A New Measure of Economic Distance

Eric O'N. Fisher
California State Polytechnic University

John Gilbert
Utah State University

Kathryn G. Marshall
California State Polytechnic University

Reza Oladi
Utah State University

[First Draft - Do Not Cite]

Abstract

This paper defines a new measure of economic distance. Using consistent input-output data, we estimate local unit costs for 35 sectors in 39 countries. The distance between two countries is the largest percentage difference in unit costs among all sectors. If all goods are traded, this distance is the uniform *ad valorem* tariff that shuts down bilateral trade. The median economic distance between the United States and its 38 trading partners is 94%. The network induced by the closest 10% of these distances has a large component with two clusters, one corresponding to the advanced economies and another to the emerging economies. China, India, and several other countries are isolated components, indicating that their unit costs are idiosyncratic. We also introduce a new measure of revealed comparative advantage. (JEL Codes: F1)

1 Introduction

This paper presents a new measure of economic distance: two economies are "close" if their local unit costs for every good are similar. This measure was developed with empirical applications in mind, and it takes full advantage of the recent flowering of internationally consistent macroeconomic data on the production sets of different countries.

There are many underlying reasons for trade costs. Some foreign countries are far away, and it costs a lot to ship some goods to a distant market. It is harder to sell a good abroad than at home because language is a barrier. Perhaps shipping across the border involves unfamiliar bureaucratic impediments. There may be an explicit tariff or non-tariff barrier. The foreign culture may well seem alien to any firm trying to expand sales by exporting. It is not easy to obtain the usual financing that facilitates the relations between local suppliers and wholesalers. The foreign legal system makes it much more difficult to enforce the norms that allow for transacting business in a repeated relationship. Exchange rate risk makes foreign sales less attractive. This list is far from exhaustive.

Still, we exploit a deep insight. *All these impediments to trade will be reflected in differences in unit costs.* Any departure from the law of one price must reflect an implicit trade cost. In fact, the only reason for two physically homogeneous goods to sell for different prices in disparate markets is that it costs money to trans-ship them. It doesn't matter if it costs money to bribe a customs officer, to pay for airfreight, or to hire a translator, if an apple sells for $1 in New York and a *pomme* sells for $2 in Paris, then you can be sure that the generalized cost of shipping an apple across markets is at least $1.

This simple insight is the essence of our measure of bilateral economic distance. Let the unobservable world price of an apple be given. Assume that the estimated local unit cost of a *pomme* in Paris is 50% above it and that of an *apfel* in Berlin is 30% above it. Then this cost differential is 20%. Do the same calculation for every good recorded in the input-output tables of France and Germany. These goods include traded and non-traded goods, but we are agnostic about this difference.[1] The maximum of the absolute value of these differences is the economic distance between France and Germany. In fact it is 68%. So a uniform tariff of 68% would shut down all trade between France and Germany; also, since we are including not traded goods in our calculation, a 68% discriminatory tax on foreigners would discourage every German from renting an apartment in Paris and *vice versa*.

Macroeconomic accounts record many more goods than factors of production. In this case, the observed vector of outputs occurs on the interior of a flat on the production possibility frontier. Small exogenous changes on the demand side will have no effect on the supporting price vector as long as all sectors remain active. Prices are determined by fixed marginal rates of technical substitution, and quantities adjust to equilibrate supply and demand.

Because we want to apply our measure, we are motivated by the data in hand. The production structure of an economy is a technology matrix that records the unit values of direct and indirect resource requirements for each sector. Our data have thirty-five sectors and four factors. Then the data on unit input requirements are actually an overdetermined system of 35 equations in the 4 unknown factor prices. Our measure exploits this insight fully.

Let there be n goods and f factors. The space of local unit costs is a cone in n-dimensional space spanned by the f vectors that describe factor uses. Consider a simple Ricardian economy where it takes one hour of labor to make Good 1 and one hour of labor to make Good 2. All possible local unit costs

[1] If the pound sterling gets strong enough and there is a price war on flights between Heathrow and JFK, a Broadway play eventually becomes a traded good for the right kind of Londoner.

are spanned by the vector $(1,1)^T$. What would local unit costs be if this economy faced world prices $p = (2,1)^T$? The usual answer is that the economy would specialize completely in the first good, and the second sector would shut down. National income accounts would show only one active local sector.

We do not see this paucity of active sectors in the data. We actually observe almost every good produced in practically every country; this phenomenon cannot be explained by simple applied general equilibrium theory. An explanation is that every national economy is somehow protected by price wedges that allow many local industries to be active. Think of iceberg transportation costs as uniform percentage price markups for every good. What are the smallest such costs that would allow both sectors to be active in this case? The least squares project of world prices onto local unit costs is $(1.5,1.5)^T$; this is the point closest to world prices that is consistent with both sectors operating. The opportunity to export the first good drives up local wages to 1.5 and thus raises the unit costs of both goods. If iceberg transportation costs were at least 50%, importing the second good and exporting the first good would no longer make sense. Our distance measure would record this country as being 0.5 from any technology that could produce at world prices. This simple example captures the essence of our projection matrices that estimate local unit costs.

Our measure deals with an important subtlety in the data. Physical technology matrices are not observable; one cannot measure hours of labor per unit of Good 1 because of the way intermediate inputs are measured. National accounts constrain empirical work in an important way. Only the value of output of Good 1 is observed; neither price nor quantity is observable separately. One can indeed record hours of labor used in the sector. The ratio of hours of labor to the value of output is a canonical element in a unit-value matrix. The full employment condition for labor implies that one observes the physical input requirement per dollar of output, but *neither the physical technology nor the price out output is identified without an ancillary assumption.* Our identifying assumption is that every country's unit value technology matrix defines *unobservable* physical units by the *unobservable* world price. The definition of a unit of Good 1 is an international dollar's worth of it. This assumption implies that different unit-value matrices record disparate physical technologies. This assumption also lends a huge advantage; the unit value of every good at world prices is unity! Then the distance of a country from world prices can be computed from the least squares projection of *the unit vector* onto its local technology as measured by its unit-value matrix. Its distance from the world is the absolute value of the largest residual from this regression.

The distance between Canada and the United States is 0.66, the largest absolute value of the differences of the thirty-five estimated local unit costs in our data. This difference occurs in "Renting of machinery and equipment and other business activities". The smallest difference in estimated unit costs occurs in "Hotels and restaurants", where it is 0.02. Since world unit costs are one for every good, each of these numbers has an interpretation as a percent; in fact renting and leasing of equipment is estimated to cost 66% more in Canada than in the United States.

Since we are constructing a measure of distance, we must use the largest of the absolute values of these estimated cost differences. This measure is conservative because it has to satisfy the triangle inequality. The average absolute value of the bilateral cost differences is only 21% for the United States and Canada. We interpret the maximal cost difference as a lower bound for a uniform bilateral prohibitive tariff; if NAFTA were abrogated and the United States put a 66% tariff or higher on every good coming from Canada and a 66% export tax or higher on every good shipped to Canada, then all bilateral trade would shut down. The larger this number is, the more that the typical consumer in Canada or the United States benefits from bilateral trade. We think of our measure as capturing generalized economic distance, since some goods are traded and others are not. The first sixteen goods recorded in our data are traded sectors, and we also apply our measure to those sixteen sectors only. The bilateral distance based on traded goods is 0.41 because Canada has a 41% cost advantage in "Wood and products of wood and cork".

Our main contribution is to define this economic distance and to bring it to the data. We define 741 bilateral economic distances among 39 countries. Using a technique from graph theory, we show that a network consisting of the 10% of closest links has a large connected component that breaks into two clusters. The first contains most of the advanced industrial countries, and the second has most of the emerging economies; the bridge between the two is the link between the Czech Republic and Poland. China, India, and Turkey are all isolated components. When we restrict our attention to the distance based on traded goods only, we see that the network described by the closest 10% of links has a single component that amalgamates the two former clusters and includes China. India, Russia, and Mexico are now isolated components. It seems that NAFTA has not equalized traded goods prices between Mexico and the United States; otherwise, this edge would be among the 10% of closest links.

Our distance measures are only as good as our projected unit costs. We examine these in detail in three ways. First, we corroborate that our costs are significantly correlated with those of other scholars who have worked with these data in a very different way. Second, we explore the real exchange rates that are implied by our projected unit costs, and show that rich countries have a high real cost of non-traded goods. Finally, we explore innovative measures of revealed comparative advantage for China, Mexico, and the United States. China has strongest revealed comparative advantage in "Basic and fabricated metals" and strong comparative disadvantage "Agriculture, forestry, and fisheries".

We hope that this brief summary of our empirical results has whetted your appetite for our theory and its applications. The rest of this paper is structured as follows.

The second section gives a review of the literature on trade costs and some discussion of the recent work in empirical trade that draws out the effects of differences in technology. The third section is the heart of our theoretical contribution; we develop our pseudo-metric and give some examples that flesh out one's economic intuition. In that section, we emphasize that one can measure physical inputs per dollar of output, but neither can be measured separately in the data; we make the identifying assumption each country's unit values are defined in terms of an unobservable vector of international prices. The fourth section gives a brief description of our data; we are very lucky to have the World Input Output Database's internationally consistent and detailed data on factor uses in thirty-five sectors in thirty-nine countries. Our measure of distance was designed precisely to take full advantage of these data.

The fifth section brings our measure to the data in four different ways. First, we show all the bilateral distances among our sample of thirty-nine countries as the graph of a network. We also give the network's minimum cost spanning tree; it is suggestive of a taxonomy of the world economy. Since our distance measures are only as good as our price projections, we then show that these are reasonable. We show that they are statistically significantly correlated with prices computed by other scholars who have used these data. Then we show that they give rise to reasonable measures of the real exchange rate, measured as the relative price of non-traded goods. Finally, we present some interesting new measures of revealed comparative advantage that arise from our price projections. The sixth section gives our conclusions and some suggestions for future research.

2 Review of the literature

If all goods are traded, then our distance measure is the uniform *ad valorem* equivalent cost that shuts down bilateral trade. It is now widely recognized that trade costs are large and consequently have a significant impact on international trade flows and economic welfare. It is also well accepted that trade costs are comprised in only small part by direct policy measures such as tariffs and the tariff equivalents of quotas (Anderson and van Wincoop, 2004), and that other costs of selling goods in foreign markets, such as transportation and freight, time costs, information costs, regulatory costs, and local distribution costs, are much more significant barriers to international trade. Indeed, of Anderson and van Wincoop's famous headline estimate of 170% average developed economy trade costs, only 8% reflects the cost of direct tariff and non-tariff barriers. The significance of broadly defined international trade costs is well-recognized in the policy world also, as the emphasis on trade facilitation in the Bali Package which resulted from the Ninth Ministerial Conference of the WTO in late 2013 indicates.

While the economic importance of trade costs is not in dispute, the measurement of trade costs remains difficult. There are two broad approaches that have been adopted in the literature. The first uses direct measures to construct estimates of the various components of trade costs; in principle, they can then be aggregated into a total measure. The second approach is indirect, using data on traded quantities and prices to infer international trade costs; Chen and Novy (2012) are a good example. As Anderson and van Wincoop (2004) note in their extensive survey of the early literature, the former approach is plagued by data inadequacies, while the latter inevitably involves the use of economic theory.

Tariff barriers are the easiest component of trade costs on which to obtain direct measures of incidence. Applied tariff barriers are available through UNCTAD's TRAINS database, and bound rates are accessible through the WTO's Consolidated Tariff Schedules. Information on the prevalence–but not the impact–of non-tariff barriers is also available through TRAINS. However, the data are far from complete, especially for developing economies. The MAcMap dataset does provide cross-sectionally complete applied tariff data, but it does so for limited years (Guimbarda et al., 2012). An extensive literature has considered the problem of aggregating the data into economically meaningful measures, following Anderson and Neary (1996). An excellent recent discussion is Kee et al. (2009).

By contrast there is no uniform source of direct measures of transportation and freight costs, and these are not widely available, even from national sources. Because trade flow data are widely available at aggregated levels, through the IMF's DOTS, and at disaggregated ones, through the UN's COMTRADE, it is possible in principle to exploit the dual reporting of each flow in FOB from the export side and CIF from the importing side to impute transportation costs. For an example see the database constructed by CEPII and described in Gaulier et al. (2008). Unfortunately, by comparing the measures obtained in this manner with directly measured freight costs for two countries where such data are available (the United States and New Zealand), Hummels and Lugovskyy (2006) find that the constructed measures are badly error-ridden and contain no useful time-series or cross-commodity variation. Hummels (2007) reviews the available sources of data on direct freight costs, and documents the changes in the cost of international transport over time.

Comprehensive data on other components of trade costs are even more difficult to obtain. While some authors like Burstein et al., 2003 and Bradford, 2005 construct local distribution costs from input-output data, it is common to use proxy data to construct indices for the components such as regulations, standards, and customs procedures that are under study. For example, one could uses counts of the average number of days that are needed for a good to cross the border or frequency counts and coverage ratios of prevailing WT0 standards that require notification. Chen and Novy, 2012 summarize data availability, along with recent studies, in the area of standards and regulation. These proxies are then typically used as covariates in a gravity model, along with other variables such as distance, border and FTA controls, and common language. Making some assumptions about a trade cost function and using knowledge of the model's elasticity of substitution,one can back out *ad valorem* tax equivalents of the relative impact of a particular component of trade cost. Anderson and van Wincoop, 2004 give the details.

Recent examples of this approach include Chen and Mattoo (2008) and Essaji (2008), who consider the role of standards and regulations, and Portugal-Perez and Wilson (2012), who consider infrastructure. Hummels and Schaur (2013) use data on air and maritime transportation of imports to the US to estimate the significance of time as a trade cost, finding that each day in transit is equivalent to an *ad valorem* tariff of between 0.6 and 2.3 percent. In addition to data availability issues, weaknesses of this general approach include the arbitrariness of the assumed trade cost function, and the potential for omitted variable bias, given the unobserved nature of many aspects of trade costs.

Given the difficulties inherent in obtaining direct measures of trade cost components and then converting them into a usable measure of incidence, a number of researchers, beginning with Head and Ries (2001), have turned to indirect measures. In essence, this approach turns gravity on its head, inferring trade costs from the trade data without specifying a trade cost function.[2] It does so using a neat algebraic trick. The gravity equation is solved for the unobservable trade costs as a function of bilateral trade flows and the multilateral resistance variables. The latter are unobservable, but can be canceled out using trade flows in the opposite direction, and domestic trade flows in each trading partner. Given knowledge of the elasticity of substitution, it is possible to back out *ad valorem* tax equivalents. Chen and Novy (2012) and Novy (2013) provide the details.

The indirect technique can be applied to a much wider range of countries and time periods, but it has some disadvantages. By its nature, the technique yields aggregate trade costs, not information on any particular component, and can only determine the geometric average of bilateral trade costs between any country pair. Moreover, because it is based on a calibration of trade flow data, any measurement error is passed through. Recent applications include Chen and Novy (2011), who introduce a correction to the measure used by Head and Ries (2001) for heterogeneity across industries by using industry-specific substitution elasticity estimates. Jacks et al. (2011) construct the measure over a long time span, and

[2] In this sense, the technique is closely related to the earlier trade potentials literature.

show that it has significant power in explaining changes in trade flows. Novy (2013) shows that the indirect trade cost measure is consistent with a wide variety of underlying trade models. The technique has also recently been used to construct a new World Bank database on trade costs for a large group of developed and developing economies (Arvis et al., 2013).

The popularity of the gravity equation is due in part to the perceived empirical limitations of the Heckscher-Ohlin-Vanek (HOV) theory, although these two perspectives are not mutually exclusive. In the HOV framework, persistent differences in prices of the same good in different countries is in itself indirect evidence of high trade costs. To the extent that these trade costs are proportional to geographical distance and size, the gravity model of trade is empirically verified. See Deardorff, 1998 for an example.

The HOV model in its simplist textbook form has been discredited by a lengthy literature, starting with Trefler (1993) and Trefler (1995). The main reason for the model's failings has to do with the lack of a common technology. In their study of ten OECD economies using data from 1985, Davis and Weinstein (2001) incorporate a gravity specification to the demand side of the HOV model and show that it improves factor content predictions. However, this specification is their seventh alteration to the standard model in a meager sample of countries, and on the margin it has a relatively small impact on several of the empirical tests they report. Measuring endowments using the value of factor services, Fisher and Marshall (2013) show strong support for the HOV model. The value of factor services predicts the factor content of trade fares well because these measures already incorporate technical differences in the local factor prices inherent in them.

On the other hand, the evidence against a common technology, especially among countries at different stages of development, is overwhelming. Recent studies, such as Maskus and Nishioka (2009) and Marshall (2012), consider how these differences between developed and developing countries explain the failure of the HOV predictions. Once differences in technology are acknowledged, a thorny question arises on how to best adjust HOV factor content predictions. See Trefler and Zhu, 2010 for the details.