TRADE AND PRODUCTIVITY*

Francisco Alcalá and Antonio Ciccone

We find that international trade has an economically significant and statistically robust positive effect on productivity. Our trade measure is imports plus exports relative to purchasing power parity GDP (*real openness*), which we argue is preferable on theoretical grounds to the nominal measure conventionally used. We also find a significantly positive aggregate scale effect. Our estimates control for proxies of institutional quality as well as geography and take into account the endogeneity of trade and institutional quality. Our analysis of the channels through which trade and scale affect productivity yields that they work through total factor productivity.

I. Introduction

How large is the effect of international trade on aggregate productivity? Answering this question requires dealing with possible reverse causation from productivity to trade. Ades and Glaeser [1999], Frankel and Romer [1999], and Alesina, Spolaore, and Wacziarg [2000] address this issue and find a significant causal effect of trade on productivity. Empirical work has since turned to examining whether estimates of the productivity gains due to trade may actually be capturing the role of institutions and geography. For example, Rodrik [2000], Rodriguez and Rodrik [2001], and Irwin and Tervio [2002] argue that trade is not a significant determinant of productivity when geography controls and proxies of institutional quality are included in the empirical analysis.

The measure of international trade used in almost all empirical work on the effect of trade on productivity is nominal imports plus exports relative to nominal GDP, usually referred to as *openness.* We argue that there are sound theoretical reasons why this measure may result in a misleading picture of the productivity gains due to trade. To see why, suppose that trade increases productivity but that productivity gains, in accordance with the

* We thank Tito Cordella, Luis Cubeddu, and Walter Garcia-Fontes for help with the data, José Garcia-Montalvo and Marek Jarocinski for their econometric expertise, Rita Almeida, Pablo Fleiss, and Israel Sancho for excellent research assistance, and David Romer for kindly providing the data and programs used in Frankel and Romer [1999]. The paper has benefited from the comments of editors, anonymous referees, and other colleagues. We gratefully acknowledge financial support from the CREI and CREA research institutes, the European Fund for Regional Development and the Fundación Caixa Galicia, the Fundación Séneca (PB/03/FS/02), and the Spanish Ministry of Science and Technology (SEC2001-0792/2002-01601).

© 2004 by the President and Fellows of Harvard College and the Massachusetts Institute of Technology.

The Quarterly Journal of Economics, May 2004

Balassa-Samuelson hypothesis, are greater in manufacturing than in the nontradable services sector. Will countries that are more productive due to trade have higher openness? Not necessarily, because the relatively greater productivity gains in manufacturing lead to a rise in the relative price of services, which may result in a decrease in openness. We show this formally in a trade model with gains from specialization.

This theoretical drawback of the openness trade measure motivates our alternative, which we refer to as *real openness*. Real openness is defined as imports plus exports in exchange rate US$ relative to GDP in purchasing power parity US$. Using real openness instead of openness as a measure of trade eliminates distortions due to cross-country differences in the relative price of nontradable goods.

When measured using real openness, we find that trade is a significant and robust determinant of aggregate productivity. For example, the elasticity of productivity with respect to real openness is around 1.2 with a standard error around 0.35 when we control for country size, geography, and proxies of institutional quality. This estimate implies that an increase in real openness taking a country from the thirtieth percentile to the median value raises productivity by 80 percent, an increase from the twentieth percentile to the median value raises productivity by 160 percent, and an increase from the twentieth percentile to the eightieth percentile raises productivity by a factor of six. Moreover, we find significantly positive aggregate scale effects, confirming the results of Frankel and Romer [1999] and Alesina, Spolaore, and Wacziarg [2000]. The elasticity of productivity with respect to population size is around 0.25 with a standard error around 0.1 when real openness, geography, and institutional quality are controlled for.

We also examine the channels through which international trade, scale, and our proxies of institutional quality affect aggregate productivity. Our findings indicate that trade and population size are significant determinants of total factor productivity but not of the capital-output ratio or the average level of human capital. Institutional quality, on the other hand, is a significant determinant of the capital-output ratio and the average level of human capital but not of total factor productivity.

Our empirical approach accounts for the endogeneity of trade and institutional quality by using instruments. The instrument for trade is constructed following Frankel and Romer [1999] and relies on trade being partly determined by characteristics of countries that are unrelated to productivity. The instruments for institutional quality considered come from Hall and Jones [1999] as well as Acemoglu, Johnson, and Robinson [2001] and are based on the link between historical European influence and the transmission of the European institutional framework. Instrumental-variables estimation of our productivity equation combining trade with institutional quality raises subtle issues regarding instrument weakness in models with multiple endogenous explanatory variables [Stock, Wright, and Yogo 2002]. We address these issues by using weak-instrument diagnostic tools, by implementing those instrumental-variables estimators that are most robust to instrument weakness, and by testing hypotheses using approaches that are valid asymptotically regardless of the weakness of instruments.

The remainder of this paper is structured in the following way. Section II explains the theoretical drawbacks of openness as a measure of trade. Section III contains the productivity equation that we estimate. Section IV discusses the data. Section V presents the results on the effect of international trade and scale on aggregate productivity. It also contains our findings on the effect of trade, scale, and institutional quality on the capital-output ratio, the average level of human capital, and total factor productivity. Section VI concludes.

II. Specialization, Trade, and Productivity

The drawbacks of openness as a measure of trade can be illustrated using a small open economies model with gains from specialization. The key of our argument is that specialization raises aggregate productivity but that productivity gains are greater in the tradable goods sector than in the nontradable goods sector. Specialization therefore results in an increase in the price of nontradable relative to tradable goods (the *trade-related* Balassa-Samuelson effect), which may result in a decrease in openness. Our model assumes that all productivity gains due to specialization occur in the tradable goods sector and that the demand for nontradable goods is completely price inelastic. We show later that these assumptions can be relaxed without affecting our main argument.

Suppose that the set of commodities each country can produce is given by the unit interval. Commodities indexed between

0 and t, $t < 1$ are tradable goods, while the remaining fraction $1 - t$ of commodities are nontradable goods. The measure of tradable goods produced in country c is denoted by d_c. As the measure of tradable goods produced domestically decreases, the country becomes *more specialized.*

Firms in tradable goods sectors $i \leq t$ are assumed to produce output y using labor l according to the constant-returns-to-scale production function $y = A_c l$, with country-specific factor efficiency A_c taken as given by firms. We assume that factor efficiency in tradable goods production is given by

$$A_c \equiv B_c g(d_c, l_c), \tag{1}$$

where B_c is an exogenous parameter, l_c is aggregate employment, and $g(\cdot)$ allows us to capture gains from specialization assuming that $\partial g / \partial d_c < 0$ and increasing returns to aggregate employment assuming that $\partial g / \partial l_c > 0$.

We suppose that gains from specialization are limited to tradable goods and that there are no increasing returns to aggregate employment in the nontradable goods sector. Firms in nontradable goods sectors are assumed to produce output s according to the constant-returns-to-scale production function $s = B_c l$.

Goods and labor markets are taken to be perfectly competitive. We also suppose that all tradable goods sell at the same price in international markets and take tradable goods as the numeraire. Symmetry in production implies that all nontradable goods produced in the same country sell at the same price in equilibrium. Nontradable goods prices vary endogenously across countries, however.

Households supply an aggregate amount of labor L_c inelastically. We assume that preferences are such that households want to consume the same quantity of each tradable and nontradable good irrespective of the price of nontradable goods.

Wages in tradable and nontradable goods sectors are equalized in labor market equilibrium. The equilibrium price of nontradable goods in country c, ρ_c, therefore reflects factor efficiency in tradable goods sectors relative to nontradable goods sectors,

$$\rho_c = g(d_c, L_c), \tag{2}$$

where we use that aggregate employment is equal to aggregate labor supply in equilibrium. Nontradable goods are therefore more expensive in countries where the production of tradable goods is more efficient relative to the production of nontradable goods. This yields a link between the degree of specialization and the price of nontradable goods that is key to our argument.

Balanced trade implies that imports $(t - d_c)x_c$, where $t - d_c$ is the range of imported tradable goods and x_c denotes consumption of each good, are equal to exports $d_c(y_c - x_c)$, where y_c denotes production of each tradable good. Hence, *GDP* is equal to aggregate consumption:

$$GDP_c = d_c y_c + \rho_c(1 - t)x_c = tx_c + \rho_c(1 - t)x_c. \tag{3}$$

Purchasing power parity (PPP) GDP differs from GDP in that the production of each good is valued using prices in a benchmark country. Hence, denoting the price of nontradable goods in the benchmark country by ρ, $PPP\ GDP = tx_c + \rho(1 - t)x_c$. To see how PPP GDP depends on the degree of specialization, we use that balanced trade and labor market clearing imply that the share of labor allocated to nontradable goods production is $(1 - t)\rho_c/(t + (1 - t)\rho_c)$. Combined with (2) and the production functions in tradable and nontradable goods sectors, this yields

$$\frac{PPP\ GDP_c}{L_c} = \frac{t + \rho(1 - t)}{g(d_c, L_c)^{-1}t + (1 - t)} B_c. \tag{4}$$

PPP average labor productivity is therefore increasing in the degree of specialization and in aggregate employment.

In equilibrium, openness (imports plus exports relative to GDP) is given by

$$Open_c = 2\frac{Imports_c}{GDP_c} = 2\frac{t - d_c}{t + (1 - t)\rho_c}. \tag{5}$$

Hence, an increase in the degree of specialization affects openness in two opposite ways. Holding the price of nontradable goods constant, a higher degree of specialization raises openness as more specialization necessarily implies a larger volume of imports. But, according to (2), a higher degree of specialization also raises the price of nontradable goods, which lowers openness. As a result, the relationship between the degree of specialization and openness may be nonmonotonic.[1] The nonmonotonicity between the degree of specialization and openness implies that higher

1. In the Appendix we discuss a model due to Rodrik, Subramanian, and Trebbi [2002] where openness also fails to be increasing in the fundamental variable driving trade because of systematic differences in nontradable goods prices.

openness is not necessarily associated with higher PPP average labor productivity.

Real openness differs from openness in that GDP is measured in PPP:

$$(6)\qquad ROpen_c = 2\,\frac{Imports_c}{PPP\ GDP_c} = 2\,\frac{t - d_c}{t + (1 - t)\rho}.$$

As the price of nontradable goods used to value production is the same across countries, real openness is a linear and increasing function of the degree of specialization. As a result, PPP average labor productivity in (4) can be written as an increasing function of real openness.[2]

Our model supposes that all gains from specialization occur in the tradable goods sector. This assumption is not necessary for specialization to increase the price of nontradable goods. The price of nontradable goods increases even if specialization in the tradable goods sector also increases productivity in the nontradable goods sector, as long as the effect is greater in the tradable than the nontradable goods sector.

In Alcalá and Ciccone [2001] we discuss the link between (real) openness and aggregate productivity when consumers substitute among consumption goods and show that the relationship between real openness and productivity is monotonically increasing whatever the elasticity of substitution among tradable and nontradable goods. We also show that the nonmonotonic relationship between openness and aggregate productivity may arise as long as the demand for nontradable goods is price inelastic. To understand this condition intuitively, it is useful to write openness in (5) as

$$(7)\qquad \frac{Open_c}{2} = \frac{\text{Imports}_c}{\text{Consumption of Tradables}_c}\,\frac{\text{Consumption of Tradables}_c}{\text{GDP}_c}.$$

An increase in the degree of specialization implies that the share of imported goods in total tradable goods consumption increases.

2. The monotonic relationship between specialization and real openness combined with the effect of specialization on the price of nontradable goods in (2) implies that the price level should be increasing in real openness (trade-related Balassa-Samuelson effect). We present empirical evidence for this relationship in the Appendix.

Hence, the first term on the right-hand side of (7) increases with specialization. Whether the second term increases or decreases with specialization depends on the elasticity of substitution between tradable and nontradable goods. If the demand for nontradable goods is price elastic, the increase in the price of nontradable goods caused by higher specialization translates into households spending a smaller share of their income on nontradable goods and a greater share on tradable goods. As a result, the second term on the right-hand side of (7) increases with specialization. Thus, specialization unambiguously increases openness when the demand for nontradable goods is price elastic. When nontradable goods demand is price inelastic, however, the second term on the right-hand side of (7) decreases with specialization and the effect of specialization on openness is ambiguous. While there is no evidence on the price elasticity for nontradable goods in general, the demand for services, which constitutes the largest part of nontradable goods demand, has been found to be very price inelastic (see Falvey and Gemmell [1996], for example).

Our analysis of the link between (real) openness and aggregate productivity has also been simplified by the assumption that the fundamental variable determining trade, the degree of specialization of a country, is given. In the Appendix we extend our theoretical framework by endogenizing specialization and analyzing how the efficient degree of specialization depends on an exogenous variable like transport costs. Our extended theoretical framework yields an inverse relationship between transport costs and specialization. This is intuitive because the efficient degree of specialization equalizes the marginal benefit of specialization, increased productive efficiency, and the marginal cost of specialization, increased total transport costs, and an exogenous increase in transport costs results in a higher marginal cost of specialization.

III. Estimation

The equation we use to estimate the effect of international trade and scale of production on average labor productivity across countries is the following straightforward extension of the specification in Frankel and Romer [1999]:

$$(8)\quad \log\left(\frac{PPP\ GDP_c}{Workforce_c}\right) = a_0 + a_1 ITrade_c + a_2 \log DScale_c + a_3 \log Area_c + a_4 IQual_c + a_5 X_c + u_c,$$

where *ITrade* stands for measures of international trade, *DScale* captures the domestic scale of production using either workforce or population, *Area* refers to the land area in square kilometers, *IQual* is a proxy of institutional quality, and *X* denotes a set of geographic control variables (FR include the first three right-hand-side variables only). The main geography controls considered are distance from the equator and continent dummies (Europe, Africa, America, Asia; the omitted continent is captured by the intercept). The variation in productivity not captured by our empirical analysis is summarized by u.

Our preferred measure of trade is log real openness, with real openness defined as imports plus exports in exchange rate US$ divided by GDP in PPP US$.[3] For comparisons with previous empirical work, we also measure trade using openness, defined as nominal imports plus exports divided by nominal GDP, and log openness.

The productivity equation cannot be estimated consistently using ordinary least squares because trade and institutional quality are endogenous (other variables in the productivity equation are assumed to be exogenous). We therefore rely on two-stage least-squares estimation. The instrument for trade is constructed following Frankel and Romer [1999], and the instruments for institutional quality considered are taken from Hall and Jones [1999] as well as Acemoglu, Johnson, and Robinson [2001].

To determine the causal effect of international trade on productivity across countries, Frankel and Romer [1999] use a two-step approach to construct an instrument for their measure of trade. The first step consists of estimating a gravity equation for bilateral trade shares that uses countries' geographic characteristics and size only as explanatory variables (i.e., the estimating equation does *not* include measures of productivity or income). The second step of the approach aggregates bilateral trade shares predicted by the gravity equation to obtain a geography-based

3. Our choice of log real openness instead of real openness is motivated by econometric specification tests as our theoretical framework does not determine the functional form of the relationship between real openness and productivity. The specification tests are documented in Alcalá and Ciccone [2001].

instrument for trade (the approach is explained in more detail in the Appendix). We use exactly the same approach, except that we employ more bilateral trade data than FR.

The instruments used by Hall and Jones [1999] to estimate the effect of institutional quality on productivity are the population share speaking English since birth, the population share speaking one of the five primary European languages (including English) since birth, the distance from the equator, and the Frankel and Romer [1999] geography-based trade variable. HJ argue, based on historical considerations, that the first three variables are correlated with past European influence and therefore with the transmission of the (growth-enhancing) European institutional framework. We use the population share speaking one of the five primary European languages since birth and distance from the equator as instruments, but drop the fraction of the population speaking English since birth, as it does not help in predicting the endogenous variables in our specifications. Acemoglu, Johnson, and Robinson [2001] use European settler mortality during the eighteenth and nineteenth centuries as an instrument for institutional quality in a sample of former colonies. They demonstrate that historic settler mortality explains a considerable amount of the variation in their proxy of institutional quality and argue that this correlation arises because the implementation of European institutions was more likely where conditions for long-term European settlements were more favorable.

We are also interested in whether trade affects average labor productivity mostly through physical capital, human capital, or factor efficiency. Our analysis follows the approach of Hall and Jones [1999] and Klenow and Rodriguez-Clare [1997]. Their starting point is the firm-level constant-returns production function $y = k^{\alpha}(E_c hl)^{1-\alpha}$, $0 \leq \alpha \leq 1$, where y denotes output, k capital, h average human capital, and l employment. E captures aggregate factor efficiency, which the firm takes as given. (Factor efficiency is related to total factor productivity by $TFP_c = E_c^{1-\alpha}$.) Aggregate externalities, if any, are captured by factor efficiency. Combining this production function with perfectly competitive product and factor markets implies that aggregate average labor productivity Y/L can be written as the product of factor efficiency, the aggregate (physical) capital-output ratio K/Y raised to the power $\alpha/(1-\alpha)$, and the aggregate average level of human capital H:

$$Y_c/L_c = E_c(K_c/Y_c)^{\alpha/(1-\alpha)}H_c. \tag{9}$$

This decomposition applied to our theoretical framework (without physical or human capital) yields that $Y/L = E$ with factor efficiency given by the right-hand side of (4). Extending our theoretical framework by assuming that physical and human capital enter tradable and nontradable goods production in the same way, and maintaining constant returns to scale at the firm level, would yield that aggregate average labor productivity could be decomposed exactly as in (9) with factor efficiency continuing to be equal to the right-hand side of (4).

To determine the channels through which trade, scale, and institutional quality affect aggregate productivity, we estimate their effect on each of the three components of average labor productivity on the right-hand side of (9). This is done by using the log of each component as the left-hand-side variable in (8).

IV. Data

Our empirical work is based on data for 1985. From the *Penn World Tables,* Mark 5.6, we obtain data on GDP per worker in PPP US$, GDP per capita in PPP US$, population, openness, and the price level (GDP in exchange rate US$ relative to GDP in PPP US$). Real openness is obtained by multiplying openness by the price level.[4] Workforce is obtained by dividing PPP GDP per capita by PPP GDP per worker and multiplying the result by population.

Our main proxy of institutional quality is constructed using indicators from Kaufmann, Kraay, and Zoido-Lobatón [1999] for the period 1997–1998. KKZ use a large amount of data to develop six different indices: government effectiveness, rule of law, graft, voice and accountability, political stability and violence, and regulatory burden (see the Appendix for details). The six indicators are measured in units ranging from −2.5 to 2.5, with higher values corresponding to better governance outcomes. We average the indicators that are closest to the government antidiversion policy (*GADP*) index used by Hall and Jones [1999] (government effectiveness, rule of law, and graft) to obtain an index that

4. The fact that real openness can be obtained by multiplying openness and the price level may give the impression that real openness depends on relative nontradable goods prices, when this operation actually undoes the dependence of openness on relative nontradable goods prices.

mimics the HJ indicator in breadth.[5] This index is referred to as *IQual*. We also consider the HJ *GDAP* index and the KKZ rule of law index as alternative proxies of institutional quality.

The bilateral trade data to obtain the geography-based trade instrument are taken from the *Direction of Trade Statistics* published by the International Monetary Fund. These statistics contain 9426 nonzero observations on bilateral trade for the countries used to estimate the productivity equation, which is approximately 2.5 times the data used by Frankel and Romer [1999]. As our geography-based trade variable relies on more bilateral trade data than the FR variable, it is more representative of the countries used to estimate the productivity equation.

Distance from the equator is taken from Hall and Jones [1999] and is measured as the absolute value of the latitude of the center of countries' most populated region.

Table I contains descriptive statistics and a correlation matrix for selected variables, including (log) openness, (log) real openness, log average labor productivity, log population, and *IQual*. It can be seen that real openness has a lower mean than openness. The lower mean is due to the average price level in our sample being 0.54. Moreover, the correlation between openness and real openness is high (0.86). Real openness correlates better with log average labor productivity than openness, however (a correlation coefficient of 0.26 in the case of openness and 0.47 in the case of real openness, and differences are even greater when the two trade measures are in logs).[6] These differences can be explained by the Balassa-Samuelson effect and are consistent with our theoretical framework. Table II looks at the productivity of countries that move at least 25 places up or down when using the real openness ranking instead of the openness ranking. It can be seen that countries moving up are on average about 100 percent more productive than the average country in the sample, while countries moving down are about 60 percent less productive than the average country in the sample. This implies that countries moving up more than 25 places are almost five times as productive as countries moving down more than 25 places. The positive relationship between productivity and the number of

5. *GADP* combines indicators of bureaucratic quality, law and order, corruption, risk of expropriation, and likelihood of government repudiation of contracts.

6. We list the ten countries with the highest and lowest (real) openness in the Appendix.

TABLE I

DESCRIPTIVE STATISTICS AND CORRELATION MATRIX FOR SELECTED VARIABLES

	Mean	Median	Maximum	Minimum	Std. Dev.	Observations
log *RGDPW*	8.81	8.87	10.55	6.56	1.04	150
Open	0.73	0.63	3.18	0.13	0.46	150
log *Open*	−0.50	−0.46	1.16	−2.03	0.64	150
ROpen	0.42	0.31	2.63	0.03	0.36	150
log *ROpen*	−1.19	−1.16	0.97	−3.38	0.82	150
log *Population*	8.61	8.78	13.87	4.17	1.93	150
IQual	0.05	−0.16	2.02	−1.83	0.90	138

	log *RGDPW*	*Open*	log *Open*	*ROpen*	log *ROpen*	log *Population*	*IQual*
log *RGDPW*	1.00	0.26	0.27	0.47	0.51	−0.02	0.70
Open		1.00	0.92	0.86	0.80	−0.63	0.39
log *Open*			1.00	0.75	0.86	−0.69	0.39
ROpen				1.00	0.87	−0.50	0.49
log *ROpen*					1.00	−0.60	0.52
log *Population*						1.00	−0.11
IQual							1.00

RGDPW stands for GDP in PPP US$ per worker, *Open* for openness, *ROpen* for real openness, and *IQual* for our proxy of institutional quality defined in the main text. Correlations are calculated for the 138 countries where all necessary data are available.

places countries move up when using the real openness ranking instead of the openness ranking (using negative numbers for countries that move down) is confirmed by the correlation coefficient, which is 0.55.

The three components on the right-hand side of (9) are calculated following Hall and Jones [1999]. Average levels of human capital at the country level are calculated by combining data on average schooling and Mincerian estimates of the individual return to schooling. The formula used by HJ to calculate the average human capital in country c is $H_c = \exp(\phi(S_c))$, where S_c is average schooling and $\phi(\cdot)$ is a piecewise linear function capturing estimated Mincerian returns (ϕ is defined assuming yearly rates of return of 13.4 percent for the first four years, 10.1 percent for years four through eight, and 6.8 percent for each additional year). We follow exactly the same approach but employ updated average schooling data [Barro and Lee 2000]. Aggregate capital is obtained by applying the perpetual inventory method used by HJ to investment data in the *PWT* and the physical capital income share α is set to ⅓. Aggregate efficiency E_c is calculated by

TABLE II
COUNTRIES MOVING UP/DOWN MORE THAN 25 PLACES IN THE REAL OPENNESS COMPARED WITH THE OPENNESS RANKING

Countries moving up more than 25 places			Countries moving down more than 25 places		
	Places up	1985 Productivity relative to average		Places down	1985 Productivity relative to average
REUNION	56	0.74	DOMINICAN REPUBLIC	25	0.66
ALGERIA	48	1.26	TOGO	25	0.14
FRANCE	43	2.54	BOTSWANA	26	0.63
IRAQ	43	1.49	SENEGAL	26	0.25
FINLAND	42	2.22	GAMBIA	27	0.15
CANADA	41	2.92	HUNGARY	28	1.01
NIGERIA	41	0.27	NAMIBIA	29	0.79
JAPAN	40	1.77	MOROCCO	29	0.60
ROMANIA	39	0.38	JAMAICA	29	0.44
SWEDEN	37	2.49	TONGA	33	0.56
U.S.A.	36	3.17	GUINEA-BISSAU	35	0.13
AUSTRALIA	36	2.72	WESTERN SAMOA	39	0.50
ITALY	35	2.55	CAPE VERDE ISLANDS	39	0.26
QATAR	34	3.44	SOLOMON ISLANDS	40	0.48
SWITZERLAND	34	2.80	MALI	45	0.16
DENMARK	34	2.24	BHUTAN	46	0.14
ANGOLA	34	0.16	SWAZILAND	47	0.49
GUINEA	34	0.15	BENIN	51	0.22
WEST GERMANY	33	2.56	SRI LANKA	56	0.52
UNITED ARAB EMIR.	30	3.58	MAURITIUS	57	0.70
ICELAND	30	2.18	LESOTHO	81	0.19
IRAN	30	1.30			
KUWAIT	27	3.29			
NORWAY	27	2.70			
SAUDI ARABIA	27	2.64			
U.K.	27	2.16			
Average	34.8	2.07	Average	38.7	0.43

Places Up/Down refers to how many places the country is moving up/down in the real openness ranking compared with the openness ranking.

combining H_c and $(K_c/Y_c)^{\alpha/(1-\alpha)}$ with data on average labor productivity and (9). The relevant 1985 data are available for 102 countries.

TABLE III
FIRST-STAGE REGRESSIONS

	Dependent variable is log *ROpen*		Dependent variable is *IQual*
	(1)	(2)	(3)
log *TFitAC*	0.539***		−0.100
	(0.158)		(0.164)
log *TFitFR*		0.329*	
		(0.187)	
log *Population*	−0.217***	−0.226***	−0.063
	(0.054)	(0.069)	(0.056)
log *Area*	0.065	0.029	−0.105**
	(0.043)	(0.049)	(0.45)
EuroLang	0.409**	0.42**	0.998***
	(0.198)	(0.204)	(0.107)
Distance from Equator	0.65	0.682	2.44***
	(0.479)	(0.49)	(0.492)
Continent Controls	All	All	All
F-stat log *TFIT*	11.66	3.06	—
R^2	0.572	0.523	0.599
Number of observations	138	138	138

First-stage (least-squares) regressions for log real openness (log *ROpen*) in columns (1) and (2) and for our proxy of institutional quality (*IQual*) in column (3). The two measures of the geography-based predicted trade (used as an instrument for log *ROpen*) are the one constructed in the present paper (*TFitAC*) and the one of Frankel and Romer [1999] (*TFitFR*). (See Sections III and IV for details on these variables.) *EuroLang* is the population share speaking one of the five primary European languages since birth (used as an instrument for institutional quality). "All" continent controls refers to four continent dummies (Africa, America, Asia, Europe). *F*-stat log *TFit* is the *F*-statistic of the hypothesis that log *TFit* can be excluded from the first-stage regression for log *ROpen* (i.e., that the coefficient of *TFit* is equal to zero). All regressions include a constant. Standard deviations are in parentheses. *** significant at 1 percent; ** 5 percent; * 10 percent.

V. RESULTS

V.A. Instrument Quality

Table III contains first-stage regression results for log real openness (log *ROpen*) and for our main proxy of institutional quality (*IQual*). The two columns with log real openness as a dependent variable allow us to compare the performance of our geography-based trade instrument (*TFitAC*) in predicting real openness with the performance of the original Frankel and Romer [1999] instrument (*TFitFR*). As mentioned earlier, the two instruments differ only in that ours uses more bilateral trade data. The sample used consists of the 138 countries where data for estimating the productivity equation are available.

In column (1) we show that our geography-based trade instrument is a highly significant determinant of log real openness, even after controlling for population, area, continent dummies, distance from the equator, and the population share speaking one of the five primary European languages since birth (*EuroLang*). In particular, the F-statistics of the hypothesis that our geography-based trade instrument can be excluded from the regression, i.e., that the coefficient on it is 0, equals 11.66. Hence, the F-statistic of the exclusion hypothesis exceeds the rule-of-thumb threshold of ten recommended by Staiger and Stock [1997] to avoid weak instrument concerns. Using workforce instead of population to measure scale yields an exclusion hypothesis F-statistic for our geography-based trade instrument of 9.16.

Column (2) contains the results of estimating the first-stage regression for log real openness using the Frankel and Romer [1999] geography-based trade instrument. The specification is identical to the one in the previous column to facilitate comparisons. It can be seen that the FR instrument is significant at the 10 percent level, but that the exclusion hypothesis F-statistic (3.06) is considerably lower than the one obtained using our instrument (and much below the Staiger and Stock [1997] rule of thumb). The exclusion hypothesis F-statistic for geography-based trade drops further when population is replaced by workforce. Overall, the first-stage regression results for log real openness indicate that using more bilateral trade data than FR to construct the geography-based trade variable has produced a considerably better instrument.

In column (3) we present the results of the first-stage regression for our main proxy of institutional quality. It can be seen that both the European languages variable and distance from the equator are highly significant, even after controlling for population, area, continent dummies, and geography-based trade. The F-statistic of the hypothesis that these two variables can be excluded from the first-stage regression is 27.35 (not in the table).

The Staiger and Stock [1997] rule of thumb has been suggested in the context of models with one endogenous variable. In models with two or more endogenous variables, instruments can be weak, although they are very significant in each first-stage regression. Intuitively, this is because endogenous explanatory variables predicted by the instruments may be close to collinear, which makes it difficult to separate the effects of these variables. Stock and Yogo [2003] provide a framework that allows testing the hypothesis of weak instruments in models with more than one endogenous variable. The hypothesis tested is that the quality of the instruments is below one of four levels.[7] For the productivity equation with all geography controls estimated using our trade instrument, the SY test rejects the hypothesis that the quality of the instruments is below the highest level at the 5 percent significance level. Therefore, weak instruments do not appear to be a concern. When we use the Frankel and Romer [1999] trade instrument, however, we cannot reject the hypothesis that the quality of the instruments is below the lowest level at the 5 percent significance level. Hence, there is strong evidence for instrument weakness when using the FR geography-based trade instrument.

We are also interested in the quality of the instruments when the analysis is restricted to the 80 former colonies in our sample because this subsample permits using the historic European settler mortality instrument of Acemoglu, Johnson, and Robinson [2001] for institutional quality instead of the Hall and Jones [1999] instruments. The first-stage regression for log real openness now yields that the effect of our log geography-based trade instrument is 0.53 with a standard error of 0.24, controlling for log population, log area, continent dummies, distance from the equator, and log historic European settler mortality (the corresponding result for the full sample is 0.54 with a standard error of 0.16, see column (1) of Table III). This compares very favorably with the performance of the Frankel and Romer [1999] geography-based trade variable, which turns out to be highly insignificant in the same regression (the FR instrument is actually a highly insignificant determinant of real openness even if we control for population and settler mortality only).[8] Still, the F-statistic of the hypothesis that our geography-based trade variable can be excluded from the first-stage regression is now 5.01 and

7. Stock and Yogo work with two definitions of weak instruments. The one that can be implemented for exactly identified models and for models with one degree of overidentification is based on the maximal size of 5 percent Wald tests of all endogenous variables. SY provide 5 percent critical values to test the hypotheses that the maximal size of such Wald tests exceeds 10 percent (rejection implies that instrument quality is not below the "highest level"), 15 percent (rejection implies that instrument quality is not below the "second-highest level"), 20 percent, or 25 percent (if this hypothesis cannot be rejected, instruments quality is below the "lowest level"). The programs to implement the tests used are available upon request.

8. Regressing log real openness on the log FR geography-based trade instrument, log population, and log historic settler mortality using least squares yields a coefficient of 0.13 with a standard error of 0.18 on the FR instrument.

therefore considerably below the Staiger and Stock [1997] rule of thumb of ten. Moreover, the Stock and Yogo [2003] test never rejects the hypothesis that the quality of instruments is below the lowest level at the 5 percent significance level. Because of the strong evidence for instrument weakness, we conclude that the former colonies sample cannot be relied upon for joint estimation of the effect of trade, scale, and institutions on productivity.

V.B. Trade and Productivity

Table IV compares the effect of trade on productivity when using the conventional openness trade measure with the effect when using the real openness trade measure. The dependent variable is log GDP per worker in PPP US$. The method of estimation is two-stage least squares (TSLS).

In column (1) we check whether openness (*Open*) is a significant determinant of productivity when institutional quality and geography are taken into account. This is done by estimating the same specification as Frankel and Romer [1999], except that we

TABLE IV

THE PRODUCTIVITY REGRESSION: COMPARING OPENNESS WITH REAL OPENNESS

	(1)	(2)	(3)	(4)
log *ROpen*			1.487**	1.028***
			(0.676)	(0.356)
Open	0.394	1.013*		
	(0.656)	(0.651)		
log *Workforce*	−0.101	−0.051	0.274	0.151
	(0.699)	(0.066)	(0.192)	(0.108)
log *Area*	0.140*	0.201***	0.146***	0.133***
	(0.073)	(0.066)	(0.056)	(0.044)
IQual	0.816***	0.780***	0.589**	0.706***
	(0.122)	(0.133)	(0.241)	(0.198)
Geo controls	Africa	Africa	Africa, Europe	Africa, Europe
Trade Instrument	*TFitFR*	*TFitAC*	log *TFitFR*	log *TFitAC*
Number of observations	138	138	138	138

Results of estimating equation (8) using openness (*Open*) and log real openness (log *ROpen*) (the dependent variable is GDP per worker in PPP US$). *IQual* denotes our proxy of institutional quality. The estimation method is two-stage least squares. The instruments used are the exogenous variables in the productivity regression plus either the geography-based trade variable of Frankel and Romer (*TFitFR*) or our geography-based trade variable (*TFitAC*), as well as the population share speaking one of the five primary European languages since birth and distance from the equator. (See Sections III and IV for details on these variables.) Only significant geography controls are included. The procedure used to eliminate insignificant geography controls is explained in footnote 9. All regressions include a constant. Standard deviations are in parentheses. *** significant at 1 percent; ** 5 percent; * 10 percent.

add our main proxy of institutional quality and significant continent dummies as determinants of productivity (following FR we measure scale using workforce).[9] The instruments used are the FR geography-based trade variable, the population share speaking one of the five primary European languages since birth, distance from the equator, and the exogenous variables included in the productivity equation.[10] The results indicate that openness and scale are insignificant determinants of productivity in this case, while the proxy of institutional quality is highly significant. This mirrors findings in Rodrik [2000], Alcalá and Ciccone [2001], and Rodrik, Subramanian, and Trebbi [2002].

Column (2) explores how results change when we use our geography-based trade instrument instead of the Frankel and Romer [1999] instrument. The specification is exactly the same as in the previous column in all other respects. It can be seen that our proxy of institutional quality remains highly significant and scale continues to be highly insignificant. Openness is significant at the 10 percent level but insignificant at the 5 percent level. The significance level of the trade variable falls when we use log openness to measure trade, or *GADP* or rule of law as alternative proxies of institutional quality [Alcalá and Ciccone 2001].

In column (3) we estimate the effect of trade on productivity using the log real openness (log ROpen) trade measure. We continue to control for log workforce, log area, our main proxy of institutional quality, and significant continent dummies. The instruments used are the Frankel and Romer [1999] geography-based trade variable, the European languages variable, distance from the equator, and the exogenous variables included in the productivity equation. The main news is that the effect of trade is now significant at the 5 percent level.

In column (4) we check how the results using the real openness trade measure change when we employ our geography-based

9. The procedure we use to select the geography controls in the productivity equation is to start out with all geography controls and eliminate the most insignificant ones as regressors as well as instruments sequentially until each remaining geography control is significant at the 10 percent level. The exception to this rule is distance from the equator, which is maintained as an instrument even when we eliminate it from the productivity equation. We then make sure that the specification is robust in the sense that the excluded geography controls taken one by one are not significant at the 10 percent level or more significant than the included geography controls when they are put back into the productivity equation.

10. Frankel and Rose [2002] also analyze the effect of openness and institutional quality on productivity but treat institutional quality as an exogenous variable.

trade instrument instead of the Frankel and Romer [1999] instrument. The specification is the same as in the previous column in all other regards. The real openness trade measure is now significant at the 1 percent level. This remains true when we proxy institutional quality using *GADP* or rule of law. Summarizing, trade is a highly significant determinant of productivity when we use the log real openness measure, but insignificant at the 5 percent level when we use the openness or the log openness measures of trade.[11]

Table V analyzes the effect of real openness on average labor productivity in some more detail. Because we are concerned about workforce measuring scale with error, we now follow Alesina, Spolaore, and Wacziarg [2000] in using population instead. Moreover, in Panel A, we include *all* geography controls in the empirical analysis. Panel B only includes significant geography controls. The method of estimation continues to be TSLS. The instruments used are our log geography-based trade variable, the European languages variable, distance from the equator, and the exogenous variables included in the productivity equation.

In column (1) it can be seen that trade is a significant determinant of productivity at the 1 percent level even when all geography controls are included in the regression. The elasticity of productivity with respect to real openness is 1.23. This implies that an increase in real openness taking a country from the twentieth percentile to the median value raises productivity by 160 percent and an increase from the twentieth percentile to the eightieth percentile raises productivity by a factor of six. The elasticity of productivity with respect to population is 0.27 and significant at the 5 percent level. The effect of trade and scale on productivity remains basically unchanged when we include significant geography controls only. The main news is that our proxy of institutional quality is significant at the 1 percent level.

Column (2) eliminates the three countries with the highest level of real openness (Hong Kong, Luxembourg, and Singapore) to see whether our findings are robust, following Rodrik [2000]. The effect of trade and scale on productivity changes very little compared with the previous column. When we include significant

11. Dollar and Kraay [2003a] use decadal growth regressions to show that increases in real openness over time raise aggregate productivity growth. They also estimate the effect of real openness, scale, and institutional quality on productivity levels but show that results cannot be relied upon because their instruments are very weak (see Dollar and Kraay [2003b] for a detailed analysis).

TABLE V
THE PRODUCTIVITY REGRESSION USING REAL OPENNESS

	Benchmark (1)	Excluding Hong Kong, Luxembourg, Singapore (2)	Excluding major oil producers (3)	Additional geo controls (4)
PANEL A: TSLS estimation with all geo controls				
log *ROpen*	1.229***	1.178***	1.135***	1.002***
	(0.339)	(0.308)	(0.371)	(0.322)
log *Population*	0.266**	0.282**	0.286***	0.239**
	(0.115)	(0.112)	(0.107)	(0.115)
log *Area*	0.062	0.013	0.038	0.054
	(0.052)	(0.048)	(0.055)	(0.045)
IQual	0.33	0.337	0.323	0.441*
	(0.261)	(0.247)	(0.251)	(0.236)
Geo controls	All continents distance equator	All continents distance equator	All continents distance equator	All continents, distance equator, 3 additional geo controls
PANEL B: TSLS estimation with significant geo controls only				
log *ROpen*	1.123***	1.072***	1.038***	0.905***
	(0.306)	(0.304)	(0.345)	(0.28)
log *Population*	0.227**	0.251**	0.249**	0.202**
	(0.101)	(0.106)	(0.099)	(0.097)
log *Area*	0.084**	0.038	0.07*	0.065*
	(0.04)	(0.043)	(0.042)	(0.033)
IQual	0.407***	0.63***	0.501***	0.518***
	(0.155)	(0.176)	(0.181)	(0.141)
Geo controls	Africa, America	Africa, Europe	Africa, America	Asia, 3 additional geo controls
P-value overidentifying restriction	0.52	0.61	0.77	0.83
Number of observations	138	135	129	138

Results of estimating equation (8) using the log of real openness (log *ROpen*) (the dependent variable is GDP per worker in PPP US$). The estimation method is two-stage least squares. *IQual* denotes our proxy of institutional quality. The instruments used are the exogenous variables in the productivity regression plus the log of our geography-based trade instrument, the population share speaking one of the five primary European languages since birth, and distance from the equator. (See Sections III and IV for details on these variables.) The procedure used to eliminate insignificant geography controls is explained in footnote 9. All regressions include a constant. "All continents" refers to four continent dummies (Europe, Africa, America, Asia). Major oil producers are listed in the Appendix. "3 additional geo controls" refers to dummies for East Asia, Latin America, and Sub-Saharan Africa. Standard deviations are in parentheses. *** significant at 1 percent; ** 5 percent; * 10 percent.

geography controls only, institutional quality becomes significant at the 1 percent level. In column (3) we follow Mankiw, Romer, and Weil [1992] in eliminating all countries that are major oil producers (these countries are listed in the Appendix). Again, results change little relative to the previous column, except that scale is now significant at the 1 percent level when we include all geography controls. Column (4) adds three geography controls (East Asia, Latin America, and Sub-Saharan Africa) to the empirical analysis. The elasticity of productivity with respect to real openness continues to be significant at the 1 percent level and is now equal to 1. The elasticity of productivity with respect to population is 0.24 and significant at the 5 percent level. Our proxy of institutional quality continues to be highly significant when we include significant geography controls only.

Figure I checks whether outliers drive the effect of trade on average labor productivity by plotting real openness predicted by the instruments against average labor productivity not explained by variables other than real openness. It can be seen that there seem to be no obvious outliers.

All specifications in Panel B of Table V maintain distance from the equator as an instrument, following Hall and Jones [1999]. The implied overidentifying restriction cannot be rejected for any of the specifications (the P-values are given at the bottom of the table). In overidentified models, the limited-information maximum-likelihood (LIML) estimator is preferable to TSLS when instruments are weak [Stock, Wright, and Yogo 2002]. When we implement the LIML estimator, we find almost identical coefficients and standard errors than using TSLS (not in the table). It is also noteworthy that the Stock and Yogo [2003] weak-instrument test for LIML estimation rejects the hypothesis that the quality of instruments is below the highest level at the 5 percent significance level. Hence, there is no evidence that LIML estimates may be distorted because of weak instruments.

The effect of real openness and population on productivity is robust to using *GADP* or rule of law as alternative proxies of institutional quality. For example, for the specifications in Table V with all geography controls, we find that the elasticity of productivity with respect to real openness is always greater than 1.1 and significant at the 1 percent level whatever the proxy for institutional quality used. The elasticity with respect to population is always greater than 0.26 and significant at the 5 percent level.

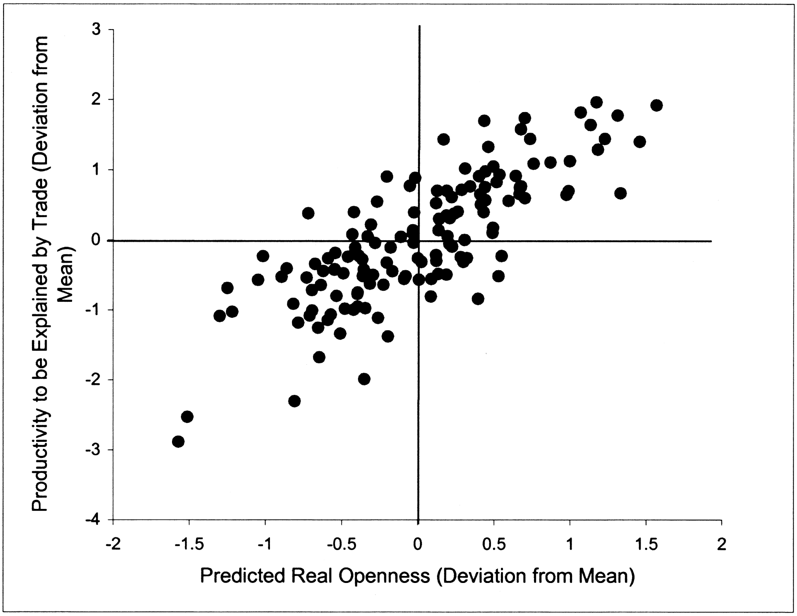


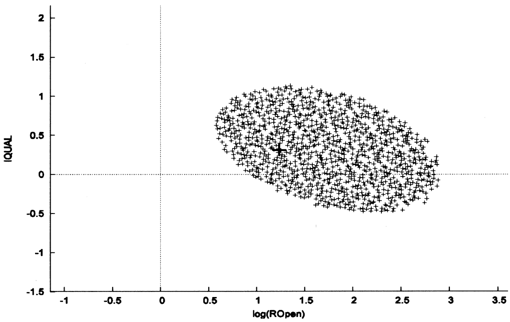
FIGURE I
Partial Scatter Plot: Real Openness-Productivity

The vertical axis measures the deviation from the mean of

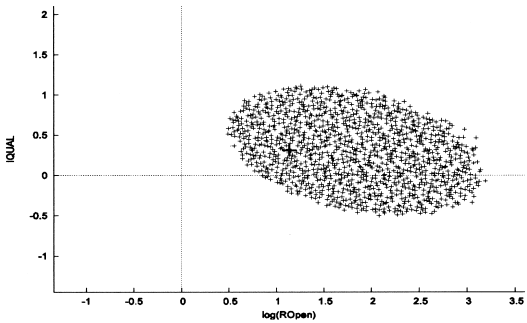
$$\log(GDP\ PPP/Workforce) - (\hat{a}_0 + \hat{a}_2 \log Population + \hat{a}_3 \log Area + \hat{a}_4 IQual + \hat{a}_5 X)$$

(notice that log real openness does not appear in the equation) with the coefficient estimates $\hat{a}_0$, $\hat{a}_2 \ldots \hat{a}_5$ taken from column (1) of Panel A of Table V. *IQual* denotes our proxy of institutional quality (see Section IV for details on this variable). The horizontal axis measures the deviation from the mean of the value of log real openness predicted by the instruments. The two observations with the lowest predicted real openness are China and India, and the one with the highest predicted real openness is Luxembourg.

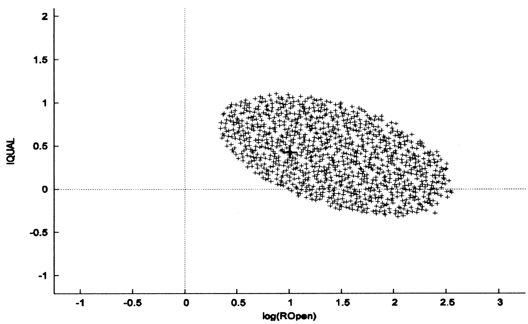
We also analyze the statistical significance of real openness as a determinant of productivity using the Moreira [2003] test statistic, which is fully robust to instrument weakness asymptotically. Figure II summarizes our results by plotting the TSLS estimates of the effect of real openness and institutional quality on productivity and the associated 95 percent confidence ellipsoids (the boundary of the cloud of crosses). The ellipsoids contain all combinations that are not significantly different from the TSLS estimates at the 5 percent level. The three graphs correspond to the specifications in Table V with all geography controls, with the exception of the specification eliminating Hong Kong,



Benchmark (Table V, Panel A column (1))



No Oil Producers (Table V, Panel A, column (3))



Additional Geo Controls (Table V, Panel A, column (4))

FIGURE II

Moreira [2003] 95 Percent Confidence Ellipsoids for the Effect of Real Openness and Institutional Quality on Productivity

ROpen denotes real openness, and *IQual* denotes our proxy of institutional quality (see Section IV for details on these variables). The clouds consist of all combinations of the effect of trade and institutions on productivity that are not significantly different at the 5 percent level from the TSLS estimates. The results are generated using Ox version 3.30.

Luxembourg, and Singapore (which would be indistinguishable from the benchmark). It can be seen that the 95 percent confidence ellipsoids only contain points where the effect of real openness on productivity is strictly positive. Moreira [2003] and Startz, Nelson, and Zivot [2001] also provide fully robust test statistics for the significance of individual endogenous variables. These statistics yield that the hypothesis that trade is not a significant determinant of productivity can be rejected at the 2 percent level for all specifications in Table V.

Our result of a significant and robust effect of real openness on productivity stands in contrast with Rodrik, Subramanian, and Trebbi's [2002] finding that institutional quality is significant and real openness insignificant when both are included in the productivity equation.[12] RST's empirical analysis of the effect of real openness and institutional quality on productivity is subject to two major limitations when compared with our approach. First, RST do not control for population or other measures of country size (the reasons why country size should be controlled for when estimating the effect of trade on productivity are detailed in Frankel and Romer [1999] and Alesina, Spolaore, and Wacziarg [2000]). Moreover, RST employ instruments that are very weak in their samples once country size is taken into account. Their specification and instruments are therefore not useful for disentangling the effects of trade and institutional quality on productivity. For example, the preferred specification of RST combines the former colonies sample and the European settler mortality instrument with the FR trade instrument. This is the combination we have shown earlier to result in the trade instrument not predicting real openness at all.[13] Our finding that population and real openness are significant determinants of productivity even when institutional quality is accounted for relies on instruments that are not weak.

12. Their empirical work focuses on showing that trade is not a robust determinant of productivity when measured using openness. But in their robustness checks they also follow Alcalá and Ciccone [2001] and Dollar and Kraay [2003a] in considering the real openness trade measure.

13. Rodrik, Subramanian, and Trebbi also estimate the effect of real openness and institutional quality on productivity by combining the Dollar and Kraay [2003a] geography-based trade instrument with the European settler mortality instrument or the European language instruments. Dollar and Kraay [2003b] show that both combinations result in very weak instruments.

TABLE VI
TRADE AND TFP

	Dependent variable is log E (1)	Dependent variable is log $(K/Y)^{1/2}$ (2)	Dependent variable is log H (3)
log *ROpen*	1.267**	0.055	−0.061
	(0.524)	(0.11)	(0.090)
log *Population*	0.341**	−0.034	−0.001
	(0.140)	(0.03)	(0.027)
log *Area*	0.044	0.054***	0.003
	(0.061)	(0.015)	(0.015)
IQual	0.189	0.170**	0.258***
	(0.23)	(0.066)	(0.054)
Geo controls	Europe America	Africa	Africa Asia
Number of observations	102	102	102

Results of estimating equation (8) using each of the three components on the right-hand side of equation (9) in logs as the left-hand-side variable. The estimation method is two-stage least squares. E denotes factor efficiency, K/Y the aggregate capital-output ratio, H the average level of human capital, *IQual* our proxy of institutional quality, and *ROpen* denotes real openness. The instruments used are the exogenous variables in the productivity regression plus the log of our geography-based trade instrument, the population share speaking one of the five primary European languages since birth, and distance from the equator. (See Sections III and IV for details on these variables.) Only significant geography controls are included. The procedure used to eliminate insignificant geography controls is explained in footnote 9. All regressions include a constant. Standard deviations are in parentheses. *** significant at 1 percent; ** 5 percent; * 10 percent.

V.C. Trade and Total Factor Productivity

Table VI summarizes our analysis of the channels through which trade affects average labor productivity. Column (1) regresses the log of factor efficiency on log real openness, log population, log area, the *IQual* proxy of institutional quality, and significant geography controls for the 102 countries where data are available. The method of estimation is TSLS and the instruments are our geography-based trade variable, the European languages variable, distance from the equator, and the exogenous variables included in the estimating equation. It can be seen that both real openness and population are significant determinants of factor efficiency at the 5 percent level while our proxy of institutional quality is highly insignificant. (As factor efficiency and total factor productivity are related by $TFP_c = E_c^{1-\alpha}$, the marginal effects of trade and scale on log total factor productivity can be obtained by multiplying the coefficients in the table by $1 - \alpha = 2/3$). Columns (2) and (3) repeat the analysis using the log of the physical capital-output ratio raised to the power 0.5 and the average level of human capital, respectively, as left-hand-side variables. Now real openness and population are insignificant, and institutional quality is significant. Hence, our results suggest that trade and scale raise average labor productivity through total factor productivity while institutional quality works through capital accumulation.

VI. Summary

Our analysis of the effect of international trade on aggregate productivity across countries emphasizes *real openness* (imports plus exports in exchange rate US$ relative to GDP in purchasing power parity US$) as a measure of trade. We argue that real openness is a better measure of trade than openness because the openness measure is distorted by cross-country differences in the price of nontradable relative to tradable goods. The distortions arise because openness is decreasing in the relative price of nontradable goods, and nontradable goods are relatively more expensive in countries where production is more efficient (the Balassa-Samuelson effect). Cross-country differences in the relative price of nontradable goods do not affect real openness because the production of nontradable goods in different countries is valued at the same prices.

Using the real openness trade measure, we find that the causal effect of trade on productivity across countries is statistically and economically significant as well as robust. We also find that productivity is affected in an economically and statistically significant way by the size of countries once international trade is taken into account. Regarding the channels through which international trade and scale affect average labor productivity, our findings indicate that they work through total factor productivity.

One of the important issues that remain to be investigated is to what extent trade policy is effective in raising productivity levels. This requires extending the empirical analysis by allowing trade to be determined by trade policy as well as geography and finding valid instruments for endogenous trade policy.

Appendix 1: Data

Obtaining the Geography-Based Trade Instrument

The gravity equation estimated to obtain the geography-based bilateral trade share of country i with country j is

$$\begin{aligned}\log\left(\frac{\tau_{ij}}{PPP\ GDP_i}\right) = {} & \alpha_0 + \alpha_1 \log Dist_{ij} + \alpha_2 \log DScale_i \\ & + \alpha_3 \log Area_i + \alpha_4 \log DScale_j + \alpha_5 \log Area_j \\ & + \alpha_6(Ldl_i + Ldl_j) + \alpha_7 Cb_{ij} + \alpha_8 Cb_{ij} \log Dist_{ij} \\ & + \alpha_9 Cb_{ij} \log DScale_i + \alpha_{10} Cb_{ij} \log Area_i \\ & + \alpha_{11} Cb_{ij} \log DScale_j + \alpha_{12} Cb_{ij} \log Area_j \\ & + \alpha_{13} Cb_{ij}(Ldl_i + Ldl_j) + v_{ij},\end{aligned}$$

where τ_{ij} denotes exports of country i to country j plus exports from j to i, $Dist_{ij}$ is the distance between the two countries (great-circle distance between countries' principal cities), Ldl_i, Ldl_j are dummies indicating whether countries i, j are landlocked, Cb_{ij} is a dummy indicating whether or not the two countries have a common border, and v_{ij} summarizes the variation in bilateral trade shares not captured by our empirical approach. The common border dummy is included by itself in the regression as well as interacted with other explanatory variables to capture trade between neighboring countries more accurately. The equation is estimated using least squares, employing the Frankel and Romer [1999] data for the right-hand-side variables and bilateral trade data from the *Direction of Trade Statistics* published by the International Monetary Fund. The ordinary least-squares estimates of the coefficients in the bilateral-trade equation are used to determine the predicted value of the bilateral trade share. Predicted bilateral trade shares are then aggregated to obtain the geography-based value of aggregate imports plus exports relative to PPP GDP for each country:

$$TFit_i \equiv \sum_j \exp\left(\text{Fitted value of } \log\left(\frac{\tau_{ij}}{PPP\ GDP_i}\right)\right).$$

Comparing the Data on Openness and Real Openness

The ten countries with the highest value of *Open* are (in this order): Singapore (3.18), Luxembourg (2.11), Hong Kong (2.09), Bahrain (1.88), Belize (1.83), St. Lucia (1.65), Malta (1.6), Lesotho (1.54), St. Vincent and Grenada (1.52), and Belgium (1.51). The ten countries with the highest value of *ROpen* are (in this order): Singapore (2.63), Bahrain (1.72), Luxembourg (1.51), Hong Kong (1.21), Puerto Rico (1.15), Belgium (1.08), Bahamas (1.0), St. Lucia (0.94), and Barbados (9.67). Evidently, there is considerable overlap. Among the bottom ten countries there is a similar amount of overlap. The ten countries with the lowest value of *Open* are Myanmar (0.13), Laos (0.14), India (0.15), Iran (0.15), Argentina (0.17), USA (0.18), USSR (0.18), Mozambique (0.18), Sierra Leone (0.19), and Brazil (0.19). The ten countries with the lowest value of *ROpen* are Bangladesh (0.03), Myanmar (0.04), India (0.04), Nepal (0.05), China (0.05), Mozambique (0.06), Laos (0.07), Sierra Leone (0.08), Brazil (0.08), and Sudan (0.08). There are many countries whose position in the ranking changes considerably, however. For example, the United States goes from being the one hundred and forty-seventh country from the top in the *Open* ranking to the one hundred and eleventh country from the top in the *ROpen* ranking. Lesotho, on the other hand, goes from ninth place in the *Open* ranking to ninetieth place in the *ROpen* ranking. Table II lists all countries moving up or down more than 25 places.

Institutional Quality Indices

Our *IQual* institutional quality proxy averages the following three indicators. Government effectiveness, which proxies mostly bureaucratic efficiency (e.g., bureaucracy/red tape, bureaucracy as an obstacle to business development, strength and expertise of the civil service to avoid drastic interruptions in government services in times of political instability) but also uses some broader data on the functioning of government (e.g., efficiency of mail delivery, quality of public health care, general condition of roads). Rule of law, which uses data on different aspects of crime and the workings of the judiciary (e.g., costs of crime, kidnapping of foreigners, independent and impartial courts) and also issues related to the enforcement of contracts and the protection of intellectual property rights (e.g., enforceability of private contracts, enforceability of government contracts). And graft, which proxies different aspects related to corruption (e.g., corruption among public officials, effectiveness of anticorruption initiatives, mentality regarding corruption). Using the *IQual* institutional quality proxy reduces the largest possible 1985 sample from 150 to 138 countries.

Major Oil Producers

We eliminate Angola, Gabon, Congo, Iraq, Oman, Kuwait, Qatar, Saudi Arabia, and the United Arab Emirates because they are major oil producers. The criterion used is that the ratio of thousands of barrels of oil produced per day to GDP exceeded 200,000 in 1985. Our list of major oil producers differs slightly from the list used by Mankiw, Romer, and Weil [1992]. Using the MRW list instead leaves results unaffected, however (we still use our list because we were unable to identify the precise definition of major oil producers of MRW).

Appendix 2: Testing for the Trade-Related Balassa-Samuelson Effect

The drawback of the openness trade measure discussed in the main text arises when the price level increases with trade. We now check for this effect by regressing the price level, P_c, on real openness and the other explanatory variables used in the productivity equation. The two-stage least squares results are (standard errors are in parentheses)

$$\log P_c = \underset{(0.15)}{0.59} \log ROpen_c + \underset{(0.05)}{0.08} \log Population_c + \underset{(0.02)}{0.06} \log Area_c - \underset{(0.11)}{0.06}\, IQual_c .$$

The regression includes all geography controls (dummies for continents and distance from the equator) and a constant. The instruments used are our geography-based trade variable, the population share speaking one of the five primary European languages since birth, and the exogenous variables included in the equation. The number of observations is 138 (the same countries used to estimate the productivity equation). It can be seen that real openness has a highly significant, positive effect on the price level. The data therefore confirm the trade-related Balassa-Samuelson effect.

The data indicate a trade-related Balassa-Samuelson effect even if we focus on geography-based trade only. Regressing the price level on our geography-based trade variable and the other exogenous variables we have been considering in our empirical work using least squares yields

$$\log P_c = \underset{(0.10)}{0.31} \log TFitAC_c - \underset{(0.03)}{0.05} \log Population_c + \underset{(0.03)}{0.10} \log Area_c + \underset{(0.13)}{0.16} \; EuroLang_c .$$

The regression includes all geography controls as well as a constant and relies on the same 138 countries used earlier. It can be seen that geography-based trade also has a highly significant, positive effect on prices. This explains the difference in performance between openness and real openness in our productivity equation.

Appendix 3: Linking Transport Costs to Efficient Specialization

We now introduce transport costs into the theoretical framework in the main text to show that an increase in transport costs lowers the efficient level of specialization. To do so, we assume an iceberg cost z of transporting tradable goods. All assumptions made in the main text are maintained. In particular, we continue to assume that the international price of all tradable goods is equal to unity, which implies that the price of imported tradable goods in terms of domestically produced, exported tradable goods is $p = (1 + z)/(1 - z)$.

The efficient degree of specialization maximizes the utility of consumers $U = x = AL/q$, where $q \equiv d + p(t - d) + (1 - t)g$. The numerator is the value of production in terms of domestic tradable goods, and the denominator is the minimum cost of purchasing one unit of each good in terms of domestic tradable goods. It is immediate that countries will specialize completely if there are no transport costs. If complete specialization or complete autarchy is not efficient, the efficient level of specialization satisfies the following first-order condition:

$$x(p - 1) = -A' l_T,$$

where the left-hand side is the *marginal cost of specialization* and the right-hand-side is the *marginal benefit of specialization,* both in terms of domestic tradable goods (l_T is the amount of labor employed in tradable goods production and $A' \equiv \partial A/\partial d$;

the assumption that specialization increases productive efficiency in tradable goods implies $A' < 0$). The MC arises because transport costs drive a wedge $p - 1$ between the price of imported tradable goods and the price of domestic tradable goods. Hence, the cost of consuming x units of a tradable good increases by $x(p - 1)$ in terms of domestic tradable goods if this good has to be imported. The MB arises because specialization implies that the tradable goods sector is more productive. Notice that the marginal effect of specialization on the MC is equal to zero when evaluated at points where utility reaches a local optimum ($\partial x/\partial d = 0$). Moreover, the marginal effect of specialization on the MB is $A''l_T$ when evaluated at points where utility reaches a local optimum (as $l_T = L - l_{NT} = L - x(1 - t)/B$). This yields that, under the concavity assumption $A'' < 0$, the MB of specialization cuts the MC from above at local optima, which implies that all local optima are local maxima. Hence, the FOC above defines a unique interior maximum. To see that an exogenous increase in transport costs raises the MC of specialization above the MB at an interior maximum, notice that an increase in specialization strictly raises utility if and only if MB − MC > 0, or making use of the definition of x and q, $(A - A'd) - p(A + A'(t - d)) > 0$. The marginal effect of an increase in p on MB − MC is therefore of the same sign as $-A - A'(t - d)$, which is of the same sign as $-A + A'd < 0$ at an interior maximum. Hence an increase in transport costs decreases the efficient degree of specialization in countries that are neither completely specialized nor completely nonspecialized (unsurprisingly, the effect of transport costs on specialization is zero in such countries).

In the case where $A = \alpha - \beta d^2$ with $\alpha \geq \beta t^2 > 0$, it is possible to solve for the efficient degree of specialization, as well as productivity and (real) openness, as a function of transport costs. If transport costs are high in the sense that $\lambda \geq t^2/(t^2 + \alpha/2\beta)$, where $\lambda = (p - 1)/p$, the country will be completely nonspecialized (and hence not trade at all). If transport costs are strictly positive but not too high in the sense that $\lambda < t^2/(t^2 + \alpha/2\beta)$, the efficient degree of specialization will be $d^* = t/\lambda - \sqrt{(t/\lambda)^2 - \alpha/\beta}$. And if transport costs are zero, the country will be completely specialized. Figure III illustrates the relationship between (real) openness, specialization, and transport costs in this example.

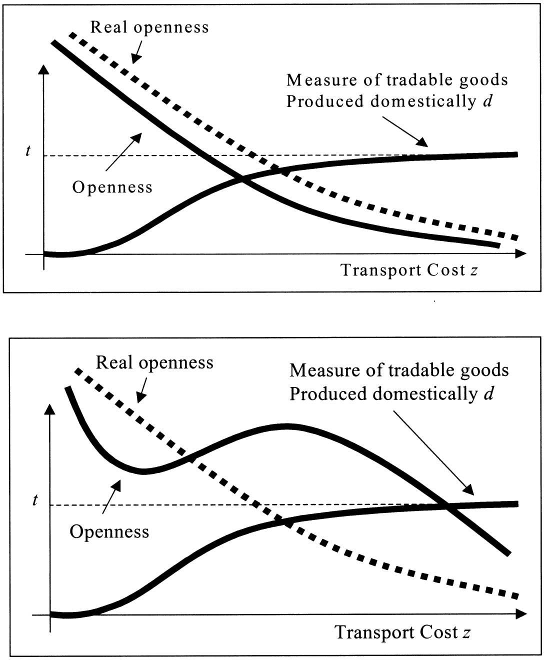


FIGURE III

The first graph corresponds to a weak trade-related Balassa-Samuelson effect. In this case, openness and real openness are positively correlated with specialization (and hence productivity). The second graph corresponds to a strong trade-related Balassa-Samuelson effect. In this case, openness is nonmonotonic in specialization, while real openness continues to be positively correlated with specialization.

APPENDIX 4: THE EXAMPLE OF RODRIK, SUBRAMANIAN, AND TREBBI [2002]

Rodrik, Subramanian, and Trebbi [2002] construct an example where openness is independent of the variable driving trade. Following their analysis, consider a multicountry trade model where countries i are endowed with N_i units of the same nontradable good and T_i units of some tradable good. Assume that the international price of all tradable goods is fixed at unity. Suppose that each country exports almost all of its tradable goods in exchange for foreign tradable goods (because each country has symmetric preferences across a very large set of tradable goods and all but one of these goods must be purchased abroad). Hence trade is driven by the amount of tradable goods a country is endowed with and the value of exports in country i is exactly equal to the value of its tradable good endowment T_i. Openness is equal to $2T_i/(T_i + p_iN_i)$ in a balanced trade equilibrium, where p_i is the relative price of the nontradable good in country i. Now assume that preferences for tradable versus nontradable goods are Cobb-Douglas, so that the nontradable good and all tradable goods combined have a fixed budget share. In this case, openness is equal to 2α, where α is the budget share of tradable goods. Hence, countries with very different amounts of traded and non-traded goods will have the same openness as long as they have the same preferences. This is because an increase in the amount of tradable goods a country is endowed with raises exports but also the demand of nontradable goods and therefore nontradable goods prices. And when preferences are Cobb-Douglas, the increase in trade is exactly offset by higher nontradable goods prices. The fact that openness does not depend on the amount of traded goods in this example makes it a counterintuitive measure of trade. It also makes openness inappropriate for cross-country empirical work trying to determine how quantities traded affect productivity.

Real openness differs from openness in that nontradable goods in different countries are valued using the price in the benchmark country. As a result, countries with a higher ratio of traded to nontraded goods have higher values of real openness, making real openness a more intuitive measure of trade than openness. Using real openness, it is therefore possible to estimate the effect of quantities traded on productivity, if any, although this requires using an instrumental-variables approach to eliminate reverse causality.

UNIVERSIDAD DE MURCIA
UNIVERSITAT POMPEU FABRA

REFERENCES

Acemoglu, Daron, Simon Johnson, and James A. Robinson, "The Colonial Origins of Comparative Development: An Empirical Investigation," *American Economic Review,* XCI (2001), 1369–1401.
Ades, Alberto, and Edward L. Glaeser, "Evidence on Growth, Increasing Returns, and the Extent of the Market," *Quarterly Journal of Economics,* CXIV (1999), 1025–1045.
Alcalá, Francisco, and Antonio Ciccone, "Trade and Productivity," CEPR Discussion Paper No. 3095, London, 2001.
Alesina, Alberto, Enrico Spolaore, Romain Wacziarg, "Economic Integration and Political Disintegration," *American Economic Review,* XC (2000), 1276–1296.
Barro, Robert, and Jong-Wha Lee, "International Data on Educational Attain-

ment: Updates and Implications," CID Working Paper No. 42, Cambridge, MA, 2000.
Dollar, David, and Aart Kraay, "Institutions, Trade, and Growth," *Journal of Monetary Economics,* L (2003a), 133–162.
Dollar, David, and Aart Kraay, "Institutions, Trade, and Growth: Revisiting the Evidence," World Bank mimeo, Washington, DC, 2003b.
Falvey, Rod, and Norman Gemmell, "Are Services Income-Elastic? Some New Evidence," *Review of Income and Wealth,* XLII (1996), 257–269.
Frankel, Jeffrey, and David Romer, "Does Trade Cause Growth?" *American Economic Review,* LXXXIX (1999), 379–399.
Frankel, Jeffrey, and Andrew Rose, "An Estimate of the Effect of Common Currencies on Trade and Income," *Quarterly Journal of Economics,* CXVII (2002), 437–466.
Hall, Robert, and Charles Jones, "Why Do Some Countries Produce More Output per Worker than Others?" *Quarterly Journal of Economics,* CXIV (1999), 83–116.
International Monetary Fund, *Direction of Trade Statistics* (Washington, DC: IMF, various years).
Irwin, Douglas, and Marko Tervio, "Does Trade Raise Income: Evidence from the Twentieth Century," *Journal of International Economics,* LVIII (2002), 1–18.
Kaufmann, Daniel, Aart Kraay, and Pablo Zoido-Lobatón, "Aggregating Governance Indicators," World Bank Working Paper No. 2195, 1999.
Klenow, Peter, and Andres Rodriguez-Clare, "The Neoclassical Revival in Growth Economics: Has It Gone Too Far?" in Ben S. Bernanke and Julio Rotemberg, eds., *NBER Macroeconomics Annual 1997* (Cambridge, MA: MIT Press, 1997).
Mankiw, Gregory, David Romer, and David Weil, "A Contribution to the Empirics of Economic Growth," *Quarterly Journal of Economics,* CVII (1992), 407–437.
Moreira, Marcelo, "A Conditional Likelihood Ratio Test for Structural Models," *Econometrica,* LXXI (2003), 1027–1048.
Rodrik, Dani, "Comments on Frankel and Rose, 'Estimating the Effects of Currency Unions on Trade and Output,' " mimeo, Kennedy School of Government, Harvard University, 2000.
Rodriguez, Francisco, and Dani Rodrik, "Trade Policy and Economic Growth: A Skeptic's Guide to the Cross-National Evidence," in Ben S. Bernanke and Kenneth Rogoff, eds., *NBER Macroeconomics Annual 2000* (Cambridge, MA: MIT Press, 2001).
Rodrik, Dani, Arvind Subramanian, and Francesco Trebbi, "Institutions Rule: The Primacy of Institutions over Geography and Integration in Economic Development," NBER Working Paper No. 9305, 2002.
Startz, Richard, Charles R. Nelson, and Eric Zivot, "Improved Inference for the Instrumental Variable Estimator," University of Washington, mimeo, 2001.
Staiger, Douglas, and James H. Stock, "Instrumental Variables Regression with Weak Instruments," *Econometrica,* LXV (1997), 557–586.
Stock, James, Jonathan Wright, and Motohiro Yogo, "A Survey of Weak Instrument and Weak Identification in Generalized Method of Moments," *Journal of Business and Economic Statistics,* XX (2002), 518–529.
Stock, James, and Motohiro Yogo, "Testing for Weak Instruments in Linear IV Regression," Harvard University mimeo, 2003.