Comment on Rodríguez and Rodrick, "Trade Policy and Economic Growth: A Skeptic's Guide to the Cross-National Evidence"

Charles I. Jones

Stanford University and NBER
Chad.Jones@Stanford.edu
http://www.stanford.edu/~chadj

June 20, 2000 – Version 2.0

Comment prepared for presentation at the NBER Macroeconomics Annual Conference, April 7-8, 2000 in Cambridge, Massachusetts.

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 relatively important in terms of 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 Rodrick 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 and 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 the Rodríguez and Rodrick 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 Rodrick 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 economic performance?" and

[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 U.S. 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 heteroscedasticity, 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.

"How confident are we about the magnitude of this effect?".

1 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 simply 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 idiosynchratic disturbance around this trend.

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

$$\bar{g}_{iT} = Constant - \beta \log y_{i0} + \beta \log(\tilde{y}_i^* A_i) + \beta \epsilon_{i0} + \frac{1}{T}(\epsilon_{it} - \epsilon_{i0}) \qquad (1)$$

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 possiblity 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 because 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 et al. (1992) explore and that Hall and Jones (1999) have emphasized recently, a specification in levels rather than growth rates:

$$\log y_{it} = Constant + \log(\tilde{y}_i^* A_i) + \epsilon_{it} - \frac{1}{\beta}\bar{g}_{\tilde{y}_i}. \tag{2}$$

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 terms: a speed-of-convergence term (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 converges to its steady state, and this rate, λ, can be calculated from the estimate of β.

2 A Closer Look at Some Results

Rodríguez and Rodrick 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: (i) 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 percent or nontariff barriers covered more than 40 percent of imports, in which case it takes a value of 0; (ii) an average tariff rate measure from Barro and Lee (1993) (*owti*); and (iii) 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 that Rodríguez and Rodrick 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 the Rodríguez and Rodrick 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 inter-

[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.

val 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 the Rodríguez and Rodrick 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 (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

[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.

in part proxying for other kinds of distortions that are omitted from the specification. On the other hand, the Knack-Keefer 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, i.e. 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 percent. For the "worst" specification, the effects are smaller: income declines by between 13 and 24 percent in the long run.

Overall, these numbers are similar to results calculated from some of the specifications reported by Rodríguez and Rodrick, such as in Table IV.1. However, at least in the conference version of their paper, they do not provide enough detail for the reader to make these calculations.

3 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 one to two percent. 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-percent 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 Rodrick. 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 percent level less than 4 percent of the time. He does find that the coefficients have the "right" sign in 60 to 80 percent 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 possiblity 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.

Table 1: Some Additional Results

	Sachs-Warner QT	Tariff Rate (*owti*)	Avg. Import Duties
	Results for All Specifications but "Worst"		
Average LR Effect	.485	-1.714	-2.758
S.D. of Variable	{0,1}	.17	.079
Average p-value	.064	.055	.005
Number of Specifications	4	4	2
Fraction w/ p-value$<$ .10	3/4	3/4	2/2
	Results from "Worst" Specification		
Long-run (LR) Effect	.158	-0.411	-0.447
p-value	.275	.509	.375
95% Conf. Interval	(-.13, .45)	(-1.6, .83)	(-3.17, 2.27)
	Proportional Reduction of SS Output per Worker from a Large Increase in Trade Restrictions		
All but "Worst"	39%	69%	58%
"Worst"	15%	24%	13%

Note: The worst specification for the Sachs-Warner QT variable occurs when the Knack and Keefer (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 "Range of LR Effects" calculations report the proportional 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., $1 - exp(\beta * 4 * stdev)$.

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 and Rodrick data set 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 the Hall and Jones (1999) paper as additional regressors.

References

Barro, Robert J. and Jongwha Lee, "International Comparisons of Educational Attainment," *Journal of Monetary Economics*, December 1993, *32*, 363–394.

___ and Xavier Sala-i-Martin, "Convergence," *Journal of Political Economy*, 1992, *100* (2), 223–251.

Frankel, Jeffrey A. and David Romer, "Does Trade Cause Growth?," *American Economic Review*, June 1999, *89* (3), 379–399.

Hall, Robert E. and Charles I. Jones, "Why Do Some Countries Produce So Much More Output per Worker than Others?," *Quarterly Journal of Economics*, February 1999, *114* (1), 83–116.

Knack, Stephen and Philip Keefer, "Institutions and Economic Performance: Cross-Country Tests Using Alternative Institutional Measures," *Economics and Politics*, November 1995, *7*, 207–227.

Levine, Ross and David Renelt, "A Sensitivity Analysis of Cross-Country Growth Regressions," *American Economic Review*, September 1992, *82* (4), 942–963.

Mankiw, N. Gregory, David Romer, and David Weil, "A Contribution to the Empirics of Economic Growth," *Quarterly Journal of Economics*, May 1992, *107* (2), 407–438.

Sala-i-Martin, Xavier, "I Just Ran Four Million Regressions," 1997. Columbia University mimeo.