Trade Theory and Trade Growth Since 1870

Mariko Klasing* University of Groningen

Petros Milionis* University of Groningen

Robert Zymek† University of Edinburgh

February 2014

Abstract

To what extent can trade theory account for the long-run evolution of world trade? Our paper offers the first systematic attempt to address this question. To this end, we consolidate existing sources on bilateral trade flows, GDP, tariffs and transportation costs for the world's major economies during the period 1870-2005. We use the data to calibrate a gravity equation of international trade which, for different parameter choices, is consistent with a broad class of quantitative trade models. Irrespective of the parameterisation, and hence the underlying model, the theory-predicted ratio of total trade to world GDP fails to match the time-series behaviour of its empirical counterpart over the last 135 years. However, changes in the relative size of bilateral and regional trading volumes are captured with a surprising degree of accuracy. In the light of our results, we explore realistic generalizations of standard models which could reconcile trade theory with the long-run data.

JEL Classification codes: F12, F15, N70

Keywords: International Trade, Trade Growth, Trade Patterns

*Department of Economics, Econometrics and Finance, Faculty of Economics & Business, University of Groningen, PO Box 800, 9700 AV Groningen, The Netherlands; m.klasing@rug.nl, p.milionis@rug.nl.

†School of Economics, University of Edinburgh, 31 Buccleuch Place, Edinburgh, EH8 9JT, United Kingdom; robert.zymek@ed.ac.uk.

We are grateful to Alejandro Cuñat, Gino Gancia, Sevi Mora, Yona Rubinstein, Thijs van Rens and Jaume Ventura as well as seminar participants at the University of Edinburgh, Brown University, the Barcelona GSE Summer Forum, the University of Vienna and the German Economists' Christmas Meeting for helpful comments and suggestions.

1 Introduction

Calibrated models of international trade struggle to replicate the observed growth in world trade. This fact was first documented by Yi (2003) who shows that the nature and magnitude of the rise in trade flows during recent decades presents a challenge for standard models. A number of papers have since confirmed Yi's observation and proposed ways to reconcile trade theory with the data.[3] Yet, this literature has focused exclusively on trade growth since World War II – a period of uninterrupted progress in global integration. We provide the first assessment of the ability of trade theory to account for the evolution of world trade over 135 years of history, covering most of the modern globalization era. Our findings show that canonical models cannot reproduce the large observed swings in the world trade share but can predict changes in bilateral trading patterns with considerable accuracy. This suggests that simple generalizations may be sufficient to bring standard trade theory in line with the long-run data.

Figure 1 graphs the ratio of world imports to world GDP between 1870 and 2005.[4] It shows that the world trade share grew steadily between 1870 and 1913, reaching 8.3% on the eve of World War I. The two world wars and a surge in protectionism in the interwar years then saw a period of de-globalization. Following the end of World War II, global trade flows recovered, surging to 13.2% of world GDP by 2005. Our paper represents the first attempt to develop a unified theoretical framework of world trading patterns over this period, capable of capturing both the dramatic rise in world trade since 1950 and the prolonged slump which preceded it.

We begin by deriving a theory-consistent gravity equation of international trade. The gravity equation predicts that the value of trade flows between any pair of countries should be positively related to their incomes and negatively to bilateral trade frictions and a weighted measure of their access to other markets. Summing over all countries, it implies that the traded share of world output is a function of the world distribution of GDPs, and prevailing trade

[3]Bergoeing and Kehoe (2003) calibrate a new trade model for the period 1960-1991 and, like Yi, conclude that it is "not capable of explaining the enormous increase in the ratio of trade to income". Yi (2003) and Bridgman (2008, 2012) show that models which allow for vertical specialization are more successful at reproducing the quantitative and qualitative features of post-1950 trade growth. By contrast, Bajona (2004) and Cuñat and Maffezzoli (2007) obtain quantitatively successful predictions from trade models in which comparative advantage responds dynamically to reductions in trade barriers. Zymek (2012) also appeals to a model of comparative advantage to argue that a large portion of the growth in world trade since 1980 can be attributed to the emergence of China.

[4]The world trade share is constructed using bilateral trade flows between a sample of 28 major economies, covering 80-90% of world GDP between 1870 and 2005. For details on data sources and assembly, see Section 2 and the Appendix.

frictions. We exploit the fact that, up to the value of one crucial parameter, the functional form of this gravity equation is the same across a broad class of trade models − including the Armington, Eaton-Kortum, Helpman-Krugman and Chaney-Melitz models, all of which are commonly used for the quantitative analysis of international trade.[5] Therefore, using long-run data on GDPs and trade barriers and exploring a range of parameter values, we can derive predictions for the evolution of the world trade share as well as broad bilateral trading patterns for a range of key theoretical models. We proceed to compare these with the observed trading patterns in the period 1870-2005.

Like the earlier literature, we find that our baseline calibration − roughly corresponding to a Melitz model − cannot capture the magnitude of variation in the global trade share predicting, for example, a 6.3 percentage-point increase between 1870 and 2005, as compared to the 9.5 percentage-point increase observed in the data. Moreover, with the exception of the two Wars, the calibration delivers a steadily rising ratio of world imports to GDP. Alternative parameterisations of the gravity equation improve the fit with the data along some dimensions, but worsen it along others. Our findings cast new light on the difficulty encountered by other authors in replicating strong post-War trade growth in quantitative models: it mirrors an inability, common to all gravity-consistent models, to match the decline in global trade between 1914 and 1945, and the initial extent of world imports during the late 19th century.

Surprisingly, given these shortcomings, the calibrations yield a near-perfect match for the evolution of regional trade *shares*. This suggests that whatever hampers the theory's ability to explain changes in overall world trade over time does not impact its capacity to capture changes in the *distribution* of aggregate and bilateral trade across regions. It leads us to conclude, tentatively, that minor modifications of baseline models could produce a successful theoretical account of the evolution of global trading patterns between 1870 and 2005. We quantitatively explore two plausible generalizations of the framework employed in this paper, and discuss a number of others.

Our paper builds on a venerable line of research seeking to test and improve the quantitative implications of trade theory. One of the motivations for the development of the "new" trade theory by Krugman (1979), Lancaster (1980) and Helpman (1981) was to provide an explanation for the concentration of world trade among a small group of industrialized economies, which appeared puzzling from the vantage point of classical theory. Helpman (1987) showed

[5]See Anderson (1979) and Bergstrand (1985), Eaton and Kortum (2002), Krugman (1979) and Helpman (1987), and Melitz (2003) and Chaney (2008). In the light of their importance for quantitative analysis, Arkolakis et al. (2012) refer to these as the set of "quantitative trade models".

that, aside from predicting bilateral trading patterns, new trade models also had novel implications for trade growth, linking it to increased similarity in countries' incomes. The relevance of this channel for post-War trade growth has been subjected to testing by means of regressions – in Hummels and Levinsohn (1995) and Baier and Bergstrand (2001) – and calibrations – in Bergoeing and Kehoe (2003). This paper is most closely related to Yi (2003) and Bergoeing and Kehoe (2003) who provide a careful evaluation of the predictive success of standard modern trade models through calibrations. What marks out our work is that we consider a much longer, more eventful time span, covering periods of expanding and of contracting global trade, and a more heterogeneous sample of countries, including large economies from the "North" as well as the "South".[6] We also use the common structure of equilibrium trade flows across a broad class of prominent models to make more general statements about the ability of quantitative trade theory to explain the long-run evolution of world trade.

The empirical literature on international trade has ventured beyond the post-War era, estimating gravity equations by means of linear regressions of log bilateral trade flows on country-pair characteristics.[7] This literature has delivered important insights about the long-run determinants of bilateral trade flows, and the responsiveness of trade to changes in tariffs and transportation costs. However, even to the extent that these regressions are consistent with theory-based gravity equations, they do not provide a particularly strong test of the ability of the underlying models to explain the historical experience: the regression-based approach allows key parameters – such as the elasticity of trade flows with respect to frictions – to assume the value which best fits the trade data, and generally employs an array of dummy variables to mitigate omitted-variable concerns. By contrast, our calibrations place greater demands on the theory, restricting parameters to be consistent with external evidence in accordance with their interpretation, and requiring some of the standard "controls" of the gravity equation (such as the multilateral resistance terms) to obey the equilibrium relationships which theory implies. With these restrictions, we find that trade models struggle to reproduce the observed evolution of world trade.

[6]The historical period following World War II has been unique in many respects. Thus, the ability of different theoretical models to account for the empirical regularities of this period might not constitute the ultimate test of these theories. This subtle point is increasingly recognized, for example, in the literature on economic growth where there is a growing consensus towards adopting longer-run perspective. See for example, Lucas (2002), Acemoglu (2008) and Galor (2011).

[7]See Eichengreen and Irwin (1995), Irwin and Terviö (2002), Estevadeordal et al. (2003), Jacks et al. (2011) and Klasing and Milionis (2012), among others.

Several authors have pointed out that allowing for vertical specialization − whereby goods cross country borders several times in the production process − may allow trade models to provide a more accurate account of the recent growth in world trade. This was first suggested by Yi (2003) and explored in further calibration analyses by Bergoeing et al. (2004) and Bridgman (2008, 2012). While there is much evidence that vertical specialization is responsible for a growing portion of international trade, this phenomenon has only arisen in recent decades as global trade costs have reached historic lows.[8] Vertical specialization cannot justify why the volume of world trade was so much larger in the 19th century than basic theory would suggest, or why rising trade costs in the interwar period caused such a dramatic decline in global trade flows. For these reasons, models of vertical specialization are of limited use in understanding the long-run evolution of world trade. Based on our calibration results, we discuss several alternative extensions of our baseline model which have the potential to bring its predictions much closer to the long-run patterns observed in the data.

The remainder of this paper is structured as follows. Section 2 describes our data on international trade, GDP and trade frictions. Section 3 outlines the assumptions underpinning the gravity equation used in our calibrations. Section 4 describes and implements the calibration strategy. Section 5 explores extensions of the model and concludes.

2 The Long-Run Evolution of World Trade

Before attempting to assess the ability of trade theory to account for the long-run evolution of world trade, it is important to establish the basic facts which need to be explained. Hence, in what follows, we carefully track the evolution of world trade since 1870, a year often associated with the beginning of the first era of globalization.[9] We focus on the evolution of both the volume of international trade relative to the level of economic activity, as well as the broad bilateral and regional patterns observed in the data. Before we discuss those, however, we begin with a brief description of our data sources.

[8] See Irwin (2005)

[9] To be precise, there is some disagreement in the literature regarding the exact starting point of the first globalization era. As O'Rourke and Williamson (2002) and Jacks (2005) have demonstrated, its beginning could also be placed somewhat earlier, in the 1850s or even in the 1820s. Yet, like the majority of the existing contributions in the literature, we take 1870 as our starting point, primarily due to the lack of comprehensive trade statistics going back even further in time.

2.1 Data Sources for Trade Flows and GDP Levels

Our strategy to accurately and consistently assess the evolution of world trade going back to 1870 is to look at trade flows among a fixed set of countries as a share of their levels of GDP. In order to calculate such figures we draw on the international trade data set assembled by Barbieri et al. (2009) as part of the Correlates of War (COW) project. We choose this data set because of its comprehensive nature in terms of both country coverage – it includes data for 194 countries – as well as year coverage – it provides annual observations from 1870 to 2009 – compared to alternative data sources.[10] Moreover, its most recent version, version 3.0, which we employ, contains information on the values of exports and imports in current-price US dollars not only at the aggregate (country) level, but also at the bilateral level.

Given the nominal nature of our trade data, in order to for us to use them for a quantitative assessment of the importance of international trade relative to GDP, the data have to be combined with equivalent nominal GDP data.[11] For the post-1950 period such GDP data can be obtained from the Penn World Tables (PWT).[12] For the pre-1950 period, however, the only available GDP figures are those of Maddison (2001, 2003) which are purchasing-power-parity (PPP) adjusted and expressed in constant 1990 prices. To obtain the necessary nominal GDP figures for the pre-1950 period, we follow the approach of Klasing and Milionis (2012) who estimate those using the "short-cut" method.

In a nutshell, the "short-cut" method approach corresponds to estimating a relationship between real and nominal GDP for a period where both types of data are available in PWT and using this estimated relationship to make out-of-sample predictions for the levels of nominal GDP prior to 1950 based on the real GDP values provided by Maddison (2001, 2003). The rationale behind this method is that a fundamental structural relationship must exist between nominal and real GDP across countries and time. Moreover, despite its "short-cut" nature, as documented in Klasing and Milionis (2012), this approach produces estimates for nominal GDP, expressed in current US Dollars, which are reasonable and which greatly expand the coverage in terms of countries

[10] Specifically, the data-set includes 24160 observations for aggregate export and import levels as well as 1121911 observations on bilateral trade flows, which exceeds by far the coverage of other related data-sets.

[11] In what follows we use the terms 'real' and 'nominal,' as it is typically done in the international comparisons literature, to refer to measures which are in PPP-adjusted versus non-PPP-adjusted terms.

[12] Although, Penn World Tables does not report non-PPP adjusted current-price GDP values in US dollars directly, this can be easily obtained by combining the available PPP-adjusted current-price GDP values with the reported PPP-factors and the U.S. dollar exchange rates.

and years compared to any alternative data set currently available.[13]

When combining these estimated nominal GDP data for the pre-1950 period together with the post-1950 PWT data, as well as when pairing them with the trade data some care is necessary to ensure consistency. This is important as over this extended period of time that our analysis covers, several countries have experienced territorial and other political changes which need to be accounted for in the data. Hence, in Appendix A1 we report the details how we accommodate for such changes given available the information. Furthermore, given the long time span of our data series and the inclusion of two world wars in our period of interest, it is natural that our series includes gaps. To enhance the representation of our sample, we make an effort, to the extent possible to eliminate, such gaps by performing a series of interpolations. Particularly, as we explain in detail in Appendix A2, during non-wars years we assume a smooth evolution of nominal trade and GDP while during wars we employ an interpolation strategy that also accounts for the potential disruptive effects of wars on trade.

2.2 The Rise, Fall and Rebound of World Trade

We focus our analysis on bilateral trade flows among major economies for which we have a reasonably long time series of observations that go back to 1870. We also exclude from our series the very recent observations after 2005, which reflect developments related to the present financial crisis that have sparked a literature in their own right. In particular, we look at a fixed sample of countries for which data on trade flows, GDP and also tariffs − as we explain later on − are available for at least 100 out of the 124 no-war years in our sample period. Moreover, we impose the additional restriction that for each of these countries we observe for at least some year bilateral trade flows with almost all other sample countries. As it turns out, such a sample can be constructed with the inclusion of 28 countries.[14] For these countries we calculate the ratio of total bilateral imports relative total GDP. The evolution of this ratio is displayed in Figure 1.

What strikes out in the figure is the clearly visible course of the rise, the fall and the subsequent rebound that world trade has followed since 1870. Specifically, our trade share series exhibits an upward trend during the first

[13] Specifically, Klasing and Milionis show that for a small set of countries for which historical nominal GDP data from national accounts statistics are available, these are not substantially different from the corresponding estimated values.

[14] The 28 countries included in our sample are: Argentina, Australia, Brazil, Canada, Chile, China, Denmark, Egypt, France, Germany, Greece, India, Italy, Japan, Mexico, Netherlands, New Zealand, Norway, Philippines, Portugal, Spain, Sri Lanka, Sweden, Thailand, Turkey, United Kingdom, United States of America, Uruguay.

globalization period from 1870 to 1913, in line with the historical narrative, which came to an end with World War I. Between 1913 and 1945, we see a full-scale globalization retreat as a result of the two world wars and the turbulent interwar era marked by the Great Depression. Following the end of World War II, however, the world witnessed a marked expansion of trade at an accelerating pace, which allowed world trade to reclaim its antebellum peak in the mid 1970s. Moreover, although this upward swing seemed to level off during the 1980s, the 1990s witnessed the start of a new period of rapid trade expansion.

What should be clarified here is that the above-described patterns are by no means an artifact of the way in which our sample of countries is constructed or our world trade share is measured. As we show in Appendix A3, we can, alternatively, broaden our sample by loosening the membership criteria and including more countries. Yet, as it can been seen in Figure A1, the evolution of the share of world trade based on this larger sample is virtually identical to what is displayed in Figure 1. Similarly the observed patterns would not be different if our measure of bilateral imports to GDP would be replaced with one based on bilateral exports or the sum of bilateral exports and imports or the average of the two.[15] Finally, instead of concentrating on bilateral imports, we could use for each country in the sample the aggregate value of imports to build our trade share series. As we show again in Appendix A3, such a modified series displays the same trends as the one shown in Figure 1, albeit with an overall higher level.

2.3 The Nature and Geography of World Trade

Moving beyond the discussion of how the volume of world trade has evolved over this long period, it is worthwhile to look into the parallel evolution of the nature and the geographical distribution of trade. To investigate the former we split our sample into a group of developed and developing economies, which following the discipline's parlance we refer to as the "North" and the "South". As typical in the literature, we consider the "North" to consist of Western Europe, Western Offshoots and Japan.[16] Based on this classification in Figure 2 we display for each year the share of imports that correspond to trade between economies within the North (North-North trade) and within

[15] We prefer to have a series based on imports, as this is what we will be using for our calibration exercise.

[16] According to this criterion we obtain a 'North' subsample which in our case includes Australia, Canada, Denmark, France, Germany, Greece, Italy, Japan, Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, United Kingdom and United States of America. This leaves Argentina, Brazil, Chile, China, Egypt, India, Mexico, Philippines, Sri Lanka, Thailand, Turkey, Uruguay to represent the 'South.'

the South (South-South trade) block as well as trade across the two blocks (North-South trade).

Breaking down the evolution of world trade in this fashion we see that from 1870 to circa 1950 South-South trade has been following a clear upward trend. This trend, though, was completely reversed between 1950 and 1990, while since then South-South trade has experienced a resurgence. Mirroring the evolution of South-South trade, it can been seen that North-North trade had been on a downward trend since 1870 with signs of complete collapse during the two world wars. Yet, since 1950 North-North trade has experienced a rapid expansion, which was only been partially reversed since 1990. Finally, the share of North-South trade had been relatively steady during the first globalization era, only experiencing an upswing during war times, likely due to trade diversion. Following the end of both world wars the upswing of North-South trade was twice reversed, reaching on-all-time low in the beginning of the 1990s. Since then, the trend has changed again and become positive as with South-South trade.

Turning to the more standard breakdown of the different countries in our sample along continental boundaries we can construct a group of European, Asian, North and South American economies. This is displayed in Figure 3. Looking at the evolution of the share of world trade corresponding to each group of countries we can clearly identify distinct regional trends.[17] While the share of Europe has been constantly declining since 1870, with a tendency of stabilization after the 1950s, Asia's share of world trade has been characterized by a steady upward trend, with only some fluctuations during war years. On the other hand, looking at the shares of North and South America it is evident that both had been constantly rising from 1870 up until the 1950s. Following this long period of expansion, however, these trend have seen then diverged. In the case of North America there has been a stabilization around 30%, whereas South America has experienced a pronounced decline during the four decades from 1950 to 1990, from which only in the recent years it seems to be recovering.[18]

[17]We should stress here that the share of a given group is based on total trade by all countries within the group, not just external imports. So trade between France and Spain, for example, is included in Europe's share of trade.

[18]Although our sample is not sufficient to construct a representative African group, we should mention that based on available information, the relative share of Africa in world trade seems to mimic that of South America.

2.4 The Evolution of Trade Frictions

Having discussed how the volume as well as the patterns of international trade have evolved since the era of the first globalization, it may be worthwhile to have them briefly juxtaposed with the evolution of trade frictions over the same period. There are many distinct manifestations of trade frictions, particularly in the context of an analysis which, like ours, takes a long-run perspective. For the moment we limit our attention to just two key factors: tariff barriers and transportation costs.

Our main source for data on tariff barriers is Clemens and Williamson (2004) who assembled a data-set on average tariff rates covering the period from 1870 to 1998. We expand their data coverage to 2005 by using more recent information from the same source that Clemens and Williamson used and supplement the pre-1950 values with updated information collected by Schularick and Solomou (2011).[19] Finally, whenever possible we fill in gaps in the series through interpolations, as we discuss in more detail in Appendix A4. To capture the evolution of transportation costs over our period of interest we employ the global index constructed by Shah Mohammed and Williamson (2004) and originally covering the period 1870 to 1997, which we extend to the year 2005. To correctly reflect the actual magnitudes of transportation costs in ad valorem terms over the period of interest, we scale this index to match the average value of the global CIF/FOB ratio for the period 1966-1995 as reported in IMF's International Financial Statistics. For details regarding the extension of the index and the scaling procedure the reader is referred to Appendix A4.

Combining the tariff data for the countries in our sample with the global transportation costs figures we can construct a country-specific measure of trade frictions for the period of interest, whose weighted average across countries is depicted in Figure 4. As it is evident from the figure, over the period from 1870 to 2005 trade costs have been on a clear downward path in ad valorem terms. This downward movement has been primarily reflecting major improvements in transportation technologies as well as important developments toward trade liberalization. Yet, as it is also visible in the figure, this path has been twice interrupted by the two world wars and the retreat to protectionism that followed the Great Depression. Now in what follows we will be using this variation in trade frictions over time to see the extent to which it can explain the above-discussed evolution of world trade in terms of volume and patterns.

[19]We are grateful to Michael Clemens and Jeffrey Williamson as well as Moritz Schularick and Solomos Solomou for being kind enough to share their data with us.

3 Baseline Model

This section describes the assumptions required to derive the "gravity" model of trade flows employed in our calibrations. As is well known, these assumptions are satisfied in a number of prominent trade models which have been used widely for quantitative analysis.

3.1 Setup

There are C countries and two goods: a homogenous good and a differentiated good, both produced under perfect competition with constant returns to scale. The homogenous good is non-tradable, but all varieties of the differentiated good are tradable.

3.1.1 Endowments and Preferences

Consumers in each $c \in C$ have preferences represented by the utility function

$$u_{ct} = \left[\int_0^{I_t} x_{ct}(i)^{\frac{\varepsilon-1}{\varepsilon}} di \right]^{\theta \frac{\varepsilon}{\varepsilon-1}} (h_{ct})^{1-\theta} \qquad \theta \in (0,1), \varepsilon > 1, \tag{1}$$

where h_{ct} represents individual consumption in c at time t of the homogenous good, $x_{ct}(i)$ denotes consumption of variety i of the differentiated good, I_t is the total mass of differentiated varieties available for consumption and ε is the elasticity of substitution between them.[20]

At time t, the total mass of consumers in c is L_{ct} Each is endowed with 1 unit of effective labour, which is the only factor of production and supplied inelastically in the local labour market at the given wage rate w_{ct}. Consumers maximize equation (1) subject to

$$\int_0^{I_t} p_{ct}(i) x_{ct}(i) di + p_{ct,h} h_{ct} \leq w_{ct}, \tag{2}$$

where $p_{ct,h}$ is the price of the homogenous good in c at t and $p_{ct}(i)$ the price of variety i of the differentiated good.

3.1.2 Trade Frictions

The homogenous good is assumed to be completely non-tradable. Varieties of the differentiated good are tradable internationally, but there is a country-specific trade friction τ_{ct} which applies in an "iceberg" fashion and is borne by

[20] Since consumers are assumed to have Cobb-Douglas preferences over the differentiated and homogenous product, the implicit elasticity of substitution between traded and non-traded goods is 1. We relax this assumption as part of our calibrations in Section 5.

the importing country. Thus, country c only receives $1 - \tau_{ct}$ units for any 1 unit of a tradable variety purchased from $c' \neq c$.[21]

3.1.3 Technologies and Market Structure

The homogenous non-traded good is produced with labour $L_{ct,h}$ under constant returns to scale and sold in a competitive domestic market. For the differentiated traded goods, we assume that technologies and market structure are such that no two countries produce the same variety i in equilibrium. Formally, defining $\delta_{ct}(i)$ as an indicator function which takes value 1 if country c produces variety i in equilibrium at time t, we consider a class of models in which

$$\delta_{ct}(i) = \begin{cases} 1 & \text{if } i \in I_{ct} \\ 0 & \text{if } i \notin I_{ct} \end{cases} \qquad \begin{array}{cc} I_{ct} \neq \emptyset & \forall c \\ \bigcup_c I_{ct} = I_t & \\ I_{ct} \bigcap I_{c't} = \emptyset & \forall c \neq c' \end{array} . \tag{3}$$

While this assumption may appear restrictive, it is in fact satisfied in four prominent modern trade models which have been used extensively for the quantitative analysis of international trade flows − the Armington model, the Eaton-Kortum model (Eaton and Kortum, 2002), the standard "new" trade model (Krugman, 1979; Helpman, 1987), and the many-country Melitz model with Pareto-distributed firm productivities and without country-pair specific fixed trade costs (Melitz, 2003; Chaney, 2008). The first two of these models assume that the differentiated goods are produced with constant returns to scale and sold in perfectly competitive markets, while the latter two assume that they are produced with increasing returns to scale under monopolistic competition with free entry.

Labour markets are perfectly competitive, and in equilibrium

$$L_{ct,h} + \int_0^{I_t} L_{ct}(i) di = L_{ct}, \tag{4}$$

where $L_{ct}(i)$ is the labour input used by country c in industry i in equilibrium at time t.

3.2 Equilibrium Trade Flows

We are now ready to derive the equilibrium value of trade flows between any two countries and, hence, an expression for the world trade share.

[21] In our calibrations in Section 4, we employ empirical estimates of trade costs which vary only by country and year, since comprehensive data on the evolution of country pair-specific trade barriers is not available. For this reason, we assume throughout that trade frictions are country-specific, instead of country *pair*-specific as is common in the literature.

3.2.1 Bilateral Trade Flows

Define country c's income as $Y_{ct} \equiv W_{ct} L_{ct}$, and let $M_{c'ct}$ denote the value of imports by country c from $c' \neq c$ at time t. In equilibrium, $M_{c'ct} = \int_{I_{c't}} P_{ct}(i) x_{ct}(i) L_{ct} di$, which can be shown to equal

$$M_{cc't} = \theta \frac{(1-\tau_{ct})^{\sigma} (1-\tau_{c't})^{\sigma}}{P_{ct}^{-\sigma} P_{c't}^{-\sigma}} \frac{Y_{ct} Y_{c't}}{Y_t}, \tag{5}$$

where $Y_t = \sum_c Y_{ct}$ and P_{ct} is implicitly defined by

$$P_{ct} \equiv \frac{1}{1-\tau_{ct}} \left[\frac{Y_{ct}/Y_t}{P_{ct}^{-\sigma}} + \sum_{c' \neq c} (1-\tau_{c't})^{\sigma} \frac{Y_{c't}/Y_t}{P_{c't}^{-\sigma}} \right]^{-\frac{1}{\sigma}} \times$$

$$\left\{ \sum_c \frac{Y_{ct}}{Y_t} \frac{\left[\sum_c (1-\tau_{ct})^{\sigma} \frac{Y_{ct}/Y_t}{P_{ct}^{-\sigma}} \right]^{-1}}{\frac{Y_{ct}/Y_t}{P_{ct}^{-\sigma}} + \sum_{c' \neq c} (1-\tau_{c't})^{\sigma} \frac{Y_{c't}/Y_t}{P_{c't}^{-\sigma}}} \right\}^{\frac{1}{2}}, \tag{6}$$

and the parameter σ captures the elasticity of trade flows with respect to trade frictions. This elasticity parameter depends on the precise microfoundations of the model's equilibrium trade flows – an issue which will be explored in greater depth in Section 4.1. For now we simply note that, as long as (3) is satisfied, the functional form of the relationship between equilibrium trade flows, country incomes and trade frictions is independent of the assumptions about production technologies and market structures, up to the value of σ.

Equations (5) and (6) together define a standard gravity equation, in which bilateral trade flows depend positively on income in c and c', and negatively on a weighted measure of trade frictions in c and c', world income as well as two multilateral resistance terms. They are symmetric in c and c', implying bilaterally balanced trade between any pair of countries.

3.2.2 World Trade Share

Let total imports by country c equal $M_{ct} \equiv \sum_{c' \neq c} M_{c'ct}$. Then the equilibrium ratio of world imports to world GDP at any t is

$$\frac{\sum_c M_{ct}}{Y_t} = \theta \sum_{c=1}^{C} \sum_{c' \neq c} \frac{(1-\tau_{ct})^{\sigma} (1-\tau_{c't})^{\sigma}}{P_{ct}^{-\sigma} P_{c't}^{-\sigma}} \left(\frac{Y_{ct}^{\frac{1}{2}} Y_{c't}^{\frac{1}{2}}}{Y_t} \right)^2. \tag{7}$$

Together with (6), equation (7) implies that, for given parameters, the value of world trade flows is determined only by the distribution of countries' incomes and international trade frictions. Assuming that the model provides a reasonably good description of the drivers of international trade – and given suitable values for the parameters θ and σ –, data on the evolution of countries' GDPs and international trade costs should thus be sufficient for the model to replicate the long-run evolution of the ratio of world imports to world GDP. We test this formally in Section 4. To build some intuition prior to the calibration, the next subsection discusses two special cases of the model.

3.2.3 Two Special Cases

i) Free Trade

Suppose first there are no barriers to international trade: $\tau_{ct} = 0 \ \forall \ c$, $\theta = 1$. Then (7) reduces to

$$\frac{\sum_c M_{ct}}{Y_t} = 1 - \sum_c \left(\frac{Y_{ct}}{Y_t}\right)^2. \tag{8}$$

Equation (8) shows that even when trade frictions are assumed away, the model-implied world trade share may change over time if there are changes in the distribution of countries' incomes. Specifically, if the variance of countries' incomes declines, the value of world imports should rise relative to world GDP.[22]

ii) Symmetric Countries

Let us now abstract from heterogeneity across countries by assuming complete symmetry: $L_{ct} = \frac{L_t}{C}$ and $\tau_{ct} = \tau_t \ \forall \ c$. Now (7) simplifies to

$$\frac{\sum_c M_{ct}}{Y_t} = \theta \frac{(1-\tau_t)^{\sigma}(C-1)}{1+(1-\tau_t)^{\sigma}(C-1)}. \tag{9}$$

In this case, only a decline in trade frictions can raise the ratio of world imports to world GDP. At the margin, declining trade frictions exert a diminishing effect on the global trade share which – for a large number of countries, C – approaches θ as τ_t falls to zero.

4 Calibrations

In the following, we show that the canonical gravity equation, described in Section 3, struggles to replicate the long-run evolution of world trade for reasonable parameter calibrations. While it has been known since Yi (2003) and Bergoeing and Kehoe (2003) that calibrated trade models underpredict the growth in world trade since World War II, we are the first to evaluate formally

[22] This feature of trade models with an implicit gravity equation was first pointed out by Helpman (1987).

the extent to which standard trade models can explain the evolution of world trade during most of the modern globalization period. Our findings cast the failings of the theory in a new light, showing that its difficulties in predicting recent trade growth reflect a broader inability to reproduce large observed swings in the world trade share.

4.1 Basic Data and Parameters

4.1.1 Data

To be able to predict trading patterns, the model described in Section 3 requires data on the evolution of countries' incomes and international trade frictions. For the former, we use countries' nominal GDP figures. For the latter, we combine data on global transportation costs with measures of countries' average tariff rates to obtain a country-time-varying measure of trade barriers. The data sources for these are as described in Section 2.

Results reported below are based exclusively on our restricted country sample, since calibrations based on our larger sample yield virtually identical results.[23] We use only trade flows between our sample countries, thereby assuming that trade flows between the sample and all non-sample countries are economically insignificant.[24] Whenever our model predicts a trade flow between any two sample countries which is missing in the data − i.e. in any country-pair year in which data on GDPs and trade barriers is available, but import data is not − we replace the model prediction with a missing value.

4.1.2 Parameter Choice

The baseline calibration requires us to choose only two key parameters, θ and σ. To determine an appropriate value for θ, the size of the traded sector, we adjust the parameter to let the theory-predicted world trade share match the observed average world trade share for the period 1870-2005. That way, we ensure that the model will predict the correct volume of trade on average, and we can focus on its ability to match the observed evolution of global trade flows.

The choice of value for σ is crucial because σ governs the impact of changes in trade frictions on the volume of trade, and tricky because the appropriate choice hinges on the precise interpretation of this parameter which, in turn, depends on the assumptions about technology and market structure underpin-

[23] These calibration results are available from the authors upon request.

[24] Our small sample of countries already accounts for 80-92% of world GDP over the whole period, making this a rather weak assumption.

ning our gravity equation. To determine an appropriate baseline value of σ, we consider four prominent trade models which are consistent with the functional form of the gravity equation in (5) and (6) as reference cases:

1. *Armington/Helpman-Krugman model* – If equations (5) and (6) were derived from an Armington or the basic Helpman-Krugman model, σ would depend on the elasticity of substitution between differentiated varieties: $\sigma = \varepsilon - 1$. As Yi (2003) points out, industry-level studies of substitution elasticities in international trade generally find values of ε in the range 2-4. Imposing an intermediate value of $\varepsilon = 3$ would imply $\sigma = 2$.

2. *Melitz model* – If equations (5) and (6) were derived Chaney's (2009) multi-country version of the Melitz model, σ would depend on ε as well as the Pareto parameter of firms' productivity distribution (which we shall define as as φ): $\sigma = \varphi \frac{\varepsilon}{\varepsilon - 1}$. Bernard et al. (2003) suggest that $\varphi \approx 3.3$, based on moments of U.S. firm-level data, yielding $\sigma \approx 5$.

3. *Eaton-Kortum model* – If equations (5) and (6) were derived from the Eaton-Kortum model, σ would equal the shape Parameter of the Fréchet distribution of industry productivities and, based on the original calibration of Eaton and Kortum (2002), $\sigma \approx 8.3$.

In our baseline calibration, we set $\sigma = 4$, inbetween the calibrated Armington and Melitz models. However, as the above makes clear, any value of σ in the range 1-9 could be justified with reference to the precedent set by an earlier calibration of one of the four models above. Therefore, we also present results for more extreme values of the trade elasticity.

All parameter values and data sources are summarized in Table 1.

4.2 Baseline Calibration

Figure 5 contrasts observed changes in the ratio of world imports to world GDP between 1870 and 2005 with the prediction of our calibration based on the model from Section 3. The figure illustrates both the quantitative and qualitative shortcomings of our baseline gravity equation.

Quantitatively, the model cannot reproduce the magnitude of variation in the trade share over the period. It predicts an increase in the world trade share of 6.3 percentage points between 1870 and 2005 (from 3.1% to 9.4%), while the actual world trade share rose by 9.5 percentage points (from 3.7% to 13.2%). Qualitatively, the model fails to capture the slump in world imports during the interwar period and the extent and shape of their resurgence since

1950. Indeed, our calibration confirms Yi's (2003) observation that standard trade models cannot replicate the magnitude and "exponential" nature of the post-War rise in the world trade share. However, it shows that this partly reflects the canonical model's corresponding obliviousness to the period of deglobalization between 1914 and 1945, from which world trade was recovering in the first three decades after World War II.

Figures 6 and 7 explore two further dimensions of our calibration exercise: the model-predicted bilateral trading patterns between rich and poor countries, and the implied geographical split of world imports. Figure 6 graphs the shares of North-North, North-South and South-South trade in total world trade, with solid lines representing the data and dashed lines the calibration outcome.[25] Although the model under-predicts the share of North-North trade in total trade, it reproduces the relative *evolution* of the trade flows between and within the two broad regions remarkably well. In particular, the calibration results − like the data − show a decline in North-North trade prior to World War I, and a surge to new heights of this trade in the post-War era. They also correctly replicate the temporary rise in South-South trade during World War II.

Figure 7 provides some further insight by displaying the share of four major geographical regions in total world imports, both as found in the data and as produced by our calibration. As can be seen from the figure, the model under-predicts the relative size of North-North trade flows largely because it assigns too small a share of world imports to Europe, while assigning too large a share to Asia. As with the bilateral flows, the calibration proves better at tracing the evolution of regional trade shares than their relative sizes. The only exception is Asia, whose share of world trade rose steadily during 1870-1930 − in contrast to the much higher, and steadily *declining* share the model indicates.

As discussed in Section 4.2.1, we let the size of the traded sector be determined by the model. Our baseline calibration delivers a value of .11 for the parameter θ. To determine whether this number is reasonable, it would be ideal to compare it with the long-run share of tradable output in world GDP. Since there is no other practicable method for separating the value of tradable output from non-tradable output for our sample, we follow a long-standing tradition and adopt manufacturing value added as a measure of the production of tradables. Based on data from the World Input-Output Database,[26] it turns out that the average share of manufacturing production in world GDP for the period 1870-2005 was .12 − virtually identical with the model-determined

[25] For simplicity, we define the "North" as all European and North American countries in our sample, plus Australia and Japan. The "South" comprises all other sample countries.

[26] For information on this source can be found at: www.wiod.org

tradables share. This is reassuring as it suggests that our measure of trade frictions combined with a reasonable parameterization of the non-traded sector delivers the correct volume of trade *on average*.

Figure 5 shows that the theory-predicted world trade share rises steadily over the whole period, with only two small slumps during 1914-1918, and 1940-1950. This is the result of the long-run decline in trade barriers over the period, as well as their diminishing marginal impact on world imports at ever-lower magnitudes. It thus seems reasonable to suppose that a larger σ − and hence a higher responsiveness of trade flows to changes in trade frictions − could improve our calibration results. Regression estimates of the tariff-elasticity of trade flows in the post-War period, based on gravity-style regressions, generally deliver values in the range of 5-10, significantly larger than the value we assign to σ,[27] and more in line with a calibrated Eaton-Kortum model. We therefore re-run our exercise setting $\sigma = 7$ and compare the outcome with the baseline calibration in Table 2.

We use three measures to assess the goodness-of-fit of the world trade share predicted by our model simulations: the Root MSE of the prediction, the share of the overall growth in world trade explained, and the implied average size of the traded sector. Raising the value of σ increases the share of overall trade growth which the model can replicate, and it leaves the implied value of θ little changed. However, the Root MSE rises as the model now undepredicts the ratio of world imports to GDP prior to 1914 more severely, while overpredicting the speed of the post-1945 recovery in global trade significantly. Conversely, lowering the value of σ to 1.5 improves the Root MSE somewhat but reduces the explained share of world trade growth. Overall, the comparison in Table 2 indicates that higher values of σ could allow the model to replicate the magnitude of variation in the world trade share, but only at the cost of worsening the overall fit of the model.

5 Extensions and Discussion

In this section we quantitatively explore two generalizations of our baseline model. We also discuss alternative extensions which may bring the model-implied evolution of the world trade share closer to the data.

[27]See Anderson and van Wincoop (2004) for a survey of gravity-based estimates of the trade elasticity.

5.1 Generalized Preferences

Suppose preferences are characterized by

$$u_{ct}=\left\{\theta_c^{\frac{1}{\eta}}\left[\int_0^{I_t}x_{ct}(i)^{\frac{\varepsilon-1}{\varepsilon}}di\right]^{\frac{\varepsilon}{\varepsilon-1}\frac{\eta-1}{\eta}}+(1-\theta_c)^{\frac{1}{\eta}}\left(x_{ct,h}\right)^{\frac{\eta-1}{\eta}}\right\}^{\frac{\eta}{\eta-1}}\quad \begin{array}{c}\theta_c\in(0,1)\ \forall c,\\ \varepsilon>1\end{array}. \tag{10}$$

Equation (10) represents a generalization of our baseline utility function: we allow for the possibility that countries may differ in their preferences over traded and non-traded goods, and that the elasticity of substitution between them − denoted η − may not equal unity. If we impose $\theta_c = \theta\ \forall\ c$, and set $\eta = 1$, we once again obtain equation (1).

Starting from equation (10), we arrive at a more complicated expression for the ratio of world imports to GDP. Equations (5) and (6) are replaced by

$$M_{cc't}=\Theta_{ct}\Theta_{c't}\frac{(1-\tau_{ct})^{\sigma}(1-\tau_{c't})^{\sigma}}{P_{ct}^{-\sigma}P_{c't}^{-\sigma}}\frac{Y_{ct}Y_{c't}}{Y_t}, \tag{11}$$

and

$$P_{ct}\equiv\frac{1}{1-\tau_{ct}}\left[\frac{\Theta_{ct}Y_{ct}/Y_t}{P_{ct}^{-\sigma}}+\sum_{c'\neq c}(1-\tau_{c't})^{\sigma}\frac{\Theta_{c't}Y_{c't}/Y_t}{P_{c't}^{-\sigma}}\right]^{-\frac{1}{\sigma}}\times$$

$$\left\{\sum_c\frac{\Theta_{ct}Y_{ct}}{Y_t}\frac{\left[\sum_c(1-\tau_{ct})^{\sigma}\frac{\Theta_{ct}Y_{ct}/Y_t}{P_{ct}^{-\sigma}}\right]^{-1}}{\frac{\Theta_{ct}Y_{ct}/Y_t}{P_{ct}^{-\sigma}}+\sum_{c'\neq c}(1-\tau_{c't})^{\sigma}\frac{\Theta_{c't}Y_{c't}/Y_t}{P_{c't}^{-\sigma}}}\right\}^{\frac{1}{2}}, \tag{12}$$

where Θ_{ct} is a non-linear function of θ_c, τ_{ct}, P_{ct} and the parameters ε and η. For $\eta = 1$, $\Theta_{ct} = \theta_c$ for all t. If $\eta > 1$, then the stronger a country's preference for tradable goods (θ_c), the smaller its barriers to trade (τ_{ct}) and the smaller its multilateral resistance (P_{ct}), the more it spends on tradables.

The functional form of Θ_{ct}, together with equations (11) and (12), provide us with a more general theory-driven mapping from countries' incomes and trade frictions into trading volumes. The next two subsections test whether this generalization can reconcile trade theory with the observed long-run evolution of world trade. Table 3 summarizes the calibration results, comparing them to our baseline assumptions.

5.2 Heterogeneous Tastes for Tradable Goods

We begin by setting $\eta = 1$, as in our baseline calibration, but allowing tradable-good preferences to vary across countries. To determine θ_c, we now let the model choose a value *for each country* which ensures that the country's average share of imports in GDP during 1870-2005 is replicated. All other features of our baseline calibration are unchanged.

As can be seen from the second column of Table 3, this change has a minor impact on the model-predicted evolution of the world trade share. If anything, the calibration results are somewhat worse than our baseline predictions. A notable improvement only occurs in the model's ability to reproduce bilateral trading patterns between the North and the South. Figure 8 shows a nearly perfect fit between the shares of bilateral flows found in the data, and the model-implied shares. On some level, this finding is not surprising. With the size of the traded sector constrained to be equal across countries, the model already performed well at tracing the evolution of the relative size of North-North, North-South and South-South trade, but underpredicted the relative significance of the North. Letting θ_c vary across countries provides enough degrees of freedom to overcome this problem.

However, the finding is remarkable on another level. It is evidently possible for the very simple model described above to account satisfactorily for the broad bilateral trading patterns during most years of our sample period. Yet, at the same time, it provides a rather poor account of changes in overall trading volumes. This suggests that whatever prevents the model from correctly predicting long-run trade *growth* could be orthogonal to the model's facility for predicting trading *patterns*.

5.3 Traded/Non-Traded Substitutability

We now turn to the elasticity of substitution between traded and non-traded goods. For ease of comparison with our baseline results, we once again impose a common θ for all countries.

To proceed, we require an appropriate value of η. Using equations (11), (12) and the function Θ_{ct}, it turns out that

$$\frac{\partial \ln \Theta_{ct}}{\partial \ln Y_{ct}/Y_t} \approx -\frac{\eta - 1}{\sigma + \eta - 1}$$

in a world with negligible trade frictions. Approximating Θ_{ct} with the trade-openness ratio, we can estimate the left-hand side of this equation using data on openness and incomes from the Penn World Tables 7.0 for our sample countries in the period 2001-05, during which global trade barriers had reached their lowest level. This yields a value of -0.14, which reflects the fact that larger countries $-$ in terms of their share of world GDP $-$ are generally less open than their smaller counterparts. Under the assumption that $\sigma = 4$, our estimate implies $\eta \approx 1.5$, i.e. a moderate degree of substitutability between traded and non-traded goods.

Setting $\eta = 1.5$ affects our calibrations in two ways. First, in addition to importing less than smaller countries as a share of tradables spending, larger countries now also spend a smaller portion of their incomes on tradable goods. However, the size of this effect is modest. Second, and more importantly, declines in trade frictions now have a bigger effect on imports because they encourage more tradables spending. Values of η greater than 1 thus raise the elasticity of trade to trade frictions.

Accordingly, the results reported in column 3 of Table 3 resemble those of from our baseline calibration with higher values of σ: while the model predicts a larger share of the overall growth in world trade, the Root MSE worsens, as the gaps between the model-implied and empirical world trade shares before World War I and in the aftermath of World War II are exaggerated. Just as with different values of σ, adjusting η appears to be of limited use in bringing the theory's predictions in line with empirical evidence.

5.4 Other Extensions

Even with the generalizations described in Section 5.1, our model remains highly stylized. There are a number of empirically appealing extensions which, if incorporated in our framework, would give the model a better chance to match the data. Two natural candidates are a relaxation of the assumption of homothetic preferences, and the incorporation of heterogeneous traded goods with different substitution elasticities.

Considering non-homothetic preferences seems highly appropriate in the analysis of long-run changes in expenditure patterns. Moreover, such preferences have been used successfully in other work to capture features of global trading patterns. For example, Markusen (1986) shows that, if poorer countries have a comparative advantage in inferior goods, the ratio of North-North to North-South trade implied by a standard model more closely resembles its empirical counterpart. Suppose, following Markusen, that spending on the goods produced by countries in the global South declines as world per-capita income rises. This would introduce a new relevant time series into our calibrations: the evolution of the distribution of real per-capita incomes across countries. More specifically, our model would predict a fall in the global trade share during periods in which the gap between Northern and Southern per-capita incomes increases, and a rise when Southern incomes catch up with the North. Of course, the empirical success of this channel strongly hinges on the appropriate extent of preference non-homotheticity, and on the observed evolution of real per-capita GDPs.

Allowing for different substitution elasticities between different types of traded goods also has considerable empirical and theoretical appeal. Suppose traded goods differ in their substitutability, and let there be a uniform fixed cost of importing a particular type of good. In that case, for a given variable trade cost, countries would import a range of low-elasticity goods but rely on their domestic production of high-elasticity goods. Lower variable trade costs now impact on the volume of trade through the extensive as well as the intensive margin: as variable trade costs fall, countries will expand both the range of sectors in which they import as well as the import volumes in any given sector. If the extensive margin dominates, this could result in "exponential" growth of trade relative to output, as observed in certain periods of our data. Moreover, the notion that falling trade barriers allowed for more, and increasingly substitutable commodities to be traded accords with narrative historical accounts of the progress of globalization.[28]

6 Conclusion

Although several papers have tried to assess the ability of modern trade models to account for the evolution of world trade over the post-War era, none has so far considered a longer time span. The present paper seeks to fill this gap, by carefully investigating how trade theory can account for the trade patterns observed in the long run.

To perform this investigation we consolidate existing sources on bilateral trade flows, GDP levels, tariffs and transportation costs for the world's major economies, constructing a comprehensive new data set for the period 1870-2005. Using this data we evaluate how well a calibrated gravity equation which could be derived from broad class of canonical trade models can predict the evolution of world trade and the observed trade patterns during that period.

The results we obtain are mixed. Our baseline calibration captures the variation in the broad bilateral patterns and in the the geographical distribution of world trade over time. However, it fails to a large degree to match the magnitudes of variation in the share of world trade over this long horizon. We extend the canonical gravity equation to allow for the possibility that countries may differ in their preferences over traded and non-traded goods, and that the elasticity of substitution among traded goods may differ from unity. Interestingly, this allows it to perform better in replicating the cross-sectional trade patterns, but not the time-series variation in world trade. Our exercise

[28]For example, O'Rourke and Williamson (2002) argue that globalization progressed in "stages", with more directly competing goods only traded once trade costs had declined sufficiently.

suggests that whatever prevents the model from correctly predicting long-run trade *growth* could be orthogonal to the model's facility for predicting trading *patterns*.

References

Acemoglu, D., 2008. Introduction to Modern Economic Growth. Princeton University Press, Princeton, NJ.

Anderson, J. E., March 1979. A theoretical foundation for the gravity equation. The American Economic Review 69 (1), 106–16.

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

Baier, S. L., Bergstrand, J. H., February 2001. The growth of world trade: Tariffs, transport costs, and income similarity. Journal of International Economics 53 (1), 1–27.

Bajona, C., 05 2004. Specific factors, learning, and the dynamics of trade. International Economic Review 45 (2), 499–521.

Barbieri, K., Keshk, O., 2012. Correlates of war project trade data set codebook, version 3.0. Online: http://correlatesofwar.org.

Barbieri, K., Keshk, O., Pollins, B., 2009. Trading data: Evaluating our assumptions and coding rules. Conflict Management and Peace Science 26 (5), 471–491.

Bergoeing, R., Kehoe, T. J., October 2003. Trade theory and trade facts, research Department Staff Report 284,Federal Reserve Bank of Minneapolis.

Bergoeing, R., Kehoe, T. J., Strauss-Kahn, V., Yi, K.-M., May 2004. Why is manufacturing trade rising even as manufacturing output is falling? The American Economic Review 94 (2), 134–138.

Bergstrand, J. H., August 1985. The gravity equation in international trade: Some microeconomic foundations and empirical evidence. The Review of Economics and Statistics 67 (3), 474–81.

Bernard, A. B., Eaton, J., Jensen, J. B., Kortum, S., September 2003. Plants and productivity in international trade. The American Economic Review 93 (4), 1268–1290.

Bridgman, B., October 2008. Energy prices and the expansion of world trade. Review of Economic Dynamics 11 (4), 904–916.

Bridgman, B., 2012. The rise of vertical specialization trade. Journal of International Economics 86 (1), 133–140.

Chaney, T., September 2008. Distorted gravity: The intensive and extensive margins of international trade. The American Economic Review 98 (4), 1707–21.

Clemens, M. A., Williamson, J. G., March 2004. Why did the tariff-growth correlation change after 1950? Journal of Economic Growth 9 (1), 5–46.

Cuñat, A., Maffezzoli, M., 04 2007. Can comparative advantage explain the growth of us trade? Economic Journal 117 (520), 583–602.

Eaton, J., Kortum, S., September 2002. Technology, geography, and trade. Econometrica 70 (5), 1741–1779.

Eichengreen, B., Irwin, D. A., February 1995. Trade blocs, currency blocs and the reorientation of world trade in the 1930s. Journal of International Economics 38 (1-2), 1–24.

Estevadeordal, A., Frantz, B., Taylor, A. M., May 2003. The rise and fall of world trade, 1870-1939. The Quarterly Journal of Economics 118 (2), 359–407.

Galor, O., 2011. Unified Growth Theory. Princeton University Press, Princenton, NJ.

Helpman, E., August 1981. International trade in the presence of product differentiation, economies of scale and monopolistic competition : A chamberlin-heckscher-ohlin approach. Journal of International Economics 11 (3), 305–340.

Helpman, E., March 1987. Imperfect competition and international trade: Evidence from fourteen industrial countries. Journal of the Japanese and International Economies 1 (1), 62–81.

Heston, A., Summers, R., Aten, B., May 2011. Penn world table version 7.0. Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania.

Hummels, D., July 1999. Have international transportation costs declined?, mimeo, University of Chicago.

Hummels, D., Summer 2007. Transportation costs and international trade in the second era of globalization. Journal of Economic Perspectives 21 (3), 131–154.

Hummels, D., Levinsohn, J., August 1995. Monopolistic competition and international trade: Reconsidering the evidence. The Quarterly Journal of Economics 110 (3), 799–836.

Irwin, D. A., 2005. Trade and globalization. In: Weinstein, M. M. (Ed.), Globalization: What's New? Columbia University Press, New York, NY.

Irwin, D. A., Terviö, M., October 2002. Does trade raise income?: Evidence from the twentieth century. Journal of International Economics 58 (1), 1–18.

Jacks, D. S., July 2005. Intra- and international commodity market integration in the atlantic economy, 1800-1913. Explorations in Economic History 42 (3), 381–413.

Jacks, D. S., Meissner, C. M., Novy, D., March 2011. Trade booms, trade busts, and trade costs. Journal of International Economics 83 (2), 185–201.

Klasing, M., Milionis, P., June 2012. Reassessing the evolution of world trade, 1870-1949, research Memorandum, Groningen Growth and Development Centre, University of Groningen.

Krugman, P. R., November 1979. Increasing returns, monopolistic competition, and international trade. Journal of International Economics 9 (4), 469–479.

Lancaster, K., May 1980. Intra-industry trade under perfect monopolistic competition. Journal of International Economics 10 (2), 151–175.

Lucas, R. E. J., 2002. Lectures on Economic Growth. Harvard University Press, Cambridge, MA.

Maddison, A., 2001. The World Economy: Volume 1 - A Millennial Perspective. Development Studies Center, OECD, Paris.

Maddison, A., 2003. The World Economy: Volume 2 - Historical Statistics. Development Studies Center, OECD, Paris.

Melitz, M. J., November 2003. The impact of trade on intra-industry reallocations and aggregate industry productivity. Econometrica 71 (6), 1695–1725.

O'Rourke, K. H., Williamson, J. G., April 2002. When did globalisation begin? European Review of Economic History 6 (1), 23–50.

Schularick, M., Solomou, S., March 2011. Tariffs and economic growth in the first era of globalization. Journal of Economic Growth 16 (1), 33–70.

Shah Mohammed, S. I., Williamson, J. G., 2004. Freight rates and productivity gains in british tramp shipping 1869-1950. Explorations in Economic History 41, 172–203.

Timmer, M. P., April 2012. The world input-output database (wiod): Contents, sources and methods, wIOD Working Paper.

Yi, K.-M., February 2003. Can vertical specialization explain the growth of world trade? Journal of Political Economy 111 (1), 52–102.

Zymek, R., 2012. Factor proportions and the growth of world trade, mimeo, University of Edinburgh.

Appendix

A1: Adjustments of Trade and GDP to Ensure Consistency

Despite the long time-span of our trade and GDP data, for the majority of the countries no real adjustment in the data is necessary. This is because both the PWT as well as the Maddison GDP data are based on contemporary rather than historical borders, which makes them directly comparable. In some cases, though, either because of complications in the nature of data or in the way the data was reported some adjustments were needed. We discuss each of this case analytically below in order to satisfy the curious reader.

1. Argentina, Egypt, Greece, Turkey: When merging our estimated nominal GDP series for the pre-1950 period with the corresponding one from PWT that starts in 1950, we noticed that in the case of these four countries the former was significantly -at least 30%- lower than the latter. This gap was due to episodes of high inflation during the 1940s, which could not have been captured by our estimation technique. To avoid an artificial downward jump in the trade shares of those countries in 1950 as a result of this problem, we drop in the pre-1950 estimated GDP series for Argentina the years 1949-1950, for Egypt the years 1945-1950, for Greece the years 1941-1949, and for Turkey the years 1942-1949.

2. China: The COW data-set reports an uninterrupted series for China in all years as well as a separate one for Taiwan starting in 1949. Moreover, as stated in the codebook, the Chinese trade figures from 1998 onward include both Hong Kong and Macao, for which, though, no trade figures are reported until 1997. Given that, we pair the trade figures for China with GDP levels that include Taiwan until 1948, exclude Taiwan after 1948 and include for Hong Kong and Macao from 1998 onward. The Maddison data include separate GDP series for Mainland China, Taiwan and Hong Kong that go back to 1870, which hence display no jumps when Taiwan was separated or when Hong Kong was annexed. Moreover, we should note that since 1950 we use the PWT version 2 series for China which is considered more reliable than the officially published data.

3. Germany: During the period from 1954 to 1989 COW reports separate series for East and West Germany. We combine those series to one after excluding bilateral trade between the two countries. The resulting figures for this period are then combined with the trade series provided for the whole of Germany until 1949 and from 1990 and on. In terms of GDP figures while Maddison provides a continuous GDP series for the whole of Germany going

back to 1870, the PWT 7.0 German GDP series starts only in 1970. To fill in the missing years, we use information from earlier versions of PWT. Specifically, we use the information from PWT 5.6, which is the last version that contains two separate GDP series for East and West Germany and which we add up to construct a series for the whole of Germany. This allows us to have a GDP series that can be pair with the available trade series discussed above.

4. Norway: In the case of Norway, the trade figures reported by COW during the years of World War I jump upwards by a factor of 10-15 which lead to the share of imports in GDP rising to 1500% for those years. Given the lack of any historical record justifying this jump, we drop the original observations during the years 1914-1918.

5. Russia/USSR: In the case of Russia, the USSR and other former Soviet countries the trade data reported by COW include one uninterrupted series labelled "Russia" starting in 1870 and separate series for the USSR successor countries from 1992 onwards. Although this is not clearly stated in the dataset codebook, we presume that the "Russia" series prior to 1992 corresponds to the Russian Empire until 1917 and the Soviet Union afterwards. Moreover, for the three Baltic countries, Estonia, Latvia and Lithuania, separate series are reported between 1919 and 1939, during the time when these countries were independent. Given that prior to 1973, the only GDP data available are for the whole USSR including the Baltics, we construct a corresponding Russia/USSR trade series. This corresponds to the available "Russia" series plus the trade figures for Estonia, Latvia and Lithuania between 1919 and 1939 and minus the total bilateral trade flows between these four countries. This series is paired with a matching Russia/USSR GDP series, which corresponds to the Soviet Union until 1992 and to present day Russia since then.

6.Yugoslavia: Yugoslavia split up in 1992. Yet, COW reports one continuous series labelled "Yugoslavia" starting from 1878 as well as additional series for Bosnia and Herzegovina, Croatia, FYROM and Slovenia starting from 1993. Hence we can infer -although this is not explicitly stated in the codebook - that starting from 1993 the "Yugoslavia" series corresponds to just Serbia, Montenegro and Kosovo. Using this reasoning, we can match the trade series with the corresponding GDP series from PWT and Maddison. PWT reports GDP levels for the successor states of former Yugoslavia starting from 1990. Given, the structure of our trade series, we add up the information from the corresponding countries in each year to produce a matching nominal GDP figure for the whole of Yugoslavia until 1992 and for Serbia, Montenegro and Kosovo starting from 1994.

Appendix A2: Data Interpolations

Both the bilateral and the aggregate data on trade flows reported by COW naturally include some gaps due to missing observations. We try to fill in those gaps through interpolations following a two step procedure. We first fill in gaps in the series during non-war years by interpolate the missing observations assuming that the trade flow evolved based on a constant annual growth rate. We then proceed to interpolated missing observations during war year. Given that the pre- and post-war trends may not be informative for the wartime evolution of bilateral trade, though, we perform these interpolations differently. For each pair of countries we identify the last available observation before the war and the first available one after the war.[29] For these years we calculate the share of the bilateral flow in the total exports of the exporting and the total imports of the importing country. We then interpolate these shares for the missing years assuming a constant growth rate. Following that, we multiply these interpolated shares with the respective aggregate exports and imports of each respective country available in the aggregate data-set. This provides us with two alternative estimates for the bilateral trade flow. Finally, in order to reduce any potential mismeasurement, we take as our estimated figure for the missing years the average of the two alternative estimates.[30]

To give an example of how we perform this war-time interpolations consider the trade flow from the United States to Britain for which no observations are available from 1939 to 1947. Our approach is to first calculate the share of American exports sent to Britain in total US exports as well as the share of British imports coming from the United States in total UK imports both in 1938 and in 1948. We then interpolate these shares and multiply the interpolated values for the former with the total value of US exports and for the latter with the total value of British imports for each year from 1939 to 1947, both of which are available. The result will be two alternative estimates for the value of the missing trade flow from the United States to Britain, with our employed estimate being the average of the two.

The advantage of this approach is that it allows us to capture the disruptive effects that both World Wars had on bilateral trade flows utilizing the information available on aggregate trade flows. It is based, however, on the assumption that the relative importance of the different trading partners of each country evolved smoothly during wars. Hence, the employed technique does not allow us to capture the effect of war-time trade diversion. As this is not

[29] If there are some observations available during the war years, we use the same approach to interpolate any gaps between them.

[30] In order to be able to perform these interpolations, we have converted some trade flow values recorded as zero to one US dollar, in order to permit the calculation of growth rates.

the focus of our current exercise, though, we do not consider this underlying assumption too restrictive.[31]

The above discussed interpolations allows us, as we explain further below, to have a relatively balanced sample of countries based on which we can conduct our analysis. Yet, we should emphasize here that in terms of the sample composition the interpolated data constitute only a small share of the available observations. First of all, the aggregate series contains very few gaps, only 29 missing observations out of a total of 114150. In the bilateral series the number of missing observations is naturally large, particularly during the years of World War I and II. However, allowing for linear interpolation during non-war years leads to an addition of only 17747 observations, which correspond to 1.58% of the original sample size of 1121911. Similarly, allowing for war-time trade interpolations provides us with another 20829 observations, which are also only 1.86% of the original sample size.

Finally, we also perform a set of interpolations based on a constant annual growth rate in order to fill in gaps observed in the Maddison's GDP data series following Klasing and Milionis (2012). Given that we only use the Maddison data to make predictions for nominal GDP, we do think that these interpolation place an important in the context of our analysis.

Appendix A3: Alternative Measures of World Trade Share

In this part of the appendix we demonstrate that the evolution of world trade and the patterns discussed in Section 2 are not simply driven by the way our baseline sample or by the choice of data used. Sticking to the same period from 1870 to 2005, we can, first of all, construct an alternative sample of countries with broader membership. Specifically, we can consider the inclusion of all countries for which data on trade flows, GDP and tariffs are available for a least 40 out of the 124 no-war years, which corresponds to just 30% of the years. From those we can add into the sample all countries for which bilateral trade flows are available with at least half of the trading partners of the United Kingdom, the country with the high number of trading partners in the data.[32] Loosening both criteria this way we end up with a sample of 44 countries.[33]

[31] Another advantage of our approach is that the interpolated bilateral flows will be the sum of all bilateral imports, while an alternative interpolation strategy might lead to the sum of bilateral exports or imports of a country to exceed the corresponding aggregate value, particularly if trade collapsed during the war.

[32] By setting the cut-off to 50%, we can reasonably argue that not many trading patterns are missing and the pairings not observed to be unimportant or non-existing.

[33] These 44 countries are: Argentina, Australia, Brazil, Canada, Chile, China, Cyprus, Denmark, Egypt, France, Germany, Ghana, Greece, Guyana, India, Italy, Japan, Kenya, Malawi, Malaysia, Mexico, Myanmar, The Netherlands, New Zealand, Nigeria, Norway,

The resulting picture for the evolution of world trade is displayed in Figure A1 which hardly differs from what can been seen in Figure 1. Despite the inclusion of more countries the series is barely less volatile, as the additions corresponds to mostly small economies.

Alternatively, we can consider having as the numerator of our world trade share not just the total bilateral imports among all the sample countries, but their total aggregate imports instead. This would also include in the calculation imports from all other countries, not included in the sample. The result of calculating the share of world trade this way is reflected in Ficture A2. Evidently the series based on aggregate data overall yields higher values for the share of world trade, as additional flows have now been included. Yet, the figure makes clear that the addition of these trade flows does not imply that a different course for world trade globally. On the contrary the series is characterized by the same trends as the original in Figure 1.

Appendix A4: Trade Frictions

To capture the evolution of trade frictions across countries and time we collect information on both tariffs barriers and transportation costs. Regarding tariffs our main source is Clemens and Williamson (2004). They provide data on average tariff rates, calculated as total import duties over the total value of imports, which span the period from 1870 to 1998. We extend the coverage of the data-set to 2005 by using more recent information from the same source that Clemens and Williamson drew from for their post-1950 series, the World Bank Development Indicators (WDI). We, furthermore, supplement these data with those from Schularick and Solomou (2011) who updated some of the values reported by Clemens and Williamson for the pre-1950 and provide additional series for the Netherlands and Switzerland, originally not covered in the data-set. Moreover, as Clemens and Williamson report separate series for the pre- and the post-1950 period, we merge the two series into one. As the two series overlap for a few years (1948-1950), when merging the two we average the reported values in the few cases where we observed discrepancies between the two. Finally, whenever possible, we fill in gaps in the series through interpolation of the missing values assuming constant annual growth rates.

To measure transportation costs over our period of interest we employ the global index constructed by Shah Mohammed and Williamson (2004) which combines historical with modern data to cover the whole period from 1870

Philippines, Portugal, Russia, Spain, Sri Lanka, Sweden, Switzerland, Tanzania, Thailand, Togo, Turkey, Uganda, United Kingdom, United States of America, Uruguay, Yugoslavia, Zambia, Zimbabwe.

to 1997. Each value of the original index corresponds to a 5-year interval starting from 1870-1874 with 1884 being the base year. To allow for year-by-year variation in transportation costs, we convert the index to an annual one as follows. We attribute each reported value to the mid-point of the interval - for example the 1870-1874 value to 1872 - and then interpolate the remaining values assuming a constant annual growth rate for the years in between. We also extend the index until 2005 to match the duration of our other data by extrapolating based on the growth rate observed during the 1990s. Given that during this period there was a constant decline of transportation costs as reported by Hummels (2007), which was not reversed at any point until the global financial crisis, we think that this extrapolation strategy is justifiable.

To make also our transportation cost figures for each year reflect correctly the actual magnitudes of transportation costs in ad valorem terms, we rebase the index and scale it based on the average value of the global CIF/FOB ratio reported in IMF's International Financial Statistics over period 1966-1995.[34] The use of this ratio as an absolute measure of transportation cost is justified by the following facts. First of all, this ratio is probably the most frequently used measure of transportation costs for the post-war period.[35] Secondly, the decline in transportation costs observed in the global CIF/FOB ratio is similar to the one reflected in the global index of Shah Mohammed and Williamson (2004). Moreover, the evolution of the CIF/FOB ratio relative to the US tariff rate, as reported by Clemens and Williamson (2004), follows the pattern discussed in Hummels (2007), namely from 100% in 1965 to 300% in 2000.

[34] We use the ratio's average value, rather than a particular annual value in order to mitigate the effect of potential measurement error in the CIF and FOB series.

[35] See for example Hummels (1999) and Baier and Bergstrand (2001).

	Value	Sources
Data		
Nominal bilateral imports	-	Barbieri, Keshk and Pollins (2009), IMF *Direction of Trade Statistics* (2012)
Nominal GDP	-	National accounts, Klasing and Milionis (2012)
Transport costs	-	Shah Mohammed and Williamson (2004), IMF *International Financial Statistics* (1996)
Tariffs	-	Clemens and Williamson (2004), Schularick and Solomou (2011)
Parameters		
θ	.11	Set to match avg. $\frac{\sum_c M_{ct}}{Y_t}$ for 1870-2005
σ	4	See text.

Table 1: Basic Data and Parameters

Model Performance				
Parameters				
σ		1.5	4	7
Goodness-of-fit				
Root MSE		1.89	2.02	2.44
$\triangle_{1870-2005} \frac{\sum_c M_{ct}}{Y_t}$	Model	4.7	6.3	8.7
	(Data)	(9.5)	(9.5)	(9.5)
θ	Model	0.10	0.11	0.14
	(Data)	(0.12)	(0.12)	(0.12)

Table 2: Baseline Calibration with Different Values of σ

Model Performance				
Parameters				
σ		4	4	4
η		1	1	1.5
θ_c		same $\forall$ c	c-specific	same $\forall$ c
Goodness-of-fit				
Root MSE		2.02	2.09	2.22
$\triangle_{1870-2005} \frac{\sum_c M_{ct}}{Y_t}$	Model	6.3	5.8	7.5
	(Data)	(9.5)	(9.5)	(9.5)
$\frac{1}{T} \sum_t \sum_c \theta_c \frac{Y_{ct}}{Y_t}$	Model	0.11	0.09	0.08
	(Data)	(0.12)	(0.12)	(0.12)

Table 3: Generalised Gravity Equation

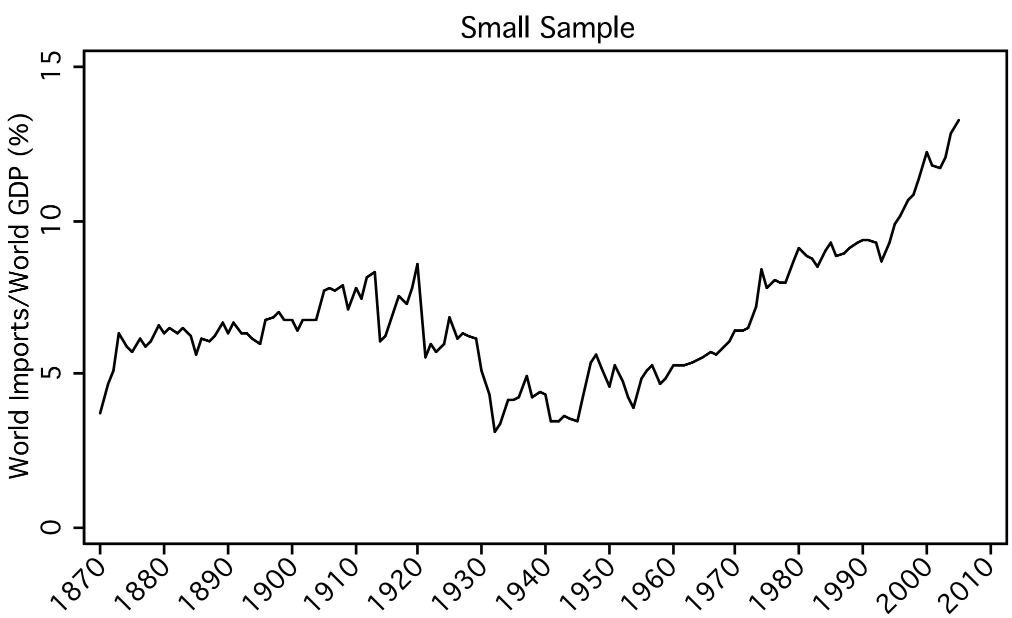

Figure 1: The Evolution of World Trade, 1870-2005

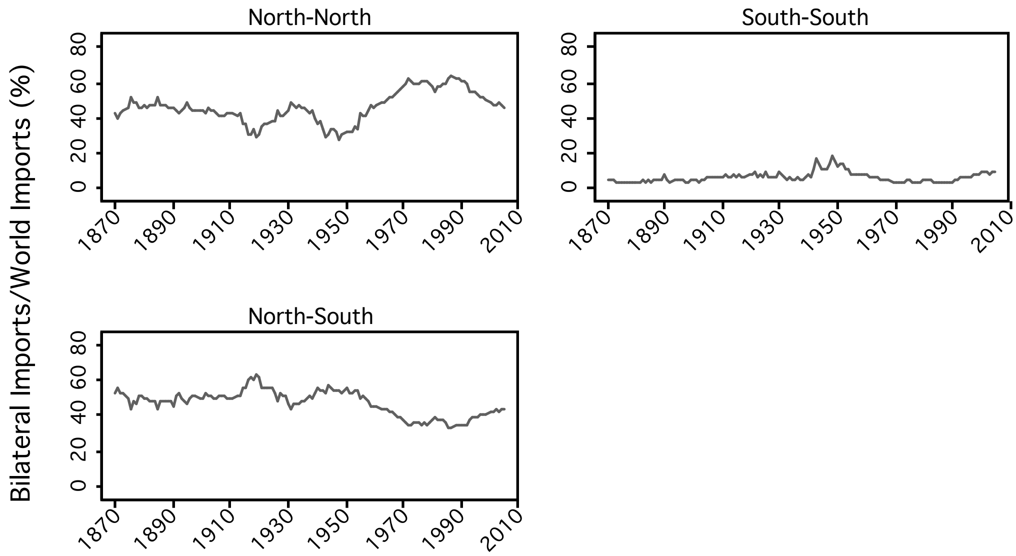

Figure 2: Bilateral Trading Patterns, 1870-2005

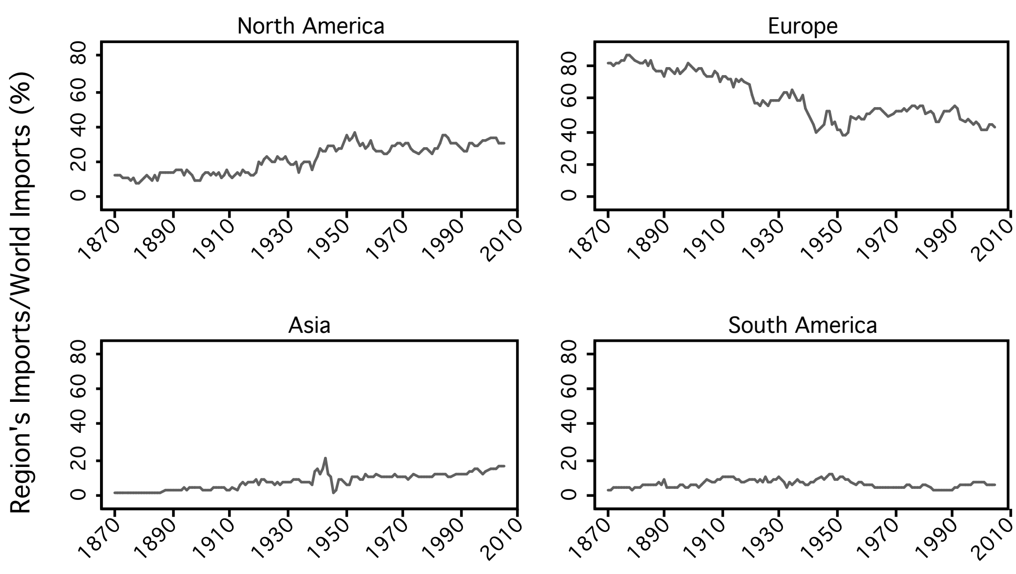

Figure 3: Regional Trade Shares, 1870-2005

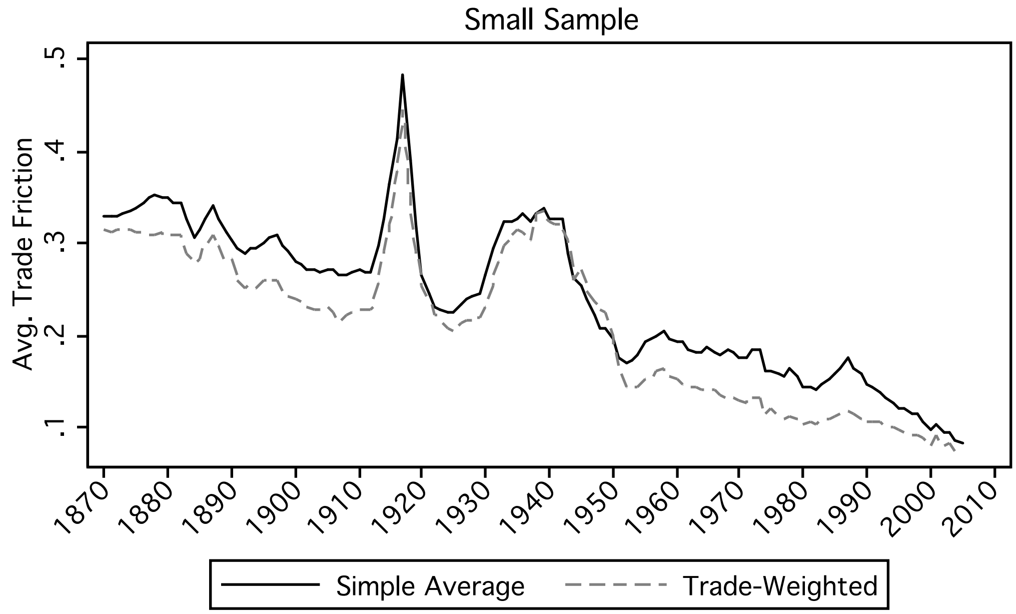

Figure 4: Magnitude of Trade Barriers, 1870-2005

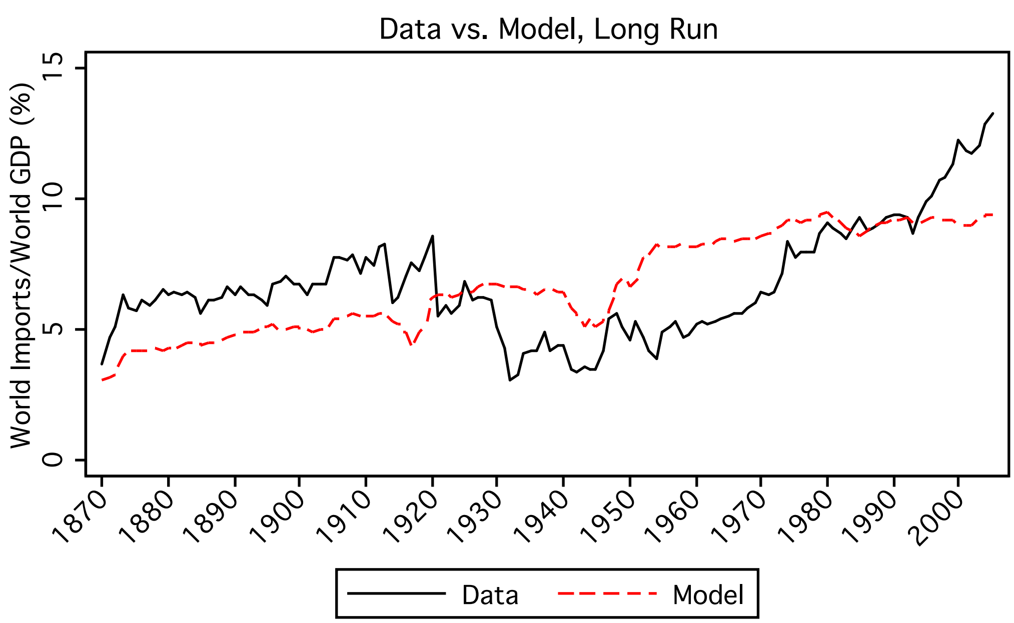

Figure 5: World Trade Share, 1870-2005 (Baseline Calibration)

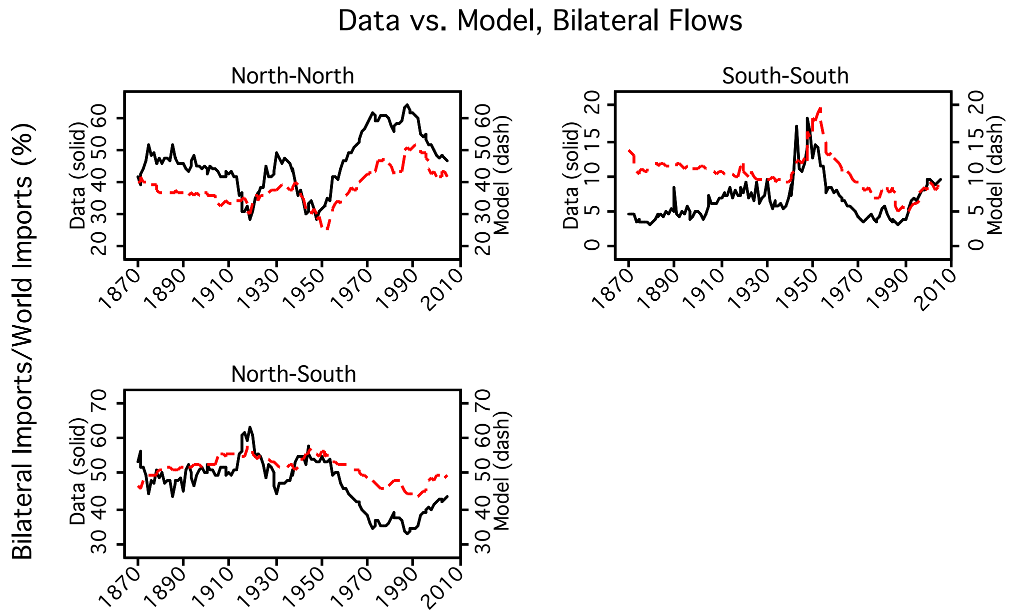

Figure 6: Bilateral Trading Patterns, 1870-2005 (Baseline Calibration)

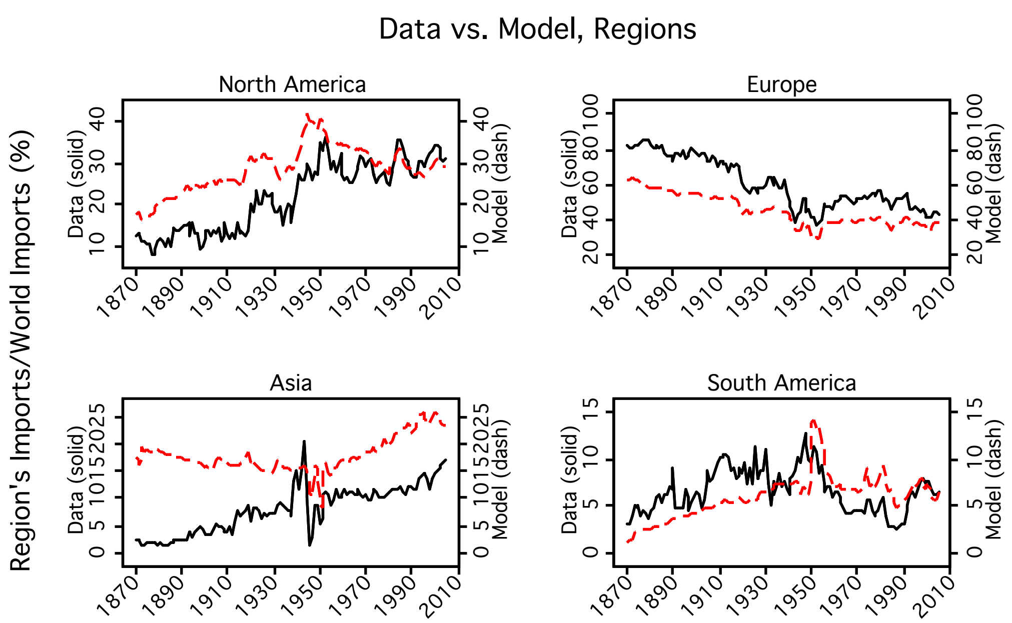

Figure 7: Regional Trade Shares, 1870-2005 (Baseline Calibration)

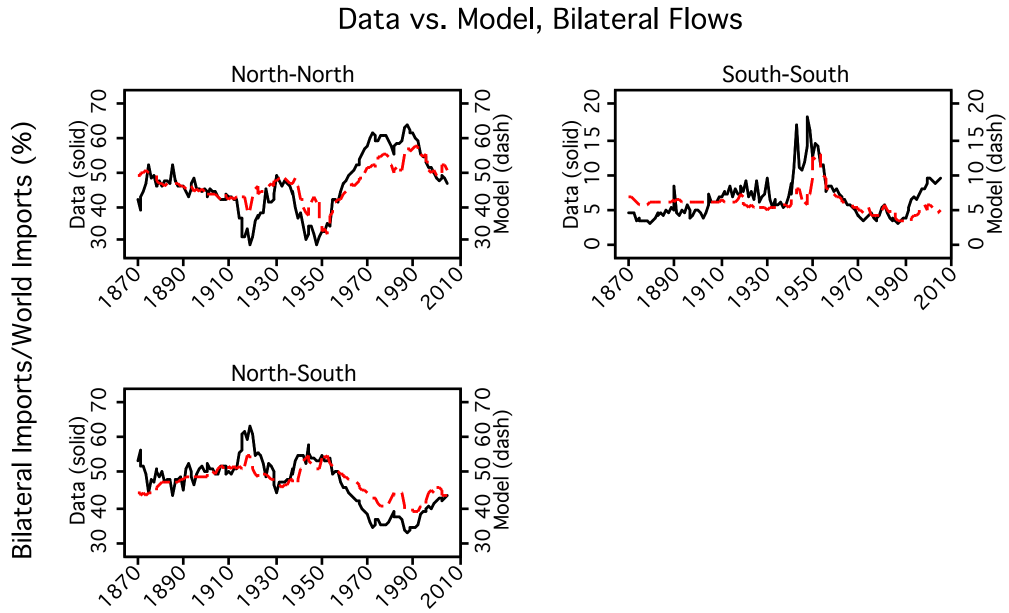

Figure 8: Bilateral Trading Patterns, 1870-2005 (Heterogeneous Tastes)

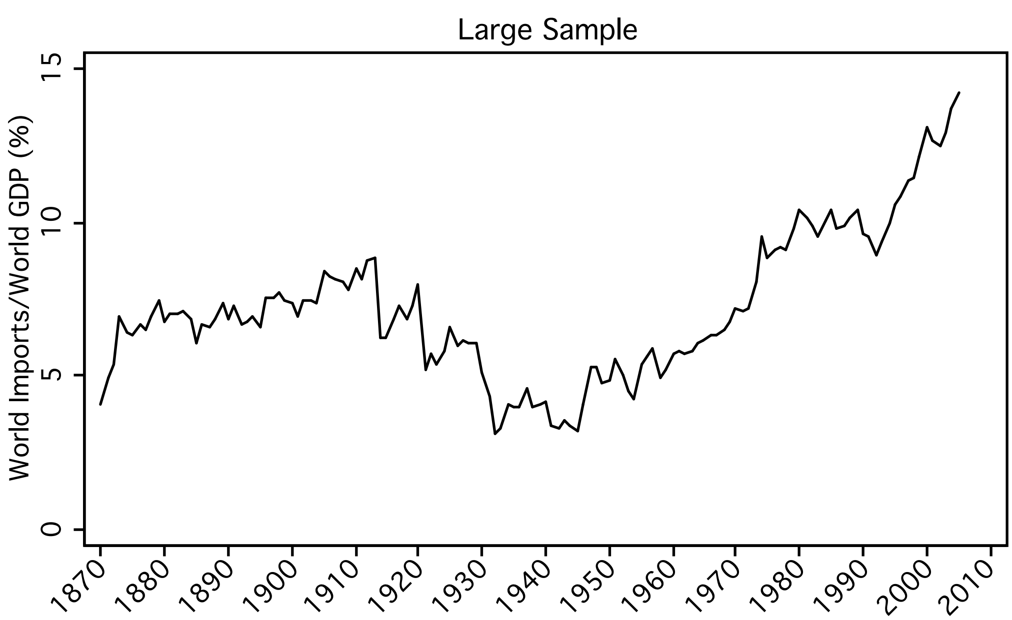

Figure A1: The Evolution of World Trade for a Large Country Sample, 1870-2005

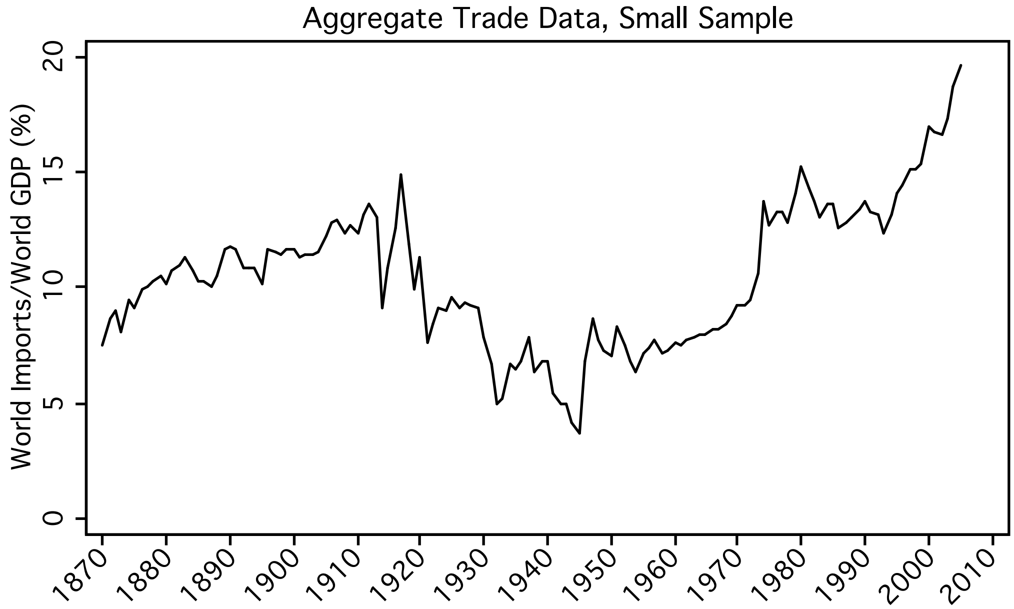

Figure A2: The Evolution of World Trade using Aggregate Trade Data, 1870-2005