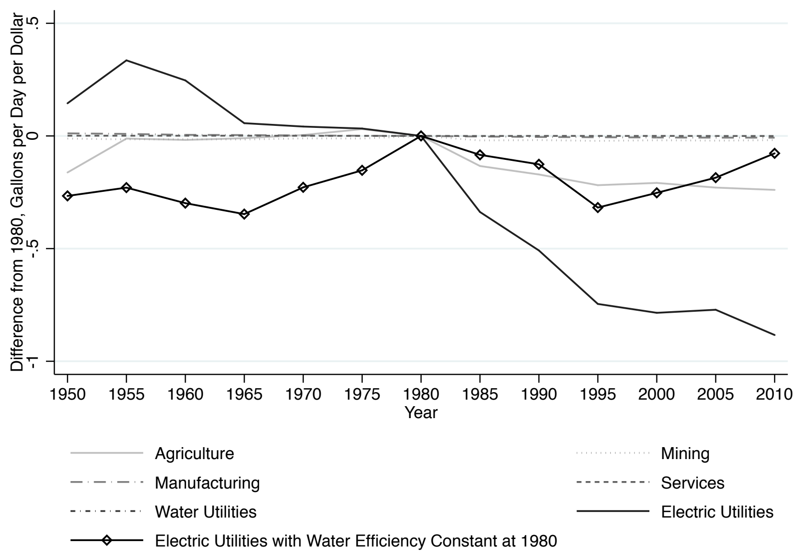
Notes: Own calculations, comparing what water use would have been producing final demand inside the manufacturing sector for 2000 with 2000 water technology.

(b) Changing product composition inside Services



Notes: Own calculations, comparing what water use would have been producing final demand inside services for 2000 with 2000 water technology.

Figure 3.10: Water Intensity Across Sectors (Relative to 1980)



Notes: This figure shows water use per dollar of gross output, **w**, relative to 1980.

References

Allan, J. (1994). Overall Perspectives on Countries and Regions. In P. Rogers, & P. Lydon (Eds.) *Water in the Arab World: Perspectives and Prognoses*, (pp. 65–100). Cambridge: Harvard University Press.

Amiti, M., & Konings, J. (2007). Trade Liberalization, Intermediate Inputs, and Productivity: Evidence from Indonesia. *The American Economic Review*, *97*(5), 1611–1638.

An, G., & Puttitanun, T. (2009). Revisiting McCallum's Border Puzzle. *Economic Development Quarterly*, *23*(2), 167–170.

Anderson, J. E., & van Wincoop, E. (2003). Gravity With Gravitas: A Solution to the Border Puzzle. *American Economic Review*, *93*(1), 170–192.

Anderson, J. E., & van Wincoop, E. (2004). Trade Costs. *Journal of Economic Literature*, *42*(3), 691–751.

Anisfeld, S. (2010). *Water Resources*. Washington, DC: Island Press.

Antràs, P., Chor, D., Fally, T., & Hillberry, R. (2012). Measuring the Upstreamness of Production And Trade Flows. *American Economic Review Papers And Proceedings*, *102*(3), 412–416.

Arkolakis, C., Costinot, A., & Rodríguez-Clare, A. (2012). New Trade Models, Same Old Gains? *The American Economic Review*, *102*(1), 94–130.

Autor, G., Dorn, D., & Hanson, G. (2013). The China Syndrome: Local Labor Market Effects of Import Competition in the United States. *American Economic Review*, *103*(6), 2121–2168.

Averyt, K., Meldrum, J., Caldwell, P., Sun, G., McNulty, S., Huber-Lee, A., & Madden, N. (2013). Sectoral Contributions to Surface Water Stress in the Coterminous United States. *Environmental Research Letters*, *8*(3), 1–9.

Baldwin, R., & Harrigan, J. (2011). Zeros, Quality, and Space: Trade Theory and Trade Evidence. *American Economic Journal: Microeconomics*, *3*(2), 60–88.

Baldwin, R. E. (2008). *The Development and Testing of the Heckscher-Ohlin Trade Models, A Review*. Cambridge: MIT Press.

Bergstrand, J. H., & Egger, P. (2010). A General Equilibrium Theory for Estimating Gravity Equations of Bilateral FDI, Final Goods Trade, and Intermediate Trade Flows. In S. Brakman, & P. van Bergeijk (Eds.) *The Gravity Model in International Trade: Advances and Applications*. Cambridge University Press, Cambridge, UK.

Berthelon, M., & Freund, C. (2008). On the Conservation of Distance in International Trade. *Journal of International Economics*, *75*(2), 310–320.

Blackhurst, M., Hendrickson, C., & i Vidal, J. S. (2010). Direct and Indirect Water Withdrawals for U.S. Industrial Sectors. *Environmental Science Technology*, *44*(6), 2126–2130.

Boulhol, H., & de Serres, A. (2010). Have Developed Countries Escaped the Curse of Distance? *Journal of Economic Geography*, *10*(1), 113–139.

Brock, W. A., & Taylor, M. S. (2005). Economic Growth and the Environment: A Review of Theory and Empirics. In P. Aghion, & S. Durlauf (Eds.) *Handbook of Economic Growth, Volume 1, Part B*, (pp. 1749–1821). Amsterdam: North Holland.

Broda, C., & Weinstein, D. (2006). Globalization and the Gains from Variety. *Quarterly Journal of Economics*, *121*(2), 541–585.

Brun, J., Carrere, C., Gillaumont, P., & de Melo, J. (2005). Has Distance Died? Evidence from a Panel Gravity Model. *World Bank Economic Review*, *19*(1), 99–120.

Buera, F., & Kaboski, J. (2012). The Rise of the Service Economy. *American Economic Review*, *102*(6), 2540–2569.

Burger, M. J., Oort, F. G. V., & Linders, G. (2009). On the Specification of the Gravity Model of Trade: Zeros, Excess Zeros and Zero-Inflated Estimation. ERIM Report Series Reference No. ERS-2009-003-ORG.

Caliendo, L., & Parro, F. (2012). Estimates of the Trade and Welfare Effects of NAFTA. NBER Working Paper 18508.

Chapagain, A. K., & Hoekstra, A. Y. (2008). The Global Component of Freshwater Demand and Supply: An Assessment of Virtual Water Flows Between Nations as a Result of Trade in Agricultural and Industrial Products. *Water International*, *33*(1), 19–32.

Chor, D. (2010). Unpacking Sources of Comparative Advantage: A Quantitative Approach. *Journal of International Economics*, *82*, 152–167.

Cicas, G., Matthews, H. S., & Hendrickson, C. (2006). The 1996 Benchmark Version of the Economic Input-Output Life Cycle Assessment (EIO-LCA) Model. Green Design Institute, Carnegie Mellon University.

Coe, D., Subramanian, A., & Tamirisa, N. (2007). The Missing Globalization Puzzle. IMF Staff Papers 54.

Cosmo, V. M. D., Hyland, M., & LLop, M. (2012). Disentangling Water Usage in the European Union: A Decomposition Analysis. MPRA Paper 42865.

Costinot, A., Donaldson, D., & Komunjer, I. (2012). What Goods Do Countries Trade? A Quantitative Exploration of Ricardo's Ideas. *Review of Economic Studies*, *79*(2), 581–608.

Costinot, A., & Rodríguez-Clare, A. (2013). Trade Theory with Numbers: Quantifying the Consequences of Globalization. NBER Working Paper 18896.

Dekle, R., Eaton, J., & Kortum, S. (2007). Unbalanced Trade. *The American Economic Review*, *97*(2), 351–355.

Dietzenbacher, E., & Hoen, A. R. (1998). Deflation of Input-Output Tables from the User's Point of View: A Heuristic Approach. *Review of Income and Wealth*, *44*(1), 111–122.

Disdier, A., & Head, K. (2008). The Puzzling Persistence of the Distance Effect on Bilateral Trade. *Review of Economics and Statistics*, *90*(1), 37–48.

Eaton, J., & Kortum, S. (2002). Technology, Geography, and Trade. *Econometrica*, *70*(5), 1741–1779.

Feenstra, R. C., Inklaar, R., & Timmer, M. P. (2015). The Next Generation of the Penn World Table. forthcoming *American Economic Review*, available for download at www.ggdc.net/pwt.

Felbermayr, G. J., & Kohler, W. (2006). Exploring the Intensive and Extensive Margins of World Trade. *Review of World Economics*, *142*(4), 642–674.

Flörke, M., Kynast, E., Börlund, I., Eisner, S., Wimmer, F., & Alcamo, J. (2013). Domestic and Industrial Water Uses of the Past 60 Years as a Mirror of Socio-Economic Development: A Global Simulation Study. *Global Environmental Change*, *23*(1), 144–156.

Gleick, P. (2003). Global Freshwater Resources: Soft-Path Solutions for the 21st Century. *Science*, *302*(5650), 1524–1528.

Gleick, P. (2013). Water Use. *Annual Review of Environmental Resources*, *28*, 275–314.

Global Value Chains and Development: Investment and Value Added Trade in the Global Economy (2013). United Nations Conference on Trade and Development.

Greene, W. H. (2012). *Econometric Analysis, Seventh Edition*. Upper Saddle River, New Jersey: Prentice Hall.

Grobovšek, J. (2011). Development Accounting with Intermediate Goods. Universitat Autònoma de Barcelona.

Grossman, G., & Krueger, A. (1991). Environmental Impacts of a North American Free Trade Agreement. NBER Working Paper 3914.

Harrigan, J. (2010). Airplanes and Comparative Advantage. *Journal of International Economics*, *82*(2), 181–194.

Head, K., & Mayer, T. (2014). Gravity Equations: Workhorse, Toolkit, and Cookbook. In G. Gopinath, E. Helpman, and K. Rogoff (Ed.) *Handbook of International Economics, Volume 4*. Elsevier Science.

Heberger, M. (2011). The World's Water Volume 7: The Biennial Report on Freshwater Resources. In P. H. Gleick (Ed.) *Australia's Millennium Drought: Impacts and Responses*, (pp. 97–126). Washington, DC: Island Press.

Hoekstra, A. Y., & Chapagain, A. (2008). *Globalization of Water: Sharing the Planet's Freshwater Resources*. Singapore: Blackwell Publishing.

Hoekstra, A. Y., & Mekonnen, M. M. (2012). The Water Footprint of Humanity. *Proceedings of the National Academy of Science*, *109*(9), 3232–3237.

Hsieh, C.-T., & Klenow, P. J. (2007). Relative Prices and Relative Prosperity. *The American Economic Review*, *97*(3), 562–585.

Hummels, D. (2007). Transportation Costs and International Trade in the Second Era of Globalization. *Journal of Economic Perspectives*, *21*(3), 131–154.

Inklaar, R., & Timmer, M. P. (2008). GGDC Productivity Level Database: International Comparisons of Output, Inputs and Productivity at the Industry Level. GGDC Research Memorandum 104. Groningen Growth and Development Centre, University of Groningen.

Jia, S., Yang, H., Zhang, S., Wang, L., & Xia, J. (2006). Industrial Water Use Kuznets Curve: Evidence from Industrialized Countries and Implications for Developing Countries,. *Journal of Water Resources Planning and Management*, *132*(3), 183–191.

Jones, C. I. (2011). Intermediate Goods and Weak Links in the Theory of Economic Development. *American Economic Journal: Macroeconomics*, *3*, 1–28.

Kenny, J., Barber, N., Hutson, S., Linsey, K., Lovelace, J., & Maupin, M. (2009). Estimated Use of Water in the United States in 2005. USGS Circular 1344, Reston, Virginia.

King, G. (1988). Statistical Models for Political Science Event Counts: Bias in Conventional Procedures and Evidence for the Exponential Poisson Regression Model. *American Journal of Political Science*, *32*(3), 838–863.

Leontief, W. (1970). Environmental Repercussions and the Economic Structure: An Input-Output Approach. *Review of Economics and Statistics*, *52*(3), 262–271.

Levchenko, A. A., & Zhang, J. (2013). The Evolution of Comparative Advantage: Measurement and Implications. NBER Working Paper 16806.

Levchenko, A. A., & Zhang, J. (2014). Ricardian Productivity Differences and the Gains from Trade. *European Economic Review*, *65*, 45–65.

Levinson, A. (2008). Environmental Kuznets Curve. Entry in New Palgrave Dictionary of Economics, 2nd Edition.

Levinson, A. (2009). Technology, International Trade, and Pollution from U.S. Manufacturing. *American Economic Review*, *99*(5), 2177–2192.

Levinson, A., & O'Brien, J. (2015). Environmental Engel Curves. NBER Working Paper 20914.

Lewis, J. B., & Linzer, D. A. (2005). Estimating Regression Models in Which the Dependent Variable Is Based on Estimates. *Political Analysis*, *13*(4), 345–364.

Liu, X. (2009). GATT/WTO Promotes Trade Strongly: Sample Selection and Model Specification. *Review of International Economics*, *17*(3), 428–446.

Maupin, M., Kenny, J., Hutson, S., Lovelace, J., Barber, N., & Linsey, K. (2014). Estimated Use of Water in the United States in 2010. USGS Circular 1405, Reston, Virginia.

Mekonnen, M. M., & Hoekstra, A. Y. (2011). National Water Footprint Accounts: The Green, Blue and Grey Water Footprint of Production and Consumption. Value of Water Research Report Series No. 50, UNESCO-IHE, Delft, the Netherlands.

Melitz, M. J., & Redding, S. J. (2014). Missing Gains from Trade? In *Papers and Proceedings of One Hundred Twenty-Sixth Annual Meeting of The American Economic Association, 104(5)*, (pp. 317–321).

Miller, R., & Blair, P. (2009). *Input-output Analysis: Foundations and Extensions*. New York: Cambridge University Press.

Miroudot, S., Lanz, R., & Ragoussis, A. (2009). Trade in Intermediate Goods and Services. OECD Trade Policy Working Papers No. 93. OECD Publishing.

O'Mahony, M., & Timmer, M. P. (2009). Output, Input and Productivity Measures at the Industry Level: The EU KLEMS Database. *Economic Journal*, *119*, F374–F403.

Ossa, R. (2012). Why Trade Matters After All. NBER Working Paper 18113.

Rauch, J. E. (1999). Networks Versus Markets in International Trade. *Journal of International Economics*, *48*(1), 7–35.

Rauch, J. E., & Trindade, V. (2003). Information, International Substitutability, and Globalization. *American Economic Review*, *93*(3), 775–791.

Reshef, A. (2013). Is Technological Change Biased Towards the Unskilled in Services? An Empirical Investigation. *Review of Economic Dynamics*, *16*(2), 312–331.

Santos Silva, J. M. C. and Tenreyro, S. (2006). The Log of Gravity. *Review of Economics and Statistics*, *88*(4), 641–658.

Shikher, S. (2011). Capital, Technology, and Specialization in the Neoclassical Model. *Journal of International Economics*, *83*, 229–242.

Shikher, S. (2012). Putting Industries into the Eaton-Kortum Model. *The Journal of International Trade and Economic Development*, *21*(6), 807–837.

Shikher, S. (2013). Determinants of Trade and Specialization in the Organisation for Economic Co-Operation and Development Countries. *Economic Inquiry*, *51*(1), 138–158.

Shiklomanov, A. (1998). World Water Resources and their Use. International Hydrological Program. http://webworld.unesco.org/water/ihp/db/shiklomanov/.

Siliverstovs, B., & Schumacher, D. (2008). Disaggregated Trade Flows and the Missing Globalization Puzzle. KOF Working Papers 209, KOF Swiss Economic Institute. Swiss Federal Institute of Technology Zurich.

Simonovska, I., & Waugh, M. E. (2014). The Elasticity of Trade: Estimates and Evidence. *Journal of International Economics*, *92*(1), 34–50.

Timmer, M., de Vries, G., & de Vries, K. (2014). Patterns of Structural Change in Developing Countries. GGDC Research Memorandum 149.

Timmer, M. P., Ypma, G., & van Ark, B. (2007). PPPs for Industry Output: A New Dataset for International Comparisons. Groningen Growth and Development Centre, University of Groningen.

Timmer, Marcel P. (ed) (2012). The World Input-Output Database (WIOD): Contents, Sources and Methods. WIOD Working Paper Number 10.

Vörösmarty, C. J., Green, P., Salisbury, J., & Lammers, R. B. (2000). Global Water Resources: Vulnerability from Climate Change and Population Growth. *Science*, *289*(5477), 284–288.

Water in a Changing World: The United Nations World Water Development Report 3 (2009). World Water Assessment Programme. London: United Nations Educational, Scientific and Cultural Organization.

Waugh, M. E. (2010). International Trade and Income Differences. *The American Economic Review*, *100*, 2093–2124.

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data, Second Edition*. Cambridge, Massachusetts: The MIT Press.

Yi, K.-M., & Zhang, J. (2011). Structural Change in an Open Economy. Federal Reserve Bank of Minneapolis Staff Report 456.

Appendix A

Appendix to Chapter 3

A.1 Data construction

In this section we discuss in detail some of the data issues.

Assigning publicly supplied water to the sectors:

We assign water use in the Public Supply category—water supplied by water utilities—to the sectors that use the water in the following way:

We first take a fraction of publicly supplied water to be residential use. We take this share to be 0.58, which is based on the relatively stable share of publicly supplied water use that goes to residential users that is given by the USGS for the years 1985-1995 and 2005. We then allocate the remaining publicly supplied water to the sectors using the share of payments by each sector to water utilities from the input-output tables.

Estimating missing water data:

The USGS does not provide data for all sectors in all years. We use the following procedures to estimate the missing values:

Mining

Mining self-supplied water use is included in self-supplied industrial water use for the years 1950-1980. We remove it by taking the ratio of self-supplied to publicly supplied (by water utilities) mining water use for the years 1985-2005 and applying it to publicly supplied mining water use for the missing years.

Commercial (Services)

We use the same procedure as for mining above (using data for the years 1985-1995) to separate commercial self-supplied water use from industrial self-supplied water use for the years 1950-1980. We also use this method to estimate commercial water use for 2000-2005, years for which the USGS does not estimate commercial water use as part of any category.

Aquaculture

We remove self-supplied aquaculture water use from self-supplied industrial water use for the years 1950-1980 by applying the growth rate in aquaculture tonnage to aquaculture water use for the years with aquaculture water use data (1985-2005).

Industrial (Manufacturing)

We subtract estimated mining, commercial, and aquaculture water use from self-supplied industrial water use for 1950-1980.

Estimating missing input-output data:

In some years, the input-output tables do not split utilities into water and electric utilities, and net exports are not split into imports and exports. We impute the missing values in the following ways:

Utilities

In 1947 and 1958, electric and water utilities are not split from utilities. We use the 1963 ratio to split the utilities category. The remaining component of utilities, gas utilities, is added back into services.

Specifically: Using the 1963 data, we compute the share of each cell that involves utilities (as producing or consuming sector) that is comprised by electric utilities or water utilities. We apply these shares to the corresponding utilities cells in the 1947 and 1958 input-output tables. To ensure that gross output balances by sector, we leave one row component of utilities empty (we use gas utilities), and compute the value as the difference between gross output computed going across rows and computed going down columns. The final cell (gas utilities x gas utilities) is calculated such that total gross output in the economy is the same as before the split was applied.

Trade

In 1958, 1963, and 1967, net exports are not divided into exports and imports.

To split net exports into exports and imports: We compute the growth rate of imports for each sector: Agriculture, Livestock, Mining, Manufacturing, and Services (Water and Electric Utilities assumed same rate of growth as Services). Goods import data are SITC Rev. 1 from the WITS database (World Integrated Trade Solution). Services import data are from the Balance of Payments from the BEA. We apply the growth rates back from 1972 to compute imports for 1958, 1963, and 1967. The trade data is not available prior to 1963, so we apply the 1963-1967 growth rate to compute imports in 1958. Exports are computed as the sum of the estimated import levels and net exports (from the input-output tables).

Concordance of input-output tables over time:

To match the sectors over time, we match the input-output sectors for the 1997 data with the NAICS categories and convert to SIC using a concordance from the BEA. We then convert the SIC categories to the 1992 input-output classification, also using a concordance from the BEA. The mapping from NAICS to the 1992 input-output sectors is not one-to-one. In cases where one NAICS category maps to many 1992 input-output sectors we distribute the value in the NAICS category according to the relative sizes of the 1992 input-output sectors within a NAICS category.

A.2 Tables

Table A.1: Scaling Factor for the Aggregate Water Content of Trade

	Exports	Imports
1950	1.21	0.95
1955	1.44	0.89
1960	1.52	0.86
1965	1.49	0.88
1970	1.17	0.85
1975	1.01	0.80
1980	0.96	0.75
1985	1.01	0.78
1990	0.98	0.79
1995	0.94	0.75
2000	0.90	0.72
2005	0.93	0.79
2010	0.83	0.75

Notes: This table reports the scaling factors used to adjust the aggregate water content of exports and imports.

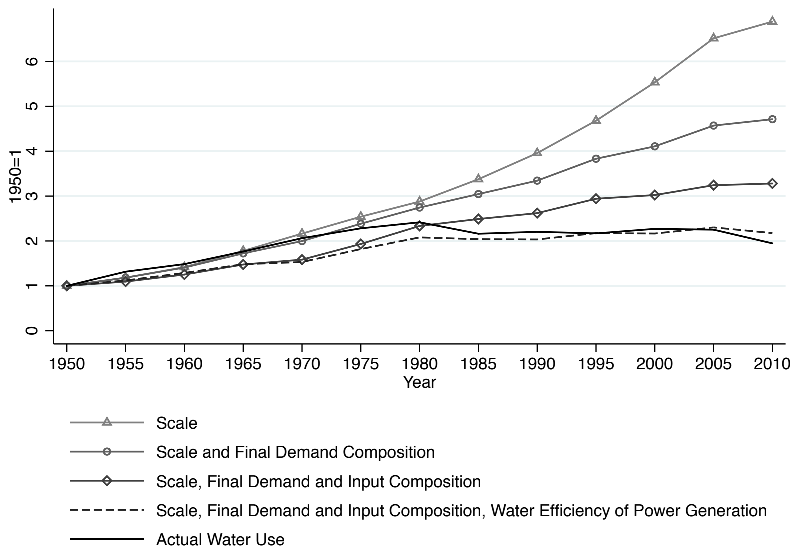
Table A.2: Decomposition with GDP Deflator

Increase in Water Use Since 1950, Allowing the following to change over time:	
Scale	6.89
Scale and Final Demand Composition	4.71
Scale, Final Demand and Input Composition	3.28
Scale, Final Demand and Input Composition, and Water Efficiency of Power Generation Sector	2.18
Scale, Final Demand and Input Composition, Water Efficiency of Power Generation Sector, and Technique (Actual Water Use)	1.95
Fraction of Water Productivity Improvement Explained by:	
Final Demand Composition	0.44
Input Composition	0.29
Water Efficiency of Power Generation Sector	0.22
Technique	0.05

Notes: The table reports the size of the effects and the fraction of the overall water productivity improvement explained by each effect for the decomposition shown in Figure A.1.

A.3 Figures

Figure A.1: Decomposition with GDP Deflator



Notes: This figure shows the decomposition using the GDP deflator. Actual water use allows all components—scale, final demand and input composition, power generation water efficiency, and technique—to change over time.