WP/04/135

IMF Working Paper

Once Again, is Openness Good for Growth?

Ha Yan Lee, Luca Antonio Ricci, and Roberto Rigobon

INTERNATIONAL MONETARY FUND

 WP/04/135

IMF Working Paper

Research Department

Once Again, is Openness Good for Growth?[1]

Prepared by Ha Yan Lee, Luca Antonio Ricci, and Roberto Rigobon

Authorized for distribution by Gian Maria Milesi-Ferretti

July 2004

Abstract

Rodriguez and Rodrik (2000) argue that the relation between openness and growth is still an open question. One of the main problems in the assessment of the effect is the endogeneity of the relation. In order to address this issue, this paper applies the identification through heteroskedasticity methodology to estimate the effect of openness on growth while properly controlling for the effect of growth on openness. The results suggest that openness would have a positive effect on growth, although small. This result stands, despite the equally robust effect from growth to openness.

JEL Classification Numbers: F10, F43, O40, C33

Keywords: Trade openness, economic growth, endogeneity, identification through heteroskedasticity

Author's E-Mail Address: hlee@mit.edu, lricci@imf.org, rigobon@mit.edu

[1] The authors are grateful to Athanasios Vamvakidis for kindly sharing his dataset. They also thank Raphael Bergoeing, Sebastian Edwards, Alejandro Mico, Andrea Repetto, and seminar participants at IASE in Santiago de Chile, and the World Bank, for very useful discussions and comments.

Contents Page

I. Introduction

A fundamental question in development and international economics is whether higher trade openness helps to improve economic growth. Even though this question has received a great deal of attention in the literature for more than a century, we still do not have, as a profession, a satisfactory answer. In a recent survey, Rodriguez and Rodrik (2000) provide a critical analysis of the main contributions of the past decade and conclude that "the nature of the relationship between trade policy and economic growth remains very much an open question" (p. 266). However, the authors interpret "the persistent interest in this area as reflecting the worry that the existing approaches haven't gotten it quite right" (p. 316).

The present paper offers some insights as to how to exploit differences across countries to address one of the main obstacles to answering the question: the problem of simultaneity, which is pervasive in econometric analysis. What most economists would consider good measures of the degree of openness of a country are, unfortunately, closely linked to the level of income. For example, measuring openness as the ratio between the sum of imports plus exports to GDP clearly is a function of the growth rate of the economy – both the numerator and the denominator are linked to the GDP growth. What this implies is that not even the sign of the bias in the standard OLS regression can be assessed.

Instrumentation via lags or other economic indicators does not offer a valid alternative if, respectively, openness is serially correlated or these other variables affect growth as much as trade. Dollar and Kraay (2003) suggest estimating the regression in first differences and instrumenting the change in openness via lagged values of openness, which appear uncorrelated with other factors influencing changes in growth. Unfortunately, the simultaneity bias can extend over time in the case under consideration. For instance, if part of the growth rate in the future is driven by investment today that requires imported goods, then the degree of openness today depends on future growth rate. Hence, using lag values of openness as instruments does not provide a reliable solution. Alternatively, other economic indicators—such as geographic ones—have been used as instruments (see, for example, Frankel and Romer (1999)). However, these instruments have often been criticized as being correlated with income as much as with trade. For example, the gravity literature has shown that geographic variables can be robustly employed to explain trade. At the same time, they can also be expected to affect growth via other channels, for example, via their relation to health conditions and productivity, to the quality of institutions, and to the availability of natural resources (see Rodriguez and Rodrik (2000) and Baldwin (2003)). To the extent the component of trade that is instrumented via geography is correlated with these other factors, the Instrumental Variables (IV) estimates are likely to be biased. Dollar and Kraay (2003) address the instrumentation problem by estimating the growth regressions in first difference and instrumenting the changes in the main explanatory variables (trade and institutions) via their lagged levels. Partly due to the different dynamic nature of the two variables, the lagged levels of trade predict well changes in trade, but not changes in institutions, and vice versa, thus allowing for a better instrumentation for trade.

In this paper, we tackle the extremely important issue of endogeneity again, but from a very different perspective. Instead of appealing to instruments that move the degree of openness but are uncorrelated with income growth (which we have already argued that are difficult to find) we solve the problem of simultaneous equations by using the relatively new literature on "identification through heteroskedasticity" (IH).

In the standard simultaneous equations problem the instrumental variables approach searches for a variable that shifts the demand (for example) to estimate the supply. In other words, we need a variable that moves the *mean* of the demand curve to estimate the supply. We now know that this procedure was introduced in Philip Wright's 1928 book in Appendix B (see Stock and Trebbi, 2003). In the same appendix, Wright indicates that we could also solve the problem of identification if instead of moving the mean we find a variable that increases the variance of one of the equations to infinity. In other words, if the volatility of the innovations to the demand schedule is infinitely large in comparison to the volatility of the innovations of the supply, in the data we only observe movements due to innovations to the demand, and therefore we can estimate the supply schedule directly. This is known today as "near identification".[2] In 1929, Leontief indicated that we did not have to move the variance to infinity, only knowing that it has changed implies that the distribution of the residuals rotate along the different schedules. Rigobon (2003a) shows that this is enough to achieve identification. [3] This approach has been defined in the literature as identification through heteroskedasticity (IH from now on) and has been recently applied to numerous issues plagued by reverse causality.[4]

Sample heteroskedasticity, both across time and across countries, is very high both in country growth rates and in their degrees of openness. It is therefore possible to use the IH method to solve for the problem of endogeneity, when analyzing our fundamental question of the effect of trade openness on growth.

---

[2] See Fisher (1976).

[3] Other theoretical derivations can be found in Sentana (1992), Sentana and Fiorentini (2001).

[4] Applications where the heteroskedasticity is modeled as a GARCH process are Caporale, et. al. (2002a), Rigobon (2002a), Rigobon and Sack (2003a). Applications where the heteroskedasticity is described by regimes shifts are Rigobon (2002b, 2003b), Rigobon and Sack (2003b), Caporale et. al. (2002b). Applications to event study estimation are developed by Rigobon and Sack (2004) and Evans and Lyons (2003). Finally, several applications to panel data can be found in the literature. Hogan and Rigobon (2003) apply the method to a very large panel data to estimate the returns to education. Rigobon and Rodrik (2004) study instead the impact of institutions on income, and how the different types of institutions are affected by income levels and the degree of openness of the country. Klein and Vella (2003) also use heteroskedasticity to estimate the returns to education. Broda and Weinstein (2004) use the inequality constraints together with the heteroskedasticity to estimate the elasticities of substitution in models of trade to evaluate the gains from variety. Pattillo, Poirson, and Ricci (2003) use the IH method to identify the impact of external debt on growth. Hviding, Nowak, and Ricci (2004) investigate the impact of official reserves on exchange rate volatility.

Our results suggest that openness has a small positive effect on growth, which is not particularly robust. They also suggest that most of the empirical works that claim to find a strong and robust result are instead likely to capture either the reverse causality effect or the effect of other economic and policy distortions that are correlated with openness (such as the black market premium, this latter point being already discussed by Rodriguez and Rodrik, 2000).

This paper is organized as follows: Section II, presents the estimation problem and derives the estimator used in the paper. In Section III, we present our results using identification through heteroskedasticity. Finally, Section IV concludes.

II. Endogeneity

As was argued, "good" measures of openness are in general closely related to the growth rate. This generates the standard endogeneity problem. To simplify the discussion we abstract from other controls and concentrate mainly on the simultaneous equations problem. Therefore, assume that openness and growth are described by the following system of equations:

$$\begin{aligned} y_t &= \alpha\ o_t + \varepsilon_t \\ o_t &= \beta\ y_t + \eta_t \end{aligned} \qquad (1)$$

where y is the growth rate, and o is the degree of openness. We are interested in estimating α, but as it is well known that with standard econometric methodologies we cannot consistently estimate α when either β is different from 0 or the variance of η is finite—unless openness and growth are cointegrated—even if we are willing to make the strong assumption that the structural shocks have finite variance and are uncorrelated. The results discussed below apply for any distribution.

Assume we estimate the first equation by OLS not taking into consideration the simultaneous equations problem. The estimation is biased because the right hand side variables are correlated with the residuals. In particular, in equation:

$$y_t = \alpha\ o_t + \nu_t$$

the OLS estimate is given by the general expression:

$$\begin{aligned} \hat{\alpha} &= (o_t' o_t)^{-1} o_t' y_t \\ &= \alpha + \beta(1-\alpha\beta)\frac{\sigma_\varepsilon^2}{\sigma_\eta^2 + \beta^2\sigma_\varepsilon^2} \end{aligned}$$

This equation encompasses all cases where the problem of simultaneous equations disappears:

- Exclusion restrictions: if $\beta = 0$, the bias goes to zero – which should be obvious given that $\beta = 0$ is assuming the problem away. This is what typically is assumed when exclusion restrictions are imposed on the system of equations. For example, most of the macro literature using VARs and identifying the model using the Choleski decompositions implicitly are making this assumption.

- Near identification: Assume that the variance $\sigma_\eta^2 \to \infty$, then it is easy to verify that if the other variance is finite then the bias goes to zero. Near identification is one of the most used assumptions in event study papers. Unfortunately, it is usually assumed implicitly, instead of explicitly.

- Cointegration or infinite variance: if the variables are cointegrated, or similarly if the observable variables (y and o) have infinite variance even though the residuals have finite variance, then the expression $(1-\alpha\beta)$ would be close to zero. In this case, it is evident that the bias goes to zero. We know from the cointegrating literature that if the variables are random walks but have a cointegrating relationship (which means that a linear combination of them is stationary), then the model is super consistent and we can run OLS. In this model, if $(1-\alpha\beta)=0$ we have a similar result. Having infinite variance for each of the variables, but finite variance for the linear combination of them is equivalent to having two cointegrating relations: $y_t = \alpha\ o_t + v_t$, and $o_t = \beta\ y_t + \pi_t$. The problem in practice is that we will not know which of the two we are estimating, but there will be no inconsistency in the estimates.

There are other assumptions that have been used to achieve identification: long-run restrictions usually assume that the sum of lag coefficients is zero for some type of shock. Additionally, partial identification can be achieved if sign restrictions are imposed. We believe that these assumptions do not apply to the problem of growth and openness, and therefore, we do not discuss them.

Observe that we can also have an estimate for the coefficient in the second equation. The estimate for β is also biased.

$$
\begin{aligned}
\hat{\beta} &= (y_t' y_t)^{-1} y_t' o_t \\
&= \beta + \alpha(1-\alpha\beta)\frac{\sigma_\eta^2}{\alpha^2\sigma_\eta^2 + \sigma_\varepsilon^2}
\end{aligned}
$$

In summary, both estimates are biased if we cannot justify exclusion restrictions, or that one of the variances is infinitely large, or that the variables are cointegrated, or that there are long-run restrictions (which is a form of exclusion restriction). We believe that the problem of simultaneity between growth and openness cannot be solved by appealing to these assumptions. Indeed, most of the literature does not appeal to them because they are impossible to rationalize in this particular framework.

A. Reversed Regressions: OLS Bounds

Before proceeding towards the IH methodology, it is interesting to discuss a very old literature studying the "bounds" of the OLS estimates in the presence of misspecification from the correlation of the explanatory variables and the residuals (this is a general problem which encompasses other issue in addition to simultaneous equations). The purpose of the bounds is to highlight or show the extent of the misspecification. The method was used by Leontief (1929) and it was recovered by Leamer (1981) and Edwards (1992).

Assume we have a general problem of misspecification that can be summarized by the simple relationship

$$y_t = a\ o_t + v_t$$

where the right hand side variable o_t is correlated with the residual v_t. Notice that this is exactly the first equation in our system of equations, but here we would like to offer the general discussion when this correlation arises from multiple sources, not just from reverse causality (and hence we use different terms for parameters and residuals than in equation 1).

It is well known that, in the presence of this misspecification, we cannot estimate *a* consistently. Indeed, there are two forms of estimating *a*.

$$y_t = a\ o_t + v_t \qquad (2a)$$

$$\begin{aligned} o_t &= \frac{1}{a}\ y_t + \tilde{v}_t \\ &= b\ y_t + \tilde{v}_t \end{aligned} \qquad (2b)$$

It is important to indicate that both regressions are *equally* wrong! Leontief studied this problem and realized that depending on the sources of the misspecification the OLS estimates in these two regressions provide bounds for the true coefficient. The estimate in equation (2a) provides one bound, and the inverse of the estimate on equation (2b) provides the other bound. A special case arises when the misspecification in the model is due to simultaneous equations. In particular, assume output and openness are given by our model equation (1). Then the OLS estimate in equation (2a) is (the same as before):

$$\hat{a}_{2a} = (o_t' o_t)^{-1} o_t' y_t = \alpha + \beta(1-\alpha\beta)\frac{\sigma_\varepsilon^2}{\sigma_\eta^2 + \beta^2\sigma_\varepsilon^2}$$

while the estimate of 1/*a* in equation (2b) is (note that the two expressions are identical):

$$\begin{aligned} \hat{b}_{2b} = (y_t' y_t)^{-1} y_t' o_t &= \beta + \alpha(1-\alpha\beta)\frac{\sigma_\eta^2}{\alpha^2\sigma_\eta^2 + \sigma_\varepsilon^2} \\ &= \frac{1}{\alpha} - \frac{1}{\alpha}(1-\alpha\beta)\frac{\sigma_\varepsilon^2}{\alpha^2\sigma_\eta^2 + \sigma_\varepsilon^2} \end{aligned}$$

Note that if we are interested in the estimation of α, we want to solve $\hat{b}_{2b}$ for $\frac{1}{\alpha}$ instead of β.

We can in fact use both estimates $\hat{a}_{2a}$ and $\hat{b}_{2b}$ to compute the range where the true coefficient α must lie if the model is correct. To illustrate the range, consider the two possible cases, where α and β have different or similar signs.

If α and β have different signs, the bias in equation (2a) makes the OLS coefficient smaller (in absolute value) than the true one. In other words,

$$|\hat{a}_{2a}| < |\alpha|$$

Similarly, it is also easy to show that in equation 2b the bias is also toward zero. Hence we can write

$$|\hat{b}_{2b}| < \left|\frac{1}{\alpha}\right|$$

Therefore,

$$|\hat{a}_{2a}| < |\alpha| < \left|\frac{1}{\hat{b}_{2b}}\right|.$$

In other words, if the two schedules have different signs, then the true coefficient lies between these two estimates.

The intuition of this result is very simple. First, it is important to realize that equation 2a is the OLS run in one direction, while equation 2b is the OLS regression in the other direction. If the schedules have different signs, simultaneous equations will bias the OLS coefficients toward zero, because the OLS coefficient is a linear combination of the two coefficients—one positive, and the other negative. Hence the OLS coefficients in both regressions are smaller in absolute terms than the true ones. However, the coefficient in the first equation (2a) attempts to estimate α and the coefficient in the second one (2b) $1/\alpha$. This is what determines the range.

When the two schedules have the same signs the range of coefficients is different. In this case, the bias in the OLS in both equations (2a and 2b) is away from zero.[5] So, if both coefficients are positive the OLS is larger than the true one, and if the coefficients are negative the OLS ones are smaller than the true ones. Nevertheless, this means that in absolute terms,

$$|\alpha| < \min\left\{|\hat{a}_{2a}|, \frac{1}{|\hat{b}_{2\mathrm{b}}|}\right\}$$

Again, this implies a range of coefficients that is admissible. The intuition in this case, follows the same reasoning as before, where the difference is due to the fact that in both equations the estimated coefficients are larger than the OLS ones.

[5] We abstract here from the case where $\alpha * \beta > 1$, i.e., when both the observable variables (y and o) as well as the residuals of 2a) and 2b) have infinite variance. This case is not very common and might arise in the presence of misspecification due to the omission of variables that are necessary to achieve cointegration.

For our purposes, if one has a prior that growth and openness positively affect each other, the reasoning above would lead to the expectation that the true coefficients are smaller than the respective OLS bounds. Note also, that each of the OLS estimates has a confidence interval. Hence, the exact bounds would need to take into account such intervals at the desired significance level.

B. Identification Through Heteroskedasticity

In this paper we appeal to a different methodology: identification through heteroskedasticity. In this section we derive the basic estimator closely following Rigobon (2003a). Assume we are interested in estimating first the following simultaneous equation system.

$$\begin{aligned} y_t &= \alpha\ o_t + \varepsilon_t \\ o_t &= \beta\ y_t + \eta_t \end{aligned}$$

where ε_t and η_t are the structural innovations. The first equation summarizes the growth equation we are interested in estimating. It measures the effect of openness on growth, and the structural residual can be interpretated as innovations to growth that are independent of all controls and other shocks. The second equation is the openness equation. It describes how growth affects the degree of openness of the economy. The innovations to this equation are interpreted as those changes in the degree of openness that are not explained by the covariates.

Assume that the innovations have mean zero, are uncorrelated, and identically distributed. Additionally, assume that the coefficients are the same across all realizations.

In this model, the only statistic we can compute from the sample is the covariance matrix of the observable variables—i.e., we can compute the variance of growth, the variance of the openness, and their covariance. However, under our assumptions this covariance matrix is explained by four unknowns: α, β, and the variances of ε_t and η_t. This is the standard identification problem in simultaneous equations—there are fewer equations (moments in this case) than the number of unknowns. Algebraically, the covariance matrix of the reduced form is

$$\Omega = \frac{1}{(1-\alpha\beta)^2}\begin{bmatrix} \sigma_\varepsilon^2 + \beta^2\sigma_\eta^2 & \alpha\sigma_\varepsilon^2 + \beta\sigma_\eta^2 \\ & \alpha^2\sigma_\varepsilon^2 + \sigma_\eta^2 \end{bmatrix}$$

where the left hand side can be estimated in the data and in the right hand side we have the theoretical moments.

Assume that the data can be split in two sets according to the heteroskedasticity of the residuals, i.e., that the residuals in these two sets have different variances. Remember that in the original model we have already imposed that the coefficients are the same across all observations. In these two sub-samples we can estimate two variance covariance matrices:

$$\Omega_1 = \frac{1}{(1-\alpha\beta)^2}\begin{bmatrix} \sigma_{\varepsilon,1}^2 + \beta^2\sigma_{\eta,1}^2 & \alpha\sigma_{\varepsilon,1}^2 + \beta\sigma_{\eta,1}^2 \\ & \alpha^2\sigma_{\varepsilon,1}^2 + \sigma_{\eta,1}^2 \end{bmatrix}$$

$$\Omega_2 = \frac{1}{(1-\alpha\beta)^2}\begin{bmatrix} \sigma_{\varepsilon,2}^2 + \beta^2\sigma_{\eta,2}^2 & \alpha\sigma_{\varepsilon,2}^2 + \beta\sigma_{\eta,2}^2 \\ & \alpha^2\sigma_{\varepsilon,2}^2 + \sigma_{\eta,2}^2 \end{bmatrix}$$

this implies that now there are six moments that can be estimated in the sample, which are explained by six coefficients: the two parameters of interest and four variances. Notice that there are as many equations as unknowns. In the standard literature of system of equations this means that the system satisfies the *order* conditions. To fully solve the problem, then, we have to verify that the six equations are linearly independent—which is known as the *rank* condition.

As is shown in Rigobon (2003a) this requires that the relative variances of the residuals shifts across the sub-samples:

$$\frac{\sigma_{\varepsilon,1}}{\sigma_{\eta,1}} \neq \frac{\sigma_{\varepsilon,2}}{\sigma_{\eta,2}}$$

The intuition why shifts in the variances achieve identification is quite simple and is closely related to the instrumental variable intuition. Consider the standard demand-supply identification problem. If we are interested in estimating the demand schedule, the IV methodology tells us that we need to find some variable or shock that shifts the supply schedule, so the slope of the demand can be computed. So, the standard intuition searches for something that moves the means.

In the IH methodology, we search for something (like a regime change) that shifts the variance instead of the mean. The different variances in the sample provide enough information to identify the coefficients in both equations (direct effect and reverse causality effect). The shift in the variance rotates the ellipse where the residuals are distributed. That rotation in the ellipse occurs along the schedules we are interested in estimating. Note that the analogy can be brought to the limit in the known case of near-identification. Assume that the variance of the supply shocks is infinitely large compared to the demand shocks. In this case, the ellipse enlarges along the demand so much that in the limit it becomes the demand. Therefore, even when we do not know when the supply moves, the likelihood that it is moving is one, hence all the variation observed is along the demand equation. This is known as near-identification, and even though there is no movement in the mean, the problem of identification has been solved.

As should be evident from the previous derivation a crucial assumption is the stability of the parameters. Even though this assumption seems unpalatable in many applications, in cross sectional panel regressions standard IV methods are implicitly assuming it already. One advantage of the method we describe is that if there are more than two regimes we can test the overidentifying restrictions.

The identification assumption in the IH methodology is that the data is heteroskedastic (which is easily testable) and that the structural shocks are uncorrelated (this is the maintained assumption). We estimate the model using GMM where the moment conditions are the zero correlation among the identified structural errors. On the other hand, estimation using MLE is more difficult given that it is hard to impose the moment condition – so crucial for identification.

C. Adding Control Variables

It should be obvious that adding additional controls has no impact on the identification problem. Assume that output and openness are described by this system of equations.

$$\begin{pmatrix} 1 & -\alpha \\ -\beta & 1 \end{pmatrix}\begin{pmatrix} y_t \\ o_t \end{pmatrix} = \phi(L)X_t + \Phi(L)\begin{pmatrix} y_t \\ o_t \end{pmatrix} + \begin{pmatrix} \varepsilon_t \\ \eta_t \end{pmatrix}$$
$$\textit{where} \qquad (3)$$
$$A \equiv \begin{pmatrix} 1 & -\alpha \\ -\beta & 1 \end{pmatrix}$$

where X are the controls or exogenous variables. In this model, if the variables are not cointegrated (which imposes some constraints on $\Phi(L)$) and we cannot impose exclusion restrictions on the exogenous variables (which imposes constraints on $\phi(L)$ not having a single term equal to zero), the problem of identification cannot be solved with standard methodologies.

The reason is that the reduced form in this model, which is

$$\begin{pmatrix} y_t \\ o_t \end{pmatrix} = A^{-1}\phi(L)X_t + A^{-1}\Phi(L)\begin{pmatrix} y_t \\ o_t \end{pmatrix} + \begin{pmatrix} \nu_{y,t} \\ \nu_{o,t} \end{pmatrix}$$
$$\textit{where} \qquad (4)$$
$$\begin{pmatrix} \nu_{y,t} \\ \nu_{o,t} \end{pmatrix} = A^{-1}\begin{pmatrix} \varepsilon_t \\ \eta_t \end{pmatrix}$$

cannot be identified from the data. To illustrate this issue assume that R is a positive definite matrix that we use to pre and post multiply each of the reduced form coefficients as follows:

$$\begin{pmatrix} y_t \\ o_t \end{pmatrix} = (AR)^{-1}(R\phi(L))X_t + (AR)^{-1}(R\Phi(L))\begin{pmatrix} y_t \\ o_t \end{pmatrix} + \begin{pmatrix} \nu_{y,t} \\ \nu_{o,t} \end{pmatrix} \qquad (5)$$

The moments from equation (4) and (5) are identical—and therefore, because there are no constraints on R other than positive definite, there exists a continuum of solutions for A that are consistent with the reduced form moments.

Notice that if we know—for example—that one coefficient of $\phi(L)$ is zero for some lag and some exogenous variable x, then pre-multiplying $\phi(L)$ for an arbitrary matrix R will violate the restriction. Therefore, if we know that one variable is not included in one of the equations, we can identify the system. This is exactly the intuition of exclusion restrictions—which in this case will imply that that variable can be used as an instrument.

One interesting aspect of the reduced form model is that $A^{-1}\phi(L)$ and $A^{-1}\Phi(L)$ are consistently estimated by using OLS. In other words, if we were to know what the matrix A is, then recovering all the coefficients of the structural model (3) would be trivial. In other words, the challenging problem is the estimation of matrix A. Notice that from equation (4) the reduced form residuals have the exact same properties as the endogenous variables. Hence, the easiest procedure is to use a two step estimator: First, it is possible to estimate the reduced form model (4)—which is similar to estimating a reduced form VAR—and recover their residuals. Second, those residuals can be used, then, to estimate the contemporaneous matrix A.

III. Results

We employ standard growth regression variables in a panel of 8 periods of 5 years each, spanning from 1961–65 to 1996–2000, and about 100 countries. The description of the variables and the corresponding main statistics are presented in the Appendix. In particular, in the IH estimation we use four measure of openness, whose sign is adjusted so that a high value means a more open regime: size of trade (share), a tariff indicator (tarind), import duties (impdut1p), and black market premium (bmp).[6] [7]

We perform some preliminary regressions of growth on the control variables and on openness with standard methodologies, such as fixed effect or difference-GMM, to derive some

[6] Vamvakidis (2002) included the black market premium among the control variables, rather than as a measure of openness. We regressions we present do not include black market premium in the list of control variables, as in several studies it has been used as a proxy for openness. Introducing it as a control variables in all the regression does not alter the thrust of the results.

[7] Edwards (1992, 1998) argues that the relation between openness and growth should be analyzed with as many measures of openness as possible and he uses nine of them. However, several such measures are unemployable in our setup as they are mainly available as a cross-section.

benchmark to assess the importance of properly controlling for endogeneity (the results are presented in the Appendix). The first method addresses the omitted variable bias by adding country-specific dummies, but cannot control for endogeneity. The second methodology implicitly accounts for fixed effects by estimating the model in first difference, but also attempts to address the endogeneity issue by instrumenting current variables with previous lags (we adopted the option that allows for all lags to be used). However, if variables are serially correlated, which is expected to be the case for openness and growth, lags are not a very good instrument. The results provide a weak evidence for the effect of openness on growth, as only the black market premium indicator is robustly associated with growth.

We now turn to the derivation of the OLS bounds discussed above and then to the implementation of the IH estimation.

A. Reversed Regression

Before estimating the IH coefficients it is instructive to analyze the bounds of the true coefficients following Leontief's reversed regressions. Controlling for fixed effects as well as the other standard variables, we computed the OLS and reversed OLS regressions for the four measures of openness: Share, Import Duties, Tariff Index, and Black Market Premium.[8] Remember from Section IIA that these two OLS estimators determine the bounds where either α or β belong. Indeed, this will be used in the second step to determine the validity of the identification. The OLS results are shown in Table 1; the results for α are of course identical to those in Table A2a in the Appendix. [9]

Table 1. OLS estimates with fixed effects

Measure of Openness	OLS eq. 2a ($\hat{a}_{2a}$)		OLS eq 2b ($\hat{b}_{2b}$)		Bounds for α	
	Point	T-stat	Point	T-stat	$\hat{a}_{2a}$	$1/\hat{b}_{2b}$
Share	0.0178	1.97	0.4583	1.97	0.0178	2.1822
Tariff Index	0.0553	0.70	0.0245	0.70	0.0553	40.8721
Import Duty	-0.0054	-0.19	-0.0263	-0.19	-0.0054	-38.0002
Black Market Premium	0.2116	5.39	0.3450	5.36	0.2116	2.8987

[8] The Sachs and Warner measure could not be employed because of the peculiar heteroskedasticity pattern that it would involve: high difference of variance across groups and minimal difference within groups.

[9] Given the presence of strong serial correlation, we run the regression also with lags (even though the literature on growth has generally not included lags). The thrust of the results is unchanged.

Notice that the point estimates for the effect of openness on growth (α) are marginally significant on the share variable, highly significant on the black market premium variable, and not significant for the others. Apart for the import duty measure (which is insignificant), they are all positive.

The last two columns of Table 1 show the bounds that the true coefficients have to satisfy.[10] As an example, let's consider the black market premium measure. The coefficients α or β have the same sign. Hence, according to Section IIA, the true coefficient would need to be smaller than the minimum of the two bounds, which in this case is 0.2116. Hence, if our IH estimator corrects properly for endogeneity, it will need to respect such a condition. The main message of Table 1 is that the bounds are very large, indicating that the endogeneity bias is potentially very large.

B. IH Estimation: Standard Setup

In this section we present the results from estimating the impact of openness on growth using the OLS estimates and the IH methodology.

The procedure is the following:

1. We estimate equation (4) and recover the residuals from the VAR. As was argued before the residuals share the exact same contemporaneous properties as our variables of interest. Initially, we will not introduce lags of growth and openness, to replicate standard growth specifications. We will subsequently add a dynamic structure to account for serial correlation.

2. We estimate the unconditional covariance matrix for each country and split the data into four groups: high-low variance of openness, and high-low variance of growth, where high and low values are defined with respect to the median. As was argued before we only need two different covariance matrices to solve the problem of identification. By appealing to four covariance matrices we have an overidentified system of equations and we can evaluate the validity of the model. In the robustness section we discuss the implications of the different splits.

3. Given the four covariance matrices we compute the contemporaneous coefficients by GMM, where the moment conditions for each regime are:

---

[10] Remember that each estimate has a confidence interval, which is not taken into account in the last two columns of Table 3, but would need to be taken into account to derive a precise bound. In our particular case, the estimates have the same sign. Hence, the precise bound would be higher—in absolute terms—than what is reported in the last two columns of Table 3 by the corresponding confidence interval at the desired significance level.

4. $$\Omega_i = \frac{1}{(1-\alpha\beta)^2}\begin{bmatrix} \sigma_{\varepsilon,i}^2 + \beta^2\sigma_{\eta,i}^2 & \alpha\sigma_{\varepsilon,i}^2 + \beta\sigma_{\eta,i}^2 \\ & \alpha^2\sigma_{\varepsilon,i}^2 + \sigma_{\eta,i}^2 \end{bmatrix}$$

5. To compute the standard errors we use the optimal weighting matrix for the GMM. We use a two-step estimation to compute the optimal weight. The estimation is as follows: From the sample we compute the covariance matrix from each of the group of countries. This provides 12 moments. These moments have to be explained with 10 parameters: 8 structural shock variances, and the two coefficients of interest. GMM reduces the weighted distance between these theoretical moments and the sample ones.

6. Alternatively, it is possible to specify the GMM by minimizing the identification assumption (that the structural shocks are uncorrelated) in the different regimes.

In Table 2, we present the estimates using the IH methodology. Because the IH methodology can estimate the impact of openness on growth and the impact of growth on openness we present both coefficients: the impact of openness on growth (α) and the impact of growth on the measurement of openness (β).

Table 2. IH estimates with fixed effects

Measure of Openness	α		β	
	Point	T-stat	Point	T-stat
Share	0.00396	1.12	0.47996	5.75
Tariff Index	0.07427	1.22	0.00997	0.38
Import Duty	-0.00083	-0.05	-0.08367	-1.31
Black Market	0.18119	7.12	0.13370	4.28

Let's analyze first the effect of openness on growth (α), which is the focus of the paper. When comparing the IH results (Table 2) and OLS ones (Table 1), we find that—for the measures that were significant in Table 1—the elimination of the endogeneity bias moves the point estimates in the direction which we would have expected. In fact, as discussed in Section IIA, for the share and black market premium measures the IH estimates should be smaller than the OLS estimates. Indeed, this is the case. The intuition lies in the fact that both variables affect each other positively and the OLS estimates capture a compounding of both effects.

Regarding the effect of growth on openness (β) for the same two measures, we find a positive and highly significant effect. These results confirm the prior that part of the positive effect of openness on growth found in the literature should actually be ascribed to the reverse causality—i.e., the one that goes from growth to openness.

The next step is to study how robust these results are to changes in the definitions of the regimes. Here, mostly, we have to estimate the coefficients using different splits. Before showing the results it is important to mention that, as it is argued in Rigobon (2003a), the estimates should be consistent to changes in the windows defining the splits. The intuition is the following: we have said that the system of equations is identified if the true data have heteroskedasticity. In other words, the heteroskedasticity provides additional equations that allow us to solve the problem of identification. Now, misspecifications of the windows (or splits) implies that the covariance matrices estimated in the data are linear combination of the true ones. This is the crucial step—that the mispecified model conforms matrices that are linear combination of the true ones! Hence, if the original system of equations has a solution, then the one from the mispecified splits is a linear combination of the original one. Therefore, the solution is the same as the one from the true system of equations, and there is only a loss in efficiency.

The standard split used in the text assumes that each country belongs to a particular group of variance, where high variance is defined as the variance above median. In a second split, we define high variance as the unconditional moment above the mean variance (this is a very small change in the definition of the windows). In a third split, we look at the different periods of 5 years: we compute the cross-sectional covariance for each 5-year period year and treat each 5-year window as a separate regime.[11] The last split is to use 5-year windows again but now group the window in four distinct groups of high-low variance of each of the two endogenous variables. The results are shown below in Table 3. The first two methods use the different volatilities across countries to create the groups, while the second two methods use the different volatility across time to split the data.

Table 3. IH estimates with fixed effects for different splits

	1	2	3	4
Share	0.003959	-0.005261	0.019170	0.018884
	1.12	-1.67	5.03	4.95
Tariff Index	0.074269	0.053254	0.045819	0.049875
	1.22	0.86	0.65	0.70
Import Duty	-0.000828	0.010274	-0.008037	-0.003624
	-0.05	0.67	-0.46	-0.20
Black Market	0.181195	0.184601	0.208993	0.210563
	7.12	7.70	7.25	7.34

t-stat below coefficients

Table 3 shows the estimates for the effect of the black market premium on growth for the four different splits. As can be seen, the estimates are generally close even though the splits involve very different arrangements of the data. Tariff index and import duty show significant stability of the coefficient across specifications.

[11] This corresponds to the methodology employed by Pattillo, Poirson, and Ricci (2003), with a program kindly provided by Rigobon.

The change in the split made the largest impact on the estimates for the Share variable. For such variable, the first two splits produce insignificant coefficients (although they have opposite signs), while the second two splits produce significant, positive, and similar coefficients. We find that first estimate is not statistically different from the other three (even though some are highly different from zero). However, it is possible to reject the hypothesis that the second coefficient is statistically the same to the third or fourth (at 5 percent confidence). This suggests a rejection of the model when the share variable is used in the estimation.
There are several possible reasons for this rejection when using share, which mainly pertain to specification issues. First, the omission of lags in the endogenous variables (to the extent these lags belong in the specification), which implies that a common shock is unaccounted for. Second, the relative importance of time-series variation versus cross-sectional variation, given that splits 1 and 2 rely mainly on the first variation and splits 3 and 4 on the second one. Other reasons relate to the existence of non-linearities, of a common shock, or of other endogenous variables unaccounted in the specification. We did not find rejections in the other three variables – which signals to us that there is something peculiar with the Share measure that requires further analysis. The next two Sections will further address this issue.

C. IH Estimation: Accounting for Serial Correlation

As was argued before in the OLS section, there are important serial correlations unaccounted for in the typical growth regression. In this section we evaluate how the estimates change when we include lags in the specification. This is not a standard growth specification but one that we were interested in exploring to be sure that the heteroskedasticity is not the result of an un-modeled lag structure.

Table 4a. OLS estimates with fixed effects for the first split

Measure of Openness	OLS eq. 2a ($\hat{a}_{2a}$)		OLS eq 2b ($\hat{b}_{2b}$)		Bounds for α	
	Point	T-stat	Point	T-stat	$\hat{a}_{2a}$	$1/\hat{b}_{2b}$
Share	0.0298	2.79	0.6317	2.76	0.0298	1.58
Tariff Index	0.1037	1.19	0.0462	1.19	0.1037	21.64
Import Duty	0.0482	1.43	0.1971	1.43	0.0482	5.07
Black Market Premium	0.1864	4.58	0.3478	4.55	0.1864	2.87

In Table 4a, we present the bounds for the case in which we allow fixed effects and lags. As can be seen, the same result as before is found. The implied bounds for the true coefficients are extremely large. For each of the openness measures it goes from less than 0.1 to more than 1.5. This indicates the severity of the endogeneity problem.

In Table 4b we present the results comparable to those in Table 2 with the split 1. The results are similar to those in Table 2, regarding the share and black market premium measure. However, the positive effect of openness as measure by tariff on growth becomes significant and the relation between import duty and growth becomes positive in both directions—although it remains insignificant.

Table 4b. IH estimates with fixed effects and lags in first step for the first split

Measure of Openness	α		β	
	Point	T-stat	Point	T-stat
Share	-0.00041	-0.10	0.65265	7.41
Tariff Index	0.18950	2.41	0.04326	1.10
Import Duty	0.02271	1.02	0.13105	1.52
Black Market	0.15881	5.59	0.13847	3.39

Table 4c reports the results comparable to those in Table 3: the effect of openness on growth, with the four splits.[12] As before, the results are very similar to those in Table 3 for share (different coefficients under the first two and the second two splits) and black market premium (very significant and positive). However, the coefficients of the tariff and import duty indices are now positive, much closer to significance, and stable.

Table 4c. IH estimates with fixed effects and lags in first step for the four splits

	1	2	3	4
Share	-0.000412	0.001675	0.028462	0.028051
	-0.10	0.39	5.82	5.71
Tariff Index	0.189496	0.152150	0.124494	0.123501
	2.41	1.95	1.47	1.46
Import Duty	0.022708	0.018349	0.037903	0.038206
	1.02	0.99	1.66	1.67
Black Market	0.158811	0.161419	0.184482	0.184519
	5.59	6.03	6.09	6.10

D. IH Estimation: Accounting for Serial Correlation and Cross-Sectional Variation

The previous two subsections show (Tables 3 and 4c) a puzzling result, which appears to be a rejection of the model when using share. The IH coefficients for the effect of share on openness appear to be different when using splits 1 and 2 or splits 3 and 4, whether lags are present or

[12] We estimated the model also using different normalizations and weighting matrices for the Generalized Method of Moments (GMM). The message from those specifications is almost identical to the one shown here (except, obviously, for the point estimates that change with each normalization). Those estimations are not shown in this paper and are available from the authors upon request.

absent in the specification. This puzzling result arises only when using share and not when using any of the other three measures of openness.

Notice that splits 1 and 2 split the data by country characteristics—i.e., countries with large variance in one variance go to some groups, and so on. Under criteria three and four we use time to split the data, either because every 5-year period is one group (split 3), or because 5-year periods with large variance in one variance go to one group (split 4), and so on. This means that in the splits 3 and 4, the IH estimation will rely mainly on the time-series heteroskedasticity to estimate the coefficients, while splits 1 and 2 will rely relatively more on cross-country variation.

As the regressions in the previous subsections were based on fixed effect, the main cross-sectional variation was automatically eliminated. This could imply that splits 1 and 2 with fixed effects would not allow the IH estimator to correct properly for endogeneity when the main source of variation is cross-sectional. It turns out that share, or the first-stage residuals for share (when not controlling for fixed effect), have much larger cross-sectional variation than time-series variation. This could explain why the problem of different coefficients for the splits 1 and 2 versus 3 and 4 was so pronounced with share. Splits 1 and 2 were unreliable for share as they could not pick its main source of variation, and heteroskedasticity is essential for the IH correction of endogeneity.

In order to avoid this problem, in this section we estimate the model using random effects. As the previous subsection has shown the importance of allowing for lags, we retain them in the specification. In Tables 5a and 5b we reproduce the OLS bounds and point estimates for the random effects case. As can be seen, the bounds continue to be extremely large. In Table 5c we present the results for all the different splits. Notice that, in comparison to the previous cases, now all variables, including share, have very stable coefficients. Openness as measured by share and black market premium have a positive and significant coefficient, even though the coefficients are smaller than the OLS estimates presented in Table 5a. Also observe that the tariff index and import duty have positive coefficients, although not significant and they are reasonably close to the results from Table 4c.

Table 5a. OLS estimates with lags and random effects for the first split

Measure of Openness	OLS eq. 2a ($\hat{a}_{2a}$)		OLS eq 2b ($\hat{b}_{2b}$)		Bounds for α	
	Point	T-stat	Point	T-stat	$\hat{a}_{2a}$	$1/\hat{b}_{2b}$
Share	0.0459	4.26	0.9345	4.18	0.0459	1.07
Tariff Index	0.1145	1.18	0.0412	1.17	0.1145	24.29
Import Duty	0.0255	0.77	0.1078	0.77	0.0255	9.27
Black Market Premium	0.1306	2.86	0.2015	2.86	0.1306	4.96

Table 5b. IH estimates with lags and random effects for the first split

Measure of Openness	α		β	
	Point	T-stat	Point	T-stat
Share	0.02706	6.83	0.82453	10.22
Tariff Index	0.13322	1.65	0.04807	1.64
Import Duty	0.01761	1.01	0.05726	0.87
Black Market	0.12945	4.75	0.07880	2.19

Table 5c. IH estimates with lags and random effects for the four splits

	1	2	3	4
Share	0.027060	0.027568	0.038623	0.039829
	6.83	6.88	9.34	9.61
Tariff Index	0.133220	0.118651	0.081517	0.096128
	1.65	1.48	1.00	1.18
Import Duty	0.017613	-0.001302	0.023872	0.023151
	1.01	-0.07	1.24	1.19
Black Market	0.129449	0.124598	0.129961	0.126375
	4.75	4.61	4.34	4.24

In summary, we find that the cross-sectional variation is crucial in understanding the effect of share on growth. However, it is important to mention that random effects allow us to use the cross-sectional variation, but may carry an omitted variable bias, as we are not controlling for country-specific effects. Nevertheless, at least we are sure we know that we are properly controlling for endogeneity and we cannot blame reverse causality (i.e., the effect of growth on openness) for the positive and significant result of share on growth.

Hence, Tables 4c and 5c should be jointly considered our preferred specifications. On the one hand, they both control for the dynamic structure. On the other hand, Table 4c properly controls for country-specific effects, while Table 5c ensures that the IH estimation works properly if we allow it to make use of the cross-sectional variation. When reading both Table 4c and 5c, we find that openness, as measured by share, black market premium, and, to a lesser extent, a tariff index, has a positive and significant effect on growth. Finally, and equally important, our estimates in Tables 4c and 5c are smaller (or not significantly larger) than those from simple OLS presented in Tables 4a and 5a, respecting the bounds conditions identified in Section IIA. This is exactly the direction we would have expected if endogeneity is a problem in the data.

IV. CONCLUSIONS

Academics and policymakers have devoted enormous energy to the question of whether openness is good for growth. Most of the evidence is based either on case studies or on regression analysis. We have learned a great deal in the last decades by studying both—but the question is still open. The main inconveniences are that case studies are always hard to replicate and are affected heavily by country idiosyncrasies, while regression analysis is plagued with endogeneity issues.

It should be clear, now, that instruments to solve the problem of simultaneous equations have been impossible to find in this case. Most of the literature moved toward proxies of openness—such as black market premium—as alternatives, with the unfortunate problem of finding variables that might be correlated with other inefficiencies not necessarily related to the degree of openness. The problem does not seem to have a solution within the standard econometric methodologies. Furthermore, the best available instrument, distance, used by Frankel and Romer (1999), cannot account for the time series variation of the openness variables and should also enter the growth equation directly because it can proxy for quality of institutions, etc. Hence, even the best instrument for openness available in the literature has several limitations.

In this paper we tackle the same question, using similar data, but resorting to a different procedure: identification through heteroskedasticity. This procedure uses instrumental variables that move the variances instead of the means. In the data, it is the case that the variation on second moments is richer than the variation on means, thus providing scope for using heteroskedasticity to estimate the contemporaneous relationships.

We find that most measures of openness would have a positive effect on growth, even when controlling for the effect from growth to openness. Furthermore, we also show that our estimates are smaller than the OLS estimates—exactly what we would have expected if endogeneity is an issue in the data.

Our results are robust to several specifications when openness is measured by trade over GDP and extremely robust when openness is measured by black market premium. However, as pointed out already by Rodriguez and Rodrik (2000) and Baldwin (2003), among others, black market premium is capturing not only openness but also reflects many other economic and policy distortions. Hence, the focus on the trade aspect of openness may be overstated. In other words, the extreme robustness of black market premium may suggest that it is openness in a broad sense—as part of the overall economic, policy, and institutional environment—that is conducive to growth. [13]

We estimated a very simple model in which other variables—notably all those which are typically available on a cross-section basis—have been excluded. Primarily, we have not

[13] Baldwin (2003) suggests that: "One can interpret openness in narrow terms to include only import and export taxes or subsidies as well as explicit nontariff distortions of trade or in varying degrees of broadness to cover such matters as exchange-rate policies, domestic taxes and subsidies, competition and other regulatory policies, education policies, the nature of the legal system, the form of government, and the general nature of institutions and culture."

included the quality of institutions in the estimation (Rigobon and Rodrik (2004) study this broader case) which could still potentially explain the positive correlation between openness and growth. Furthermore, we have treated some of the control variables as exogenous when some of them could perhaps be considered endogenous. Future research should extend the current methodology to include those aspects.

From the methodological point of view, this paper shows, once again, that the variation that exists in the data can be used to solve identification issues affected by endogeneity. It can therefore be used to investigate other unanswered questions in the growth literature, especially those related to the impact of policies, as these are likely to be dependent on the level of development and growth of a country. These questions are not only extremely important from the theoretical point of view, but they are crucial policy issues that need our attention.

APPENDIX

We mainly employ the dataset used recently by Vamvakidis (2002) and we complement it with two measures on trade openness from the Economic Freedom Network. The panel encompasses 8 periods of 5 years each, spanning from 1961-65 to 1996-2000, and about 100 countries. The main statistics of all variables are presented in Table A1.

Growth is measured by real GDP growth per capita (ypcg). The set of control variables encompasses initial real GDP per capita (ypc0), the ratio of investment to GDP (iy), the level of inflation (infl), the ratio of M2 to GDP (m2), the growth rate of population (popg), the natural logarithm of the level of education (Ledu, note that in the Table A1 the variable is presented in levels) and the age dependency ratio, i.e., the ratio of dependents to working-age population (age).

We have five openness measures: the ratio of the sum of imports and exports to GDP (share), import duties as a percentage of imports (impdut1p), the average years of openness indicated by the Sachs and Warner index (swyo), the difference between official exchange rate and black market rate (bmp) and a tariff indicator which is the average of revenue from taxes on international trade as a percentage of exports plus imports, the mean tariff rate, and the standard deviation of the tariff rates (tarind). The last two measures are from the Economic Freedom Network dataset. All openness measures are adjusted so that a high value means a more open regime.

The Appendix also reports the results that are obtained by regressing growth on the control variables and on openness, using standard methodologies such as fixed effects and difference GMM (Tables A2a and A2b).[14] The results provide a weak evidence for the effect of openness on growth. Only the black market premium indicator is robustly associated with growth (a positive coefficient represent negative effect of the premium on growth), while the Sachs and Warner indicator is significant only in the fixed effect estimation. Note that the Sachs and Warner index is highly dominated by its black market premium component, and is therefore likely to capture the same effect.

[14] We run the difference-GMM with the one step robust estimator (the one step is preferred for inference on the coefficients). The tests reject the null of no first order autocorrelation, but do not reject the null of no second order autocorrelation (in the presence of second order autocorrelation the estimates would be biased). The robust option does not deliver the Sargan test for the overidentifying restrictions. When we run the one step homoskedastic estimator, the Sargan test often rejects the null hypothesis that the overidentifying restrictions are valid. However, when we run the difference-GMM with the two step estimator (which would partially account for heteroskedasticity but is not reliable for inference on coefficients) or with lags (which would account for serial correlation but is not standard in growth literature), the Sargan test accepts the null hypothesis that the overidentifying restrictions are valid. We do not report such additional results, as heteroskedasticity and—to a smaller extent—serial correlation are discussed carefully in the paper.

Table A1. Descriptive Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
ypcg	779	2.03	2.94	-8.49	13.55
ypc0	762	7.51	1.57	4.58	10.74
iy	730	21.64	7.12	3.58	52.43
infl	696	34.50	203.73	-3.01	3357.61
m2	658	33.55	25.14	2.45	180.27
popg	808	1.93	1.12	-0.77	6.25
edu	767	44.00	32.70	1.00	147.83
age	801	0.77	0.18	0.37	1.21
share	752	64.36	44.32	5.71	393.75
bmp	555	7.50	3.48	0.00	10.00
tarind	521	5.83	2.59	0.00	10.00
swyo	736	0.42	0.48	0.00	1.00
impdut1p	439	-10.71	8.53	-47.90	0.00

Table A2a. Growth Regressions – OLS – Fixed Effects

share	0.0178				
	(1.64)				
tarind		0.0553			
		(0.62)			
swyo			0.8113		
			(2.02)**		
impdut1p				-0.0054	
				(0.18)	
bmp					0.2116
					(5.36)***
ypc0	-3.9256	-5.9618	-4.2697	-5.7131	-5.5037
	(6.62)***	(7.38)***	(6.94)***	(6.54)***	(7.33)***
iy	0.1639	0.1674	0.2032	0.2216	0.1669
	(5.86)***	(5.56)***	(7.72)***	(6.18)***	(5.74)***
infl	-0.0011	-0.0013	-0.0012	-0.0012	-0.0011
	(2.26)**	(2.80)***	(2.67)***	(2.99)***	(2.68)***
m2	-0.0129	0.0199	0.0104	0.0284	0.0236
	(0.74)	(1.18)	(0.70)	(1.73)*	(1.48)
popg	-0.0755	0.0685	-0.2970	-0.7784	0.0849
	(0.19)	(0.17)	(0.91)	(1.77)*	(0.24)
Ledu	-0.9522	-1.6403	-1.0918	-0.1224	-1.8478
	(2.92)***	(2.01)**	(3.05)***	(0.15)	(2.91)***
age	-5.7065	-8.9837	-3.9850	-3.2989	-6.6932
	(2.96)***	(3.28)***	(2.12)**	(1.24)	(3.11)***
Constant	45.9811	68.8750	46.6285	56.2517	62.1187
	(7.99)***	(7.72)***	(7.80)***	(6.32)***	(7.85)***
Observations	476	362	445	265	370
R-squared	0.52	0.59	0.54	0.56	0.60

Notes: Robust *t* statistics in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Table A2b. Growth Regressions – difference-GMM

LD.ypcgmm5	0.6192	0.5843	0.6153	0.4976	0.6130
	(9.58)***	(8.69)***	(8.26)***	(5.70)***	(9.08)***
D.share	0.0386				
	(1.54)				
D.tarind		0.1117			
		(0.89)			
D.swyo			0.9671		
			(1.28)		
D.impdut1p				0.0137	
				(0.34)	
D.bmp					0.1448
					(2.77)***
D.iy	0.1143	0.1836	0.1777	0.2396	0.1484
	(2.76)***	(3.92)***	(4.71)***	(4.11)***	(3.50)***
D.infl	-0.0014	-0.0005	-0.0009	-0.0015	-0.0009
	(1.79)*	(0.55)	(1.22)	(1.95)*	(1.32)
D.m2	-0.0092	0.0394	0.0488	0.0350	0.0410
	(0.24)	(0.96)	(1.26)	(1.23)	(1.06)
D.popg	-0.3421	-0.3495	-0.3362	-1.2418	-0.0733
	(0.95)	(0.70)	(1.07)	(2.76)***	(0.18)
D.Ledu	-3.2490	-3.7085	-2.8249	-1.9806	-2.6286
	(4.31)***	(3.62)***	(3.94)***	(2.37)**	(3.42)***
D.age	-5.6697	-6.4731	-8.1916	-1.6137	-8.0952
	(1.60)	(1.38)	(2.11)**	(0.38)	(2.06)**
Constant	0.6563	0.6483	0.3080	0.7743	0.3398
	(2.74)***	(2.52)**	(1.16)	(3.01)***	(1.43)
Observations	312	240	291	174	249
Number of NO	70	67	65	51	66

Notes: Robust *z* statistics in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

REFERENCES

Baldwin, Robert E., 2003, "Openness and Growth: What is the Empirical Relationship? NBER Working Paper No. 9578 (Cambridge: National Bureau of Economic Research).

Broda, C. and Weinstein D., 2004, "Globalization and the Gains from Variety," Federal Reserve Bank of New York, No. 180.

Caporale, G. M., Cipollini, A., and Demetriades, P., 2002a, "Monetary Policy and the Exchange Rate During the Asian Crisis: Identification Through Heteroskedasticity," CEMFE, *forthcoming*.

Caporale, G. M., Cipollini, A., and Spagnolo, N., 2002b, "Testing for Contagion: A Conditional Correlation Analysis," CEMFE, unpublished.

Dollar, David and Aart Kraay, 2003, "Institutions, Trade, and Growth," *Journal of Monetary Economics*, Vol. 50, pp. 133–62.

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

———, 1998, "Openness, Productivity and Growth: What Do We Really Know?" *Economic Journal,* Vol. 108, No. 447, pp. 383–98.

———, 1999, "How Effective Are Capital Controls?" *Journal of Economic Perspectives,* Vol. 13, No. 4, pp. 65–84.

Evans, Martin D.D. and Richard K. Lyons, 2003, "How is Macro News Transmitted to Exchange Rates?" NBER Working Paper No. 9433 (Cambridge: National Bureau of Economic Research).

Fisher, Franklin M., 1976, "The Identification Problem in Econometrics," (New York: Robert E. Krieger Publishing Co., 2nd ed.).

Frankel, Jeffrey A. and David Romer, 1999, "Does Trade Cause Growth?" *American Economic Review*, Vol. 89, No. 3, pp. 379–99.

Hogan, Vincent and Roberto Rigobon, 2002, "Using Unobserved Supply Shocks to Estimate the Returns to Education," NBER Working Paper No. 9145 (Cambridge: National Bureau of Economic Research).

Hviding, K., Nowak M., and Ricci L. A., 2004, "Can Higher Reserves Help Reduce Exchange Rate Volatility?" IMF Working Paper, unpublished.

Klein, R., and Vella F., 2003, "Identification and Estimation of the Triangular Simultaneous Equations Model in the Absence of Exclusion Restrictions Through the Presence of Heteroskedasticity," Rutgers University, mimeo.

Leamer, Edward, 1981, "Is It A Demand Curve, or Is It a Supply Curve? Partial Identification Through Inequality Constraints," *Review of Economics and Statistics*, Vol. 63, No. 3, pp. 319–27.

Leontief, W., 1929, "Ein Versuch zur Statistischen Analyse von Angebot und Nachfrage," Weltwirtschaftliches Archiv, Vol. 30, Heft 1, pp. 1–53.

Pattillo, Catherine, Hélène Poirson, and Luca A. Ricci, 2003, "Through What Channels Does External Debt Affect Growth," *Brookings Trade Forum* (Washington, DC: Brookings Institution Press) pp. 228–78.

Rigobon, Roberto, 2000, "A Simple Test for Stability of Linear Models Under Heteroskedasticity, Omitted Variable, and Endogenous Variable Problems," (September) MIT, mimeo.

———, 2002a, "The Curse of Non-Investment Grade Countries," *Journal of Development Economics*, Vol. 69, Issue 2, December, pp. 423–49.

———, 2002b, "Contagion: How to Measure It?" in *Preventing Currency Crises in Emerging Markets,* ed. by Sebastian Edwards and Jeffrey Frankel (Chicago: The University Chicago Press) pp. 269–334.

———, 2003a, "Identification Through Heteroskedasticity," *Review of Economics and Statistics*, Vol. 85, Issue 4 (Cambridge: The MIT Press) pp. 777–92.

———, 2003b, "On the Measurement of the International Propagation of Shocks: Are They Stable?" *Journal of International Economics*, Vol. 61, Issue 2, pp. 261–83.

——— and Dani Rodrik, 2004, "The Determinants of Institutions, Openness, and Income," MIT, mimeo.

Rigobon, Roberto and Brian Sack, 2004, "The Impact of Monetary Policy on Asset Prices," *Journal of Monetary Economics*, NBER Working Paper No. 8794 (Cambridge: National Bureau of Economic Research).

———, 2003a, "Spillovers Across U.S. Financial Markets," NBER Working Paper No. 9640 (Cambridge: National Bureau of Economic Research).

———, 2003b, "Measuring the Reaction of Monetary Policy to the Stock Market," *Quarterly Journal of Economics*, Vol. 118, Issue 2, pp. 639–69.

Rodríguez, Francisco and Dani Rodrik, 2000 "Trade Policy and Economic Growth: A Skeptic's Guide to the Cross-National Evidence," in *NBER Macroeconomics Annual 2000*, ed. by Ben S. Bernanke and Kenneth Rogoff (Cambridge: The MIT Press).

Sentana, Enrique, 1992, "Identification of Multivariate Conditionally Heteroskedastic Factor Models," FMG Discussion Paper No. 139 (London: London School of Economics).

———, and Gabriele Fiorentini, 2001, "Identification, Estimation and Testing of Conditional Heteroskedastic Factor Models," *Journal of Econometrics*, Vol. 102, No. 2, pp. 143–64.

Stock, James and Francesco Trebbi, 2003, "Who Invented Instrumental Variable Regression?" *Journal of Economic Perspectives,* Vol.17, pp. 177–94.

Vamvakidis, Athanasios, 2002, "How Robust is the Growth Openness Connection? Historical Evidence," *Journal of Economic Growth*, Vol. 7, pp. 57–80.

Wright, Philip Green, 1928, "The Tariff on Animal and Vegetable Oils," (New York: The Macmillan Company).