Does Trade Raise Income?

Evidence from the Twentieth Century

Douglas A. Irwin*   Marko Terviö

Department of Economics   Department of Economics
Dartmouth College   MIT
Hanover, NH 03755   50 Memorial Drive
and NBER   Cambridge, MA 02142

This Draft: July 18, 2001

Abstract:

Efforts to estimate the effects of international trade on a country's real income have been hampered by the failure to account for the endogeneity of trade. Frankel and Romer recently use a country's geographic attributes – notably its distance from potential trading partners – to construct an instrument to identify the effects of trade on income in 1985. Using data from the pre-World War I, the interwar, and the post-war periods, we find that the main result of Frankel and Romer is confirmed throughout the whole century: countries that trade more as a proportion of their GDP have higher incomes even after controlling for the endogeneity of trade. We also find that the OLS estimate of trade's effect on income is biased downwards in almost every sample year. However, this result is not robust to the inclusion of distance from equator (latitude).

*Address for correspondence: Department of Economics, Dartmouth College, Hanover, NH 03755, USA. Fax: +1-603-646-2122. E-mail address: douglas.irwin@dartmouth.edu

Keywords: trade, income, distance, history

JEL Codes: F1, O4

1. Introduction

Perhaps the most fundamental proposition of international trade theory is that trade allows a country to achieve a higher real income than would otherwise be possible.[1] Yet this basic axiom has eluded satisfactory empirical treatment in part because ordinary least squares (OLS) regressions of income on trade neglect the endogeneity of trade and therefore fail to identify its true effect on income. Thus, the positive correlation between trade and income could mean that countries with higher incomes engage in more trade rather than the other way around.

Frankel and Romer (1999) recently overcame this simultaneity problem by constructing an instrument using a country's geographic attributes — notably its distance from other trade partners — that are correlated with trade but uncorrelated with income. In a cross-section of countries in 1985, they find that the effect of trade on income is considerably higher (although imprecisely estimated) with instrumental variables (IV) estimation, suggesting that the positive association of trade and income is not entirely due to high income countries trading more extensively than others.

This paper evaluates the Frankel-Romer finding across different time periods to see if there are systematic differences between the OLS and the IV estimates of trade's impact on income. We examine the trade-income relationship using data from the pre-World War I period (1913), the interwar period (1928), the Great Depression (1938), the early postwar period (1954), and several years in the later post-war period (1964, 1975, 1985, 1990). These data span various periods in which the relationship between trade and income could potentially differ from that in the late postwar period.

Using these data and the Frankel-Romer method, our findings are consistent with

theirs: in most periods during the twentieth century, instrumenting for trade with geographic characteristics raises the estimated positive effect of trade on income by a substantial margin and, in some cases, the precision of those estimates. With the exception of the interwar years, the IV estimate is greater than the OLS estimate. In addition, we can reject the hypothesis that the OLS and IV estimates are equal in three of the eleven samples (including the two most recent samples, 1985 and 1990).

We then also use a slightly different method in constructing the instrument, which yields an instrument with higher partial correlation with the endogenous variable, at the cost of reducing the sample and the applicability of the results. This alternative method yields similar but more precise IV results. We also briefly investigate the effect of including another geographic variable, latitude (distance from the equator). The inclusion of latitude does not significantly affect the OLS estimates of trade on income but destroys the consistent relationship between trade and income in the IV regressions. We conclude by discussing the implications of these findings.

2. Empirical Framework

The key insight of the Frankel-Romer paper is that a country's geographic attributes, especially distance from one's trading partners, yield important information on the "natural" or "expected" volume of its trade with other countries. While highly correlated with trade, these attributes are not important determinants of country's income (except as they operate through trade), nor does a country's income affect these attributes. Therefore, these geographic factors can be used as exogenous instruments for identifying the impact of trade on income.

These geographic characteristics, however, are also important determinants of within country trade. That is, countries with a relatively large population or land area, such as the United States, tend to have smaller trade to GDP ratios than smaller countries, such as Finland, because there are more opportunities for within-country trade in larger countries. Therefore, country size must also be taken into account when evaluating the effect of trade on income. Opportunities for within country transactions cannot be measured using the same monetary metric as international trade, but they can be proxied by population and area.

We adopt the empirical procedure devised by Frankel and Romer with some minor alterations. First, a geographical prediction of countries' bilateral trade flows is estimated using a variant of the standard gravity equation (omitting the income variables). More formally, we estimate:

(1) J_{ij} / log (t_{ij}/GDP_i) =

$a_0 + a_1$ log(distance$_{ij}$) + a_2 log(pop$_i$) + a_3 log(area$_i$) + a_4 log(pop$_j$) + a_5 log(area$_j$) + a_6 (border$_{ij}$) + a_7 (landlocked$_i$ + landlocked$_j$) + e_{ij}.

This equation states that the value of bilateral trade between country i and country j (denoted t_{ij}) relative to country i's GDP is related to the distance between them, their respective areas and populations, whether or not they share a common border, and whether one or both of the countries is landlocked.[2]

The instrument is constructed by aggregating the predicted values of country i's trade with countries j ($\hat{\tau}_{ij}$) over all j, yielding a predicted value of a country i's trade share with all sample countries j:[3]

(2) $\hat{T}_i = 3_{j..i}$ exp ($\hat{\tau}_{ij}$).

-4-

The simplest way to evaluate the impact of trade on income is to run the OLS

regression that explains a country's per capita income as dependent upon the trade ratio and

the country's area and population:

(3) $\log (GDP_i/pop_i) = b_0 + b_1 (T_i) + b_2 \log (area_i) + b_3 \log (pop_i) + u_i,$

where T_i is the actual trade to GDP ratio and where b_1 captures the contribution of trade on per

capita income. However, there are reasons to suspect that the estimate of b_1 is biased. High

income countries are better able to afford infrastructure that is conducive to trade (such as

better ports and airports) and have more resources to overcome the informational search costs

associated with trade. Rich countries may also simply demand relatively more tradable goods.

Open trade policies might be correlated with other successful domestic economic policies that

raise income, or alternatively poorer countries may be more likely to imposes taxes that

discourage trade but are relatively simple to administer. In these cases, the OLS estimate of b_1

would be biased upward because of a positive correlation between u_i and T_i, and IV estimation

would reduce the estimated effect of trade on income.

Alternatively, recorded trade could be subject to measurement error, in which case the

IV estimate would raise the estimated effect of trade on income. One type of measurement

error is that trade serves as an imperfectly measured proxy for a host of beneficial transactions

between countries, such as productivity-enhancing knowledge and R&D spillovers that are

thought to be transmitted via trade flows, as described by Helpman (1997). If these income-

enhancing transactions are only partially captured by recorded trade flows, this measurement

error would produce a downward biased estimate of b_1. It is impossible to distinguish between

these two cases just from OLS results.

To overcome the endogeneity problem of trade, the above OLS estimation is replaced

by two stage least squares (2SLS), where the geographical component of a country's trade is used to instrument for a country's actual trade share. A country's geographical features are not correlated with u_i and therefore (equation 2) gives an appropriate instrument for trade. In this case, the first stage regression is:

(4a) $T_i = c_0 + c_1 (\hat{T}i) + c_2 \log (area_i) + c_3 \log (pop_i) + v_i$,

The predicted values from this equation, $\widetilde{T}i$, are then used to estimate the second stage regression:

(4b) $\log (GDP_i/pop_i) = d_0 + d_1 (\widetilde{T}i) + d_2 \log (area_i) + d_3 \log (pop_i) + w_i$,

The coefficient estimate of d_1 can then be compared to the b_1 estimate to obtain information about the direction and magnitude of the possible OLS bias. Other independent determinants of per capita income are not likely to be correlated with the instrument and are thus relegated to the error term.

An important data issue at this point is the reliability of bilateral trade between small and distant countries. Sometimes trade is recorded as zero between these countries, although the actual trade is more plausibly very small rather than exactly zero. However, a predicted J_{ij} between such countries can be calculated on the basis of the coefficients estimated in equation (1) instead of using observations with the value of zero. This yields n(n-1) predicted values of J_{ij}, where n is the number of countries with bilateral trade data. We follow this procedure throughout and impute the geographically predicted trade shares for country pairs whose recorded trade is zero. Reassuringly, the predicted trade shares for country pairs with zero recorded trade are very small as they are typically small countries that are quite distant from one another (the notable exception being Israel and Egypt).

This imputation procedure can be taken a step further. Frankel and Romer calculate the predicted values of J_{ij} not only for all pairs of their 62 sample countries, but also for another 88 countries for which data on bilateral trade is not available. This extended sample comprises essentially the whole world, so the aggregated trade share variable is a country's total trade with the outside world divided by its GDP. In the income regressions, they then use these predicted trade shares to instrument for the trade share variable from Penn World Tables. We follow this procedure in the basic specification as closely as possible, but for the prewar samples the restricted availability of GDP data means that many countries that are used in this imputation stage are actually missing from the estimation of the income equations.

We also consider a slightly different method of constructing the instrument, in which the bilateral trade shares are not imputed for countries for which we do not have bilateral trade data (but we always impute them for country pairs that have zero reported bilateral trade). This alternative instrument gives the geographically predicted trade share with other sample countries, so the GDP share of actual trade with other sample countries only is used as the instrumented variable. If the gravity relation is systematically different for countries in the sample than for those that are added in the imputation stage, then this procedure can improve the quality of the instrument and thus the precision of the estimated effects of trade on income. This improvement however comes at the cost of excluding most smaller less developed countries from the income stage, and the results with the alternative instrument must be interpreted keeping in mind this narrower applicability.

3. Data and Empirical Results

The key data requirements are the values of bilateral trade and a measure of income.

Frankel and Romer use data from the International Monetary Fund's Direction of Trade Statistics for bilateral trade, a source which we also use for the years 1954, 1964, 1975, 1985, and 1990.[4] (The sample countries for each year are listed in the appendix.) Bilateral trade flows for 1913 were compiled from various national sources similar to those used by Estevadeordal (1997) and for 1928 and 1938 are from the League of Nations's publication Network of World Trade (1942) as used by Eichengreen and Irwin (1995) for analyzing interwar trade flows. Total trade figures for 1913, 1928, and 1938 are available in the League of Nation's annual Statistical Yearbook.

Frankel and Romer use the Penn World Tables for data on GDP for the year 1985. We use the Penn World Tables for nominal GDP in 1954, 1964, 1975, 1985, and 1990. Nominal GDP is used since trade is a nominal variable, and the units of measurement (U.S. dollars) cancel out from the trade/GDP ratio to facilitate a comparison of coefficients across the years. We use data from Prados de la Escosura (2000) for nominal GDP in the years 1913, 1928, and 1938. However, these figures are only available for a few, principally European countries and reduce the sample size significantly. As an alternative measure that allows for a larger sample, we use Maddison's (1995) real GDP data inflated to approximate nominal GDP by multiplying it by the U.S. price deflator.

The sources for the other variables are as follows. Frankel and Romer use the Penn World Tables for data on population, which we also use for the years 1954, 1964, 1975, 1985, and 1990.[5] Maddison (1995) presents population statistics for the years 1913, 1928, and 1938. Pre-World War II area is available in the League of Nations (1927). The distance data for the years 1913, 1928, and 1938 is based on Linnemann (1966), as was used in Eichengreen and Irwin (1995), and for the postwar years is the same data used by Frankel

and Romer.[6]

Finally, we note that the city-states of Hong Kong and Singapore (for which data were available since 1964 and 1975, respectively) are excluded from all of our samples because of concerns that they may be influential outliers.

Table 1 reports the estimated results from equation 1. (The columns for 1913-A, 1928-A, and 1938-A refer to results using the Prados de la Escosura GDP data, while those with the suffix B refer to results using the Maddison data.) The findings are unsurprising: the bilateral trade share is decreasing in the distance between the two countries, in the size of the home country's population, and is increasing in the population of the other country (as there are more consumers for the home country's products). Beyond this, sharing a common border has a positive effect on trade and being landlocked has a negative effect on trade, although these effects are significant only half the time. These results are generally consistent across all years. Geographic characteristics alone explain about 30 to 40 percent of the variance of the log of the bilateral trade share.

Table 2 reports the coefficients on the instrument from the first stage regression (equation 4a) that can be used to determine whether the geographic determinants of trade help explain the actual trade share beyond the information in country size and population. The message here is that the constructed trade share does indeed provide information about the actual trade share even after controlling for population and area, with the caveat that in 1928 and 1964 the predicted trade share is not a significant explanatory variable of the actual trade share.

Table 3 presents the complete results of the income regressions for the basic specification. With the exception of 1928 and 1938-B, the 2SLS estimates of the effect of

trade on income are larger than the OLS estimates. Frankel and Romer find that the IV coefficient is 2.3 times greater than the OLS estimate. We find that the 2SLS estimate exceeds the OLS estimate by a factor of 2.6, averaging all cases. Thus, the magnitude of our ratio across the century is quite comparable to that found by Frankel and Romer for 1985.

Frankel and Romer's estimates suggest a one percentage point increase in the trade share increases per capita income by 2 or 3 percent. In our samples, a one percentage point increase in the trade share increases per capita income by 3.0 percent, on average. Our estimated OLS effect of trade on income is somewhat larger than Frankel and Romer's, which helps to account for the larger 2SLS coefficient estimates that we find.

Frankel and Romer find that the OLS coefficient on the trade share is statistically significant but the IV coefficient is only marginally significant for 1985.[7] We also find that the OLS estimate is generally significant (the exceptions being 1913, 1938-A, and 1954), but also that the 2SLS estimate is significant at the 10 percent level in four of the eleven cases (1913-B, 1954, 1985, and 1990). The Hausman test indicates that we can reject the hypothesis that the 2SLS and OLS estimates are equal in each of these cases except 1913-B. In the other cases, we cannot reject the hypothesis that the 2SLS and OLS estimates are equal.

Data for Hong Kong and Singapore are available for the last three sample years. In results we do not report, including these two countries clearly raises the 2SLS estimates and lowers the OLS estimates. These results are driven by the anomalously high actual trade share of Singapore.

Table 4 provides a summary of the key results regarding the coefficient on trade in the income regression. Columns 1 and 2 present the coefficient as just reported. Columns 3

and 4 report the results using the slightly different method of constructing the instrument, in which the bilateral trade shares are not imputed for countries for which we do not have bilateral trade data (but is imputed for country pairs that have zero reported bilateral trade). This alternative instrument gives the geographically predicted trade share with other sample countries, so if the gravity relation is systematically different for countries in the sample than for those that are added in the imputation stage, then this procedure can improve the quality of the instrument and thus the precision of the estimated effects of trade on income (at the cost of excluding most smaller less developed countries from the income stage). With the alternative instrument the results are somewhat stronger. There the 2SLS estimate is significant in six of the eight years (the exceptions being the interwar years), and exceeds the OLS estimate on average by a factor of 3.3. This is because the alternative instrument is a clearly better predictor of the actual trade share in the postwar samples, and is insignificant only in 1928-A (this difference is not due to the different sample of countries in the estimation of the income equation, but holds also if the countries without bilateral trade data are excluded from the first stage regression of the basic specification).

As previously noted, there are several reasons why the 2SLS estimates of trade's impact on income could be expected to be lower than the OLS estimates. Yet the 2SLS coefficients are in fact higher in every year, except 1928 and 1938-B, and significantly so in the most recent years. The results lend credibility to the basic Frankel-Romer finding that the often noted positive relationship between trade on income is not an artifact of the endogeneity of trade. They also found that the IV estimate was higher than the OLS estimate, but since it was on the margin of statistical significance, they concluded that sampling variation (by which they mean a chance positive correlation between the

instrument and the residual) was "the most plausible explanation" for the result.

The uniformity of our results suggest that it is not chance sampling error at work, but perhaps their alternative explanation: that OLS is biased down because trade is subject to measurement error. The higher 2SLS coefficients could simply indicate that trade is consistently measured with error (without it being a proxy for other interactions) or that countries with poorly measured income also have poorly measure trade. An economically relevant, special case of measurement error is that trade is an imperfect proxy for a host of income-enhancing interactions with other countries.

While the Frankel-Romer approach yields results that appear to be consistent across time, another question is whether the results are robust to the inclusion of other geographic variables that could potentially enter into the gravity or income regression. Rodriguez and Rodrik (2000) have criticized the Frankel-Romer results on the grounds that they are not robust to the inclusion of a country's latitude (distance from equator).[8] The last columns of Table 4 present our results from including latitude in the income regression. The inclusion of latitude slightly diminishes the size of the OLS coefficient without affecting its statistical significance. However, the 2SLS coefficients lose their significance and are negative in half of the cases.[9]

There is an ongoing debate about the impact of latitude in these regressions. Using somewhat different control factors, Hall and Jones (1999) and Frankel and Rose (2001) find that the geographic-determined portion of trade remains a significant determinant of income in 1985 even after controlling for latitude. Unlike other geographic variables such as distance from trading partners, however, latitude does not have a natural economic interpretation.[10] Clearly, the precise role of latitude in determining economic performance is

unsettled and an issue for further research. However one wants to interpret the Rodrik-Rodriguez critique, it affects these regressions over the whole century.

4. Conclusions

A well-known stylized fact is that countries that trade more (as a proportion of their GDP) tend to have higher incomes, other things being equal. A perennial issue in interpreting this fact is that countries may trade more precisely because they are rich, i.e. the trade share is endogenous. Following Frankel and Romer, we use a plausible geography-based instrument to control for the endogeneity of trade. We confirm and strengthen their main conclusion over the twentieth century: that OLS is not an upward-biased indicator of trade's impact on per capita income and, if anything, OLS appears to be downward biased, sometimes significantly so. In other words, countries that trade more have higher incomes, even after controlling for the endogeneity of trade. This could mean that trade is measured with substantial error and/or that it is an imperfect proxy for other income-enhancing interactions between nations.

Appendix: Sample Countries

	1913	1928	1938	1954	1964	1975	1985	1990
AFGHANISTAN	Z	Z	Z					
ALBANIA	Z	Z	Z					
ALGERIA	Z	Z	Z		Y	Y	X	Y
ANGOLA	Z	Z	Z		Y	Y	Y	
ARGENTINA	A	A	A	X	X	X	X	X
AUSTRALIA	A	A	A	X	X	X	X	X
AUSTRIA	A	A	A	X	X	X	X	X
BAHAMAS							Y	
BAHRAIN						Y	Y	
BANGLADESH					Y	Y	Y	Y
BARBADOS					Y	Y	Y	
BELGIUM	A	A	A	X	X	X	X	X
BELIZE							Y	Y
BENIN					Y	Y	Y	Y
BHUTAN							Y	
BOLIVIA	Z	Z	Z	Y	X	Y	X	Y
BOTSWANA					Y	Y	Y	
BRAZIL	B	B	B	X	X	X	X	X
BULGARIA	A	A	A			Y	Y	Y
BURKINA FASO					Y	Y	Y	Y
BURUNDI					Y	Y	Y	Y
CAMEROON	Z	Z	Z		Y	Y	Y	Y
CANADA	A	A	A	X	X	X	X	X
CAPE VERDE IS.					Y	Y	Y	Y
CENTRAL AFRICAN REPUBLIC					Y	Y	Y	Y
CHAD					Y	Y	Y	Y
CHILE	B	B	B	X	X	X	X	X
CHINA	B	B	B		X	Y	X	X
COLOMBIA	B	B	B	X	X	X	X	X
COMOROS					Y	Y	Y	Y
CONGO					Y	Y	Y	Y
COSTA RICA	Z	Z	Z	Y	X	Y	Y	Y
CUBA	Z	Z	Z					
CYPRUS				Y	Y	Y	Y	Y
CZECHOSLOVAKIA		A	A		X	Y	Y	X
DENMARK	A	A	A	X	X	X	X	X
DJIBOUTI						Y	Y	
DOMINICA							Y	
DOMINICAN REP	Z	Z	Z	Y	Y	Y	Y	Y
ECUADOR	Z	Z	Z	Y	X	Y	X	Y
EGYPT	B	Z	Z	X	X	X	X	X
EL SALVADOR	Z	Z	Z	Y	Y	Y	Y	Y
ETHIOPIA	Z	Z	Z	X	X	X	X	
FIJI					Y	Y	Y	Y
FINLAND	A	A	A	X	X	X	X	X
FRANCE	A	A	A	X	X	X	X	X
GABON					Y	Y	Y	Y
GAMBIA	Z	Z	Z		Y	Y	Y	Y
GHANA (GOLD COAST)	Z	Z	Z		X	X	X	X
GREECE	A	A	A	X	X	X	X	X
GRENADA							Y	Y
GUATEMALA	Z	Z	Z	Y	Y	Y	Y	Y
GUINEA					Y	Y	Y	Y
GUINEA-BISSAU					Y	Y	Y	Y
GUYANA				Y	Y	Y	Y	Y
HAITI	Z	Z	Z		Y	Y	Y	
HONDURAS	Z	Z	Z	Y	Y	Y	Y	Y
HUNGARY	Z	A	A			Y	X	Y
ICELAND				Y	X	Y	X	Y

Country	1	2	3	4	5	6	7	8
INDIA	B	B	B	X	X	X	X	X
INDO-CHINA	Z	Z	Z					
INDONESIA (DUTCH EAST INDIA)	B	B	B		X	X	X	X
IRAN (PERSIA)	Z	Z	Z		X	X	X	X
IRAQ	Z	Z	Z	Y	Y	Y	Y	
IRELAND	Z	A	A	X	X	X	X	X
ISRAEL (PALESTINE)	Z	Z	Z	X	X	X	X	X
ITALY	A	A	A	X	X	X	X	X
IVORY COAST					Y	Y	Y	Y
JAMAICA				Y	X	Y	Y	Y
JAPAN	A	A	A	X	X	X	X	X
JORDAN (TRANSJORDANIA)	Z	Z	Z	Y	Y	Y	Y	Y
KENYA	Z	Z	Z	X	X	X	X	X
KUWAIT							X	
LAOS							Y	Y
LESOTHO					Y	Y	Y	Y
LIBERIA	Z	Z	Z		Y	Y	Y	
LIBYA	Z	Z	Z					
LUXEMBOURG				Y	Y	Y	Y	Y
MADAGASCAR	Z	Z	Z		Y	Y	Y	Y
MALAWI				Y	Y	Y	Y	Y
MALAYSIA	Z	Z	Z		Y	Y	X	Y
MALI					Y	Y	Y	Y
MALTA				Y	Y	Y	Y	
MAURITANIA					Y	Y	Y	Y
MAURITIUS				Y	Y	Y	Y	Y
MEXICO	B	B	B	X	X	X	X	X
MONGOLIA							Y	Y
MOROCCO	Z	Z	Z	Y	X	Y	X	Y
MOZAMBIQUE	Z	Z	Z		Y	Y	Y	Y
MYANMAR				Y	Y	Y	Y	
NAMIBIA					Y	Y	Y	Y
NEPAL	Z	Z	Z		Y	Y	Y	
NETHERLANDS	A	A	A	X	X	X	X	X
NEW ZEALAND	A	A	A	X	X	X	X	X
NICARAGUA	Z	Z	Z	Y	Y	Y	Y	Y
NIGER					Y	Y	Y	
NIGERIA	Z	Z	Z	X	X	X	X	X
NORWAY	A	A	A	X	X	X	X	X
OMAN							Y	Y
PAKISTAN				X	X	X	X	X
PANAMA	Z	Z	Z	Y	X	Y	Y	Y
PAPUA NEW GUINEA					Y	Y	Y	Y
PARAGUAY	Z	Z	Z	Y	X	Y	X	Y
PERU	B	B	B	X	X	X	X	X
PHILIPPINES	B	B	B	X	X	X	X	X
POLAND		A	A		Y	Y	X	Y
PORTUGAL	A	A	A	X	X	X	X	X
PUERTO RICO					Y	Y	Y	
QATAR							Y	
REUNION					Y	Y	Y	

Country								
ROMANIA	Z	A	A		X	X	Y	
RWANDA					Y	Y	Y	Y
SAUDI-ARABIA (ARABIA AUTONOMOUS STATES)	Z	Z	Z		X	Y	X	
SENEGAL					Y	Y	Y	Y
SEYCHELLES					Y	Y	Y	Y
SIERRA LEONE	Z	Z	Z		Y	Y	Y	Y
SOMALIA	Z	Z	Z		Y	YY	Y	
SOUTH AFRICA	B	Z	Z	X	X	X	X	X
(SOUTH) KOREA	Z	Z	Z	Y	X	X	X	X
SPAIN	A	A	A	X	X	X	X	X
SRI LANKA (CEYLON)	Z	Z	Z	Y	Y	Y	Y	Y
ST. LUCIA							Y	
SUDAN	Z	Z	Z			Y	X	Y
SURINAME					Y	Y	Y	
SWAZILAND					Y	Y	Y	
SWEDEN	A	A	A	X	X	X	X	X
SWITZERLAND	A	A	A	X	X	X	X	X
SYRIA	Z	Z	Z		Y	Y	Y	Y
TAIWAN (FORMOSA)	Z	Z	Z	X	X	X	X	X
TANZANIA	Z	Z	Z		X	X	Y	
THAILAND (SIAM)	B	B	B	Y	X	X	X	X
TOGO					Y	Y	Y	Y
TONGA							Y	
TRINIDAD & TOBAGO				Y	Y	Y	Y	Y
TUNISIA	Z	Z	Z		X	X	X	X
TURKEY	A	A	A	X	X	X	X	X
UGANDA	Z	Z	Z	X	X	X	Y	X
UNITED KINGDOM	A	A	A	X	X	X	X	X
UNITED ARAB EMIRATE							Y	
URUGUAY	Z	Z	Z	X	X	X	X	X
USA	A	A	A	X	X	X	X	X
USSR/RUSSIA	B	B	B		Y	Y	Y	
VANUATU							Y	Y
VENEZUELA	Z	B	B	X	X	X	X	X
WESTERN SAMOA							Y	Y
WEST GERMANY	A	A	A	X	X	X	X	X
YEMEN						Y	Y	
YUGOSLAVIA		A	A		X	X	X	X
ZAIRE (CONGO)	Z	Z	Z	Y	Y	Y	Y	
ZAMBIA	Z	Z	Z		Y	Y	Y	Y
ZIMBABWE (RHODESIA)	Z	Z	Z	Y	X	Y	Y	X

A - country with GDP data from both Prados de la Escosura and Maddison. B - country with GDP data only from Maddison. Z - country without bilateral trade data used in trade share imputation. X - country with bilateral trade data and GDP data. Y - country with GDP data but without bilateral trade data. Note: in post-war years all (still existing) countries are used in trade share imputation.

Table 1: Bilateral Trade

Dependent Variable: Log of Bilateral Trade Share (t_{ij}/GDP$_i$)
(standard errors in parenthesis)

	1913-A	1913-B	1928-A	1928-B	1938-A	1938-B	1954	1964	1975	1985	1990
Log of Distance	-0.67*	-0.88*	-0.60*	-0.59*	-0.55*	-0.60*	-0.74*	-0.87*	-1.14*	-0.90*	-1.04*
	(0.10)	(0.09)	(0.06)	(0.05)	(0.06)	(0.05)	(0.06)	(0.04)	(0.06)	(0.04)	(0.05)
Log of Population i	-0.02	-0.10*	-0.39*	-0.42*	-0.43*	-0.46*	-0.24*	-0.44*	-0.29*	-0.33*	-0.28*
	(0.07)	(0.06)	(0.05)	(0.04)	(0.04)	(0.03)	(0.04)	(0.03)	(0.04)	(0.03)	(0.04)
Log of Area i	-0.12*	-0.13*	0.05	0.04	0.02	0.03	-0.03	0.05*	-0.01	-0.01	-0.11*
	(0.06)	(0.05)	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)	(0.02)	(0.04)	(0.02)	(0.03)
Log of Population j	0.82*	0.78*	0.57*	0.48*	0.52*	0.38*	0.58*	0.55*	0.51*	-0.56*	0.51*
	(0.07)	(0.06)	(0.05)	(0.04)	(0.04)	(0.04)	(0.04)	(0.27)	(0.04)	(0.03)	(0.04)
Log of Area j	0.03	-0.19*	0.19*	-0.05	0.12*	-0.03	-0.02	-0.03	-0.04	-0.14*	-0.16*
	(0.06)	(0.05)	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)	(0.02)	(0.04)	(0.02)	(0.03)
Landlocked	-	-	-0.15	-0.43*	-0.27*	-0.50*	-0.10	-0.09	-0.06	0.31*	-0.63*
			(0.10)	(0.10)	(0.09)	(0.09)	(0.13)	(0.08)	(0.13)	(0.08)	(0.10)
Border	0.74*	0.66*	0.70*	0.43*	0.28*	0.10	0.31	0.05	0.15	0.58*	0.63*
	(0.32)	(0.29)	(0.18)	(0.16)	(0.17)	(0.15)	(0.25)	(0.19)	(0.25)	(0.18)	(0.23)
Constant	-7.29*	-3.66*	-3.27*	-1.30*	-2.41*	0.01	-3.19*	-1.08*	1.21	0.25	3.21*
	(1.20)	(0.92)	(0.77)	(0.56)	(0.71)	(0.53)	(0.69)	(0.47)	(0.67)	(0.44)	(0.59)
No. of Countries	23	36	29	41	29	41	41	62	50	60	50
N (t_{ij} > 0)	454	994	652	1134	660	1122	1356	2886	2304	3264	2364
N (t_{ij} = 0)	52	266	160	506	152	518	284	896	146	276	86
R^2	0.42	0.33	0.46	0.37	0.44	0.39	0.29	0.33	0.27	0.33	0.31
MSE	1.64	2.03	1.21	1.33	1.10	1.22	1.63	1.71	2.00	1.67	1.92

* indicates significant at the 10 percent level.

Note: A- Prados de la Escosura GDP data. B - Maddison GDP data.

Table 2: Relationship Between Actual and Constructed Trade Share

Dependent Variable: Trade to GDP Ratio
(standard errors in parenthesis)

	1913-A	1913-B	1928-A	1928-B	1938-A	1938-B	1954	1964	1975	1985	1990
Constructed Trade Share	2.10*	1.50*	0.24	0.58	0.29*	0.41*	1.50*	0.47	0.82*	1.32*	0.58*
	(0.85)	(0.40)	(0.20)	(0.27)	(0.13)	(0.11)	(0.61)	(0.30)	(0.20)	(0.27)	(0.16)
Log of Population	-0.08	-0.10*	-0.03	-0.02	-0.01	-0.00	-0.05*	-0.08*	-0.09*	-0.06*	-0.07*
	(0.09)	(0.04)	(0.05)	(0.02)	(0.03)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Log of Area	0.09	0.08*	-0.03	0.01	0.00	0.01	-0.03*	-0.02	-0.02	-0.01	-0.01
	(0.11)	(0.04)	(0.02)	(0.01)	(0.01)	(0.00)	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)
Constant	-0.12	0.48	0.63	0.17	0.05	-0.02	1.16*	1.33*	1.45*	1.06*	1.23*
	(1.00)	(0.33)	(0.67)	(0.33)	(0.32)	(0.12)	(0.29)	(0.23)	(0.21)	(0.24)	(0.25)
N	23	36	29	41	29	41	69	124	131	146	113
F-test	6.04*	13.80*	1.35	4.58*	4.65*	14.27*	5.98*	2.41	16.74*	23.44*	13.08*
(p-value)	(0.02)	(0.00)	(0.26)	(0.04)	(0.04)	(0.00)	(0.02)	(0.12)	(0.00)	(0.00)	(0.00)
R^2	0.28	0.39	0.43	0.42	0.62	0.61	0.52	0.47	0.54	0.57	0.45
MSE	0.46	0.23	0.14	0.10	0.05	0.04	0.20	0.22	0.26	0.26	0.26

* indicates significance at the 10 percent level.

Note: The F statistic tests whether the constructed trade share coefficient is significantly different from zero.

Note: A- Prados de la Escosura GDP data. B - Maddison GDP data.

Table 3: Relationship of Trade and Income

Dependent Variable: Log of Per Capita GDP
(standard errors in parenthesis)

A. OLS

	1913-A	1913-B	1928-A	1928-B	1938-A	1938-B	1954	1964	1975	1985	1990
Trade Share	0.18	0.44	2.83*	2.98*	3.80	6.54*	0.73	1.11*	1.21*	1.13*	0.99*
	(0.31)	(0.42)	(0.81)	(0.80)	(2.45)	(1.86)	(0.53)	(0.37)	(0.31)	(0.30)	(0.36)
Log of Population	-0.16	-0.10	0.09	0.02	0.07	-0.00	0.01	0.21*	0.24*	0.22*	0.24*
	(0.13)	(0.10)	(0.13)	(0.09)	(0.19)	(0.10)	(0.10)	(0.07)	(0.08)	(0.07)	(0.09)
Log of Area	0.17	-0.05	0.31*	-0.04	0.16	-0.01	0.03	-0.05	-0.04	-0.07	-0.08
	(0.10)	(0.08)	(0.08)	(0.06)	(0.09)	(0.06)	(0.08)	(0.05)	(0.06)	(0.06)	(0.07)
Constant	-1.46	-0.87	-5.12*	-1.87*	-3.64	-1.78*	5.38*	4.41*	4.74	5.98*	6.25*
	(1.26)	(0.92)	(1.45)	(0.89)	(2.11)	(0.97)	(1.16)	(0.77)	(0.82)	(0.77)	(0.93)
N	23	36	29	41	29	41	69	124	131	146	113
R^2	0.01	0.09	0.36	0.34	0.06	0.37	0.01	0.07	0.10	0.10	0.07
MSE	0.71	0.66	0.58	0.52	0.70	0.51	0.88	0.89	0.95	0.99	1.04

* indicates significant at the 10 percent level.

B. 2SLS

	1913-A	1913-B	1928-A	1928-B	1938-A	1938-B	1954	1964	1975	1985	1990
Trade Share	0.65	1.68*	2.37	1.28	7.62	2.70	4.91*	3.54	2.24	2.85*	3.30*
	(0.69)	(0.86)	(3.93)	(2.61)	(7.24)	(3.76)	(2.62)	(3.13)	(0.96)	(0.91)	(1.33)
Log of Population	-0.15	-0.03	0.05	-0.08	0.29	-0.13	0.30	0.45	0.36	0.40*	0.43*
	(0.14)	(0.12)	(0.33)	(0.18)	(0.43)	(0.14)	(0.22)	(0.32)	(0.13)	(0.12)	(0.14)
Log of Area	0.23	0.00	0.29	-0.04	0.18	-0.01	0.26	0.01	-0.00	0.00	0.02
	(0.13)	(0.09)	(0.17)	(0.06)	(0.10)	(0.06)	(0.18)	(0.09)	(0.07)	(0.07)	(0.10)
Constant	-2.15	-2.34*	-4.49	-0.47	-6.39	-0.20	-1.96	0.42	2.54	2.37	1.69
	(1.61)	(1.35)	(5.43)	(2.24)	(5.35)	(1.67)	(4.70)	(5.16)	(2.11)	(1.98)	(2.72)
N	23	36	29	41	29	41	69	124	131	146	113
Hausman F (p-value)	0.70	4.13*	0.02	0.55	0.44	0.68	6.25	0.88	1.48	5.62*	5.10*
	(0.41)	(0.05)	(0.90)	(0.46)	(0.51)	(0.20)	(0.20)	(0.35)	(0.23)	(0.02)	(0.03)
MSE	0.75	0.75	0.58	0.55	0.73	0.54	1.23	0.95	0.99	1.10	1.22
Ratio of 2SLS/OLS Coefficients on Trade	3.6	3.8	0.8	0.4	2.0	0.4	6.7	3.2	1.9	2.5	3.3

* indicates significant at the 10 percent level.

Note: A- Prados de la Escosura GDP data. B - Maddison GDP data.

Table 4: Coefficient on Trade Share in Income Equation

	Basic Specification		Reduced Imputation		Basic With Latitude	
	OLS	2SLS	OLS	2SLS	OLS	2SLS
1913-A	0.18	0.65	0.52	2.12	0.11	1.30
	(0.31)	(0.69)	(0.48)	(1.30)	(0.38)	(0.83)
1913-B	0.44	1.68*	0.37	1.61*	0.15	0.61
	(0.42)	(0.86)	(0.36)	(0.80)	(0.32)	(0.76)
1928-A	2.83*	2.37	1.57*	7.87	2.33*	-7.19
	(0.81)	(3.93)	(0.44)	(11.08)	(0.90)	(13.63)
1928-B	2.98*	1.28	1.07*	1.32	2.69*	-4.69
	(0.80)	(2.61)	(0.34)	(1.02)	(0.85)	(24.69)
1938-A	3.80	7.62	0.21	2.84	4.26*	-4.32
	(2.45)	(7.24)	(0.56)	(2.39)	(1.97)	(5.88)
1938-B	6.54*	2.70	0.81*	0.67	2.97	-0.10
	(1.86)	(3.76)	(0.32)	(0.50)	(2.38)	(7.38)
1954	0.73	4.91*	4.61*	12.03*	0.49	0.17
	(0.53)	(2.62)	(1.93)	(4.83)	(0.40)	(1.42)
1964	1.11*	3.54	2.66*	5.02*	0.80*	-4.43
	(0.37)	(3.13)	(1.07)	(2.23)	(0.29)	(5.30)
1975	1.21*	2.24*	1.83*	2.72*	0.82*	-0.63
	(0.31)	(0.96)	(0.66)	(0.96)	(0.24)	(0.94)
1985	1.13*	2.85*	2.98*	3.03*	0.47*	0.88
	(0.30)	(0.91)	(0.64)	(0.92)	(0.24)	(0.82)
1990	0.99*	3.30*	2.50*	2.14*	0.46*	1.01
	(0.36)	(1.33)	(0.48)	(0.60)	(0.26)	(0.95)

* indicates significant at the 10 percent level. Standard errors in parenthesis.

Note: A- Prados de la Escosura GDP data. B - Maddison GDP data.

Figure 1: The Instruments and the Actual Trade Shares
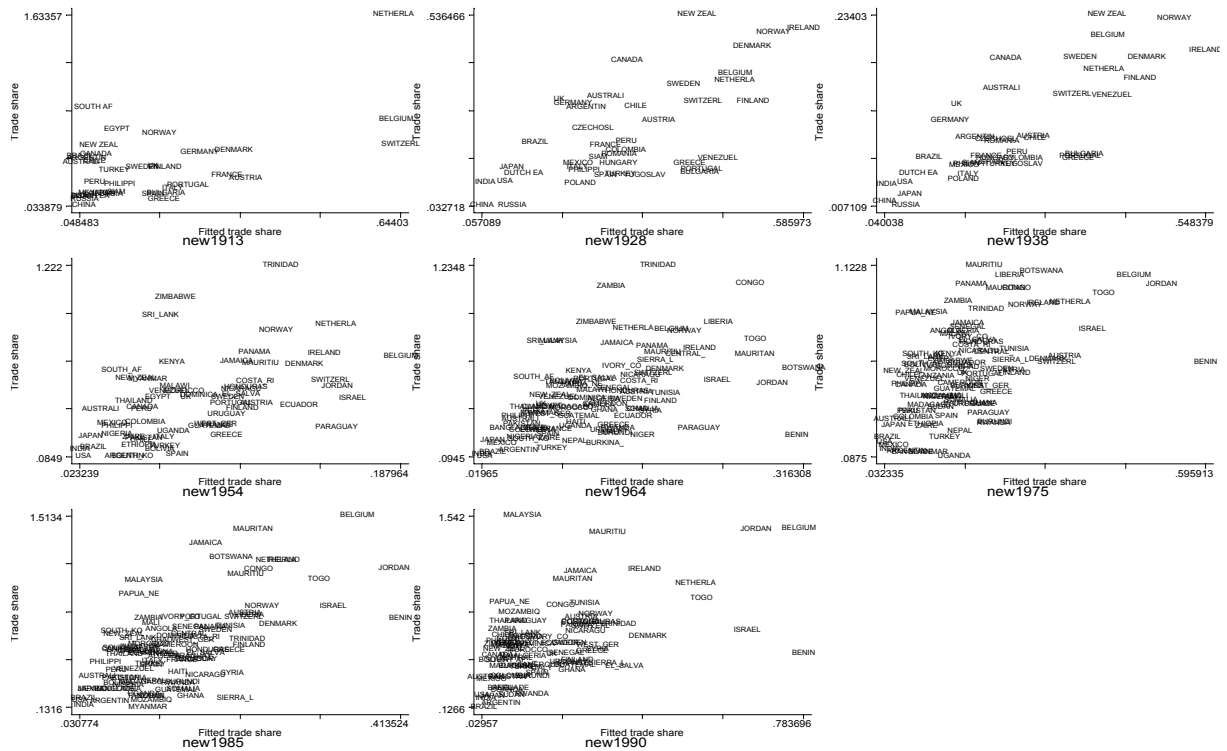
The Instruments and the Actual Trade Shares

References

Acemoglu, Daron, Simon Johnson, James A. Robinson. 2000. "The Colonial Origins of Comparative Development: An Empirical Investigation." MIT working paper.

Eichengreen, Barry, and Douglas A. Irwin. 1995. "Trade Blocs, Currency Blocs, and the Reorientation of World Trade in the 1930s." Journal of International Economics 38, 1-24.

Estevadeordal, Antoni. 1997. "Measuring Protection in the Early Twentieth Century." European Review of Economic History 1, 89-125.

Frankel, Jeffrey and David Romer. 1999. "Does Trade Cause Growth?" American Economic Review 89, 379-399.

Frankel, Jeffrey, and Andrew Rose. 2001. "An Estimate of the Effect of Currency Unions and Boards on Trade and Output." Working paper, Harvard University and University of California at Berkeley.

Hall, Robert E., and Charles I. Jones. 1999. "Why Do Some Countries Produce So Much More Output Per Worker than Others?" Quarterly Journal of Economics 114, 83-116.

Helpman, Elhanan. 1997. "R&D and Productivity: The International Connection." NBER Working Paper No. 6106.

Jones, Charles I. 2000. "Comment on Rodriguez and Rodrik, Trade Policy and Economic Growth: A Skeptic's Guide to the Cross-National Evidence." NBER Macroeconomics Annual 2000. MIT Press, Cambridge. Available from www.stanford.edu/~chadj.

League of Nations. 1927. International Statistical Yearbook, 1926. League of Nations, Geneva.

League of Nations. 1942. The Network of World Trade. League of Nations, Geneva.

Linnemann, Hans. 1966. An Econometric Study of International Trade Flows. North-Holland, Amsterdam.

Maddison, Angus. 1995. Monitoring the World Economy, 1820-1992. OECD, Paris.

Prados de la Escosura, Leandro. 2000. "International Comparisons of Real Product, 1820-1990: An Alternative Data Set." Explorations in Economic History 37, 1-41.

Rodriguez, Francisco, and Dani Rodrik. 2000. "Trade Policy and Economic Growth: A Skeptic's Guide to the Cross-National Evidence." NBER Macroeconomics Annual 2000 MIT Press, Cambridge.

Samuelson, Paul A. 1939. "The Gains from International Trade." Canadian Journal of Economics and Political Science 5, 195-205.

Samuelson, Paul A. 1962. "The Gains from International Trade Once Again." Economic

Journal 72, 820-829.

Summers, Robert and Alan Heston. 1991. "The Penn World Tables (Mark 5): An Expanded

Set of International Comparisons, 1950-1988." Quarterly Journal of Economics 106, 327-68.

Endnotes

1. Samuelson's (1939, 1962) two papers are among the classic articles on the gains from trade, but even Adam Smith discussed how a country open to international trade would increase "the exchangeable value of its annual produce."

2. Unlike Frankel and Romer, we do not include interaction terms with the border dummy as this simply introduces six new statistically insignificant variables and does not improve the equation's fit.

3. Aggregation would appear to bias the level of estimated T_i because the error term is aggregated through the exponential function, so $E[\exp(e_{ij})] > 0$. But under a homoskedasticity assumption this just multiplies the expectation of the predicted shares by a constant. This is not a problem here because the purpose is simply to obtain an instrument that is correlated with the underlying "natural" trade share of a country. The mean value of the instrument does not matter for this purpose, and a casual comparison of predicted and actual trade shares would be misleading.

4. Specifically, bilateral trade between countries i and j is measured as $M_{ij} + M_{ji}$, i.e. the sum of bilateral imports.

5. For a description of the Penn data, see Summers and Heston (1991). We use the Mark 5.6 version of the data available from the NBER website. Frankel and Romer's population variable is actually the working population constructed from Penn World Tables. We use total population and GDP at current international prices (cgdp) for GDP.

6. From the point of view of the gravity model, an ideal distance measure would give the (harmonic) average distance between any two inhabitants in the two countries. Here the distance is that between principal cities.

7. As Frankel and Romer note in their footnote 15, the standard errors of the estimated

coefficients in the gravity equations should also be accounted for in the 2SLS standard errors

because the instrument is constructed from the aggregated fitted values from the gravity

equations. This is done by using the delta-method and adding $\hat{V} = \hat{G}\hat{S}\hat{G}'$ to the usual 2SLS

covariance matrix $\hat{Z}$, where $\hat{S}$ is the estimated covariance matrix of **a**, the estimated vector of

coefficients of the gravity equation, and $\hat{G}$ is the gradient matrix $(d\hat{a} / d\hat{b})$. The revised

standard errors are then obtained as the square roots of the diagonal elements of the matrix sum

$\hat{Z} + \hat{V}$. As in Frankel and Romer, this procedure never increases the standard errors by more

than 10 percent.

8. The data for latitude are available at Chad Jones's web site: http://www.stanford.edu/~chadj

9. Here latitude was only added to the income stage, and the instrument was taken from the

basic specification. Latitudes are also statistically significant variables in the gravity equation,

but their inclusion would not change the results.

10. Hall and Jones (1999) justify using distance from the equator in an income regression by

arguing that it is a proxy for Western influence. Acemoglu, Johnson, and Robinson (2000) argue

that European countries colonized those parts of Asia and Africa with amenable climates and

environments, and once the effect of these western institutions is taken into account then latitude

does not play a role in determining per capita income.