IMF Staff Papers
Vol. 56, No. 3
© 2009 International Monetary Fund

Trade Openness and Growth: Pursuing Empirical *Glasnost*

ANDREAS BILLMEIER and TOMMASO NANNICINI*

Studies of the impact of trade openness on growth are based either on cross-country analysis—which lacks transparency—or case studies—which lack statistical rigor. This paper applies a transparent econometric method drawn from the treatment evaluation literature (matching estimators) to make the comparison between treated (that is, open) and control (that is, closed) countries explicit while remaining within a statistical framework. Matching estimators highlight that common cross-country evidence is based on rather far-fetched country comparisons, which stem from the lack of common support of treated and control countries in the covariate space. The paper therefore advocates paying more attention to appropriate sample restriction in cross-country macro research. [JEL C21, C23, F43, O57]

IMF Staff Papers (2009) **56,** 447–475. doi:10.1057/imfsp.2008.39;
published online 3 March 2009

The relationship between trade openness or economic liberalization on the one hand, and income or growth on the other, is one of the main

*Andreas Billmeier is an economist in the IMF's Middle East and Central Asia Department. Tommaso Nannicini is an assistant professor of economics at Bocconi University; he is also affiliated with IGIER and IZA. Large parts of this paper were written while Nannicini was a visiting scholar at the IMF, and he would like to express his gratitude for the hospitality and support extended to him. An online appendix to this article is posted at http://www.tommasonannicini.eu/Portals/0/trade_matching_only_appendix.pdf. The authors would like to thank Guido Tabellini, Athanasios Vamvakidis, and Marla Ripoll for sharing their data; Klaus Enders, Christian Keller, Luca Ricci, Athanasios Vamvakidis, seminar participants at the IMF and ASSET 2007, and an anonymous referee for helpful suggestions; as well as David Einhorn, Anna Maripuu, and Judith Rey for their help with the data and editorial issues.

©International Monetary Fund. Not for Redistribution

conundrums in the economics profession, especially when it comes to combining theoretical and policy-related findings with empirical findings. The theoretical advantages of trade for growth are known at least since Ricardo: international trade enables a country to specialize using its comparative advantage and benefit both statically and dynamically from the international exchange of goods.[1] From a policy perspective, the continuing efforts to liberalize international trade on a multilateral basis—first under the leadership of the General Agreement on Tariffs and Trade (GATT), now the World Trade Organization (WTO)—have contributed to better market access and rates of growth of international current account transactions much above worldwide economic growth. From an empirical point of view, however, the trade-growth link is still under discussion, both from a methodological angle and regarding the size and significance of the estimated effects.

Testing the empirical relevance of theoretical predictions in macroeconomics, growth theory, and political economics builds on cross-country evidence. In the attempt to detect correlations or causal relationships between aggregate variables, within-country variation is usually not sufficiently large to estimate the parameters of interest in a significant way, or it is so peculiar to the countries under consideration that the estimates do not hold more generally. At the end of the day, one must use some degree of cross-sectional variation to make inference on macro variables.

There is, however, widespread skepticism regarding the possibility of making sound inferences based on cross-country data. The empirical debate over the trade-growth nexus is a paradigmatic case. As Bhagwati and Srinivasan (2001) point out, both globalization supporters and foes rely on cross-country estimates, which dramatically suffer from specification problems, endogeneity, and measurement errors. According to them, cross-country regression estimates are completely unreliable, and robust evidence on the relationship between trade openness and growth ‘‘can come only from careful case studies of policy regimes of individual entries’’ (p. 19). Case studies, however, also suffer from apparent weaknesses as they lack statistical rigor and are exposed to arbitrary case selection. Instead of throwing out the baby (that is, cross-country statistical analysis) with the bath water, we propose to use case-study considerations as a sensitivity analysis of conventional cross-country estimates.

This paper evaluates the impact of a binary treatment—trade openness or economic liberalization—on the outcome—changes in per capita income. We

[1]Some theoretical models developed in the literature imply negative (or at least not necessarily positive) growth effects from trade; see the short discussion in Rodriguez and Rodrik (2001). By and large, however, macro theory has identified international exchange as a potential source of growth, theoretical exceptions being often associated with market failures that should be corrected by national policies different from protectionism; see, for example, the discussion in Bhagwati (2002). The December 2004 issue of the *Journal of International Trade and Economic Development* provides a recent review of the debate.

©International Monetary Fund. Not for Redistribution

use microeconometric matching estimators from the treatment evaluation literature that are based on the same identifying assumption as ordinary least squares (OLS)—conditional independence; that is, the selection into treatment is fully determined by observable characteristics—to make the estimation procedure more transparent; in other words, to bring *glasnost* to muddied waters. In doing so, we are able to identify an additional weakness of cross-country estimates: we show that the country comparisons that lie behind simple cross-sectional results are often more than far-fetched.

Based on the analysis of these data-driven country comparisons, we argue that it is important to control for continent or macro-region dummies to make the matches more sensible. We also show, however, that this remedy may not be enough if open and closed countries are not evenly distributed across regions—that is, they lack common support. For example, for a prominent openness indicator from the literature (the Sachs-Warner-Wacziarg-Welch openness dummy, SWWW), developed countries should not be used to investigate the trade-growth link after 1980, as all countries in this group are open and do not provide the necessary within-group variation to estimate the counterfactual outcome in the case of no treatment.[2]

Cleaning the sample of countries outside the area of common support with respect to geographic areas and other important covariates, we confirm a positive and significant association between openness and growth within selected regions and after 1970. Using an alternative measure of trade barriers, we find instead inconclusive evidence, broadly confirming results in the literature that the SWWW is biased toward finding a positive effect of trade liberalization.

Matching estimates are of course subject to the same endogeneity issues of OLS, but this does not undermine the basic message of our results. Even estimators that deal more credibly with endogeneity—such as diff-in-diff panel strategies—suffer from the same common-support issue that we have identified as long as they rely to some degree on cross-sectional variation. This is the reason why, also in panel setups, more transparent estimators that control for a distributional overlap of treated and control countries in the covariate space should be preferred.

In the end, we wind up with the usual statistical trade-off between internal and external validity: although dropping countries outside the common support produces more sound statistical inference, these results cannot be extrapolated to make general statements that go beyond the sample effectively used in the estimation. In other words, it is unlikely that the effect of openness on growth can be robustly estimated for a worldwide sample of countries, casting doubt on much of the cross-country growth literature that strives to cover an ever-increasing set of countries.

[2]See Section II for a description of the SWWW indicator.

©International Monetary Fund. Not for Redistribution

I. Literature Review

Empirical Studies on Trade Openness and Growth

Providing conclusive empirical evidence on the intuitively positive causal effect of trade on growth has been a challenging endeavor, complicated by a multiplicity of factors; see, for example, Winters (2004) for an overview. Most of the literature has used cross-country evidence that suffers from numerous shortcomings, related to both the measurement of openness and econometric modeling.

Following Barro's (1991) seminal paper on growth regressions, several prominent cross-country studies established a positive link between trade openness and growth; these studies include Dollar (1992), Sachs and Warner (1995), and Edwards (1992, 1998). Similarly, Vamvakidis (2002) finds, in a historical context, evidence that trade is associated positively with growth after 1970, but not before. In a Stern Review of the cross-sectional literature on trade and growth, Rodriguez and Rodrik (2001) criticize the choices of openness measure and weak econometric strategies. They find little evidence that open trade policies as measured in the aforementioned contributions are significantly associated with economic growth once they correct for the weaknesses they point out.[3] Harrison (1996) shows that most of the explanatory power of the composite openness dummy assembled in Sachs and Warner (1995) comes from the nontrade components of this measure.

DeJong and Ripoll (2006) take up one of the suggestions voiced in Rodriguez and Rodrik (2001) and construct an alternative measure of direct trade barriers—*ad valorem* tariff rates—that is arguably more immune to the Rodriguez-Rodrik critique. They find that the relationship between trade barriers and income is nonlinear for a panel of 60 countries. In particular, the correlation between trade barriers and income is negative for rich countries but positive (albeit statistically weaker) in poorer countries. Salinas and Aksoy (2006), however, criticize this indicator on the grounds that it is not a rigorous representation of the tariff structure and that it does not capture other trade barriers, such as nontariff barriers—in fact they find a low correlation between cross-country measures of unweighted average tariffs and the frequency of nontariff barriers.

From a methodological perspective, deep skepticism has been brought to bear against cross-country evidence on the trade-growth issue. In addition to the citation in the previous section, Bhagwati and Srinivasan (2002, p. 181) point out that "cross-country regressions are a poor way to approach this question" and that "the choice of period, of the sample, and of the proxies, will often imply many degrees of freedom where one might almost get what one wants if one only tries hard enough!" Levine and Renelt (1992) and Temple (2000) apply extreme-bounds analysis to show that the results of

[3] Rodriguez (2006) takes stock of recent developments in this respect.

©International Monetary Fund. Not for Redistribution

cross-country growth regressions are not robust to even small changes in the conditioning information set (that is, right-hand side variables).

Focusing on identification issues, cross-country studies suffer from two major weaknesses: reverse causality (that is, liberalized trade causes higher economic growth versus more trade being the result of economic growth) and endogeneity (for example, country-specific omitted characteristics affecting both openness and growth). Dealing with endogeneity has triggered a substantial amount of interest in the use of instrumental variables (IV). This family of models suggests using regressors that have an impact on openness, but are uncorrelated with income. Using gravity models, Frankel and Romer (1999) and Irwin and Tervio (2002) find a positive effect running from trade to growth by isolating geographical components of openness that are assumed independent of economic growth, including population, land area, borders, and distances. But even these presumably exogenous instruments could have indirect effects on growth, thereby biasing the estimates.[4] Dollar and Kraay (2003) suggest estimating the regressions in differences and using lagged openness as instrument. However, the simultaneity bias in the trade-growth context could extend over time—trade today may depend on growth tomorrow via imports for investment purposes—and using lagged variables as instruments is unlikely to fully correct for the bias.

As an alternative approach to classic IV, Lee, Ricci, and Rigobon (2004) use identification through heteroskedasticity in a panel framework, and find that openness has a small, positive, but not particularly robust effect on growth. They have to rely, however, on the nontestable assumption that the structural shocks in the system of simultaneous equations are uncorrelated. Using the same technology, Rigobon and Rodrik (2005) find that trade openness (defined as the trade share in GDP) has a significant negative effect on income.

Another strand in the trade and growth literature seeks to improve upon cross-country regressions by employing panel techniques, geared at controlling for (time-invariant) unobservable country effects. An early example is Harrison (1996), who uses fixed-effect estimators and finds a stronger impact of various openness indicators in a panel setup compared with standard cross-country regressions. Wacziarg and Welch (2003) further the discussion in the literature in three directions: they update, expand, and correct the trade openness indicator in Sachs and Warner (1995); they show that the Sachs and Warner (1995) results of a positive effect of trade on growth break down if extended to the 1990s in a cross-sectional setup; and they provide evidence in a panel context that, even in the 1990s, there is a positive effect of trade on growth when the analysis is limited to

[4]In fact, Irwin and Tervio (2002) and Rodriguez and Rodrik (2001) find that geographical latitude has a significant effect on growth, casting doubt on the identifying assumption used by Frankel and Romer (1999). Furthermore, these instruments relate primarily to trade volumes, not trade policies, as discussed by Rodriguez and Rodrik.

©International Monetary Fund. Not for Redistribution

within-country effects.[5] Slaughter (2001) uses a diff-in-diff approach to infer the effect of four very specific trade liberalization events on income growth dispersion and finds no systematic link between trade liberalization and per capita income convergence. Giavazzi and Tabellini (2005) also apply a diff-in-diff approach to study the interactions between economic and political liberalizations. They find a positive and significant effect of economic liberalization on growth, but they claim that this effect cannot be entirely attributed to international trade, as liberalizations tend to be accompanied by other policy improvements.

Empirical Studies Applying Matching Estimators to Macro Data

A limited, but growing, strand of aggregate empirical literature—particularly in political economics—applies microeconometric estimators developed in the treatment evaluation literature to cross-country data to overcome the weaknesses of OLS in cross-sectional setups. Persson and Tabellini (2003) use propensity-score matching methods to estimate the effects of political institutions (proportional against majoritarian electoral rule; presidential against parliamentary regime) on a set of relevant economic variables. Edwards and Magendzo (2003) apply matching estimators to analyze the macroeconomic record of dollarized economies. Atoyan and Conway (2006) use matching estimators to evaluate the impact of IMF programs.

All these studies point to the fact that nonparametric (or semiparametric) matching estimators allow the OLS linearity assumption to be relaxed. This is not their only merit, however, as the linearity assumption can be also relaxed in the OLS framework by specifying a fully saturated model. The major advantage of matching techniques is that they allow the researcher to carefully check for the existence of a common support in the distributions of treated and control units across the covariate space. And this advantage can be even greater in a small sample of countries, as the ‘‘matched’’ treated and control units can be easily identified. This ‘‘transparency’’ attribute of matching estimators is described and exploited in Section III with respect to the estimated effect of trade openness on growth.

II. Data and Variables of Interest

Under the trade and growth umbrella, a whole set of relationships have been analyzed in the literature. As the dependent variable, GDP levels, changes, GDP per capita, and relative incomes (or dispersion thereof) have been used as outcome measures, mainly to distinguish between level, growth, and convergence effects. We employ the difference of (log) per capita GDP, as we are interested in the dynamic impact of trade openness over time, not only in its one-off effects on the individual income level.

[5]Wacziarg and Welch (2003) essentially conduct difference regressions in growth, or diff-in-diff regressions in log income.

©International Monetary Fund. Not for Redistribution

For trade and openness, two major groups of indicators have emerged in the literature, addressing somewhat different questions. On the one hand, there are simple measures of trade volumes that are particularly subject to endogeneity problems (especially if normalized by GDP), and have in fact been used within an IV framework (for example, Frankel and Romer, 1999; Rigobon and Rodrik, 2005). On the other hand, there have been repeated efforts to identify the impact of trade policy and lower trade barriers on economic growth. To this end, a variety of indicators have been constructed, the most notable among them being the binary indicator by Sachs and Warner (1995), extended, updated, and revised by Wacziarg and Welch (2003); short SWWW.[6] According to this indicator, a country is considered closed to international trade in any given year if at least one of the following conditions is satisfied: (i) average tariffs exceed 40 percent; (ii) nontariff barriers cover more than 40 percent of its imports; (iii) it has a socialist economic system; (iv) the black market premium on the exchange rate exceeds 20 percent; and (v) much of its exports are controlled by a state monopoly. A country is open if none of these conditions applies. As our binary indicator of openness—or economic liberalization in the language of Giavazzi and Tabellini (2005)—we use the SWWW trade openness policy dummy. Our main treatment indicator thus intends to capture policy changes that reduce the constraints on market operations below a critical threshold along these five dimensions.[7]

In the third part of Section III, we repeat the analysis to the extent feasible with an alternative trade barrier indicator suggested by Rodriguez and Rodrik (2001) and applied in DeJong and Ripoll (2006).[8] This annual indicator is available for the period 1975 until 2000 and represents, in the words of DeJong and Ripoll (2006, p. 630), "ad-valorem tariffs, measured using import duties as a percentage of imports, as reported by the World Bank." This indicator is, hence, essentially a subindicator of SWWW—corresponding to (i) above—but with an additional degree of freedom: to obtain a binary treatment consistent with our treatment evaluation framework, we need to split the country sample according to the degree of protection by choosing an appropriate tariff threshold to indicate whether a country is open or closed.

To anchor our results in the existing literature, we draw on two data sets used recently in a related context. Vamvakidis (2002) presents historical evidence of the connection between openness and growth over the period 1870–1990; we focus on the post-1950 part of his data set.[9] The data set

[6]For a comparison of various indicators, see Harrison (1996).

[7]The SWWW dummy captures, in fact, more than just openness to trade, for example, also the socialist origin. Nevertheless, we base our initial analysis on this dummy, given the prominence it has achieved in the literature. Sachs and Warner (1995, p. 25) note that the socialism indicator serves as a proxy for central planning, which could be viewed as a substitute for overt trade policies such as tariffs.

[8]We are grateful to an anonymous referee for this suggestion.

[9]See Vamvakidis (2002) for a detailed description of the data sources.

©International Monetary Fund. Not for Redistribution

consists of repeated country cross-sections for the intervals 1950–70, 1970–90, and 1990–98. Besides the average GDP per capita growth and the openness dummy, the data set contains information on the initial GDP, investment share, population growth, secondary school enrollment, inflation, and black market premium.

The other data set we use has been analyzed in Giavazzi and Tabellini (2005) and Persson and Tabellini (2006). Of this very rich panel data set covering about 180 countries over the period 1960–2000, we use decade averages (1961–70, 1971–80, 1981–90, and 1991–2000) only for a few variables that are related to the question at hand: the updated SWWW, the log change in per capita GDP, and the same control variables mentioned above (with the only two exceptions that inflation is not reported, while a democracy dummy is present).

III. Matching Estimators and Cross-Country Analysis

Methodology

The common aim of most of the empirical studies reviewed in Section I is to assess whether a pro-openness trade policy has a causal effect on either the level or the growth rate of GDP. This problem of inference involves "what if" statements and thus counterfactual outcomes. Hence, it can be translated into a treatment-control situation and analyzed within Rubin's (1974) potential-outcome framework for causal inference. The essential feature of this approach is to define the causal effect of interest as the comparison of the potential outcomes for the same unit measured at the same time: $Y(0) =$ (the value of GDP growth Y if the country is exposed to treatment $T = 0$, that is, if it is closed to trade), and $Y(1) =$ (the value of GDP growth if the same country is exposed to treatment $T = 1$, that is, it is open to trade). Only one of these two potential outcomes can be observed—specifically, the one corresponding to the treatment the country received—but the causal effect is defined by their comparison, that is, $Y(1) - Y(0)$. This highlights that estimating the causal relationship between T and Y is hampered by a problem of missing data—the counterfactual outcomes $Y(0)$ for open countries and $Y(1)$ for closed countries.

In this setting, the aim of statistical analysis is usually that of estimating some features of the distribution of $Y(1) - Y(0)$, like

$$E[Y(1) - Y(0)], \tag{1}$$

which is called the average treatment effect (ATE). Alternatively, one can be interested in the ATE for the subpopulation of the treated observations:

$$E[Y(1) - Y(0)|T = 1], \tag{2}$$

which is called the average effect of treatment on the treated (ATT). In the present context, the ATE corresponds to the counterfactual question: what would have been the growth rate of the countries in our sample had they

©International Monetary Fund. Not for Redistribution

decided to switch their trade regime? On the contrary, the ATT focuses on the counterfactual question for treated units only: what would have been the growth rate of open countries had they decided to close their economies?

Problems for the identification of these ATE may arise from the existence of country-specific unobservables affecting both the two potential outcomes (or just one of them) and the treatment indicator. The fact that the treatment might be *endogenous* reflects the idea that the outcomes are jointly determined with the treatment, or that there are omitted confounders related to both the treatment and the outcomes. One of the assumptions that allow the identification of the ATE is the "unconfoundedness" condition, also referred to as "selection on observables" or "conditional independence assumption," which is the rationale behind common estimation strategies such as regression modeling and matching.[10] This assumption considers the conditioning set of all relevant pretreatment variables X and assumes that

$$Y(1), Y(0) \perp T | X \tag{3}$$

$$0 < Pr(T = 1|X) < 1. \tag{4}$$

That is, conditioning on observed covariates X, the treatment assignment is independent of potential outcomes.[11] Unconfoundedness says that treatment assignment is independent of potential outcomes after accounting for a set of observable characteristics X. In other words, exposure to treatment is random within cells defined by the variables X.

Under unconfoundedness, one can identify the ATE within subpopulations defined by X:

$$E[Y(1) - Y(0)|X] = E[Y(1)|T = 1, X] - E[Y(0)|T = 0, X], \tag{5}$$

and also the ATT as

$$E[Y(1) - Y(0)|T = 1, X] = E[E[Y(1)|T = 1, X] - E[Y(0)|T = 0, X]|T = 1], \tag{6}$$

where the outer expectation is over the distribution of X in the subpopulation of treated units. In other words, thanks to unconfoundedness, one can use the observed outcome of treated (control) units, conditional on X, to estimate the counterfactual outcome of control (treated) units.

An implication of the above results is that, if we could divide the sample into cells determined by the exact values of the variables X, then we could just

[10]See Imbens (2004) for a review of nonparametric estimation methods under this assumption.

[11]To identify the ATT, a weaker version of these conditions suffices: $Y(0) \perp T|X$ and $Pr(T = 1|X) < 1$.

©International Monetary Fund. Not for Redistribution

take the average of the within-cell estimates of the ATE. Often the variables X are continuous, so that smoothing techniques are needed; under unconfoundedness several estimation strategies can serve this purpose. Regression modeling and matching are viable alternatives, which rely on the same identification condition. The main advantage of matching with respect to linear regression is that the latter obscures information on the distribution of covariates in the two treatment groups. In principle, one would like to compare countries that have the same values of all covariates; but unless there is a substantial overlap between the two covariates distributions, a regression model relies heavily on model specification—that is, on extrapolation—for the estimation of (treatment) effects. It is thus crucial to check how much the distributions of the treated and control units overlap across covariates, and which is the region of common support for the two distributions.

In contrast to other studies that apply the propensity-score version of matching to macro data—see Persson and Tabellini (2003)—we implement the above strategy by using the ‘‘nearest neighbor’’ algorithm for covariate matching.[12] Matching estimators impute the country's missing counterfactual outcome by using average outcomes for countries with ‘‘similar’’ values of the covariates. The nearest neighbor algorithm uses the following simple approach to estimate the pair of potential outcomes. The potential outcome associated to the treatment that country A received is simply equal to the observed outcome of A. The potential outcome associated to the treatment that country A did not receive is equal to the outcome of the nearest country that received the opposite treatment (country B), where ‘‘nearest’’ means that the vector of covariates of B shows the smallest distance from the vector of covariates of A according to some predetermined distance measure.

Formally, define $||x||_V = (x'Vx)^{1/2}$ as the vector norm with positive definite weight matrix V, and let $||x-z||_V$ be the distance between vectors x and z.[13] Let $d(i)$ be the smallest distance from the covariates of country i, X_i, with respect to the covariates of all other countries with the opposite treatment. Allowing for the possibility of ties, define $J(i)$ as the set of indices for the countries that are at least as close to country i as its nearest neighbor:

$$J(i) = \{k = 1, ..., N | T_k = 1 - T_i, ||X_k - X_i||_V = d(i)\}. \tag{7}$$

The pair of potential outcomes for country i are estimated as

$$\hat{Y}_i(l) = Y_i \quad \text{if } T_i = l \tag{8}$$

[12]See Abadie and others (2004) for a description of this algorithm and the program that implements it in Stata.

[13]Following Abadie and others (2004), we let V be the diagonal matrix with the inverses of the variances of the covariates on the main diagonal. All the estimates presented in this section are robust to the utilization of a different distance metric—the Mahalanobis distance suggested by Rubin (1980).

©International Monetary Fund. Not for Redistribution

$$\hat{Y}_i(l) = \frac{1}{\# J(i)} \sum_{k \in J(i)} Y_k \quad \text{if } T_i = 1 - l, \tag{9}$$

where $\# J(i)$ is the numerosity of the set $J(i)$. The ATE and ATT are thus estimated as

$$\tau_{ATE} = \frac{1}{I} \sum_{i=1}^{I} [\hat{Y}_i(1) - \hat{Y}_i(0)] \tag{10}$$

$$\tau_{ATT} = \frac{1}{I_T} \sum_{i:T_i=1} [Y_i - \hat{Y}_i(0)], \tag{11}$$

where I and I_T are the sample size and the number of treated countries, respectively. These nearest-neighbor matching estimators both allow for the identification of the ATE and ATT under unconfoundedness and are fully transparent, as the list of country matches underlying, the results can be displayed in small samples (see the next two subsections and the online appendix).

Summing up, applying matching estimators to (small) cross-country samples comes with a disadvantage and an advantage. The disadvantage is that unconfoundedness is unlikely to hold, as it is often implausible to assume that country-specific unobservable characteristics do not play any role in treatment assignment. The advantage is that they allow us to transparently check for the existence of common support. Consequently, matching estimators are not used in this section as a magic bullet able to produce more reliable estimates than regression modeling, as both estimation strategies rest on the same identification condition and are therefore subject to the same specification problems. They are used, instead, to highlight the country comparisons that are behind cross-sectional results, to assess their plausibility, and to check whether the distributions of treated and control countries display sufficient overlap in the covariate space. After these steps, the cross-sectional results are improved by restricting the estimates to the region of common support. Even though these refined results must also rely on the conditional independence assumption, their plausibility can be further assessed by a careful inspection of the new country matches produced by the nearest neighbor algorithm. In other words, as the estimation process is no longer a black box but based on a transparent match of countries, case-study considerations along the lines of Bhagwati and Srinivasan (2001) can be introduced to assess the robustness of the results.

The Unbearable Lightness of Cross-Country Estimates

We now turn to the data sets introduced in Section II and apply matching estimators to shed light on the country comparisons underlying cross-sectional estimates.

©International Monetary Fund. Not for Redistribution

Table 1 presents results for the Vamvakidis (2002) data set. We confirm his results that openness—as represented by the Sachs and Warner (1995) dummy—has a significant effect on growth after 1970, but not before. The coefficients indicate that an open country grows, on average, by 1.5 to 2 percentage points per year faster than a closed economy. The results for both types of matching estimates, ATE and ATT, are qualitatively and quantitatively similar to the standard OLS results. The estimates are robust to the introduction of regional dummies among control variables. Unfortunately, the data set comes with several drawbacks: (i) the data are pooled for 20-year intervals; (ii) the information stops in 1998, too early to meaningfully capture the countries of the former Soviet Union territory; and (iii) the sample size is very small in the 1950s and 1960s—mainly developed and Latin American countries.

In Table 2, we repeat the exercise switching to the Persson and Tabellini (PT) data set. This data set contains more countries and extends until 2000, using the Wacziarg and Welch (2003) update of the Sachs-Warner dummy. We produce pooled estimates by decades for the whole data set. Again, the matching results are very similar to the OLS results, and we find a significant effect of trade on growth for the 1990s and 1970s. In these decades, open countries grew on average by 1.5 to 2 percentage points faster than closed countries, whether we control for regional dummies or not. The growth effect of openness is not significantly different from zero in the 1980s and 1960s.

So far, matching does not add anything to OLS results, as the estimates are very similar and based on the same identification assumption. We now turn to the transparency advantage of the nearest-neighbor matching estimator to reveal the country comparisons underlying the estimates from the PT data set. Tables 3 and 4 display the full list of treated (that is, open to trade for more than half of the decade) and control (that is, closed) countries and their nearest neighbors in the ATE estimation for the 1990s. The online appendix contains the full lists of treated and control countries and their nearest neighbors for the 1990s, 1980s, 1970s, and 1960s.[14] In all of these tables, the first column (*Country*) indicates the country under consideration; the second column (*Baseline*) shows the nearest neighbor used to estimate the counterfactual outcome of the country in the first column for the ATE estimation *without* area dummies; the third column (*Area*) shows the nearest neighbor used to estimate the counterfactual outcome of the country in the first column for the ATE estimation *with* area dummies. For example, the *Baseline* matches in Tables 3 and 4 for the 1990s are the country comparisons underlying the 1.505 coefficient in Table 2 (that is, the effect of openness on growth without controlling for area dummies), whereas the *Area* matches in

[14]Online appendix available at http://www.tommasonannicini.eu/Portals/0/trade_matching_only_appendix.pdf. Tables 1 through 8 display the matches for the PT data set.

©International Monetary Fund. Not for Redistribution

Table 1. Openness and Growth, Cross-Country Evidence (I), 1950–98
OLS: $E[Y|X]=\alpha+\tau T+\beta X$; ATE: $E[Y(1)-Y(0)|X]$; ATT: $E[Y(1)-Y(0)|T=1,X]$

	1990–98	1970–90	1950–70
OLS without area dummies			
Estimate of τ	1.630*	1.041*	0.041
SE	(0.772)	(0.492)	(0.473)
Adjusted R^2	0.36	0.52	0.41
OLS with area dummies			
Estimate of τ	1.357*	1.883**	−0.175
SE	(0.625)	(0.537)	(0.538)
Adjusted R^2	0.52	0.57	0.42
Matching without area dummies			
Estimate (ATE)	1.731*	3.089**	−0.347
SE	(0.808)	(0.749)	(0.490)
Estimate (ATT)	1.974*	1.578**	−0.423
SE	(0.968)	(0.616)	(0.724)
Matching with area dummies			
Estimate (ATE)	1.604*	3.180**	−0.414
SE	(0.749)	(0.853)	(0.522)
Estimate (ATT)	1.681	1.421**	−0.726
SE	(0.962)	(0.487)	(0.761)
Treated	73	24	22
Controls	36	64	24
Observations	109	88	46

Source: Authors' calculations based on data in Vamvakidis (2002).

Note: Dependent variable (Y): real GDP per capita growth. Treatment indicator (T): trade openness dummy (Sachs and Warner, 1995). Control variables (X) as understood in Vamvakidis (2002) include initial GDP per capita, secondary school enrollment, population growth, investment share, black market premium, and inflation for 1990–98 and 1970–90; initial GDP per capita, illiteracy rate, population growth, and investment share for 1950–70. Area dummies refer to Africa, Asia, Latin America, Middle East, developed countries, and transition economies. ATE and ATT stand for average treatment effect and average treatment effect on the treated, respectively, and are estimated by nonparametric nearest-neighbor matching. *corresponds to 5 percent significance level; **corresponds to 1 percent significance level.

the same tables lie behind the 1.318 coefficient in Table 2 (that is, the effect of openness after controlling for area dummies).

Tables 3 and 4 indicate that a few *Baseline* matches appear to work reasonably "well"—for example, in the 1990s for Bulgaria and Egypt, which are matched with Ukraine and Algeria, respectively. Arguably, this intuitive appreciation is based on the implicit assumption that there are region-specific unobservable effects, for example, a common language, colonization, level of development, geographic proximity, or legal origin. For others—for example,

©International Monetary Fund. Not for Redistribution

Table 2. Openness and Growth, Cross-Country Evidence (II), 1961–2000
OLS: $E[Y|X]=\alpha+\tau T+\beta X$; ATE: $E[Y(1)-Y(0)|X]$; ATT: $E[Y(1)-Y(0)|T=1,X]$

	1991–2000	1981–90	1971–80	1961–70
OLS without area dummies				
Estimate of τ	1.576*	0.594	2.117**	–0.058
SE	(0.644)	(0.578)	(0.809)	(0.559)
Adjusted R^2	0.23	0.32	0.20	0.29
OLS with area dummies				
Estimate of τ	1.514**	0.773	1.610**	–0.089
SE	(0.553)	(0.425)	(0.516)	(0.574)
Adjusted R^2	0.33	0.46	0.28	0.30
Matching without area dummies				
Estimate (ATE)	1.505*	0.056	2.094*	–0.254
SE	(0.646)	(0.695)	(0.933)	(0.838)
Estimate (ATT)	1.453*	–0.391	2.815**	0.138
SE	(0.705)	(1.208)	(0.878)	(0.941)
Matching with area dummies				
Estimate (ATE)	1.318*	0.641	2.399**	–0.442
SE	(0.672)	(0.462)	(0.825)	(0.829)
Estimate (ATT)	1.130	0.826	1.895*	0.213
SE	(0.742)	(0.579)	(0.831)	(0.943)
Treated	87	43	33	31
Controls	26	66	74	75
Observations	113	109	107	106

Source: Authors' calculations based on data in Persson and Tabellini (2006).

Note: Dependent variable (Y): real GDP per capita growth. Treatment indicator (T): trade openness dummy (Sachs and Warner, 1995; Wacziarg and Welch, 2003). Control variables (X) include initial GDP per capita, secondary school enrollment, population growth, and investment share. Area dummies refer to Africa, Asia, Latin America, Middle East, developed countries, and transition economies. ATE and ATT stand for average treatment effect and average treatment effect on the treated, respectively, and are estimated by nonparametric nearest-neighbor matching. *corresponds to 5 percent significance level; **corresponds to 1 percent significance level.

Albania and Sri Lanka, which are matched with Central African Republic and Algeria, respectively—the matches are somewhat less meaningful. In particular, all (treated) developed countries give rise to very poor matches (for example, Italy and the United Kingdom with Russia, or the United States and Canada with China). In other words, most of the baseline matches do not appear robust to area-specific unobservables.

Therefore, we construct country groups that may capture some of these area-specific unobservables. We divide the world into six groups: Africa, Asia, Latin America, Middle East, developed economies, and transition

©International Monetary Fund. Not for Redistribution

Table 3. Cross-Country Matches, Treated Countries, 1991–2000

Country	Baseline	Area	Refined	Country	Baseline	Area	Refined
Albania	C.A.R.	Belarus	Belarus	Latvia	Belarus	Belarus	Belarus
Argentina	Iran	Venezuela	*dropped*	Luxembourg	Iceland	Iceland	*dropped*
Australia	Belarus	Iceland	*dropped*	Madagascar	Chad	Chad	Senegal
Austria	Belarus	Iceland	*dropped*	Malaysia	Iran	China	China
Belgium	Belarus	Iceland	*dropped*	Mali	Chad	Chad	Chad
Benin	Malawi	Malawi	Malawi	Mauritania	Togo	Syria	Syria
Botswana	Belarus	Zimbabwe	Zimbabwe	Mauritius	Zimbabwe	Zimbabwe	Zimbabwe
Brazil	India	Venezuela	*dropped*	Mexico	Iran	Venezuela	*dropped*
Bulgaria	Ukraine	Ukraine	Ukraine	Morocco	Venezuela	Algeria	Algeria
Cameroon	Senegal	Senegal	Senegal	Mozambique	Rwanda	Rwanda	Rwanda
Canada	China	Iceland	*dropped*	Nepal	Zimbabwe	P.N.G.	P.N.G.
Cape Verde	Togo	Togo	Togo	Netherlands	Belarus	Iceland	*dropped*
Chile	Iran	Venezuela	*dropped*	New Zealand	Belarus	Iceland	*dropped*
Colombia	Algeria	Venezuela	*dropped*	Nicaragua	Gabon	Venezuela	*dropped*
Costa Rica	Zimbabwe	Venezuela	*dropped*	Niger	Chad	Chad	Chad
Cyprus	Iceland	Iceland	*dropped*	Norway	Belarus	Iceland	*dropped*
Czech R.	Belarus	Belarus	Belarus	Panama	Iceland	Venezuela	*dropped*
Denmark	Belarus	Iceland	*dropped*	Paraguay	Zimbabwe	Venezuela	*dropped*
Dominican R.	Zimbabwe	Venezuela	*dropped*	Peru	Iran	Venezuela	*dropped*
Ecuador	Zimbabwe	Venezuela	*dropped*	Philippines	Algeria	India	India
Egypt	Algeria	Algeria	Algeria	Poland	Belarus	Belarus	Belarus
El Salvador	P.N.G.	Venezuela	*dropped*	Portugal	Belarus	Iceland	*dropped*
Ethiopia	Senegal	Senegal	Senegal	Romania	Ukraine	Ukraine	Ukraine
Finland	Belarus	Iceland	*dropped*	Singapore	Lesotho	China	China
France	Russia	Iceland	*dropped*	Slovak R.	Belarus	Belarus	Belarus
Gambia	Togo	Togo	Togo	Slovenia	Belarus	Belarus	Belarus
Germany	Russia	Iceland	*dropped*	South Africa	Algeria	Zimbabwe	Zimbabwe
Ghana	Gabon	Gabon	Gabon	South Korea	China	China	China
Greece	Belarus	Iceland	*dropped*	Spain	Belarus	Iceland	*dropped*
Guatemala	Senegal	Venezuela	*dropped*	Sri Lanka	Algeria	Pakistan	Pakistan
Guinea	Senegal	Senegal	Senegal	Sweden	Belarus	Iceland	*dropped*
Honduras	Zimbabwe	Venezuela	*dropped*	Switzerland	Belarus	Iceland	*dropped*
Hong Kong SAR	Iran	China	China	Tanzania	Burkina F.	Burkina F.	Burkina F.
Hungary	Belarus	Belarus	Belarus	Thailand	China	China	China
Indonesia	Iran	India	India	Trinidad T.	Zimbabwe	Venezuela	*dropped*
Ireland	Belarus	Iceland	*dropped*	Tunisia	Zimbabwe	Algeria	Algeria
Israel	Iran	Iceland	*dropped*	Turkey	Iran	Iceland	*dropped*
Italy	Russia	Iceland	*dropped*	Uganda	Chad	Chad	Chad
Côte d'Ivoir	Senegal	Senegal	Senegal	United Kingdom	Russia	Iceland	*dropped*
Jamaica	Iceland	Venezuela	*dropped*	United States	China	Iceland	*dropped*
Japan	China	Iceland	*dropped*	Uruguay	Belarus	Venezuela	*dropped*
Jordan	Syria	Syria	Syria	Yemen	Angola	Syria	Syria
Kenya	Senegal	Senegal	Senegal	Zambia	Togo	Togo	Togo

Source: Authors' calculations based on data in Persson and Tabellini (2006).

Note: *Baseline* and *Area* refer to the nearest-neighbor match without and with area dummies, respectively, in the ATE estimation (see Table 2). *Refined* refers to the nearest-neighbor match without Latin America and developed countries in the ATE estimation (see Table 5). C.A.R. and P.N.G. stand for Central African Republic and Papua New Guinea, respectively.

©International Monetary Fund. Not for Redistribution

Table 4. Cross-Country Matches, Control Countries, 1991–2000

Country	Baseline	Area	Refined	Country	Baseline	Area	Refined
Algeria	Colombia	Morocco	Tunisia	Lesotho	Singapore	Cape Verde	Cape Verde
Angola	Tanzania	Tanzania	Guinea	Malawi	Mali	Mali	Mali
Belarus	Czech R.	Czech R.	Czech R.	Nigeria	Guatemala	Kenya	Kenya
Burkina Faso	Tanzania	Tanzania	Tanzania	Pakistan	Morocco	Indonesia	Indonesia
Burundi	Mozambique	Mozambique	Mali	P.N.G.	Guinea	Nepal	Nepal
C.A.R.	Mozambique	Mozambique	Benin	Russia	Italy	Poland	Poland
Chad	Mali	Mali	Mali	Rwanda	Mozambique	Mozambique	Mozambique
China	Brazil	Indonesia	Indonesia	Senegal	Guinea	Guinea	Guinea
Congo	Nicaragua	Ghana	Ghana	Syria	Paraguay	Morocco	Morocco
Gabon	Ghana	Ghana	Ghana	Togo	Benin	Benin	Benin
Iceland	Cyprus	Cyprus	*dropped*	Ukraine	Poland	Poland	Poland
India	Brazil	Indonesia	Indonesia	Venezuela	Morocco	Colombia	*dropped*
Iran	Argentina	Tunisia	Tunisia	Zimbabwe	Ecuador	Zambia	Ghana

Source: Authors' calculations based on data in Persson and Tabellini (2006).

Note: *Baseline* and *Area* refer to the nearest-neighbor match without and with area dummies, respectively, in the ATE estimation (see Table 2). *Refined* refers to the nearest-neighbor match without Latin America and developed countries in the ATE estimation (see Table 5). C.A.R. and P.N.G. stand for Central African Republic and Papua New Guinea, respectively.

economies (where being a developed country takes precedence over geographic region).[15] The matches underlying the estimation with these area dummies are reported in the *Area* column. For Albania and Sri Lanka in the 1990s, this step appears to work reasonably well, as they are now matched with Belarus and Pakistan, which are certainly perceived as more similar than the baseline-nearest neighbors. There are, however, certain surprising findings: for example, all developed countries are matched with Iceland! A similar result obtains for Latin America, where Venezuela is the only control that is picked to be a match. This is due to the fact that Iceland and Venezuela are, according to the SWWW classification, the only closed economies in the group of developed countries and Latin America in the 1990s.[16] In other words, there is no common support between treated and control countries in those two regions. Introducing area dummies is not

[15]We label as *developed* all countries that joined the Organization for Economic Cooperation and Development (OECD) between its foundation in 1961 and 1973—the end of the initial participation wave that concluded with the accession of New Zealand. In addition, we add Cyprus and Israel for lack of better options. Countries that joined the OECD more recently (starting with Mexico in 1994) are allocated to their geographic region. The label *transition* is used for all countries in central and eastern Europe that are contained in the sample, including for the period before 1990.

[16]See the online appendix mentioned in footnote 14 for a precise list of treated and untreated countries across regions.

©International Monetary Fund. Not for Redistribution

enough to control for area-specific unobservables, unless there is a sufficient overlap of treated and untreated countries in all areas.

Summing up, the matches listed in the *Baseline* and *Area* columns of Tables 3 and 4 for the 1990s (and in the online appendix for the other decades) show that country comparisons underlying cross-country analysis are often more than far-fetched. This unbearable lightness of cross-country analysis extends from matching to other cross-sectional estimators that rely on the unconfoundedness assumption, such as plain regression modeling. This is due to the fact that OLS estimates are based either on the same implicit but far-fetched country comparisons or—even worse—on parametric extrapolation beyond the region of common support. In fact, if treated and control countries are very different from each other with respect to covariates, the OLS estimate of the counterfactual outcome of the treated is constructed by linearly extrapolating the observed outcome of control units, and vice versa.

Refined Evidence in Selected Samples

The above discussion shows that—as long as we want to control for area-specific unobservable characteristics—we should restrict the analysis of the trade-growth nexus to regions with enough treatment variation. In other words, to improve the quality of the country matches underlying the results, we should drop regions with no common support between treated and control units.

In Table 5, we reestimate the pooled specification eliminating countries that lack common support with respect to regional affiliation. As a criterion, we establish that the ratio between treated and control countries (or vice versa) should be below 10; that is, for every 10 treated countries, we require more than 1 potential control country (or vice versa). As shown in the online appendix, this requirement eliminates the group of developed economies and Latin America in the 1990s because (almost) all of them are open. The same holds for other regions and other decades, however: in the 1980s, almost all developed economies are already open. In the 1970s and 1960s, almost all African economies (except Mauritius in the 1970s) are closed according to the SWWW dummy. Moreover, transition economies are excluded prior to the 1990s, as none of them is in fact in transition (that is, open).

Table 5 reports matching estimates—of both the ATE and ATT—restricted to countries that meet the common-support condition for geographic areas. Comparing these estimates to the previous ones for the unrestricted sample (Table 2), the coefficients appear to be slightly more significant and also larger in magnitude in the 1990s and 1970s. Moreover, we now find stronger evidence of a marginally significant positive effect of openness on growth in the 1980s, especially for the countries that were open to trade (ATT estimate). All the estimated effects lie in the 1.5 to 2.5 percentage point range. For the 1960s, again, the coefficients are never

©International Monetary Fund. Not for Redistribution

significantly different from zero. The *Refined* column in Tables 3 and 4 reports the country matches underlying these results. Counterintuitive matches are now considerably reduced.

We conclude from this exercise that it is important to check for the existence of common support. In fact, in small samples of countries, the advantage of matching estimators lies in the guidance for appropriately restricting the analysis to specific subsamples. Unlike the estimates in Table 2—which replicate common results from growth regressions in the literature—the estimates presented in Table 5 fully control for area-specific unobservables and are based on more plausible country comparisons. There is no free lunch, however, as the external validity of the estimates is now

Table 5. Openness and Growth, Refined Evidence (I), 1961–2000
ATE: $E[Y(1)-Y(0)|X]$; ATT: $E[Y(1)-Y(0)|T=1,X]$

	1991–2000	1981–90	1971–80	1961–70
Matching without area dummies				
Estimate (ATE)	1.921**	0.949	1.841**	0.084
SE	(0.647)	(0.706)	(0.634)	(0.666)
Estimate (ATT)	1.760**	2.358*	2.229**	0.613
SE	(0.633)	(0.959)	(0.625)	(0.915)
Matching with area dummies				
Estimate (ATE)	1.389*	0.809	2.240**	−0.143
SE	(0.708)	(0.526)	(0.684)	(0.680)
Estimate (ATT)	0.849	1.252*	1.986*	0.544
SE	(0.738)	(0.524)	(0.770)	(0.924)
Africa	Yes (19–14)	Yes (5–30)	No (1–33)	No (0–35)
Asia	Yes (9–4)	Yes (6–8)	Yes (6–8)	Yes (4–10)
Latin America	No (19–1)	Yes (6–17)	Yes (4–19)	Yes (5–18)
Middle East	Yes (6–3)	Yes (2–6)	Yes (1–7)	Yes (1–7)
Developed countries	No (25–1)	No (24–2)	Yes (21–4)	Yes (21–4)
Transition economies	Yes (9–3)	No (0–3)	No (0–3)	No (0–1)
Treated	43	19	32	31
Controls	24	61	38	39
Observations	67	80	70	70

Source: Authors' calculations based on data in Persson and Tabellini (2006).

Note: Dependent variable (Y): real GDP per capita growth. Treatment indicator (T): trade openness dummy (Sachs and Warner, 1995; Wacziarg and Welch, 2003). Control variables (X) include initial GDP per capita, secondary school enrollment, population growth, investment share, and area dummies (as indicated). The two numbers in parentheses after each area refer to the number of treated and control countries, respectively. Samples restricted to certain areas to meet the common-support condition for area dummies. ATE and ATT stand for average treatment effect and average treatment effect on the treated, respectively, and are estimated by nonparametric nearest-neighbor matching. *corresponds to 5 percent significance level; **corresponds to 1 percent significance level.

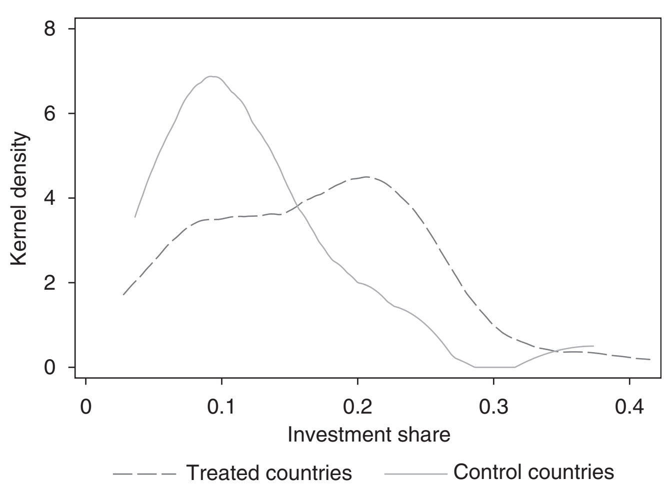
©International Monetary Fund. Not for Redistribution

reduced. The results recommend refraining from commenting on the effect of trade openness on growth in developed countries after 1980, in Africa before 1980, and in Latin America after 1990.

The estimates in Table 5 control for the existence of common support with respect to a set of covariates that we deem important to capture unobservable regional characteristics associated to geography, level of development, culture, or legal origins—that is, area dummies for Africa, Middle East, Asia, Latin America, transition economies, and developed economies. The common support, however, should also be checked for other covariates. In principle, we would like to match countries that are very similar with respect to all covariates, but this is impossible if treated and control units are not evenly distributed across all the ranges of variation of covariates. Figures 1–8 show that, for example, this condition is not often met for investment share and secondary school enrollment. These figures report the kernel density of treated and control countries over the ranges of variation of these two variables. For instance, Figure 5 shows that the common support for investment share in the 1970s ranges from 0.11 to 0.39, with 27 (control) countries below this region and 1 (treated) country above. To meet the common-support condition, these 28 countries should be dropped from the estimation sample.

Table 6 reports matching estimates for samples restricted to the regions of common support identified in Figures 1–8. This evidence is consistent with the one described in Table 5. When carefully matching only countries that lie in the common support, cross-sectional estimates using the SWWW dummy

Figure 1. Common Support for Investment Share, 1991–2000

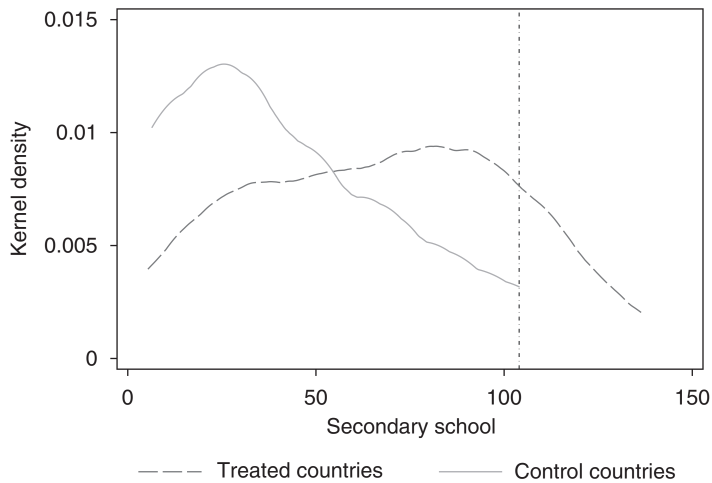


Source: Authors' calculations based on data in Persson and Tabellini (2006).

Note: Kernel density of investment share in the 1990s for both treated and control observations. Treated countries: 87. Control countries: 26. All countries in common support.

©International Monetary Fund. Not for Redistribution

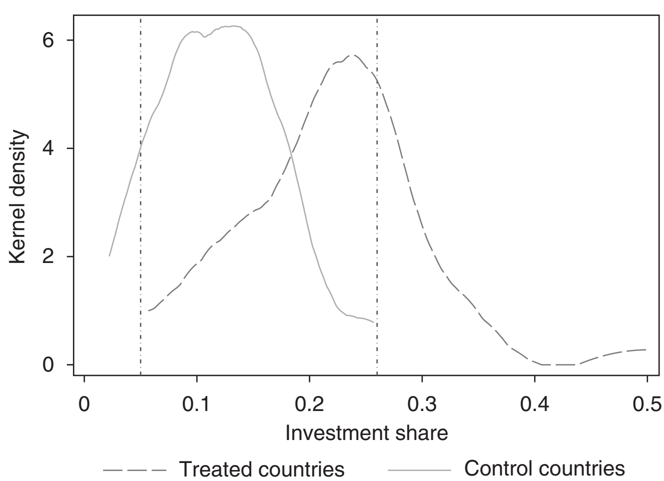
Figure 2. Common Support for Secondary School Enrollment, 1991–2000



Source: Authors' calculations based on data in Persson and Tabellini (2006).

Note: Kernel density of secondary school enrollment in the 1990s for both treated and control observations. Treated countries: 87. Control countries: 26. Common support: (0, 104). Countries above common support: 14 treated.

Figure 3. Common Support for Investment Share, 1981–90



Source: Authors' calculations based on data in Persson and Tabellini (2006).

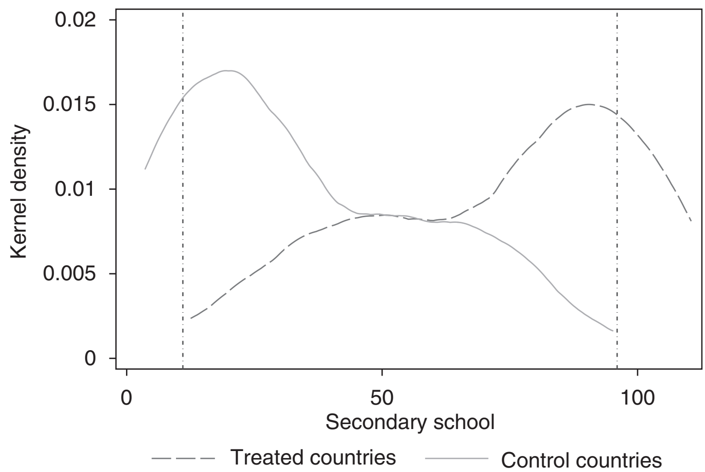
Note: Kernel density of investment share in the 1980s for both treated and control observations. Treated countries: 43. Control countries: 66. Common support: (0.05, 0.26). Countries above common support: 10 treated. Countries below common support: 6 controls.

continue to detect a positive and significant association between openness and growth in the 1970s, 1980s, and 1990s, but not in the 1960s.

Finally, we are aware that—even though our refined estimates improve the internal validity of cross-country results by checking for common

©International Monetary Fund. Not for Redistribution

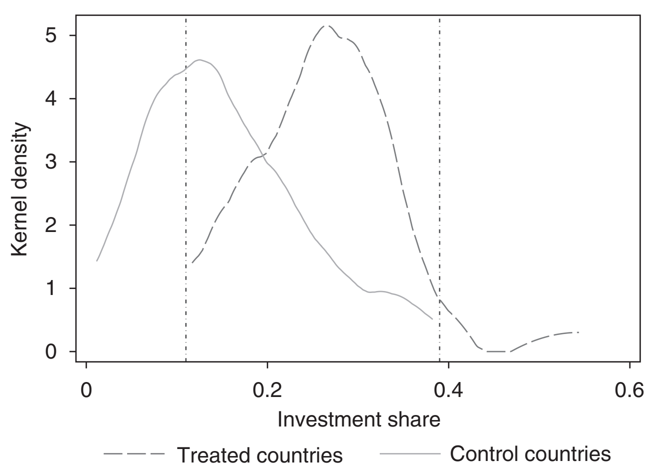
Figure 4. Common Support for Secondary School Enrollment, 1981–90



Source: Authors' calculations based on data in Persson and Tabellini (2006).

Note: Kernel density of secondary school enrollment in the 1980s for both treated and control observations. Treated countries: 43. Control countries: 66. Common support: (11, 96). Countries above common support: 11 treated. Countries below common support: 11 controls.

Figure 5. Common Support for Investment Share, 1971–80



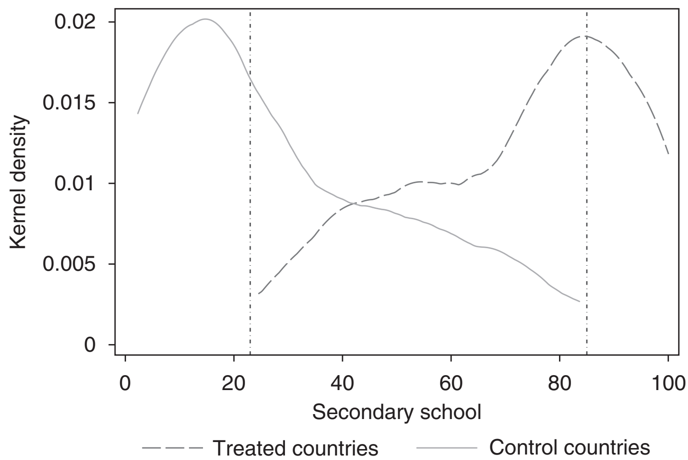
Source: Authors' calculations based on data in Persson and Tabellini (2006).

Note: Kernel density of investment share in the 1970s for both treated and control observations. Treated countries: 33. Control countries: 74. Common support: (0.11, 0.39). Countries above common support: 1 treated. Countries below common support: 27 controls.

support—our results still suffer from the fact that country-specific unobservables (that is, endogenous selection into treatment) might violate the conditional independence assumption. By the same token, if conditional independence is not met, matching estimates should not be interpreted as

©International Monetary Fund. Not for Redistribution

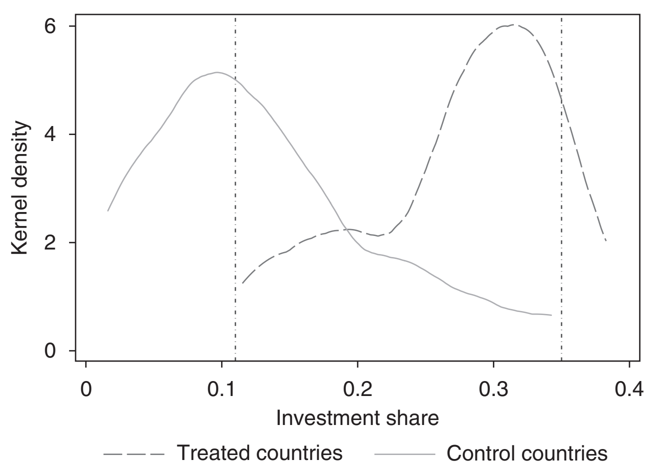
Figure 6. Common Support for Secondary School Enrollment, 1971–80



Source: Authors' calculations based on data in Persson and Tabellini (2006).

Note: Kernel density of secondary school enrollment in the 1970s for both treated and control observations. Treated countries: 33. Control countries: 74. Common support: (23, 85). Countries above common support: 11 treated. Countries below common support: 38 controls.

Figure 7. Common Support for Investment Share, 1961–70



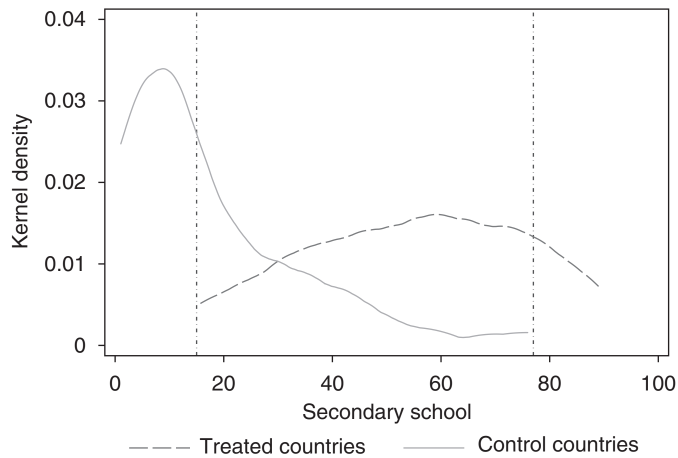
Source: Authors' calculations based on data in Persson and Tabellini (2006).

Note: Kernel density of investment share in the 1960s for both treated and control observations. Treated countries: 31. Control countries: 75. Common support: (0.11, 0.35). Countries above common support: 4 treated. Countries below common support: 32 controls.

causal effects, because the direction of causality is unclear. The main message of this paper is somewhat different, however: matching estimators clarify the importance of controlling for common support in the covariate space, especially in macro samples, as they tend to be comparatively small.

©International Monetary Fund. Not for Redistribution

Figure 8. Common Support for Secondary School Enrollment, 1961–70



Source: Authors' calculations based on data in Persson and Tabellini (2006).

Note: Kernel density of secondary school enrollment in the 1960s for both treated and control observations. Treated countries: 31. Control countries: 75. Common support: (15, 77). Countries above common support: 5 treated. Countries below common support: 45 controls.

This point extends to state-of-the-art estimation techniques. Panel methods, for example, can overcome some of the OLS weaknesses by using within-country (that is, time-series) information to control for unobservable time-invariant country characteristics. However, as long as they still use some cross-sectional variation—as does, for example, the diff-in-diff estimator—they still suffer from insufficient transparency that could hide a lack of common support.

More Evidence Using an Alternative Indicator of Trade Restrictions

The trade restriction measure compiled by DeJong and Ripoll (2006) consists of annual observations from 1975 to 2000 of *ad valorem* tariffs—import duties as a share of imports—for 74 countries. To match the structure of our 10-year pooled data set, we discard the beginning of the sample (which also displays a number of missing observations) and use the averages for 1981–90 and 1991–2000, consistent with the framework above. Contrary to the SWWW dummy, the trade barrier measure is not binary, and a threshold needs to be introduced to distinguish ‘‘open'' (that is, treated) from ‘‘closed'' (that is, control) countries. In the baseline specification, we use the median tariff rate as the threshold; that is, countries with a lower average protection over the decade are regarded as open; countries with a higher import duties ratio as closed. We also discuss results that emerge if we set the threshold at the first and third quartile of the tariff level distribution. After merging the PT and DeJong-Ripoll data sets, we are left with 68 countries for the 1990s

©International Monetary Fund. Not for Redistribution

Table 6. Openness and Growth, Refined Evidence (II), 1961–2000
ATE: $E[Y(1)-Y(0)|X]$; ATT: $E[Y(1)-Y(0)|T=1,X]$

	1991–2000	1981–90	1971–80	1961–70
Matching with common support A				
Estimate (ATE)	1.318*	0.884	1.916**	−0.273
SE	(0.672)	(0.510)	(0.743)	(0.705)
Estimate (ATT)	1.130	0.813	1.645*	0.513
SE	(0.742)	(0.627)	(0.686)	(1.049)
Matching with common support B				
Estimate (ATE)	1.644*	1.139*	2.007**	−0.360
SE	(0.731)	(0.524)	(0.717)	(0.641)
Estimate (ATT)	1.407	2.019**	1.936**	0.034
SE	(0.819)	(0.747)	(0.690)	(0.833)
Matching with common support C				
Estimate (ATE)	1.644*	1.141*	1.774*	−0.203
SE	(0.731)	(0.531)	(0.757)	(0.678)
Estimate (ATT)	1.407	1.691*	1.619*	0.304
SE	(0.819)	(0.725)	(0.651)	(0.843)

Source: Authors' calculations based on data in Persson and Tabellini (2006).

Note: Dependent variable (Y): real GDP per capita growth. Treatment indicator (T): trade openness dummy (Sachs and Warner, 1995; Wacziarg and Welch, 2003). Control variables (X) include initial GDP per capita, secondary school enrollment, population growth, investment share, and area dummies. Restricted samples to meet the common-support condition for investment share (A), secondary school enrollment (B), or both (C). See Figures 1–8 for the numbers of treated and control countries dropped because outside of common supports A, B, and C. ATE and ATT stand for average treatment effect and average treatment effect on the treated, respectively, and are estimated by nonparametric nearest-neighbor matching. *corresponds to 5 percent significance level; **corresponds to 1 percent significance level.

and 71 for the 1980s, significantly less than for the SWWW measure (which contains 113 and 109 observations in the 1990s and 1980s, respectively).

Although not key to our argument, we briefly discuss the estimation results based on this alternative indicator of trade liberalization. Table 7 shows the cross-country evidence for the DeJong-Ripoll measure, corresponding to Table 2 for the SWWW indicator. For *tariff 1*, the baseline (median) tariff threshold, the results, although statistically insignificant, would indicate that in a treated country—that is, a country with a tariff-to-imports ratio below the median—the annual growth rate of real per capita GDP in the 1990s is up to 1.2 percent lower than in the control country. Compared with the results in Table 2, there is a somewhat larger discrepancy between OLS and matching estimates. In the 1980s, the link between openness and trade flips to become mostly positive, but again largely insignificant, except for the ATT estimator without area dummies.

©International Monetary Fund. Not for Redistribution

Table 7. Tariffs and Growth, Cross-Country Evidence, 1981–2000
OLS: $E[Y|X]=\alpha+\tau T+\beta X$; ATE: $E[Y(1)-Y(0)|X]$; ATT: $E[Y(1)-Y(0)|T=1,X]$

	1991–2000			1981–90		
	tariff1	tariff2	tariff3	tariff1	tariff2	tariff3
OLS without area dummies						
Estimate of τ	−1.093	−0.534	−1.484*	−0.318	−0.470	−1.200
SE	(0.556)	(0.534)	(0.659)	(0.732)	(0.772)	(0.687)
Adjusted R^2	0.22	0.19	0.26	0.31	0.31	0.34
OLS with area dummies						
Estimate of τ	−0.178	0.074	−1.139	0.044	−0.577	−0.450
SE	(0.467)	(0.459)	(0.887)	(0.633)	(0.717)	(0.724)
Adjusted R^2	0.41	0.41	0.44	0.52	0.52	0.52
Matching without area dummies						
Estimate (ATE)	−0.311	0.832	−1.561*	1.285	1.149	1.201
SE	(0.533)	(0.673)	(0.720)	(0.992)	(1.098)	(1.456)
Estimate (ATT)	−0.034	0.241	−1.656	3.234*	−0.093	2.155
SE	(0.675)	(0.582)	(0.845)	(1.375)	(1.049)	(1.775)
Matching with area dummies						
Estimate (ATE)	−0.789	1.226*	−1.542*	0.311	1.427	1.505
SE	(0.523)	(0.620)	(0.709)	(0.735)	(0.966)	(1.348)
Estimate (ATT)	−1.188	0.318	−1.865*	1.023	0.262	2.450
SE	(0.675)	(0.614)	(0.810)	(0.878)	(0.865)	(1.629)
Treated	34	17	51	35	18	53
Controls	34	51	17	36	53	18
Observations	68	68	68	71	71	71

Source: Authors' calculations based on data in Persson and Tabellini (2006) and DeJong and Ripoll (2006).

Note: Dependent variable (Y): real GDP per capita growth. Treatment indicators (T): *tariff1* equal to 1 if import duties as a percentage of imports lower than the sample median; *tariff2* equal to 1 if import duties as a percentage of imports lower than the sample 25th percentile; *tariff3* equal to 1 if import duties as a percentage of imports lower than the sample 75th percentile. Control variables (X) include initial GDP per capita, secondary school enrollment, population growth, and investment share. Area dummies refer to Africa, Asia, Latin America, Middle East, developed countries, and transition economies. ATE and ATT stand for average treatment effect and average treatment effect on the treated, respectively, and are estimated by nonparametric nearest-neighbor matching. *corresponds to 5 percent significance level; **corresponds to 1 percent significance level.

The other threshold levels also offer conflicting and mainly weak signals: for *tariff 2*—only countries with a tariff-to-imports ratio below the 25th percentile of the distribution are defined as open—the results broadly indicate a positive effect of a lower tariff level on growth in the 1990s and are mixed for the 1980s. For *tariff 3*—only countries with a tariff-to-imports ratio below the 75th percentile of the distribution are defined as open—the

©International Monetary Fund. Not for Redistribution

Table 8. Tariffs and Growth, Refined Evidence, 1981–2000
ATE: $E[Y(1)-Y(0)|X]$; ATT: $E[Y(1)-Y(0)|T=1,X]$

	1991–2000			1981–90		
	tariff1	tariff2	tariff3	tariff1	tariff2	tariff3
Matching without area dummies						
Estimate (ATE)	−0.320	0.341	−0.449	−0.663	0.296	−0.274
SE	(0.604)	(0.693)	(0.693)	(0.980)	(0.835)	(0.735)
Estimate (ATT)	−0.353	0.087	−0.451	0.327	0.308	0.344
SE	(0.686)	(0.589)	(0.865)	(1.322)	(0.867)	(0.846)
Matching with area dummies						
Estimate (ATE)	−0.018	—	−0.163	−0.452	0.188	−1.090
SE	(0.506)	—	(0.667)	(0.751)	(0.878)	(0.642)
Estimate (ATT)	0.054	—	0.374	0.708	0.328	−0.658
SE	(0.620)	—	(0.720)	(0.760)	(0.953)	(0.878)
Africa	No (1–11)	No (0–12)	Yes (4–8)	Yes (2–13)	No (0–15)	Yes (4–11)
Asia	Yes (4–7)	No (1–10)	Yes (8–3)	Yes (4–7)	No (1–10)	Yes (9–2)
Latin America	Yes (5–9)	No (0–14)	No (13–1)	Yes (5–9)	No (0–14)	No (13–1)
Middle East	No (0–6)	No (0–6)	Yes (1–5)	No (0–6)	No (0–6)	Yes (2–4)
Developed countries	No (23–0)	Yes (16–7)	No (23–0)	No (22–1)	Yes (15–8)	No (23–0)
Transition economies	Yes (1–1)	No (0–2)	No (2–0)	No (2–0)	Yes (1–1)	No (2–0)
Treated	10	16	13	11	16	15
Controls	17	7	16	29	9	17
Observations	27	23	29	40	25	32

Source: Authors' calculations based on data in Persson and Tabellini (2006) and DeJong and Ripoll (2006).

Note: Dependent variable (Y): real GDP per capita growth. Treatment indicators (T): *tariff1* equal to 1 if import duties as a percentage of imports lower than the sample median; *tariff2* equal to 1 if import duties as a percentage of imports lower than the sample 25th percentile; *tariff3* equal to 1 if import duties as a percentage of imports lower than the sample 75th percentile. Control variables (X) include: initial GDP per capita, secondary school enrollment, population growth, investment share, and area dummies (as indicated). The two numbers in parentheses after each area refer to the number of treated and control countries, respectively. Samples restricted to certain areas to meet the common-support condition for area dummies. ATE and ATT stand for average treatment effect and average treatment effect on the treated, respectively, and are estimated by nonparametric nearest-neighbor matching. *corresponds to 5 percent significance level; **corresponds to 1 percent significance level.

results indicate a significant negative effect of trade liberalization on growth in the 1990s and mixed effects in the 1980s.

More importantly for our purposes, the country pairings stemming from the matching exercise are similarly far-fetched in the case of the DeJong-Ripoll tariff measures as in the case of the SWWW trade liberalization

©International Monetary Fund. Not for Redistribution

dummy.[17] In Table 8, we proceed to eliminate regions that lack common support with respect to regional affiliation; that is, where there are not enough treated compared with control countries (or vice versa) according to the criterion established above. Due to the smaller sample size of the DeJong-Ripoll data set, the number of observations drops further, and amounts now to less than half of the observations for the SWWW dummy (see Table 5). For example, the group of developed countries is excluded from the estimation for the median and 75th percentile threshold specification, whereas it is included for the 25th percentile threshold of the tariff-to-GDP ratio, implying that developed economies on average have a rather low tariff level and can only be included in the restricted sample if countries with a tariff ratio above the 25th percentile are coded closed. All estimates are insignificant and there is no clear directional effect: in precisely 50 percent of the cases (11), the impact is positive, in the rest negative. Using *tariff 2*—that is, estimating a country sample based exclusively on the group of developed countries in the 1990s and almost exclusively in the 1980s—the impact of trade liberalization is unambiguously positive (but insignificant), qualitatively confirming a result in DeJong and Ripoll (2006), who find evidence of a negative relationship between tariffs and growth only among the world's rich countries.

To sum up, for the DeJong-Ripoll measure of tariff barriers, and after controlling for common support, the sample size drops drastically and makes sound statistical inference difficult. Although the unrestricted sample shows a more pronounced negative effect of trade openness (in the sense of a low tariff level) on growth, the appropriately restricted sample does not convey any strong message and is even consistent with the opposite affirmation.

IV. Conclusions

In this paper, we take another look at the openness-growth nexus in international macroeconomics. To add empirical *glasnost* to the results obtained in the literature, we examine classic pooled cross-country regressions and show the pitfalls related to the underlying country comparisons. Employing matching estimators from the treatment evaluation literature, we show that the country matches behind the estimates are often far-fetched—the unbearable lightness of cross-country estimates. We explain this problem as a lack of overlap between open and closed countries in the covariate space.

We show that restricting the sample to treated and control countries that share a common support is not always feasible due to data restrictions related to the openness measure or the data set employed. When this restriction can be applied, as is the case of the SWWW openness indicator, we confirm a positive correlation between trade openness and growth in selected regions after 1970. As the conditional independence assumption is likely not to hold

[17]See Tables 9 through 12 in the online appendix.

©International Monetary Fund. Not for Redistribution

also in samples restricted to meet the common-support condition, however, we cannot interpret this correlation as a causal effect.

The main argument of the paper, however, goes far beyond the evidence presented. The lack of *glasnost* identified above is not limited to OLS-type estimates but extends to any econometric framework that uses at least some cross-sectional variation. Controlling for common support should, hence, be part of any empirical strategy in macro cross-country investigations.

REFERENCES

Abadie, A., D. Drukker, J.L. Herr, and G.W. Imbens, 2004, ‘‘Implementing Matching Estimators for Average Treatment Effects in Stata,’’ *Stata Journal*, Vol. 4, No. 3, pp. 290–311.

Atoyan, R., and P. Conway, 2006, ‘‘Evaluating the Impact of IMF Programs: A Comparison of Matching and Instrumental-Variable Estimators,’’ *Review of International Organizations*, Vol. 1, No. 2, pp. 99–124.

Barro, R.J., 1991, ‘‘Economic Growth in a Cross Section of Countries,’’ *Quarterly Journal of Economics*, Vol. 106, No. 2, pp. 407–33.

Bhagwati, J., 2002, *Free Trade Today* (Princeton, New Jersey, Princeton University Press).

______ and T.N. Srinivasan, 2001, ‘‘Outward-Orientation and Development: Are Revisionists Right?’’ in *Trade, Development and Political Economy: Essays in Honor of Anne Krueger*, ed. by D.K. Lal and R. Snape (Basingstoke, United Kingdom and New York, Palgrave).

______, 2002, ‘‘Trade and Poverty in the Poor Countries,’’ *American Economic Review Papers and Proceedings*, Vol. 92, No. 2, pp. 180–3.

DeJong, D.N., and M. Ripoll, 2006, ‘‘Tariffs and Growth: An Empirical Exploration of Contingent Relationships,’’ *Review of Economics and Statistics*, Vol. 88, No. 4, pp. 625–40.

Dollar, D., 1992, ‘‘Outward Oriented Developing Economies Really Do Grow More Rapidly: Evidence from 95 LDCs, 1976–1985,’’ *Economic Development and Cultural Change*, Vol. 40, No. 3, pp. 523–44.

______, and A. Kraay, 2003, ‘‘Institutions, Trade, and Growth,’’ *Journal of Monetary Economics*, Vol. 50, No. 1, pp. 133–62.

Edwards, S., 1992, ‘‘Trade Orientation, Distortions, and Growth in Developing Countries,’’ *Journal of Development Economics*, Vol. 39, No. 1, pp. 31–57.

______, 1998, ‘‘Openness, Productivity, and Growth: What Do We Really Know?’’ *Economic Journal*, Vol. 108, No. 2, pp. 383–98.

______, and I. Magendzo, 2003, ‘‘Dollarization and Economic Performance: What Do We Really Know?’’ *International Journal of Finance & Economics*, Vol. 8, No. 4, pp. 351–63.

Frankel, J.A., and D. Romer, 1999, ‘‘Does Trade Cause Growth?’’ *American Economic Review*, Vol. 89, No. 3, pp. 379–99.

Giavazzi, F., and G. Tabellini, 2005, ‘‘Economic and Political Liberalizations,’’ *Journal of Monetary Economics*, Vol. 52, No. 7, pp. 1297–330.

©International Monetary Fund. Not for Redistribution

Harrison, A., 1996, ‘‘Openness and Growth: A Time-Series, Cross-Country Analysis for Developing Countries,’’ *Journal of Development Economics*, Vol. 48, No. 2, pp. 419–47.

Imbens, G.W., 2004, ‘‘Nonparametric Estimation of Average Treatment Effects under Exogeneity: A Review,’’ *Review of Economics and Statistics*, Vol. 86, No. 1, pp. 4–29.

Irwin, D.A., and M. Tervio, 2002, ‘‘Does Trade Raise Income? Evidence from the Twentieth Century,’’ *Journal of International Economics*, Vol. 58, No. 1, pp. 1–18.

Lee, H.Y., L.A. Ricci, and R. Rigobon, 2004, ‘‘Once Again, Is Openness Good for Growth?’’ *Journal of Development Economics*, Vol. 75, No. 2, pp. 451–72.

Levine, R., and D. Renelt, 1992, ‘‘A Sensitivity Analysis of Cross-Country Growth Regressions,’’ *American Economic Review*, Vol. 82, No. 4, pp. 942–63.

Persson, T., and G. Tabellini, 2003, *The Economic Effects of Constitutions* (Cambridge, Massachusetts, MIT Press).

———, 2006, ‘‘Democracy and Development: The Devil in the Details,’’ *American Economic Review, Papers and Proceedings*, Vol. 96, No. 2, pp. 319–24.

Rigobon, R., and D. Rodrik, 2005, ‘‘Rule of Law, Democracy, Openness, and Income,’’ *Economics of Transition*, Vol. 13, No. 3, pp. 533–64.

Rodriguez, F., 2006, ‘‘Openness and Growth: What Have We Learned?’’ Wesleyan Economics Working Papers No. 2006-11 (Middletown, Connecticut, Wesleyan University).

———, and D. Rodrik, 2001, ‘‘Trade Policy and Economic Growth: A Skeptics Guide to the Cross-National Evidence,’’ in *NBER Macroeconomics Annual 2000*, ed. by B. Bernanke and K. Rogoff (Cambridge, Massachusetts, MIT Press).

Rubin, D., 1974, ‘‘Estimating Causal Effects of Treatments in Randomised and Non-Randomised Studies,’’ *Journal of Educational Psychology*, Vol. 66, No. 5, pp. 688–701.

———, 1980, ‘‘Bias Reduction Using Mahalanobis-Metric Matching,’’ *Biometrics*, Vol. 36, No. 2, pp. 293–8.

Sachs, J.D., and A. Warner, 1995, ‘‘Economic Reform and the Process of Global Integration,’’ *Brookings Papers in Economic Activity*: Vol. I, pp. 1–18.

Salinas, G., and A. Aksoy, 2006, ‘‘Growth before and after Trade Liberalization,’’ World Bank Policy Research Working Paper 4062 (Washington, World Bank).

Slaughter, M.J., 2001, ‘‘Trade Liberalization and Per Capita Income Convergence: A Difference-in-Differences Analysis,’’ *Journal of International Economics*, Vol. 55, No. 1, pp. 203–28.

Temple, J., 2000, ‘‘Growth Regressions and What the Textbooks Don’t Tell You,’’ *Bulletin of Economic Research*, Vol. 52, No. 3, pp. 181–205.

Vamvakidis, A., 2002, ‘‘How Robust Is the Growth Openness Connection? Historical Evidence,’’ *Journal of Economic Growth*, Vol. 7, No. 1, pp. 177–94.

Wacziarg, R., and K.H. Welch, 2003, ‘‘Trade Liberalization and Growth: New Evidence,’’ NBER Working Paper 10152 (Cambridge, Massachusetts, National Bureau of Economic Research).

Winters, L.A., 2004, ‘‘Trade Liberalization and Economic Performance: An Overview,’’ *Economic Journal*, Vol. 114 (February), pp. F4–F21.

©International Monetary Fund. Not for Redistribution