No. 12 june 2005

Jan-Egbert Sturm
Jakob de Haan

Determinants of Long-term Growth: New Results Applying Robust Estimation and Extreme Bounds

Research Paper Series
Thurgauer Wirtschaftsinstitut

THURGAUER
WIRTSCHAFTSINSTITUT
an der Universität Konstanz

Determinants of Long-term Growth: New Results Applying RobustEstimation and Extreme Bounds

Jan-Egbert Sturm
sturm@twi-kreuzlingen.ch
TWI and University of Konstanz

Jakob de Haan
J.de.Haan@eco.rug.nl
University of Groningen
and CESifo

ABSTRACT

june 2005

Two important problems exist in cross-country growth studies: outliers and model uncertainty. Employing Sala-i-Martin's (1997a,b) data set, we first use robust estimation and analyze to what extent outliers influence OLS regressions. We then use both OLS and robust estimation techniques in applying the Extreme Bounds Analysis (EBA) to deal with the problem of model uncertainty. We find that the use of robust estimation affects the list of variables that are significant determinants of economic growth. Also the magnitude of the impact of these variables differs sometimes under the various approaches.

Keywords: sensitivity analysis, outliers, economic growth, robust estimation

JEL: O4, C21, C52

Address: Jan-Egbert Sturm, Hauptstraße 90, 8280 Kreuzlingen, Switzerland

1. Introduction

There are two important problems in the empirical cross-country growth literature: (1) outliers and parameter heterogeneity and (2) model uncertainty. The second problem has been discussed extensively, the first problem has been largely ignored (Temple, 2000). [1]

Although economists engaged in estimating cross-country growth models often test the residuals of their regressions for heteroskedasticity and structural change, they hardly ever test for *unusual observations*. Still, their data sets probably frequently contain unusual observations. Some authors therefore suggest to use so-called robust estimation techniques (see, for example, Temple, 1998).[2] Robust estimators can be thought of as trying to seek out the most coherent part of the data, the part best approximated by the model being estimated. These estimators will not be led astray by outliers.

Employing the data set of Sala-i-Martin (1997a,b), we first use robust estimation to examine to what extent outliers influence OLS outcomes. We use the Least Trimmed Squares (LTS) estimator of Rousseeuw (1984, 1985) to identify outlying observations. This technique can cope extremely well with data sets containing outliers. The basic principle of LTS is to fit the majority of the data, after which outliers may be identified as those points that lie far away from the robust fit. Note that all available information, i.e. including the outlying observations, is used to come to a robust fit. LTS by itself is not suited for inference. As proposed by Rousseeuw (1984), this can be resolved by using re-weighted least squares (RWLS). We find that some variables become significant after the outliers are reweighted. Also in terms of their impact on growth, the results for several variables are sensitive to the inclusion of outliers.

As far as *model uncertainty* is concerned, the central difficulty is that several different models may all seem reasonable given the data, but yield different conclusions about the parameters of interest. Unfortunately, economic theory does not provide enough guidance for the proper specification of the empirical model. Various methods have been proposed to deal with this problem, including Extreme Bounds Analysis (EBA). The central idea of EBA is to run a whole range of possible models and to examine how sensitive parameter estimates are to different model specifications. The next step in the paper is that we replicate the results of Sala-i-Martin (1997a,b) employing both Leamer's (1983) and Sala-i-Martin's variants of the EBA using OLS.

---

[1] Excellent surveys of cross-country growth studies are provided in Durlauf and Quah (1999) and Temple (1999).

[2] As we will explain in somewhat more detail in Section 2, robustness in this case is defined here in terms of the observations included in the regression. Of course, model uncertainty and the role of outliers may be related as outliers can have consequences for the choice of variables. See also Temple (2000).

Recently, Temple (2000) has forcefully argued to use robust estimation techniques when applying EBA. This is the final step in our analysis. We apply both Leamer's (1983) EBA and Sala-i-Martin's (1997a,b) variant of EBA using robust estimation. We find that the EBA based on OLS and the EBA based on robust estimation yield similar lists of variables that are significant determinants of economic growth, even though the magnitude of the impact of these variables differs sometimes under both approaches. Based on the (preferred) Sala-i-Martin variant of the EBA estimated with robust estimation, we conclude that the following variables should be considered as important determinants of economic growth: number of years open, equipment investment, Latin-American dummy, Sub-Saharan dummy and fraction Muslim.

The remainder of this paper is organized as follows. Section 2 discusses the concepts of outlying observations and robust estimation in greater detail. Section 3 shows how sensitive OLS regressions in cross-country growth models are to outliers. Section 4 reviews the problem of model uncertainty, while section 5 presents the results of the EBA analysis, using both OLS and robust estimation. Section 6 compares the outcomes of the EBA analyses focusing on the ranking of the variables in terms of their significance and impact on economic growth and offers some concluding remarks.

2. Outliers, robust estimation and parameter heterogeneity

In the usual presentation of outliers it is stressed that one or more observations may be measured with a substantial degree of error. Barro (2000) argues that in particular less developed countries tend to have a lot of measurement error in national accounts and other data. This may have affected the conclusions of cross-country growth models. As Schwartz and Welsch (1986, p. 171) put it: "OLS and many other commonly used maximum likelihood techniques have an unbounded influence function; any small subset of the data can have an arbitrarily large influence on their coefficient estimates. In a world of fat-tailed or asymmetric error distributions, data errors and imperfectly specified models, it is just those data in which we have the least faith that often exert the most influence on the OLS estimates."

Following Barnett and Lewis (1994, p. 316) we define an outlier as an observation 'lying outside' the typical relationship between the dependent and explanatory variables revealed by the remaining data. For instance, point A in figure 1(a) is clearly an outlier. Outliers in the dependent variable – i.e. in the y-direction – often possess large positive or large negative residuals, which are easy to detect by plotting the residuals.[3] Observations may be outlying for several reasons. The most obvious one involves problems with the quality of the data. Outliers in the explanatory variables may be more problematic than outliers in the dependent variable. As figure 1(b) shows, an unusual observation in the *x*-direction (B) can actually tilt the OLS regression line. In such a case we call the outlier a (bad) leverage point. Note that looking at the OLS residuals cannot discover bad leverage points. If a leverage point tilts the regression line, deleting the points with the largest OLS residuals implies that some 'good' points would be deleted instead of the 'bad' leverage point.

[INSERT FIGURE 1 ABOUT HERE]

Basically, there are two ways to deal with outliers: regressions diagnostics and robust estimation. Diagnostics are certain statistics mostly computed from the OLS regression estimates with the purpose of pinpointing outliers and leverage points.[4] Often the unusual observations are then removed or corrected after which an OLS analysis on the remaining observations follows. When there is only one unusual observation, some of these methods work quite well. However, single-case diagnostics are well known to be inadequate in the presence of multiple outliers or leverage points (Temple, 2000).

Take for instance figure 1c. Deleting either of the two outliers will have little effect on the regression outcome and will therefore not be spotted by the single-case diagnostics. The potential effect of one outlying observation is clearly masked by the presence of the other. Testing for groups of observations to be influential might solve this masking effect problem. However, a serious problem in the multiple observation case is how to determine the size of the subset of jointly influential observations. Suppose we are interested in detecting all subsets of size m=2,3,…, of observations that are considered to be jointly outliers and/or high-leverage. A sequential method might be useful, but where to stop? In the multiple observation case the number of possible subsets for which each diagnostic measure of interest can be computed is: $\frac{n!}{m!(n-m)!}$ where n is number of observations. For m=5 and n=50 this results in over 2 million

[3] Note, however, that if x_i is near the centre of the set of explanatory observations, as is the case in figure 1(a), it will mainly affect the constant and hardly alter the slope.
[4] See, for instance, Belsley et al. (1980) and Chatterjee and Hadi (1988).

diagnostics. Therefore we prefer so-called robust regression techniques that employ estimators that are not strongly affected by (groups of) outliers.

Two closely related methods are the Least Median of Squares (LMS) and Least Trimmed Squares (LTS) introduced by Rousseeuw (1984). LMS minimizes the median of the squared residuals. LTS typically minimizes the sum of squares over half the observations, the chosen half being the combination which gives the smallest residual sum of squares.[5] According to Temple (2000), LTS is generally thought preferable to LMS.[6]

As shown by Rousseeuw (1984), robust estimators have an abnormally slow convergence rate and hence perform poorly from the point of view of asymptotic efficiency. Because of its low finite-sample efficiency, LTS is not suited for inference. As proposed by Rousseeuw (1984), this can be resolved by using re-weighted least squares (RWLS). A simple, but effective, way is to put weight zero if the observation is an outlier and weight one otherwise.

When we take a parameter heterogeneity perspective, it is clear that we can think about outliers in another way (Temple, 2000). Some observations may be entirely correct, but drawn from a different regime. In most research, it is implicitly assumed that only one regime generates the data, or that the parameters in the different regimes vary randomly. However, both assumptions may not be correct (see e.g. Durlauf and Johnson, 1995). Some authors therefore suggest to use so-called robust estimation techniques to deal with these problems (Temple, 1998). Robust estimators can be thought of as trying to seek out the most coherent part of the data, the part best approximated by the model being estimated. We can illustrate this by the following example. Figure 2 shows a data set for two variables (x and y) in which 40 per cent of the (50) observations follow a different distribution than the rest of the observations. Assume a researcher would not know this and would simply estimate a linear relationship between x and y. As the OLS model assumes all observations are drawn from one single distribution, the OLS regression line estimated by this researcher as shown in figure 2 will not reveal any valuable information. In contrast, the LTS regression line looks for a linear relationship, which fits the majority of the data. As 60 per cent of the data set follows a linear relationship, LTS will reveal that relationship. The remaining 40 per cent of the observation will have large negative residuals which are easily

---

[5] Note that the choice of minimising half of the observations can be altered and the proportion increased if required (Temple, 2000).

[6] Robust estimation techniques are increasingly being used in practice, e.g. in finance, chemistry, electrical engineering, process control and computer vision (Meer et al., 1991). For a survey of robust methods and some applications, see Rousseeuw (1997). Still, robust estimators have been criticised. See e.g. the

depicted by graphing the standardized residuals. As this example shows, that does not mean that those observations should be ignored and simply thrown away. Those observations reveal that the linear model is not adequate for the entire data set as not all observations follow the same regime.

[INSERT FIGURE 2 ABOUT HERE]

3. Outlier Sensitivity

This section presents the results of a cross-country growth model using robust estimation as outlined in the previous section. In the analysis we stick as closely as possible to Sala-i-Martin (1997a,b), both in terms of the variables taken into account and in using his data set.[7] As in Sala-i-Martin (1997a,b), the variables in the basic model are the level of income in 1960, life expectancy in 1960 and the primary-school enrollment rate in 1960.[8] For comparison purposes we first apply OLS. The results for the basic model are shown in column (2) of Table 1, while column (1) shows some statistics (mean and standard deviation). Then the LTS regression technique is used to detect outlying countries, i.e. countries that do not follow the general pattern of the data. After having detected the outliers, we apply reweighted least squares (RWLS). We reweigh those countries whose robust residual is greater than 2.5 times the robust standard error.[9,10] As explained in the previous section, the identified outliers are dropped. The results for the basic model are shown in column (3) of Table 1.

[INSERT TABLE 1 ABOUT HERE]

---

exchange between Rousseeuw (1993) and Hettmansperger and Sheather (1992) and Stromberg (1993). See also Ruppert and Simpson (1990) and Rousseeuw et al. (1999) for further contributions.

[7] The data set is available at http://www.columbia.edu/~xs23/data/millions.htm.

[8] This set of variables differs from our preferred specification (see e.g. De Haan and Sturm, 2000). Still, as we want to stay as closely as possible to Sala-i-Martin's model, we employ his specification.

[9] In order to decide whether a residual is large, we have to compare it with an estimate of σ, which has to robust itself. For this purpose we have used: $\sigma = c(1+\frac{5}{n-k})\sqrt{\frac{1}{h}\sum_{i=1}^{h}(e^2)_{i:n}}$ where $(e^2)_{1:n} \leq \ldots \leq (e^2)_{n:n}$ are the ordered squared residuals, n is the number of observation, k is the number of regressors, and c is set at 1.4826.

[10] To check the robustness of all our findings to this somewhat arbitrary yardstick we have also used others. The main results are not very sensitive to this and, hence, the qualitative conclusions do not change. All results are available on request.

It follows from Table 1 that the results for the basic model are not very sensitive to the inclusion of outliers: the OLS estimation outcomes are very similar to the LTS/RWLS estimates. Note that LTS identifies 5 countries that follow a different pattern than the bulk of the other countries.[11] However, dropping these countries affects the coefficients only marginally. The coefficient of initial income is somewhat lower in absolute terms in the LTS/RWLS estimates, while the coefficient of the school enrollment rate is somewhat higher and estimated with more precision. In order to check to what extent the choice of weight zero for outliers has influenced our results, we have also experimented with a weight of 0.5 in the RWLS estimates; this gave very similar results (which are available on request).

Next, each of the other 59 variables used by Sala-i-Martin (1997a,b) is added as an additional explanatory variable to the basic model. Table 2 shows the results. Columns (3)-(5) show the OLS results (number of observations, estimated coefficients and *t*-statistics) of adding a particular variable to the basic model, while columns (8)-(10) present the LTS/RWLS outcomes (number of identified outliers, estimated coefficients and *t*-statistics).

In Table 2, the variables are ordered by their absolute *t*-values in the RWLS regressions. In column (6) the ranking on the basis of the *t*-statistics of the OLS regression are shown. In other words, here we focus on the statistical significance of the estimates. However, from a growth perspective not only the precision of the estimates – as part of the statistical significance – is relevant, but especially the magnitude of the impact of the variables. In columns (7) and (11) we therefore show the ranking of the variables in terms of standardized coefficients, i.e. the ranking of the explanatory variables in terms of the growth effects of a one standard deviation change in the variables.[12]

[INSERT TABLE 2 ABOUT HERE]

In the full sample we have 103 countries. As not all variables are available for all countries, the sample is sometimes reduced to even only 65. As follows from column (8) of Table 2, the number of outlying observations as indicated by the LTS technique varies between 1 and 15.

There are some noteworthy differences between the OLS and RWLS estimates. For instance, the variable terms of trade growth is definitely not significant in the standard OLS equation. However, after reweighing the outlying countries it becomes highly significant. More or less the same holds for the following variables: democratic freedom, ratio workers to population,

[11] These outlying countries are Botswana, Guyana, Hong Kong, Singapore and South Korea.

inflation, the fraction of the population speaking a foreign language, revolutions and coups and the public consumption share. Apparently, the insignificance of these variables in the OLS regression reflects the influence of a limited number of outlying observations. Interestingly, the reverse situation (significant in OLS, insignificant in LTS/RWLS) does not occur.

Also in terms of the size of the estimated coefficients there are some noteworthy differences between the OLS and the LTS/RWLS estimation results. The effect of the standard deviation of the black market exchange rate premium, for instance, is much higher according to the robust estimation outcomes than according to the OLS results. Similar results show up for the following variables: fraction Buddhist, non-equipment investment, democratic freedom, public consumption share and primary exports in 1970. Except for the latter variable, the coefficients found with robust estimation are higher than those in the OLS regressions.

4. Model uncertainty

The basic problem of model uncertainty is that economic theory does not provide enough guidance to select the proper specification of an empirical growth model. Or, as Brock and Durlauf (2001), put it: growth theories are open-ended, i.e. the validity of one causal theory of growth does not imply the falsity of another. Sala-i-Martin (1997a,b) identifies, for instance, around 60 variables that have been suggested to be correlated with economic growth. Levine and Renelt (1992) and Sala-i-Martin (1997a,b) investigate the 'robustness' of regressions by checking how sensitive the estimated coefficient of each variable of interest is to the inclusion of additional explanatory variables. Using different methodologies they come up with different sets of explanatory variables which are robustly related to economic growth.[13] Levine and Renelt (1992) conclude that only very few regressors pass their version of Leamer's (1983, 1985) Extreme Bounds Analysis (EBA).

The EBA can be exemplified as follows. Equations of the following general form are estimated:

---

[12] We thank one of the referees for this suggestion.

[13] As pointed out by Temple (2000), robustness of a variable (in the sense that its significance is not depending on the choice of conditioning variables) is neither a necessary nor a sufficient condition for an interesting finding. Especially if causality is indirect (e.g. a variable affects investment or human capital), a finding that a variable is fragile in a growth model should be interpreted extremely carefully. Furthermore,

$$Y= \alpha M + \beta F + \gamma Z + u \quad (1)$$

where Y is the dependent variable; M is a vector of 'standard' explanatory variables; F is the variable of interest; Z is a vector of up to three (here we follow Levine and Renelt, 1992) possible additional explanatory variables, which according to the literature may be related to the dependent variable; and u is an error term. The extreme bounds test for variable F says that if the lower extreme bound for β – i.e. the lowest value for β minus two standard deviations – is negative, while the upper extreme bound for β – i.e. the highest value for β plus two standard deviations – is positive, the variable F is not robustly related to Y.

Sala-i-Martin (1997a,b) rightly argues that the test applied in the extreme bounds analysis is too strong for any variable to really pass it. If the distribution of the parameter of interest has some positive and some negative support, then one is bound to find one regression for which the estimated coefficient changes sign if enough regressions are run. Instead of analyzing the extreme bounds of the estimates of the coefficient of a particular variable, Sala-i-Martin (1997a,b) suggests to analyze the entire distribution of the estimates of the parameter of interest. Broadly speaking, if the averaged 90 per cent confidence interval of a regression coefficient does not include zero, Sala-i-Martin classifies the corresponding regressor as a variable that is strongly correlated with economic growth. He identifies a much larger set of 'robust' variables by looking at the distribution of the coefficient estimates. In our empirical analysis we will use both versions of the EBA. However, as we agree with the critique of Sala-i-Martin (1997a,b) on Leamer's version of the EBA, our conclusions will be based on the Sala-i-Martin variant of the EBA.

Recently, Doppelhofer et al. (2000) proposed a so-called Bayesian Averaging of Classical Estimates (BACE) approach to check the robustness of different explanatory variables in growth regressions. [14] This approach builds upon the approach as suggested by Sala-i-Martin (1997a,b) in the sense that different specifications are estimated (by OLS) to check the sensitivity of the coefficient estimate of the variable of interest. The major innovation of BACE as compared to the

---

a robust variable may not be very interesting as robustness is defined in terms of significance of coefficients. A robust variable may therefore be of little quantitative importance.

[14] Other authors have come up with other alternatives. For instance, Fernández et al. (2001) adopt a Bayesian framework that explicitly allows for the specification uncertainty. Their findings broadly support the more 'optimistic' conclusion of Sala-i-Martin (1997a,b) and they find a roughly similar set of variables that can be classified as 'important' for growth regressions. Hoover and Perez (2001) apply a cross-sectional version of the general-to-specific search methodology associated with the LSE approach to econometrics, which leads to a different and smaller set of variables as compared to Sala-i-Martin (1997a,b). Kalaitzidakis et al. (2002) propose another more direct specification testing approach and compare it with the EBA. They find that both methods are complementary. None of these methods takes the problem of outliers into account.

Sala-i-Martin's approach is that there is no set of fixed variables included and the number of explanatory variables in the specifications is flexible. This is where the Bayesian nature comes in: a prior needs to be specified with respect to the number of explanatory variables likely to be optimal in economic growth model. Furthermore, to penalize the use of additional regressors, Doppelhofer et al. (2000) use weights based on the Schwarz criterion. The biggest disadvantages of the BACE approach are the need of having a balanced data set, i.e. an equal number of observations for all regressions (due to the chosen weighting scheme), the restriction of limiting the list of potential variables to be less than the number of observations and the computational burden.[15] Furthermore, this approach does not take the problem of outliers and parameter heterogeneity into account. In the next section we will therefore employ the EBA, using both OLS and LTS/RWLS. As Temple (2000, p. 195) argues: "It is well known in the statistics literature that the presence of a few influential outliers can either hide a relationship, or create the appearance of one where none exists… This suggests that any good approach to model uncertainty should ideally be robust to observations that are measured with error, or drawn from a different regime… I propose using a simple variant of EBA in which each regression is first estimated by robust methods".

5. Extreme Bounds Analysis using OLS and robust estimation

We first apply the two versions of the EBA using OLS. The results are shown in Table 3. For Leamer's variant of the EBA, we not only report the extreme bounds, but – given the critique of Leamer's EBA as discussed the previous section – also the percentage of the regressions in which the coefficient of the variable F is significantly different from zero at the 5 percent level. Following Temple (2000), we also checked the results of various diagnostic tests of the models yielding the extreme bounds. To be precise, we performed the following tests on the regressions yielding the extreme bounds: Jarque-Bera (1987) test on normality of residuals, White's (1980) test for general heteroskedasticity and the Ramsey RESET test of functional form (see Ramsey, 1969; Granger and Terasvirta, 1993; Lee et al., 1993).[16]

---

[15] With k explanatory variables there are 2^k possible combinations to be tested. For k=32 (as in Doppelhofer et al., 2000), this boils down to around 4.3 billion regressions; an infeasibly large number. Doppelhofer et al. (2000), therefore, apply a so-called random sampling algorithm, which reduces the number of regressions to approximately 21 million.

[16] The results of these tests are not reported, but are available on request.

As pointed out, instead of analyzing only the extreme bounds of the estimates of the coefficient of a particular variable in Leamer's variant of the EBA, we follow Sala-i-Martin's suggestion to analyze the entire distribution. We report the unweighted parameter estimate of β and its standard deviation, as well as the outcomes of the cumulative distribution function (CDF) test. The CDF test is based on the fraction of the cumulative distribution function lying on each side of zero. CDF(0) indicates the larger of the areas under the density function either above or below zero; in other words, regardless of whether this is CDF(0) or 1-CDF(0). So CDF(0) will always be a number between 0.5 and 1.0. However, in contrast to Sala-i-Martin, we use the unweighted instead of the weighted CDF(0).[17] The variables in Table 3 are ordered on the basis of the CDF(0). We not only focus on the significance but also on the impact of the variables on economic growth. Therefore, in column (7) the ranking of the variables on the basis of standardized coefficients is shown.

[INSERT TABLE 3 ABOUT HERE]

It follows from Table 3 that a list of 16 variables should be considered as robustly related to economic growth if we apply the criterion that CDF(0) > 0.95.[18] Note that according to Leamer's (1983) EBA, quite a few of these variables should not be considered as robust.[19] In fact, only the variable "fraction Confucian" is robust according to Leamer's EBA. A variable like "equipment investment" is not robust according to this test even though its coefficient is significantly different from zero in almost all regressions that we ran. This further illustrates how stringent Leamer's EBA is.

Even though we favor Sala-i-Martin's variant of the EBA, the coefficients of some variables that are considered to be robust are significantly different from zero in only slightly more (and for the standard deviation of the variable "black-market premium" even less than) 50 per cent of the regressions. If we add as an additional criterion that the estimated coefficients should be significantly different from zero in at least 90 per cent of the regressions, the following

---

[17] Sala-i-Martin (1997a) proposes using the (integrated) likelihood to construct a weighted CDF(0). However, the varying number of observations in the regressions due to missing observations in some of the variables poses a problem. Sturm and De Haan (2002) show that as a result this goodness of fit measure may not be a good indicator of the probability that a model is the true model and the weights constructed in this way are not equivariant for linear transformations in the dependent variable. Hence, changing scales will result in rather different outcomes and conclusions. We therefore restrict our attention to the unweighted version.

[18] Note that Sala-i-Martin considers a variable to be robust if the CDF(0) > 0.90, but we consider this to be too low given the one-sidedness of the test.

[19] The Jarque-Bera, White and RESET tests for the extreme bound regressions are all insignificant.

variables should be considered as important determinants of economic growth: fraction Confucian, equipment investment, number of years open, fraction Buddhist, primary exports 1970 and the rule of law. Of course, this additional criterion is rather arbitrary and may be replaced by a more or less stringent criterion, which will affect the list of variables that are considered to be important determinants of economic growth.

It is quite interesting that the ranking of the variables in terms of their impact on economic growth (as shown in column (7) of Table 3) is quite similar to the ranking based on the CDF(0). In fact, the correlation coefficient of the rankings based on CDF(0) and standardized coefficients is 0.93. Of the variables that we identified to be important determinants of economic growth, only the fraction Buddhist has a relatively small impact.

As explained in the previous section, Temple (2000) suggests to use robust estimation when the Extreme Bounds Analysis is employed.[20] The idea to combine robust estimation and EBA – either in Leamer's version or, preferably, Sala-i-Martin's variant of it – is intuitively very appealing. However, it is very (computer) time intensive. Therefore, we apply this method to a limited number of variables. We also use somewhat less control variables, i.e. 45/44 instead of 58. Variables which are highly correlated with variables in the basic model or other potential explanatory variables have been deleted. The results of the EBA based on robust estimation are shown in Table 4.

[INSERT TABLE 4 ABOUT HERE]

It follows from Table 4 that 13 variables have a CDF(0) > 0.95, so that these variables should be considered as robustly related to economic growth. However, the coefficients of some of these variables – notably democratic freedom and political rights – are often insignificant. Adding again the criterion that the coefficients should be significantly different from zero in at least 90 per cent of the regressions, the following variables should be considered as important determinants of economic growth: number of years open, equipment investment, Latin-American dummy, Sub-Saharan dummy and fraction Muslim. It is interesting that this list includes investment, which is often considered to be sensitive to outliers.

The ranking in terms of the standardized coefficient of the variables is again very similar to the ranking based on the CDF(0). Exceptions include the standard deviation of the black market

---

[20] Using the data set of Levine and Renelt (1992), he finds that the bounds for the variable real exchange rate distortions are narrowed sufficiently that classifying the relationship as 'fragile' may no longer be justified.

premium (CDF(0) rank 18, standardized coefficient rank 8) and – as opposed to the results presented in Table 3 – fraction Buddhist (CDF(0) rank 10, standardized coefficient rank 2). The correlation of both rankings amounts to 0.86.

Finally, as pointed out by one of the referees, some of the variables in our data set have a 0-1 range which may have an effect on the LTS algorithm. Table 5 therefore shows the results if we drop all variables that have large number of observations equal to zero. Our cut-off point is 75%. The variables concerned are: Fraction Confucian, Fraction Jewish, Fraction Buddhist, Fraction Hindu, dummy indicating that the country is a previous French colony, dummy indicating that the country is a previous Spanish colony, Age, Fraction of the population speaking English, dummy for Latin American countries, and the number of political assassinations. The results in Table 5 are very similar to those reported in Table 4.

[INSERT TABLE 5 ABOUT HERE]

6. Discussion and Concluding Remarks

As argued by Temple (2000), the EBA is an intuitively appealing way to deal with the problem of model uncertainty. As we prefer Sala-i-Martin's variant of the EBA, our conclusions concerning the variables that are robust determinants of economic growth are based on this version of the EBA. As the OLS regression technique is very sensitive to outlying observations, in theory one country might drive all the results. We therefore prefer robust estimation.

[INSERT TABLE 6 ABOUT HERE]

The columns in the left-hand side of Table 6 show the ranking of the variables according to their CDF(0) using OLS and LTS/RWLS. The rankings are remarkably similar. With a few exceptions ("fraction Confucian", "fraction Buddhist" and "fraction Catholic") the ranking of the variables with a high CDF(0) do not differ more than 5 positions.

The columns in the right-hand side of Table 5 show the rankings based on standardized coefficients. The rankings are again remarkably similar. Variables that have diverging ranking are "fraction Confucian", "fraction Buddhist" and "standard deviation of the black market premium". Note, however, that some of the estimated coefficients differ under both approaches. For instance, the coefficient of equipment investment is lower according to the LTS/RWLS estimates than under the OLS estimates, while the reverse result shows up for non-equipment investment.

We conclude that it is useful to employ the EBA using robust estimation instead of OLS. This approach can deal with the problems of outliers and parameter heterogeneity and model uncertainty. Our analysis suggests that some variables that are considered to be robust according to the EBA based on OLS are not robust according to EBA based on LTS/RWLS, while the impact of some variables also differs somewhat under both approaches.

References

Barnett, Vic and Toby Lewis (1994), *Outliers in Statistical Data*, New York: John Wiley and Sons.

Barro, Robert J. (2000), “Inequality and Growth in a Panel of Countries”, *Journal of Economic Growth*, 5 (1), 5-32.

Belsley, David A., Edwin Kuh and Roy E. Welsch (1980), *Regression Diagnostics*, New York: John Wiley and Sons.

Brock, W.A and S.N. Durlauf (2001), “What Have We Learned From a Decade of Empirical Research on Growth? Growth Empirics and Reality”, *World Bank Economic Review*, 15(2), 229-272.

Chatterjee, Samprit and Ali S. Hadi (1988), *Sensitivity Analysis in Linear Regression*, New York: John Wiley and Sons.

Doppelhofer, G., R.I. Miller and X. Sala-i-Martin (2000), “Determinants of Long-term Growth: A Bayesian Averaging of Classical estimates (BACE) Approach”, NBER Working Paper, No. 7750.

Durlauf, S.N. and P.A. Johnson (1995), “Multiple regimes and cross-country growth behaviour”, *Journal of Applied Econometrics*, 10, 365-84.

Durlauf, S.N. and Danny T. Quah (1999), “The new empirics of economic growth”, in Taylor, J. and M. Woodford (eds), *Handbook of Macroeconomics*, North-Holland, New York.

Fernández, Carmen, Eduardo Ley and Mark F.J. Steel (2001), “Model uncertainty in cross-country growth regressions”, *Journal of Applied Econometrics*, 16, 563-576.

Granger, C.W.J. and T. Terasvirta (1993), *Modelling nonlinear economic relationships*, Oxford: Oxford University Press.

Haan, Jakob de and Jan-Egbert Sturm (2000), “On the relationship between economic freedom and economic growth”, *European Journal of Political Economy*, 16, 215-241.

Hettmansperger, Thomas P. and Simon J. Sheather (1992), A cautionary note on the method of Least Median Squares, *American Statistician*, 46:2, 79-83.

Hoover, Kevin D. and Stephen J. Perez (2001), " Evaluating Specification Search in Cross-country Growth Regressions: Extreme Bounds vs. General-to-Specific Methodologies", unpublished manuscript.

Jarque, C.M. and A.K. Bera (1987), "A Test for Normality of Observations and Regression Residuals", *International Statistical Review*, 55, 163-172.

Kalaitzidakis, Pantelis, Theofanis P. Mamuneas and Thanasis Stengos (2002), "Specification and Sensitivity Analysis of Cross-Country Growth Regressions", *Empirical Economics*, 27, 645-656.

Leamer, Edward E. (1983), "Let's take the Con Out of Econometrics", *American Economic Review*, 73(3), 31-43.

Leamer, Edward E. (1985), "Sensitivity Analysis Would Help", *American Economic Review*, 75, 308-313.

Lee, T.H., H. White and C.W.J. Granger (1993), "Testing for neglected non-linearities in time series models: a comparison of neural network methods and alternative tests", *Journal of Econometrics*, 56, 269-290.

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

Meer, P., D. Mintz, A. Rosenfeld and D. Kim (1991), "Robust regression methods in computer vision: a review", *International Journal of Computer Vision*, 6, 59-70.

Ramsey, J. (1969), "Tests for specification errors in classical least-squares regression analysis", *Journal of the Royal Statistical Society B*, 31, 350-371.

Rousseeuw, Peter J. (1984), "Least Median of Squares Regression." *Journal of the American Statistical Association*, 79, 871-880.

Rousseeuw, Peter J. (1985), "Multivariate Estimation with High Breakdown Point." in Grossmann, W., G. Pflug, I. Vincze and W. Wertz (eds), *Mathematical Statistics and Applications*, Dordrecht: D. Reidel Publishing Co.

Rousseeuw, P.J. (1993), "Comment on Hettmansperger and Sheather", *American Statistician*, 47, 162-163.

Rousseeuw, Peter J. (1997), Introduction to positive-breakdown methods. In: Maddala, G.S. and C.R. Rao (eds), *Handbook of Statistics*, Amsterdam: Elsevier.

Rousseeuw, Peter J., S. van Aelst and M. Hubert (1999), “Rejoinder to discussion of Regression depth”, *Journal of the American Statistical Association*, 94, 419-433.

Ruppert, David and Douglas G. Simpson (1990), “Comment”, *Journal of the American Statistical Association*, 85, 644-646.

Sala-i-Martin, Xavier (1997a), “I Just Ran Four Millions Regressions.” Mimeo, Columbia University.

Sala-i-Martin, Xavier. (1997b), “I Just Ran Two Millions Regressions.” *American Economic Review*, 87(2), pp. 178-183.

Schwartz, S. and Roy E. Welsch (1986), “Applications of Bounded-influence and Diagnostic Methods in Energy Modeling.” In: Belsley, David A. and Edwin Kuh (eds), *Model Reliability*, Cambridge MA: MIT Press.

Stromberg, A.J. (1993), “Computing the exact least median of squares estimate and stability diagnostics in multiple linear regression”, *SIAM Journal of Scientific Computing*, 14, 1289-1299.

Sturm, Jan-Egbert and Jakob de Haan (2002), How robust is Sala-i-Martin’s robustness analysis?, mimeo.

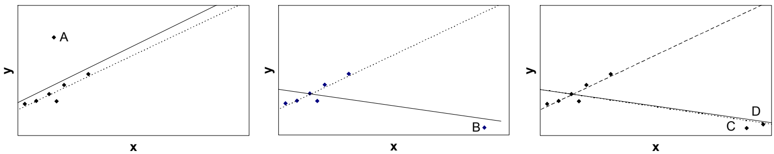
Temple, Jonathan (1998), “Robustness Tests of the Augmented Solow Model”, *Journal of Applied Econometrics*, 13, 361-375.

Temple, Jonathan (1999), “The New Growth Evidence”, *Journal of Economic Literature*, 37, 112-156.

Temple, Jonathan (2000), “Growth Regressions and What the Textbooks Don't Tell You”, *Bulletin of Economic Research*, 52 (3), 181-205.

White, H. (1980), “A heteroskedasticity-consistent covariance matrix and a direct test for heteroskedasticity”, *Econometrica*, 48, 817-838.

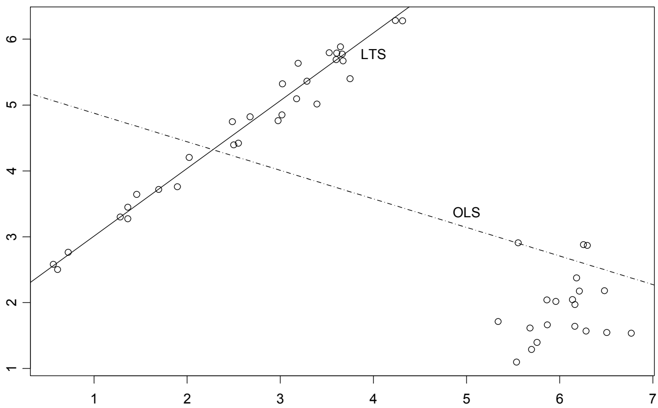
Figure 1: Outlying observations and bad leverage points



(a) (b) (c)

The solid lines represent the OLS estimates including the unusual observation(s). The dotted lines represent the OLS estimates without the unusual observations A, B, or C. The dashed line represents the OLS estimate without observations C and D.

Figure 2: Hypothetical example



Hypothetical data set for series x and y of 50 observations. The first 30 observations have the following characteristics: series x is uniformly distributed between 0.5 and 4.5. Series y is distributed around the line y=2+x+e, where e is normally distributed with mean zero and standard deviation 0.2. The remaining 20 observations follow a Normal distribution: x~N(6,0.5) and y~N(2,0.5). The solid line represents the LTS estimate, whereas the dotted line represents the OLS estimate.

Table 1: Descriptive statistics and estimation results of the base model

	(1) Mean (st.dev.)	(2) OLS	(3) LTS/RWLS
GDP per capita growth 1960-1992	1.73		
	(1.81)		
Constant		4.30	3.24
		(3.18)	(2.73)
log GDP per capita in 1960	7.33	-1.47	-1.20
	(0.89)	(-5.17)	(-4.79)
Life expectancy in 1960	53.09	0.13	0.11
	(12.25)	(4.92)	(4.76)
Primary School Enrollment in 1960	0.71	1.65	1.79
	(0.30)	(1.95)	(2.42)
Adjusted R^2		0.42	0.46
Number of observations		103	103
Outliers / reweighted observations			5

Table 2: Descriptive statistics and estimation results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
				OLS					LTS/RWLS		
Variable	Mean	St.dev.	Obs.	coeff.	*t*-stat.	rank *t*-stat.	rank impact	Outl.	coeff.	*t*-stat.	rank impact
1 Latin American dummy	0.21	0.41	103	-1.18	3.63	6	9	11	-1.59	6.10	8
2 Fraction Muslim	0.21	0.34	103	1.33	2.97	10	11	11	2.02	5.92	6
3 Number of years open economy	0.35	0.35	103	2.60	5.50	2	2	3	2.46	5.66	3
4 Fraction Buddhist	0.04	0.16	103	2.91	3.31	7	13	8	8.34	5.29	2
5 Fraction Confucian	0.01	0.07	103	8.03	4.56	3	6	3	8.42	5.17	10
6 Rule of law	0.52	0.33	92	2.29	4.03	4	3	4	2.57	4.94	4
7 Equipment investment	0.04	0.03	82	27.94	5.57	1	1	6	23.04	4.83	5
8 Non-Equipment investment	0.14	0.06	82	8.32	2.69	12	10	7	11.87	4.70	7
9 S.D. of black-market premium	0.45	0.89	95	-0.33	2.08	17	20	15	-2.05	4.38	1
10 Democratic freedom	0.62	0.29	93	-0.96	1.46	29	23	11	-2.18	4.11	9
11 Spanish colony dummy	0.16	0.36	103	-0.86	2.27	16	18	10	-1.27	4.06	17
12 Sub-Saharan dummy	0.32	0.47	103	-0.90	2.33	15	15	7	-1.23	3.86	12
13 Liquid liabilities	0.32	0.19	65	3.12	3.02	9	5	5	3.16	3.82	11
14 Absolute latitude	0.23	0.16	103	3.07	2.67	13	8	9	3.20	3.41	14
15 Public consumption share	0.10	0.07	96	-3.70	1.66	25	24	7	-6.82	3.09	15
16 Terms of trade growth	0.05	2.62	89	-0.02	0.36	48	49	9	-0.16	2.96	18
17 Fraction Catholic	0.35	0.37	103	-1.24	3.24	8	12	6	-0.99	2.89	24
18 Degree of capitalism	3.35	1.49	103	0.26	2.69	11	16	1	0.27	2.87	21
19 Exchange rate distortions	1.24	0.40	102	-0.75	1.97	18	19	6	-0.95	2.84	22
20 Political rights	3.96	2.00	103	-0.25	2.36	14	7	7	-0.26	2.83	13
21 War dummy	0.39	0.49	102	-0.52	1.79	21	26	8	-0.63	2.56	28
22 Ratio workers to population	-0.92	0.24	100	-0.31	0.51	44	46	8	-1.36	2.46	25
23 Primary exports 1970	0.73	0.28	100	-2.21	3.63	5	4	7	-1.32	2.29	23
24 Average inflation	0.24	0.60	98	-0.25	1.05	34	37	6	-0.66	2.21	20
25 Fraction speaking foreign language	0.30	0.40	103	0.47	1.31	30	34	7	0.68	2.19	31
26 Revolutions and coups	0.20	0.25	103	-0.89	1.51	28	30	3	-1.22	2.15	29
27 Fraction of pop. speaking English	0.08	0.25	103	-1.15	1.92	20	22	8	-1.04	2.10	34
28 Civil liberties	3.91	1.82	103	-0.24	1.95	19	14	6	-0.22	2.07	19
29 Public defense share	0.03	0.02	96	9.90	1.74	22	29	5	11.12	2.03	32
30 S.D. of inflation rate	0.05	0.24	98	-0.29	0.49	46	47	6	-2.01	1.94	16
31 Outward orientation	0.36	0.48	102	0.36	1.26	31	36	3	0.45	1.74	36
32 Ethnolinguistic fractionalization	0.42	0.29	98	-0.16	0.28	52	50	10	0.81	1.64	35
33 Fraction Protestant	0.16	0.23	103	-1.10	1.73	24	27	9	-0.77	1.54	40
34 Black Market Premium	0.17	0.33	93	-0.38	0.86	37	40	3	0.97	1.45	26
35 Urbanization rate	0.34	0.25	101	0.80	0.78	38	33	3	1.26	1.34	27
36 Tariff restrictions	0.03	0.02	83	-5.47	0.70	42	41	2	-9.20	1.25	38
37 Secondary school enrollment	0.21	0.21	100	1.22	0.93	35	25	4	1.43	1.24	30
38 Size labor force	7.42	21.40	100	0.00	0.74	39	44	6	0.01	1.15	42
39 Average years of primary schooling	2.72	1.90	85	-0.17	1.09	32	17	8	-0.14	1.11	33
40 Free trade	0.23	0.07	83	4.02	1.56	26	21	2	2.81	1.10	37
41 Population growth	0.02	0.01	103	0.05	0.00	59	59	5	-20.42	1.03	39
42 Public investment share	0.05	0.04	95	2.53	0.71	40	43	4	3.04	0.98	44
43 Age	0.21	0.36	103	-0.69	1.74	23	28	7	-0.31	0.88	45
44 Political assassinations	0.01	0.04	95	-2.90	0.92	36	39	6	-2.13	0.80	46
45 Human capital * GDP per capita	0.27	0.22	85	-0.82	0.59	43	35	5	-0.79	0.69	41

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
			OLS						LTS/RWLS		
Variable	Mean	St.dev.	Obs.	coeff.	*t*-stat.	rank *t*-stat.	rank impact	Outl.	coeff.	*t*-stat.	rank impact
46 S.D. of domestic credit	0.36	0.74	88	-0.06	0.30	50	53	4	-0.12	0.67	48
47 Average years of education	3.36	2.45	85	-0.06	0.50	45	38	7	-0.05	0.55	43
48 Political instability	0.06	0.11	95	-0.39	0.29	51	52	6	0.61	0.51	49
49 Fraction GDP in mining	0.05	0.08	103	2.77	1.54	27	31	7	1.09	0.44	47
50 Domestic credit growth	0.23	0.21	85	0.17	0.26	53	54	6	0.20	0.38	51
51 French colony dummy	0.18	0.39	103	0.09	0.22	55	55	4	0.12	0.34	50
52 Average years of higher education	0.07	0.09	94	0.88	0.48	47	45	7	0.44	0.30	52
53 Public education share	0.02	0.01	98	11.17	0.71	41	42	5	-3.05	0.21	53
54 Area	0.84	1.77	101	0.00	0.04	57	57	4	0.01	0.20	54
55 Fraction Hindu	0.04	0.14	103	-0.01	0.01	58	58	7	0.15	0.18	55
56 Higher education enrollment	0.03	0.05	102	1.08	0.26	54	48	5	0.39	0.11	56
57 British colony dummy	0.34	0.48	103	0.05	0.18	56	56	6	-0.02	0.08	57
58 Average years of secondary schooling	0.54	0.61	93	0.35	1.06	33	32	8	0.01	0.05	58
59 Fraction Jewish	0.01	0.08	103	-0.58	0.34	49	51	6	0.05	0.03	59

Columns (1) and (2) show the mean and standard deviation, respectively, of the explanatory variables. Columns (3)-(5) show the OLS results (number of observations, estimated coefficients and *t*-statistics) of adding a particular variable to the basic model, while columns (8)-(10) present the LTS/RWLS outcomes (number of identified outliers, estimated coefficients and *t*-statistics). The variables are ordered by their absolute *t*-values in the RWLS regressions. In column (6) the ranking on the basis of the *t*-statistics of the OLS regression is shown. In columns (7) and (11) the ranking of the variables in terms of standardized coefficients is shown.

Table 3: Extreme Bounds Analysis using OLS

	Leamer EBA		Sala-i-Martin EBA				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Variable	Extreme bounds Lower	Upper	%significant at 5%	Unweighted CDF(0)	Unweighted Beta	Unweighted St.error	Rank stand.impact
1 Fraction Confucian	0.38	12.66	100.00	1.00	7.35	1.76	6
2 Equipment investment	-3.96	52.68	99.97	1.00	25.96	5.56	1
3 Number of years open economy	-0.25	4.38	99.97	1.00	2.32	0.50	2
4 Fraction Buddhist	-1.42	5.54	92.67	0.99	2.51	0.86	14
5 Primary exports 1970	-4.50	1.67	93.73	0.99	-1.93	0.64	5
6 Rule of law	-1.24	5.99	92.19	0.99	2.17	0.62	3
7 Fraction Muslim	-1.55	3.68	88.95	0.99	1.52	0.49	7
8 Latin American dummy	-3.54	1.12	88.50	0.99	-1.13	0.36	9
9 Fraction Catholic	-3.05	1.20	84.38	0.98	-1.13	0.41	13
10 Sub-Saharan dummy	-3.77	1.74	76.83	0.98	-1.16	0.43	4
11 Non-Equipment investment	-6.33	24.68	76.72	0.98	7.57	3.19	11
12 Absolute latitude	-4.18	8.68	66.72	0.97	2.72	1.21	10
13 Liquid liabilities	-3.35	8.02	69.25	0.96	2.46	1.10	8
14 Fraction Protestant	-4.80	1.72	57.81	0.96	-1.39	0.67	19
15 Exchange rate distortions	-2.91	1.34	55.07	0.96	-0.88	0.40	15
16 S.D. of black-market premium	-0.96	0.32	36.50	0.96	-0.29	0.16	24
17 Degree of capitalism	-0.36	0.89	51.71	0.94	0.22	0.11	18
18 Political rights	-1.34	0.77	34.49	0.93	-0.22	0.13	12
19 Fraction of pop. speaking English	-3.50	1.66	28.14	0.92	-0.99	0.61	26
20 Public consumption share	-27.54	10.40	36.24	0.92	-4.69	2.54	17
21 Age	-2.26	1.26	15.90	0.92	-0.63	0.41	27
22 Democratic freedom	-4.02	2.33	18.21	0.89	-0.99	0.70	23
23 Spanish colony dummy	-2.58	2.86	46.13	0.89	-0.68	0.41	25
24 Public defense share	-21.20	32.69	16.80	0.88	8.34	5.94	30
25 War dummy	-1.68	1.26	17.63	0.87	-0.42	0.31	29
26 Civil liberties	-1.13	1.30	17.41	0.85	-0.17	0.15	20
27 Fraction speaking foreign language	-1.62	2.45	19.72	0.84	0.51	0.39	31
28 Free trade	-10.42	14.90	4.01	0.82	2.79	2.64	32
29 Average years of primary schooling	-351.24	797.11	4.05	0.81	-0.18	0.19	16
30 Average years of secondary schooling	-350.62	797.94	3.05	0.80	0.35	0.36	28
31 Outward orientation	-0.97	1.57	2.69	0.79	0.28	0.30	38
32 Ratio workers to population	-4.91	4.56	22.02	0.77	-0.78	0.71	36
33 Average inflation	-10.05	4.64	6.67	0.75	-0.30	0.30	37
34 Urbanization rate	-4.38	5.20	1.12	0.75	0.79	1.06	34
35 Size labor force	-0.03	0.08	0.44	0.74	0.01	0.01	40
36 Fraction GDP in mining	-16.54	13.61	17.69	0.73	1.66	2.23	39
37 S.D. of domestic credit	-3.57	1.50	4.86	0.71	-0.26	0.28	35
38 Secondary school enrollment	-5.98	7.71	2.52	0.71	0.94	1.37	33
39 Black Market Premium	-2.28	3.81	3.51	0.71	-0.30	0.53	42
40 Revolutions and coups	-3.77	3.58	2.98	0.70	-0.40	0.68	41
41 Public investment share	-23.09	27.14	1.24	0.69	2.19	3.92	44
42 Human capital * GDP per capita	-54.72	39.59	2.33	0.69	-1.29	2.09	22
43 Public education share	-68.54	75.19	0.56	0.68	8.74	16.42	45
44 Political assassinations	-18.33	15.87	0.02	0.68	-1.63	3.17	48
45 Tariff restrictions	-57.15	43.44	0.55	0.67	-4.13	8.17	43

	Leamer EBA			Sala-i-Martin EBA			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Variable	Extreme bounds Lower	Upper	%significant at 5%	Unweighted CDF(0)	Unweighted Beta	Unweighted St.error	Rank stand.impact
46 French colony dummy	-2.38	3.06	0.12	0.65	0.19	0.45	47
47 Average years of higher education	-350.38	797.26	0.01	0.64	0.77	1.88	46
48 Terms of trade growth	-0.34	0.23	0.06	0.63	-0.02	0.06	50
49 Average years of education	-797.39	351.02	0.89	0.62	0.00	0.20	58
50 Fraction Jewish	-184.90	101.18	0.17	0.62	-3.79	5.49	21
51 Ethnolinguistic fractionalization	-3.02	2.53	0.68	0.61	-0.21	0.60	49
52 British colony dummy	-1.80	1.33	1.29	0.61	-0.11	0.31	51
53 Political instability	-6.84	10.24	0.32	0.59	0.40	1.45	52
54 S.D. of inflation rate	-16.64	23.22	3.91	0.56	-0.03	0.79	59
55 Higher education enrollment	-18.33	23.23	0.02	0.54	0.56	4.38	53
56 Domestic credit growth	-5.47	8.48	0.02	0.54	0.07	0.80	57
57 Area	-0.31	0.43	0.03	0.54	0.01	0.08	55
58 Population growth	-128.17	116.51	0.31	0.53	-2.17	24.61	54
59 Fraction Hindu	-31.00	10.18	0.52	0.52	-0.12	1.10	56

Columns (1)-(3) show the outcomes of Leamer's variant of EBA (lower and upper extreme bounds and percentage of the regression in which the variable is significantly different from zero). Columns (4)-(6) show the outcomes of Sala-i-Martin's variant of EBA (unweighted CDF(0), estimated coefficient and standard error). Variables are ordered on the basis of the CDF(0). In column (7) the ordering of the variables on the basis of standardized coefficients is shown.

Table 4: Extreme Bounds Analysis using robust estimation (LTS/RWLS)

	Leamer EBA			Sala-i-Martin EBA			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Variable	Extreme bounds Lower	Upper	%significant at 5%	Unweighted CDF(0)	Unweighted Beta	Unweighted St.error	Rank stand.impact
1 Number of years open economy	-0.79	5.05	99.54	1.00	2.13	0.45	4
2 Equipment investment	-12.68	54.81	97.85	1.00	22.10	5.15	3
3 Fraction Muslim	-6.71	5.49	98.75	1.00	1.97	0.39	7
4 Latin American dummy	-4.41	3.66	93.21	0.99	-1.33	0.32	9
5 Rule of law	-2.25	7.20	89.81	0.99	2.22	0.57	5
6 Sub-Saharan dummy	-5.58	2.77	92.26	0.98	-1.46	0.37	6
7 Fraction Confucian	-7,500.43	3,813.67	86.42	0.98	34.59	18.14	1
8 Non-Equipment investment	-8.60	25.68	87.11	0.98	9.31	2.69	10
9 Primary exports 1970	-5.22	4.12	77.94	0.97	-1.60	0.59	12
10 Fraction Buddhist	-842.37	630.63	89.57	0.97	5.60	2.91	2
11 Fraction Protestant	-6.88	2.12	70.74	0.97	-1.39	0.56	18
12 Liquid liabilities	-3.30	9.31	76.72	0.95	2.46	0.88	11
13 Exchange rate distortions	-3.75	2.15	66.30	0.95	-0.83	0.35	17
14 Political rights	-1.29	0.82	48.48	0.94	-0.20	0.09	14
15 Absolute latitude	-5.83	11.36	70.36	0.93	2.73	1.06	13
16 Fraction Catholic	-3.63	1.53	68.39	0.93	-0.92	0.36	16
17 Democratic freedom	-6.05	3.38	42.89	0.93	-1.03	0.49	19
18 S.D. of black-market premium	-11.77	6.88	59.77	0.92	-0.74	0.27	8
19 Degree of capitalism	-0.64	1.02	44.53	0.90	0.18	0.10	21
20 Spanish colony dummy	-4.34	3.63	60.99	0.89	-0.78	0.36	20
21 Age	-2.82	1.15	19.08	0.88	-0.51	0.35	24
22 Public consumption share	-33.22	15.43	52.40	0.86	-4.99	2.37	15
23 Public defense share	-34.63	57.56	28.10	0.84	8.08	5.48	23
24 War dummy	-1.90	1.63	19.38	0.78	-0.29	0.27	26
25 Terms of trade growth	-0.46	0.37	40.61	0.78	-0.08	0.05	22
26 Public investment share	-27.88	36.12	21.67	0.74	3.83	3.68	25

Columns (1)-(3) show the outcomes of Leamer's variant of EBA (lower and upper extreme bounds and percentage of the regression in which the variable is significantly different from zero). Columns (4)-(6) show the outcomes of Sala-i-Martin's variant of EBA (unweighted CDF(0), estimated coefficient and standard error). Variables are ordered on the basis of the CDF(0). In column (7) the ordering of the variables on the basis of standardized coefficients is shown.

Table 5: Extreme Bounds Analysis using robust estimation (LTS/RWLS), excl. variables that have large number of observations equal to zero

	Leamer EBA			Sala-i-Martin EBA			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Variable	Extreme bounds		%significant	Unweighted	Unweighted	Unweighted	Rank
	Lower	Upper	at 5%	CDF(0)	Beta	St.error	Stand.impact
1 Number of years open economy	-0.81	5.12	99.36	1.00	2.12	0.45	3
2 Equipment investment	-16.77	51.92	99.03	1.00	22.21	5.21	2
3 Fraction Muslim	-7.01	4.73	99.06	1.00	2.01	0.40	4
4 Rule of law	-2.45	7.41	92.06	0.99	2.37	0.57	1
5 Sub-Saharan dummy	-5.53	2.76	88.87	0.98	-1.41	0.38	5
6 Non-Equipment investment	-7.47	22.82	90.69	0.99	9.52	2.70	7
7 Primary exports 1970	-4.86	3.80	80.00	0.97	-1.66	0.59	10
8 Fraction Protestant	-5.83	1.96	68.42	0.96	-1.38	0.56	17
9 Liquid liabilities	-3.61	9.67	81.73	0.98	2.60	0.89	9
10 Exchange rate distortions	-4.01	1.70	62.08	0.93	-0.81	0.36	16
11 Political rights	-1.19	0.90	49.34	0.90	-0.21	0.10	12
12 Absolute latitude	-6.14	9.32	66.33	0.91	2.57	1.08	11
13 Fraction Catholic	-3.55	1.56	72.73	0.93	-0.96	0.35	13
14 Democratic freedom	-7.28	6.58	48.74	0.92	-1.17	0.57	15
15 S.D. of black-market premium	-10.54	6.94	56.50	0.92	-0.70	0.26	6
16 Degree of capitalism	-0.62	1.03	43.93	0.90	0.18	0.10	18
17 Age	-2.44	1.11	17.18	0.87	-0.50	0.36	21
18 Public consumption share	-30.88	17.66	48.50	0.84	-4.89	2.44	14
19 Public defense share	-34.12	50.29	25.30	0.83	7.64	5.53	20
20 War dummy	-1.86	1.97	21.47	0.77	-0.29	0.28	23
21 Terms of trade growth	-0.47	0.32	39.20	0.79	-0.08	0.05	19
22 Public investment share	-26.94	34.70	23.58	0.74	4.03	3.78	22

Columns (1)-(3) show the outcomes of Leamer's variant of EBA (lower and upper extreme bounds and percentage of the regression in which the variable is significantly different from zero). Columns (4)-(6) show the outcomes of Sala-i-Martin's variant of EBA (unweighted CDF(0), estimated coefficient and standard error). Variables are ordered on the basis of the CDF(0). In column (7) the ordering of the variables on the basis of standardized coefficients is shown.

Table 6: A comparison of the EBA estimation results

	Ranking according to CDF(0):		Ranking according to impact:	
	(1)	(2)	(3)	(4)
Variable	EBA with OLS(CDF(0))	EBA with LTS/RWLS(CDF(0))	EBA with OLS(coef.)	EBA with LTS/RWLS(coef.)
Number of years open economy	**3 (1.00)**	**1 (1.00)**	2 (2.32)	4 (2.13)
Equipment investment	**2 (1.00)**	**2 (1.00)**	1 (25.96)	3 (22.10)
Fraction Muslim	7 (0.99)	**3 (1.00)**	7 (1.52)	7 (1.97)
Latin American dummy	8 (0.99)	**4 (0.99)**	9 (-1.13)	9 (-1.33)
Rule of law	**6 (0.99)**	5 (0.99)	3 (2.17)	5 (2.22)
Sub-Saharan dummy	10 (0.98)	**6 (0.98)**	4 (-1.16)	6 (-1.46)
Fraction Confucian	**1 (1.00)**	7 (0.98)	6 (7.35)	1 (34.59)
Non-Equipment investment	11 (0.98)	8 (0.98)	11 (7.57)	10 (9.31)
Primary exports 1970	**5 (0.99)**	9 (0.97)	5 (-1.93)	12 (-1.60)
Fraction Buddhist	**4 (0.99)**	10 (0.97)	14 (2.51)	2 (5.60)
Fraction Protestant	14 (0.96)	11 (0.97)	18 (-1.39)	18 (-1.39)
Liquid liabilities	13 (0.96)	12 (0.95)	8 (2.46)	11 (2.46)
Exchange rate distortions	15 (0.96)	13 (0.95)	15 (-0.88)	17 (-0.83)
Political rights	18 (0.93)	14 (0.94)	12 (-0.22)	14 (-0.20)
Absolute latitude	12 (0.97)	15 (0.93)	10 (2.72)	13 (2.73)
Fraction Catholic	9 (0.98)	16 (0.93)	13 (-1.13)	16 (-0.92)
Democratic freedom	21 (0.89)	17 (0.93)	19 (-0.99)	19 (-1.03)
S.D. of black-market premium	16 (0.96)	18 (0.92)	20 (-0.29)	8 (-0.74)
Degree of capitalism	17 (0.94)	19 (0.90)	17 (0.22)	21 (0.18)
Spanish colony dummy	22 (0.89)	20 (0.89)	21 (-0.68)	20 (-0.78)
Age	20 (0.92)	21 (0.88)	22 (-0.63)	24 (-0.51)
Public consumption share	19 (0.92)	22 (0.86)	16 (-4.69)	15 (-4.99)
Public defense share	23 (0.88)	23 (0.84)	24 (8.34)	23 (8.08)
War dummy	24 (0.87)	24 (0.78)	23 (-0.42)	26 (-0.29)
Terms of trade growth	26 (0.63)	25 (0.78)	26 (-0.02)	22 (-0.08)
Public investment share	25 (0.69)	26 (0.74)	25 (2.19)	25 (3.83)

Columns (1) and (2) show the rankings of the variables based on their CDF(0) in the EBA estimated by OLS and LTS/RWLS, respectively. The CDF(0) is shown in parentheses. Variables in bold have CDF(0) of at least 0.95 and are also significant in more than 90 per cent of the regressions. Columns (3) and (4) show the rankings of the standardized coefficients according to the EBA estimated by OLS and LTS/RWLS, respectively. The (unscaled) estimated coefficients are shown in parentheses.

Already published

No.	Title	Authors
1	IMF and Economic Growth: The Effects of Programs, Loans, and Compliance with Conditionality	Axel Dreher
2	Do gasoline prices converge in a unified Europe with non-harmonized tax rates?	Axel Dreher, Tim Krieger
3	Is There A Causal Link between Currency and Debt Crisis?	Axel Dreher, Bernhard Herz, Volker Karb
4	What Determines Differences in Foreign Bank Efficiency? Australien Evidence	Jan-Egbert Sturm, Barry Williams
5	Market oriented institutions and policies and economic growth: A critical survey	Jakob de Haan, Susanna Lundström, Jan-Egbert Sturm
6	Does Globalization Affect Growth? Evidence from a new Index of Globalization	Axel Dreher
7	Transformation nicht-gehandelter in handelbare Kreditrisiken	Günter Franke
8	Student Flows and Migration: An Empirical Analysis	Axel Dreher, Panu Poutvaara
9	Foreign Exchange Intervention and the Political Business Cycel: A Panel Data Analysis	Axel Dreher, Roland Vaubel
10	M&A-Transaktionen - Fluch oder Segen der Realoptionstheorie?	Günter Franke, Christian Hopp
11	Wie hat sich die universitäre volkswirtschaftliche Forschung in der Schweiz seit Beginn der 90er Jahre entwickelt?	Miriam Hein

THURGAUER
WIRTSCHAFTSINSTITUT
an der Universität Konstanz

Hauptstr. 90
CH-8280 Kreuzlingen 2

Telefon: +41 (0)71 677 05 10
Telefax: +41 (0)71 677 05 11

info@twi-kreuzlingen.ch
www.twi-kreuzlingen.ch