DOES TRADE MATTER FOR GROWTH WHEN THE GEOGRAPHICAL INSTRUMENTS ARE RANDOMLY GENERATED?[1]

Sabine Deij, Jakob B. Madsen, and Laura Puzzello

Department of Economics, Monash University

Abstract. In their highly influential paper 'Does Trade Cause Growth?,' Frankel and Romer (FR) estimate a trade equation to predict bilateral trade shares, which are in turn used to construct an instrument for trade openness in income regressions. Several papers have followed the FR approach; however, they rarely state whether out-of-sample predictions of bilateral trade flows are included in the instrument set. Using bilateral trade shares predicted from randomly generated geographical characteristics to form instruments for trade openness, this paper shows that the results are highly sensitive to whether out-of-sample predictions are included in the instrument set. We show analytically and empirically that the coefficient of trade openness in income regressions is severely upward biased when out-of-sample predictions are excluded from the instrument set because the instrument captures the number of trading partners and, therefore, violates the exclusion restriction. Thus, out-of-sample predictions should always be included in the instrument set to eliminate mechanical endogeneity.

JEL: F14, F43; O40
Key words: Trade-growth nexus; Randomly generated instruments; Endogeneity

1. Introduction

In their seminal paper Frankel and Romer (1999, FR henceforth) propose a novel IV approach in which the geographic characteristics of countries are used to construct an instrument for trade in per capita income cross-country regressions. FR use a two-step approach for identification. In the first-step, they generate the instrument. They estimate a trade equation to predict bilateral trade shares based on geographic characteristics that are unrelated to income, such as bilateral distance, common border, size and landlockedness. They sum up predicted bilateral shares, including imputed shares (out-of-sample) corresponding to missing bilateral trade flows, to obtain, for each country, the share of trade to GDP predicted by its geography. In the second-step they employ the generated instrument to examine the relationship between trade and per capita income. They find that trade has a large and robust positive impact on per capita income.

The FR IV approach has gained widespread popularity in the literature on the growth effects of trade, migration and FDI (see, for analysis and discussion, Rodriguez and Rodrik, 2001; Irwin and

[1] We gratefully acknowledge the comments and suggestions of Kevin Staub, Vladimir Tyazhelnikov, Ken Clements, Mark Razhev and participants at the Melbourne Trade Workshop, Melbourne University, August 2017 and the Economics Workshop, Melbourne University, November 2017. Jakob Madsen acknowledges financial support from the Australian Research Council, grant DP150100061.

Terviö, 2002; Alcalá and Ciccone, 2004; Noguer and Siscart (2005); Harrison and Rodríguez-Clare, 2010; Feyrer, 2009; Andersen and Dalgaard, 2011; Ortega and Peri, 2014; Alesina *et al*. 2016; Pascali, 2017, among others).[2] However, the literature is often strikingly silent about the prediction method used to generate the instruments for trade in the per capita income regressions; particularly, whether out-of-sample bilateral trade flows are included in the instrument set for trade. Important exceptions are Noguer and Siscart (2005) and Gervais (2015), who argue that *only* in-sample predictions should be included in the instrument set because out-of-sample observations introduce noise and, therefore, reduce the strength of the instrument.

In this paper we show that the answer to the question 'does trade affect income' is highly sensitive to whether out-of-sample predictions are included in the instrument set. Specifically, we generate bilateral trade shares predicted from randomly generated geographical characteristics to construct instruments for trade openness. These are then used to test whether trade causes growth in the second-stage regressions (Section 2 and 3). Excepting type I errors, randomly generated instruments should be weak and result in insignificant relationships between trade and income, regardless of whether any relationship exists. However, this is not what we find. The coefficients of trade openness are, on average, significantly positive in 96-100% of the counterfactual regressions when only in-sample predictions are used as instruments for trade openness in the second-stage regressions. This casts serious doubt on this IV procedure. Conversely, the coefficients of trade openness are, on average, insignificant in 99% of the simulations when out-of-sample predictions are included in the instrument variable (IV) set as we would expect in a randomized experiment.

Why does the exclusion of out-of-sample predictions in the IV set create spuriously positive relationships between income and trade? In Section 4 we show analytically that this result is driven by endogeneity of the instrument using in-sample only predictions. This instrument captures the number of distinct partners a country trades with, which, in turn, is directly affected by its income. Indeed, low-income countries have fewer trading partners than high-income countries because they face higher trading costs due to their institutions, infrastructure and business environment (Djankov *et al*., 2002). Thus, the coefficient of trade openness is upward biased when only in-sample bilateral trade flows are included in the instrument.

The paper proceeds as follows. The empirical strategy and estimates are presented in Sections 2 and 3. Section 4 shows analytically and empirically that the positive correlation between trade openness and the number of trade partners creates an upward bias in the coefficient of trade

[2] The FR IV approach has also been applied to study the effects of trade on child labor (Edmonds and Pavcnik, 2006), the environment (Frankel and Romer, 2005) and volatility. The literature focusing on the effect of trade on volatility employs fitted measures of trade openness as an instrument for trade at the sector-level (e.g. Di Giovanni and Levchenko, 2009; and Ardelean *et al*., 2017).

openness in the income regression. In Section 5 it is demonstrated that the precision of the original FR instrument increases by expanding the sample of countries used in the estimation of the bilateral trade and the income equations, or by using real trade openness and income data from the most recent version of the Penn World Tables (PWT v9.0). Section 6 concludes.

2. Empirical Strategy

Consider the following income regression model:

$$lnY_i = \alpha_0 + \alpha_1 T_i + \alpha_2 lnN_i + \alpha_3 lnA_i + \boldsymbol{X'\alpha} + e_i \quad (1)$$

where Y_i is country i's income per capita; T_i is country i's total trade to GDP, i.e. trade openness; N_i and A_i is country i's population and area and is a proxy for within-country trade; and $\boldsymbol{X}$ is a vector of control variables. Identifying the effect of trade on income is complicated because of the two-way causal relationship between these two variables. FR address this issue by proposing a two-step procedure, which we follow as described below.

In the first-step FR generate instruments for trade openness by regressing bilateral trade openness on the following set of geographic characteristics:

$$\begin{aligned} \ln\left(\tau_{ji}/GDP_i\right) = {} & \beta_0 + \beta_1 \ln D_{ij} + \beta_2 lnN_i + \beta_3 lnA_i + \beta_4 lnN_j + \beta_5 lnA_j + \beta_6\left(L_i + L_j\right) \\ & + \beta_7 B_{ij} + \beta_8 B_{ij} D_{ij} + \beta_9 B_{ij} N_j + \beta_{10} B_{ij} A_i + \beta_{11} B_{ij} N_i + \beta_{12} B_{ij} A_j + \beta_{13} B_{ij}\left(L_i + L_j\right) + \varepsilon_{ij} \quad (2) \end{aligned}$$

where τ_{ji} is the total bilateral trade between country i and country j; D_{ij} is the geographic distance between country i and j; L is a dummy variable that takes on the value of 1 for landlocked countries and zero otherwise; B_{ij} is an indicator variable taking the value of 1 if countries i and j share a border and zero otherwise; and ε is a stochastic error term.

The estimates of Eq. (2) are used to form two instruments for country i's trade openness, T_i:

$$\hat{T}_i^{ISX} = \sum_{j \in \Omega_{ij}} e^{ln\left(\widehat{\frac{\tau_{ij}}{GDP_i}}\right)}, \quad (3)$$

$$\hat{T}_i^{TSX} = \sum_{j \in \Psi_{ij}} e^{ln\left(\widehat{\frac{\tau_{ij}}{GDP_i}}\right)}, \quad (4)$$

where Ω_{ij} is the set of countries with which i actively trades; Ψ_{ij} is the set of all countries with which i can potentially trade (i.e., those with which it does and does not trade). In words, $\hat{T}_i^{ISX}$ only includes in-sample predictions from Eq. (2), while $\hat{T}_i^{TSX}$ includes predictions over the total sample of possible trade flows (in-sample plus out-of-sample predictions).

In the second-step we employ the generated instruments to investigate the relationship between trade and income using Two-Stage Least Squares (2SLS). Accordingly, we estimate the following first-stage regressions:

$$T_i = \gamma_0 + \gamma_1 \hat{T}_i^{ISX} + \gamma_2 \ln N_i + \gamma_3 \ln A_i + \boldsymbol{X}'\boldsymbol{\gamma} + e_{1i}, \tag{5}$$

$$T_i = \mu_0 + \mu_1 \hat{T}_i^{TSX} + \mu_2 \ln N_i + \mu_3 \ln A_i + \boldsymbol{X}'\boldsymbol{\mu} + e_{2i}, \tag{6}$$

where e is a stochastic error term. These regressions yield the instruments $\hat{T}_i^{IS}$ (Eq. (5)) and $\hat{T}_i^{TS}$ (Eq. (6)). We estimate the following second-stage income regressions:

$$lnY_i = a_0 + a_1 \hat{T}_i^{IS} + a_2 \ln N_i + a_3 \ln A_i + \boldsymbol{X}'\boldsymbol{\xi} + e_{3i}, \tag{7}$$

$$lnY_i = b_0 + b_1 \hat{T}_i^{TS} + b_2 \ln N_i + b_3 \ln A_i + \boldsymbol{X}'\boldsymbol{\varsigma} + e_{4i}. \tag{8}$$

Using Eq. (8) without the $\boldsymbol{X}$ control variables as their baseline regression, FR find per capita income to be a significantly increasing function of trade openness. However, Rodriguez and Rodrik (2001) show that the coefficient of trade openness becomes insignificant when geographic and institutional controls are added to the baseline regression. Thus, we include the geographic and institutional controls suggested by Rodriguez and Rodrik (2001, RR henceforth) in the regressions below. The controls considered by Noguer and Siscart (2005) are included in the regressions in the online Appendix C to ensure that our results are robust to this consideration.

The key question asked in this paper is whether we should include or exclude unobserved bilateral trade pairs from the instruments; i.e., whether we should use either $\hat{T}_i^{ISX}$ or $\hat{T}_i^{TSX}$ as instruments. This is not a trivial issue because the maximum number of bilateral trading partners is significantly higher than the number of recorded trade flows and the results in most samples are influenced by this choice, as we show below. We would have a maximum number of bilateral trade flows of 15,778 in our 98-country sample if all countries traded with each other and every other possible partner, noting that there are 161 possible trading partners for each country. Instead, we observe 9,757 positive trade flows, which are used to estimate Eq. (2).[3] If the relationship between bilateral trade and geographic characteristics is very different for the non-trading or unobserved pairs, then more precise estimated trade effects can be achieved by excluding out-of-sample predictions from the instrument; a point originally made by Irwin and Terviö (2002). However, a much greater concern than efficiency is whether the coefficients of trade openness are biased in any of the income models given by Eqs. (7) and (8).

[3] Helpman *et al*. (2008) show a similar incidence of bilateral trade 'zeros' for each year between 1970 and 1997.

Our empirical strategy is as follows. First we estimate the first- and second-stage regressions using actual data to ensure that our results are consistent with those of FR and RR. Thereafter, we repeat the exercise using trade openness measures predicted from randomly generated geographic characteristics. In each round we estimate Eqs (7) and (8) in which 1) controls, ***X***, are excluded; 2) distance to equator is included; 3) the percentage of land in the tropics in included; and 4) continental dummies are included. The last three specifications follow RR and have been widely used as controls in the literature. OLS and 2SLS/ IV regressions are presented in all cases.

3. Empirical Analysis

3.1 Data

Following FR we use bilateral trade flows in 1985 from the IMF Direction of Trade Statistics between the 98 countries in Mankiw *et al.*'s (1992) sample and 161 possible trading partners (98 - 1 = 97 partners within the sample and 64 countries in the rest of the world). These countries tend to have the most reliable data of the world's countries, be large, and have per capita income levels that are less likely to be determined by idiosyncratic factors. Population, income (real GDP per capita) and trade openness are from PWT Mark 5.6. The CEPII GeoDist database is used as the source for the geographic variables: area, the landlocked dummy, latitudinal coordinates, bilateral distance (population-weighted) and dummy variables for common border. Data on the percentage of land or population in the tropics, and regional dummies (per continent) is from the Centre for International Development (CID). In the robustness section we present results for a larger sample of countries and use higher quality data from the most recent version of PWT. More details on the data are provided in online Appendix A.

3.2 FR and RR Replications

Table 1 shows the estimates for the bilateral trade equation, Eq. (2). The coefficients of the geographic characteristics are almost all statistically significant, while the coefficients of the interaction terms are almost all insignificant, results that are in line with those of FR.

Turning to Table 2, which reports the income regression results for our main four model specifications each estimated by OLS, and 2SLS/IV using $\hat{T}_i^{IS}$ and $\hat{T}_i^{TS}$ as instruments for trade, respectively. The following conclusions emerge from the regressions: First, the null hypothesis of weak instruments is rejected in all cases, suggesting that $\hat{T}_i^{IS}$ and $\hat{T}_i^{TS}$ are both potentially good instruments. Second, the coefficients of trade openness in the second-stage regression are positive and statistically significant, regardless of whether $\hat{T}_i^{IS}$ and $\hat{T}_i^{TS}$ are used as instruments in the baseline

FR regression (Model (1)) and when continental fixed effects are included in the regressions (Model (4)).

Table 1. Estimates of the Bilateral Trade Equation (Eq.(2))

	Variable (1)	Border Interaction (2)
Constant	-6.264***	6.687**
	(0.597)	(2.617)
Ln distance$_{ij}$	-1.110***	0.286
	(0.034)	(0.338)
Ln population$_i$	-0.134***	-0.284**
	(0.024)	(0.140)
Ln area$_i$	-0.141***	0.052
	(0.017)	(0.144)
Ln population$_j$	0.933***	-0.091
	(0.020)	(0.130)
Ln area$_j$	-0.233***	-0.056
	(0.017)	(0.157)
Landlocked$_{ij}$	-0.671***	0.159
	(0.053)	(0.181)
Observations	9757	
R^2	0.318	

Note. The dependent variable is ln(τ_{ji}/GDP_i). Colum (1) reports the coefficient of the variable listed, and column (2) shows the coefficient of the interaction between the variable in the first column and border. Heteroscedasticity consistent standard errors are in parentheses. **, *** Significant at 5 and 1 percent, respectively.

Consistent with the findings of RR, the coefficients of trade openness become insignificant in the IV-$\hat{T}_i^{TS}$-regressions when the out-of-sample predictions are included in the instrument set and the share of the fraction of land within the tropics or distance to the equator are included as controls (Models (2) and (3)). However, when out-of-sample predictions are excluded from the IV-set (IV-$\hat{T}_i^{IS}$-regressions), the coefficient of trade-openness becomes significant at least at the 5% level; a key result of Noguer and Siscart (2005). From these conflicting results it can be inferred that the coefficients of trade openness in the income equations must be biased in either the IV-$\hat{T}_i^{IS}$-regressions or the IV-$\hat{T}_i^{TS}$-regressions. Thus it can be concluded that the growth-trade nexus cannot be resolved before we know 1) which of the sampling procedures yield biased parameter estimates; and 2) the source of the bias. To identify which sampling procedure produces biased estimates we first generate both instruments for trade aggregating bilateral trade shares predicted from randomly generated geographical characteristics. Then we analyze the randomized instruments to identify the source of the bias, which we show is systematically related to per capita income.

Table 2. Estimates of the Income Equation using Actual Data

Income regressions:	**Model (1)**			**Model (2)**		
	OLS	IV- $\hat{T}_i^{TS}$	IV-$\hat{T}_i^{IS}$	OLS	IV- $\hat{T}_i^{TS}$	IV-$\hat{T}_i^{IS}$
Trade share$_i$	0.911***	2.454***	2.743***	0.578***	0.463	0.702**
	(0.305)	(0.686)	(0.736)	(0.204)	(0.377)	(0.339)
Ln population$_i$	0.271***	0.381***	0.402***	0.106	0.097	0.1163
	(0.102)	(0.131)	(0.140)	(0.072)	(0.0741)	(0.0730)
Ln area$_i$	-0.087	0.084	0.116	-0.087	-0.100	-0.074
	(0.088)	(0.129)	(0.131)	(0.065)	(0.074)	(0.0726)
Distance to equator$_i$				-0.087***	4.190***	4.124
				(0.065)	(0.3318)	(0.3252)
Obs.	98	98	98	98	98	98
R^2	0.145	-	-	0.600	-	-
First-stage regressions:						
$\hat{T}_i$	-	6.818***	7.166***	-	7.606***	8.484***
		(1.356)	(1.427)		(1.931)	(2.095)
Partial R^2	-	0.284	0.321	-	0.282	0.335
KP *rk* Wald *F*-stat	-	25.27	25.20	-	15.51	16.40
Income regressions:	**Model (3)**			**Model (4)**		
	OLS	IV- $\hat{T}_i^{TS}$	IV-$\hat{T}_i^{IS}$	OLS	IV- $\hat{T}_i^{TS}$	IV-$\hat{T}_i^{IS}$
Trade share$_i$	0.636***	0.643	1.083***	0.704***	1.073**	1.217***
	(0.205)	(0.416)	(0.382)	(0.254)	(0.507)	(0.442)
Ln population$_i$	0.072	0.073	0.109	-0.037	0.018	0.040
	(0.076)	(0.077)	(0.078)	(0.104)	(0.109)	(0.103)
Ln area$_i$	-0.082	-0.081	-0.033	0.040	0.065	0.074
	(0.070)	(0.082)	(0.082)	(0.065)	(0.076)	(0.073)
% Land in tropics$_i$	-1.580***	-1.536***	-1.536***			
	(0.167)	(0.167)	(0.167)			
Sub-Saharan Africa$_i$				-1.889***	-1.830***	-1.806***
				(0.206)	(0.210)	(0.206)
East Asia$_i$				-0.626*	-0.776**	-0.834**
				(0.340)	(0.367)	(0.348)
Latin America$_i$				-0.581**	-0.472*	-0.430*
				(0.221)	(0.250)	(0.233)
Obs.	98	98	98	98	98	98
R^2	0.547	-	-	0.594	-	-
First-stage regressions:						
$\hat{T}_i$	-	7.673***	8.128***	-	6.745***	7.843***
		(1.729)	(1.861)		(1.435)	(1.588)
Partial R^2	-	0.289	0.331	-	0.230	0.305
KP *rk* Wald *F*-stat	-	19.70	19.09	-	22.08	24.39

Note. Heteroscedastic consistent standard errors are in parentheses. The standard errors in the IV-regressions are corrected for the errors created from the generated regressors. $\hat{T}_i^{IS}$ is the predicted trade openness based on in-sample observations only. $\hat{T}_i^{TS}$ is the predicted trade openness based on the total sample, i.e., including out-of-sample predictions. The KP *rk* Wald *F*-stat is the Kleibergen-Paap *rk* Wald *F*-statistic. Exogenous variables are included in the first-stage regressions but not shown. *, **, *** Significant at 10, 5 and 1 percent, respectively.

3.3 Random Generated Instruments

In-sample and total-sample *random* generated trade openness, $\tilde{T}_i^{ISX}$ and $\tilde{T}_i^{TSX}$, respectively, are created for each Monte-Carlo replication b = 1,…, 1000, by randomly drawing bilateral distances, areas and populations from normal distributions with means and standard deviations equal to those observed in the data. For each replication we ensure that geographic distances are symmetric across bilateral trading partners, $D_{ij}(b) = D_{ji}(b)$, and that country i's area and population do not change whether i is the origin or the destination country, i.e., $A_i(b) = A_{j=i}(b)$ and $N_i(b) = N_{j=i}(b)$. The landlocked status is drawn from a random variable where the probability of drawing 1 (e.g. landlocked) equals the observed frequency of landlocked countries in the data. For each replication we ensure that country i's landlocked status does not change, regardless of whether i is the origin or the destination country, i.e., $L_i(b) = L_{j=i}(b)$. Finally, we draw symmetric borders from a random variable where the probability of drawing 1 equals the observed incidence of a border in the data.

First- and second-stage regressions are estimated for each replication. Table 3 summarizes the income regression results when the instruments are randomly generated (1000 replications for each model). The table reports 1) the average coefficients and the corresponding standard deviation (in parentheses); 2) the number of replications for which the coefficients of trade openness in the income regressions are statistically significant at the 10% [in square brackets] and 5% {in curly brackets} levels; 3) the number of Kleibergen-Paap *rk* Wald weak identification tests for which the F-statistic in the first-stage regression is greater than 10 <in angle brackets>.

Table 3. Estimates of the Income Equation using Randomized Instruments (1000 replications)

Second-stage results:	Model (1)		Model (2)		Model (3)		Model (4)	
	IV- $\tilde{T}_i^{TS}$	IV-$\tilde{T}_i^{IS}$	IV- $\tilde{T}_i^{TS}$	IV-$\tilde{T}_i^{IS}$	IV- $\tilde{T}_i^{TS}$	IV-$\tilde{T}_i^{IS}$	IV- $\tilde{T}_i^{TS}$	IV-$\tilde{T}_i^{IS}$
Trade share$_i$	-4.000	6.691	2.266	3.632	0.531	4.289	-1.938	6.062
	(88.754)	(0.892)	(89.180)	(0.676)	(62.945)	(0.715)	(66.014)	(1.297)
	[9]	[1000]	[7]	[997]	[11]	[1000]	[8]	[983]
	{1}	{1000}	{2}	{989}	{3}	{997}	{3}	{960}
Obs.	98	98	98	98	98	98	98	98
First-stage regressions:								
$\tilde{T}_i$	7.463	25.675	6.542	30.105	6.748	28.462	6.262	20.696
	(25.862)	(4.235)	(25.659)	(6.385)	(25.729)	(5.608)	(24.294)	(4.616)
	[136]	[1000]	[114]	[998]	[116]	[1000]	[125]	[980]
	{72}	{1000}	{63}	{979}	{64}	{996}	{67}	{951}
Partial R^2	0.013	0.125	0.013	0.112	0.013	0.114	0.013	0.074
	(0.019)	(0.029)	(0.019)	(0.033)	(0.019)	(0.032)	(0.019)	(0.026)
KP *rk* Wald F-stat	1.196	18.489	1.147	9.462	1.166	13.080	1.217	11.547
	<4>	<955>	<5>	<434>	<5>	<798>	<5>	<630>

Notes. The table reports average values from 1000 replications. The standard deviation of this average is reported in parentheses. For the estimated coefficients, the number of replications that produce an estimate significant at least at the 10% level is in [square brackets] and the number of replications in which the estimate is significant at least at the 5% level is in {curly brackets}. The number of times the KP *rk* Wald F-stat is greater than 10 is in <angle brackets>. Model (1) controls for area and population (in logs). Models (2), (3) and (4) add distance to the equator, percentage of land in the tropics and continental dummies, respectively, as control variables. Exogenous variables are included in the first-stage regressions but not shown. Full results for the second-stage regressions are reported in Appendix Table B1.

The results are remarkably sensitive to whether out-of-sample predictions are included in the IV set. Considering the results from the first-stage regressions, when only in-sample predictions are included in the regressions, IV-$\tilde{T}_i^{IS}$, the instruments turn out, in most cases, to be potentially strong. The KP *rk* Wald *F*-stats are larger than 10 in 43-95% of the cases and range, on average, between 9.5 and 18.5. Furthermore, the simulated coefficients of trade are statistically significant at the 5% level in at least 98% of the replications in the income regressions.

The results are quite different when the $\tilde{T}_i^{TS}$ instrument is used. In only 1% of the cases, at most, the coefficient of trade openness is significant at conventional significance levels. This suggests that the coefficients of trade openness are unbiased when $\tilde{T}_i^{TS}$ is used as an instrument for trade openness and, therefore, that causality is not found where it does not exist. Similarly, the *F*-tests of excluded restrictions are, on average, extremely low and the KP *rk* Wald *F*-statistic is, on average, very low and is greater than 10 only in 0.5% of the simulations. Again this suggests that trade openness is independent of geographic characteristics when these are randomly generated, as expected.

Overall, the simulations show that research relying on in-sample predictions will, almost surely, find a positive relationship between trade and income even if such a relationship does not exist; a relationship that disappears when out-of-sample predictions are included in the instrument set. This implies that the estimated effect of trade is biased when out-of-sample observations are excluded from the instrument set because of feedback effects from income to trade openness – a result we prove analytically in the next section.

4. The nexus between per capita income and number of trade partners

So what is giving these seemingly paradoxical results in the previous section? To answer this question we need to focus on the first-step, in which the instruments are generated.

When geographic characteristics are randomly generated, the bilateral trade equation (Eq. (2)) approximately predicts the logs of the average bilateral trade openness. More formally,

$$\widehat{ln\frac{\tau_{ij}}{GDP_i}(b)} \cong \overline{ln\frac{\tau_{ij}}{GDP_i}} = k,$$

where k is a constant equal to the average trade openness. Substituting this expression into Eqs. (3) and (4), yields the following two distinctive instruments:

$$\tilde{T}_i^{ISX}(b) = \sum_{j\in\Omega_{ij}} e^{ln\left(\widehat{\frac{\tau_{ij}}{GDP_i}(b)}\right)} \cong \sum_{j\in\Omega_{ij}} e^{k} = NP_i e^{k},$$

$$\tilde{T}_i^{TSX}(b) = \sum_{j\in\Psi_{ij}} e^{ln\left(\widehat{\frac{\tau_{ij}}{GDP_i}(b)}\right)} \cong \sum_{j\in\Psi_{ij}} e^{k} = 161e^{k},$$

where NP_i is the number of countries with which country i trades actively, and 161 is the maximum number of potential trade partners country i can trade with in our data. In other words, in each replication, $\tilde{T}_i^{ISX}$ captures the number of effective trade partners, which vary from country to country. However, when out-of-sample predictions are included in the data, $\tilde{T}_i^{TSX}(b)$ captures stochastic values that are scattered around $161e^k$ for all countries; where $161e^k$ is close to the values recovered from the estimates in this paper.[4]

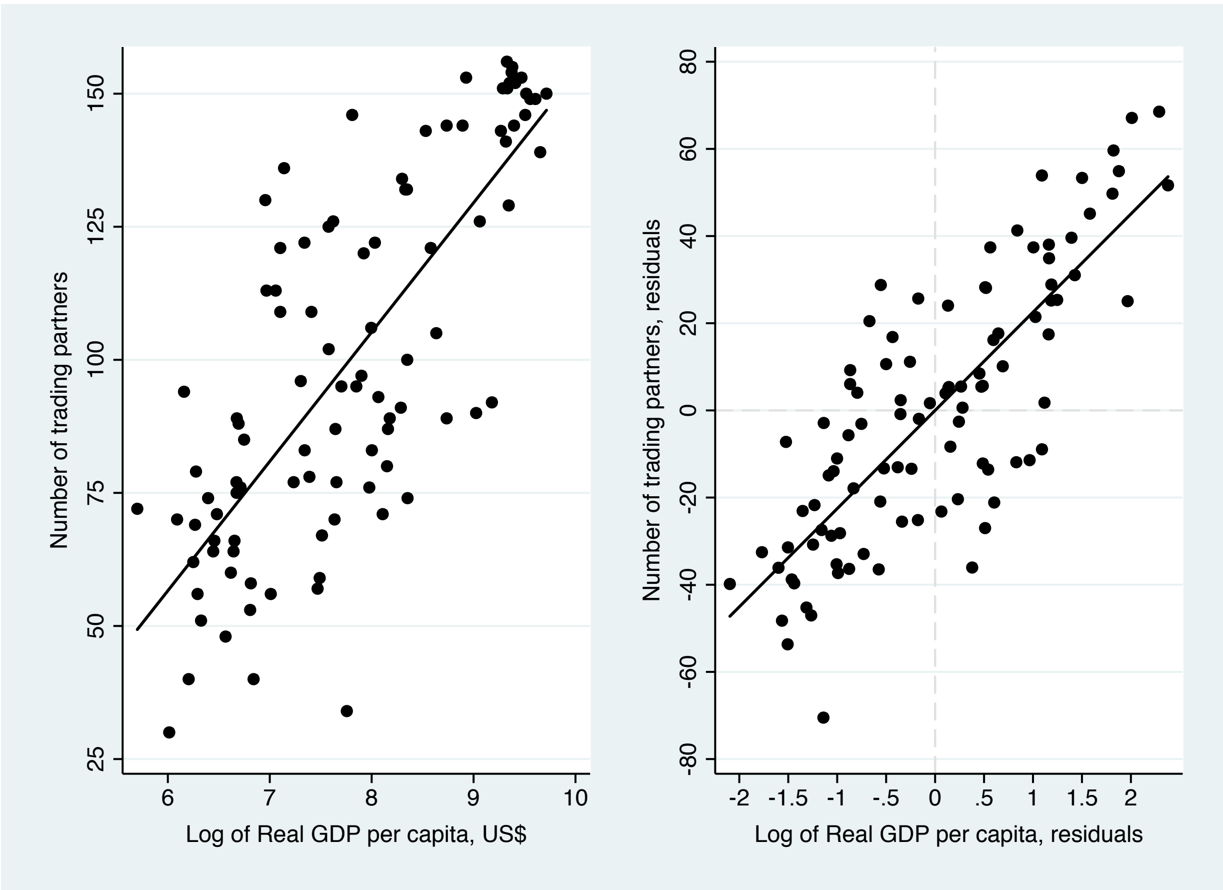

Figure 1. The relationship between the number of trading partners and per capita income. The left-hand-side panel plots the actual observations while the right-hand-side panel plots the residuals for each variable after accounting for the logs of population and area.

[4] In our data, the average log of the bilateral trade share is -8.922, which implies a value for e^k of 0.0013. With an average of 99.56 partners, the approximated average values for $\tilde{T}_i^{ISX}$ and $\tilde{T}_i^{TSX}$ are 0.0133 and 0.0214, respectively. These numbers are close to the average values for $\tilde{T}_i^{ISX}$ and $\tilde{T}_i^{TSX}$ of 0.0157 and 0.0231 across all countries and the replications in this paper.

What is the implication of these considerations for the 2SLS results? The estimated coefficient of $\tilde{T}_i^{IS}$ in the first-stage regression, γ_1, in Eq. (5), tends to be significantly positive because a country's trade openness and number of trading partners are positively correlated. At the same time, income per capita and the number of trading partners are positively correlated, as shown in Figure 1 (the correlation coefficient is 0.77 in the left-hand side panel).[5] The nexus between per capita income and the number of trading partners becomes even tighter when land area and population are controlled for, as shown in the right panel of Figure 1.

Intuitively, since the second-stage estimated coefficient of income can be derived as the ratio of the coefficients of the instruments from the reduced form and the first-stage regressions, the positive correlation between income per capita and the number of trading partners implies that the estimated coefficient of trade openness, $\hat{\alpha}_1$, in the second-stage Eq. (7), tends to be positive and significant too. If, by contrast, the out-of-sample observations, are included in the instrument set, $\tilde{T}_i^{TSX}$ does not have any identifying variation, the significance of the estimated μ_1 in the first-stage regressions, Eq. (6), and that of the estimated trade effect a_2 in the second-stage, Eq. (8), tends towards zero. This is exactly what the results in Table 3 show.

To show more explicitly that the coefficients of trade openness in the income regressions are mostly driven by the number of trade partners when the out-of-sample observations are excluded from the data, the variation in $\hat{T}_i^{ISX}$ is decomposed into trade openness (intensive margin) and the number of countries that country i trades with, NP_i (extensive margin):

$$\hat{T}_i^{ISX} = \sum_{j\in\Omega_{ij}} e^{ln\left(\widehat{\frac{\tau_{ij}}{GDP_i}}\right)} = NP_i * \frac{\sum_{j\in\Omega_{ij}} e^{ln\left(\widehat{\frac{\tau_{ij}}{GDP_i}}\right)}}{NP_i} = NP_i * \frac{\hat{T}_i^{ISX}}{NP_i} = NP_i * \overline{\hat{T}_i^{ISX}}$$

where $\overline{\hat{T}_i^{ISX}}$ is each country's average predicted bilateral trade openness.

The results of decomposing each margin into separate instruments for trade openness are presented in Table 4, where actual data are used. The regressions reveal a very distinct pattern. When NP_i is used as the only instrument for trade, the estimated coefficients of trade in the income regressions are very close in magnitude to the average coefficients obtained from the instruments when randomized, $\tilde{T}_i^{ISX}$ (see Table 3). This result confirms our hypothesis that the identifying variation in $\tilde{T}_i^{ISX}$ is driven solely by the number of trade partners of country i.

[5] Poor countries such as Botswana, Burkina Faso, Chad and Nepal, for example, have at most 40 trading partners, while advanced countries, such as Australia and the US, have at least 150 partners. Even small advanced economies such as Denmark, have a large number of trading partners (153), suggesting that the positive relationship is not driven by the size of the population or land area.

Table 4. 2SLS Estimates of the Income Equation Using Actual Data in the Extensive and Intensive Margins of $\hat{T}^{ISX}$

Second-stage results:		Model (1)		Model (2)		
Trade share$_i$	4.902***	6.762***	-0.778	2.236***	3.616***	-1.814
	(1.223)	(1.445)	(1.839)	(0.673)	(0.993)	(1.955)
Ln population$_i$	0.555***	0.688***	0.151	0.240**	0.351**	-0.086
	(0.206)	(0.261)	(0.173)	(0.114)	(0.175)	(0.212)
Ln area$_i$	0.355	0.561*	-0.274	0.093	0.242*	-0.346*
	(0.233)	(0.331)	(0.222)	(0.103)	(0.133)	(0.208)
Distance to equator$_i$				3.700***	3.318***	4.820***
				(0.487)	(0.792)	(0.856)
Obs.	98	98	98	98	98	98
First-stage regressions:						
NP_i	0.004***	0.004***		0.006***	0.005***	
	(0.001)	(0.001)		(0.002)	(0.002)	
$\overline{\hat{T}_i^{ISX}}$	385.26**		320.12*	442.69**		303.03*
	(145.67)		(177.66)	(170.57)		(177.84)
Partial R^2	0.197	0.139	0.040	0.212	0.136	0.037
KP rk Wald F-stat	12.71	23.00	3.25	5.765	10.97	2.903
Hansen J-statistic	7.28			10.86		
[p-value]	[0.007]			[0.001]		
Second-stage results:	**Model (3)**			**Model (4)**		
Trade share$_i$	2.622***	4.336***	-2.493	3.116***	6.256***	-4.272
	(0.763)	(1.019)	(2.603)	(0.948)	(1.588)	(4.632)
Ln population$_i$	0.236*	0.377**	-0.185	0.324**	0.795***	-0.783
	(0.125)	(0.192)	(0.277)	(0.164)	(0.289)	(0.723)
Ln area$_i$	0.134	0.320*	-0.422	0.202*	0.412*	-0.294
	(0.123)	(0.175)	(0.278)	(0.115)	(0.223)	(0.309)
% Land in tropics$_i$	-1.384***	-1.215***	-1.888***			
	(0.212)	(0.308)	(0.359)			
Sub-Saharan Africa$_i$				-1.502***	-0.998**	-2.687***
				(0.272)	(0.424)	(0.713)
East Asia$_i$				0.129	1.053*	-2.045
				(0.331)	(0.604)	(1.340)
Latin America$_i$				-1.606***	-2.882**	1.396
				(0.544)	(1.132)	(1.973)
Obs.	98	98	98	98	98	98
First-stage regressions:						
NP_i	0.006***	0.005***		0.005***	0.003***	
	(0.001)	(0.001)		(0.001)	(0.001)	
$\overline{\hat{T}_i^{ISX}}$	446.14**		297.042	444.18**		228.22
	(170.28)		(181.12)	(168.68)		(173.36)
Partial R^2	0.208	0.133	0.035	0.156	0.078	0.024
KP *rk* Wald F-stat	7.81	16.05	2.69	9.05	15.40	1.733
Hansen J-statistic	11.743			12.736		
[p-value]	[0.001]			[0.000]		

Note. Heteroscedastic consistent standard errors are in parentheses. The standard errors in the second-stage are corrected for the errors created from the generated regressors. $\hat{T}_i^{IS}$ is the predicted trade openness based on in-sample observations only. NP_i and $\overline{\hat{T}_i^{ISX}}$ are i's number of trading partners and average bilateral predicted trade openness, respectively. The KP *rk* Wald F-stat is the Kleibergen-Paap *rk* Wald F-statistic. Exogenous variables are included in the first-stage regressions but not shown. *, **, *** Significant at 10, 5 and 1 percent, respectively.

Turning to the second-stage results in Table 4, the estimated trade effect is large and significant *only* when the number of trading partners, for each individual country, NP_i, is included as an instrument for trade. This implies that NP_i and $\overline{\hat{T}_i^{ISX}}$ identify different vectors of parameters – a result that is not revealed by tests of overidentifying restrictions (Parente and Santos Silva, 2012). In addition, comparing the results in the first two columns for each model panel of Table 4, the estimated trade effects are not statistically different[6]; however, NP_i by itself is the strongest of the two instruments for trade (the corresponding *KP* statistics are greater than 10 for all models). Finally, note that the null hypotheses of overidentifying restrictions are rejected at the 1% level in all cases, underscoring that the validity of the overidentifying restrictions are not sufficient conditions for the model to be identified.

These findings suggest that the cross-country variation in NP_i is what makes the coefficient of $\hat{T}_i^{ISX}$ more significant than that of $\hat{T}_i^{TSX}$ in the income regressions (see results in Table 2, and online Appendix Tables C1 and C2). However, it is the same variation that makes $\hat{T}_i^{ISX}$ an invalid instrument for trade because the number of trading partners is endogenous to income, i.e., the Cov $(\hat{T}^{ISX}, e_3) \neq 0$; a violation of the exclusion restriction. Indeed, more developed countries have access to better institutions, infrastructure and business environments (Djankov *et al*., 2002) so that the cost of engaging in trade (exporting and importing) tends to be lower for them than for poor countries allowing them to trade with more partners.

The model by Helpman *et al*. (2008) offers a simple framework that addresses this issue. Extending the model of Melitz (2003) to include fixed costs of exporting and bounded productivity distributions, they show that some countries do not trade with each other because the firms are not sufficiently productive to penetrate each other's markets. In this framework destinations with lower fixed cost of exporting are, *ceteris paribus*, more likely to trade with any other country. Along the same vein using product level export data, Baldwin and Harrigan (2011) show that richer countries are more likely than poor countries to import from the US.

5. Does trade really matter for growth?

Our findings imply that even though the FR baseline regressions yield unbiased results, their efficiency and consistency properties may be compromised by the small country sample and data quality. In this section we show that the precision of the original FR instrument improves when we

[6] The values for the *t*-statistics are: -0.9825; 1.1504; -1.3464; -1.6978. Thus, none of the differences are significant at the 5% level; however, the last restriction is marginally rejected at the 10% level.

use a larger sample of countries, real trade openness as opposed to nominal trade openness, and improved quality of GDP data provided in the latest version of PWT v9.0.

Table 5. Estimates of the Income Equation using Actual Data: Larger Sample of Countries

Second-stage results	**Model (1)**	**Model (2)**	**Model (3)**	**Model (4)**
	IV- $\hat{T}_i^{TS}$	IV- $\hat{T}_i^{TS}$	IV- $\hat{T}_i^{TS}$	IV- $\hat{T}_i^{TS}$
Trade share$_i$	3.058***	1.157**	1.097*	1.736***
	(0.796)	(0.566)	(0.635)	(0.541)
Ln population$_i$	0.330***	0.084	0.064	0.156*
	(0.117)	(0.064)	(0.078)	(0.084)
Ln area$_i$	0.126	0.001	0.005	0.092
	(0.120)	(0.068)	(0.072)	(0.065)
Distance to equator$_i$		3.601***		
		(0.346)		
% Land in tropics$_i$			-1.335***	
			(0.195)	
Sub-Saharan Africa$_i$				-1.544***
				(0.167)
East Asia$_i$				-0.311
				(0.190)
Latin America$_i$				-0.990***
				(0.357)
Obs.	147	147	146	147
First stage regressions:				
$\hat{T}_i$	3.346***	3.357***	3.352***	3.317***
	(0.604)	(0.774)	(0.824)	(0.656)
Partial R^2	0.169	0.145	0.139	0.151
KP *rk* Wald *F*-stat	30.706	18.789	16.566	25.565

Note. The dependent variable is log of real GDP per capita reported by PWT Mark 5.6 for the year 1985. Heteroscedastic consistent standard errors are in parentheses. The standard errors in income regressions are corrected for the errors created from the generated regressors. $\hat{T}_i$ is the predicted trade openness based on the total sample; i.e. including out-of-sample predictions. The KP *rk* Wald *F*-stat is the Kleibergen-Paap *rk* Wald *F*-statistic. Exogenous variables are included in the first-stage regressions but not shown. *, **, *** Significant at 10, 5 and 1 percent, respectively.

5.1 Using a larger country sample

Table 5 reports the first- and second-stage estimates of the income equation when the sample of countries is increased (from 98) to 146/147 and $\hat{T}_i^{TS}$ is used as instrument for trade. The data are for trade openness and real GDP are identical to that used earlier, from PWT Mark 5.6. The coefficient of trade openness is significantly positive in all models; even when the fraction of land in the tropics or the distance to the equator are included as controls. Furthermore, in online Appendix Table D1, when the sample of countries is expanded (from 98 to 104-146, where the number of observations are limited by the availability of controls), it is shown that the statistical significance of the coefficient of $\hat{T}_i^{TS}$ improves, relative to the results in Table C1, in 8 out of the 10 additional specifications in which additional controls are included in the regressions. Thus, the increased number of countries in our sample yields significant trade effects in 12 of the 14 estimated models, which is a considerable improvement over the regression results in Table C1.

When the instruments are randomly generated, online Appendix Table D2 shows that the results in Section 3.3 are robust to the larger sample of countries. Again, the randomly generated instrument using only in-sample predictions, $\tilde{T}_i^{IS}$, consistently generates positive and significant income-effects of trade openness, while its counterpart, $\tilde{T}_i^{TS}$, which includes out-of-sample predictions, only produces significant income-effects in 1%, or less, of the replications. These results reinforce the results in the previous sections that $\hat{T}_i^{TS}$ yields unbiased estimates of the income effects of trade.

5.2 Using real openness and improved data

Trade openness has thus far been measured as nominal imports plus exports divided by nominal GDP. However, Alcalá and Ciccone (2004) point out that *real* openness (nominal imports plus exports divided by purchasing power parity (PPP) GDP) is the appropriate openness variable to use because it eliminates distortions in nominal openness induced by cross-country differences in relative prices of non-tradable products. Supposing that specialization increases productivity in the tradable sector more than it does in the non-tradable sector, then the relative price of non-tradable goods increases due to the Balassa-Samuelson effect. However, as non-tradable goods enter the calculation of GDP, nominal trade openness might not necessarily increase even though the increased specialization should have been echoed in an increasing trade openness. Alcalá and Ciccone (2004) argue that a monotonic relationship between specialization and openness is restored if one expresses trade as a percentage of PPP GDP instead of GDP. In the absence of data for real trade in PPP, Alcalá and Ciccone (2004) measure real openness as exports plus imports in US$ relative to GDP in US$ PPP. However, this measure mixes up nominal and real values in the numerator and denominator and is sensitive to the level of the exchange rate in the time at which real openness is measured.

Real openness is used as the independent variable in the regressions in Table 6, where real openness is based on the most recent PWT v9.0, which not only provides new data on real trade in PPP, but also improved income data (see, for an in-depth analysis, Feenstra *et al*., 2009 and Feenstra *et al*., 2015). The measure of income per capita we take from PWT v9.0 is the real GDP^E per capita. This measure is closest to the one available in previous versions of the PWT, including the PWT Mark5.6 used by FR, and it does not account for differences in the price of exports and imports. Data for 1985 are used in all regressions for comparability purposes, even if 2014 data are available.

The 1985-data are probably of better quality as data are often adjusted several years after they are first published.

Table 6. Estimates of the Income Equation using Real Openness

Second-stage results:	**Model (1)**		**Model (2)**	
	GDP v5.6	GDP^E	GDP v5.6	GDP^E
Real trade share$_i$	2.270***	2.364***	0.585	0.830*
	(0.492)	(0.505)	(0.398)	(0.434)
Ln population$_i$	0.251**	0.218*	0.089	0.071
	(0.102)	(0.113)	(0.070)	(0.084)
Ln area$_i$	0.087	0.144	-0.085	-0.013
	(0.103)	(0.104)	(0.075)	(0.079)
Distance to equator$_i$			3.945***	3.592***
			(0.396)	(0.402)
Obs.	96	96	96	96
First-stage regressions:				
$\widehat{T_i}$	8.075***	8.075***	7.463***	7.463***
	(1.439)	(1.439)	(2.161)	(2.161)
Partial R^2	0.335	0.335	0.255	0.255
KP rk Wald F-stat	31.478	31.478	11.922	11.922
Second-stage results:	**Model (3)**		**Model (4)**	
	GDP v5.6	GDP^E	GDP v5.6	GDP^E
Real trade share$_i$	0.753*	0.998**	0.919**	1.223**
	(0.413)	(0.452)	(0.430)	(0.506)
Ln population$_i$	0.055	0.042	-0.073	-0.049
	(0.072)	(0.084)	(0.093)	(0.115)
Ln area$_i$	-0.064	0.007	0.078	0.131*
	(0.080)	(0.085)	(0.068)	(0.077)
% Land in tropics$_i$	-1.486***	-1.339***		
	(0.179)	(0.192)		
Sub-Saharan Africa$_i$			-1.777***	-1.503***
			(0.216)	(0.261)
East Asia$_i$			-0.538**	-0.463*
			(0.225)	(0.262)
Latin America$_i$			-0.451	-0.453
			(0.318)	(0.351)
Obs.	96	96	96	96
First-stage regressions:				
$\widehat{T_i}$	8.020***	8.020***	7.480***	7.480***
	(1.860)	(1.860)	(1.681)	(1.681)
Partial R^2	0.285	0.285	0.237	0.237
KP *rk* Wald F-stat	18.586	18.586	19.796	19.796

Note: The country sample is the 98 sample excluding Papua New Guinea and Somalia, for which data are not available in PWT v9.0. $\widehat{T_i}$ is the predicted trade openness based on the total sample; i.e. including out-of-sample predictions. GDP v5.6 refers to the measure of real GDP per capita reported by the PWT Mark 5.6, while GDP^E refers to the measure constructed using the expenditure-side GDP at current PPPs reported in PWT v9.0. Real trade openness is from PWT v9.0. Heteroscedastic consistent standard errors are in parentheses. The standard errors in income regressions are corrected for the errors created from the generated regressors. The KP *rk* Wald F-stat is the Kleibergen-Paap *rk* Wald F-statistic. Exogenous variables are included in the first-stage regressions but not shown. *, **, *** Significant at 10, 5 and 1 percent, respectively.

Even though 33 years have passed since 1985, the per capita GDP data used by FR PWT Mark5.6 for 1985 has been improved in the PWT v9.0 in terms of retrospective adjustments made by national statistical agencies and PPP conversions. The second-stage regressions using per capita GDP from PWT Mark5.6 (data used by FR) and real GDP^E per capita in 1985 from PWT v9.0 as dependent variables are presented in Table 6. The coefficients of real openness are statistically significant in seven of the eight regressions, which is an improvement over the results in Table 2, where (nominal) openness is significantly positive in only two of the four cases. In online Appendix Table D3 we show that the significance of the coefficient of $\hat{T}_i^{TS}$ improves in 7 out of the 10 additional specifications relative to the regressions using the FR data in Table C1, when both real openness and more recent income data are used; thus strengthening the conclusion of FR that per capita income is positively related to trade. Finally, as shown in the online Appendix Table D4, the results from random generated instruments in Section 3.3 are robust to the use of real openness and new measures of GDP per capita, GDP^E.

5.3 Economic significance

Thus far we have not discussed the magnitude of the coefficient of trade openness. The average coefficient estimates of trade openness for Models 2-4, using only the IV- $\tilde{T}_i^{TS}$ regression results, is 0.73 (Table 2) when nominal trade openness is used, while it is 0.91 when real openness is used (Table 6). The size of the coefficients indicates that trade has been influential for the income growth over the past last globalization wave, 1960-2009, in the OECD countries. Using the updated data for 21 OECD countries of Madsen (2009), the nominal trade openness increased by approximately 10 percentage points as a simple average over the period 1960-2009; thus contributing to a 9.1% increase in real per capita income when the real openness elasticities are used. Conversely, the approximate 12 percentage point trade collapse over the period 1913-1932 resulted in an income contraction of 10.9% compared to what it would have been had trade openness stayed at the 1913 level. Thus, while these counterfactual simulations show that trade is influential for growth, the effects are comparatively small relative to technological progress, noting that are growth is driven by technological progress in steady state. Per capita income grew 209% over the period 1960-2009, on average, in the OECD countries (Madsen, 2009), suggesting that the 9.1% trade-induced growth has not made a comparatively large contribution when factoring in that the trade expansion in the period 1960-2009 was probably the largest 50-year expansion in the OECD countries' history.

6. Concluding remarks

Following the influential paper of FR the causal effects of trade openness on per capita income has not yet been resolved. In this paper we show that the diverging results, to a large degree, are driven by the choice of a seemingly innocent sampling procedure – so innocent that most papers do not even mention it. This paper shows that the statistical significance of trade openness in income regressions is highly sensitive to whether out-of-sample bilateral trade share predictions are included in the IV-set, because the instrument for trade openness is endogenous to income when out-of-sample predictions are omitted from the IV-set.

Using bilateral trade shares predicted from randomly generated geographical characteristics to form instruments for trade openness, we show that a significantly positive relationship between income and trade is spuriously created when *only* in-sample bilateral trade flows are included in the instrument set. However, the significance of randomly created trade openness disappears once the out-of-sample predictions are included in the IV-set, suggesting that the estimates can only be unbiased if out-of-sample predictions are included in the IV-set.

Why is an apparently innocent truncation of the IV-set so influential for the significance of trade openness in the income equation? The answer lies in the fact that the truncation of the IV-set to include in-sample bilateral trade only causes the instrument to capture each country's number of trading partners. We show analytically and empirically that a significantly positive relationship between income and the instrument for trade is spuriously created because high-income countries have more trade partners than low-income countries. However, inclusion of out-of-sample predictions in the sample eliminates this endogeneity bias. From this it can be concluded that the results of Rodriguez and Rodrik (2001), in which out-of-sample observations are included in the regressions, still stands.

Is this finding a setback for FR's finding of positive growth effects of trade? Based on the 98-sample regressions of Rodriguez and Rodrik (2001) it would seem so. However, we find positive income-effects of trade when the country sample is expanded and when real openness and income data from the most recent PWT are used - even when various controls are included in the regressions. The coefficients of trade openness are significantly positive in 12 out of 14 model specifications in which various controls are included, when the country sample is expanded, and in 11 out of 14 cases when openness is measured in real terms, even when the number of countries is limited to 98. As a precaution, it has to be noted that the statistical significance of openness in the income equation is generally not high, suggesting that the openness-income nexus is still not sufficiently strong to make firm conclusions about the causal effects of trade on income. Furthermore, counterfactual simulations suggest that the last globalization wave, over the period 1960-2009, contributed only a 9.1% increase in per capita income in the OECD countries, which is less than 5% of the 209% increase for the average OECD countries over this period. Thus, while trade is likely to enhance productivity, its contribution to growth is small relative to that of technological progress.

Our results have wide-spread implications for empirical trade modelling. An implication of our analysis is in that out-of-sample predictions from the bilateral trade equation should always be included in the instrument set for trade openness in regressions in which outcome variables are positively related to the number of trade partners, such as per capita income, investment, saving, education, R&D-intensity, etc. The same result applies to cross-border flows based on the FR framework, such as foreign direct investment, migration, foreign patenting, and portfolio investment, because the number of bilateral flows for these variables is also likely to be endogenous to income.

References

Alcalá, F., Ciccone, A., 2004. Trade and Productivity. *Quarterly Journal of Economics* 119, 613–646.

Alesina, A., Harnoss, J., Rapoport, H., 2016. Birthplace Diversity and Economic Prosperity. *Journal of Economic Growth* 21, 101–138.

Andersen, T.B., Dalgaard, C.-J., 2011. Flows of People, Flows of Ideas, and the Inequality of Nations. *Journal of Economic Growth* 16, 1–32.

Ardelean, A., León-Ledesma, M., Puzzello, L., 2017. Industry Volatility and International Trade. School of Economics, University of Kent, Working Paper No. 1709.

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

Di Giovanni, J., Levchenko, A.A., 2009. Trade Openness and Volatility. *Review of Economics and Statistics 91*, 558–585.

Djankov, S., La Porta, R., Lopez-De-Silanes, F., Shleifer A., 2002. The Regulation of Entry. *Quarterly Journal of Economics* 117, 1–37.

Edmonds, E. V., Pavcnik, N., 2006. International Trade and Child Labor: Cross-Country Evidence. *Journal of International Economics* 68, 115-140.

Feenstra, R.C., Heston, A., Timmer, M.P., Deng, H., 2009. Estimating Real Production and Expenditures Across Nations: A Proposal For Improving The Penn World Tables. *Review of Economics and Statistics* 91, 201–212.

Feenstra, R.C., Inklaar, R., Timmer, M.P., 2015. The Next Generation of the Penn World Table. *American Economic Review* 105, 3150–3182.

Feyrer, J., 2009. Trade and Income – Exploiting Time Series in Geography. National Bureau of Economic Research Working Paper No. 14910.

Frankel, J.A., Romer, D., 1999. Does Trade Cause Growth? *American Economic Review* 89, 379–399.

Frankel, J.A., Rose, A.K., 2005. Is Trade Good or Bad for the Environment? *Review of Economics and Statistics* 87, 85-91.

Gervais, A., 2015. Trade and Growth: A Gravity Approach. *Southern Economic Journal* 82, 453–470.

Harrison, A., Rodríguez-Clare, A., 2010. *Trade, Foreign Investment, and Industrial Policy for Developing Countries, in: Handbook of Development Economics*. Elsevier, pp. 4039–4214.

Helpman, E., Melitz, M., Rubinstein, Y., 2008. Estimating Trade Flows: Trading Partners and Trading Volumes. *Quarterly Journal of Economics* 123, 441–487.

Irwin, D.A., Terviö, M., 2002. Does Trade Raise Income?: Evidence from the Twentieth Century. *Journal of International Economics* 58, 1–18. doi:10.1016/S0022-1996(01)00164-7

Madsen, J.B., 2009. Trade Barriers, Openness, and Economic Growth. *Southern Economic Journal* 76, 397–418.

Mankiw, N.G., Romer, D., Weil, D.N., 1992. A Contribution to the Empirics of Economic Growth. *Quarterly Journal of Economics* 107, 407–437.

Melitz, M.J., 2003. The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. *Econometrica* 71, 1695–1725.

Noguer, M., Siscart, M., 2005. Trade Raises Income: A Precise and Robust Result. *Journal of International Economics* 65, 447–460.

Ortega, F., Peri, G., 2014. Openness and Income: The Roles of Trade and Migration. *Journal of International Economics* 92, 231–251.

Parente, P.M.D.C., Santos Silva, J.M.C., 2012. A Cautionary Note on Tests of Overidentifying Restrictions. *Economics Letters* 115, 314–317.

Pascali, L., 2017. The Wind of Change: Maritime Technology, Trade, and Economic Development. *American Economic Review* 107, 2821–2854.

Rodriguez, F., Rodrik, D., 2001. Trade Policy and Economic Growth: A Skeptic's Guide to the Cross-National Evidence, in: *NBER Macroeconomics Annual 2000*. MIT Press, pp. 261–338.

ONLINE APPENDIX

NOT TO BE PUBLISHED

Appendix A. Data

A.1 Bilateral data set

The bilateral data set includes bilateral data for the 98 countries from the Mankiw (1992) sample and 162 partner countries, i.e., each country has 161 partners. Despite relevant data available for a larger set of partner countries, the analysis follows FR and limits partner countries outside the sample to those countries whose population is greater than 100,000.

The 98 countries in the Mankiw (1992) sample include: Algeria, Angola, Argentina, Australia, Austria, Bangladesh, Belgium, Benin, Bolivia, **Botswana**, Brazil, Burkina Faso, Burundi, Cameroon, Canada, Central African Republic, Chad, Chile, Colombia, Congo, Costa Rica, Denmark, Dominican Republic, Ecuador, Egypt, El Salvador, Ethiopia, Finland, France, Germany (unified), Ghana, Greece, Guatemala, Haiti, Honduras, Hong Kong, India, Indonesia, Ireland, Israel, Italy, Ivory Coast, Jamaica, Japan, Jordan, Kenya, South Korea, Liberia, Madagascar, Malawi, Malaysia, Mali, Mauritania, Mauritius, Mexico, Morocco, Mozambique, Myanmar, Nepal, Netherlands, New Zealand, Nicaragua, Niger, Nigeria, Norway, Pakistan, Panama, Papua New Guinea, Paraguay, Peru, Philippines, Portugal, Rwanda, Senegal, Sierra Leone, Singapore, Somalia, South Africa, Spain, Sri Lanka, Sudan, Sweden, Switzerland, Syria, Tanzania, Thailand, Togo, Trinidad & Tobago, Tunisia, Turkey, U.K., U.S.A., Uganda, Uruguay, Venezuela, Zaire, Zambia and Zimbabwe.

The 162 partner countries include the 98 countries listed in the previous paragraph and: Afghanistan, Albania, Bahamas, Bahrain, Barbados, Belize, Bhutan*, Brunei Darussalam, Bulgaria, Cambodia, Cape Verde, China, Comoros, Cuba, Cyprus, Czechoslovakia, Djibouti, East Timor*, Eritrea*, Equatorial Guinea, Fiji, French Polynesia*, Gabon, Gambia, Guinea, Guinea-Bissau, Guyana, Hungary, Iceland, Iran, Iraq, Kuwait, Laos, Lebanon, Libya, **Lesotho**, Luxembourg, Maldives, Malta, Mongolia, **Namibia**, New Caledonia, North Korea, Oman, Poland, Puerto Rico*, Qatar, Reunion, Romania, Sao Tome and Principe, Saudi Arabia, Solomon Islands, St. Lucia, St. Vincent & Grenadines, Suriname, **Swaziland**, Taiwan*, U.S.S.R., United Arab Emirates, Vanuatu, Vietnam, Western Samoa, Yemen and Yugoslavia.

In the country lists above, countries in bold font are non-reporting in the DOTS but enter our dataset because of the symmetry imposed on bilateral trade flows. There is no record of bilateral trade for countries that are starred.

Bilateral trade data from the DOTS for the year 1985 is used to construct symmetric bilateral trade flows. Bilateral trade shares are calculated by dividing bilateral trade in nominal terms by the destination country's GDP. The latter is the product of real GDP per capita (base year 1985) and a country's population both from the Penn World Tables (PWT) Mark 5.6.

The real bilateral trade share, used in the robustness analysis, is the sum of real bilateral export share and real bilateral import share. The real bilateral export share is the nominal export share (equal to nominal exports as a share of GDP) divided by the reporter country's ratio of price level of exports to its price level of Output-side GDP. To get the real bilateral import share, the nominal import share is divided by the reporter's ratio of price level of imports to the price level of Output-side GDP. More formally:

$$real\ bilateral\ trade\ share_{ji} = \frac{exports_{ji}/GDP_i}{pl_{exports,i}/pl_{GDP^o,i}} + \frac{imports_{ji}/GDP_i}{pl_{imports,i}/pl_{GDP^o,i}}$$

All data for price levels ($pl_{x,i}$) are sourced from PWT version 9.0.

The data set is completed with data on area, bilateral distance, border and landlocked status from the CEPII GeoDist database. Population data are from the PWT Mark 5.6 and, when missing, the World Development Indicators (WDI).

A.2 Country data set

Real income per capita, trade openness as well as population data are taken from PWT Mark5.6. Area is sourced from CEPII. PWT v9.0 is the source for real GDP^E per capita, and the real export share and import share for 1985.

Data on the percentage of land or population in the tropics, and continents is from the Centre for International Development (CID). Latitude and distance to the equator are sourced from the CEPII. Legal origin is from La Porta *et al*. (2008) and, when missing, from the CIA World Factbook. The index of ethno-linguistic fractionalization is from Easterly and Levine (1997). Data on constraint on executive is from the Polity IV Project (2014). Finally, data on corruption and the quality of governance come from the International Country Risky Guide (ICRG) provided by the Political Risk Services Group.

A.3. Extended country sample

The extended sample consists of 147 countries and 166 partner countries. The extended sample includes four countries with populations less than 100,000 in 1985. They are Dominica, Grenada, Seychelles and Tonga.

Countries in the 147 sample include the 98 countries in the Mankiw (1992) sample and:
Bahamas, Bahrain, Barbados, Belize, Bulgaria, Cape Verde, China, Comoros, Cyprus, Czechoslovakia, Djibouti, Dominica, Fiji, Gabon, Gambia, Grenada, Guinea, Guinea-Bissau, Guyana, Hungary, Iceland, Iran, Iraq, Kuwait, Laos, **Lesotho**, Luxembourg, Malta, Mongolia, **Namibia**, Oman, Poland, Qatar, Reunion, Romania, Saudi Arabia, Seychelles, Solomon Islands, St.Lucia, St.Vincent & Grenadines, Suriname, **Swaziland**, Tonga, U.S.S.R., United Arab Emirates, Vanuatu, Western Samoa, Yemen, Yugoslavia.

The partner countries include the147 countries listed above and:
Afghanistan, Albania, Bhutan*, Brunei Darussalam, Cambodia, Cuba, East Timor*, Eritrea*, Equatorial Guinea, French Polynesia*, Lebanon, Libya, Maldives, New Caledonia, North Korea, Puerto Rico*, Sao Tome and Principe, Taiwan* and Vietnam.

In the country lists above, countries in bold font are non-reporting in the DOTS but enter our dataset because of the symmetry imposed on bilateral trade flows. There is no record of bilateral trade for countries that are starred.

References to online Appendix

Easterly, W., Levine, R., 1997. Africa's Growth Tragedy: Policies and Ethnic Divisions. *Quarterly Journal of Economics* 112, 1203–1250.

La Porta, R., Lopez-de-Silanes, F., Shleifer, A., 2008. The Economic Consequences of Legal Origins. *Journal of Economic Literature* 46, 285–332.

Mankiw, N.G., Romer, D., Weil, D.N., 1992. A Contribution to the Empirics of Economic Growth. *Quarterly Journal of Economics* 107, 407–437.

Mayer, T., Zignago, S., 2011. Notes on CEPII's Distances Measures: the GeoDist Database, CEPII Working Papers No. 2011-25.
Polity IV Project, 2014. Polity IV Project, Political Regime Characteristics and Transitions, 1800-2015. Centre for Systemic Peace, Virginia.

Table A1. Overview of variables used in the analysis

Variable	**Description**
Real GDP per capita	Real GDP per capita, chain weighted, US $, base year=1985.
Population	Total population.
Area	Area in km^2.
Bilateral Trade	Sum of bilateral exports and imports, in millions of US $.
Distance	Distance between two main cities, weighted for the geographic distribution of the population within the country, in km.
Landlocked	Dummy variable set equal to 1 for landlocked countries.
Border	Dummy variable set equal to 1 for country pairs sharing a border.
Latitude	Calculated as the latitude of the main city, scaled to take values between -1 and 1.
Distance to the equator	Calculated using the absolute value of the latitude, scaled to take values between 0 and 1.
% Land in tropics	The percentage of land area located in the tropics.
Continental Dummies	Dummy variables for Latin America, Sub-Saharan Africa, and East Asia.
Expenditure-side GDP	Expenditure-side real GDP at current PPPs (in mil. 2011US$)
Price level of Output-side GDP	Price level of GDP^O (PPP/XR); price level of USA GDPo in 2011=1
Price level of exports	Price level of exports, price level of USA GDP^o in 2011=1
Price level of imports	Price level of imports, price level of USA GDP^o in 2011=1
Real export share	Share of merchandise exports at current PPPs
Real import share	Share of merchandise imports at current PPPs
% Population in tropics	The percentage of the population living in a tropical area.
ICRG index	An index constructed as the sum of five variables: corruption, bureaucratic quality and rule of law, each multiplied by 5/3, as well as repudiation of contracts and expropriation risk. The index is normalized to vary between 0 and 1.
Corruption	Assessment of corruption within the political system; rescaled to take values between 0 and 1.
Executive constraint	Index of the extent to which decision making power of the executive is constrained by institutionalized procedure.
Ethno-linguistic fractionalization	Index that measures the probability that two randomly selected people from a given country do not belong to the same ethno-linguistic group.
Legal origin	Variable that takes on 1 if a country's legal origin is English, 2 if it is French, 3 if it is German and 4 if it is Scandinavian.

Appendix B. Full results for Table 3

Table B1. Estimates of the Income Equation using Randomized Instruments: Full Results

	Model (1)		Model (2)		Model (3)		Model (4)	
	IV- $\widetilde{T}_i^{TS}$	IV-$\widetilde{T}_i^{IS}$	IV- $\widetilde{T}_i^{TS}$	IV-$\widetilde{T}_i^{IS}$	IV- $\widetilde{T}_i^{TS}$	IV-$\widetilde{T}_i^{IS}$	IV- $\widetilde{T}_i^{TS}$	IV-$\widetilde{T}_i^{IS}$
Trade share$_i$	-4.000	6.691	2.266	3.632	0.531	4.289	-1.938	6.062
	(88.754)	(0.892)	(89.180)	(0.676)	(62.945)	(0.715)	(66.014)	(1.297)
	[9]	[1000]	[7]	[997]	[11]	[1000]	[8]	[983]
	{1}	{1000}	{2}	{989}	{3}	{997}	{3}	{960}
Ln population$_i$	-0.078	0.683	0.242	0.352	0.064	0.373	-0.433	0.766
	(6.317)	(0.064)	(7.185)	(0.055)	(5.185)	(0.059)	(9.893)	(0.194)
	[125]	[1000]	[7]	[962]	[3]	[940]	[0]	[948]
	{64}	{1000}	{1}	{343}	{0}	{139}	{0}	{857}
Ln area$_i$	-0.631	0.553	0.096	0.244	-0.093	0.315	-0.137	0.399
	(9.832)	(0.099)	(9.662)	(0.073)	(6.834)	(0.078)	(4.429)	(0.087)
	[3]	[505]	[21]	[540]	[11]	[608]	[0]	[755]
	{0}	{19}	{6}	{38}	{4}	{68}	{0}	{162}
Distance to equator$_i$			3.692	3.314				
			(24.668)	(0.187)				
			[644]	[992]				
			{607}	{990}				
% Land in tropics$_i$					-1.590	-1.219		
					(6.211)	(0.071)		
					[647]	[994]		
					{604}	{991}		
Sub-Saharan Africa$_i$							-2.313	-1.029
							(10.594)	(0.208)
							[570]	[834]
							{521}	{727}
East Asia$_i$							0.448	-2.803
							(26.826)	(0.527)
							[41]	[961]
							{15}	{890}
Latin America$_i$							-1.358	0.996
							(19.428)	(0.382)
							[54]	[188]
							{22}	{0}
Observations	98	98	98	98	98	98	98	98

Notes. The table reports average values from 1000 replications. The standard deviation of this average is reported in parentheses. For the estimated coefficients, the number of replications that produce an estimate significant at least at the 10% level is in [square brackets] and the number of replications in which the estimate is significant at least at the 5% level is in {curly brackets}.

Appendix C. Regressions of the Income Equation with Various Controls

Table C1. Estimates of the Income Equation using Actual Data: Additional Controls Included

Second-stage results:	Model (5)		Model (6)		Model (7)		Model (8)		Model (9)	
	IV- $\hat{T}_i^{TS}$	IV-$\hat{T}_i^{IS}$	IV- $\hat{T}_i^{TS}$	IV-$\hat{T}_i^{IS}$	IV- $\hat{T}_i^{TS}$	IV-$\hat{T}_i^{IS}$	IV- $\hat{T}_i^{TS}$	IV-$\hat{T}_i^{IS}$	IV- $\hat{T}_i^{TS}$	IV-$\hat{T}_i^{IS}$
Trade share$_i$	0.693*	1.023***	0.613	1.078***	0.451	0.710**	0.524	0.909***	0.841*	1.110***
	(0.386)	(0.353)	(0.440)	(0.407)	(0.329)	(0.288)	(0.379)	(0.330)	(0.490)	(0.413)
Ln population$_i$	-0.030	0.001	0.067	0.108	0.006	0.028	-0.026	0.010	-0.024	0.017
	(0.077)	(0.081)	(0.085)	(0.088)	(0.066)	(0.066)	(0.073)	(0.076)	(0.097)	(0.090)
Ln area$_i$	0.146**	0.176**	-0.081	-0.033	0.026	0.053	0.060	0.097	0.124*	0.140**
	(0.070)	(0.071)	(0.083)	(0.082)	(0.063)	(0.063)	(0.069)	(0.069)	(0.069)	(0.070)
Latitude$_i$	0.609**	0.568**	0.058	0.010			0.212	0.169	0.296	0.280
	(0.273)	(0.287)	(0.328)	(0.339)			(0.260)	(0.271)	(0.379)	(0.396)
% Population in tropics$_i$	-2.012***	-1.979***			-1.304***	-1.290***	-1.480***	-1.447***	-1.293***	-1.276***
	(0.203)	(0.208)			(0.219)	(0.222)	(0.225)	(0.232)	(0.285)	(0.284)
Distance to equator$_i$					2.483***	2.431***				
					(0.368)	(0.370)				
% Land in tropics$_i$			-1.565***	-1.533***			-0.719***	-0.712***		
			(0.186)	(0.185)			(0.206)	(0.207)		
Sub-Saharan Africa$_i$									-0.865***	-0.839**
									(0.329)	(0.333)
East Asia$_i$									-0.413	-0.529
									(0.394)	(0.374)
Latin America$_i$									-0.123	-0.054
									(0.312)	(0.303)
Observations	98	98	98	98	98	98	98	98	98	98
First stage regressions:										
$\tilde{T}_i$	8.655***	9.554***	8.853***	9.659***	7.602***	8.487***	8.825***	9.649***	7.364***	8.597***
	(2.216)	(2.302)	(2.211)	(2.294)	(1.935)	(2.091)	(2.222)	(2.327)	(1.787)	(1.831)
Partial R^2	0.310	0.369	0.317	0.374	0.282	0.336	0.315	0.371	0.244	0.328
KP *rk* Wald *F*-stat	15.253	17.223	16.040	17.734	15.431	16.468	15.774	17.193	16.988	22.057

Note. The dependent variable is log of real GDP per capita reported by PWT Mark 5.6 for the year 1985. Heteroscedastic consistent standard errors are in parentheses. The standard errors in the second-stage are corrected for the errors created from the generated regressors. $\hat{T}_i^{IS}$ is the predicted trade openness based on in-sample observations only. $\hat{T}_i^{TS}$ is the predicted trade openness based on the total sample, i.e., including out-of-sample predictions. The KP *rk* Wald *F*-stat is the Kleibergen-Paap *rk* Wald *F*-statistic. Exogenous variables are included in the first-stage regressions but not shown. *,**, *** Significant at 10, 5 and 1 percent, respectively.

Table C2. Estimates of the Income Equation using Actual Data: Additional Controls Included

Second-stage results:	Model (10)		Model (11)		Model (12)		Model (13)		Model (14)	
	IV- $\hat{T}_i^{TS}$	IV-$\hat{T}_i^{IS}$	IV- $\hat{T}_i^{TS}$	IV-$\hat{T}_i^{IS}$	IV- $\hat{T}_i^{TS}$	IV-$\hat{T}_i^{IS}$	IV- $\hat{T}_i^{TS}$	IV-$\hat{T}_i^{IS}$	IV- $\hat{T}_i^{TS}$	IV-$\hat{T}_i^{IS}$
Trade share$_i$	0.133	0.356	0.094	0.396	0.782**	0.895***	0.773*	1.099***	0.823**	1.097***
	(0.351)	(0.322)	(0.402)	(0.346)	(0.306)	(0.284)	(0.406)	(0.377)	(0.393)	(0.352)
Ln population$_i$	-0.066	-0.058	-0.043	-0.034	-0.007	0.005	0.032	0.062	0.021	0.051
	(0.053)	(0.054)	(0.058)	(0.058)	(0.077)	(0.077)	(0.078)	(0.083)	(0.087)	(0.091)
Ln area$_i$	0.074	0.108	0.066	0.112	0.132**	0.140**	0.146*	0.178**	0.137**	0.159**
	(0.070)	(0.069)	(0.083)	(0.077)	(0.059)	(0.060)	(0.076)	(0.077)	(0.069)	(0.070)
Latitude$_i$	0.255	0.287	0.450**	0.475**	0.500**	0.483**	0.477*	0.426	0.322	0.258
	(0.189)	(0.198)	(0.221)	(0.224)	(0.209)	(0.213)	(0.278)	(0.289)	(0.323)	(0.335)
% Population in tropics$_i$	-1.499***	-1.527***	-1.622***	-1.647***	-1.268***	-1.261***	-1.715***	-1.661***	-1.944***	-1.915***
	(0.182)	(0.180)	(0.203)	(0.198)	(0.207)	(0.209)	(0.254)	(0.255)	(0.221)	(0.226)
IGRC-Index$_i$	2.425***	2.259***								
	(0.343)	(0.354)								
Corruption$_i$			1.669***	1.520***						
			(0.278)	(0.277)						
Executive constraint$_i$					0.210***	0.210***				
					(0.027)	(0.028)				
Ethno-ling. fract.$_i$							-0.716**	-0.785**		
							(0.319)	(0.319)		
Legal Origin$_i$									0.230*	0.255**
									(0.119)	(0.119)
Observations	90	90	90	90	94	94	95	95	96	96
First stage regressions:										
$\tilde{T}_i$	6.812***	7.614***	6.996***	8.003***	8.063***	8.949***	8.786***	9.748***	8.895***	9.813***
	(1.708)	(1.903)	(1.679)	(1.842)	(2.675)	(2.859)	(2.204)	(2.297)	(2.220)	(2.277)
Partial R^2	0.238	0.279	0.228	0.281	0.273	0.328	0.319	0.383	0.334	0.399
KP *rk* Wald *F*-stat	15.905	16.017	17.363	18.884	9.084	9.797	15.890	18.013	16.053	18.578

Note. The dependent variable is log of real GDP per capita reported by PWT Mark 5.6 for the year 1985. Heteroscedastic consistent standard errors are in parentheses. The standard errors in the second-stage are corrected for the errors created from the generated regressors. $\hat{T}_i^{IS}$ is the predicted trade openness based on in-sample observations only. $\hat{T}_i^{TS}$ is the predicted trade openness based on the total sample, i.e., including out-of-sample predictions. The KP *rk* Wald *F*-stat is the Kleibergen-Paap *rk* Wald *F*-statistic. Exogenous variables are included in the first-stage regressions but not shown. *, **, *** Significant at 10, 5 and 1 percent, respectively.

Table C3. Estimates of the Income Equation using Randomized Instruments: Additional Controls Included

Second-stage results:	Model (5)		Model (6)		Model (7)		Model (8)		Model (9)	
	IV- $\widetilde{T_i^{TS}}$	IV-$\widetilde{T_i^{IS}}$	IV- $\widetilde{T_i^{TS}}$	IV-$\widetilde{T_i^{IS}}$	IV- $\widetilde{T_i^{TS}}$	IV-$\widetilde{T_i^{IS}}$	IV- $\widetilde{T_i^{TS}}$	IV-$\widetilde{T_i^{IS}}$	IV- $\widetilde{T_i^{TS}}$	IV-$\widetilde{T_i^{IS}}$
Trade share$_i$	7.533	4.176	2.925	4.407	-2.988	3.302	0.946	3.830	1.954	5.601
	(213.193)	(0.751)	(86.356)	(0.766)	(115.918)	(0.611)	(36.876)	(0.693)	(34.725)	(1.266)
	[7]	[999]	[10]	[1000]	[9]	[998]	[9]	[999]	[11]	[981]
	{2}	{992}	{3}	{994}	{1}	{989}	{2}	{992}	{2}	{958}
Ln population$_i$	0.613	0.298	0.272	0.403	-0.285	0.247	0.013	0.283	0.145	0.699
	(20.039)	(0.071)	(7.643)	(0.068)	(9.799)	(0.052)	(3.449)	(0.065)	(5.273)	(0.192)
	[0]	[1]	[1]	[549]	[0]	[10]	[0]	[1]	[0]	[914]
	{0}	{0}	{0}	{1}	{0}	{0}	{0}	{0}	{0}	{716}
Ln area$_i$	0.776	0.467	0.156	0.309	-0.328	0.319	0.100	0.371	0.192	0.417
	(19.664)	(0.069)	(8.876)	(0.079)	(11.914)	(0.063)	(3.463)	(0.065)	(2.139)	(0.078)
	[44]	[993]	[9]	[351]	[0]	[991]	[3]	[973]	[40]	[856]
	7.533	4.176	2.925	4.407	-2.988	3.302	0.946	3.830	1.954	5.601
Latitude$_i$	-0.232	0.180	-0.183	-0.338			0.165	-0.161	0.230	0.017
	(26.225)	(0.092)	(9.017)	(0.080)			(4.164)	(0.078)	(2.037)	(0.074)
	[247]	[0]	[0]	[0]			[0]	[0]	[0]	[0]
	{123}	{0}	{0}	{0}			{0}	{0}	{0}	{0}
% Population in tropics$_i$	-1.326	-1.663			-1.501	-1.141	-1.444	-1.192	-1.224	-0.998
	(21.392)	(0.075)			(6.632)	(0.035)	(3.211)	(0.060)	(2.150)	(0.078)
	[694]	[998]			[701]	[988]	[653]	[968]	[655]	[344]
	{656}	{993}			{658}	{970}	{608}	{894}	{606}	{152}
Distance to equator$_i$					3.176	1.908				
					(23.366)	(0.123)				
					[587]	[818]				
					{536}	{468}				
% Land in tropics$_i$			-1.409	-1.309			-0.712	-0.662		
			(5.826)	(0.052)			(0.641)	(0.012)		
			[704]	[999]			[653]	[579]		
			{682}	{995}			{585}	{155}		

Continues

Second-stage results:	Model (5)		Model (6)		Model (7)		Model (8)		Model (9)	
	IV- $\widetilde{T_i^{TS}}$	IV-$\widetilde{T_i^{IS}}$	IV- $\widetilde{T_i^{TS}}$	IV-$\widetilde{T_i^{IS}}$	IV- $\widetilde{T_i^{TS}}$	IV-$\widetilde{T_i^{IS}}$	IV- $\widetilde{T_i^{TS}}$	IV-$\widetilde{T_i^{IS}}$	IV- $\widetilde{T_i^{TS}}$	IV-$\widetilde{T_i^{IS}}$
Sub-Saharan Africa$_i$									-0.758	-0.408
									(3.334)	(0.122)
									[419]	[1]
									{306}	{0}
East Asia$_i$									-0.895	-2.474
									(15.041)	(0.549)
									[4]	[963]
									{0}	{870}
Latin America$_i$									0.163	1.101
									(8.932)	(0.326)
									[0]	[121]
									{0}	{0}
Observations	98	98	98	98	98	98	98	98	98	98
First stage regressions:										
$\tilde{T}_i$	6.645	27.750	6.396	28.326	6.785	30.220	6.661	28.578	6.153	20.743
	(26.091)	(5.795)	(25.892)	(5.663)	(26.004)	(6.490)	(26.252)	(5.962)	(24.938)	(4.754)
	[109]	[998]	[109]	[998]	[111]	[997]	[110]	[997]	[117]	[978]
	{59}	{986}	{64}	{988}	{61}	{977}	{60}	{987}	{63}	{935}
Partial R^2	0.013	0.105	0.013	0.110	0.013	0.111	0.013	0.107	0.013	0.072
	(0.019)	(0.031)	(0.019)	(0.032)	(0.019)	(0.033)	(0.020)	(0.032)	(0.019)	(0.027)
KP *rk* Wald *F*-stat	1.156	10.928	1.145	11.415	1.153	9.660	1.153	10.819	1.197	9.924
	<5>	<624>	<4>	<681>	<6>	<464>	<5>	<611>	<6>	<487>
CD Wald *F*-stat	1.235	10.928	1.218	11.556	1.232	11.660	1.234	11.068	1.217	6.944

Notes. The table reports average values from 1000 replications. The standard deviation of this average is reported in parentheses. For the estimated coefficients, the number of replications that produce an estimate significant at least at the 10% level is in [square brackets] and the number of replications in which the estimate is significant at least at the 5% level is in {curly brackets}. The number of times the KP *rk* Wald *F*-stat is greater than 10 is in <angle brackets>. Exogenous variables are included in the first stage regressions but not shown.

Table C4. Estimates of the Income Equation using Randomized Instruments: Additional Controls Included

Second stage results:	Model (10)		Model (11)		Model (12)		Model (13)		Model (14)	
	IV- $\widetilde{T_i^{TS}}$	IV-$\widetilde{T_i^{IS}}$	IV- $\widetilde{T_i^{TS}}$	IV-$\widetilde{T_i^{IS}}$	IV- $\widetilde{T_i^{TS}}$	IV-$\widetilde{T_i^{IS}}$	IV- $\widetilde{T_i^{TS}}$	IV-$\widetilde{T_i^{IS}}$	IV- $\widetilde{T_i^{TS}}$	IV-$\widetilde{T_i^{IS}}$
Trade share$_i$	-1.294	6.503	-2.268	4.353	-2.285	2.992	-0.822	3.921	0.774	4.006
	(30.720)	(35.110)	(94.978)	(2.799)	(76.557)	(0.656)	(50.191)	(0.703)	(97.508)	(0.683)
	[0]	[473]	[1]	[874]	[11]	[988]	[7]	[999]	[6]	[998]
	{0}	{173}	{0}	{764}	{1}	{957}	{2}	{991}	{3}	{992}
Ln population$_i$	-0.117	0.160	-0.110	0.079	-0.333	0.229	-0.115	0.323	0.016	0.371
	(1.091)	(1.246)	(2.712)	(0.080)	(8.153)	(0.070)	(4.643)	(0.065)	(10.710)	(0.075)
	[0]	[0]	[0]	[0]	[0]	[0]	[0]	[6]	[0]	[151]
	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{0}	{1}
Ln area$_i$	-0.142	1.041	-0.293	0.713	-0.088	0.291	-0.011	0.457	0.133	0.397
	(4.662)	(5.328)	(14.434)	(0.425)	(5.498)	(0.047)	(4.955)	(0.069)	(7.957)	(0.056)
	[1]	[451]	[5]	[840]	[67]	[968]	[35]	[992]	[35]	[991]
	{0}	{186}	{0}	{678}	{21}	{827}	{8}	{961}	{5}	{936}
Latitude$_i$	0.046	1.187	0.255	0.800	0.947	0.177	0.729	-0.020	0.333	-0.415
	(4.495)	(5.138)	(7.818)	(0.230)	(11.172)	(0.096)	(7.921)	(0.111)	(22.573)	(0.158)
	[0]	[0]	[193]	[17]	[234]	[0]	[74]	[0]	[2]	[0]
	{0}	{0}	{38}	{0}	{141}	{0}	{6}	{0}	{0}	{0}
% Population in tropics$_i$	-1.319	-2.302	-1.425	-1.977	-1.454	-1.133	-1.976	-1.199	-1.949	-1.601
	(3.873)	(4.426)	(7.925)	(0.234)	(4.646)	(0.040)	(8.223)	(0.115)	(10.496)	(0.073)
	[631]	[888]	[716]	[979]	[722]	[994]	[585]	[967]	[685]	[998]
	{585}	{866}	{687}	{973}	{691}	{989}	{541}	{939}	{653}	{996}
IGRC-Index$_i$	3.489	-2.322								
	(22.895)	(26.166)								
	[291]	[14]								
	{226}	{6}								
Corruption$_i$			2.833	-0.431						
			(46.828)	(1.380)						
			[253]	[6]						
			{196}	{5}						

Continues

Second stage results:	Model (10)		Model (11)		Model (12)		Model (13)		Model (14)	
	IV- $\widetilde{T_i^{TS}}$	IV-$\widetilde{T_i^{IS}}$	IV- $\widetilde{T_i^{TS}}$	IV-$\widetilde{T_i^{IS}}$	IV- $\widetilde{T_i^{TS}}$	IV-$\widetilde{T_i^{IS}}$	IV- $\widetilde{T_i^{TS}}$	IV-$\widetilde{T_i^{IS}}$	IV- $\widetilde{T_i^{TS}}$	IV-$\widetilde{T_i^{IS}}$
Executive constraint$_i$					0.231	0.196				
					(0.508)	(0.004)				
					[773]	[999]				
					{741}	{997}				
Ethno-ling. fract.$_i$							-0.378	-1.384		
							(10.640)	(0.149)		
							[172]	[999]		
							{77}	{985}		
Legal Origin$_i$									0.226	0.514
									(8.704)	(0.061)
									[87]	[999]
									{22}	{995}
Observations	90	90	90	90	94	94	95	95	96	96
First-stage regressions-										
$\tilde{T}_i$	3.123	14.443	3.651	20.131	6.694	27.736	6.177	28.711	6.584	28.198
	(24.242)	(5.918)	(25.086)	(6.145)	(26.033)	(6.481)	(27.032)	(6.138)	(26.167)	(5.747)
	[106]	[463]	[113]	[865]	[115]	[977]	[106]	[995]	[114]	[997]
	{55}	{202}	{62}	{755}	{55}	{893}	{54}	{984}	{55}	{989}
Partial R^2	0.013	0.030	0.013	0.056	0.014	0.096	0.013	0.110	0.013	0.110
	(0.019)	(0.020)	(0.019)	(0.026)	(0.020)	(0.033)	(0.020)	(0.033)	(0.020)	(0.032)
KP *rk* Wald *F*-stat	1.095	2.762	1.135	5.891	1.157	6.097	1.133	11.103	1.151	10.555
	<3>	<2>	<2>	<63>	<3>	<22>	<3>	<633>	<3>	<594>

Notes. The table reports average values from 1000 replications. The standard deviation of this average is reported in parentheses. For the estimated coefficients, the number of replications that produce an estimate significant at least at the 10% level is in [square brackets] and the number of replications in which the estimate is significant at least at the 5% level is in {curly brackets}. The number of times the KP *rk* Wald *F*-stat is greater than 10 is in <angle brackets>. Exogenous variables are included in the first stage regressions but not shown.

Appendix D. Additional Results

Table D1. Estimates of the Income Equation using Actual Data: Additional Controls Included with a Larger Sample of Countries

Second-stage results:	Model (5)	Model (6)	Model (7)	Model (8)	Model (9)
	IV- $\hat{T}_i^{TS}$	IV- $\hat{T}_i^{TS}$	IV- $\hat{T}_i^{TS}$	IV- $\hat{T}_i^{TS}$	IV- $\hat{T}_i^{TS}$
Trade share$_i$	1.451***	0.789	1.483**	1.279**	1.244**
	(0.497)	(0.525)	(0.677)	(0.514)	(0.491)
Ln population$_i$	-0.026	0.020	0.029	-0.024	-0.006
	(0.089)	(0.071)	(0.095)	(0.086)	(0.093)
Ln area$_i$	0.219***	-0.005	0.151*	0.168**	0.151**
	(0.080)	(0.066)	(0.090)	(0.085)	(0.069)
Latitude$_i$	0.762**	0.407		0.535	0.161
	(0.307)	(0.352)		(0.354)	(0.343)
% Population in tropics$_i$	-1.738***		-1.247***	-1.379***	-1.112***
	(0.227)		(0.254)	(0.248)	(0.249)
Distance to equator$_i$			1.869***		
			(0.461)		
% Land in tropics$_i$		-1.253***		-0.457*	
		(0.217)		(0.260)	
Sub-Saharan Africa$_i$					-1.058***
					(0.256)
East Asia$_i$					-0.610*
					(0.352)
Latin America$_i$					-0.163
					(0.253)
Obs.	128	146	128	128	128
First-stage regressions:					
$\tilde{T}_i$	4.499***	3.758***	3.745***	4.519***	4.172***
	(1.039)	(0.911)	(0.996)	(1.102)	(0.877)
Partial R^2	0.176	0.144	0.144	0.173	0.145
KP *rk* Wald *F*-stat	18.742	17.023	14.150	16.811	22.619

Note. The dependent variable is log of real GDP per capita reported by PWT Mark 5.6 for the year 1985. Heteroscedastic consistent standard errors are in parentheses. The standard errors in the second-stage are corrected for the errors created from the generated regressors. $\hat{T}_i^{TS}$ is the predicted trade openness based on the total sample; i.e., including out-of-sample predictions. The KP *rk* Wald *F*-stat is the Kleibergen-Paap *rk* Wald *F*-statistic. Exogenous variables are included in the first-stage first-stage regressions but not shown. *, **, *** Significant at 10, 5 and 1 percent, respectively.

Table D1. (Cont'd)

Second-stage results:	Model (10)	Model (11)	Model (12)	Model (13)	Model (14)
	IV- $\hat{T}_i^{TS}$	IV- $\hat{T}_i^{TS}$	IV- $\hat{T}_i^{TS}$	IV- $\hat{T}_i^{TS}$	IV- $\hat{T}_i^{TS}$
Trade share$_i$	0.770	0.841*	1.415**	1.139***	1.588***
	(0.531)	(0.504)	(0.688)	(0.413)	(0.558)
Ln population$_i$	-0.168***	-0.157***	-0.050	0.029	0.007
	(0.057)	(0.057)	(0.109)	(0.074)	(0.103)
Ln area$_i$	0.222**	0.226**	0.223***	0.208**	0.215**
	(0.090)	(0.093)	(0.082)	(0.083)	(0.084)
Latitude$_i$	0.486**	0.605***	0.699***	0.668**	0.525
	(0.212)	(0.217)	(0.264)	(0.299)	(0.350)
% Population in tropics$_i$	-1.681***	-1.755***	-1.282***	-1.617***	-1.689***
	(0.214)	(0.213)	(0.239)	(0.266)	(0.246)
IGRC-Index$_i$	1.622***				
	(0.516)				
Corruption$_i$		1.069***			
		(0.359)			
Executive constraint$_i$			0.167***		
			(0.032)		
Ethno-ling. fract.$_i$				-0.743**	
				(0.299)	
Legal Origin$_i$					0.185
					(0.123)
Obs.	112	112	121	104	124
First-stage regressions:					
$\tilde{T}_i$	3.810***	3.955***	3.815***	5.042***	4.445***
	(0.840)	(0.807)	(1.101)	(1.216)	(1.047)
Partial R^2	0.168	0.169	0.138	0.226	0.174
KP *rk* Wald *F*-stat	20.565	24.026	11.998	17.185	18.015

Note. Heteroscedastic consistent standard errors are in parentheses. The standard errors in the second-stage are corrected for the errors created from the generated regressors. $\hat{T}_i^{TS}$ is the predicted trade openness based on the total sample; i.e., including out-of-sample predictions. The KP *rk* Wald *F*-stat is the Kleibergen-Paap *rk* Wald *F*-statistic. Exogenous variables are included in the first-stage first-stage regressions but not shown. *, **, *** Significant at 10, 5 and 1 percent, respectively.

Table D2. Estimates of Income Equation using Randomized Instruments: Larger Sample of Countries

Second -tage results:	Model (1)		Model (2)		Model (3)		Model (4)	
	IV- $\widetilde{T_i^{TS}}$	IV-$\widetilde{T_i^{IS}}$	IV- $\widetilde{T_i^{TS}}$	IV-$\widetilde{T_i^{IS}}$	IV- $\widetilde{T_i^{TS}}$	IV-$\widetilde{T_i^{IS}}$	IV- $\widetilde{T_i^{TS}}$	IV-$\widetilde{T_i^{IS}}$
Trade share$_i$	1.335	6.638	-2.710	4.561	1.889	5.267	2.111	4.697
	(82.041)	(0.628)	(71.290)	(0.593)	(64.782)	(0.668)	(42.970)	(0.501)
	[8]	[1000]	[5]	[998]	[3]	[998]	[6]	[1000]
	{1}	{1000}	{0}	{993}	{0}	{994}	{3}	{1000}
Ln population$_i$	0.205	0.591	-0.235	0.364	0.133	0.424	0.198	0.485
	(5.965)	(0.046)	(5.873)	(0.049)	(5.591)	(0.058)	(4.771)	(0.056)
	[31]	[1000]	[0]	[25]	[0]	[9]	[1]	[1000]
	{6}	{980}	{0}	{0}	{0}	{0}	{0}	{988}
Ln area$_i$	-0.014	0.417	-0.303	0.269	0.066	0.327	0.115	0.272
	(6.671)	(0.051)	(5.606)	(0.047)	(4.993)	(0.051)	(2.622)	(0.031)
	[2]	[959]	[3]	[935]	[2]	[921]	[2]	[999]
	{0}	{145}	{1}	{141}	{0}	{65}	{0}	{938}
Distance to equator$_i$			4.844	2.507				
			(22.909)	(0.191)				
			[581]	[957]				
			{537}	{855}				
% Land in tropics$_i$					-1.223	-0.743		
					(9.200)	(0.095)		
					[541]	[294]		
					{481}	{102}		
Sub-Saharan Africa$_i$							-1.511	-1.284
							(3.779)	(0.044)
							[693]	[1000]
							{660}	{1000}
East Asia$_i$							-1.114	-1.966
							(14.155)	(0.165)
							[58]	[1000]
							{25}	{911}
Latin America$_i$							-0.248	0.191
							(7.287)	(0.085)
							[117]	[0]
							{60}	{0}
Observations	147	147	147	147	146	146	147	147
First stage regression (selected results):								
$\tilde{T}_i$	3.343	21.610	3.168	21.744	3.092	19.975	3.102	22.049
	(21.410)	(2.713)	(20.721)	(3.436)	(20.878)	(3.003)	(20.485)	(3.179)
	[115]	[1000]	[105]	[998]	[106]	[998]	[108]	[1000]
	{57}	{1000}	{56}	{984}	{57}	{994}	{52}	{1000}
Partial R^2	0.007	0.107	0.007	0.082	0.007	0.077	0.008	0.099
	(0.011)	(0.018)	(0.011)	(0.019)	(0.011)	(0.018)	(0.011)	(0.019)
KP *rk* Wald *F*-stat	1.079	15.081	1.055	7.015	1.044	7.448	1.114	12.268
	<4>	<976>	<3>	<20>	<2>	<31>	<4>	<844>

Notes. The table reports average values from 1000 replications. The standard deviation of this average is reported in parentheses. For the estimated coefficients, the number of replications that produce an estimate significant at least at the 10% level is in [square brackets] and the number of replications in which the estimate is significant at least at the 5% level is in {curly brackets}. The number of times the KP *rk* Wald *F*-stat is greater than 10 is in <angle brackets>. Exogenous variables are included in the first stage regressions but not shown.

Table D3. Estimates of the Income Equation using Actual Data: Additional Controls Included, using Real Trade Openness and GDP^E

Second stage results:	Model (5)	Model (6)	Model (7)	Model (8)	Model (9)
	IV- $\hat{T}_i^{TS}$	IV- $\hat{T}_i^{TS}$	IV- $\hat{T}_i^{TS}$	IV- $\hat{T}_i^{TS}$	IV- $\hat{T}_i^{TS}$
Real Trade share$_i$	0.888*	0.863*	0.798*	0.753	0.928*
	(0.461)	(0.493)	(0.419)	(0.470)	(0.540)
Ln population$_i$	-0.069	0.020	-0.001	-0.055	-0.096
	(0.087)	(0.088)	(0.084)	(0.086)	(0.113)
Ln area$_i$	0.189***	0.007	0.079	0.112	0.182**
	(0.072)	(0.087)	(0.072)	(0.075)	(0.072)
Latitude$_i$	0.740**	0.264		0.386	0.557
	(0.321)	(0.325)		(0.290)	(0.415)
% Population in tropics$_i$	-1.585***		-0.945***	-1.111***	-1.034***
	(0.238)		(0.257)	(0.273)	(0.401)
Distance to equator$_i$			2.344***		
			(0.466)		
% Land in tropics$_i$		-1.287***		-0.658***	
		(0.207)		(0.243)	
Sub-Saharan Africa$_i$					-0.649
					(0.432)
East Asia$_i$					-0.090
					(0.350)
Latin America$_i$					-0.076
					(0.380)
Obs.	96	96	96	96	96
First stage regression (selected results):					
$\tilde{T}_i$	8.811***	9.025***	7.440***	8.879***	7.863***
	(2.449)	(2.455)	(2.161)	(2.467)	(2.079)
Partial R^2	0.300	0.303	0.255	0.300	0.243
KP *rk* Wald *F*-stat	12.945	13.510	11.849	12.954	14.304

Note: The country sample is the 98 sample excluding Papua New Guinea and Somalia, for which data are not available in PWT v9.0. $\hat{T}_i$ is the predicted trade openness based on the total sample; i.e. including out-of-sample predictions. GDP^E refers to the measure constructed using the expenditure-side GDP at current PPPs reported in PWT v9.0 for the year 1985. Real trade openness is from PWT v9.0. Heteroscedastic consistent standard errors are in parentheses. The standard errors in income regressions are corrected for the errors created from the generated regressors. The KP *rk* Wald *F*-stat is the Kleibergen-Paap *rk* Wald *F*-statistic. Exogenous variables are included in the first-stage regressions but not shown. *, **, *** Significant at 10, 5 and 1 percent, respectively.

Table D3. (Cont'd)

Second stage results:	Model (10)	Model (11)	Model (12)	Model (13)	Model (14)
	IV- $\hat{T}_i^{TS}$	IV- $\hat{T}_i^{TS}$	IV- $\hat{T}_i^{TS}$	IV- $\hat{T}_i^{TS}$	IV- $\hat{T}_i^{TS}$
Real trade share$_i$	0.131	0.037	0.851***	0.985**	1.008**
	(0.404)	(0.486)	(0.321)	(0.477)	(0.464)
Ln population$_i$	-0.086	-0.050	-0.057	0.001	-0.022
	(0.073)	(0.073)	(0.095)	(0.084)	(0.092)
Ln area$_i$	0.084	0.064	0.171***	0.186**	0.178**
	(0.073)	(0.089)	(0.062)	(0.078)	(0.071)
Latitude$_i$	0.306	0.526*	0.664**	0.605*	0.536
	(0.236)	(0.269)	(0.261)	(0.337)	(0.395)
% Population in tropics$_i$	-1.070***	-1.183***	-0.816***	-1.266***	-1.495***
	(0.200)	(0.233)	(0.273)	(0.305)	(0.249)
IGRC-Index$_i$	2.872***				
	(0.455)				
Corruption$_i$		2.115***			
		(0.395)			
Executive constraint$_i$			0.211***		
			(0.037)		
Ethno-ling. fract.$_i$				-0.690**	
				(0.349)	
Legal Origin$_i$					0.167
					(0.125)
Obs.	88	88	92	93	95
First stage regression (selected results):					
$\tilde{T}_i$	7.028***	7.119***	8.687***	8.837***	8.998***
	(1.767)	(1.856)	(2.991)	(2.429)	(2.463)
Partial R^2	0.254	0.233	0.284	0.296	0.306
KP *rk* Wald *F*-stat	15.812	14.719	8.433	13.233	13.346

Note: The country sample is the 98 sample excluding Papua New Guinea and Somalia, for which data are not available in PWT v9.0. $\hat{T}_i$ is the predicted trade openness based on the total sample; i.e. including out-of-sample predictions. GDP^E refers to the measure constructed using the expenditure-side GDP at current PPPs reported in PWT v9.0 for the year 1985. Real trade openness is from PWT v9.0. Heteroscedastic consistent standard errors are in parentheses. The standard errors in income regressions are corrected for the errors created from the generated regressors. The KP *rk* Wald *F*-stat is the Kleibergen-Paap *rk* Wald *F*-statistic. Exogenous variables are included in the first-stage regressions but not shown. *, **, *** Significant at 10, 5 and 1 percent, respectively.

Table D4. Estimates of Income Equation using Randomized Instruments: Real Trade Openness and GDP[E]

Second-stage results:	Model (1)		Model (2)		Model (3)		Model (4)	
	IV- $\widetilde{T_i^{TS}}$	IV-$\widetilde{T_i^{IS}}$	IV- $\widetilde{T_i^{TS}}$	IV-$\widetilde{T_i^{IS}}$	IV- $\widetilde{T_i^{TS}}$	IV-$\widetilde{T_i^{IS}}$	IV- $\widetilde{T_i^{TS}}$	IV-$\widetilde{T_i^{IS}}$
Real trade share$_i$	0.294	5.346	1.019	4.526	3.542	4.348	0.567	5.774
	(40.785)	(0.626)	(73.951)	(1.039)	(60.469)	(0.739)	(35.900)	(1.268)
	[48]	[1000]	[15]	[986]	[30]	[1000]	[32]	[989]
	{18}	{1000}	{3}	{947}	{6}	{998}	{11}	{977}
Ln population$_i$	0.169	0.289	0.080	0.233	0.158	0.194	-0.131	0.522
	(0.971)	(0.015)	(3.229)	(0.045)	(2.754)	(0.034)	(4.504)	(0.159)
	[177]	[176]	[0]	[0]	[0]	[0]	[0]	[776]
	{4}	{0}	{0}	{0}	{0}	{0}	{0}	{345}
Ln area$_i$	-0.101	0.496	0.008	0.407	0.298	0.390	0.090	0.416
	(4.816)	(0.074)	(8.410)	(0.118)	(6.912)	(0.084)	(2.245)	(0.079)
	[2]	[963]	[3]	[934]	[1]	[959]	[54]	[925]
	{0}	{767}	{0}	{820}	{0}	{841}	{11}	{753}
Distance to equator$_i$			3.470	1.210				
			(47.653)	(0.669)				
			[372]	[115]				
			{308}	{64}				
% Land in tropics$_i$					-0.825	-0.662		
					(12.213)	(0.149)		
					[431]	[639]		
					{367}	{381}		
Sub-Saharan Africa$_i$							-1.689	-0.204
							(10.243)	(0.362)
							[347]	[76]
							{281}	{40}
East Asia$_i$							-0.234	-1.974
							(11.998)	(0.424)
							[9]	[384]
							{1}	{79}
Latin America$_i$							-0.662	0.926
							(10.951)	(0.387)
							[32]	[361]
							{9}	{3}
Observations	96	96	96	96	96	96	96	96
First stage regressions:								
$\tilde{T}_i$	9.749	36.081	6.523	29.941	7.713	33.764	7.767	26.438
	(22.859)	(7.032)	(22.191)	(7.995)	(22.931)	(8.000)	(20.588)	(6.706)
	[140]	[1000]	[111]	[976]	[120]	[998]	[131]	[982]
	{72}	{1000}	{53}	{917}	{61}	{995}	{72}	{964}
Partial R^2	0.015	0.192	0.014	0.099	0.014	0.134	0.014	0.092
	(0.022)	(0.041)	(0.022)	(0.038)	(0.022)	(0.040)	(0.021)	(0.033)
KP *rk* Wald *F*-stat	1.262	30.716	1.129	7.107	1.197	16.284	1.283	14.082
	<1>	<992>	<0>	<97>	<0>	<863>	<2>	<728>

Notes. The table reports average values from 1000 replications. The standard deviation of this average is reported in parentheses. For the estimated coefficients, the number of replications that produce an estimate significant at least at the 10% level is in [square brackets] and the number of replications in which the estimate is significant at least at the 5% level is in {curly brackets}. The number of times the KP *rk* Wald *F*-stat is greater than 10 is in <angle brackets>. Exogenous variables are included in the first stage regressions but not shown.