

Francisco Rodríguez and Dani Rodrik

UNIVERSITY OF MARYLAND AND HARVARD UNIVERSITY

Trade Policy and Economic Growth: A Skeptic's Guide to the Cross-National Evidence

It isn't what we don't know that kills us. It's what we know that ain't so.

—Mark Twain

1. Introduction

Do countries with lower barriers to international trade experience faster economic progress? Few questions have been more vigorously debated in the history of economic thought, and none is more central to the vast literature on trade and development.

The prevailing view in policy circles in North America and Europe is that recent economic history provides a conclusive answer in the affirmative. Multilateral institutions such as the World Bank, the IMF, and the OECD regularly promulgate advice predicated on the belief that openness generates predictable and positive consequences for growth. A recent report by the OECD (1998, p. 36) states: “More open and outward-oriented economies consistently outperform countries with restrictive trade and [foreign] investment regimes.” According to the IMF (1997, p.

We thank Dan Ben-David, Sebastian Edwards, Jeffrey Frankel, David Romer, Jeffrey Sachs, and Andrew Warner for generously sharing their data with us. We are particularly grateful to Ben-David, Frankel, Romer, Sachs, Warner, and Romain Wacziarg for helpful email exchanges. We have benefited greatly from discussions in seminars at the University of California at Berkeley, University of Maryland, University of Miami, University of Michigan, MIT, the Inter-American Development Bank, Princeton, Yale, IMF, IESA, and the NBER. We also thank Ben Bernanke, Roger Betancourt, Allan Drazen, Gene Grossman, Ann Harrison, Chang-Tai Hsieh, Doug Irwin, Chad Jones, Frank Levy, Douglas Irwin, Rick Mishkin, Arvind Panagariya, Ken Rogoff, James Tybout, and Eduardo Zambrano for helpful comments, Vladimir Kliouev for excellent research assistance, and the Weatherhead Center for International Affairs at Harvard for partial financial support.

84): "Policies toward foreign trade are among the more important factors promoting economic growth and convergence in developing countries."

This view is widespread in the economics profession as well. Krueger (1998, p. 1513), for example, judges that it is straightforward to demonstrate empirically the superior growth performance of countries with "outer-oriented" trade strategies. According to Stiglitz (1998, p. 36), "[m]ost specifications of empirical growth regressions find that some indicator of external openness—whether trade ratios or indices or price distortions or average tariff level—is strongly associated with per-capita income growth." According to Fischer (2000), "[i]ntegration into the world economy is the best way for countries to grow."

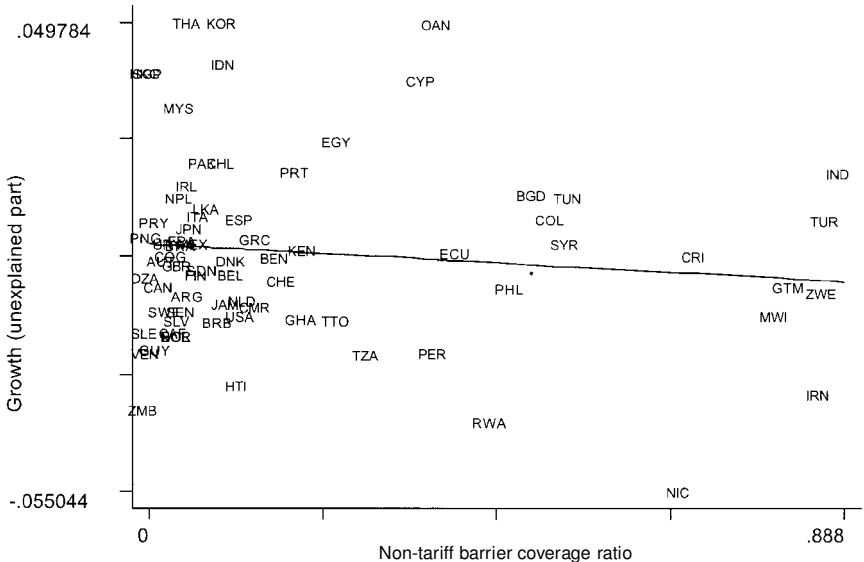
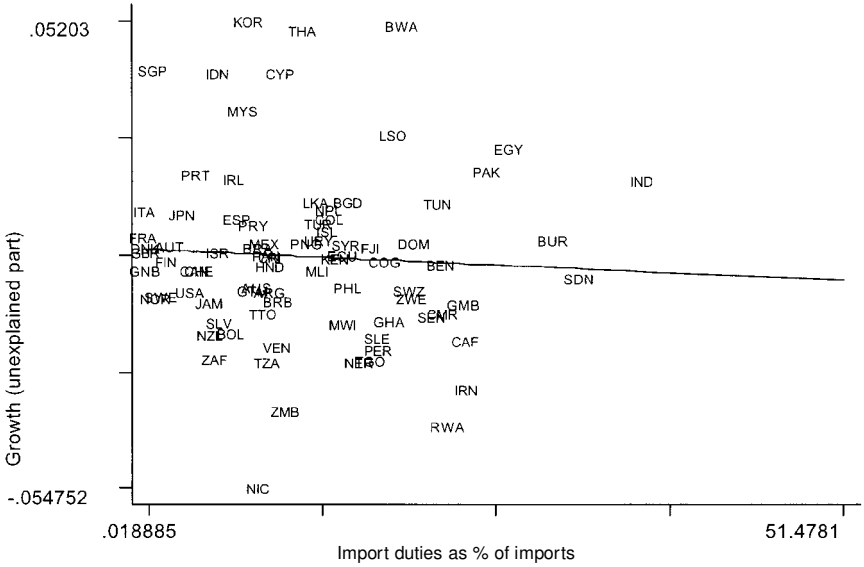
Such statements notwithstanding, if there is an inverse relationship between trade barriers and economic growth, it is not one that immediately stands out in the data. See for example Figure 1. The figure displays the (partial) associations over 1975–1994 between the growth rate of per capita GDP and two measures of trade restrictions. The first measure is an average tariff rate, calculated by dividing total import duties by the volume of imports. The second is a coverage ratio for nontariff barriers to trade.¹ The figures show the relationship between these measures and growth after controlling for levels of initial income and secondary education. In both cases, the slope of the relationship is only slightly negative and nowhere near statistical significance. This finding is not atypical. Simple measures of trade barriers tend not to enter significantly in well-specified growth regressions, regardless of time periods, subsamples, or the conditioning variables employed.

Of course, neither of the two measures used above is a perfect indicator of trade restrictions. Simple tariff averages underweight high tariff rates because the corresponding import levels tend to be low. Such averages are also poor proxies for overall trade restrictions when tariff and nontariff barriers are substitutes. As for the nontariff coverage ratios, they do not do a good job of discriminating between barriers that are highly restrictive and barriers with little effect. And conceptual flaws aside, both indicators are clearly measured with some error (due to smuggling, weaknesses in the underlying data, coding problems, etc.).

In part because of concerns related to data quality, the recent literature on openness and growth has resorted to more creative empirical strategies. These strategies include: (1) constructing alternative indicators of openness (Dollar, 1992; Sachs and Warner, 1995); (2) testing robustness by using a wide range of measures of openness, including subjective indica-

1. Data for the first measure come from World Bank (1998). The second is taken from Barro and Lee (1994), and is based on UNCTAD compilations.

Figure 1 PARTIAL ASSOCIATION BETWEEN GROWTH AND DIRECT MEASURES OF TRADE RESTRICTIONS



tors (Edwards, 1992, 1998); and (3) comparing convergence experience among groups of liberalizing and nonliberalizing countries (Ben-David, 1993). This recent round of empirical research is generally credited with having yielded stronger and more convincing results on the beneficial consequences of openness than the previous, largely case-based literature. Indeed, the cumulative evidence that has emerged from such studies provides the foundation for the previously noted consensus on the growth-promoting effects of trade openness. The frequency with which these studies are cited in international economics textbooks and in policy discussions is one indicator of the influence that they have exerted.

Our goal in this paper is to scrutinize this new generation of research. We do so by focusing on what the existing literature has to say on the following question: *Do countries with lower policy-induced barriers to international trade grow faster, once other relevant country characteristics are controlled for?* We take this to be the central question of policy relevance in this area. To the extent that the empirical literature demonstrates a positive causal link from openness to growth, the main operational implication is that governments should dismantle their barriers to trade. Therefore, it is critical to ask how well the evidence supports the presumption that doing so would raise growth rates.

Note that this question differs from an alternative one we could have asked: *Does international trade raise growth rates of income?* This is a related, but conceptually distinct question. Trade policies do affect the volume of trade, of course. But there is no strong reason to expect their effect on growth to be quantitatively (or even qualitatively) similar to the consequences of changes in trade volumes that arise from, say, reductions in transport costs or increases in world demand. To the extent that trade restrictions represent policy responses to real or perceived market imperfections or, at the other extreme, are mechanisms for rent extraction, they will work differently from natural or geographical barriers to trade and other exogenous determinants. Frankel and Romer (1999) recognize this point in their recent paper on the relationship between trade volumes and income levels. These authors use the geographical component of trade volumes as an instrument to identify the effects of trade on income levels. They appropriately caution that their results cannot be directly applied to the effects of trade policies.

From an operational standpoint, it is clear that the relevant question is the one having to do with the consequences of trade *policies* rather than trade *volumes*. Hence we focus on the recent empirical literature that attempts to measure the effect of trade policies. Our main finding is that this literature is largely uninformative regarding the question we posed above. There is a significant gap between the message that the consum-

ers of this literature have derived and the facts that the literature has actually demonstrated. The gap emerges from a number of factors. In many cases, the indicators of openness used by researchers are problematic as measures of trade barriers or are highly correlated with other sources of poor economic performance. In other cases, the empirical strategies used to ascertain the link between trade policy and growth have serious shortcomings, the removal of which results in significantly weaker findings.

The literature on openness and growth through the late 1980s was usefully surveyed in a paper by Edwards (1993). This survey covered detailed multicountry analyses (such as Little, Scitovsky, and Scott, 1970, and Balassa, 1971) as well as cross-country econometric studies (such as Feder, 1983, Balassa, 1985, and Esfahani, 1991). Most of the cross-national econometric research that was available up to that point focused on the relationship between exports and growth, and not on trade policy and growth. Edwards's evaluation of this literature was largely negative (1993, p. 1389):

[M]uch of the cross-country regression-based studies have been plagued by empirical and conceptual shortcomings. The theoretical frameworks used have been increasingly simplistic, failing to address important questions such as the exact mechanism through which export expansion affects GDP growth, and ignoring potential determinants of growth such as educational attainment. Also, many papers have been characterized by a lack of care in dealing with issues related to endogeneity and measurement errors. All of this has resulted, in many cases, in unconvincing results whose fragility has been exposed by subsequent work.

Edwards argued that such weaknesses had reduced the policy impact of the cross-national econometric research covered in his review.

Our paper picks up where Edwards's survey left off. We focus on a number of empirical papers that either were not included in or have appeared since that survey. Judging by the number of citations in publications by governmental and multilateral institutions and in textbooks, this recent round of empirical research has been considerably more influential in policy and academic circles.² Our detailed analysis covers the

2. We gave examples of citations from international institutions above. Here are some examples from recent textbooks. Yarbrough and Yarbrough (2000, p. 19) write "[o]n the trade-growth connection, the empirical evidence is clear that countries with open markets experience faster growth," citing Edwards (1998). Caves, Frankel, and Jones (1999, pp. 256–257) warn that "[r]esearch testing this proposition is not unanimous" but then continue to say "productivity growth does seem to increase with openness to the international economy and freedom from price and allocative distortions in the domestic econ-

four papers that are probably the best known in the field: Dollar (1992), Sachs and Warner (1995), Ben-David (1993), and Edwards (1998). We also include an analysis of Frankel and Romer (1999), and shorter discussions of Lee (1993), Harrison (1996), and Wacziarg (1998).

A few words about the selection of papers. The paper by Dollar (1992) was not reviewed in Edwards's survey, perhaps because it had only recently been published. We include it here because it is, by our count, the most heavily cited empirical paper on the link between openness and growth. Sachs and Warner (1995) is a close second, and the index of openness constructed therein has now been widely used in the cross-national research on growth.³ The other two papers are also well known, but in these cases our decision was based less on citation counts than on the fact that they are representative of different types of methodologies. Ben-David (1993) considers income convergence in countries that have integrated with each other (such as the European Community countries). Edwards (1998) undertakes a robustness analysis using a wide range of trade-policy indicators, including some subjective indicators. Some of the other recent studies on the relationship between trade policy and growth will be discussed in the penultimate section of the paper.

Our bottom line is that the nature of the relationship between trade policy and economic growth remains very much an open question. The issue is far from having been settled on empirical grounds. We are in fact skeptical that there is a general, unambiguous relationship between trade openness and growth waiting to be discovered. We suspect that the relationship is a contingent one, dependent on a host of country and external characteristics. Research aimed at ascertaining the circumstances under which open trade policies are conducive to growth (as well as those under which they may not be) and at scrutinizing the channels through which trade policies influence economic performance is likely to prove more productive.

omy," citing Sachs and Warner (1995) and Dollar (1992). Husted and Melvin (1997) cite Ben-David (1993) in support of the FPE theorem (p. 111), and Sachs and Warner (1995) in support of the statement that "[o]nly a few countries have followed outward-oriented development strategies for extensive periods of time, but those that have done so have been very successful" (p. 287). Krugman and Obstfeld (1997, 260) write that by the late 1980s "[s]tatistical evidence appeared to suggest that developing countries that followed relatively free trade policies had on average grown more rapidly than those that followed protectionist policies (although this statistical evidence has been challenged by some economists)."

3. From its date of publication, Dollar's paper has been cited at least 92 times, according to the *Social Science Citations Index*. Sachs and Warner (1995) is a close second, with 81 citations. Edwards (1992), Ben-David (1993), and Lee (1993) round off the list, with 57, 38, and 17 citations, respectively.

Finally, it is worthwhile reminding the reader that growth and welfare are not the same thing. Trade policies can have positive effects on welfare without affecting the rate of economic growth. Conversely, even if policies that restrict international trade were to reduce economic growth, it does not follow that they would necessarily reduce the level of welfare. Negative coefficients on policy variables in growth regressions are commonly interpreted as indicating that the policies in question are normatively undesirable. Strictly speaking, such inferences are invalid.⁴ Our paper centers on the relationship between trade policy and growth because this is the issue that has received the most attention in the existing literature. We caution the reader that the welfare implications of empirical results regarding this link (be they positive or negative) must be treated with caution.

The outline of this paper is as follows. We begin with a conceptual overview of the issues relating to openness and growth. We then turn to an in-depth examination of each of the four papers mentioned previously (Dollar, 1992; Sachs and Warner, 1995; Edwards, 1998; and Ben-David 1993), followed by a section on Frankel and Romer (1999). The penultimate section discusses briefly three other papers (Lee, 1993; Harrison, 1996; and Wacziarg 1998). We offer some final thoughts in the concluding section.

2. Conceptual Issues

Think of a small economy that takes world prices of tradable goods as given. What is the relationship between trade restrictions and real GDP in such an economy? The modern theory of trade policy as it applies to such a country can be summarized in the following three propositions:

1. In static models with no market imperfections and other pre-existing distortions, the effect of a trade restriction is to reduce the level of real GDP at world prices. In the presence of market failures such as externalities, trade restrictions *may* increase real GDP (although they are hardly ever the first-best means of doing so).
2. In standard models with exogenous technological change and diminishing returns to reproducible factors of production (e.g., the neo-
4. Some of the main problems with economic growth as a measure of welfare are that: (1) the empirically identifiable effect of policies on rates of growth—especially over short intervals—could be different from their effect on levels of income; (2) levels of per capita income may not be good indicators of welfare because they do not capture the distribution of income or the level of access to primary goods and basic capabilities; and (3) high growth rates could be associated with suboptimally low levels of present consumption.

classical model of growth), a trade restriction has no effect on the long-run (steady-state) rate of growth of output.⁵ This is true regardless of the existence of market imperfections. However, there may be growth effects during the transition to the steady state. (These transitional effects may be positive or negative, depending on how the long-run level of output is affected by the trade restriction.)

3. In models of endogenous growth generated by nondiminishing returns to reproducible factors of production or by learning-by-doing and other forms of endogenous technological change, the presumption is that lower trade restrictions boost output growth in the world economy as a whole. But a subset of countries may experience diminished growth, depending on their initial factor endowments and levels of technological development.

Taken together, these points imply that there should be no theoretical presumption in favor of finding an unambiguous, negative relationship between trade barriers and growth rates in the types of cross-national data sets typically analyzed.⁶ The main complications are twofold. First, in the presence of certain market failures, such as positive production externalities in import-competing sectors, the long-run levels of GDP (measured at world prices) can be higher with trade restrictions than without. In such cases, data sets covering relatively short time spans will reveal a positive (partial) association between trade restrictions and the growth of output along the path of convergence to the new steady state. Second, under conditions of endogenous growth, trade restrictions may also be associated with higher growth rates of output whenever the restrictions promote technologically more dynamic sectors over others. In dynamic models, moreover, an increase in the growth rate of output is neither a necessary nor a sufficient condition for an improvement in *welfare*.

Since endogenous-growth models are often thought to have provided the missing theoretical link between trade openness and long-run growth, it is useful to spend a moment on why such models in fact provide an ambiguous answer. As emphasized by Grossman and Helpman (1991), the general answer to the question “Does trade promote

5. Strictly speaking, this statement is true only when the marginal product of the reproducible factors (“capital”) tends to zero in the limit. If this marginal product is bounded below by a sufficiently large positive constant, trade policies can have an effect on long-run growth rates, similar to their effect in the more recent endogenous growth models (point 3 below). See the discussion in Srinivasan (1997).

6. See Buffie (1998) for an extensive theoretical discussion of the issues from the perspective of developing countries.

innovation in a small open economy?" is "It depends."⁷ In particular, the answer depends on whether the forces of comparative advantage push the economy's resources in the direction of activities that generate long-run growth (via externalities in research and development, expanding product variety, upgrading product quality, and so on) or divert them from such activities. Grossman and Helpman (1991), Feenstra (1990), Matsuyama (1992), and others have worked out examples where a country that is behind in technological development can be driven by trade to specialize in traditional goods and experience a reduction in its long-run rate of growth. Such models are in fact formalizations of some very old arguments about infant industries and about the need for temporary protection to catch up with more advanced countries.

The issues can be clarified with the help of a simple model of a small open economy with learning-by-doing. The model is a simplified version of that in Matsuyama (1992), except that we analyze the growth implications of varying the import tariff, rather than simply comparing free trade with autarky. The economy is assumed to have two sectors, agriculture (a) and manufacturing (m), with the latter subject to learning-by-doing that is external to individual firms in the sector but internal to manufacturing as a whole. Let labor be the only mobile factor between the two sectors, and normalize the economy's labor endowment to unity. We can then write the production functions of the manufacturing and agricultural sectors, respectively, as

$$\begin{aligned} X_t^m &= M_t n_t^\alpha, \\ X_t^a &= A(1 - n_t)^\alpha, \end{aligned}$$

where n_t stands for the labor force in manufacturing, α is the share of labor in value added in the two sectors (assumed to be identical for simplicity), and t is a time subscript. The productivity coefficient in manufacturing, M_t , is a state variable evolving according to

$$\dot{M}_t = \delta X_t^m,$$

where an overdot represents a time derivative and δ captures the strength of the learning effect.

We assume the economy has an initial comparative disadvantage in manufacturing, and normalize the relative price of manufactures on world markets to unity. If the ad valorem import tariff on manufactures is τ , the domestic relative price of manufactured goods becomes $1 + \tau$.

7. This is a slight paraphrase of Grossman and Helpman (1991, p. 152).

Instantaneous equilibrium in the labor market requires the equality of value marginal products of labor in the two sectors:

$$A(1 - n_t)^{\alpha-1} = (1 + \tau)M_t n_t^{\alpha-1}$$

It can be checked that an increase in the import tariff has the effect of allocating more of the economy's labor to the manufacturing sector:

$$\frac{dn_t}{d\tau} > 0.$$

Further, for a constant level of τ , n_t evolves according to

$$\hat{n}_t = \frac{\delta}{1 - \alpha} (1 - n_t) n_t^\alpha,$$

where $\hat{\cdot}$ denotes proportional changes.

Let Y_t denote the value of output in the economy evaluated at world prices:

$$Y_t = M_t n_t^\alpha + A(1 - n_t)^\alpha.$$

Then the instantaneous rate of growth of output at world prices can be expressed as follows:

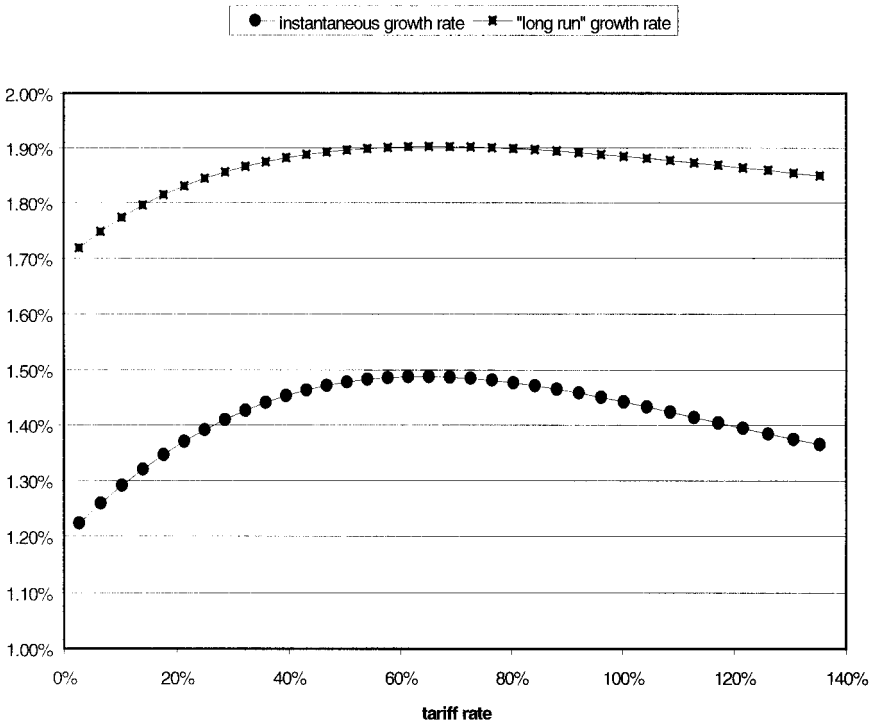
$$\hat{Y}_t = \delta \left(\lambda_t + \frac{\alpha}{1 - \alpha} (\lambda_t - n_t) \right) n_t^\alpha,$$

where λ_t is the share of manufacturing output in total output when both are expressed at world prices (i.e., $\lambda_t = X_t^m/Y_t$).

Consider first the case when $\tau = 0$. In this case, it can be checked that $\lambda_t = n_t$ and the expression for the instantaneous growth rate of output simplifies to $\hat{Y}_t = \delta \lambda_t n_t^\alpha$, which is strictly positive whenever $n_t > 0$. Growth arises from the dynamic effects of learning, and is faster the larger the manufacturing base n_t . A small tariff would have a positive effect on growth on account of this channel because it would enlarge the manufacturing sector (raise n_t).

When $\tau > 0$, the manufacturing share of output at world prices is less than the labor share in manufacturing, and $\lambda_t < n_t$. Now the second term in the expression for \hat{Y}_t is negative. The intuition is as follows. The tariff imposes a production-side distortion in the allocation of the economy's resources. For any given gap between λ_t and n_t , the productive efficiency cost of this distortion rises as manufacturing output (the base of the distortion) gets larger.

Figure 2 GROWTH RATES OF GDP AT WORLD PRICES



Hence the tariff exerts two contradictory effects on growth. By pulling resources into the manufacturing sector, it enlarges the scope for dynamic scale benefits, thereby increasing growth. But it also imposes a static efficiency loss, the cost of which rises over time as the manufacturing sector becomes larger.⁸ Figure 2 shows the relationship between the tariff and the rate of growth of output (at world prices) for a particular parameterization of this model. Two curves are shown, one for the instantaneous rate of growth (based on the expression above), and the other for the average growth rate over a twenty-year horizon [calculated as $\frac{1}{20}(\ln Y_{20} - \ln Y_0)$]. In both cases, growth increases in τ until a critical level, and then diminishes in τ . This pattern is, however, by no means

8. We emphasize once again that these results on the growth of *output* do not translate directly into welfare consequences. In this particular model, the *level* effect of a tariff distortion also has to be taken into account before a judgment on welfare can be passed. Hence it is possible for welfare to be reduced (raised) even though the growth rate of output is permanently higher (lower).

general, and other types of results can be obtained under different parameterizations.

The model clarifies a number of issues. First, it shows that it is relatively straightforward to write a well-specified model that generates the conclusions that many opponents of trade openness have espoused—namely, that free trade can be detrimental to some countries' economic prospects, especially when these countries are lagging in technological development and have an initial comparative advantage in “nondynamic” sectors. More broadly, the model illustrates that there is no determinate theoretical link between trade protection and growth once real-world phenomena such as learning, technological change, and market imperfections (here captured by a learning-by-doing externality) are taken into account. Third, it highlights the exact sense in which trade restrictions distort market outcomes. A trade barrier has resource-allocation effects because it alters a *domestic* price ratio: it raises the domestic price of import-competing activities relative to the domestic price of exportables, and hence introduces a wedge between the domestic relative-price ratio and the opportunity costs reflected in relative border prices.⁹ While this point is obvious, it bears repeating, as some of the empirical work reviewed below interprets openness in a very different manner.

3. *David Dollar (1992)*

As mentioned previously, the paper by Dollar (1992) is one of the most heavily cited studies on the relationship between openness and growth. The principal contribution of Dollar's paper lies in the construction of two separate indices, which Dollar demonstrates are each negatively correlated with growth over the 1976–1985 period in a sample of 95 developing countries. The two indices are an “index of real exchange-rate distortion” and an “index of real exchange-rate variability” (henceforth DISTORTION and VARIABILITY). These indices relate to “outward orientation,” as understood by Dollar (1992, p. 524), in the following way:

Outward orientation generally means a combination of two factors: first, the level of protection, especially for inputs into the production process, is

9. Some authors have stressed the effects that the high levels of discretion associated with trade policies can have on rent seeking and thus on economic performance (Krueger, 1974; Bhagwati, 1982). These effects go beyond the direct impact on resource allocation that we discuss. They are however related more directly to the discretionary nature of policies than to their effect on the economy's openness. Discretionary export promotion policies—which will make an economy more open—should in principle be just as conducive to rent seeking as protectionist policies.

relatively low (resulting in a sustainable level of the real exchange rate that is favorable to exporters); and second, there is relatively little variability in the real exchange rate, so that incentives are consistent over time.

We shall argue that DISTORTION has serious conceptual flaws as a measure of trade restrictions, and is in any case not a robust correlate of growth, while VARIABILITY, which appears to be robust, is a measure of instability more than anything else.

In order to implement his approach, Dollar uses data from Summers and Heston (1988, Mark 4.0) on comparative price levels. Their work compares prices of an identical basket of consumption goods across countries. Hence, letting the United States be the benchmark country, these data provide estimates of each country i 's price level (RPL_i) relative to the United States: $RPL_i = 100 \times P_i / (e_i P_{US})$, where P_i and P_{US} are the respective consumption price indices, and e_i is the nominal exchange rate of country i against the U.S. dollar (in units of home currency per dollar).¹⁰ Since Dollar is interested in the prices of tradable goods only, he attempts to purge the effect of systematic differences arising from the presence of nontradables. To do this, he regresses RPL_i on the level and square of GDP per capita and on regional dummies for Latin America and Africa, as well as year dummies. Let the predicted value from this regression be denoted \widehat{RPL}_i . Dollar's index DISTORTION is RPL_i / \widehat{RPL}_i , averaged over the ten-year period 1976–1985. VARIABILITY is in turn calculated by taking the coefficient of variation of the annual observations of RPL_i / \widehat{RPL}_i for each country over the same period.

Dollar interprets the variation in the values of DISTORTION across countries as capturing cross-national differences in the restrictiveness of trade policy. He states: "the index derived here measures the extent to which the real exchange rate is distorted away from its free-trade level by the trade regime" and "a country sustaining a high price level over many years would clearly have to be a country with a relatively large amount of protection" (Dollar 1992, p. 524). Since this type of claim is often made in other work as well,¹¹ we shall spend some time on it before reviewing Dollar's empirical results. We will show that a comparison of price indices for tradables is informative about levels of trade protection only under very restrictive conditions that are unlikely to hold in practice.

10. Our notation differs from Dollar's (1992). In particular, the exchange rate is defined differently.

11. For example, in Bhalla and Lau (1992), whose index is also used in Harrison (1996). We will discuss Harrison's paper in the penultimate section.

3.1 TRADE POLICIES AND PRICE LEVELS

We will not discuss further Dollar's method for purging the component of nontradable-goods prices that is systematically related to income and other characteristics.¹² Assuming the method is successful, the DISTORTION measure approximates (up to a random error term) the price of a country's tradables relative to the United States. Letting P^T stand for the price index for tradables and neglecting the error, the DISTORTION index for country i can then be expressed as $P_i^T/(e_i P_{US}^T)$.

Let us, without loss of generality, fix the price level of tradables in the United States, P_{US}^T , and assume that free trade prevails there. The question is under what conditions trade restrictions will be associated with higher levels of $P_i^T/(e_i P_{US}^T)$. Obviously, the answer depends on the effect of the restrictions on P_i^T (and possibly on e_i).

Note that P_i^T is an aggregate price index derived from the domestic prices of two types of tradables: import-competing goods and exportables. Hence P_i^T can be expressed as a linearly homogenous function of the form

$$P_i^T = \pi(p_i^m, p_i^x),$$

where p_i^m and p_i^x are the domestic prices of import-competing goods and exportables, respectively. Since Summers–Heston price levels are estimated for an identical basket of goods, the price-index function $\pi(\cdot)$ applies equally to the United States:

$$P_{US}^T = \pi(p_{US}^m, p_{US}^x).$$

Next, define t_i^m and t_i^x as the ad valorem equivalent of import restrictions and export restrictions, respectively. Assume that the law of one price holds (we shall relax this below). Then, $p_i^m = e_i p_{US}^m (1 + t_i^m)$ and $p_i^x = e_i p_{US}^x / (1 + t_i^x)$. Consequently, the domestic price of tradables relative to U.S. prices can be expressed as

$$\frac{P_i^T}{e_i P_{US}^T} = \frac{\pi(P_{US}^m (1 + t_i^m), p_{US}^x / (1 + t_i^x))}{\pi(p_{US}^m, p_{US}^x)} = \frac{(1 + t_i^m) \pi\left(p_{US}^m, \frac{p_{US}^x}{(1 + t_i^x)(1 + t_i^m)}\right)}{\pi(p_{US}^m, p_{US}^x)},$$

12. For a good recent discussion of the problems that may arise on this account see Falvey and Gemell (1999).

where we have made use of the linear homogeneity of $\pi(\cdot)$. Note that the nominal exchange rate has dropped out thanks to the assumption of the law of one price.

Consider first the case where there are binding import restrictions, but no export restrictions ($t_i^m > 0$ and $t_i^x = 0$). In this instance, it is apparent that $P_i^T > e_i P_{US}^T$, and trade restrictions do indeed raise the domestic price of tradables (relative to the benchmark country). Judging from the quotations above, this is the case that Dollar seems to have in mind.

On the other hand, consider what happens when the country in question rescinds all import restrictions and imposes instead export restrictions at an ad valorem level that equals that of the import restrictions just lifted ($t_i^m = 0$ and $t_i^x > 0$). From the Lerner (1936) symmetry theorem, it is evident that the switch from import protection to export taxation has no resource-allocation and distributional effects for the economy whatsoever. The relative price between tradables, p_i^m/p_i^x , remains unaffected by the switch. Yet, because export restrictions reduce the domestic price of exportables *relative to world prices*, it is now the case that $P_i^T < e_i P_{US}^T$. The country will now appear, by Dollar's measure, to be outward-oriented.

One practical implication is that economies that combine import barriers with export taxes (such as many countries in sub-Saharan Africa) will be judged less protected than those that rely on import restrictions alone. Conversely, countries that dilute the protective effect of import restrictions by using export subsidies ($t_i^x < 0$) will appear more protected than countries that do not do so.

Hence the DISTORTION index is sensitive to the form in which trade restrictions are applied. This follows from the fact that trade policies work by altering relative price *within* an economy; they do not have unambiguous implications for the level of prices in a country relative to another. A necessary condition for Dollar's index to do a good job of ranking trade regimes according to restrictiveness is that export policies (whether they tax or promote exports) play a comparatively minor role. Moreover, as we show in the next section, this is not a sufficient condition.

3.2 HOW RELEVANT IS THE LAW OF ONE PRICE IN PRACTICE?

The discussion above was framed in terms that are the most favorable to Dollar's measure, in that we assumed the law of one price (LOP) holds. Under this maintained hypothesis, the prices of tradable goods produced in different countries can diverge from each other, when expressed in a common currency, only when there exist trade restrictions (or transport costs).

However, there is a vast array of evidence suggesting that LOP does not accurately describe the world we live in. In a recent review article,

Rogoff (1996, p. 648) writes of the “startling empirical failure of the law of one price.” Rogoff concludes: “commodities where the deviations from the law of one price damp out very quickly are the exception rather than the rule” (Rogoff, 1996, p. 650). Further, the evidence suggests that deviations from LOP are systematically related to movements in nominal exchange rates (see references in Rogoff, 1996). Indeed, it is well known that (nominal) exchange-rate policies in many developing countries are responsible for producing large and sustained swings in *real* exchange rates. Trade barriers or transport costs typically play a much smaller role.

Dollar (1992, p. 525) acknowledges that “there might be short-term fluctuations [unrelated to trade barriers] if purchasing-power parity did not hold continuously,” but considers that these fluctuations would average out over time. Rogoff (1996, p. 647) concludes in his survey that the speed of convergence to purchasing-power parity (PPP) is extremely slow, of the order of roughly 15% per year. At this speed of convergence, averages constructed over a time horizon of 10 years (the horizon used in Dollar’s paper) would exhibit substantial divergence from PPP in the presence of nominal shocks.

Under this interpretation, a significant portion of the cross-national variation in price levels exhibited in DISTORTION would be due not to trade policies, but to monetary and exchange-rate policies. Unlike trade policies, nominal exchange-rate movements have an unambiguous effect on the domestic price level of traded goods relative to foreign prices when LOP fails: an appreciation raises the price of both import-competing and exportable goods relative to foreign prices, and a depreciation has the reverse effect. Countries where the nominal exchange rate was not allowed to depreciate in line with domestic inflation would exhibit an appreciation of the real exchange rate (a rise in domestic prices relative to foreign levels), and correspondingly would be rated high on the DISTORTION index. Countries with aggressive policies of devaluation (or low inflation relative to the trend depreciation of their nominal exchange rate) would receive low DISTORTION ratings.

Transport costs provide another reason why DISTORTION may be unrelated to trade policies, especially in a large cross-section of countries. Dollar’s index would be influenced by geographic variables such as access to sea routes and distance to world markets, even when LOP—appropriately modified to allow for transport costs—holds. Hence in practice DISTORTION is likely to capture the effects of geography as well as of exchange-rate policies. Indeed, when we regress Dollar’s DISTORTION index on the black-market premium (a measure of exchange-rate policy), a set of continent dummies, and two trade-related geo-

Table 1 EFFECT OF GEOGRAPHICAL AND EXCHANGE-RATE POLICY VARIABLES ON DOLLAR'S DISTORTION INDEX

<i>Variable</i>	(1)	(2)
bmpav	0.07*** (1.971)	0.083** (2.47)
rcoast	-0.045* (-3.321)	-0.053* (-3.032)
tropics	0.209*** (1.829)	0.145 (1.004)
Latin America	0.012 (0.097)	-0.037 (-0.257)
SSA	0.451* (3.319)	0.46** (2.43)
East Asia	-0.12 (-0.921)	-0.145 (-0.889)
TAR		-0.017 (-0.08)
NTB		-0.276*** (-1.851)
R ²	0.52	0.58
N	89	71

Heteroskedasticity-corrected *t*-statistics in parentheses. See appendix for variable definitions. Regressions include a constant term and cover only developing countries. Levels of statistical significance indicated by asterisks: * 99%, ** 95%, *** 90%.

graphic variables (the coastal length over total land area and a dummy for tropical countries), we find that these explain more than 50% of the variation in Dollar's distortion index. Furthermore, two trade-policy variables (tariffs and quotas) enter with the wrong sign (Table 1)!

To summarize, DISTORTION is theoretically appropriate as a measure of trade restrictions when three conditions hold: (1) there are no export taxes or subsidies in use, (2) LOP holds continuously, and (3) there are no systematic differences in national price levels due to transport costs and other geographic factors. Obviously, all of these requirements are counterfactual. Whether one believes that DISTORTION still provides useful empirical information on trade regimes depends on one's priors

regarding the practical significance of the three limitations expressed above.¹³ Our view is that the second and third of these—the departure from LOP and the effect of geography—are particularly important in practice. We regard it as likely that it is the variation in nominal exchange-rate policies and geography, and not the variation in trade restrictions, that drives the cross-sectional variation of DISTORTION.

3.3 WHY VARIABILITY?

As mentioned previously, Dollar (1992) uses his measure of distortion in conjunction with a measure of variability, the latter being the coefficient of variation of DISTORTION measured on an annual basis. He is driven to do this because the country rankings using DISTORTION produce some “anomalies.” For example, “Korea and Taiwan have the highest distortion measures of the Asian developing economies” and “the rankings within the developed country groups are not very plausible” (Dollar, 1992, pp. 530–531). The ten least-distorted countries by this measure include not only Hong Kong, Thailand, and Malta, but also Sri Lanka, Bangladesh, Mexico, South Africa, Nepal, Pakistan, and Syria! Burma’s rating (90) equals that of the United States. Taiwan (116) is judged more distorted than Argentina (113). Our discussion above indicated that DISTORTION is highly sensitive to the form in which trade policies are applied and to exchange-rate policies as well as omitted geographic characteristics. So such results are not entirely surprising.

Dollar states that the “number of anomalies declines substantially if the real exchange rate distortion measure is combined with real exchange rate variability to produce an outward orientation index” (Dollar, 1992, p. 531). He thus produces a country ranking based on a weighted average of the DISTORTION and VARIABILITY indices. Since these two indices are entered separately in his growth regressions, we shall not discuss this combined index of “outward orientation” further.

However, we do wish to emphasize the obvious point that the VARIABILITY index has little to do with trade restrictions, as commonly understood, or with inward or outward orientation per se. What does VARIABILITY really measure? The ten countries with the highest VARIABILITY scores are Iraq, Uganda, Bolivia, El Salvador, Nicaragua, Guyana, Somalia, Nigeria, Ghana, and Guatemala. For the most part, these are countries that have experienced very high inflation rates and/or se-

13. The sensitivity of Dollar’s index to these assumptions highlights a generic difficulty with regression-based indices which use the residual from a regression to proxy for an excluded variable: such indices capture variations in the excluded variable accurately only as long as the model is correctly and fully specified. If some variables are excluded from the estimated equation, they will form part of the index.

vere political disturbances during 1976–1985. It is plausible that VARIABILITY measures economic instability at large. In any case, it is unclear to us why we should think of it as an indicator of trade orientation.

3.4 EMPIRICAL RESULTS

The first column of Table 2 shows our replication of the core Dollar (1992) result for 95 developing countries. Dollar's benchmark specification includes on the right-hand side the investment rate (as a share of GDP, averaged over 1976–1985) in addition to DISTORTION and VARIABILITY. As shown in column (1), DISTORTION and VARIABILITY both enter with negative and highly significant coefficients using this specification. [Our results are virtually identical to those in Dollar (1992), with the

Table 2 REPLICATION AND EXTENSION OF DOLLAR'S (1992) RESULTS

	(1)	(2)	(3)	(4)	(5)
DISTORTION	-0.018* (-3.128)	-0.008 (-1.009)	-0.003 (-0.406)	-0.004 (-0.514)	-0.008 (-0.899)
VARIABILITY	-0.080* (-2.64)	-0.080** (-2.084)	-0.103* (-3.3)	-0.107* (-3.51)	-0.099* (-3.212)
Investment/GDP	0.137* (3.515)	0.100** (2.278)			
Latin America		-0.015** (-2.34)	-0.016* (-2.65)	-0.014** (-2.362)	-0.019* (-3.337)
East Asia		0.007 (0.747)	0.010 (0.937)	0.011 (0.976)	0.004 (0.382)
SSA		-0.018** (-2.419)	-0.026* (-3.824)	-0.029* (-4.129)	-0.028* (-3.411)
Log initial income				-0.004 (-1.097)	-0.011** (-2.182)
Schooling, 1975					0.005** (2.531)
N	95	95	95	95	80
R ²	0.38	0.45	0.40	0.41	0.49

Dependent variable: growth of real GDP per capita, 1976–1985. Heteroskedasticity-corrected *t* statistics in parentheses. Regressions include a constant term and cover only developing countries. Levels of statistical significance indicated by asterisks: * 99%; ** 95%; *** 90%.

difference that our t -statistics are based on heteroskedasticity-corrected standard errors.]

None of Dollar's runs include standard regressors such as initial income, education, and regional dummies. The other columns of Table 2 show the results as we alter Dollar's specification to make it more compatible with recent cross-national work on growth (e.g., Barro, 1997). First, we add regional dummies for Latin America, East Asia, and sub-Saharan Africa to ensure that the results are not due to omitted factors correlated with geographical location (column 2). Next we drop the investment rate (column 3), and add in succession initial income (column 4) and initial schooling (column 5).¹⁴ The dummies for Latin America and sub-Saharan Africa are negative and statistically significant. Initial income and education also enter significantly, with the expected signs (negative and positive, respectively).

We find that the VARIABILITY index is robust to these changes, but that DISTORTION is not. In fact, as soon as we introduce regional dummies in the regression, the estimated coefficient on DISTORTION comes down sizably and becomes insignificant. Whatever DISTORTION may be measuring, this raises the possibility that the results with this index are spurious, arising from the index's correlation with (omitted) regional effects.

Dollar's original results were based on data from Mark 4.0 of the Summers–Heston database (Summers and Heston, 1988). We have recalculated Dollar's DISTORTION and VARIABILITY indices using the more recent version (Mark 5.6) of the Summers–Heston data, confining ourselves to the same period examined by Dollar (1976–1985). The revised data allow us to generate these indices for 112 developing countries. We have also rerun the regressions for cross sections over different periods, as well as in panel form with fixed effects. We do not report these results here, for reasons of space (see the working-paper version of this paper, Rodríguez and Rodrik, 1999). The bottom line that emerges is similar to the conclusion just stated: the estimated coefficient on VARIABILITY is generally robust to alterations in specifications; the coefficient on DISTORTION is not.

4. *Jeffrey Sachs and Andrew Warner (1995)*

We turn next to the paper "Economic reform and the process of global integration" by Jeffrey Sachs and Andrew Warner (1995). This extremely

14. The income variable comes from the Summers–Heston (Mark 4.0) data set used in Dollar (1992). Schooling is from Barro and Lee (1994).

influential paper¹⁵ is an ambitious attempt to solve the measurement-error problem in the literature by constructing an index of openness that combines information about several aspects of trade policy. The Sachs–Warner (SW) openness indicator (OPEN) is a zero–one dummy, which takes the value 0 if the economy was closed according to *any* one of the following criteria:

1. it had average tariff rates higher than 40% (TAR);
2. its nontariff barriers covered on average more than 40% of imports (NTB);
3. it had a socialist economic system (SOC);
4. it had a state monopoly of major exports (MON);
5. its black-market premium exceeded 20% during either the decade of the 1970s or the decade of the 1980s (BMP).¹⁶

The rationale for combining these indicators into a single dichotomous variable is that they represent different ways in which policymakers can close their economy to international trade. Tariffs set at 50% have exactly the same resource-allocation implications as quotas at a level that raised domestic market prices for importables by 50%. To gauge the effect of openness on growth, it is necessary to use a variable that classifies as closed those countries that were able to effectively restrict their economies' integration into world markets through the use of different combinations of policies that would achieve that result. Furthermore, if these openness indicators are correlated among themselves, introducing them separately in a regression may not yield reliable estimates, due to their possibly high level of collinearity.

The SW dummy has a high and robust coefficient when inserted in growth regressions. The point estimate of its effect on growth (in the original benchmark specification) is 2.44 percentage points¹⁷: economies that pass all five requirements experience on average economic growth two and a half percentage points higher than those that do not. The *t*-statistic is 5.50 (5.83 if estimated using robust standard errors). This coefficient appears to be highly robust to changes in the list of controls: in a recent paper which subjects 58 potential determinants of growth to

15. A partial listing of papers that have made use of the Sachs–Warner index includes Hall and Jones (1998), Wacziarg (1998), Sala-i-Martin (1997), Burnside and Dollar (1997), and Collins and Bosworth (1996).

16. Sachs and Warner use data from the following sources: Lee (1993) for nontariff barriers, Barro and Lee (1993) for tariffs, World Bank (1994) for state monopoly of exports, Kornai (1992) for the classification of socialist and nonsocial countries, and International Currency Analysis (various years) for black-market premia.

17. In the long run, such an economy would converge to a level of per capita GDP 2.97 times as high as if it had remained closed.

an exhaustive sensitivity analysis, the average p -value for the SW index is less than 0.1%.¹⁸

In this section we ask several questions about Sachs and Warner's results. First, we ask which, if any, of the individual components of the index are responsible for the strength of the SW dummy. We find that the SW dummy's strength derives mainly from the combination of the black-market premium (BMP) and the state-monopoly-of-exports (MON) variables. Very little of the dummy's statistical power would be lost if it were constructed using only these two indicators. In particular, there is little action in the two variables that are the most direct measures of trade policy: tariff and nontariff barriers (TAR and NTB).

We then ask to what extent the black-market premium and state-monopoly variables are measures of trade policy. We suggest that their significance in explaining growth can be traced to their correlation with other determinants of growth: macroeconomic problems in the case of the black-market premium, and location in sub-Saharan Africa in the case of the state-monopoly variable. We conclude that the SW indicator serves as a proxy for a wide range of policy and institutional differences, and that it yields an upward-biased estimate of the effects of trade restrictions proper.

4.1 WHICH INDIVIDUAL VARIABLES ACCOUNT FOR THE SIGNIFICANCE OF THE SW DUMMY?

We start by contrasting Sachs and Warner's result with the results of controlling separately for individual components of their index. Column 1 of Table 3 reproduces their baseline regression, and column 2 shows what happens when each of the components of the SW index is inserted separately into the same specification.¹⁹ The variables BMP and MON are highly significant, whereas the rest are not. An F -test for the joint significance of the other three components (SOC, TAR, and NTB) yields a p -value of 0.25.

18. Sala-i-Martin (1997). The variable used by Sala-i-Martin is the number of years an economy was open according to the SW criteria, whereas here we follow Sachs and Warner's (1995) original article and use a dummy which captures whether or not the economy was open during 1970–1989.

19. We use the same set of controls used by Sachs and Warner. These are log of GDP in 1970, secondary schooling in 1970, primary schooling in 1970, government consumption as a percentage of GDP, number of revolutions and coups per year, number of assassinations per million population, relative price of investment goods, and ratio of investment to GDP. However, our results are highly robust to changes in the list of controls. For example, the simple correlations of TAR, NTB, and SOC with growth are, respectively, $-.048$, $-.083$ and $-.148$. Our result is also not due to multicollinearity: the R^2 's from regressions of any one of SOC, NTB, and TAR on the other two are, respectively, 0.02, 0.05, and 0.05.

Table 3 EFFECT OF DIFFERENT OPENNESS INDICATORS ON GROWTH

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
OPEN	2.44* (5.83)						
BMP		-1.701* (-3.65)					
MON		-2.020* (-2.84)					
SOC		-1.272 (-1.39)					
NTB		-0.453 (-0.81)					
TAR		-0.134 (-0.18)					
BM			2.086* (4.82)		2.119* (5.09)	2.519* (5.94)	2.063* (4.64)
SQT				0.877*** (1.82)	0.735 (1.59)	0.663 (1.30)	
SOC							.389 (.56)
QT							.657 (1.28)
R ²	0.593	0.637	0.522	0.455	0.617	0.522	0.619
N	79	71	78	75	74	74	74

Dependent variable: growth of GDP per capita, 1970–1989. All equations except that for column 6 include the following controls: log of GDP in 1970, investment rate in 1970, government consumption/GDP, assassinations per capita, deviation from world investment prices, secondary-schooling ratio, primary-schooling ratio, revolutions and coups, and a constant term. Column 6 drops the investment rate and deviation from world investment prices. Numbers in parentheses are *t*-statistics based on Huber–White heteroskedasticity-consistent standard errors.

To check whether it is mainly the combination of BMP and MON that drives the Sachs and Warner's result, we ask the following question: suppose that we had built a dummy variable, in the spirit of Sachs and Warner, which classified an economy as closed only if it was closed according to BMP and MON. That is, suppose we ignored the information the other three variables give us as to the economy's openness.

How significant would the coefficient of our variable be in a growth regression? How different would the partition between open and closed economies that it generates be from that generated by the SW dummy? Suppose alternatively that we also constructed an openness dummy based only on the information contained in SOC, NTB, and TAR. How significant would that variable be in a growth regression? And how correlated would it be with the SW index?

Columns (3)–(6) of Table 3 address the question of significance. We denote by BM a variable that takes the value 1 when the economy is open according to criteria 4 and 5 above, whereas SQT equals 1 when the economy passes criteria 1, 2, and 3. We substitute these variables for the SW openness index in the regression Sachs and Warner present in their paper. Entered on its own, BM is highly significant, with an estimated coefficient that is very close to that on OPEN (2.09 vs. 2.44; see column 3). When SQT is substituted for BM, the estimated coefficient on SQT is much smaller (0.88) and significant only at the 90% level (column 4). We next enter BM and SQT simultaneously: the coefficient of SQT now has a *t*-statistic of 1.59, whereas the coefficient on BM retains a *t*-statistic of 5.09 and a point estimate (2.12) close to that on the openness variable in the original equation (column 5). Once the investment rate and investment prices, which are likely to be endogenous, are taken out of the equation, the *t*-statistic on SQT drops to 1.30 and that on BM rises to 5.94 (column 6).

The comparability of the results in Table 3 is hampered by the fact that the sample size changes as we move from one column to the next. This is because not all of the 79 countries in the sample have data for each of the individual SW components. To check whether this introduces any difficulties for our interpretation, we have also run these regressions holding the sample size fixed. We restricted the sample to those countries which have the requisite data for all the components, using both the original specification ($n = 71$) and a specification where we drop two of the SW regressors with *t*-statistics below unity (primary schooling, and revolutions and coups) to gain additional observations ($n = 74$). In both cases, our results were similar to those reported above: Regardless of whether BM and SQT are entered separately or jointly, the coefficient on BM is highly significant (with a point estimate that is statistically indistinguishable from that on OPEN) while the coefficient on SQT is insignificant.²⁰

Hence, once BM is included, there is little additional predictive power

20. The largest *t*-statistic we obtained for SQT in these runs is 1.4. These results are not shown, to save space, but are available on request.

coming from regime type (socialist or not), level of tariffs, or coverage of nontariff barriers.²¹ The strength of the SW index derives from the low growth performance of countries with either high black-market premia or state export monopolies (as classified by Sachs and Warner).²²

The reason why BM performs so much better than SQT is that BM generates a partition between closed and open economies that is much closer to that generated by OPEN than the partition generated by SQT. Only six economies are classified differently by BM and by OPEN, while OPEN and SQT disagree in 31 cases. The disagreement between OPEN and SQT is concentrated in 15 African and 12 Latin American economies which SQT fails to qualify as closed but BM (and therefore OPEN) does: the African economies are found to be closed because of their state monopolies of exports, and those of Latin America because of their high black-market premia. The average rate of growth of these economies is 0.24, much lower than the sample average of 1.44.²³

In view of the overwhelming contribution of the black-market premium and the dummy for state monopoly of exports to the statistical performance of the SW openness index, it is logical to ask what exactly it is that these two variables are capturing. To what extent are they indicators of trade policy? Could they be correlated with other variables that have a detrimental effect on growth, therefore not giving us much useful information on trade openness per se? We turn now to these questions, first with an analysis of the state-monopoly-of-exports variable, and then with a discussion of the black-market premium variable.

21. A different form in which the "horse race" can be run, suggested to us by Jeffrey Sachs, is to introduce OPEN and BM together in the regression, to see if OPEN clearly "wins." When we do this, we find that the point estimate of the coefficient on OPEN is generally larger than that on BM, but that the two coefficients are statistically indistinguishable from each other, because OPEN and BM are highly collinear with each other (as we discuss further below). On the other hand, when OPEN and SQT are entered together, SQT has the wrong (negative) sign and the equality of coefficients can easily be rejected.
22. Harrison and Hanson (1999) have studied the SW dummy and reach a similar conclusion, namely that the effect of trade-policy indicators (tariffs and quotas) on the strength of the SW dummy is small and not significant. The key difference between our work and Harrison and Hanson's is that they introduce the subcomponents of the SW index separately in their regression whereas we construct the subindexes described in the text.
23. Our result is not due to an arbitrary distinction between BM and SQT. SQT performs more poorly than any other openness index constructed on the basis of three of the five indicators used by Sachs and Warner, and BM performs more strongly than any index constructed with two of these five indicators. A similar result applies to partitions along other dimensions: those constructed using four indicators which exclude either BMP or MON do more poorly than any of those which include them; and either BMP or MON individually does better than any of the other indicators. Details of these exercises can be found in the working-paper version of our paper (Rodríguez and Rodrik, 1999).

4.2 WHAT DOES THE STATE-MONOPOLY-OF-EXPORTS VARIABLE REPRESENT?

Sachs and Warner's rationale for using an indicator of the existence of a state monopoly on major exports is the well-known equivalence between import and export taxes (Lerner, 1936). The variable MON is meant to capture cases in which governments taxed major exports and therefore reduced the level of trade (exports *and* imports). Sachs and Warner use an index of the degree of distortion caused by export marketing boards, taken from the World Bank study *Adjustment in Africa: Reforms, Results, and the Road Ahead* (World Bank, 1994).²⁴

We note that the World Bank study covers only 29 African economies that were under structural adjustment programs from 1987 to 1991. This results in a double selection bias. First, non-African economies with restrictive policies towards exports automatically escape scrutiny. Second, African economies with restrictive export policies but not undergoing adjustment programs in the late 1980s are also overlooked. Since Africa was the slowest-growing region during the period covered, and since economies that need to carry out structural adjustment programs are likely to be doing worse than those that do not, the effect is to bias the coefficient on openness upwards on both accounts.

How this selection bias affects the country classification can be illustrated by two examples: Indonesia and Mauritius. Both of these economies are rated as open in Sachs and Warner's sample. Both are excluded from the sample used to construct the state-monopoly-on-exports variable: Indonesia because it is not in Africa, and Mauritius because it was doing well and was not undergoing a World Bank adjustment program during the period covered by the World Bank study. Yet both of these economies would seem to satisfy the conditions necessary to be rated as closed according to the export-monopoly criterion: Indonesian law restricts oil and gas production to the state oil company, Pertamina; and Mauritius sells all of its export sugar production through the Mauritius Sugar Syndicate.²⁵ Indonesia and Mauritius are also among the ten fastest-growing economies in Sachs and Warner's sample.

24. Sachs and Warner (1995) cite a different source in their paper, but World Bank (1994) appears to be the correct source.

25. See Pertamina (1998) for Indonesia, and Gulhati and Nallari (1990, p. 22) as well as World Bank (1989, p. 6) for Mauritius. Oil represented 61.2% of Indonesian exports and sugar represented between 60–80% of Mauritius exports during the period covered by Sachs and Warner's study (see World Bank, 1983, Table E, and 1998). Although manufactures have recently outstripped sugar as Mauritius's main export, this is a recent development: in 1980 sugar represented 65% of Mauritius's total exports, and agriculture was surpassed by manufacturing as the main source of exports only in 1986 (World Bank, 1998).

One of the problems that this selection bias causes in Sachs and Warner's estimation is that it makes the variable MON virtually indistinguishable from a sub-Saharan Africa dummy.²⁶ There are 13 African countries (out of 47) in Sachs and Warner's study that are not rated as closed according to MON. (Twelve of these were not included in World Bank study.) But for all but one of these observations MON adds no additional information, either because they are dropped from the sample due to unavailability of other data or because they are rated as closed by other trade-policy indicators used to construct the index. The result is that the only difference between having used an export-marketing-board variable to construct the SW index and having used a sub-Saharan Africa dummy is a single observation. That observation is Mauritius, the fastest-growing African economy in the sample.²⁷

We conclude that the export-marketing-board variable, as implemented, is not a good measure of trade policy and creates a serious bias in the estimation. Except for Mauritius, whose classification as open seems to us to be due exclusively to selection bias, the inclusion of MON in the SW dummy is indistinguishable from the use of a sub-Saharan Africa dummy. In that respect, the only information that we can extract from it is that African economies grew more slowly than the rest of the world during the seventies and eighties.

4.3 WHAT DOES THE BLACK-MARKET PREMIUM VARIABLE MEASURE?

The second source of strength in the SW openness variable is the black-market premium. Indeed, the simple correlation between the openness dummy and BMP is 0.63. A regression of growth on the black-market premium dummy and all the other controls gives a coefficient of -1.05 with a t -statistic of nearly 2.5 in absolute value. How good an indicator of openness is the black-market premium?

The black-market premium measures the extent of rationing in the market for foreign currency. The theoretical argument for using the black-market premium in this context is that, under certain conditions, foreign exchange restrictions act as a trade barrier. Using our notation from the previous section (but omitting country subscripts), the domes-

26. This is true despite the fact that the SW dummy's coefficient is still significant after the estimation is carried out controlling for a sub-Saharan Africa dummy. The reason is that the SW dummy still has substantial explanatory power left due to its use of the black-market premium variable.
27. Both Lesotho and Botswana had higher growth rates than Mauritius, but Lesotho was not rated due to insufficient data (Sachs and Warner 1995, p. 85), and Botswana is dropped from their sample because of unavailability of government-consumption data.

tic price of import-competing goods relative to exportables can be expressed as follows:

$$\frac{p^m}{p^x} = \frac{e^m p^{m*} (1 + t^m)(1 + t^x)}{e^x p^{x*}},$$

where an asterisk refers to border prices. We now allow for the possibility that the exchange rates applicable to import and export transactions (e^m and e^x , respectively) can differ. Foreign-currency rationing can drive a wedge between these two exchange rates.

Suppose the form that rationing takes is as follows: all imports are financed at the margin by buying foreign currency in the black market, while all export receipts are handed to the central bank at the official exchange rate. In this case, $e^m/e^x = 1 + \text{BMP}$, and the presence of a black-market premium has the same resource-allocation consequences as a trade restriction. On the other hand, if at the margin exporters can sell their foreign-currency receipts on the black market as well, then the wedge between e^m and e^x disappears. In this case, the black-market premium does not work like a trade restriction.²⁸ Neither does it do so when the premium for foreign currency is generated by restrictions on capital-account (as opposed to current-account) transactions.

But there is a deeper problem with interpreting the black-market premium as an indicator of trade policy. Sachs and Warner rate an economy closed according to BMP if it maintains black-market premia in excess of 20% for a whole decade (the 1970s or the 1980s). Such levels of the black-market premium are indicative of sustained macroeconomic imbalances. Overvaluation of this magnitude is likely to emerge (1) when there is a deep inconsistency between domestic aggregate-demand policies and exchange-rate policy, or (2) when the government tries to maintain a low exchange rate in order to counteract transitory confidence or balance-of-payments crises. Such imbalances may be sparked by political conflicts, external shocks, or sheer mismanagement, and would typically manifest themselves in inflationary pressures, high and growing levels of external debt, and a stop-go pattern of policymaking. In addition, since black-market premia tend to favor government officials who can trade exchange-rate allocations for bribes, we would expect them to be high wherever there are high levels of corruption. Therefore, countries with greater corruption, a less reliable

28. In one respect, Sachs and Warner (1995) treat BMP differently from a trade restriction: the cutoff for tariffs (TAR) is set at 40%, while that for BMP is set at 20%.

bureaucracy, and lower capacity for enforcement of the rule of law are also likely have higher black-market premia.

Hence it is reasonable to suppose that the existence of sizable black-market premia over long periods of time reflects a wide range of policy failures. It is also reasonable to think that these failures will be responsible for low growth. What is more debatable, in our view, is the attribution of the adverse growth consequences exclusively to the trade-restrictive effects of black-market premia.

Many of the relationships just discussed are present in the data. The simple correlations of black-market premia with the level of inflation, the debt/exports ratio, wars, and institutional quality are all sufficiently high to warrant preoccupation. Indeed, of the 48 economies ranked as closed according to the BMP criteria, 40 had one or more of the following characteristics: average inflation over 1975–1990 higher than 10%, debt-to-GNP ratio in 1985 greater than 125%, a terms-of-trade decline of more than 20%, an institutional-quality index less than 5 (on a scale of 1 to 10), or involvement in a war.

We also view the fact that there exist important threshold effects in the black-market premium as indicative that this variable may simply be capturing the effect of widespread macroeconomic and political crises. If we insert the values of the black-market premium in the 1970s and 1980s as continuous variables in the regression, the estimated coefficients are extremely weak, and they fail to pass an *F*-test for joint significance at 10%. The strength of Sachs and Warner's result comes in great part from the dichotomous nature of the variable BMP and from the fact that the 20% threshold allows more weight to be placed on the observations for which the black-market premia—and probably also the underlying macroeconomic imbalances—are sufficiently high.

That the effect of the black-market premium is highly sensitive to the macroeconomic and political variables that one controls for is shown in Table 4, where we present the results of controlling for each of the indicators of macroeconomic and political distress that we have mentioned. In three out of five cases, each of these variables individually is enough to drive the coefficient on BMP below conventional levels of significance. If we insert all our controls together, the estimated coefficient on BMP goes down by more than half and the *t*-statistic drops below 1.

This kind of evidence does not by itself prove that higher black-market premia are unrelated to growth performance. The results in Table 4 might be due to high multicollinearity between the black-market premium and the indicators of macroeconomic and political distress that we have chosen. But what they do show is that there is very little in the data

Table 4 EFFECT OF BLACK-MARKET PREMIUM ON GROWTH BEFORE AND AFTER CONTROLLING FOR MEASURES OF MACROECONOMIC AND POLITICAL DISEQUILIBRIUM;

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Black-market premium	-1.044** (-2.47)	-0.727 (-1.57)	-0.768 (-1.62)	-1.200* (-2.84)	-0.945** (-2.31)	-0.551 (-1.66)	-0.438 (-0.98)
Inflation, 1975-1990		-3.201*** (-1.78)					-1.024 (-0.58)
Debt/GDP ratio in 1985			-0.015* (-5.75)				-0.011* (-3.21)
Terms-of-trade shock				1.038 (0.42)			3.894 (1.48)
War					-1.378** (-2.32)		-0.135 (-0.15)
Quality of institutions						0.441* (2.86)	0.433*** (2.00)
Summary statistics							
R ²	0.476	0.382	0.589	0.496	0.507	0.567	0.703
N	80	76	54	77	80	75	46

Dependent variable: growth of GDP per capita, 1970-1989. All equations include the following controls: log of GDP in 1970, investment rate in 1970, government consumption/GDP, assassinations per capita, deviation from world investment prices, secondary-schooling ratio, primary-schooling ratio, revolutions and coups, and a constant term. Numbers in parentheses are *t*-statistics based on Huber-White heteroskedasticity-consistent standard errors.

to help us distinguish the effect of high black-market premia from those of other plausible right-hand-side variables relating to macroeconomic distress. In other words, they show that the black-market premium is not a good measure of trade policy, because it is also a proxy for many other variables unrelated to trade policy.

4.4 SENSITIVITY AND GENERAL IMPLICATIONS

The interpretational problems with the state-monopoly-of-exports and black-market premium variables would not be so important if these two variables were responsible for only part of the effect of the SW index on growth. But the fact that they seem to be its overwhelming determinant makes us worry about the extent to which the results speak meaningfully about the role of trade policies.

The arguments in the previous two sections have shown that the individual coefficients on MON and BMP are not very robust to controlling for variables such as an Africa dummy or indicators of macroeconomic and political distress. However, much of the force of the SW variable comes from its combination of the effects of MON and BMP. The reason is that the SW dummy uses MON to classify as closed all but one of the economies in sub-Saharan Africa and then uses BMP to classify as closed a set of economies with macroeconomic and political difficulties. It thus builds a "supervariable" which is 1 for all non-African economies without macroeconomic or political difficulties. This variable will be statistically stronger than either an African dummy or macroeconomic controls, because it jointly groups information from both.²⁹

In the working-paper version of this paper (Rodríguez and Rodrik, 1999) we show that the coefficient on the SW variable, although generally robust to changes in the list of controls, is particularly sensitive to the inclusion of other summary indicators of macroeconomic and political crises. In particular, both the summary indicator of institutional quality developed by Knack and Keefer (1995) and a dummy variable that captures the effect of being in Africa and high macroeconomic disequilibria can easily drive the coefficient of the SW dummy below conventional significance levels. This sensitivity is important not because it shows the existence of a specification in which the SW dummy's significance is not robust, but because this lack of robustness shows up precisely when it is other indicators of political and macroeconomic imbalances that are introduced in the regression. This appears to suggest that

29. If MON and BMP are inserted separately, together with an Africa dummy and a measure of institutional quality, then neither MON nor BMP is individually significant, and the p -value for a joint significance test is 0.09 (0.31 after controlling for NTB, TAR, and SOC), but OPEN gets a t -statistic of 3.06 and BM one of 2.93 (SQT gets 1.46).

the SW variable may be acting as a proxy for these imbalances rather than as an indicator of trade policy.

We do not pretend to have a good answer to the question of whether it is macroeconomic and political distress that drive trade policy or the other way around.³⁰ Nor do we give an answer to the question of whether all of these are determined in turn by some other underlying variables such as poor institutions or antimarket ideology. What we believe we have established is that the statistical power of the SW indicator derives not from the direct indicators of trade policy it incorporates, but from two components that we have reasons to believe will yield upward-biased estimates of the effects of trade restrictions. The SW measure is so correlated with plausible groupings of alternative explanatory variables—macroeconomic instability, poor institutions, location in Africa—that it is risky to draw strong inferences about the effect of openness on growth based on its coefficient in a growth regression.

5. *Sebastian Edwards (1998)*

The third paper that we discuss is Sebastian Edwards's recent *Economic Journal* paper "Openness, productivity and growth: What do we really know?" (Edwards, 1998). The papers by Dollar and by Sachs and Warner deal with data problems by constructing new openness indicators. Edwards takes the alternative approach of analyzing the robustness of the openness–growth relationship to the use of different existing indicators. Edwards writes: "the difficulties in defining satisfactory summary indexes suggest that researchers should move away from this area, and should instead concentrate on determining whether econometric results are robust to alternative indexes" (1998, p. 386). The presumption is that the imperfections in specific indicators would not seem quite as relevant if the estimated positive coefficient on openness were found to be robust to differences in the way openness is measured.

To carry out this robustness analysis, Edwards runs regressions of total factor productivity growth on nine alternative indicators of openness. (Initial income and a measure of schooling are used as controls.³¹)

30. Sachs and Warner's view is that causality goes from restrictive trade policies to macroeconomic instability (personal communication with Sachs). For the purposes of the present paper, we are agnostic about the existence or direction of any causality. An argument that macroeconomic imbalances are largely unrelated to trade policies is not difficult to make, and receives considerable support from cross-national evidence (see Rodrik, 1999, Chap. 4).

31. In an earlier and heavily cited paper, Edwards (1992) carried out a similar analysis for growth rates of real GDP per capita using a somewhat different set of nine alternative indicators of trade-policy distortions. We focus here on Edwards (1998) because it is more recent and the data set used in the earlier paper was not available.

His estimates of total factor productivity growth are the Solow residuals from panel regressions of growth on changes of capital and labor inputs. The nine indicators of openness he uses are: (1) the SW openness index; (2) the World Bank's subjective classification of trade strategies in *World Development Report 1987*; (3) Leamer's (1988) openness index, built on the basis of the average residuals from regressions of trade flows; (4) the average black-market premium; (5) the average import tariffs from UNCTAD via Barro and Lee (1994); (6) the average coverage of nontariff barriers, also from UNCTAD via Barro and Lee (1994); (7) the subjective Heritage Foundation index of distortions in international trade; (8) the ratio of total revenues on trade taxes (exports + imports) to total trade; and (9) Holger Wolf's (1993) regression-based index of import distortions for 1985.

The results Edwards presents are weighted least squares (WLS) regressions of TFP growth on indicators (1)–(9), where the weighting variable is GDP per capita in 1985. They are shown in column 1, rows 1–9, of Table 5: six of the nine indicators are significant, and all but one have the “expected” sign. He repeats the analysis using instrumental weighted least squares (column 2), and finds five of nine indicators significant at 10% (three at 5%) and all having the “correct” sign.³² He also builds an additional indicator as the first principal component of indicators (1), (4), (5), (6), and (9), which he finds to be significant in WLS estimation (row 10). He concludes that “these results are quite remarkable, suggesting with tremendous consistency that there is a significantly positive relationship between openness and productivity growth.”

We will argue that Edwards's evidence does not warrant such strong claims. The robustness of the regression results, we will show, is largely an artifact of weighting and identification assumptions that seem to us to be inappropriate. Of the 19 different specifications reported in Edwards (1998), only three produce results that are statistically significant at conventional levels once we qualify these assumptions. Furthermore, the specifications that pass econometric scrutiny are based on data that suffer from serious anomalies and subjectivity bias.

5.1 THE PROBLEM WITH WEIGHTING

The justification for the resort to WLS estimation is not provided in the paper, but it is presumably to correct for possible heteroskedasticity in the residuals. If disturbances are not homoskedastic, ordinary least-squares estimates will be inefficient. If the form of the skedastic function

32. In his paper, Edwards erroneously claims that two additional variables are significant in the IV-2SLS estimation: Leamer's index and tariffs. This mistake was apparently due to two typographical errors in his Table 4, p. 393.

Table 5 ALTERNATIVE WEIGHTING ASSUMPTIONS

<i>Openness indicator</i>	(1) <i>Weighted least squares</i> (<i>weight=GDP</i>)	(2) <i>Weighted 2SLS</i> (<i>weight=GDP</i>)	(3) <i>Weighted least squares</i> (<i>weight=ln GDP</i>)	(4) <i>Weighted 2SLS</i> (<i>weight=ln GDP</i>)	(5) <i>Robust standard errors</i>	(6) <i>2SLS, Robust Standard Errors</i>
1. Sachs–Warner	0.0094** (2.12)	0.0089*** (1.84)	0.0101*** (1.81)	0.0080 (1.28)	0.0102 (1.54)	0.0078 (1.06)
2. World Development Report	0.0075* (3.57)	0.0131* (3.36)	0.0070** (2.45)	0.0126** (2.64)	0.0068* (3.67)	0.0126** (2.13)
3. Leamer	0.0010 (1.03)	0.0123 (1.40)	0.0041 (0.82)	-0.0013 (-0.20)	0.0041 (0.82)	-0.0033 (-0.32)
4. Black-market premium	-0.0217* (-3.59)	-0.0192*** (-1.95)	-0.0108** (-2.57)	-0.0035 (-0.56)	-0.0098*** (-1.79)	-0.0027 (-0.54)
5. Tariffs	-0.0450* (-2.77)	-0.1001 (-1.52)	0.0065 (0.51)	0.0013 (0.03)	0.0114 (0.88)	0.0079 (0.28)

6. Quotas	-0.0047 (-0.45)	-0.0398 (-0.42)	0.0029 (0.35)	0.0461 (0.68)	0.0036 (0.43)	0.0401 (0.79)
7. Heritage Foundation	-0.0074* (-4.50)	-0.0133* (-3.75)	-0.0066** (-3.02)	-0.0195* (-3.30)	-0.0064* (-2.87)	-0.0202* (-3.24)
8. Collected-trade-taxes ratio	-0.4849* (3.04)	-1.6668** (-2.15)	-0.2808** (-2.15)	-1.8256 (-1.23)	-0.2676** (-2.25)	-1.8368 (-1.06)
9. Wolf's index of import distortions	3.5E-05 (0.27)	-2.6E-04 (-0.72)	4.8E-05 (0.41)	-3.7E-04 (-0.99)	4.1E-05 (0.36)	-3.3E-04 (-1.21)
10. Principal-components factor	-0.0070** (-2.38)		-0.0047 (-1.61)		-0.0043 (-1.37)	

Dependent variable: TFP growth, 1980–1990. These are the estimated coefficients from regressions where each of the trade-policy indicators is entered separately. Each equation also includes log GDP per capita in 1965 and schooling in 1965 as regressors [as in the original Edwards (1998) specification]. *t*-statistics are in parentheses (based on heteroskedasticity-consistent standard errors in column 3).

is known, then it is appropriate to use WLS. This is indeed what Edwards implicitly assumes when he uses GDP per capita as his weighting variable. If it is unknown, White's (1980) covariance-matrix estimator allows for the calculation of heteroskedasticity-robust standard errors that are invariant to the form of the skedastic function.

When there is heteroskedasticity, the standard deviation of the disturbance in the growth equation varies systematically across countries. Edwards's decision to weight his observations by the level of GDP per capita implies an assumption that the standard deviation of the disturbances in the growth equation is inversely proportional to the square root of the level of GDP per capita in 1985. In other words, if the United States is—as it in effect was in 1985 according to the Summers–Heston data—59 times wealthier than Ethiopia, the standard deviation of the growth rate conditional on having the United States's income is 7.7 ($59^{1/2}$) times lower than conditional on having Ethiopia's income. Using the estimates of the residuals' standard deviation from one of Edwards's equations, we can calculate the implied root-mean-square error of the growth rate conditional on having the incomes of the United States and of Ethiopia. The former is 0.8 percentage points, whereas the latter is 6 percentage points. It may be reasonable to suppose that growth data for poor countries are less reliable than those for rich countries, but the errors implied by Edwards's weighting assumption for poor countries' growth data seem to us to be unreasonably high. As a matter of fact, it is hard to think of a reason to be doing regression analysis on a broad cross section of primarily poor countries if we believe that underdeveloped nations' economic data are this uninformative.

Columns 3 and 4 of Table 5 repeat Edwards's regressions using the natural log of 1985 per capita GDP as the weighting variable. In terms of our calculations above, the ratio between the U.S. and Ethiopian standard deviations would now be a more reasonable 1.31. This set of regressions results in six of the eighteen coefficients having the “wrong” sign. Five out of nine coefficients are significant among the least-squares regressions (four at 5%), and two out of nine in the instrumental variables (IV) regressions. The coefficient on the principal-components variable now becomes insignificant.³³

33. Why does weighting by GDP give such different results? The reason seems to be that there is a relationship between the openness indices used by Edwards and TFP growth at high levels of income. This relationship in itself is apparently driven by the fact that the great majority of economies with restrictive trade practices and high levels of GDP per capita in 1985 were oil exporters. Because of their high incomes, these economies are weighted very heavily in the WLS regressions. It is well known that oil-exporting economies had very low rates of growth during the 1980s (see for example the studies in Gelb, 1988). If one redoes regressions 1–19 using GDP per capita weights but includ-

One way to put aside doubts about the appropriateness of alternative assumptions regarding the nature of the skedastic function is to use White's (1980) heteroskedasticity-consistent standard errors, which are robust to the form of heteroskedasticity. We show these estimates in column 5 and 6 of Table 5. Four out of nine coefficients are now significant among the least-squares regressions (three at 5%), and two out of nine among the IV regressions. Only twelve of the eighteen coefficients have the correct sign. The principal components variable is also insignificant.

5.2 THE PROBLEM WITH IDENTIFICATION

The two significant IV coefficients in Table 5 are moreover quite sensitive to the specification of the instrument lists. In particular, the IV versions of equations 2 and 7 in Table 5 are two of the only three equations in which the Heritage Foundation Index of Property Rights Protection is used as an instrument by Edwards.³⁴ If this instrument is not excludable from the second-stage regression, Edwards's IV estimation will give biased estimates of the coefficient of openness on growth. Theoretically, it seems to us unreasonable to assert that the protection of property rights can effectively be assumed not to be an important determinant of growth, given the extensive literature concerned precisely with such an effect.³⁵ In Table 6, columns 1–4, we show that, if property rights are included in the second-stage regression for these two equations, this term gets a significant coefficient in indicator 2 (World Development Report index) and a positive albeit insignificant coefficient in indicator 7 (Heritage Foundation index). Chi-squared tests of the overidentifying restrictions also reject the null hypothesis that these restrictions hold for indicator 2. Furthermore, in both indicators the *t*-statistic on the openness proxy falls to well below 0.5 in absolute value.

If we take seriously the fact that property rights are not excludable from the productivity growth regressions, we are left with the conclusion that, among 17 different specifications in Tables 5 and 6, we find

ing a dummy for oil exporters, one gets very similar results to those in column 3 of Tables 5 and 6. Only the coefficients for the World Development Report index, the Heritage Foundation index, and the least-squares estimate of the collected-taxes ratio remain significant, and the least-squares coefficient on quotas changes sign.

34. His other instruments include TFP growth in the 1970s and the black-market premium, export/GDP, import/GDP and terms-of-trade changes for 1975–1979.

35. Barro (1997) names "the importance of institutions that ensure property rights and free markets" for economic growth as one of the "dominant themes" of his recent research (p. xiv). For examples of the literature emphasizing the importance of property rights for economic growth, see Clague et al. (1996), Acheson and McFetridge (1996), Jodha (1996), Tornell (1997), Park and Ginarte (1997), Grossman and Kim (1996), and Thompson and Rushing (1996).

Table 6 SENSITIVITY TO IDENTIFICATION ASSUMPTIONS AND CHOICE OF TRADE TAX INDICATOR

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
World Development Report index	0.0126** (2.13)	0.0023 (0.40)						
Heritage Foundation index			-0.0202* (-3.24)	-0.003 (-0.24)				
Property rights		-0.0107* (-2.91)		-0.010 (-1.43)				
Collected-taxes ratio (Edwards)					-0.2676 (-2.25)**			
Average duty (World Bank)						0.0225 (1.01)		
Average import duty (World Bank)							0.0007 (2.30)**	0.0003 (0.884)
Average export duty (World Bank)							-0.0003 (-1.09)	
Test of overidentifying restrictions	29.3244		5.4072					
p-value	6.72E-06		0.2480					
N	30	30	56	56	45	43	43	66

Dependent variable: TFP Growth, 1980–1990. Each equation also includes log GDP per capita in 1965 and schooling in 1965 as regressors. *t*-statistics based on heteroskedasticity-consistent standard errors in parentheses.

evidence of a negative and statistically significant correlation between trade-restricting policies and productivity growth in only three cases. Those are the ones that use the collected-taxes ratio, the World Development Report index, or the Heritage Foundation index. We take up some problems with these indexes in the next subsection.

5.3 DATA ISSUES

Edwards reports that the collected-taxes ratio (which measures trade tax revenue as a proportion of total trade) is calculated from raw data provided by the IMF. We are puzzled by these data, because many of the numbers for developing countries are implausible. India, a country with one of the world's highest tariff rates, is listed as having an average ratio of 2.4% lower than the sample average and barely above the value for Chile (2.3%). The mean value of the collected-taxes ratio in the sample is 2.8%, which strikes us as very low.

We have attempted to replicate Edwards's results using data from the World Bank's World Development Indicators (1998). This source, which was not available at the time Edwards's analysis was first conducted, provides collected trade tax ratios for imports and exports separately, which we have combined to derive an index in the spirit of Edwards's variable.³⁶ According to this index, India's average trade tax is 37.3% (a more plausible figure than Edwards's 2.4%). We replicate equation 8 of Table 5 with these data, and the results are shown in columns 6–8 of Table 6. The coefficient on average duties is now insignificant and has the "wrong" sign (column 6). If we introduce import and export duties separately (column 7), then import duties in fact get a positive and *significant* coefficient (contrary to the expected negative coefficient), and export duties are insignificant.

One shortcoming of these specifications (including Edwards's) is the small sample size (between 43 and 45). Since export duties are not reported for many countries, one way of increasing the sample size is to introduce *only* the import-duty variable from the World Development Indicators database. This increases the sample size to 66 countries. The estimated coefficient on import duties is once again positive and insignificant (column 8).

These results are in line with others we have reported earlier: there is

36. As our earlier discussion showed, when imports and exports are both taxed, their distortionary effect is multiplicative rather than additive. So instead of summing import and export taxes, we use the formula $(1 + \text{mdut})(1 + \text{xdut}) - 1$, where mdut (xdut) is import (export) duties as a percentage of imports (exports). We take the average of observations for 1980–1985. Our results (on the sign and insignificance of the coefficient on trade taxes) are unchanged, however, when we take the simple sum $\text{mdut} + \text{xdut}$.

little evidence that simple averages of trade taxes are significantly and negatively correlated with growth.

The other two variables that are significant are the subjective indexes constructed by the World Bank and the Heritage Foundation. It is striking that two subjective indexes are the only variables that are robust to our econometric analysis, since subjective indexes are well known to suffer from judgment biases. Indeed, a look at the two indexes reveals some striking contrasts. In the Heritage Foundation Index, for example, Chile and Uganda are in the same category (4 on a scale of 1 to 5, where 5 is most protected). Perhaps even more problematic is the fact that the Heritage Foundation index rates policies in 1996, well after the end of Edwards's sample period (1980–1990). Similar problems are present in the World Bank index, where high-growth Korea is rated as more open than moderate-growth Malaysia despite having higher tariff rates and nontariff-barrier coverage as well as a lower export/GDP ratio, and moderate-growth Tunisia—which had average tariffs of 21% and average nontariff coverage of 54%—is classified in the same group as Chile, Malaysia, and Thailand. In fact, in his 1993 literature review, Edwards (1993, pp. 1386–1387) himself drew attention to serious problems with this index. As he noted, Chile, which in other studies is rated as the most open economy in the developing world, was grouped in the second category (moderately outward-oriented); Korea was classified in the group of most open economies for both 1963–1973 and 1973–1985 despite the fact that in the former period the Korean trade regime was considerably more restrictive than in the latter.

In the working-paper version of this paper we report the results of recomputing these subjective indexes using the quantitative information on which they are purportedly based. Given that these underlying data are no different from those used in some of the other empirical work that we have discussed in this and other sections of the paper, it should come as no surprise that these attempts generally yielded insignificant coefficients. The natural conclusion from these results appears to be that either the mismatch in time periods or subjectivity biases or both are the fundamental causes for the significance of the Heritage Foundation and World Bank indexes.

In sum, we do not concur with Edwards's assertion that the cross-country data reveal the existence of a robust relationship between openness and productivity or GDP growth.³⁷ In our view, there is little evi-

37. Our results are basically unaltered if we use growth of GDP per capita from 1980 to 1990 instead of TFP growth as the dependent variable. In this case the World Bank and Heritage Foundation indexes remain significant, but the collected-trade-taxes ratio is now only significant at a 10% level and the black-market premium is insignificant. Similar results emerge for IV estimation.

dence to support such an assertion. The results reviewed in this section are for the most part highly dependent on questionable weighting and identification assumptions. The trade-policy indicators whose significance is not affected by these assumptions either are subjective indexes apparently highly contaminated by judgement biases or lack robustness to the use of more credible information from alternative data sources.

6. Dan Ben-David (1993)

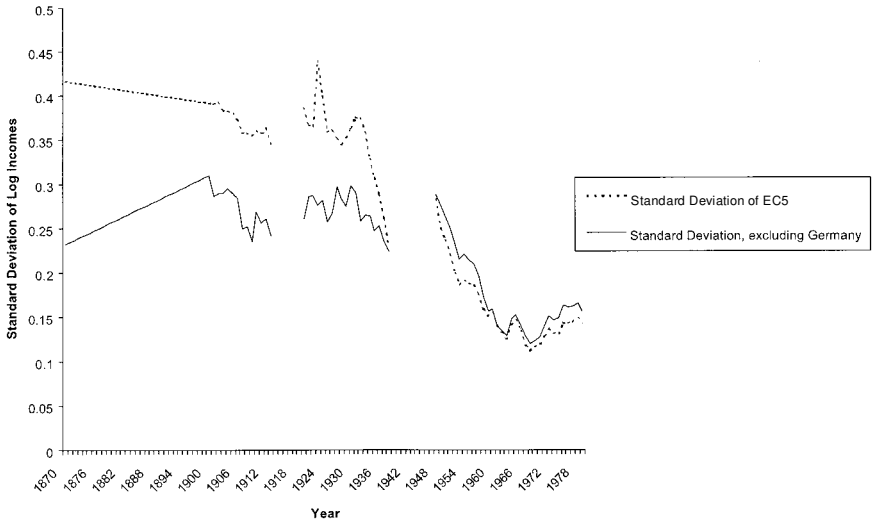
Ben-David's (1993) *QJE* paper "Equalizing exchange: Trade liberalization and income convergence" takes an altogether different approach to studying the effect of openness on economic growth. Ben-David analyzes the effect of trade policies on income by asking whether trade liberalization leads to a reduction in the dispersion of income levels among liberalizing countries (i.e., whether it contributes to what has been called σ -convergence). We pick Ben-David as an example of a strand of the literature which has centered on studying the effect of trade on convergence. Another distinctive aspect of Ben-David's work is that it is nonparametric and not regression-based.

The expectation that trade liberalization might lead to income convergence is grounded in the factor price equalization (FPE) theorem. According to trade theory, free trade in goods leads to the equalization of factor prices under certain conditions (including an equal number of goods and factors, identical technologies, and absence of transport costs). As barriers to trade are relaxed (and assuming in addition that differences in capital-labor ratios and labor-force participation ratios do not countervail), a tendency towards FPE can be set into motion, resulting in convergence in per capita incomes.

There is no necessary relationship between the level of dispersion in incomes and the growth rate. Countries could in principle be converging to lower levels of GDP per capita. But in the case of the European Community, on which Ben-David concentrates, the convergence experienced was indeed to higher incomes. Overall growth from 1945 to 1994 of the EC5 (Belgium, France, the Netherlands, Italy, and Germany) was 3.45% compared to 1.21% percent from 1900 to 1939 and 1.16% from 1870 to 1899. Therefore, if Ben-David's claim is right, convergence in the EEC was achieved by raising the income of poor countries rather than by lowering that of rich countries.

Ben-David's argument goes beyond simply ascertaining that a decrease in dispersion occurred during the postwar era. He tries to show that trade liberalization caused this decrease by discarding other plausible alternatives. Thus he argues (1) that the observed convergence was

Figure 3 EFFECT OF EXCLUDING GERMANY IN DISPERSION CALCULATIONS



not simply a continuation of a long-term convergence trend unrelated to postwar economic integration; (2) that the European countries that chose not to enter a free-trade agreement did not experience the same extent of convergence as the EEC; and (3) that other subsets of economies in the world that were not economically integrated did not experience convergence. We examine each of his arguments in turn.

6.1 WAS EUROPEAN CONVERGENCE A CONTINUATION OF A LONG-TERM TREND?

In support of the argument that the reduction in dispersion was not simply the continuation of a long-run trend, Ben-David argues that the series of per capita income dispersion (solid line in Figure 3) does not show any visible downward tendency before the postwar era. When presenting this series, Ben-David excludes Germany from the calculations,³⁸ arguing that not doing so would bias the conclusion in favor of convergence:

Germany was always among the poorest, in per capita terms, of the six countries. Today, it is one of the wealthiest countries in Europe. As a result of its heightened prosperity, it might be claimed that all of the convergence that

38. Luxembourg is also excluded, because Maddison (1982) does not provide data for it.

has been witnessed within the EEC is due to the behavior of Germany. Thus, its exclusion should bias the results away from convergence. (Ben-David, 1993, p. 662)

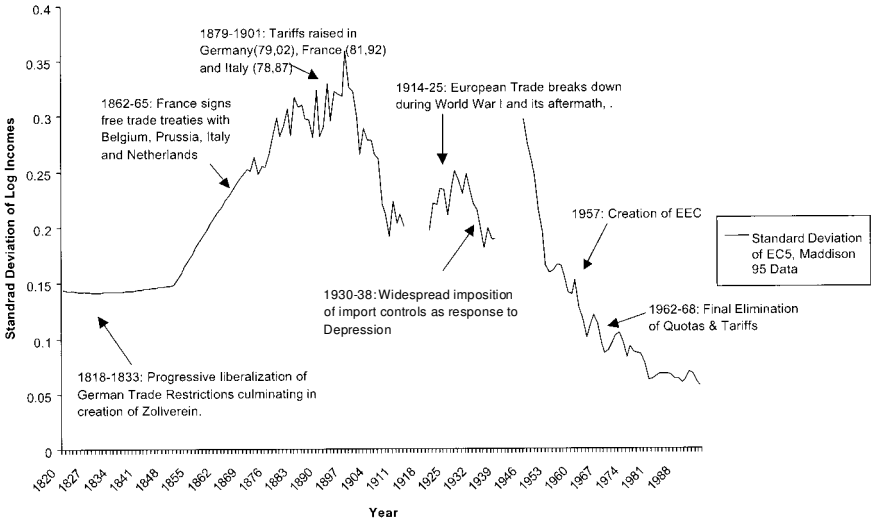
Note however that the purpose of Figure 3 (Figure VII in Ben-David's paper) is not only to establish the existence of convergence following postwar liberalization, but also to establish the *absence* of a long-term trend in convergence predating it. Thus the exclusion of Germany from the series, which biases the results against convergence, would also bias the results in favor of the hypothesis that there was no prewar convergence trend, had Germany's convergence occurred before the postwar period.

That is indeed what happened. Between 1870 and the eve of World War II, Germany's income went from less than 50% to 75% of the average for the remaining members of the EEC. And by 1958, one year after the EEC was formed, Germany had surpassed Belgium as the leader of the five. The exclusion of Germany therefore has the effect of understating the fall in dispersion before the creation of the EEC. The dashed line in Figure 7, which displays the dispersion of log per capita incomes including Germany, shows this. Once Germany is included in the sample, it appears that dispersion has been on a downward trend since 1870. The hypothesis that postwar convergence was simply a continuation of a long-term trend can no longer be rejected easily, raising doubts about the conclusion that convergence was caused by postwar trade policies.³⁹

Figure 4 plots the standard deviation of log incomes for the original members of the EEC, now using Maddison's more recent (1995) estimates and including Germany. We reach the same conclusion as in Figure 3: dispersion has followed a downward trend since the beginning of the twentieth century. From a peak of 0.36 in 1897, dispersion had fallen to 0.25 in 1930, and 0.19 in 1939. By the time the EEC was created, it had fallen to 0.16. It appears therefore that the further reduction in disper-

39. Ben-David (in personal communication) has pointed out to us that much of the prewar convergence is due to the fact that "while the other countries were in the Depression, Germany surged ahead as Hitler built his war machine." Indeed, dispersion appears trendless from 1900 to 1932, and starts falling only as Germany's income rises during the National Socialist period. But we are not sure of what to make of that fact. Germany's income remained high after the war—compared to other European countries—suggesting that not all of the convergence was due to the policies of the Nazi period or to the buildup of the war machine. In any case, Nazi Germany pursued highly protectionist policies, so that its experience sheds doubt on the argument that poor countries that close their economies experience slower growth. Finally, the observation for 1870 in Figure 7 suggests that dispersion was much higher in the late nineteenth century than in 1930. The last point is confirmed when we examine Maddison's (1995) more recent estimates (see Figure 8), which provide a fuller picture of trends in dispersion since 1820. These estimates were not available to Ben-David at the time his paper was written.

Figure 4 DISPERSION OF PER CAPITA INCOMES AND TRADE POLICY EVENTS



sion that followed the creation of the EEC (to 0.06 by 1994) was a continuation of a long-term trend that predated European integration. Moreover, this conclusion is not sensitive to whether Germany is included in the sample: that is because Maddison's (1995) revised estimates suggest that there was a uniform pattern of convergence during the pre-World War I period, with Italy, France, and Germany *all* catching up with Belgium and the Netherlands.

A closer look at Figure 4 suggests that there is in fact very little association between episodes of economic integration and σ -convergence over time. The period leading up to 1878 was an era of continuous trade liberalization, at the level of both national markets and international ones. This period witnessed the creation of the German *Zollverein* (1833) and the unification of Italy (1860), as well as the signing of free-trade agreements between Prussia and Belgium (1844), France and Belgium (1842), France and Prussia (1862), France and Italy (1863), and France and the Netherlands (1865).⁴⁰ Most of these bilateral agreements had

40. The discussion in this and the following two paragraphs borrows heavily from Chapter V of Pollard (1974). Above we list treaties between countries included in Figure 4, but the extent of trade liberalization from 1820 to 1878 in Europe was impressive. Prussia signed free-trade treaties with Britain (1841 and 1860), Turkey (1839), Greece (1840), Austria (1868), Spain (1868), Switzerland (1869), Mexico (1869), and Japan (1869); France with Britain (1860), Switzerland (1864), Sweden, Norway, the Hanse Towns,

most-favored-nation clauses, extending the benefits of bilateral liberalization to third countries. Yet, despite increasing economic integration, dispersion more than doubled from 1820 to 1880 (from 0.14 to 0.29).⁴¹

The retreat from free trade started during the 1880s, with Germany's Tariff Act of 1879. Italy raised tariffs in 1878 and 1887, France in 1881 and 1892.⁴² This rise in protection followed the depression of the 1870s and was motivated by the desire to protect European farmers from the influx of cheap American grain imports (which began to undersell German grain in 1875) while at the same time compensating industry for the increased wages of workers.⁴³ Nevertheless, as Figure 4 shows, the period from the 1880s to World War I was, if anything, one of convergence.⁴⁴

The breakdown in world trade that followed World War I and the spread of beggar-thy-neighbor protectionist policies adopted during the Great Depression seem also to have had very little effect on dispersion. Even though fascist governments in Italy and Germany raised agricultural tariffs and other protectionist barriers, and in France the power of agricultural groups was large enough to drive the French price of wheat in 1939 to three times its price in London (Cobban, 1965, p. 156), on the eve of World War II dispersion stood at its lowest level since the 1860s.

In sum, Figure 4 shows no long-run tendency for trade liberalization to be associated with greater convergence in per capita incomes. If anything, it shows increasing dispersion during the nineteenth century and falling dispersion during the twentieth century. While one can interpret this evidence in different ways, we find the most straightforward read-

and Spain (1865), Austria (1866), and Portugal (1867); Belgium with Britain (1862); Italy with Britain (1862) and Turkey and Greece (1839–1940). Aside from the MFN clause, measures were taken to ease international trade such as the inclusion in the Treaties of Berlin of clauses extending commercial freedoms to foreign citizens (1878, 1885). There were even attempts to create customs unions between France and Germany and between France and its neighbors.

41. A caveat applies here: for 1820–1850 we rely on just two observations: one for 1820, and another one for 1850. Since the 1850 observation for Italy was not available, we constructed it as the result of a linear interpolation between the 1820 and the 1870 observation. Even if we disregard the evidence before 1870, the yearly data from 1870–1880 indicate that the increase in dispersion predated the first protectionist measures.
42. Again, tariff adoption was widespread, with only Holland and the United Kingdom resisting the reversion towards protectionism.
43. In effect, high tariffs worked to the detriment of labor in what came to be known in Germany as the “compact of rye and iron.” See Gerschenkron (1943) and Rogowski (1989) for detailed discussions of this era. As Rogowski points out, the reversion towards protectionism was more accentuated in capital-poor countries such as Germany, Italy, and France than in capital-rich countries such as Belgium and the Netherlands.
44. O'Rourke's (1997) econometric study of this period (1975–1914), covering a panel of 10 countries, finds that higher tariffs were correlated with *faster* economic growth, and that the estimated effects are quantitatively large.

ing to be that post-World War II convergence was in fact a continuation of a long-run trend that got started around the turn of the twentieth century.

6.2 DID NON-EEC EUROPEAN COUNTRIES EXPERIENCE CONVERGENCE?

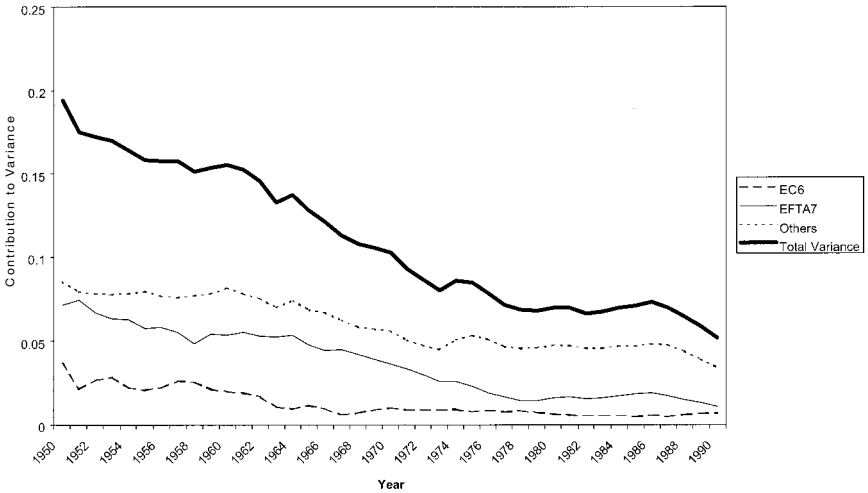
Ben-David also claims that countries in Europe that did not undertake trade liberalization failed to experience convergence. He supports his argument by showing that (a) there was no convergence among the United Kingdom, Denmark, and Ireland until they began to relax their trade restrictions vis-à-vis Europe, and that (b) EFTA countries experienced significant convergence with the EEC as trade barriers among them were liberalized.

To demonstrate (a), Ben-David plots the standard deviation among the United Kingdom, Denmark, and Ireland, all of which started liberalizing trade with the EEC in the mid-1960s. He shows that their dispersion *among themselves* started falling only after 1965. It is not clear to us why this is the relevant test, since the trade liberalization in question took place between these countries and Europe as well as amongst themselves. In Figure 5, we show that even if there is an indication of convergence among these three countries after 1965, it is not caused by conver-

Figure 5 GDP OF UNITED KINGDOM, DENMARK, AND IRELAND, RELATIVE TO EEC MEAN



Figure 6 CONTRIBUTION TO VARIANCE AROUND EUROPEAN MEAN



gence to the mean income of EEC members. Ireland has shown very little convergence to the EEC until recent years, and Denmark has oscillated close to the EEC average since the 1950s. The United Kingdom has been converging—downward—to the EEC level steadily (at least) since the 1950s. None of the three countries seem to experience different patterns of convergence after they relaxed trade restrictions with the EEC in 1965.

As regards (b), there has indeed been substantial convergence by EEC and EFTA member countries to the European mean since the 1950s. But we are skeptical whether such convergence can be attributed to trade liberalization. In Figure 6, we plot the contribution to the variance around the European mean⁴⁵ of three subsets of European countries: the six members of the European Economic Community, the seven members of the European Free Trade Association,⁴⁶ and six remaining European countries which did not join either EFTA or the EEC.⁴⁷ It is evident from

45. This is defined as $(1/N_{\text{EUROPE}}) \sum_{i \in J} [(y_i - \bar{y}_{\text{EUROPE}}) / \bar{y}_{\text{EUROPE}}]^2$ for $J = \{\text{EEC6, EFTA6, others}\}$. Normalization by the mean achieves the same purpose as calculating the variance of log incomes (and is more appropriate for large income differences), and putting the expression in terms of the variance (not the standard deviation) ensures that the three components sum to the total.

46. Austria, Switzerland, Sweden, Denmark, Norway, Finland, and the United Kingdom. Even though Portugal was officially a member of EFTA, it was allowed to implement tariffs and to deviate from EFTA policies, so we follow Ben-David in treating it as a non-EFTA country.

47. Cyprus, Greece, Iceland, Ireland, Portugal, and Spain.

Figure 6 that all subgroups have experienced substantial convergence. The non-EFTA and non-EEC countries have seen their contribution to the variance around the European mean fall from 0.085 to 0.034 from 1950 to 1992.⁴⁸ European convergence seems to be the result of factors largely unrelated to trade liberalization.

6.3 DID OTHER AREAS OF THE WORLD EXPERIENCE CONVERGENCE?

To add plausibility to the story that trade liberalization was behind the European trend towards convergence in the postwar era, Ben-David shows that subsets of countries that have not become integrated have experienced no tendency to converge. He points to the well-known fact that the dispersion of world incomes has not decreased in the postwar era (it has actually increased). He also shows that the dispersion of incomes among the world's 25 richest countries (excluding the EEC6) has not decreased either. He compares these experiences with those of economically integrated Europe and U.S. states to show that convergence seems to occur only when there is substantial trade liberalization.

There is an asymmetry in his selection of diverging and converging areas, however. Whereas the regions he shows to be converging are all close to each other geographically, those which are diverging are not. To have a fair standard of comparison, one must ask whether trade liberalization—or its absence—among geographically adjacent economies would lead towards convergence or divergence.

Did subsets of geographically adjacent economies that liberalized trade tend to observe convergence? There are at least two important cases in which the trends in convergence go counter to what we would expect on the basis of Ben-David's argument. Consider the experiences of East Asia and Latin America, two regions with radically different trade policies and which constitute the canonical examples of open and closed economies. If the liberalization–convergence view is right, the relatively open East Asian economies should have converged, whereas the relatively closed Latin American economies should have diverged. In fact, countries in East Asia have steadily diverged since the 1960s, with the standard deviation of their log incomes going from 0.47 in 1960 to 0.81 in 1989.⁴⁹ As for

48. If one includes Turkey as a seventh country in this group, the contribution to dispersion goes from 0.103 in 1950 to 0.053 in 1992. An alternative measure of dispersion around the European mean is the standard deviation of log incomes around the mean log income. The latter measure for the non-EEC, non-EFTA countries falls from 0.15 in 1950 to 0.05 in 1990 (0.20 to 0.10 if Turkey is included).

49. The East Asian countries are Hong Kong, Indonesia, Japan, South Korea, Malaysia, Philippines, Singapore, Taiwan, and Thailand. Data are from Summers and Heston (1994). If the Philippines is excluded, the rise in dispersion is from 0.50 to 0.73.

Latin America, there has been a steady *decrease* in dispersion during the period of import substitution, from 0.55 before the Great Depression to 0.20 in the late 1980s.⁵⁰ More striking, dispersion has sharply risen since the late 1980s, just as Latin American countries liberalized their trade. (See Rodríguez and Rodrik, 1999, for more details.)

Another important counterexample comes from the historical experience of the United States. Figure 7 plots the ratio of GDP per capita for the United States to the average GDP per capita for its three main European trading partners (the United Kingdom, France, and Germany) up to 1938.⁵¹ Trade with Europe was approximately two-thirds of total U.S. trade during the nineteenth century,⁵² and the bulk of that was with those three countries. It is however evident from Figure 7 that despite declining levels of import duties, the United States and Europe steadily diverged between 1820 and 1938. Again, there seems to be no evident relationship between trade liberalization and income convergence.⁵³

We close by drawing attention to Slaughter's (2000) recent examination of the same issue. Slaughter undertakes a systematic analysis by compar-

50. The Latin American countries are Argentina, Brazil, Chile, Colombia, Mexico, and Peru. Data are from Maddison (1995), Summers and Heston (1994), and World Bank (1998). Latin American import substitution policies started rather spontaneously as a response to the collapse of world-wide demand for raw materials in 1929 and the adoption of protectionist measures by the United States and Britain in 1930 and 1931. Most countries abandoned convertibility and imposed trade barriers during this period and did not liberalize until recent years (see Díaz-Alejandro, 1981).
51. The cutoff date of 1938 is chosen because during World War II the Americas overtook Europe as the main destination for U.S. exports. The Americas overtook Europe as the main source of imports much earlier, during World War I. Including observations after 1940 would not change our results: the GDP per capita in 1994 for the United States was still 27% higher than that of its three main European trading partners, despite the fact that after 1944 tariff rates stayed well into the single digits (Bureau of the Census, 1989). Choosing the Americas instead of Europe as a standard of comparison would strengthen our results, as the divergence between U.S. and Latin American incomes during the nineteenth and twentieth centuries has been extremely high (see Haber, 1997), and Canada represents only about half of U.S. trade with the Americas.
52. Before World War II, exports to Europe were 43% of total exports and imports from Europe were 29% of total imports (Bureau of the Census, 1989).
53. Our broader conclusion is not necessarily inconsistent with Ben-David's own reading of the evidence. Ben-David (in personal communication) writes that the main conclusions that can be drawn from his research are that "trade liberalization is associated with income convergence only when (a) the liberalization is comprehensive and (b) the liberalization occurs between countries that trade extensively with each other," and that "there is no evidence that these outcomes hold for poor countries." In fact, Ben-David (1999) has argued that trade flows will be of little use in transferring knowledge to countries with low levels of human capital. This contrasts strongly with much of the discussion in the literature, which has interpreted Ben-David as making the much stronger claim that liberalization leads developing countries to converge with their richer trading partners. A few examples are IMF (1997, p. 84), World Bank (1996, p. 32), Vamvakidis (1996, p. 251), and Richardson *et al.* (1997, p. 100), all of which refer to Ben-David in discussions about developing economies.

Figure 7 RATIO OF U.S. TO EUROPEAN GDP AND IMPORT DUTIES, 1820–1938



ing convergence patterns among liberalizing countries before and after liberalization with the convergence pattern among randomly chosen control countries before and after liberalization. As he emphasizes, this difference-in-differences approach avoids the pitfalls of before-and-after comparisons (nonliberalizing countries too may exhibit the same pattern before and after) or of comparing liberalizing countries with non-liberalizing ones (the liberalizing countries may have been converging prior to the liberalization as well). Hence Slaughter's approach amounts to a more systematic version of the kind of exercise we have carried out above by way of specific illustrations (but using only post-World War II data). Slaughter focuses specifically on four instances of trade liberalization: formation of the EEC, formation of EFTA, liberalization between EEC and EFTA, and Kennedy Round tariff cuts under GATT. His conclusion is that there is no systematic link between trade liberalization and convergence. In fact, he reports that much of the evidence suggests trade liberalization diverges incomes among liberalizers. This parallels our results above.

7. Jeffrey Frankel and David Romer (1999)

Frankel and Romer's (1999) very recent *AER* paper on trade and incomes has received considerable attention since its publication. This paper ana-

lyzes the relationship between trade and income by estimating cross-country regressions of income per capita on the trade–GDP ratio and two measures of country size (population and land area). The authors’ aim is to address the problem of the likely endogeneity of trade with respect to income. So the trade share is instrumented by first estimating a gravity equation, where bilateral trade flows are regressed on geographic characteristics (countries’ size, their distance from each other, whether they share a common border, and whether they are landlocked). The fitted trade values are then aggregated across partners to create an instrument for the actual trade share. An earlier version of Frankel and Romer’s paper included initial income among the regressors in the second-stage equation, so that the results could also be given a growth interpretation. The main finding of the paper is that the IV estimate of the effect of trade on income is if anything greater than the OLS estimate.

As we mentioned in the introduction, this paper is concerned with the relationship between incomes and the *volume* of trade, and does not have immediate implications for trade *policy*. The reason is that the implications of geography-induced differences in trade, on the one hand, and policy-induced variations in trade, on the other, can be in principle quite different. Selective trade policies work as much by altering the structure of trade as they do by reducing its volume. To the extent that policy is targeted on market failures, trade restrictions can augment incomes (or growth rates) even when indiscriminate barriers in the form of geographical constraints would be harmful. Of course, to the extent that selective trade policies are subject to rent seeking, it is also possible that geography-induced variations in trade *underestimate* the real costs of trade restrictions. Ultimately, whether on balance trade policies are used towards benign ends or malign ends is an empirical question, on which Frankel and Romer’s paper is silent.

With regard to the role of trade flows proper, we are concerned that Frankel and Romer’s geographically constructed trade share may not be a valid instrument. The reason is that geography is likely to be a determinant of income through a multitude of channels, of which trade is (possibly) only one. Geography affects public health (and hence the quality of human capital) through exposure to various diseases. It influences the quality of institutions through the historical experience of colonialism, migrations, and wars. It determines the quantity and quality of natural endowments, including soil fertility, plant variety, and the abundance of minerals. The geographically determined component of trade may be correlated with all these other factors, imparting an upward bias to the IV estimate unless these additional channels are explicitly controlled for in the income equation.

As there is a single instrument used in Frankel and Romer's regressions, conventional exclusion–restriction tests performed conditional on a subset of the instruments being excludable from the second-stage regression cannot be carried out. To check whether Frankel and Romer's result can be attributed to nontrade effects of geography, we simply test whether some summary statistics of the geographical factors influencing trade can be excluded from the second-stage regression. We rerun Frankel and Romer's income regressions, adding three summary indicators of geography: (1) distance from the equator (used in Hall and Jones, 1998); (2) the percentage of a country's land area that is in the tropics (from Radelet, Sachs, and Lee, 1997); and (3) a set of regional dummies.

Table 7 shows the results. Columns 1 and 5 replicate Frankel and Romer's (1999) results in their Table 3, for the OLS and IV versions of the income equation, respectively. The other columns show the consequences of introducing the geography variables. The results are highly

Table 7 FRANKEL–ROMER REGRESSIONS WITH ADDITIONAL GEOGRAPHICAL VARIABLES

<i>Variable</i>	1	2	3	4	5	6	7	8
Trade share	0.85 (3.47)	0.57 (3.00)	0.46 (2.36)	0.61 (3.88)	1.97 (1.99)	0.34 (0.41)	0.21 (0.26)	0.25 (0.41)
Disteq		3.58 (9.26)				3.65 (7.98)		
Tropics			-1.42 (-9.84)				-1.46 (-8.03)	
East Asia				-1.21 (-7.71)				-1.21 (-7.59)
Latin America				-0.67 (-4.48)				-0.74 (-3.83)
Sub-Saharan Africa				-1.94 (-14.72)				-1.99 (-12.82)
Method	OLS	OLS	OLS	OLS	IV	IV	IV	IV
<i>n</i>	150	150	145	150	150	150	145	150
<i>R</i> ²	0.0949	0.4312	0.4628	0.66	0.43	0.44	0.4563	0.65

The dependent variable is log of income per person in 1985. IV standard errors include adjustment for generated regressors. All equations include the logs of population and land area. Disteq is distance from equator, as measured by Hall and Jones (1998). Tropics is fraction of country's area in tropics, as measured by Radelet, Sachs, and Lee (1997).

suggestive. The new variables enter with highly significant coefficients, indicating that they belong in the income equation. Moreover, once the additional geography variables are included, (1) the IV coefficient estimates on trade become statistically insignificant (with *t*-statistics around 0.4 or below), and (2) the IV point estimates on trade are reduced below their OLS counterparts. These findings are consistent with the hypothesis that nontrade effects of geography are the main driving force behind the findings of Frankel and Romer.⁵⁴

8. Other Recent Work

Before we close, we mention briefly some other recent papers that have examined the connection between openness and economic growth. We focus on three papers in particular: Lee (1993), Harrison (1996), and Wacziarg (1998). These papers are of interest because they contain some methodological innovations.

Lee (1993) reasons, on the basis of an analytical model, that the distortionary effects of trade restrictions should be larger in economies that, in the absence of trade restrictions, would be more exposed to trade. Hence he interacts an indicator of trade policy with a measure of what he calls “free trade openness” (FREEOP).⁵⁵ The latter is constructed by regressing observed import shares on land area, distance from major trading partners, import tariffs, and black-market premia, and then calculating the predicted value of imports when the actual values of tariffs black-market premia are replaced by zeros. He finds that this composite measure (FREETAR) enters a growth regression with an estimated coefficient that is negative and statistically significant.

Lee uses two indicators of trade policy: an import-weighted tariff average and the black-market premium. We have discussed above the shortcomings of the latter as a measure of trade policy (when reviewing Sachs and Warner, 1995). The problem with Lee’s tariff variable, as Lee (1993, p. 320) acknowledges, is that the underlying tariff data are from “various years in the 1980s”—the tail end of the 1960–1985 period over which his growth regressions are run. This raises the possibility of reverse causation: countries that perform well tend to liberalize their trade regime eventually. To check for this possibility, we have repeated Lee’s regression, using the same specification and tariff variable, but over the subse-

54. We have carried out this exercise for various other samples [e.g., the higher-quality 98-country sample used by Frankel and Romer (1999), and samples excluding possible outliers such as Luxembourg and Hong Kong] and reach identical conclusions.

55. Specifically, the composite measure is constructed as $FREETAR = FREEOP \log(1 + \text{tariff})$.

quent time period 1980–1994.⁵⁶ While the estimated coefficient on FREETAR is negative for this later period, it is nowhere near significant (t -statistic -0.80).

Harrison's (1996) main methodological contribution is to examine the relationship between trade policy and growth in a *panel* setting, using fixed effects for countries. This approach has the advantage that it enables the analyst to look for evidence of the effects of trade liberalization *within* countries.⁵⁷ But it has the disadvantage that the available time series are necessarily short, requiring the use of annual data or (at most) five-year averages. It may be a lot to ask of such data to reveal much about the relationship between trade policy and growth, because of the likely lags involved and the contamination from business-cycle effects.⁵⁸

Harrison uses seven indicators of trade policy, and finds that three of these "exhibit a robust relationship with GDP growth" (1996, p. 443). These three are the following: (1) the black-market premium, (2) a measure based on the price level of a country's tradables (relative to international prices), and (3) a subjective measure of trade liberalization constructed at the World Bank. We have already discussed at length the problems involved in interpreting measures of each of these types as indicators of trade policy.

Finally, the paper by Wacziarg (1998) is an ambitious attempt to uncover the channels through which openness affects economic growth. Wacziarg's index of trade policy is a linear combination of three indicators: (1) the average import duty rate, (2) the NTB coverage ratio, and (3) the SW indicator.⁵⁹ The weights used to construct the combined index come from a regression of trade volumes (as a share of GDP) on these three indicators plus some other determinants. Using a panel made up of five-year averages for 57 countries during 1970–1989, Wacziarg finds that investment is the most important channel through which openness increases growth, accounting for more than 60% of the total effect.

We have two worries about this paper. First, we are not sure that the

56. Since Summers–Heston data are not available for the 1990s, we used World Bank data on GDP per capita (at constant prices).

57. Harrison (1996) cites disappointing results with cross-section regressions as a motivation for going the panel route.

58. Indeed, when Harrison (1996) controls for some business-cycle conditions, about half of her significant coefficients (on openness-related variables) disappear. The empirical evidence on the *short-run* relationship between trade liberalization and economic growth is judiciously reviewed in Greenaway, Morgan, and Wright (1998), who point to both positive and negative findings. These authors attempt to trace out the dynamics of the output response using three different indicators of policy (including the SW index), and report finding a J-curve effect: output first falls and then increases.

59. More specifically, Wacziarg uses the timing of trade liberalization in Sachs and Warner (1995) to assign a value to each country for any given five-year period.

regularities revealed by the data over time horizons of five years or less are particularly informative about the relationship between trade policy and long-run economic performance. It would be interesting to see if the results hold up with averages constructed over a decade or more. Second, as discussed previously, we are skeptical that the SW measure, on which the Wacziarg indicator is partly based, is a meaningful indicator of trade policy. Wacziarg remarks in a footnote (1998, footnote 9) that the “exclusion of [the SW indicator] from the trade policy index reduced the precision of the estimates . . . but did not change the qualitative nature of the results.” We would have preferred to see estimates based only on tariff and NTB indicators.

9. Concluding Remarks

We have scrutinized in this paper the most prominent recent empirical studies on the relationship between trade barriers and economic growth. While we do not pretend to have undertaken an exhaustive survey, we believe that the weaknesses we have identified are endemic to this literature.

We emphasize that our difficulty with this literature is *not* a variant of the standard robustness criticism often leveled at cross-country growth empirics. Going back at least to Levine and Renelt (1992), a number of authors have pointed to the sensitivity of growth regressions to changes in the list of controls, and to the failure of these coefficients to pass the test of “extreme-bounds analysis.” Whatever position one takes on this debate, the general point that we wish to make about the empirical literature on openness and growth is much simpler. For the most part, the strong results in this literature arise either from obvious misspecification or from the use of measures of openness that are proxies for other policy or institutional variables that have an independent detrimental effect on growth. When we do point to the fragility of the coefficients, it is to make the point that the coefficients on the openness indicators are particularly sensitive to controls for these other policy and institutional variables. To the extent that these objections can be conceptualized as variants of the robustness criticism, it is robustness at a much more basic level than that typically discussed in the Bayesian literature.

Still, in view of the voluminous research on the subject, a natural question that arises is whether we shouldn’t take comfort from the fact that so many authors, using varying methods, have all arrived at the same conclusion. Don’t we learn something from the cumulative evidence, even if individual papers have shortcomings?

We take a different message from this large literature. Had the nega-

tive relationship between trade restrictions and economic growth been convincingly demonstrated, we doubt that this issue would continue to generate so much empirical research. We interpret the persistent interest in this area as reflecting the worry that the existing approaches haven't gotten it quite right. One indication of this is that the newer papers are habitually motivated by exegeses on the methodological shortcomings of prior work.

We are especially struck and puzzled by the proliferation of indexes of trade restrictions. It is common to assert in this literature that simple trade-weighted tariff averages or nontariff coverage ratios—which we believe to be the most direct indicators of trade restrictions—are misleading as indicators of the stance of trade policy. Yet we know of no papers that document the existence of serious biases in these direct indicators, much less establish that an alternative indicator performs better (in the relevant sense of calibrating the restrictiveness of trade regimes).⁶⁰ An examination of simple averages of taxes on imports and exports and NTB coverage ratios leaves us with the impression that these measures in fact do a decent job of rank-ordering countries according to the restrictiveness of their trade regimes. In the working-paper version of this paper, we provide a simple measure of import duties for a large sample of countries and three different periods, so that the reader can form his/her judgement on this (Rodríguez and Rodrik, 1999, Table VIII.1).⁶¹

As we mentioned in the introduction, we are skeptical that there is a strong negative relationship in the data between trade barriers and economic growth, at least for levels of trade restrictions observed in practice.⁶² We view the search for such a relationship as futile. We think there are two other fruitful avenues for future research.

60. Pritchett (1996) comes closest. The point of his paper, however, is to document the weak correlation between commonly used indicators of trade restrictions, and not to argue for the superiority of one indicator over the others.

61. This is the measure of import tariffs we used in Figure 1 (top panel).

62. In his comment on this paper, Chad Jones acknowledges the fragility of many of the results in the literature, but reports a range of exercises that leads him to conclude, as a best estimate, that trade restrictions are harmful to long-run incomes and that the effects are potentially large. We caution the reader about regressions where the *level* of per capita income is regressed on measures of trade restrictions. It is well known that countries reduce their trade barriers as they get richer, so levels regressions are subject to problems of reverse causality. It is difficult to overcome this problem via instrumentation, since adequate instruments (exogenous variables that are correlated with trade restrictions, but are otherwise uncorrelated with incomes) are particularly difficult to find in this context (as our discussion in Section 7 highlights). When regressions are run in growth form, we find that none of the available continuous measures of trade restrictions (simple tariff averages or nontariff coverage ratios) enter significantly in the vast majority of reasonable specifications. Some dichotomous measures based on the continuous variables do somewhat "better," but only if the break point is set at a sufficiently high level (e.g., a tariff rate or nontariff coverage ratio in excess of 40%).

First, in cross-national work, it might be productive to look for *contingent* relationships between trade policy and growth. Do trade restrictions operate differently in low- vs. high-income countries? In small vs. large countries? In countries with a comparative advantage in primary products vs. those with comparative advantage in manufactured goods? In periods of rapid expansion of world trade vs. periods of stagnant trade? Further, it would help to disaggregate policies and to distinguish the possibly dissimilar effects of different types of trade policies (or of combinations thereof). Are tariff and nontariff barriers to imports of capital goods more harmful to growth than other types of trade restrictions? Does the provision of duty-free access to imported inputs for exporters stimulate growth? Are export-processing zones good for growth? Does the variation in tariff rates (or NTBs) across sectors matter? The cross-national work has yet to provide answers to such questions.

Second, we think there is much to be learned from microeconomic analysis of plant-level datasets. These datasets constitute a rich source for uncovering the ways in which trade policy influences the production, employment, and technological performance of firms (see Roberts and Tybout, 1996). Recent research by Bernard and Jensen (1995, 1998), Aw, Chung, and Roberts (1998), and Clerides, Lach and Tybout (2000) has already shed new light on the relationship between trade and firm performance. For example, these papers (based on the experiences of countries as diverse as the United States, Taiwan, and Mexico) find little evidence that firms derive technological or other benefits from exporting per se; the more common pattern is that efficient producers tend to self-select into export markets. In other words, causality seems to go from productivity to exports, not vice versa. Relating these analyses to trade *policies* is the obvious next step in this line of research.

Let us close by restating our objective in this paper. We do not want to leave the reader with the impression that we think trade protection is good for economic growth. We know of no credible evidence—at least for the post-1945 period—that suggests that trade restrictions are systematically associated with higher growth rates. What we would like the reader to take away from this paper is some caution and humility in interpreting the existing cross-national evidence on the relationship between trade policy and economic growth.

The tendency to greatly overstate the systematic evidence in favor of trade openness has had a substantial influence on policy around the world. Our concern is that the priority afforded to trade policy has generated expectations that are unlikely to be met, and it may have crowded out other institutional reforms with potentially greater payoffs. In the real world, where administrative capacity and political capital are

scarce, having a clear sense of policy priorities is of utmost importance. The effects of trade liberalization may be on balance beneficial on standard comparative-advantage grounds; the evidence provides no strong reason to dispute this. What we dispute is the view, increasingly common, that integration into the world economy is such a potent force for economic growth that it can effectively substitute for a development strategy.

Data Appendix

SECTION 1

1. Import duties as a percentage of imports. Source: World Bank (1998).
2. Nontariff barriers. Source: Barro and Lee (1994).

SECTION 3

3. bmpav: average black-market premium. Source: Sachs and Warner (1995).
4. rcoast: coastal length over total land area. Source: Radelet, Sachs, and Lee (1997).
5. tropics: dummy for tropical countries. Source: Radelet, Sachs, and Lee (1997).
6. Latin America: dummy for countries in Latin America and the Caribbean.
7. SSA: dummy for countries in sub-Saharan Africa.
8. East Asia: dummy for countries in East Asia.
9. TAR: own-import-weighted ratio of tariff revenues to trade. Source: Barro and Lee (1994).
10. NTB: own-import-weighted nontariff frequency on capital goods and intermediates. Source: Barro and Lee (1994).
11. DISTORTION: ratio of consumption price level to U.S. price level, measured in identical currencies, divided by the fitted value of a regression on GDP, GDP squared, year dummies, and continent dummies. Source: Dollar (1992).
12. VARIABILITY: Coefficient of variation of DISTORTION. Source: Dollar (1992).
13. Investment/GDP: Source: Summers and Heston (1988) for Table 2.
14. Log initial income: Source: Summers and Heston (1988) for Table 2.
15. Schooling, 1975: Barro and Lee (1994).

SECTION 4

16. BMP: Dummy variable equal to 1 if black-market premium exceeds 20% during either the 1970s or the 1980s. Source: Sachs and Warner (1995).
17. BMP70, BMP80: Black-market premium during (respectively) 1970s and 1980s. Source: Sachs and Warner (1995).
18. MON: Dummy variable equal to 1 if the country had a score of 4 (highest score) on the Export Marketing Index in World Bank (1994). Source: Sachs and Warner (1995).
19. SOC: Dummy variable equal to 1 if the country was classified as socialist in Kornai (1992). Source: Sachs and Warner (1995).
20. TAR: own-import-weighted ratio of tariff revenues to trade. Source: Barro and Lee (1994).
21. NTB: own-import-weighted nontariff frequency on capital goods and intermediates. Source: Barro and Lee (1994).
22. OPEN: Variable equal to 0 if the country had $BMP = 1$, $MON = 1$, $SOC = 1$, $TAR > 0.4$, or $NTB > 0.4$. Source: Sachs and Warner (1995).
23. BM, SQT, QT, etc.: Openness indexes constructed using subsets of the Sachs and Warner's information. The label for each index denotes the openness indicators used to construct that index: M = state monopoly of main export, S = socialist economic system, Q = nontariff barriers, T = tariffs, B = black-market premium. For example, SMQT is set to 0 if it is closed according to either of the criteria for S, M, Q, or T, and to 1 otherwise.
24. Inflation, 1975–1990. Source: World Bank (1998).
25. Debt/exports, 1985. Source: World Bank (1998).
26. Change in terms of trade. Source: Barro and Lee (1994).
27. War: dummy for countries that participated in at least one external war over the period 1960–1985. Source: Barro and Lee (1994).
28. Quality of institutions: Institutional quality index from Knack and Keefer (1995).
29. Government budget surplus, 1970–1990. Source: World Bank (1998).
30. Population growth. Source: World Bank (1998).

SECTION 5

31. Sachs–Warner: Same as OPEN in Section 4.
32. World Development Report: World Development Report outward-orientation index, 1973–1985. Source: Edwards (1998).

33. Leamer: Openness index estimated by Leamer (1988) using residuals from disaggregated trade-flow regressions. Source: Edwards (1998).
34. Black-market premium: same as BMP80 in Section 4.
35. Tariffs: Same as TAR in Section 4.
36. Quotas: Same as NTB in Section 4.
37. Heritage Foundation: Subjective index of the extent to which government policies distort trade, from Johnson and Sheehy (1996). Source: Edwards (1998).
38. Collected-trade-taxes ratio: Average for 1980–1985 of ratio of total revenues on international trade taxes to total trade. Source: Edwards (1998).
39. Wolf's index of import distortions: A regression-based index from Wolf (1993). Source: Edwards (1998).
40. Principal-components factor: First principal component of OPEN, black-market premium, tariffs, quotas, and Wolf's index. The equation used to calculate it is

$$COM = -0.469 \times OPEN + 0.320 \times BLACK + 0.494 \times TARIFF + 0.553 \times QR + 0.354 \times WOLF.$$
41. Log of GDP per capita, 1985. From Summers and Heston (1994). Source: Edwards (1998).
42. Property rights: Heritage Foundation index of property-rights protection, from Johnson and Sheehy (1996). Source: Edwards (1998).
43. Average import and export duties (World Bank): From World Bank (1998). Average duty is calculated as $(1 + \text{export duty}) \times (1 + \text{import duty}) - 1$.
44. Merged duty index: Simple average of average duty (43) and (38).
45. Trade distortion index based on Lee's data. Analog of Heritage index using data from Lee (1993) in Barro and Lee (1994). Countries are rated on a score of 1 to 5 according to the maximum of tariff rate and nontariff-barrier coverage ratio: higher than 20%: "very high" (a rating of 5); between 15 and 20%: "high" (4); between 10% and 15%: "moderate" (3); between 5% and 10%: "low" (2); and between 0 and 5%: "very low" (1).

SECTION 6

46. Contributions to variance around EC mean, from Summers and Heston (1994).
47. GDP per capita (Figure 3): Maddison (1982). Source: Ben-David (1993).
48. GDP per capita (Figure 4, Table 7): Maddison (1995).

49. GDP per capita (Figures 5–7): Summers and Heston (1994).
 50. Ratio of import duties to imports, United States, from Bureau of the Census (1989, Series U211).

REFERENCES

- Acheson, A. L. K. and D. McFetridge. (1996). Intellectual property and endogenous growth. In *The Implications of Knowledge-Based Growth for Microeconomic Policies*. P. Howitt (ed.). Calgary: University of Calgary Press.
- Aw, B. Y., S. Chung, and M. J. Roberts. (1998). Productivity and turnover in the export market: Micro evidence from Taiwan and South Korea. Pennsylvania State University.
- Balassa, B. (1971). *The Structure of Protection in Developing Countries*. Baltimore: Johns Hopkins University Press, 1971.
- . (1985). Exports, policy choices, and economic growth in developing countries after the 1973 oil shock. *Journal of Development Economics* 18(2): 23–35.
- Barro, R. J. (1997). *Determinants of Economic Growth: A Cross-Country Empirical Study*. Cambridge, MA: The MIT Press.
- , and J.-W. Lee. (1993). International comparisons of educational attainment. *Journal of Monetary Economics* 32(3):363–394.
- , and ———. (1994). Data set for a panel of 138 countries. Harvard University.
- Ben-David, D. (1993). Equalizing exchange: Trade liberalization and income convergence. *Quarterly Journal of Economics* 108(3).
- . (1999). Teach your children well: Planting the seeds of education and harvesting the benefits of trade. Tel Aviv University.
- Bernard, A. B., and J. B. Jensen. (1995). Exporters, jobs, and wages in U.S. manufacturing, 1975–1987. In *Brookings Papers on Economic Activity: Microeconomics 1995*, pp. 67–112.
- , and ———. (1998). Exporting and productivity. Presented at 1998 Summer Institute, NBER, Cambridge, MA.
- Bhagwati, J. N. (1982). Directly unproductive, rent-seeking (DUP) Activities. *Journal of Political Economy* 90(5):988–1002.
- Bhalla, S., and L. J. Lau. (1992). Openness, technological progress, and economic growth in developing countries. World Bank.
- Buffie, E. (1998). *Trade Policy in Developing Countries*. Unpublished manuscript.
- Bureau of the Census. (1989). *Historical Statistics of the United States, Colonial Times to 1970*. White Plains, NY: Kraus International Publications.
- Burnside, C., and D. Dollar. (1997). Aid, policies, and growth. World Bank. Policy Research Working Paper 1777.
- Caves, R. E., J. A. Frankel, and R. W. Jones. (1999). *World Trade and Payments: An Introduction*. 8th ed. Reading, MA: Addison-Wesley, 1999.
- Clague, C., P. Knack, S. Keefer, and M. Olson. (1996). Property and contract rights in autocracies and democracies. *Journal of Economic Growth* 1(2):243–276.
- Clerides, S., S. Lach, and J. Tybout. (2000). Is “Learning-by-exporting” important? Micro-dynamic evidence from Colombia, Mexico, and Morocco. *Quarterly Journal of Economics*, forthcoming.
- Cobban, A. (1965). *A History of Modern France*. New York: Braziller.
- Collins, S. M. and B. P. Bosworth. (1996). Economic growth in East Asia:

- Accumulation versus assimilation. *Brookings Papers on Economic Activity* 1996 (2):135–191.
- Díaz-Alejandro, C. (1981). Southern Cone stabilization plans. In *Economic Stabilization in Developing Countries*, W. Cline and S. Weintraub (eds.). Washington: The Brookings Institution.
- Dollar, D. (1992). Outward-oriented developing economies really do grow more rapidly: Evidence from 95 LDCs, 1976–85. *Economic Development and Cultural Change* 1992:523–544.
- Edwards, S. (1992). Trade orientation, distortions, and growth in developing countries. *Journal of Development Economics* 39(1):31–57.
- . (1993). Openness, trade liberalization, and growth in developing countries. *Journal of Economic Literature* XXXI(3):1358–1393.
- . (1998). Openness, productivity and growth: What do we really know? *Economic Journal* 108(March):383–398.
- Esfahani, H. S. (1991). Exports, imports, and growth in semi-industrialized countries. *Journal of Development Economics* 35(1):93–116.
- Falvey, R., and N. Gemell. (1999). Factor endowments, nontradables prices and measures of “openness.” *Journal of Development Economics* 58(February):101–122.
- Feder, G. (1983). On exports and economic growth. *Journal of Development Economics* 12(1/2):59–73.
- Feenstra, R. (1990). Trade and uneven growth. Cambridge, MA: National Bureau of Economic Research. NBER Working Paper 3276.
- Fischer, S. (2000). Lunch address given at the conference on “Promoting Dialogue: Global Challenges and Global Institutions,” Washington: American University.
- Frankel, J., and D. Romer. (1999). Does trade cause growth? *American Economic Review* 89(3):379–399.
- Gelb, A. H., and Associates. (1988). *Oil Windfalls: Blessing or Curse?* New York: Oxford University Press.
- Gerschenkron, A. (1943). *Bread and Democracy in Germany*, Berkeley and Los Angeles: University of California Press.
- Greenaway, D., W. Morgan, and P. Wright. (1998). Trade reform, adjustment and growth: What does the evidence tell us? *The Economic Journal* 108 (September):1547–1561.
- Grossman, G., and E. Helpman. (1991). *Innovation and Growth in the Global Economy*. Cambridge, MA: The MIT Press.
- Grossman, H., and M. Kim. (1996). Predation and accumulation. *Journal of Economic Growth* 1(3):333–350.
- Gulhati, R. and R. Nallari. (1990). *Successful Stabilization and Recovery in Mauritius*. Washington: The World Bank.
- Haber, S. (1997). *How Latin America Fell Behind: Essays on the Economic Histories of Brazil and Mexico, 1800–1914*. Stanford, CA: Stanford University Press.
- Hall, R. E., and C. I. Jones. (1998). Why do some countries produce so much more output per worker than others? Cambridge, MA: National Bureau of Economic Research. NBER Working Paper W6564.
- Harrison, A. (1996). Openness and growth: A time-series, cross-country analysis for developing countries. *Journal of Development Economics* 48:419–447.
- , and G. Hanson. (1999). Who gains from trade reform? Some remaining puzzles. *Journal of Development Economics* 50:125–154.

- Husted, S., and M. Melvin. (1997). *International Economics*, 4th ed. Reading, MA: Addison-Wesley.
- IMF (1997). *World Economic Outlook*, Washington.
- International Currency Analysis. (1995). *World Currency Yearbook*. Brooklyn, NY: International Currency Analysis, Inc.
- Jodha, N. (1996). Property rights and development. In *Rights to nature: Ecological, economic, cultural and political principles of institutions for the environment*, S. Hanna (ed.). Washington: Island Press.
- Johnson, B. T., and T. P. Sheehy. (1996). *1996 Index of Economic Freedom*. Washington: The Heritage Foundation.
- Knack, S., and P. Keefer. (1995). Institutions and economic performance: Cross-country tests using alternative institutional measures. *Economics & Politics*, November, pp. 207–228.
- Kornai, J. (1992). *The Socialist System: The Political Economy of Communism*. Princeton, NJ: Princeton University Press.
- Krueger, A. O. (1974). The political economy of the rent-seeking society. *American Economic Review* 64(June):291–303.
- . (1998). Why trade liberalisation is good for growth. *The Economic Journal* 108(September):1513–1522.
- Krugman, P. R., and M. Obstfeld. (1997). *International Economics: Theory and Policy*, 4th ed. Reading, MA: Addison-Wesley, 1997.
- Leamer, E. (1988). Measures of openness. In *Trade Policy and Empirical Analysis*, R. Baldwin (ed.). Chicago: University of Chicago Press.
- Lee, J.-W. (1993). International trade, distortions, and long-run economic growth. *International Monetary Fund Staff Papers* 40(2):299–328.
- Lerner, A. P. (1936). The symmetry between import and export taxes. *Economica* 11:306–313.
- Levine, R., and D. Renelt. (1992). A sensitivity analysis of cross-country growth regressions. *American Economic Review* 82(4):942–963.
- Little, I., T. Scitovsky, and M. Scott. (1970). *Industry and Trade in Some Developing Countries*. London and New York: Oxford University Press.
- Maddison, A. (1982). *Phases of Capitalist Development*. Oxford and New York: Oxford University Press.
- . (1995). *Monitoring the World Economy: 1820–1992*. Development Centre Studies. Paris and Washington: Organisation for Economic Co-operation and Development, 1995.
- Matsuyama, K. (1992). Agricultural productivity, comparative advantage, and economic growth. *Journal of Economic Theory* 58(2):317–334.
- OECD. (1998). *Open Markets Matter: The Benefits of Trade and Investment Liberalisation*. Paris: OECD, 1998.
- O'Rourke, K. H. (1997). Tariffs and growth in the late nineteenth century. Center for Economic Policy Research. Discussion Paper 1700.
- Park, W. and J. C. Ginarte. (1997). Intellectual property rights and economic growth. *Contemporary Economic Policy* 15(3):51–61.
- Pertamina. (1998). <http://www.pertamina.co.id/ptmhp.htm>.
- Pollard, S. (1974). *European Economic Integration, 1815–1970*. London: Thames and Hudson.
- Pritchett, L. (1996). Measuring outward orientation: Can it be done? *Journal of Development Economics* 49(2).

- Radelet, S., J. D. Sachs, and J.-W. Lee. (1997). Economic growth in Asia. Harvard Institute for International Development. Development Discussion Paper.
- Richardson, P., et al. (1997). Globalization and linkages: Macro-structural challenges and opportunities. OECD Economic Surveys 28.
- Roberts, M. J., and J. R. Tybout, eds. (1996). *Industrial Evolution in Developing Countries*. New York: Oxford University Press for the World Bank.
- Rodríguez, F., and D. Rodrik. (1999). Trade policy and economic growth: A skeptic's guide to the cross-national literature. Cambridge, MA: National Bureau of Economic Research. NBER Working Paper 7081.
- Rodrik, D. (1999). *The New Global Economy and Developing Countries: Making Openness Work*. Washington: Overseas Development Council.
- Rogoff, K. (1996). The purchasing power parity puzzle. *Journal of Economic Literature* XXXIV (June):647–668.
- Rogowski, R. (1989). *Commerce and Coalitions: How Trade Affects Domestic Political Alignments*. Princeton: Princeton University Press.
- Sachs, J., and A. Warner. (1995). Economic reform and the process of global integration. *Brookings Papers on Economic Activity* 1995(1):1–118.
- Sala-i-Martin, X. (1997). I just ran two million regressions. *American Economic Review* 87(2):178–183.
- Slaughter, M. J. (2000). Trade liberalization and per capita income convergence: A difference-in-differences analysis. *Journal of International Economics*, forthcoming.
- Srinivasan, T. N. (1997). As the century turns: Analytics, empirics, and politics of development. Economic Growth Center, Yale University.
- Stiglitz, J. E. (1998). Towards a new paradigm for development: Strategies, policies, and processes. 1998 Prebisch Lecture, UNCTAD, Geneva.
- Summers, R., and A. Heston. (1988). A new set of international comparisons of real product and price levels: Estimates for 130 countries, 1950–1985. *Review of Income and Wealth* 34(March):1–25.
- , and ———. (1994). *Penn World Tables Mark 5.6*. <http://www.nber.org>.
- Thompson, M., and F. Rushing. (1996). An empirical analysis of the impact of patent protection on economic growth. *Journal of Economic Development* 21(2): 61–79.
- Tornell, A. (1997). Economic growth and decline with endogenous property rights. *Journal of Economic Growth* 2(3):219–250.
- Vamvakidis, A. (1996). Regional integration and economic growth. *The World Bank Economic Review* 12(2):251–270.
- Wacziarg, R. (1998). Measuring the dynamic gains from trade. *World Bank Working Paper* 2001.
- White, H. L. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica* 48:817–838.
- Wolf, H. (1993). Trade orientation: Measurement and consequences. *Estudios de Economía* 20(20):52–72.
- World Bank. (1983). *Mauritius Economic Memorandum: Recent Developments and Prospects*. Washington: The World Bank.
- . (1989). *Mauritius: Managing Success*. Washington: The World Bank.
- . (1994). *Adjustment in Africa: Reforms, Results, and the Road Ahead*. Washington: The World Bank.
- . (1996). *Global Economic Prospects and the Developing Countries*. Washington: The World Bank.

-
- . (1998). *World Development Indicators CD-ROM*. Washington: The World Bank.
- Yarbrough, B. M., and R. M. Yarbrough. (2000). *The World Economy: Trade and Finance*, 5th ed. Fort Worth, TX: Harcourt College Publishers.

Comment

CHANG-TAI HSIEH
Princeton University

Francisco Rodríguez and Dani Rodrik argue that the conventional wisdom among multilateral institutions in Washington (and many economists) that lower trade barriers results in significantly faster growth is based on weak empirical evidence. Their main point is that the empirical evidence that purportedly shows a negative correlation between trade barriers and growth typically relies on measures that are either measures of macroeconomic imbalances or bad institutions, and are not actually measures of trade barriers. For example, they argue that a widely used measure of trade restrictions—deviation of domestic prices of tradables from world prices—reflects deviations from PPP due to overvalued exchange rates, and is not a measure of trade barriers. To take another example, Rodríguez and Rodrik argue that the widely used Sachs–Warner openness index is largely a dummy variable for sub-Saharan Africa and countries with large macroeconomic imbalances, which again is not a measure of trade barriers.

However, the fact that trade barriers are not robustly correlated with growth once controls for macroeconomic imbalances and bad institutions are introduced does not imply that trade barriers do not have deleterious effects of growth. There is a fundamental identification problem in separating the effects of trade restrictions from those of macroeconomic imbalances and bad institutions, since countries with bad macroeconomic policies and weak institutions also have severe trade restrictions. And when countries liberalize their trade regimes, it typically takes place along with a macroeconomic stabilization program. Therefore, there may not be enough cross-country variation in trade restrictions orthogonal to macroeconomic imbalances to identify the effect of trade on growth, even if trade restrictions do have significant negative effects on growth. For example, even if the Sachs–Warner index is a dummy for sub-Saharan Africa, it is still the case that most sub-Saharan countries have in fact imposed significant trade restrictions.

Nonetheless, I find their main point—that there is a large standard error around precisely how much trade barriers matter for growth—

-
- . (1998). *World Development Indicators CD-ROM*. Washington: The World Bank.
- Yarbrough, B. M., and R. M. Yarbrough. (2000). *The World Economy: Trade and Finance*, 5th ed. Fort Worth, TX: Harcourt College Publishers.

Comment

CHANG-TAI HSIEH
Princeton University

Francisco Rodríguez and Dani Rodrik argue that the conventional wisdom among multilateral institutions in Washington (and many economists) that lower trade barriers results in significantly faster growth is based on weak empirical evidence. Their main point is that the empirical evidence that purportedly shows a negative correlation between trade barriers and growth typically relies on measures that are either measures of macroeconomic imbalances or bad institutions, and are not actually measures of trade barriers. For example, they argue that a widely used measure of trade restrictions—deviation of domestic prices of tradables from world prices—reflects deviations from PPP due to overvalued exchange rates, and is not a measure of trade barriers. To take another example, Rodríguez and Rodrik argue that the widely used Sachs–Warner openness index is largely a dummy variable for sub-Saharan Africa and countries with large macroeconomic imbalances, which again is not a measure of trade barriers.

However, the fact that trade barriers are not robustly correlated with growth once controls for macroeconomic imbalances and bad institutions are introduced does not imply that trade barriers do not have deleterious effects of growth. There is a fundamental identification problem in separating the effects of trade restrictions from those of macroeconomic imbalances and bad institutions, since countries with bad macroeconomic policies and weak institutions also have severe trade restrictions. And when countries liberalize their trade regimes, it typically takes place along with a macroeconomic stabilization program. Therefore, there may not be enough cross-country variation in trade restrictions orthogonal to macroeconomic imbalances to identify the effect of trade on growth, even if trade restrictions do have significant negative effects on growth. For example, even if the Sachs–Warner index is a dummy for sub-Saharan Africa, it is still the case that most sub-Saharan countries have in fact imposed significant trade restrictions.

Nonetheless, I find their main point—that there is a large standard error around precisely how much trade barriers matter for growth—

largely convincing. But we shouldn't find the results surprising. Starting from Levine and Renelt (1992), there is overwhelming evidence that very little—not even factors such as increases in human capital that *a priori* would seem to be important in explaining growth—is robust in the empirical growth literature. Therefore, there is no reason to expect measures of trade barriers to be robustly correlated with growth, even if we were to obtain accurate measures of trade restrictions. Furthermore, given the diversity of countries around the world and the different forms which trade barriers take, it is silly to think that one can find a consistent cross-country relationship between trade restrictions and growth. First of all, trade barriers take many different forms. We do not expect there to be significant deleterious growth effects from a well-administered uniform 20–30% tariff. In contrast, a country in which trade barriers are set in a discretionary manner with rampant rent seeking will probably have poor growth performance. Second of all, countries are very different. Small countries probably benefit more from trade than large countries. Countries that are more specialized benefit more from trade than countries that are already well diversified. Finally, we know that trade barriers introduce distortions, but so does every form of government intervention, and there is no reason to believe that the costs of trade distortions are significantly different from the costs of other government interventions. So there is a sense in which the empirical studies that attempt to find a robust cross-country correlation between trade restrictions and growth are as sensible as a cross-country regression of growth on, say, sales taxes or income taxes.

One way to make progress in understanding how trade restrictions affect growth is to differentiate between the effects of different types of trade barriers. Here, the empirical growth literature gives us some guidance on what to look for. Specifically, starting from De Long and Summers (1991), many authors have found that investment in machinery and equipment is the only variable (other than a dummy for sub-Saharan Africa) that is robustly correlated with growth. This is sensible. After all, countries that have grown rapidly are ones that have invested resources in using the machines that embody the technologies of the industrial revolution. Trade policy—specifically, restrictions on imports of capital goods—can affect machinery and equipment investment by increasing the price of imported machinery and equipment. Restrictions on capital-good imports are even more harmful in a developing country that has little domestic production of capital goods and would thus benefit the most from purchasing capital goods embodying the most advanced technologies.

As far as I am aware, there is no study that specifically studies the growth effects of restrictions on capital-good imports. There is, however,

suggestive evidence that such restrictions have important effects. It is well known, for example, that Taiwan and Korea have had highly distorted trade regimes, but which nonetheless always kept domestic prices of capital goods close to world prices. Consequently, despite their distorted trade regimes, the share of imports and investment in machinery and investment in these two countries were among the highest in the world. Another piece of evidence is from Charles Jones's (1994) intriguing paper—the dual of De Long and Summers's paper—that shows that the relative price of capital differs enormously between rich and poor countries (by a factor of four). Further, this relative price is significantly correlated with growth even after controlling for initial income, so we know the correlation is not driven by reverse causation due to a Balassa–Samuelson effect. Clearly, the way in which the relative price of capital affects growth is by lowering the amount of capital equipment; if a country increases the relative price of capital and thus of growth, there is going to be less of both.

The main problem is that we do not know whether the large differences in relative price of capital (orthogonal to GDP/worker) are due to differences between trade barriers for capital goods and barriers for consumption goods, or due to domestic distortions that affect the relative price of all capital goods. Clearly, if capital goods are mostly imports, the distinction is moot. However, a simple way to test this is to see whether the share of imports in total machinery and equipment investment is correlated with the relative price of capital. The idea is that if differences in the relative price of capital are due entirely to domestic distortions, then they should affect the aggregate quantity of investment in machinery and equipment but should have no effect on the *composition* of investment between imports and domestically produced capital goods. Table 1 shows that, controlling for initial income and the manufacturing share of GDP, a doubling in the relative price of capital (about the difference between Korea and India) lowers the import share of investment by almost 6 percentage points in the full sample. The effect is even stronger in developing countries, where a similar increase in the relative price of capital lowers the import share by almost 10 percentage points.¹ It would obviously be better to get direct measures of restrictions on imports of capital goods. In addition, the sample, particularly that for the non-OECD countries, is small, since we are restricted to the countries that

1. The data on initial income and manufacturing share are from the *Penn World Tables* (Mark 5.6). The relative price of capital is from Charles Jones's Web site (<http://www.stanford.edu/~chadj/RelPrice.asc>), and imports of machinery and equipment relative to total investment in machinery and equipment were graciously provided by Lee Jong-Wha (who compiled them from the OECD's trade-statistics datatapes).

Table 1 CROSS-COUNTRY DIFFERENCES IN
RELATIVE PRICE OF CAPITAL AFFECT
IMPORTS OF CAPITAL GOODS

<i>Variable</i>	<i>Full sample</i>	<i>Non-OECD countries</i>
log(relative price of capital)	-0.0574 (0.0359)	-0.0986 (0.0453)
Manufacturing share of GDP	0.3881 (0.2639)	0.7954 (0.3049)
log(initial income per capita)	-0.1040 (0.1085)	-0.0738 (0.1685)
N	52	35
SEE	0.1498	0.1490
R ²	0.12	0.26

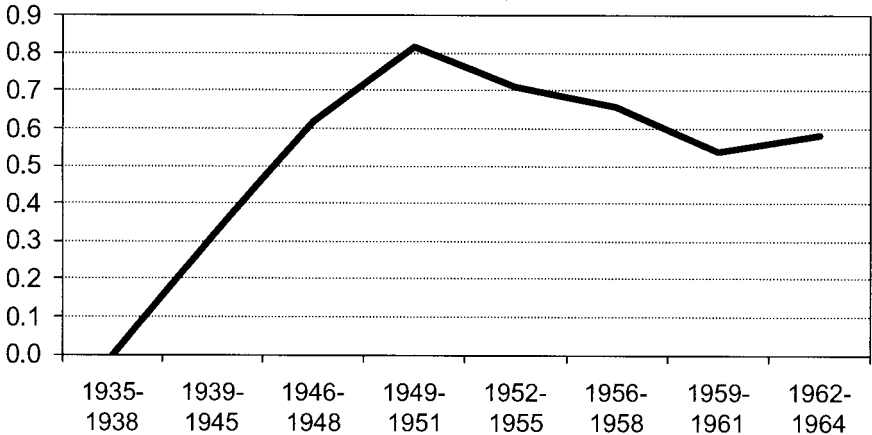
Dependent variable is imports of capital goods from OECD countries/
total investment in machinery and equipment.

have participated in the benchmark surveys of the United Nations International Comparisons Project. Nonetheless, these results provide suggestive evidence that part of the cross-country difference in the relative price of capital is due to trade barriers.

One can also turn to narrative histories of particular countries for evidence of the impact of capital-good restrictions. For example, I have always found Carlos Diaz-Alejandro's (1970) story of Argentina's economic decline particularly compelling and disturbing. Starting with the Great Depression, Argentina sought to redistribute wealth from rural landowners and exporting elites to the urban working class by making imports of consumer goods freely available but severely restricting imports of capital goods. This policy of redistribution doubled the relative price of capital in Argentina from the late 1930s to the late 1940s (see Figure 1), which led to anemic rates of investment in machinery and equipment in Argentina since the end of World War II. Consequently, a country that was among the wealthiest nations in the world in the early twentieth century is now decidedly a Third World nation.

In the other direction, the experience of India in the 1990s provides evidence that the removal of restrictions on capital-good imports can have significant positive effects on growth. Specifically, India liberalized imports of capital goods in the early 1990s without lowering barriers on imports of consumer goods (which it has done only recently). Due to the

Figure 1 CHANGE IN RELATIVE PRICE OF MACHINERY AND EQUIPMENT:
ARGENTINA, 1935-1964



Source: Diaz-Alejandro (1970).

removal of trade barriers on capital goods, there has been a surge of capital-good imports in India over the last decade. Although it is difficult to disentangle the effect of this policy change from that of other policy reforms introduced by the Indian Government at the same time, the fact is that India has experienced high growth rates over the last decade.

Ultimately then, this paper should not change one's prior idea that trade restrictions are bad for growth, but it is useful to point out that there is a large standard error surrounding the point estimate of its negative effect. If we want to narrow this error band, it is important to differentiate between the very different types of trade restrictions that countries have put into place. For example, I have provided some suggestive evidence that restrictions on capital-good imports have important adverse effects on growth. To the extent that this paper prompts us to ask (and attempt to answer) more refined questions about how trade restrictions affect growth, it serves a useful purpose. Nonetheless, I worry about the potential misuse of the authors' fine work by opponents of free trade in the political arena. After all, there are many vested interests that benefit from trade restrictions and much fewer interest groups that actively support free trade. It would be a shame if opponents of free trade (wrongly) interpret this paper as claiming that trade restrictions do not have adverse effects on growth, rather than as saying we don't precisely know how much trade barriers affect growth.

REFERENCES

- De Long, J. B., and L. Summers. (1991). Equipment investment and economic growth. *Quarterly Journal of Economics* 106(2):445–502.
- Diaz-Alejandro, C. (1970). *Essays on the Economic History of the Argentina Republic*. New Haven, CT: Yale University Press.
- Jones, C. (1994). Economic growth and the relative price of capital. *Journal of Monetary Economics* 34(3):359–382.
- Levine, R., and Renelt, D. (1992). A sensitivity analysis of cross-country growth regressions. *American Economic Review* 82(4):942–963.

Comment

CHARLES I. JONES
Stanford University and NBER

1. Introduction

Rodríguez and Rodrik replicate and check for robustness the results of several of the most influential papers in the cross-country growth literature on trade policy and economic growth. These studies suggest that policies that distort trade are associated with reduced growth rates over some period of time, and that the effects are fairly important in magnitude and relatively robust in terms of statistical significance.

Interpreted narrowly, the findings of Rodríguez and Rodrik suggest that the results of these existing studies are not as strong as the papers indicate. First, Rodríguez and Rodrik remind us that theory provides no clear indication of the net effect: trade restrictions could reduce income levels or growth rates through the usual channels such as specialization, but the common infant-industry argument, for example, suggests that trade restrictions could in some circumstances promote long-run performance. Second, we do not know exactly how we should measure trade restrictions, which leads to a large number of different approaches in the literature. However, it is not obvious that the variables used in these studies truly capture policy restrictions on trade, making the evidence difficult to interpret. Finally, Rodríguez and Rodrik argue that the results of these studies are not particularly robust. Including additional variables that plausibly belong in the specification, especially some measure of macroeconomic distortions (such as the black-market premium) or some measure of institutional quality or property rights [such as the Knack–Keefer (1995) measure], typically reduces the magnitude of the effect and enlarges the confidence interval substantially so that the trade-policy variable is not statistically significant at traditional levels.

REFERENCES

- De Long, J. B., and L. Summers. (1991). Equipment investment and economic growth. *Quarterly Journal of Economics* 106(2):445–502.
- Diaz-Alejandro, C. (1970). *Essays on the Economic History of the Argentina Republic*. New Haven, CT: Yale University Press.
- Jones, C. (1994). Economic growth and the relative price of capital. *Journal of Monetary Economics* 34(3):359–382.
- Levine, R., and Renelt, D. (1992). A sensitivity analysis of cross-country growth regressions. *American Economic Review* 82(4):942–963.

Comment

CHARLES I. JONES
Stanford University and NBER

1. Introduction

Rodríguez and Rodrik replicate and check for robustness the results of several of the most influential papers in the cross-country growth literature on trade policy and economic growth. These studies suggest that policies that distort trade are associated with reduced growth rates over some period of time, and that the effects are fairly important in magnitude and relatively robust in terms of statistical significance.

Interpreted narrowly, the findings of Rodríguez and Rodrik suggest that the results of these existing studies are not as strong as the papers indicate. First, Rodríguez and Rodrik remind us that theory provides no clear indication of the net effect: trade restrictions could reduce income levels or growth rates through the usual channels such as specialization, but the common infant-industry argument, for example, suggests that trade restrictions could in some circumstances promote long-run performance. Second, we do not know exactly how we should measure trade restrictions, which leads to a large number of different approaches in the literature. However, it is not obvious that the variables used in these studies truly capture policy restrictions on trade, making the evidence difficult to interpret. Finally, Rodríguez and Rodrik argue that the results of these studies are not particularly robust. Including additional variables that plausibly belong in the specification, especially some measure of macroeconomic distortions (such as the black-market premium) or some measure of institutional quality or property rights [such as the Knack–Keefer (1995) measure], typically reduces the magnitude of the effect and enlarges the confidence interval substantially so that the trade-policy variable is not statistically significant at traditional levels.

Interpreted broadly, the paper seems to suggest that trade-policy restrictions may not be particularly harmful to long-run economic performance, and that other factors could be much more important.

In preparing my discussion, I contacted several of the authors of four of the papers discussed by Rodríguez and Rodrik to get their general reactions. Because the issues are complicated and it would constitute a paper in itself, I have decided not to report and discuss their comments point by point. Suffice it to say that there are disagreements about a number of the criticisms among the parties involved.¹ Related to the “broad” interpretation of the paper, these authors reminded me that the belief among some economists that trade restrictions are harmful in the long run is based on many kinds of evidence, including case studies and micro studies. However, because this broader discussion is not my area of expertise, and because surely cross-country regressions are one piece of evidence upon which these beliefs are based, I will limit the scope of my discussion in the way the paper is limited.

My comment on Rodríguez and Rodrik’s paper will focus on the magnitude of the effect of trade restrictions on economic performance, providing a slightly different emphasis from that presented in the paper. First, I would like to review a useful way that cross-country growth regressions can be interpreted, focusing especially on the magnitude of the estimated effects in the long run. Second, I will attempt to interpret in this framework some specifications that Rodríguez and Rodrik seem to approve of most. In particular, I’d like to look at two questions: “What is our best estimate of the effect of trade restrictions on long-term eco-

1. I will report my interpretation of a few of the most interesting ones, though I surely will not do the authors justice. Andrew Warner pointed out to me that the “monopolizes exports” component of the Sachs–Warner index is *not* a dummy for sub-Saharan Africa. It is based on a careful analysis of the subject by the World Bank. It may closely resemble an Africa dummy, but maybe that is a good thing! One could include an Africa dummy with the Sachs–Warner openness measure to check for robustness; in my tests, the openness measure survives. Also, the spirit of their index is that a country can close itself off in a number of different ways that may differ across countries, and Sachs and Warner try to provide an index to capture this phenomenon. This nonlinearity means that running a horse race among the components of the index will not capture the same forces. Dan Ben-David reminded me of Figures XII and XIII in his paper, which provide an additional piece of evidence supporting his view: the reduction in tariffs between the United States and Canada in the late 1960s associated with the Kennedy round, and the associated behavior of incomes. He also noted that the breakdown of European trade in the interwar period is associated with a cessation of convergence, and the resumption of convergence occurs with the reduction of tariffs and quotas after the war. Sebastian Edwards noted that he has tried in earlier work to address measurement-error concerns by running “reverse” regressions. With respect to heteroskedasticity, he also commented that there are conceptual concerns about White-robust errors and that different weightings give different results (for example, weighting by exports per capita gives results like those he obtained). David Dollar provided a broader perspective that is incorporated throughout my comment.

conomic performance?" and "How confident are we about the magnitude of this effect?"

2. *Interpreting Cross-Country Growth Regressions*

The interpretation of cross-country growth regressions that I find most useful is provided by Mankiw, Romer, and Weil (1992) and Barro and Sala-i-Martin (1992). These papers derive a basic cross-country growth specification from a neoclassical growth model. The derived specification suggests that the growth rate of a particular country over some time period, like thirty years, is a function (often linearized) of the gap between where the country starts out and the country's steady state. To be more accurate, the simplest neoclassical growth model has one state variable, such as the ratio of per capita income to the technology index ($\tilde{y} \equiv y/A$), and the model predicts that the growth rate of this state variable is approximately proportional to the gap between its current value and its steady-state value:

$$\frac{\dot{\tilde{y}}_{it}}{\tilde{y}_{it}} = -\lambda(\log \tilde{y}_{it} - \log \tilde{y}_i^*),$$

where λ is commonly called the speed of convergence. The technology index is often assumed to follow some simple process, such as

$$\log A_{it} = \log A_i + \log Z_t + \epsilon_{it}.$$

That is, we assume that a country's technology index is the product of a parameter A_i indexing a country's long-run productivity level, the world technology index (which is assumed to grow at a constant rate g), and an idiosyncratic disturbance around this trend.

The first equation can be integrated and combined with the second to yield a cross-country growth specification:

$$\bar{g}_{iT} = \text{constant} - \beta \log y_{i0} + \beta \log (\tilde{y}_i^* A_i) + \beta \epsilon_{i0} + \frac{1}{T} (\epsilon_{iT} - \epsilon_{i0}), \quad (1)$$

where $\bar{g}_{iT} \equiv (1/T)(\log y_{iT} - \log y_{i0})$ and $\beta \equiv (1/T)(1 - e^{-\lambda T})$.

A difficulty with this approach is that one does not observe directly the steady state to which countries are converging, nor the total factor productivity parameter. Variables such as investment rates in physical or human capital can be connected to \tilde{y}^* theoretically, but of course these

variables are typically endogenous as well. This leads to the difficult situation in which the econometrician does not know the correct specification but has a large number of candidate regressors at hand. An additional problem with this approach is the possible correlation of the candidate regressors with the error term(s), including the possibility of omitted-variable bias and endogeneity.

What I'd like to point out about this specification, however, is that the reason variables like trade policy or the quality of institutions are thought to enter these regressions is that they are potential determinants of the steady-state income level (detrended by the world technology index) toward which an economy is converging. This suggests an alternative specification of the regression that Mankiw, Romer, and Weil (1992) explore and that Hall and Jones (1999) have emphasized recently, a specification in levels rather than growth rates:

$$\log y_{it} = \text{constant} + \log(\tilde{y}_i^* A_i) + \epsilon_{it} - \frac{1}{\beta} \bar{\delta}_{\tilde{y}_i} \quad (2)$$

If levels of output per worker at time t are randomly distributed around their steady-state values, then this specification has the potential to work well. Notice that it uses different variation in the data, in that the estimation does not first condition on an earlier level of output per worker. One advantage is that more precise estimates may be obtained as a result. Of course, there are still endogeneity and omitted-variable problems, but these issues are also relevant for the specification in terms of growth rates; in some ways, they are simply made more explicit by the levels specification.

In terms of interpretation, the coefficients from the cross-country growth specification are really the product of two factors: a speed-of-convergence factor (β) and the coefficient that relates the particular variable to the steady-state level of income. One can interpret this product of coefficients as the effect on average growth rates over a particular period, but when the length of the time period is changing, as it is across these studies, the size of the coefficient will change for this reason (note that β depends on T), making comparisons across specifications difficult.

An alternative useful interpretation is obtained by calculating the long-run effect on the steady state, either by dividing by the coefficient on initial income or simply by running the levels regression directly.² One may of course also care about the rate at which the economy con-

2. These two methods will generally yield different results, since different variation in the data is used to estimate the effects; both are useful in practice.

verges to its steady state, and this rate, λ , can be calculated from the estimate of β .

3. *A Closer Look at Some Results*

Rodríguez and Rodrik examine a large number of measures of trade restrictions in their evaluation of the literature. Many are criticized for reasons discussed briefly above, but a few are put forward as being reasonable measures. These are typically the most direct measures of tariff rates or nontariff barriers. I will focus on three particular measures: (1) the QT component of the Sachs–Warner openness measure, which takes a value of 1 unless the country had average tariff rates higher than 40% or nontariff barriers covered more than 40% of imports, in which case it takes a value of 0; (2) an average tariff rate measure from Barro and Lee (1993) (owti); and (3) the simple average of the available statistics on import duties as a percentage of imports, which are reported in Table VIII of the conference version of their paper and which Rodríguez and Rodrik refer to in their conclusion. For some reason that I do not understand, they do not use this import-duties variable in any of their robustness checks in the paper.

I should make clear from the beginning that a narrow version of Rodríguez and Rodrik’s conclusion survives my analysis of these data: estimates using these variables are not completely robust, in the sense that confidence intervals are large in some specifications. However, I’d like to go further and examine the magnitude of the effects and the confidence interval itself. What is our best guess about the effect of trade restrictions on long-run economic performance, and what is our range of uncertainty?

Table 1 summarizes my findings from estimating approximately 100 specifications; from among these, I’ve selected the 13 that strike me as most appropriate, and I’ve further summarized these 13 specifications by averaging the coefficients and p -values and reporting some statistics. A few of the specifications are growth regressions, replicating results in Rodríguez and Rodrik’s paper; most are levels regressions of the same basic specifications, which generally improved the precision of the estimates.³ One possible problem with these levels regressions is reverse causality: poor countries may resort to tariffs to raise revenue more than rich countries, e.g., because their tax systems are not well developed. In results not reported, I made some attempt to address issues of endogeneity by instrumenting with the variables used in Hall and Jones

3. The growth-regression specifications produced estimates of the long-run effect of 0.535 for QT and -1.80 for owti, roughly in line with the results from the levels regressions.

Table 1 SOME ADDITIONAL RESULTS

	<i>Sachs–Warner</i> <i>QT</i>	<i>Tariff</i> <i>rate, owti</i>	<i>Avg. import</i> <i>duties</i>
<i>Results for All Specifications but Worst</i>			
Average long-run effect	0.485	-1.714	-2.758
S.d. of variable	{0,1}	0.17	0.079
Average <i>p</i> -value	0.064	0.055	0.005
Number of specifications	4	4	2
Fraction with <i>p</i> < .10	3/4	3/4	2/2
<i>Results from Worst Specification</i>			
Long-run effect	0.158	-0.411	-0.447
<i>p</i> -value	0.275	0.509	0.375
95% conf. interval	(-0.13, 0.45)	(-1.6, 0.83)	(-3.17, 2.27)
<i>Proportional Reduction of SS Output per Worker from a Large Increase in Trade Restrictions</i>			
All but "worst"	39%	69%	58%
"Worst"	15%	24%	13%

The worst specification for the Sachs–Warner QT variable occurs when Knack and Keefer's (1995) quality-of-institutions variable (*icrge*) is added to the specification. The worst specification for the tariff rate (*owti*) occurs when both *icrge* is added and simultaneously the outlier India is dropped. The worst specification for the average-import-duties variable occurs when an indicator variable for the African continent is added to the specification. The calculations of long-run effects report the proportionality factor by which incomes would be reduced in the long run if a hypothetical country increased trade restrictions by 4 standard deviations (or went from a 1 to a 0 in the Sachs–Warner case). It is calculated as, e.g., $1 - \exp(\beta \times 4 \times \text{stdev})$. All but two of the regression results are from levels regressions; a growth regression is run for each of these first two variables (and is the specification with the largest *p*-value in the first part of the table). The first two columns use the Rodríguez–Rodríguez dataset *Sw.dat* and include *gvxdxe*, *assassp*, *revcoup*, and be as additional regressors, sometimes adding *africa* and *icrge*. Results for the last column include variables from Hall and Jones (1999) as additional regressors.

(1999); in general, the point estimates were actually a little larger in magnitude, perhaps because of measurement error, but the estimates were less precise. A similar result is found by Frankel and Romer (1999).

The table is divided into three parts. In the first, I report the average effect on steady-state incomes from two to four specifications that exclude the specification that is *worst* in the sense of having the least-significant (and, it turns out, smallest) estimate. In the second, I report this worst specification.

In general, there are a number of reasonable specifications that lead to precisely estimated effects, as summarized in the first part of the table. In my brief experience, however, there were typically one or two key things that could be added to these specifications that led to problems (see the notes to the table). For example, adding the quality-of-institutions variable from Knack and Keefer (1995) often led the trade-policy variable to be estimated imprecisely. This could mean that the trade-policy variable is in

part proxying for other kinds of distortions that are omitted from the specification. On the other hand, the Knack–Keefe variable is itself not without problems, as it is a subjective measure constructed by a consulting firm.

The third section of the table examines the magnitude of the effects estimated in the previous two parts. Specifically, I calculate the change in steady-state income associated with a large change in trade policy, viz. a movement of 4 standard deviations, or a movement from 1 to 0 for the Sachs–Warner variable. For all but the worst specifications, our best estimate of the size of the effect is substantial—a decline in income by 40% to 70%. For the worst specification, the effects are smaller: income declines by between 13% and 24% in the long run.

Overall, these numbers are similar to results calculated from some of the specifications reported by Rodríguez and Rodrik, such as in Table 3. However, at least in the conference version of their paper, they do not provide enough detail for the reader to make these calculations.

4. Final Thoughts

There are two other recent papers that I think should be mentioned in this context. The first is an omission from the conference version of the paper that has to some extent been addressed in the published version: the study of openness and income levels by Frankel and Romer (1999). Frankel and Romer’s measure of openness is the trade share of GDP rather than a policy variable, and their general finding is a relatively robust relationship between openness and income levels: a change that increases the trade share by one percentage point raises income levels by 1% to 2%. A key contribution of the paper is to show that this finding is robust to endogeneity concerns by using the geographical determinants of trade as an instrument. Another finding, however, is that the magnitude of the effect is somewhat imprecisely estimated, and 95% confidence intervals include zero in a number of specifications.

Another paper that I’ve found helpful is Sala-i-Martin (1997). People sometimes conclude from the cross-country growth regression literature that virtually none of the relationships are robust, a statement that would seem to receive support from Rodríguez and Rodrik. Sala-i-Martin builds on the robustness work by Levine and Renelt (1992) by examining the entire distribution of coefficient estimates on particular variables from running more than 32,000 permutations of growth regressions. As a general matter, Sala-i-Martin highlights a number of variables that are robust across specifications, including the Sachs–Warner openness measure. On the other hand, consistent with the present paper—and with the original results of Levine and Renelt (1992)—Sala-i-Martin

finds that the other measures of trade policy he examines are among the least robust variables in his study, being statistically significant at the 95% level less than 4% of the time. He does find that the coefficients have the “right” sign in 60% to 80% of the specifications he considers, depending on the measure.

In conclusion, it seems to me that the cross-country growth regression evidence leads to the following results. Our best estimate is that trade restrictions are harmful to long-run incomes, and that the effects are potentially large. For this reason, I worry a little about the “broad” interpretation of the paper that I provided at the beginning of my remarks. In addition, however, there is a large amount of uncertainty regarding the magnitude of the effect; it could be small, and there are some specifications that allow for the possibility that the effect works in the opposite direction. Cross-country growth regressions appear to be a coarse tool for this particular question, and, at least so far, are unable to provide a more precise answer.

REFERENCES

- Barro, R. J., and J. Lee. (1993). International comparisons of educational attainment. *Journal of Monetary Economics* 32(December):363–394.
- , and X. Sala-i-Martin. (1992). Convergence. *Journal of Political Economy* 100(2):223–251.
- Frankel, J. A., and D. Romer. (1999). Does trade cause growth? *American Economic Review* 89(3):379–399.
- Hall, R. E., and C. I. Jones. (1999). Why do some countries produce so much more output per worker than others? *Quarterly Journal of Economics* 114(1):83–116.
- Knack, S., and P. Keefer. (1995). Institutions and economic performance: Cross-country tests using alternative institutional measures. *Economics and Politics* 7(November):207–227.
- Levine, R., and D. Renelt. (1992). A sensitivity analysis of cross-country growth regressions. *American Economic Review* 82(4):942–963.
- Mankiw, N. G., D. Romer, and D. Weil. (1992). A contribution of the empirics of economic growth. *Quarterly Journal of Economics* 107(2):407–438.
- Sala-i-Martin, X. (1997). I just ran four million regressions. *American Economic Review* 87(2):178–183.

Discussion

Andrew Warner disputed the authors’ conclusion that there are no strong results in the literature linking trade openness to growth. He asserted that the simplest possible measure of protectionism, average tariff rates, is significantly and negatively correlated with growth, controlling for initial income and growth and even excluding Africa from the

finds that the other measures of trade policy he examines are among the least robust variables in his study, being statistically significant at the 95% level less than 4% of the time. He does find that the coefficients have the “right” sign in 60% to 80% of the specifications he considers, depending on the measure.

In conclusion, it seems to me that the cross-country growth regression evidence leads to the following results. Our best estimate is that trade restrictions are harmful to long-run incomes, and that the effects are potentially large. For this reason, I worry a little about the “broad” interpretation of the paper that I provided at the beginning of my remarks. In addition, however, there is a large amount of uncertainty regarding the magnitude of the effect; it could be small, and there are some specifications that allow for the possibility that the effect works in the opposite direction. Cross-country growth regressions appear to be a coarse tool for this particular question, and, at least so far, are unable to provide a more precise answer.

REFERENCES

- Barro, R. J., and J. Lee. (1993). International comparisons of educational attainment. *Journal of Monetary Economics* 32(December):363–394.
- , and X. Sala-i-Martin. (1992). Convergence. *Journal of Political Economy* 100(2):223–251.
- Frankel, J. A., and D. Romer. (1999). Does trade cause growth? *American Economic Review* 89(3):379–399.
- Hall, R. E., and C. I. Jones. (1999). Why do some countries produce so much more output per worker than others? *Quarterly Journal of Economics* 114(1):83–116.
- Knack, S., and P. Keefer. (1995). Institutions and economic performance: Cross-country tests using alternative institutional measures. *Economics and Politics* 7(November):207–227.
- Levine, R., and D. Renelt. (1992). A sensitivity analysis of cross-country growth regressions. *American Economic Review* 82(4):942–963.
- Mankiw, N. G., D. Romer, and D. Weil. (1992). A contribution to the empirics of economic growth. *Quarterly Journal of Economics* 107(2):407–438.
- Sala-i-Martin, X. (1997). I just ran four million regressions. *American Economic Review* 87(2):178–183.

Discussion

Andrew Warner disputed the authors’ conclusion that there are no strong results in the literature linking trade openness to growth. He asserted that the simplest possible measure of protectionism, average tariff rates, is significantly and negatively correlated with growth, controlling for initial income and growth and even excluding Africa from the

sample. With respect to his own work with Jeff Sachs, Warner argued that there are many ways to close an economy to trade—hence the strategy of constructing a composite variable which treats a country as being closed to trade if any of a number of criteria are met. Dani Rodrik said that it is difficult to interpret the bivariate relationship of tariff rates and growth, as rich countries tend to lower tariff rates, which leads to the possibility of reverse causality; he also questioned whether the estimated relationship between tariff rates and growth holds up in more recent data. Rodrik agreed in general with Sachs and Warner’s strategy of combining indicators. However, given the paper’s finding that much of the statistical effect of the Sachs–Warner indicator is due to only two of the variables that make it up, he argued that one must be careful to determine whether these key variables truly measure trade policies or instead reflect other country characteristics.

Alberto Alesina argued that growth rates may be an especially poor measure of the benefits of trade; for example, trade permits people to enjoy a wide variety of products not produced at home. On the other hand, Alesina and Allan Drazen both emphasized the point that trade policy is not made by social planners but by lobbies and interest groups. It may be that interest groups fight harder to protect their income shares through trade protection when income is growing slowly overall; this is yet another possible source of reverse causation. Pursuing the political-economy issue, Daron Acemoglu pointed out that the correlation between restrictive trade policies and corrupt, rent-seeking governments may not be an accident; the two may be mutually supporting. Thus, one benefit of more open trade is that it may reduce the scope for governmental corruption.

Marvin Goodfriend differentiated between the classical static efficiency benefits of trade and the dynamic gains associated with the diffusion of knowledge and technology. Possibly, he suggested, improving communications (including developments such as the Internet) will reduce the importance of trade policy for information flows.

Greg Mankiw was not surprised by the lack of robustness in the cross-country results, given the large number of candidate variables relative to the number of country observations. He conjectured that economists support free trade because they believe Ricardo, not because they have been convinced by regressions. Rodrik agreed that there is a strong presumption that trade restrictions are distortionary, but that magnitudes are important. For example, if the growth effects of trade liberalization are small, economic advisors may do better by giving a higher priority to other types of reforms.