# Truth and Robustness in Cross-country Growth Regressions

by

Kevin D. Hoover

and

Stephen J. Perez

31 August 2000

Kevin D. Hoover
Department of Economics
University of California
Davis, California 95616-8578

Tel. (530) 752-2129
Fax (530) 752-9382
E-mail kdhoover@ucdavis.edu

Stephen J. Perez
Department of Economics
Washington State University
Pullman, Washington 99164-4741

Tel. (509) 335-1740
Fax (509) 335-4362
E-mail sjperez@wsu.edu

We thank Oscar Jorda, Judith Giles, Wayne Joerding, and participants in seminars at the University of California, Irvine, and the University of Victoria for helpful comments on an earlier draft. We also thank Orley Ashenfelter for his help in getting this project off the ground.

## Abstract

The work of Levine and Renelt (1992) and Sala-i-Martin (1997a, b) which attempted to test the robustness of various determinants of growth rates of per capita GDP among countries using two variants of Edward Leamer's extreme-bounds analysis is reexamined. In a realistic Monte Carlo experiment in which the universe of potential determinants is drawn from those in Levine and Renelt's study, both versions of the extreme-bounds analysis are evaluated for their ability to recover the true specification. Levine and Renelt's method is shown to have low size and extremely low power: nothing is robust; while Sala-i-Martin's method is shown to have high size and high power: it is undiscriminating. Both methods are compared to a cross-sectional version of the general-to-specific search methodology associated with the LSE approach to econometrics. It is shown to have size near nominal size and high power. Sala-i-Martin's method and the general-to-specific method are then applied to the actual data from the original two studies. The results are consistent with the Monte Carlo results and are suggestive that the factors that most affect differences of growth rates are ones that are beyond the control of policymakers.

**JEL Classification:** C4, C8, O4

**Keywords:** growth, cross-country growth regressions, extreme-bounds analysis, general-to-specific specification search

**Truth and Robustness in Cross-country Growth Regressions**

## 1. The variety of determinants of cross-country growth

Economists typically prefer theoretically informed empirical investigations. Sometimes, however, we face questions for which there is no generally agreed constrained-optimization model to which the empirical researcher can turn. The question, what explains the differences in growth rates among nations?, is a case in point.

While the neoclassical growth model (Robert Solow 1956) tells us that in steady state, growth rates depend on the rates of growth of the labor force and of technological progress, it gives us little notion of what might determine technological progress, especially when technology must be conceived to include all aspects of social organization that might relate to the effectiveness of production. Out of steady-state, increasing the rate of utilization of factors of production or increasing capital investment can temporarily increase growth rates. Also, the higher the gap between the current level of output and the steady-state level, the higher the growth rates. Models of growth with increasing returns suggest that we cast a wider net, looking at industrial organization, research and development, investment in education and other factors (Paul Romer 1986, Robert Lucas 1988 and many others ably surveyed in Charles Jones 1998 and Robert Barro and Xavier Sala-i-Martin 1995). These models generally assume that some factor is important and try to work out the mechanisms of its influence, but they give little guidance to which of the many possible factors really influence growth.

Two – often related – responses to this situation are common. First, economists sometimes take a broader view of the "theory" that they aim to test, to include the less formal considerations of political science or sociology that might bear on the determinants of growth. Second, they sometimes take what, from the point of view traditional econometrics, is an atheoretical approach. This is well exemplified in a series of empirical investigations that are referred to as "cross-country growth regressions." In this literature, cross-sectional regression or panel data techniques are used to try to identify which of a large number of factors are statistically and economically significant determinants of growth rates.[1]

One problem with the literature is that different studies reach different conclusions depending on what combination of regressors the investigator chooses to put into his regression. In an attempt to put some order into the literature, Ross Levine and David Renelt (1992) assembled a cross-sectional data set with a large number of potential regressors and subjected it to a variety of Edward Leamer's "extreme-bounds analysis" (1983, 1985).[2] The central idea in Leamer's analysis is that a coefficient of interest is robust only to the degree that it displays a small variation to the presence or absence of other regressors. Leamer and Leonard (1983) define the extreme-bounds for the coefficient of a particular variable within a search universe as ranging between the lowest estimate of its value minus two times its standard error to the highest estimate of its value plus two times its standard error, where the extreme values are drawn from the set of every possible subset of regressors that include the variable of interest. A variable is said to be *robust* if its extreme bounds lie strictly to one side or the other of zero. And the

[1] The literature is huge. Important contributions are due to Roger Kormendi and Phillip Meguire (1985), Kevin Grier and Gordon Tullock (1989), Barro (1991), J. Bradford DeLong and Lawrence Summers (1991) and Jeffrey Sachs and Andrew Warner (1995, 1996). A recent book by Barro (1997) gives a good overview to the literature.

[2] Temple (1999) reexamines Levine and Renelt's data set from a perspective largely sympathetic to extreme-bounds analysis, including Sala-i-Martin's variant.

narrower the extreme bounds, the more confidence one is supposed to have in the coefficient estimate.

Employing a modified version of Leamer's approach that reduces the number of regressions needed to compute the extreme bounds, Levine and Renelt hold that a variable is not robust if its extreme bounds contain zero. They find that few variables can be regarded as robust determinants of economic growth (that is, almost all coefficient estimates are "fragile").

Sala-i-Martin (1997a, b) argues that Levine and Renelt employ too strict a standard of robustness. He suggests that if *most* of the distribution of the coefficient estimates (plus or minus two standard errors) lie to one side of zero, then the rest might be regarded as irrelevant outliers, and the variable should be regarded to be robust. By analogy with the ordinary practice with significance tests, he suggests that we regard a variable to be robust if 95 percent of the distribution of the coefficient estimates lies to one side or the other of zero. Predictably, on this more permissive standard, Sala-i-Martin finds considerably more variables to be robust determinants of economic growth.

Leamer's notion of robustness strikes us as an odd one. There is no reason to believe that a variable that is robust in Leamer's sense is thereby guaranteed to be a true determinant of economic growth or that a true determinant of economic growth is guaranteed to be a robust one, or even that there is a high correspondence of any kind between truth and robustness (see Kevin D. Hoover 1995 and Hoover and Stephen J. Perez 2000). Leamer (in David F. Hendry *et al.* 1990, p. 188) rejects the notion of a true specification: "I . . . don't think there is a true *data-generating process*. . ." . But then it is puzzling what one is supposed to do with one's robust coefficient estimates.

Advocates of extreme-bounds analysis sometimes argue that the concern with truth is misplaced because to reject a variable as fragile is not to deny that it might be a true determinant of the dependent variable. Rather it is to deny that we have good evidence for its truth. That is, the claim is about us and our certainty, not about the world. But this looks at the wrong side of the issue. The extreme-bounds procedure also identifies some variables as robust. If this says that our evidence is good, that we are more certain, what is the evidence evidence for? what are we more certain about, if not that these variables are true determinants of the dependent variable?

The required notion of truth is not a metaphysical puzzle. A variable is a true determinant of economic growth if variations in that variable (induced by policy or accident) can be relied upon to yield predictable variations in the rate of economic growth. To discover such determinants we seek to convince ourselves that particular variables predictably explain past growth rates and then we hope that the relationship is an enduring one that can be used to explain future growth rates. Whether we are warranted in claiming that particular determinants are true or not is a nice question for the philosophy of science perhaps. Nevertheless, when we use estimated relationships instrumentally, we must be assuming that they are true in this sense and not just correlations or data-summaries.

A practical question to ask of any search methodology, such as Levine and Renelt's or Sala-i-Martin's versions of extreme-bounds analysis is: when there is a truth to be found, do their methodologies discover it? One goal of this paper is to evaluate the success of these two versions of extreme-bounds analysis in a realistic Monte Carlo setting in which we know in fact what the true determinants are. The setting is "realistic" in the sense that it uses actual variables, rather than ones fabricated from random-number generators, as the true determinants and the other variables in the extreme-bounds analysis search universe. The dependent variable, which is calibrated to act like the rate of economic growth, is generated through a bootstrap procedure. The realistic setting ensures that the problem faced in the Monte Carlo is similar to the one that an actual investigator of the determinants of economic growth faces.

We contrast the two extreme-bounds procedures with a mechanized version of the general-to-specific specification search methodology associated with David Hendry and others and often referred to as the LSE [London School of Economics] methodology.[3] In the linear context typical of the cross-country growth literature, the general-to-specific methodology begins with the idea that the truth can be characterized by a sufficiently rich regression: the *general regression*. In particular, if every possible variable is included in the regression, then the regression must contain all the information about the true determinants. It may, however, not provide it in a perspicacious form. The information content might be sharpened by a more parsimonious regression – the *specific regression*. This specific regression is acceptable if it has the properties (a) it is statistically well specified (for example, it has white noise errors); (b) that it is a valid restriction of the general regression, and (c) that it encompasses every other parsimonious regression that is a valid restriction of the general regression.[4]

---

[3] The adjective "LSE" is, to some extent, a misnomer. It derives from the fact that there is a tradition of time-series econometrics that began in the 1960s at the London School of Economics; see Mizon (1995) for a brief history. The practitioners of LSE econometrics are now widely dispersed among academic institutions throughout Britain and the world. The LSE approach is described sympathetically in Gilbert (1986), Hendry (1987,1995, esp. chs. 9-15), Pagan (1987), Phillips (1988), Ericsson, Campos and Tran (1990), and Mizon (1995). For more sceptical accounts, see Hansen (1996) and Faust and Whiteman (1995, 1997) to which Hendry (1997) replies.

[4] The general-to-specific methodology is explained in detail by *inter alia* Phillips (1988) and Hendry (1995).

One regression *encompasses* another if it contains all of the information of the other regression.[5] LSE econometricians have developed various ways of implementing encompassing tests, but there is an easy way to understand the general notion (and a simple way to implement a test). Consider two competing models of the same dependent variable. If we form a third model which uses the union (excluding redundant variables) of both sets of regressors, then each original regression can be seen as nested in the joint regression. If one of the original regressions can be shown to be a valid restriction of the joint regression, then it encompasses the other regression. Of course, it may turn out that neither regression encompasses the other. There may be some third regression, more parsimonious than the joint regression that encompasses them both or it may be that the joint regression is as parsimonious as the joint set of regressors permits.

The LSE approach has been used almost exclusively in time-series contexts. There is, however, nothing in its conceptual structure that prevents its extension to cross-sectional data.

Previously, we have evaluated the efficacy of the general-to-specific approach (Hoover and Perez 1999). Our goal was to determine whether common objections to the LSE approach had practical merit. One objection is that any path of simplifications from the general to the specific is just one of many and there is no guarantee that a particular simplification will be the true specification. We acknowledge the problem, but we showed that simple methods of generating a feasible number of competing specifications

---

[5] For general discussions of encompassing, see, for example, Mizon (1984), Mizon and Richard (1986), Hendry and Richard (1987), Hendry (1988; 1995, ch. 14).

and then choosing among them on the basis of encompassing tests was an effective strategy.[6]

A second objection is that the general-to-specific searches involve multiple testing with unknown distributional properties. In particular, many conjecture that the size of the whole search procedure is vastly larger than the conventional sizes of the underlying specification tests (see, for example, Michael Lovell 1983, who evaluates a variety of search procedures, though not general-to-specific). This is the objection that is usually associated with a general condemnation of "data mining." In fact, we found in a realistic Monte Carlo setting that the size of the search was very close to the nominal size of the underlying specification tests. This suggests that a distinction must be drawn between undisciplined or wrongly disciplined data mining, which is invidious, and well disciplined data mining, which is useful in many contexts.

The LSE approach is not a mechanical one. Instead it relies on a combination of system and econometrician's art. Our evaluation was necessarily based on a mechanical approximation to what LSE econometricians actually do. Some aspects of the approach that might make it more successful were ignored. Nevertheless, the overall assessment was positive.

In this paper, we adapt the LSE approach (described in the next section) to a cross-sectional context. We then investigate the relative success of it and the two extreme-bounds methodologies. Once again, the LSE approach is found to have good properties. In particular, it dominates both the other methodologies. After presenting a

[6] We regarded multiple search paths as an innovation relative to the LSE approach. In response to our suggestion, Krolzig and Hendry (1999) adopt our approach and Hendry and Hans-Martin Krolzig (2000) point out a precedent (Grayham Mizon 1977).

careful assessment of the three methodologies, we turn again to the practical problem at hand: what really determines economic growth? We investigate this question through a comparative investigation of the data sets compiled by Levine and Renelt and by Sala-i-Martin. Our conclusions are substantially different from theirs.

## 2. The effectiveness of three search methodologies in a realistic Monte Carlo study

### 2.1 Data and simulated data

In order to understand the effectiveness of search methodologies, it is essential that the data of the universe of variables used in the simulations display the same sort of intercorrelations as actual data. To achieve this we start with the data set used by Levine and Renelt (1992). Their original data set contains 119 countries for 40 variables. Reporting is not complete so that the 119 x 40 matrix of variables has many missing values. A number of variables and countries are, therefore, deleted from the data set with the goal of producing the largest complete matrix possible. The result is a 107 country by 36 variable data set. The original Levine-and-Renelt data set and the reduced, complete data set are described in detail in Appendix 1. The average rate of growth of GDP per capita 1960-89 (GYP) is the variable of interest in the simulations. Each simulation replaces actual GDP growth with a simulated variable constructed from a linear combination of a subset of the other variables in the data set.

### 2.2 Alternative specifications and the criteria of success

In earlier work on time-series models, Hoover and Perez (1999) used the particular set of models suggested by Lovell (1983) as the “true” specifications. The particular models were suggested by competing theoretical approaches in macroeconomics and reflected several possible dynamic specifications. That approach is less natural in this context, because the literature is motivated largely by *a priori* ignorance of the empirical factors that might explain growth.

One issue that arose in the time-series context does carry over. Any search procedure may fail to select a variable for one of three reasons: (1) it, in fact, is not a true determinant of the dependent variable or (2) the search method is unsuccessful or (3) the signal-to-noise ratio is low (that is, there is insufficient variance in the independent variable relative to the dependent variable). The first type of failure is desirable, the second clearly not. The third, however, is unavoidable. Since there is no reason to believe that a true specification would necessarily have only those variables that are easy to detect, there is no reason in evaluating different search procedures to favor specifications in which all the variables have a high signal-to-noise ratio. The failure to identify a variable with a low signal-to-noise ratio should not be regarded as a failure of the method. Of course, this is a matter of degree. The strategy we adopt is to simulate models with randomly chosen specifications and to evaluate their success relative to norms that depend on the signal-to-noise ratios for each independent variable.

Each simulation is based on a true specification that relates the growth rate to zero, three, seven, or fourteen independent variables. Call the growth rate $\mathbf{y}$ (a 107 x 1 vector). Call the complete set of variables in the data set $\mathbf{X}$ (an 107 x 34 matrix). Let $\mathbf{X}_j$ be a randomly selected $j$-element subset of the variables of $\mathbf{X}$, where $j = 0, 3, 7, 14$. There is a degree of unavoidable arbitrariness in the choice of values for $j$. We justify our choices as follows: Zero independent variables is a baseline for checking the size of the search. Levine and Renelt (1992) and Sala-i-Martin (1997a, b) consider regressions with no fewer than three and no more than seven independent variables. We, therefore, study simulations in which there are three and seven variables as well. When the simulated true specification has either three or seven variables, one of Levine and Renelt's or Sala-i-Martin's test specifications can in fact coincide precisely with the truth. Levine and Renelt's and Sala-i-Martin's search algorithms do not prevent *more than seven* variables from being identified as robust. In fact, Sala-i-Martin (1997a) identifies 21 variables as robust (and maintains an additional 3 as free variables in all regressions, for a total of 24 selected determinants of country growth rates). True specifications with fourteen variables (chosen because it is twice seven), therefore, allow for the implied cases in which the test specifications are, to different degrees, underspecified relative to the truth.

For each $j$, twenty different *specifications* are chosen. And for each specification, 100 *simulations* are run.[7] We proceed as follows:

1. Select a $j$-element subset. This defines the specification.

2. Run the regression, $\mathbf{y} = \mathbf{X}_j\mathbf{B} + \mathbf{u}$ and retain the estimates of the coefficient matrix, $\hat{\mathbf{B}}$, and the estimated residuals, $\hat{\mathbf{u}}$.

3. For each simulation search $i$ = 1, 2, . . ., 100 construct a simulated dependent variable $\mathbf{y}_i^* = \mathbf{X}_j\hat{\mathbf{B}} + \hat{\mathbf{u}}_i^*$. Since the Monte Carlo is based on actual data, which

[7] Simulations use Matlab 5.2 running on PC with 300 Mhz. We would have preferred to examine both more specifications and more simulations of each specification. Unfortunately, each run of 100 simulations for one specification using all three search procedures takes about one and a half days of computing time.

may be heteroscedastic, we construct the elements of the vector $\hat{\mathbf{u}}_i^*$ by sampling from $\hat{\mathbf{u}}$ using a wild bootstrap.[8]

4. The three search procedures are run for the $i^{\text{th}}$ simulation. The successes and failures at identifying the true variables are recorded for each search procedure.

5. The procedure begins again with a new simulation at step 2 until 100 simulations are recorded. The type I and type II errors are recorded.

The proportion of type I error, the empirical *size*, is calculated as the ratio of the incorrect variables included (significantly at the 5 percent critical level for general-to-specific searches) to the total possible incorrect variables. Since the null of the relevant test statistics is that a variable equals zero (or is absent), size measures the likelihood of identifying a variable as a true regressor when it is not. We report the *size ratio*, defined as *size*/0.05. The denominator is the nominal size used in the tests. A size ratio of unity implies that the size of the search procedure is exactly the nominal size.

The empirical *power* is the fraction of the replications in which a true variable is picked out by the search procedure (significantly at the 5 percent critical level for the general-to-specific procedure); that is, it is the complement of the proportion of type II error. In order to control for variations in the signal-to-noise ratio, we compute the *true (simulated) power* from the proportion of type II error for each specification over the 100 bootstrap simulations without search (that is, with knowledge of the correct regressors). The true (simulated) power is the power one would estimate if God revealed the correct specification. If the signal-to-noise ratio is low (high), the true (simulated) power will

[8] Our implementation follows Brownstone and Kazimi (1999), section 2. The wild bootstrap is due to Wu (1986). Horowitz (1997) shows that the wild bootstrap is superior to the more familiar paired bootstrap when data are heteroscedastic.

also be low (high). The *power ratio* is defined as $\frac{power}{true\,(simulated)\ power}$. A power ratio of unity indicates that a search algorithm does as well at picking out the true variables as one would do, given the signal-to-noise ratio, with full knowledge of the true specification.

There is, of course, always a balancing between type I and type II error. If one put no weight on type I error, a search algorithm can achieve 100 percent power by selecting every variable in the data set. The power ratio in that case could easily be much greater than unity, as the true (simulated) power is sometimes very low. The cost, of course, is that the size of such an algorithm is large. Similarly, if one puts no weight on type II error, a search algorithm can achieve a low size by selecting nothing. The cost is that the power of such an algorithm is zero.

### 2.3 Extreme-bounds analysis

We assess a particular variant of Leamer's extreme-bounds analysis related to that used by Levine and Renelt (1992) and Sala-i-Martin (1997a, b). A practical problem in implementing extreme-bounds analysis is the large number of regressors. For example, Levine and Renelt's (1992) data set has 39 variables excluding the dependent variable GYP. There are $2^{39} = 5.498\text{x } 10^{11}$ linear combinations of the regressors. At one second per regression, it would take 17,433 years to try them all.

Levine and Renelt simplify the problem by adopting Leamer's notion that some variables should be included in every regression on the assumption that they are known to be robust *a priori*. The variables RGDP60 (real per capita GDP in 1960), PRI (primary school enrollment rate in 1960), and INV (the average investment share of GDP 1960 to

1989) are included in every regression. In describing Leamer's approach, Michael McAleer, Adrian Pagan and Paul A. Volcker (1983) divide the universe of regressors into *free variables*, which theory dictates should be in the regression; *focus variables*, a subset of free variables that are of immediate interest; and *doubtful variables*, which competing theories suggest might be important. Levine and Renelt treat the three variables included in every regression as free variables, and they let every other variable in turn play the part of a focus variable, while linear combinations of the remaining variables play the part of doubtful variables. They restrict the number of subsets of the doubtful variables further by considering only subsets with three or fewer variables. The largest regression for Levine and Renelt, then, has seven independent variables, exclusive of the constant term: one focus variable, three free variables, and (at most) three doubtful variables.

Sala-i-Martin's (1997a, b) approach modifies Levine and Renelt's search procedure in two ways. First, he considers only regressions of *exactly* seven independent variables: one focus variable, three free variables, and (exactly) three doubtful variables. He tries every linear combination of three doubtful variables in the search universe. Second, he looks at a different criterion for robustness. The estimate of a coefficient on a focus variable is robust in Sala-i-Martin's sense if 95 percent or more of the estimates lie to one side of zero.

The results reported here follow Sala-i-Martin's evaluation procedure modified to eliminate the free variables. To compute the extreme bounds each variable in the search universe is allowed to be the focus variable in turn and regressions that include it and every subset of exactly three other variables (plus a constant) is computed. From these estimates, variables are identified as robust on the Levine and Renelt and the Sala-i-Martin criteria.

Our procedure differs from both Levine and Renelt, because we do not maintain that we have a priori knowledge of *any* of the true regressors. This seems reasonable in the simulations since the true specifications are chosen randomly. We did, however, examine another set of simulations (not reported in detail here) in which the three free variables are part of every true specification and are maintained in every search. There is no qualitative difference between these simulations and the ones reported here.

### 2.4 General-to-specific

The precise details of the general-to-specific algorithm are given in Appendix 2a. Here we provide an outline of the procedure. The search procedure proposed is a modification in the cross-sectional context of the general-to-specific search procedure evaluated in Hoover and Perez (1999) in a time series context.

There are five principal elements in the search procedure.

First, the data are divided into two overlapping samples. The data are ordered according to the size of the dependent variable (i.e., per capita GDP). One sample consists of the first 90 percent of the ordered sample, the other of the last 90 percent. A search is conducted over each subsample and only those variables that are selected in both subsamples are part of the final specification.[9]

Second, each search begins with a general specification in which all of the variables in the search universe are included as regressors. The general specification is

[9] This procedure was an innovation in Hoover and Perez (1999) relative to the LSE methodology, but has been adopted in, for example, Krolzig and Hendry (1999) and Hendry and Hans-Martin Krolzig (2000).

simplified sequentially by removing variables with low *t*-statistics one at a time. Initially, five simplification paths are tried in which each of the variables with the five lowest *t*-statistics is the first variable to be removed along a simplification path. After that – with the exceptions noted below – variables with the lowest *t*-statistic are removed one at a time until all the remaining variables are significant on a 5-percent test. After removal of each variable, a battery of specification tests is performed. The test battery includes a Breusch-Pagan (1980) test for heteroscedasticity, a subsample stability test using an equality of variance test (a cross-sectional analogue to a Chow test), and an F-test of the restrictions from the general model. The number of tests failed is recorded for each step.

Third, after all variables in a specification are significant, the test battery is run. If all tests are passed, this is the terminal specification. If any are failed, the last specification passing all tests becomes the current specification.

Fourth, if a terminal specification is not found in the last step, a new round of variable elimination proceeds with the removal of the variable with the lowest *t*-stat in the current specification. At each step the test battery is run. If a specification fails one of the tests, the last removed variable is replaced, the variable with the next lowest *t*-statistic is removed and the test battery is run again. This process continues until a variable can be removed without failing any of the tests or all variables are tried.

Fifth, once all search paths have ended in a terminal specification, the *final specification* is chosen through a sequence of encompassing tests. We form the non-redundant joint model from each of the different terminal specifications; take all candidate specifications and perform the F-test for encompassing the other specifications. If only one specification passes, it is the final specification. If more than one specification passes, the specification with the minimum Schwarz criterion is the *overall terminal specification* for the subsample. If no model passes, reopen the search on the non-redundant joint model (including testing against the general specification) using only a single search path and take the resulting model as the *overall terminal specification* for the subsample. The *final specification* is, as noted above, the intersection of the regressors of the overall terminal specifications from the two 90-percent subsamples.

### 2.5 Results of the Simulations

The results of the simulations are presented in Table 1. Recall that for both the size and the power ratio a value of unity is a useful reference point. A size ratio of unity indicates that the algorithm incorrectly accepts a variable at the same rate that independent tests of a 5-percent nominal size would do. A power ratio of unity indicates that the algorithm chooses the true variables at the same frequency that one would if one knew the true specification and used a *t*-test with a 5-percent critical value to decide whether a variable should be retained.

The extreme-bounds analysis using Leamer's original criterion shows an extremely low size, which becomes lower as the number of variables in the true specification becomes higher. When there are 14 true variables, this algorithm almost never selects a variable that does not belong. The trade off, however, is that its power ratio is low in all cases, and it too approaches zero as the number of variables in the true specification becomes equal to and then larger than the number of variables in the specifications used to estimate the extreme bounds. While the extreme-bounds algorithm almost never commits type I error, it almost always commits type II error. This confirms generically the criticism made by Sala-i-Martin of Levine and Renelt's use of extreme-bounds analysis. It is overly strict. It says, "nothing is robust;" and, in so saying, it is unable to find the truth at all.

In contrast, the modified extreme-bounds analysis does almost exactly the reverse. Its size ratio is 42 percent greater than the 5-percent nominal test size when there are no true variables. It rises to an astonishing 540 percent greater than the nominal test size for specifications with 14 true variables. Compared to the standard extreme-bounds analysis it picks out too many variables that do not belong in the true specification. Of course, this increases the power ratio. When there are three variables in the true specification, the power ratio of 1.07 suggests that it is more likely to pick the true variables than even knowing the true specification would suggest. As the number of variables in the true specification becomes larger, the power ratio falls to less than 2/3. Although Sala-i-Martin shows some success in correcting the overly strict character of Levine and Renelt's method, the cure comes at the price of going way too far in the other direction. His method is overly lax. It says, "many variables are robust" and, in so saying, it is unable to discriminate the true from the false.

The general-to-specific algorithm finds the middle ground. Its size ratio is only a little greater than unity and becomes larger only slowly. With 14 true variables it is only 27 percent greater than unity. Its power ratio is always a little less than unity, but except for the case of three true variables, substantially larger than that for the modified extreme-bounds analysis. In comparison to the other two methods, the general-to-specific algorithm not only usually finds the truth nearly as well as one would if God had whispered the true specification in one's ear, but it also is able to discriminate between true and false variables extremely well.

The fact that the empirical size is well behaved (that is, near unity) for the general-to-specific search algorithm is perhaps the most striking thing about these findings. Many critics of data mining in general, and the general-to-specific methodology in particular, express *a priori* scepticism of the practice of multiple, sequential testing using conventional critical values. Invariably, they predict that such test procedures are bound to understate the true size of the joint test implicit in the search procedure. The evidence here runs in the other direction altogether. Far from the simulations showing that the empirical size is very high, it is, in fact, no more than 28% larger than the nominal test size. These results are broadly consistent with the earlier findings of Hoover and Perez (1999), who found empirical sizes for the general-to-specific algorithm that were greater – but only a little greater – than the nominal sizes of the tests. One way to understand this result is that the disciplines imposed by the various encompassing tests in the search procedure tend to force the final specification to be close to the true specification. And, if one had known the true specification *a priori*, the nominal test sizes would have been correct. Tests based on a specification that is near the true specification have similar size.

## 3. Re-examining the data

### 3.1 Adapting to Real-world Data

The previous section cast doubt on the efficacy of extreme-bounds methods in identifying the true determinants of a dependent variable in a case in which those true determinants were in fact known. The general-to-specific methodology did substantially better. What implications would that have for reasonable conclusions about the actual determinants of cross-country growth differentials? To investigate this question, we apply the general-to-specific methodology to the larger of Levine and Renelt's (1992) data sets and to Sala-i-Martin's (1997a, b) data set. We compare the results in each case to those using Sala-i-Martin's modified extreme-bounds analysis.

In the Monte Carlo simulations of the last section, we worked with "nice" data. In particular, our set of dependent variables was transformed to be homoscedastic and we eliminated a carefully chosen set of countries and variables in order to produce a data matrix without missing values. The two data sets that we use here contain missing values, and we cannot be assured that they are homoscedastic. Neither Levine and Renelt nor Sala-i-Martin concern themselves with heteroscedasticity. They handle missing values through, what is sometimes called *casewise deletion*: for any regression, if a country does not report the values for each of the variables required for that regression, that country is omitted from the regression. As a result, regressions are run over a shifting set of countries. While this may raise some questions of the legitimacy of comparing coefficient estimates across regressions conducted on different samples, it does not pose any barrier to the mechanical implementation of extreme-bounds analysis.

This is, unfortunately, not true of the general-to-specific methodology. Since it involves constant comparison of more parsimonious specifications to the general specification, it would seem to require that both were estimated on the same set of observations. This is, of course, difficult when there are many missing values. If we restrict ourselves to countries for which the data set is complete, then we have very few degrees of freedom for the general specification. If we eliminate variables for which the data for some country is missing, then we lose variables of significant interest. Our solution to this problem is pragmatic and two-pronged. On the one hand, we do eliminate some countries and variables from the data set when they have too many missing values (though in some cases, we are able to consider them partly at a later stage). On the other hand, we make some modifications to the search algorithm to adapt to the missing data.

We do not regard these pragmatic steps as in any way arguing in favor of the other methodologies. The essential problem of incomplete information affects all three methodologies equally. In the practice of casewise deletion, information is thrown away. The extreme-bounds methodologies can be implemented in a way that one is not reminded of that fact, while the general-to-specific methodology does not permit one to forget it: that is the only difference.[10]

While, the details of the handling of the data and the modifications to the algorithm are slightly different for each of the two data sets (and are described separately), there are common elements.[11]

First, because the degrees of freedom vary from regression to regression as simplification proceeds, it is not possible to test encompassing using a standard *F*-test. Instead, following the suggestion of Hendry and Krolzig (1999) and Bruce Hansen (1999), we use an information criterion. The usual Schwarz criterion (or Bayesian Information Criterion or BIC) is calculated as: $SC = -2L/T + k\log(T)/T$, where $L$ is the maximized value of the log-likelihood function and $T$ is number of observations and $k$ the number of coefficients estimated in the regression. Case-wise deletion poses a problem:

---

[10] We have attempted to implement special techniques for handling data with missing observations (see Griliches 1986, Little and Rubin 1987,and Schaefer 1997), with little success so far.

the number of observations may vary depending on which variables appear in the regression, making valid comparison using the Schwarz criterion difficult. For example, suppose that two regressions have the same number of, but different, regressors. And that there are $T_1$ observations on regression 1 and $T_2 < T_1$ on regression 2. Then even if each regression had the identical maximized value of the log-likelihood function, regression 1 would have a lower Schwarz criterion than regression 2. But that is just an artifact of the missing observations. A simple way to correct for the problem is observe that the first term in the Schwarz criterion formula is just the average value of the maximized log-likelihood function. With more observations, this term is likely to be more precisely estimated, but it should *on average* be the same independent of the sample size. The problem is the second term. It aims to penalize overparameterized regressions. To place the comparison on a fairer basis, we evaluate this term at the *same* number of observations for all regressions compared, using the minimum number for any regression in the comparison set. Consider *n* regressions with numbers of observations $T_j$, $j = 1, 2, \ldots$ n. We calculate an *adjusted Schwarz criterion* for each regression $j$: $ASC_j = -2L_j/T_j + k_j \log(T_{min})/T_{min}$, where $T_{min} = \min\{T_j\}$ is the fewest number of observations among the regressions being compared.[12] The regression with the smallest adjusted Schwarz criterion is preferred.

The second modification common to both data sets is that we conduct only a full-sample search rather than searches over two overlapping 90 percent subsamples.[13]

---

[11] Step-by-step details of the search procedure are found in Appendix 2a.

[12] We thank Oscar Jorda for suggesting this adjustment. Note that the value of the adjusted Schwarz criterion for any regression depends in general on which other regressions are in the comparison set.

[13] In the time-series case, Hoover and Perez (1999) show that even without the overlapping subsamples the size of the algorithm is close to the nominal size of the test and has little effect on the power. The overlapping subsamples substantially decrease the size in that case. The cost of eliminating it in this case is, then, likely to be an increase in size, that is an increase in variables that are selected incorrectly. Recall

The third modification common to both data sets is that the choice of subsample equality test used in the test battery differs when the degrees of freedom are short.

The fourth modification common to both data sets is justified on the grounds that, owing to the missing values, the general-to-specific algorithm cannot usually operate on the same sample as the modified extreme-bounds analysis on these data sets. It, therefore, seems reasonable to consider the set of robust regressors chosen by the extreme-bounds analysis as a possible alternative specification. The LSE methodology suggests that serious alternative specifications – of whatever provenance – ought to be subjected to encompassing tests. To implement the comparison, we treat the modified extreme-bounds specification as one of the terminal specifications (in addition to the five derived from simplification of the general specification) and permit the algorithm to choose it or to extract non-redundant information from it as from any other terminal specification.

### 3.2 The Levine-and-Renelt Data Set

We begin with the complete data set described in Appendix 1. The variable TRD (the ratio of total trade to GDP) is defined to be the sum of X (the export share in GDP) and M (the import share in GDP). The three are, therefore, perfectly collinear. We therefore eliminate from the search TRD and GTRD (the growth rate of TRD), which is nearly collinear with XSG (the growth rate of X) and MSG (the growth rate of M). The data set,

from Table 1, however, that the size for the general-to-specific algorithm is uniformly lower than nominal test size anyway and that it is much lower than the size of the modified extreme-bounds algorithm, so there is likely to be some margin for error on this dimension.

including the dependent variable, now comprises 36 variables (the dependent variable, GYP (the growth of real GDP per capita 1960-89) and 35 independent variables) for 119 countries.

Initial searches using a 5-percent significance level for the test battery resulted in premature stopping in the sense that an unreasonably large number of regressors were retained. We believe heteroscedasticity is the cause. Following the suggestion of Krolzig and Hendry (1999) and Hendry and Krolzig (1999) the search procedure is run with a significance level of 1 percent on all tests. Since a lower critical value makes it harder to reject the nulls of specification tests, it is a more permissive standard.

A general specification was run using all the variables and eliminating those countries with incomplete data. For this regression $\overline{R}^2 = 0.99$. The left-hand specification in Table 2 shows the regression that corresponds to the final specification. It retains only 8 of the original 35 variables; its $\overline{R}^2 = 0.98$. The adjusted Schwarz criterion for the general regression is 0.503 and for the final specification is 0.198. Not counting the constant, 3 of the 8 coefficients are statistically insignificant at the 5 percent critical value, but only 2 at the 10 percent critical value.

For comparison, we conducted a modified extreme-bounds search on the Levine and Renelt data. Following Levine and Renelt (1992) the three variables, INV (investment share in GDP), PRI (primary school enrollment rate in 1960) and RGDP60 (logarithm of real GDP in 1960) are retained as free variables in every regression. Table 3 shows the results of the modified extreme-bound analysis in a format that corresponds to Table 1 of Sala-i-Martin (1997b), which is an expanded version of Table 1 of Sala-i-Martin (1997a). Running the eye quickly down the columns headed “Lower Extreme” and "Upper Extreme" confirms Levine and Renelt's (1992) original conclusions that none of the variables, including the three free variables, is robust: each set of extreme bounds straddles zero. Nevertheless, on the modified robustness criterion – based, following Sala-i-Martin's (1997a) preference on a non-normal, weighted cumulative distribution function – 7 doubtful variables are robust.[14]

The right-hand specification in Table 2 shows the regression that corresponds to the modified extreme-bounds search using the 7 robust variables identified in Table 3, along with the three free variables. Of these 10 variables, 7 are statistically significant using a 10 percent critical value, 6 are significant using a 5 percent critical value, and 1 is significant using a 1 percent critical value. Three, however, are not significant at any conventional level. The regression fits badly, $\overline{R}^2 = 0.69$, relative to the general-to-specific regression – $\overline{R}^2 = 0.98$. The adjusted Schwarz criterion is lower for the general-to-specific regression (0.156) than for the modified-extreme-bound regression (3.37). Since the samples of the two regressions in Table 2 are different, we cannot directly check for encompassing using a *F*-test. If, however, we eliminate the minimum number of countries to give them a common sample, then it is possible to test them each against a model that contains the non-redundant union of their regressors. Against, this joint model, neither specification encompasses the other. The *p*-value of the *F*-test of the exclusion restrictions implied by the modified extreme-bounds specification is $p = 0.02$, while that for the general-to-specific is $p = 0.00$**.** The balance of evidence is that the general-to-specific specification encompasses the modified extreme-bounds specification,

[14] The construction of the cumulative distribution function and the weighting scheme are described in Sala-i-Martin (1997a, b).

although on the much restricted common sample we can also reject the exclusion of the variables unique to the extreme-bounds specification.

The results of comparing the two specifications are consistent with the conclusions summarized in Table 1. Two variables are chosen by both the general-to-specific algorithm and the modified extreme-bounds analysis. Eight are chosen by the modified extreme-bounds analysis only, while six are chosen by the general-to-specific algorithm only. A pattern in which the general-to-specific specification is very nearly nested in the modified extreme-bounds specification would be consistent with the large size and relatively large power ratios of the modified extreme-bounds analysis and the small size and relatively large power ratios of the general-to-specific algorithm. This pattern is only weakly suggested in this case, though, as we shall see, more strongly suggested in the case of the Sala-i-Martin data set.

How important are the various determinants of economic growth economically? We evaluate the contribution of each statistically significant (at the 5-percent level) variable in the general-to-specific search in the following way: first, ignoring parameter uncertainty we compute the product of the corresponding coefficient and the median value of the variable. This provides some reference point for a typical degree of importance. We then consider what would happen *ceteris paribus* if, instead of taking the median value, the variable took the most favorable and least favorable values in the data set (depending on the sign of the coefficient this could be the maximum or the minimum value in either case). These favorable or unfavorable extremes are reported as the increments above or below the median, so that they indicate how much more and how much less growth a country would display if it were able to adopt the best or suffer the worst value of the variable actually experienced by any country in the data set.

The economic importance of the determinants of growth are reported in Table 4. The variables are arranged in ascending order of the effect on the growth rate attributable to each variable evaluated at its median value. The five variables divide into two groups: real capita GDP in 1960 and a group of trade-related variables. Real per capita GDP in 1960 has a small negative effect on growth substantiating the finding of conditional convergence (see Mankiw, Romer, and Weil 1992). The trade variables appear to have larger effects. The import share has the largest negative effect at the median and at both the favorable and unfavorable extremes. The growth of imports has the largest positive effect – again at the median and the extremes. The two export variables have positive but more moderate effects.

Sala-i-Martin (1997a,b) points out that Levine and Renelt's data set includes variables as potential regressors that may be endogenous. The trade variables are likely to be endogenous related to economic growth. Endogenous regressors call into question a causal reading, not only of any final regressions based on the data set, but also the validity of the ordinary-least-squares regressions in all the intermediate stages of both search procedures. To account for this in a new data set, Sala-i-Martin, to a greater degree than Levine and Renelt, collected variables that were likely to be predetermined, so that a causal reading of their relationship to the rate of growth of per capita GDP is more plausible.

### 3.3 The Sala-i-Martin Data Set

The data set is described in Appendix 3. It contains 64 variables for 138 countries with 14.5 percent of the values missing. The dependent variable is the growth rate of real per capita GDP 1960-92 (GR56092). We omit completely HUMANYL (missing values for 66 countries) and LLY1 (missing values for 47 countries). We also omit age AGE, because the data do not appear to correspond to its definition: "average age of the population."[15] In order to have a sufficient number of degrees of freedom to run the general regression we also omit temporarily the variables FREEOP, FREETAR, PYR60, HUMAN60, and TOT1. These are chosen because the patterns of their missing values are such that relatively few omissions permits us to increase the number of countries with complete coverage for the remaining variables substantially. Once we obtain the final specification from the general-to-specific search omitting these variables, we test it against the specification with them added as additional regressors. Using the adjusted Schwarz criterion, the final specification does not encompass the specification with the omitted regressors reintroduced. However, an $F$-test of the final specification against the more general model using the maximum common sample (50 countries) cannot reject their insignificance ($p$-value = 0.30). They are, therefore, omitted in the final specification. The search is conducted with the nominal size of all tests set at 1 percent.

The adjusted Schwarz criterion for the general regression is –6.427 and for the final specification is –6.644. The left-hand specification of Table 5 shows the regression that corresponds to the final specification. Of its 17 variables (other than the constant) 14

[15] The data descriptions in Sala-i-Martin (1997a, b) and the data posted on his website are poorly and inaccurately documented. We cannot verify in every case the correspondence to their definitions or the indicated sources. We were, however, able to reproduce his modified extreme-bounds estimates using the

are significant at the 5-percent level and 10 at the 1-percent level. The modified extreme-bounds analysis for this data set is reported as Table 1 in Sala-i-Martin (1997b) and in an abbreviated form as Table 1 in Sala-i-Martin (1997a). The right-hand specification of Table 5 reports the regression that corresponds to the set of variables that meet Sala-i-Martin's preferred robustness criterion. Of its 24 variables (other than the constant), 17 are not statistically significant at the 10 percent level. Of the remaining 7 variables, 4 are significant at the 1 percent level; and an additional 2 at the 5 percent level; and one more at the 10 percent level.

On the adjusted Schwarz criterion, the general-to-specific specification encompasses the modified extreme-bounds specification. The same conclusion is reached through testing the two models against the minimally-nesting joint specification on a common sample. $F$-tests cannot reject the validity of the exclusion restrictions implied by the general-to-specific specification ($p$-value = 0.83), while rejecting the validity of the restrictions implied by the modified extreme-bounds specification ($p$-value = 0.08) at the 10-percent level.

Comparison of the specifications in Table 5 is consistent with the results of the Monte Carlo study summarized in Table 1. Thirteen variables are chosen by both search methodologies; 11 additional regressors are chosen only by modified extreme-bounds analysis; 4 are chosen only by the general-to-specific algorithm. These are just the patterns one would expect given the size and power ratios for the two search methodologies reported in Table 1.

---

data. AGE was not robust in the modified extreme-bounds analysis and recomputing the modified extreme-bounds analysis omitting it has little effect on the robustness of the remaining variables.

Table 6 provides an assessment of the economic importance of the various determinants of economic growth, again with respect to those variables reported as significant (at better than a 5-percent level) in the general-to-specific search reported in Table 5. Again, these are arranged in ascending order according to their importance for growth at their median values.

GDP per capita in 1960 is the most important variable reported in Table 6. At the median, it is a huge negative factor that is only partly offset in the most favorable case. This finding is strong evidence for the conditional convergence hypothesis (see Mankiw, Romer, and Weil 1992).

Five religious indicators are selected (the fraction of the population in each country that is Buddhist, Confucist, Jew, Muslim, and Protestant). The first four have a small effect at the median and a large effect at the favorable extreme. Protestant has a negative effect at the median with a small favorable and moderate unfavorable extreme.

Three financial measures are selected. A measure of financial instability (the standard deviation of domestic credit) is relatively important at both the median and the unfavorable extreme; while a measure of exchange-rate distortion is a more important negative factor at the median, though less important at extremes. Surprisingly, the standard deviation of inflation is related positively to growth and shows a moderately significant effect at the favorable extreme.

Investment is relatively important. Interesting, the search eliminates nonequipment investment and public investment and retains only equipment investment. Equipment investment is the second highest positive influence at the median, and its influence increases by a factor of five at the favorable extreme. The negative value at the unfavorable extreme nullifies the contribution of investment to growth.

Political variables are also important. Wars have a moderate influence only at the unfavorable extreme. Revolutions and coups, perversely, register a positive influence, which suggests that they proxy for some more complex social or political factors. The index marking the rule of law is moderately positively important at the median and the extremes.

**Conclusion**

There are two main points to this study. The first is methodological. Despite the fact that we do not have good *a priori* theory of the determinants of differences in growth rates between countries, we would like to identify the true determinants. Robustness in Leamer's sense bears little practical relationship to truth. Extreme-bounds approaches in the form advocated by Levine and Renelt are too stringent and reject the truth too frequently (small size, but low power), while those advocated by Sala-i-Martin are not discriminating and accept the false too frequently along with the true (high power, but large size). In contrast, the general-to-specific specification search methodology is – like Little Bear's bed in the tale of *Goldilocks* – just right: it maintains a size near (and even a little below) the nominal size of the tests used in the search and has power approaching the true power one should find if the specification were not in doubt.

It is sometimes objected that the advantage of the general-to-specific search is illusory because it presupposes (wrongly, it is asserted) that the true specification is nested in the search universe, and that this is unlikely, since the search universe never includes every variable that matters to the dependent variable in any way. This misunderstands both the exercise conducted in this paper and the underlying strategy of the LSE methodology. Of course, the general-to-specific search cannot locate the true specification if the true variables are not available to the search. But equally, there is no reason to suppose that extreme-bounds analysis is any more informative when variables are omitted from its search universe, than when they are included. The argument is that if any of the methods fail to find the truth when it is in fact there to be found, the method is *a fortiori* unsuccessful. If robustness does not correspond to truth when truth is to be had, why should it be regarded as a desirable characteristic when truth is required but unavailable? We can never guarantee that the specifications selected by the general-to-specific approach are true. But the approach is part of a critical, indeed dialectical, methodology. If anyone seriously argues that an important variable has been omitted from the specification, the appropriate response is to add that variable to the search universe and, then, to rerun the search. It is worth noting that this same critical spirit can be applied to the general-to-specific search algorithm itself. We have presented only a single version of a general approach. While we have shown that it is superior to the two alternatives that we studied, it is not necessarily the best implementation of that approach. We look forward to further refinements and developments – and perhaps to further horse-races against other search methodologies.

The second main conclusion from the study is that in practice extreme-bounds methods are misleading about the determinants of growth. Sala-i-Martin was right to criticize Levine and Renelt (1982) for rejecting too many potential determinants of growth as non-robust. What is more, he is right to question the exogeneity of a number of the determinants of growth that they consider. On the other hand, the evidence of the general-to-specific approach is that his approach selects many variables that probably do not truly determine differences in growth rates. The spuriously included variables are a mixture, but a number of political and political economics variables (civil liberties, political rights, economic organization, and openness to trade) are particularly important among them. Three political variables (the rule of law, revolutions and coups, and wars) survive the general-to-specific search and are economically important. Variables reflecting macroeconomic and financial conditions are relatively important in the general-to-specific specification and ignored as non-robust by the modified extreme-bounds analysis.

The general-to-specific search, therefore, reaches different conclusions about the determinants of differences in growth rates among countries than does the modified extreme-bounds analysis. There are two messages. First, initial conditions matter: religious factors and the initial level of economic development, which, unfortunately, are not amenable to public policy, are relatively important. Second, resource endowments, religious and geographic factors, initial life expectancy, and initial level of economic development. Third, it is best to be law-abiding, to have high private investment, to have a stable macroeconomic environment, and to avoid wars, but these are, at best, a partial offset to starting off on the right foot.

## Appendix 1. The Levine and Renelt (1992) Data Set

The Levine and Renelt data set (L&R) was downloaded from the World Bank's website (http:www.worldbank.org/html/prdmg/grthweb/ddlevren.htm) on 7/30/98. Levine and Renelt use two data sets: one for the years 1960-89 and one for 1974-89. The coverage of the later data set is less complete and we do not use it in this paper. The data used in the paper and the documentation here correspond to the file posted on the website. The appendix to Levine and Renelt (1992) refers to 13 variable names which do not appear in in the downloaded data set. The data set includes 3 variable names that do not appear in the appendix. Two of these are synonyms (IMP in the appendix is M in the data set, and GSG in the appendix is GGOV in the data set).

**Definitions of Variables in the Levine and Renelt (1992) Data Set**

| No. | Variable | Definition | Source |
|---|---|---|---|
| 1 | AFRICA | Dummy Variable for Subsaharan African Countries | |
| 30 | BMP | Black Market Exchange Rate Premium | Picks Currency Yearbook; World Bank Updates |
| 31 | BMS | Standard Deviation of BMP | |
| 2 | CIVL | Index of Civil Liberties | Barro (1991) |
| 3 | DCPYI[2] | Initial Value of the Ratio of Gross Claims on the Private Sector by the Central Bank to GDP | |
| 4[1] | GDC | Growth Rate of Domestic Credit | IMFIFS |
| 37[1] | GGOV[2,3] | Growth Rate of GOV | |
| 5 | GM | Growth of Imports | WBNA |
| 6 | GOV | Government Consumption Share of Gross Domestic Product | WBNA |
| 40 | GPI[2] | Growth Rate of PI | |
| 7 | GPO | Growth of Population | WBSI |
| 8 | GSG | Growth of the Share of Government Consumption (GOV) | WBNA |
| 39 | GTRD[2] | Growth Rate of TRD | |
| 9 | GX | Growth of Exports | WBNA |
| 38[1] | GX[2,4] | Growth Rate of X | |
| 10 | GYP | Growth of Real per Capita Gross Domestic Product | WBNA |
| 11 | INV | Investment Share of Gross Domestic Product | WBNA |
| 12 | LAAM | Dummy Variable for Latin American Countries | |
| 13 | LIT60 | Literacy Rate in 1960 | WBSI |
| 32 | M | Import Share of GDP | WBNA |
| 35 | MIX | Dummy Variable for Mixed Government | Barro (1991) |
| 33 | MSG | Growth of Import Share | WBNA |
| 14 | OECD | Dummy for OECD Countries (members of the Organization for Economic Cooperation and Development) | |
| 15 | OIL | Dummy for OPEC Countries (members of the Organization of Petroleum Exporting Countries) | |
| 16 | PI | Average Inflation of GDP Deflator | WBNA |
| 17 | POP70 | Population in 1970 | Summers and Heston (1988) |
| 18 | PRI | Primary Enrollment Rate 1960 | Barro (1991) |
| 34 | PRJ | Primary Enrollment Rate 1970 | Barro (1991) |
| 19 | REVC | Number of Revolution and Coups per Year | Barro (1991) |
| 20 | RGDP60 | Real GDP per Capita in 19xx | Summers and Heston (1988) |
| 21 | SCOUT | Dummy for Outward Orientation | Syrquin and |

| | | | |
|---|---|---|---|
| | | | Chenery (1988) |
| 22 | SEC | Secondary Enrollment Rate 1960 | Barro (1991) |
| 23 | SED | Secondary Enrollment Rate 1970 | Barro (1991) |
| 36 | SOC | Dummy for Socialist Economy | Barro (1991) |
| 24 | STDC[5] | Standard Deviation of GDC (Growth of Domestic Credit) | IMFIFS |
| 25 | STPI[6] | Standard Deviation of PI (inflation) | WBNA |
| 26[1] | TRD | Ratio of Total Trade (exports + imports) to GDP | WBNA |
| 28 | X | Export Share of GDP | WBNA |
| 27 | XSG | Growth of Export Share of GDP | WBNA |
| 29 | YRSCH[2] | Average Years of Schooling at 1980 | |

Sources: *WBNA: World Bank National Accounts*
*WBSI: World Bank Social Indicators*
*IMFIFS: International Monetary Fund, International Finance Statistics*
*IMFGFS: International Monetary Fund, Government Finance Statistics Yearbook*

Notes:
No. indicates the variable number in the Levine and Renelt (1992) data set.
[1]Indicates that variable is omitted from the data set used in the Monte Carlo simulation and in the specification search in Section 3.2.
[2]These variables do not appear in the documentation for the Levine and Renelt data set, but have been identified from the documentation to King and Levine (1993), downloaded from the World Bank website on 10/6/97.
[3]GGOV is the same as GSG
[4]GX is the same as XSG
[5]Appears in documentation on website and in Levine and Renelt (1992) as STDD, but appears in the data set posted on the website as STDC.
[6]Appears in documentation on website and in Levine and Renelt (1992) as STDI, but appears in the data set posted on the website as STPI.

**Country Coverage for the Levine and Renelt Data Set**

| No. | Country | No. | Country | No. | Country |
|---|---|---|---|---|---|
| 1[1] | Afghanistan | 40 | Haiti | 80 | Paraguay |
| 2 | Algeria | 41 | Honduras | 81 | Peru |
| 3 | Angola | 42 | Hong Kong | 82 | Philippine |
| 4 | Argentina | 43 | Iceland | 83 | Portugal |
| 5 | Australia | 44 | India | 84 | Rwanda |
| 6 | Austria | 45 | Indonesia | 85 | Saudi Arabia |
| 7 | Bangladesh | 46 | Iran | 86 | Senegal |
| 8 | Barbados | 47[1] | Iraq | 87 | Sierra Leone |
| 9 | Belgium | 48 | Ireland | 88[1] | Singapore |
| 10 | Bolivia | 49 | Israel | 89 | Somalia |
| 11 | Botswana | 50 | Italy | 90 | South Africa |
| 12 | Brazil | 51 | Jamaica | 91 | Spain |
| 13 | Burundi | 52 | Japan | 92 | Sri Lanka |
| 14 | Cameroon | 53 | Jordan | 93 | Sudan |
| 15 | Canada | 54 | Kenya | 94 | Swaziland |
| 16 | Central African Republic | 55 | Korea | 95 | Swedend |
| 17 | Chad | 56 | Kuwait | 96 | Switzerland |
| 18 | Chile | 57 | Lesotho | 97 | Syria |
| 19 | Colombia | 58 | Liberia | 98[1] | Taiwan |
| 20 | Congo | 59 | Luxembourg | 99 | Tanzania |
| 21 | Costa Rica | 60 | Madagascar | 100 | Thailand |
| 22 | Cote D'Ivoire | 61 | Malawi | 101 | Togo |
| 23 | Cyprus | 62 | Malaysia | 102 | Trinidad and Tobago |
| 24 | Denmark | 63 | Mali | 103 | Tunisia |
| 25 | Dominican Republic | 64 | Malta | 104 | Turkey |
| 26 | Ecuador | 65 | Mauritania | 105[1] | Uganda |
| 27 | Egypt | 66 | Mauritius | 106 | Great Britain |
| 28 | El Salvador | 67 | Mexico | 107 | United States |
| 29 | Ethiopia | 68 | Morocco | 108 | Uruguay |
| 30 | Fiji | 69[1] | Mozambique | 109 | Venezuela |
| 31 | Finland | 70 | Netherlands | 110 | Yemen |
| 32 | France | 71 | New Zealand | 111 | Zaire |
| 33 | Gabon | 72 | Nicaragua | 112 | Zambia |
| 34 | Gambia | 73 | Niger | 113[1] | Zimbabwe |
| 35 | Germany | 74 | Nigeria | 114[1] | Burma |
| 36 | Ghana | 75 | Norway | 115[1] | Guyana |
| 37 | Greece | 76 | Oman | 116 | Benin |
| 38 | Guatemala | 77 | Pakistan | 117 | Burkina Faso |
| 39[1] | Guinea-Bissau | 78 | Panama | 118[1] | Nepal |
| | | 79 | Papua New Guinea | 119[1] | Suriname |

Note:
No. indicates the country number in the Levine and Renelt (1992) data set.
[1]Indicates that variable is omitted from the data set used in the Monte Carlo simulations.

**Appendix 2a. The General-to-Specific Search Algorithm Used in the Simulations**

A. The data run are generated according to the simulated equation setup with either 0, 3, 7, or 14 true variables included. *Candidate* variables include a constant and all variables in Appendix 1, with the exceptions noted in the main text and in the footnotes to the appendix. A *replication* is creation of a simulated dependent variable using one of the simulated models and one draw from the bootstrapped random errors. *Nominal size* governs the conventional critical values used in all of the tests employed in the search: it is 5 percent.

B. Two sub-samples are created: one is the first 90% of the data set (ordered using the dependent variable) the other is the last 90% of the data set. Independent searches are run on the two sub-samples.

*C.* A *general specification* is estimated on a replication using a full set of candidate variables.

D. Five search paths are examined. Each path begins with the removal of one of the candidate variables with the five lowest *t*-statistics in the current general specification. All *t*-statistics are calculated using White's heteroscedasticity-corrected standard errors. The first search begins by re-estimating the regression. This re-estimated regression becomes the *current specification*. The search continues until it reaches a *terminal specification*

E. The current specification is estimated and all searchable variables are ranked according to their *t*-statistic. The searchable variable with the lowest *t*-statistic is removed.

F. Each current specification is subjected to the following battery of tests:

 i. Breusch-Pagan test for heteroscedasticity
 ii. subsample stability test: an *F*-test for the equality of the variances of the first half versus the second half of the sample in which the data are ordered according to the value of the dependent variable. (This is analogous to a Chow test in a time-series context.) This test compares the regressions over each subsample to the regression over the full sample. If the degrees of freedom do not permit splitting the sample into equal subsamples, the test is replaced by one that compares a regression over the first k + (n-k)/2 observations to the one over the full sample. On both tests, see Chow (1960).
 iii. An *F*-test of the hypothesis that the current specification is a valid restriction of the current general specification

G. The number of tests failed is recorded and the new specification becomes the current specification. Return to C until all remaining variables have a significant *t*-statistic.

H. If all variables are significant, and all of the tests in the test battery are passed, the current specification is the terminal specification and go to J. If any of the tests fails return to the last specification for which all the tests are passed and go to I.

I. The variable with the lowest *t*-statistic is eliminated. The resulting current specification is then subjected to the battery of tests.

   i. If the current specification fails any one of these tests, the last variable eliminated is replaced, and the current specification is re-estimated eliminating the variable with the next lowest insignificant *t*-statistic.
   ii. If the current specification passes all tests, re-estimate and return to I.
   iii. The process of variable elimination ends when a current specification passes the battery of tests and either has all variables significant or cannot eliminate any remaining insignificant variable without failing one of the tests.

J. After a terminal specification has been reached, it is recorded and the next search path is tried until all have been searched.

K. Once all search paths have ended in a terminal specification, the *final specification* is chosen through a sequence of encompassing tests. We form the non-redundant joint model from each of the different terminal specifications; take all candidate specifications and perform the F-test for encompassing the other specifications. If only one specification passes, it is the final specification. If more than one specification passes, the specification with the minimum Schwarz criterion is the final specification. If no model passes, reopen the search on the non-redundant joint model (including testing against the general specification) using only a single search path and take the resulting model as the final specification.

L. The final specification is the intersection of the two specifications from each sub-sample.

## Appendix 2b. The General-to-Specific Search Algorithm used on the full data set

We perform a general-to-specific search that does not account for missing data. In other words, whenever a regression is run, any country that has missing data for any of the variables in the regression is eliminated. This presents a problem for the standard general-to-specific search procedure that maintains an encompassing hierarchy through $F$-tests of specific models to the general model. It may be impossible to perform an $F$-test for the restrictions associated with a specific regression (one in which there are fewer regressors than the general regression) because it may have a different number of observations. Therefore, an *adjusted Schwarz criterion*.

Suppose $n$ regressions are run with $T_i$ ($j$ = 1, 2, …, $n$) observations. The adjusted Schwarz criterion is calculated as:

$$ASC_j = -2L_j/T_j + k_j \log(T_{min})/T_{min},$$

where $T_{min} = \min\{T_j,\}$ is the fewest number of observations among the regressions being compared

For the equivalent $F$-test, a set of restrictions is not rejected if the $ASC$ is lower in the restricted regression. Therefore the procedure is the same as for the simulations (we continue to do encompassing tests to choose the final specification). But, the $ASC$ is substituted anywhere an $F$-test would be run if the number of observations was the same throughout.

*A.* Five search paths are examined. Each path begins with the removal of one of the candidate variables with the (5) lowest $t$-statistics in the current general specification. All $t$-statistics are calculated using White's heteroscedasticity-corrected standard errors. The first search begins by re-estimating the regression. This re-estimated regression becomes the *current specification*. The search continues until it reaches a *terminal specification*

B. The current specification is estimated and all searchable variables are ranked according to their $t$-statistic. The searchable variable with the lowest $t$-statistic is removed.

C. Each current specification is subjected to the following battery of tests:

   i. Breusch-Pagan test for heteroscedasticity

ii. subsample stability test: an $F$-test for the equality of the variances of the first half versus the second half of the sample in which the data are ordered according to the value of the dependent variable. (This is analogous to a Chow test in a time-series context.) This test compares the regressions over each subsample to the regression over the full sample. If the degrees of freedom do not permit splitting the sample into equal subsamples, the test is replaced by one that compares a regression over the first $K$+2 observations one over the full sample. On both tests, see Chow (1960).
iii. An $F$-test of the hypothesis that the current specification is a valid restriction of the current general specification

D. The number of tests failed is recorded and the new specification becomes the current specification. Return to C until all remaining variables have a significant $t$-statistic.

E. If all variables are significant, and all of the tests in the test battery are passed, the current specification is the terminal specification and go to G. If any of the tests fails return to the last specification for which all the tests are passed and go to F.

F. The variable with the lowest $t$-statistic is eliminated. The resulting current specification is then subjected to the battery of tests.

i. If the current specification fails any one of these tests, the last variable eliminated is replaced, and the current specification is re-estimated eliminating the variable with the next lowest insignificant $t$-statistic.
ii. If the current specification passes all tests, return to F.
iii. The process of variable elimination ends when a current specification passes the battery of tests and either has all variables significant or cannot eliminate any remaining insignificant variable without failing one of the tests.

G. After a terminal specification has been reached, it is recorded and the next search path is tried until all have been searched.

H. Once all search paths have ended in a terminal specification, the *final specification* is chosen through a sequence of encompassing tests. We form the non-redundant joint model from each of the different terminal specifications; take all candidate specifications and perform the ASC test for encompassing (lower ASC than the union implies encompassing) the other specifications. If only one specification passes, it is the final specification. If more than one specification passes, the specification with the minimum ASC is the final specification. If no model passes, reopen the search on the non-redundant joint model (including testing against the general specification) using only a

**Appendix 3. The Sala-i-Martin (1997a,b) Data Set.**

The Sala-i-Martin data set was downloaded from his website (http://www.columbia.edu/cu/economics/) on 9/19/98. The data refer to a period of analysis of 1960-1992. Following instructions on the website, the documentation for sources and definitions given below is based on Sala-i-Martin (1997a,b). We believe that the documentation is not completely accurate. Some comments are recorded in square brackets, but we believe that there may be other discrepancies.

**Definitions of Variables in the Sala-i-Martin (1997a,b) Data Set**

| Variable | | | Definition | Source or Reference | Comments |
|---|---|---|---|---|---|
| A | B | Name[1] | | | |
| 39 | 17 | ABSLATIT | Absolute Lattitude. | Barro (1996) | |
| 40 | 26 | AGE | Age | | Average age of the population. [reported data does not appear to correspond to this description]. |
| 16 | 58 | AREA | Area (Scale Effect). | Barro and Lee (1993) | Total area of the country. |
| 32 | 39 | ASSASSP2 | Political Assassinations | Barro and Lee (1993) | Number of political assassinations. |
| 15 | 31 | BMP1 | Black Market Premium | Barro and Lee (1993) | log (1+Black Market Premium). |
| 9 | 12 | BMS6087 | Standard Deviation of the Black Market Premium | Levine & Renelt (1992). | 1960-89. |
| 41 | 55 | BRIT | British Colony | Barro (1996) | Dummy variable for former British colonies. |
| 42 | 20 | BUDDHA | Fraction of Buddhist | Barro (1996) | |
| 43 | 21 | CATH | Fraction of Catholic | Barro (1996) | |
| 37 | 9 | CIVLIBB | Civil Liberties | Knack and Keefer (1995) | Index. |
| 44 | 3 | CONFUC | Fraction of Confucius | Barro (1996) | |
| 45 | 47 | DEMOC65 | Index of Democracy | Knack and Keefer (1995) | 1965; qualitative index of democratic freedom. |
| 57 | 33 | DPOP6090 | Growth Rate of Population | Barro and Lee (1993) | 1960-90. |
| 2 | 14 | ECORG | Degree of Capitalism | Hall and Jones (1996). | Index of degree in which economies favor capitalist forms of production |
| 4 | 24 | ENGFRAC | Fraction of Population Able to Speak English | Hall and Jones (1996) | |
| 7 | 1 | EQINV | Equipment Investment | Delong and Summers (1991) | |
| 46 | 45 | FRAC | Ethnolinguistic Fractionalization | Easterly and Levine (1997)[2] | Probability two random people in a country do not speak same language. |
| 17 | 49 | FREEOP | Free Trade Openness | Barro and Lee (1993) | |
| 18 | 48 | FREETAR | Tariff Restrictions | Barro and Lee (1993) | Degree of tariff barriers. |
| 47 | 38 | FRENCH | French Colony | Barro (1996) | Dummy variable for former French colonies. |
| 10 | 57 | GDC | Growth of Domestic Credit | Levine and Renelt (1992) | 1960-90. |
| 31 | 25 | GDE1 | Defense Spending Share | Barro and Lee (1993) | Fraction of GDP |
| 58 | * | GDPSH60L | log(GDP per capita 1960). | Barro and Lee (1993) | log(Summers-Heston GDP per capita in 1960). |

| | | | | | |
|---|---|---|---|---|---|
| 30 | 53 | GEEREC1 | Government. Education Spending Share | Barro and Lee (1993) | Fraction of GDP |
| 29 | 23 | GGCFD3 | Public Investment Share | Barro and Lee (1993) | Fraction of GDP |
| 56 | ** | GR56092 | Growth Rate of GDP per capita | Summers/Heston Data; Penn World Tables | 1960-90; the dependent variable. |
| 64 | 27 | GVXDXE52 | Public Consumption Share | Barro and Lee (1993) | Public consumption minus education and defense as fraction of GDP |
| 27 | 54 | H60 | Higher Education. Enrollment | Barro and Lee (1993) | 1960 |
| 48 | 42 | HINDU | Fraction of Hindu | Barro (1996) | |
| 23 | 50 | HYR60 | Average Years of Higher School | Barro and Lee (1993) | . Average years of higher education of total population in 1960. |
| 24 | 43 | HUMAN60 | Average Years of Schooling | Barro and Lee (1993) | 1960; called "H" in definition of HUMANYL. |
| 59 | 41 | HUMANYL | H*log(GDP60). | Barro and Lee (1993) | Product of average years of schooling and log(GDP per capita in 1960). |
| 49 | 35 | JEW | Fraction of Jewish | Barro (1996) | |
| 19 | 7 | LAAM | Latin American Dummy | | Dummy for Latin American countries. |
| 62 | 29 | LFORCE60 | Size Labor Force (Scale Effect). | Barro and Lee (1993) | |
| 28 | * | LIFEEO60 | Life Expectancy | Barro and Lee (1993) | 1960 |
| 60 | 36 | LLY1 | Liquid Liabilities to GDP | King and Levine (1993) | Ratio of liquid liabilities to GDP (a measure of financial development). |
| 1 | 11 | MINING | Fraction of GDP in Mining | Hall and Jones (1996) | |
| 50 | 5 | MUSLIM | Fraction of Muslim | Barro (1996) | |
| 8 | 16 | NONEQINV | Non-Equipment Investment | Delong and Summers (1991) | |
| 5 | 30 | OTHFRAC | Fraction of Population Able to Speak a Foreign Language | | |
| 25 | * | P60 | Primary School Enrollment | Barro and Lee (1993) | 1960. |
| 11 | 28 | PI | Average Inflation Rate | Levine and Renelt (1992) | 1960-90. |
| 34 | 52 | PINSTAB2 | Political Instability | Knack and Keefer (1995) | |
| 51 | 13 | PRIEXP70 | Primary Exports | Sachs and Warner (1996) | Fraction of primary exports in total exports in 1970. |
| 36 | 6 | PRIGHTSB | Political Rights | Barro (1996) | |
| 52 | 19 | PROT | Fraction of Protestant | Barro (1996) | |
| 21 | 37 | PRY60 | Average Years of Primary School | Barro and Lee (1993) | 1960 |
| 6 | 18 | RERD | Exchange Rate Distortions | Barro and Lee (1993) | [Actual Source: Levine and Renelt(1992)] |
| 33 | 10 | REVCOUP | Revolutions and Coups | Barro and Lee (1993) | Number of military coups and revolutions. |
| 53 | 4 | RULELAW | Rule of Law | Barro (1996) | |
| 26 | 44 | S60 | Secondary School Enrollment | Barro and Lee (1993) | 1960. |
| 20 | 8 | SAAFRICA | Sub-Sahara African Dummy | | Dummy for sub-Sahara African countries. |
| 12 | 46 | SCOUT | Outward Orientation | Levine and Renelt (1992) | |

| | | | | | |
|---|---|---|---|---|---|
| 54 | 22 | SPAIN | Spanish Colony | Barro (1996 | Dummy variable for former Spanish colonies |
| 13 | 40 | STDC | Standard Deviation of Domestic Credit | King and Levine (1993) | 1960-90. |
| 14 | 32 | STPI | Standard Deviation of Inflation | Levine and Renelt (1992) | 1960-90. |
| 22 | 51 | SYR60 | Average Years of Secondary School | Barro and Lee (1993) | 1960 |
| 38 | 59 | TOT1 | Terms of Trade Growth | Barro and Lee (1993). | 1960-90. |
| 55 | 56 | URB60 | Urbanization Rate | Barro and Lee (1993) | Fraction of population living in cities. |
| 35 | 15 | WARDUM | War Dummy | Barro and Lee (1993) | Dummy for countries that have been involved in war any time between 1960 and 1990. |
| 61 | 34 | WORK60L | Ratio Workers to Population | Barro and Lee (1993) | [Apparently logged |
| 63 | | X0 | | | Duplicates Variable 56. |
| 3 | 2 | YRSOPEN | Number of Years Open Economy | Sachs and Warner (1996) | Index. |

Notes:

Column A under "Variables" indicates the order of the variable in the data set downloaded from Sala-i-Martin's website.

Column B under "Variables" indicates the order in which the variables appears in Sala-i-Martin (1997b) "Appendix 1: Descriptions and Sources of Variables."

*One of three "free variables" included in all of Sala-i-Martin's regressions, but not numbered in Appendix 1 of Sala-i-Martin (1997b).

**Dependent variable, not numbered in Appendix 1 of Sala-i-Martin (1997b).

[1]Variable names taken from the headers of downloaded data set, converted to upper-case.

Sala-i-Martin (1997a,b) indicates the references to the following sources imply that data were downloaded as indicated:

Barro and Lee (1993): National Bureau of Economic Research Web Page.
Delong and Summers (1991): World Bank Research Department Web Page.
Hall and Jones (1996): Charles Jones's Web Page.
Knack and Keefer (1996): supplied by Robert Barro.
Levine and Renelt (1992): World Bank Research Department Web Page.
Sachs and Warner (1996): supplied by Andrew Warner.

[2]Given originally as Easterly and Levine (1996), which is now published as Easterly and Levine (1997).

**Country Coverage for the Sala-i-Martin Data Set**

| No. | Country | No. | Country | No. | Country |
|---|---|---|---|---|---|
| 1 | Algeria | 41 | Tanzania | 81 | Bangladesh |
| 2 | Angola * | 42 | Togo | 82 | Myanmar (Burma) * |
| 3 | Benin | 43 | Tunisia | 83 | China * |
| 4 | Botswana | 44 | Uganda | 84 | Hong Kong |
| 5 | Burkina Faso | 45 | Zaire | 85 | India |
| 6 | Burundi | 46 | Zambia | 86 | Indonesia |
| 7 | Cameroon | 47 | Zimbabwe | 87 | Iran |
| 8 | Cape Verde * | 48 | Bahamas, The * | 88 | Iraq |
| 9 | Central African Republic | 49 | Barbados * | 89 | Israel |
| 10 | Chad | 50 | Canada | 90 | Japan |
| 11 | Comoros * | 51 | Costa Rica | 91 | Jordan |
| 12 | Congo | 52 | Dominica * | 92 | Korea |
| 13 | Egypt | 53 | Dominican Republic | 93 | Kuwait * |
| 14 | Ethiopia | 54 | El Salvador | 94 | Malaysia |
| 15 | Gabon | 55 | Grenada * | 95 | Nepal |
| 16 | Gambia | 56 | Guatemala | 96 | Oman * |
| 17 | Ghana | 57 | Haiti | 97 | Pakistan |
| 18 | Guinea * | 58 | Honduras | 98 | Philippines |
| 19 | Guinea Bissau | 59 | Jamaica | 99 | Saudi Arabia * |
| 20 | Cote d'Ivoire * | 60 | Mexico | 100 | Singapore |
| 21 | Kenya | 61 | Nicaragua | 101 | Sri Lanka |
| 22 | Lesotho | 62 | Panama | 102 | Syria |
| 23 | Liberia * | 63 | St. Lucia * | 103 | Taiwan |
| 24 | Madagascar | 64 | St. Vincent & Grenadines * | 104 | Thailand |
| 25 | Malawi | 65 | Trinidad & Tobago | 105 | United Arab Emirates * |
| 26 | Mali | 66 | United States | 106 | Yemen N.Arab * |
| 27 | Mauritania | 67 | Argentina | 107 | Austria |
| 28 | Mauritius | 68 | Bolivia | 108 | Belgium |
| 29 | Morocco | 69 | Brazil | 109 | Cyprus |
| 30 | Mozambique | 70 | Chile | 110 | Denmark |
| 31 | Niger | 71 | Colombia | 111 | Finland |
| 32 | Nigeria | 72 | Ecuador | 112 | France |
| 33 | Rwanda | 73 | Guyana | 113 | Germany, West |
| 34 | Senegal | 74 | Paraguay | 114 | Greece |
| 35 | Seychelles * | 75 | Peru | 115 | Hungary * |
| 36 | Sierra Leone * | 76 | Suriname * | 116 | Iceland * |
| 37 | Somalia * | 77 | Uruguay | 117 | Ireland |
| 38 | South Africa | 78 | Venezuela | 118 | Italy |
| 39 | Sudan * | 79 | Afghanistan * | 119 | Luxembourg * |
| 40 | Swaziland | 80 | Bahrain * | 120 | Malta * |

continued next page

| No. | Country |
|---|---|
| 121 | Netherlands |
| 122 | Norway |
| 123 | Poland * |
| 124 | Portugal |
| 125 | Spain |
| 126 | Sweden |
| 127 | Switzerland |
| 128 | Turkey |
| 129 | United Kingdom |
| 130 | Yugoslavia |
| 131 | Australia |
| 132 | Fiji |
| 133 | New Zealand |
| 134 | Papua New Guinea |
| 135 | Solomon Islands * |
| 136 | Tonga * |
| 137 | Vanuatu * |
| 138 | Western Samoa * |

Note: Country number corresponds to the order in the downloaded data set, identical to the order in Barro and Lee (1993).
*Indicates country is removed because either dependent variable or one of the free variables is missing for it.

## References


Barro, Robert J. (1991) "Economic Growth in a Cross-section of Countries," *Quarterly Journal of Economics* 106(2), May, 407-433.

Barro, Robert J. (1996) "Determinants of Democracy," unpublished typescript, Harvard University.

Barro, Robert J. (1997) *Determinants of Economic Growth.* Cambridge, MA: MIT Press.

Barro, Robert J. and Jong-Wha Lee. (1993) "International Comparisons of Educational Attainment," *Journal of Monetary Economics* 32(3), December, 363-394.

Barro, Robert J. and Xavier X. Sala-i-Martin. (1995) *Economic Growth.* New York: McGraw Hill.

Breusch, T.S. and A.R. Pagan. (1980).The Lagrange multiplier test and its application to model specification in econometrics, *Review of Economics Studies* 47, 239-53.

Brownstone, David and Camilla Kazimi. (1999) "Applying the Bootstrap," unpublished typescript, University of California, Irvine.

Chow, Gregory C. (1960) "Tests of Equality Between Sets of Coefficients in Two Linear Regressions," *Econometrica*, 28, 591-605.

Delong, J. Bradford and Lawrence H. Summers. (1991) "Equipment Investment and Economic Growth," *Quarterly Journal of Economics* 106(2), May, 445-502.

Easterly, William and Ross Levine (1997) "Africa's Growth Tragedy: Policies and Ethnic Divisions," *Quarterly Journal of Economics* 112(4), November, 1203-50.

Ericsson, Neil R., Julia Campos and H-A Tran (1990) "PC-GIVE and David Hendry's Econometric Methodology," *Revista de Econometrica* 10, 7-117.

Faust, Jon and Charles H. Whiteman. (1995) "Commentary [on Grayham E. Mizon "Progressive Modeling of Macroeconomic Times Series: The LSE Methodology]," in Kevin D. Hoover (ed.) *Macroeconometrics: Developments, Tensions and Prospects.* Boston: Kluwer, 1995, pp. 171-180.

Faust, Jon and Charles H. Whiteman. (1997) "General-to Specific Procedures for Fitting a Data-Admissible, Theory-Inspired, Congruent, Parsimonious, Encompassing, Weakly-Exogenous, Identified, Structural Model to the DGP: A Translation and Critique," *Carnegie-Rochester Conference Series on Economic Policy*, 47, December, 121-61.

Gilbert, Christopher L. (1986) "Professor Hendry's Econometric Methodology," *Oxford Bulletin of Economics and Statistics*, 48, August, 283-307. Reprinted in Granger (1990).

Greene, William H. (1997) *Econometric Analysis*, 3rd ed. Upper Saddle River, N.J.: Prentice Hall.

Grier, Kevin B. and Gordon Tullock. (1989) "An Emprical Analysis of Cross-national Economic Growth," *Journal of Monetary Economics* 24(2), September, 259-276.

Griliches, Zvi (1986) "Economic Data Issues," in Zvi Griliches (ed.) and Michael D. Intriligator (ed.) *Handbook of Econometrics*, vol. 3. Amsterdam: North-Holland, pp. 1466-1514.

*Government Finance Statistics Yearbook*. Washington, D.C.: International Monetary Fund, various years.

Hall, Robert and Charles Jones. (1996) "The Productivity of Nations," NBER Working Paper No. 5812, November.

Hansen, Bruce E. (1996) "Methodology: Alchemy or Science?" *Economic Journal* 106(438), September, 1398-1431.

Hansen, Bruce E. (1999) "Discussion of 'Data Mining Reconsidered'" *Econometrics Journal*, 2(2), 1-25.

Hendry, David F. (1987) "Econometric Methodology: A Personal Viewpoint," in Truman Bewley (ed.) *Advances in Econometrics*, Vol. 2. Cambridge: Cambridge University Press.

Hendry, David F. (1988) "Encompassing," *National Institute Economic Review*, (August), 88-92.

Hendry, David F. (1995) *Dynamic Econometrics*. Oxford: Oxford University Press.

Hendry, David F. (1997) "On Congruent Econometric Relations: A Comment." *Carnegie-Rochester Conference Series on Public Policy, Vol. 47, 163-190.*

Hendry, David F. and Hans-Martin Krolzig. (1999) "Improving on 'Data Mining Reconsidered' by K.D. Hoover and S.J. Perez," *Econometrics Journal*, 2(2), 1-25.

Hendry, David F., Edward E. Leamer and Dale J. Poirier. (1990) "The ET Dialogue: A Conversation on Econometric Methodology," *Econometric Theory* 6, 171-261.

Hendry, David F. and Jean-François Richard. (1987) "Recent Developments in the Theory of Encompassing," in Bernard Cornet and Henry Tulkens (eds.) *Contributions to Operations Research and Economics: The Twentieth Anniversary of Core*. MIT Press: Cambridge, MA.

Hoover, Kevin D. (1995) "In Defense of Data Mining: Some Preliminary Thoughts," in Kevin D. Hoover and Steven M. Sheffrin (eds.) *Monetarism and the Methodology of Economics: Essays in Honour of Thomas Mayer*. Aldershot: Edward Elgar, 1995.

Hoover, Kevin D. and Stephen J. Perez. (1999) "Data Mining Reconsidered: Encompassing and the General-to-Specific Approach to Specification Search," *Econometrics Journal* 2(2), 1-25.

Hoover, Kevin D. and Stephen J. Perez. (2000) "Three Attitudes Towards Data-mining," *Journal of Economic Methodology* 7(2), 195-210.

Horowitz, J.L. (1997) "Bootstrap Methods in Econometrics: Theory and Numerical Performance," in R. Kreps and K. Wallis (eds.) *Advances in Economics and Econometrics: Seventh World Congress of the Econometric Society,* vol. 3. Cambridge: Cambridge University Press.

*International Financial Statistics*. Washington, D.C. International Monetary Fund, various years.

Jones, Charles I. (1998) *Introduction to Economic Growth.* New York: Norton.

King, Robert G. and Ross Levine. (1993) "Finance, Entrepeneurship, and Growth: Theory and Evidence," *Journal of Monetary Economics* 32(3), December 1993, 513-42.

Knack, Stephen and Philip Keefer. (1995) "Institutions and Economic Performance: Cross-Country Tests Using Alternative Institutional Measures," *Economics and Politics* 7(3), November, 207-27.

Kormendi, Roger C. and Philip G. Meguire. (1985) "Macroeconomic Determinants of Growth," *Journal of Monetary Economics* 16(2), September, 141-163.

Krolzig, Hans-Martin and David F. Hendry. (1999) "Computer Automation of General-to-Specific Model Selection Procedures," unpublished typescript Institute of Economics and Statistics and Nuffield College, Oxford University.

Leamer, Edward E. (1983) "Let's Take the Con Out of Econometrics," *American Economic Review* 73(1), 31-43.

Leamer, Edward E. (1985) "Sensitivity Analysis Would Help," *American Economic Review* 75(3), 308-313.

Leamer, Edward E. and Herman Leonard. (1983) "Reporting the Fragility of Regression Estimates," *Review of Economics and Statistics* 65(2), May, 306-317.

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

Little, Roderick J. A. and Donald B. Rubin. (1987) *Statistical Analysis with Missing Data.* New York : Wiley.

Lovell, Michael C. (1983) "Data Mining," *The Review of Economics and Statistics*, Vol. 65, No. 1 (February), 1-12.

Lucas, Robert E., Jr. (1988) “On the Mechanics of Economic Development,” *Journal of Monetary Economics* 22(1), July, 3-42.

Mankiw, N. Gregory, David Romer and David N. Weil. (1992) “A Contribution to the Empirics of Economic Growth.” *Quarterly Journal of Economics*, 107, May, 407-437.

McAleer, Michael, Adrian Pagan and Paul A. Volcker. (1985) “What Will Take the Con Out of Econometrics,” *American Economic Review* 75(3), June, 293-307.

Mizon, Grayham E. (1977) “Model Selection Procedures,” in M.J. Artis and R.A. Nobay (eds.) *Studies in Modern Economic Analysis*, ch. 4. Oxford: Blackwell.

Mizon, Grayham E. (1984) “The Encompassing Approach in Econometrics,” in D.F. Hendry and K.F. Wallis (eds.) *Econometrics and Quantitative Economics.* Oxford: Basil Blackwell, pp. 135-172.

Mizon, Grayham E. (1995) “Progressive Modelling of Macroeconomic Time Series: The LSE Methodology,” in Kevin D. Hoover (ed.) *Macroeconometrics: Developments, Tensions and Prospects.* Boston: Kluwer, 1995, pp. 107-170.

Mizon, Grayham E. and Jean-François Richard (1986) “The Encompassing Principle and Its Application to Testing Non-nested Hypotheses,” *Econometrica* 54, 657-678.

Pagan, Adrian (1987) "Three Econometric Methodologies: A Critical Appraisal," *Journal of Economic Surveys,* Vol. 1, No. 1, 3-24. Reprinted in Granger (1990).

*Pick’s Currency Yearbook*. New York: Pick Publishing Corp., various years.

Phillips, Peter C. B. (1988) "Reflections on Econometric Methodology," *Economic Record* 64, 334-359.

Romer, Paul M. (1986) “Increasing Returns and Long-run Growth,” *Journal of Political Economy* 94(5), October, 1002-1037.

Sachs, Jeffrey D. and Andrew M. Warner (1995) “Economic Reform and the Process of Economic Integration,” *Brookings Papers on Economic Activity* No. 1, August, 1-95.

Sachs, Jeffrey D. and Andrew M. Warner. (1996) “Natural Resource Abundance and Economic Growth,” unpublished typescript, Harvard Institute International Development.

Sala-i-Martin, Xavier X. (1997a) “I Have Just Run Two Million Regressions,” *American Economic Review* 87(2), 178-183.

Sala-i-Martin, Xavier X. (1997a) “I Have Just Run Four Million Regressions,” unpublished typescript, Economic Department , Columbia University.

Schafer, J. L. (1997) *Analysis of Incomplete Multivariate Data*. London and New York: Chapman and Hall.

Solow, Robert M. (1956) “A Contribution to the Theory of Economic Growth,” *Quarterly Journal of Economics* 70(1), February, 65-94.

Summers, Robert and Alan Heston. (1988) “A New Set of International Comparisons of Real Product and Price Levels Estimates for 130 Countries, 1950-1985,” *Review of Income and Wealth* 34 (1), March, 1-25.

Syrquin, Moshe and Chenery Hollis. (1989) “Patterns of Development, 1950-1983,” World Bank Working Paper Series No. 41.

Temple, Jonathan. (1999) “Growth Regressions and What the Textbooks Don’t Tell You,” unpublished typescript, Nuffield College and the Institute of Economics and Statistics, Oxford University, 7 October 1999.

*World Bank National Accounts*, Washington, D.C., various years.

*World Bank Social Indicators*, Washington, D.C., various years.

Wu, C.F.J. (1986) “Jackknife, Bootstrap and Other Resampling Methods in Regression Analysis,” *Annals of Statistics* 14(4), 1261-1295.

## Table 1. The Efficacy of Three Search Algorithms

| Models with: | Extreme-bounds analysis | | Modified Extreme-bounds analysis | | General-to-Specific | |
|---|---|---|---|---|---|---|
| | Size Ratio[1] | Power Ratio[2] | Size Ratio[1] | Power Ratio[2] | Size Ratio[1] | Power Ratio[2] |
| **0 true variables** | 0.31 | | 1.42 | | 1.08 | |
| **3 true variables** | 0.01 | 0.47 | 5.44 | 1.07 | 1.07 | 0.90 |
| **7 true variables** | 0.00 | 0.15 | 6.06 | 0.79 | 1.20 | 0.88 |
| **14 true variables** | 0.00 | 0.05 | 6.40 | 0.61 | 1.28 | 0.86 |

Notes: The basic data are a pool of 34 variables described in Appendix 1. For each number of true variables, 20 models are specified by choosing the indicated number of regressors at random from the pool. Coefficients are calibrated from a regression of the chosen regressors on the actual average growth rate. 100 dependent variables are created from the same regressors and coefficients and error terms constructed with a wild bootstrap procedure from the errors of the calibrating regression. Specification searches are then conducted by each of the three methods and the number of type I and type II errors are recorded. Statistics reported here average over each of the 100 simulations for each of the 20 models. Details of the simulations and the search procedures are found in Section 2 and Appendix 2a.

[1]Size is calculated as the proportion of incorrect variables included (significantly for general-to-specific) to the total possible incorrect variables. The Size Ratio is average ratio of the size to the nominal size (0.05) used as the critical value in all the hypothesis tests in the search procedures. A Size Ratio of 1.00 indicates that on average the size is equal to the nominal size (0.05).

[2]Power is calculated as the proportion of times a true variables is included (significantly for the general-to-specific procedure). The true (simulated) power is based on the number of type II errors made in 100 simulations of the true model without any search. The Power Ratio is the average ratio of power to true (simulated) power. A Power Ratio of 1.00 indicates that on average the power is equal to the true (simulated) power. The Power Ratio is not relevant when there are no true variables.

**Table 2. Regressions of the Growth Rate of GDP Per Capita (GYP) Using Levine and Renelt 1992 Data**

| Variable | | Specification Search Method | | | | | |
|---|---|---|---|---|---|---|---|
| | | General-to-Specific | | | Modified Extreme Bounds | | |
| Mnemonic | Name | Coefficient | Standard Error | *p*-value | Coefficient | Standard Error | *p*-value |
| Constant | | 0.189 | 0.137 | 0.170 | -0.922 | 0.940 | 0.329 |
| DCPYL | Ratio Private Central Bank Claims to GDP | 0.283 | 0.153 | 0.0701 | | | |
| GPO | Population Growth | -0.998 | 0.0347 | 0.000 | | | |
| M | Import Share of GDP | -0.665 | 0.564 | 0.243 | | | |
| MSG | Growth of Import Share | -0.912 | 0.0249 | 0.000 | | | |
| X | Export Share of GDP | 1.24 | 0.865 | 0.157 | | | |
| XSG | Growth of Export Share of GDP | 0.0416 | 0.0137 | 0.00350 | | | |
| GM | Growth of Imports | 0.892 | 0.0160 | 0.000 | 0.0668 | 0.0749 | 0.375 |
| RGDP60 | Real Per Capita GDP in 1960 | -0.0482 | 0.0187 | 0.0125 | -0.238 | 0.103 | 0.0233 |
| AFRICA | Subsaharan African Country | | | | -1.13 | 0.454 | 0.0145 |
| BMP | Black Market Premium | | | | -0.00346 | 0.00169 | 0.0434 |
| GX | Growth of Export Share | | | | 0.146 | 0.0681 | 0.0353 |
| INV | Investment Share | | | | 6.83 | 2.81 | 0.0173 |
| LAAM | Latin American Country | | | | -1.08 | 0.370 | 0.0045 |
| PRI | Primary School Enrollment in 1960 | | | | 1.15 | 0.667 | 0.0874 |
| SEC | Secondary Enrollment Rate 1960 | | | | 0.328 | 1.25 | 0.793 |
| SED | Secondary Enrollment Rate 1970 | | | | 1.15 | 1.26 | 0.364 |
| **Summary Statistics** | | | | | | | |
| $\bar{R}^2$ | | | 0.98 | | | 0.69 | |
| Standard Error of Regression | | | 0.214 | | | 1.00 | |
| Sum of Squared Residuals | | | 2.95 | | | 86.9 | |
| Number of Observations | | | 73 | | | 97 | |
| Mean of the Dependent Variable | | | 2.14 | | | 2.00 | |
| Standard Deviation of the Dependent Variable | | | 1.65 | | | 1.80 | |
| Log-Likelihood | | | 13.6 | | | -132 | |
| Adjusted Schwarz Criterion[1] | | | 0.156 | | | 3.37 | |
| F-statistic for All Regressors | | | 527 | | | 22.2 | |

Note: Shaded areas identify variables common to both search procedures.

[1] $ASC = -2L/T + k\log(T_{min})/T_{min}$, where $T_{min}$ is the fewest number of observations among the regressions being compared. In this case, regressions using the variables from the general-to-specific and modified extreme bounds searches are compared and $T_{min} = 73$.

**Table 3: Extreme Bounds Analysis and Modified Extreme Bounds Analysis of Levine and Renelt (1992) Data**

| Mnemonic | Lower Extreme | Upper Extreme | Fraction Significant | Beta | Standard Deviation | CDF Normal[1] | CDF Non-Normal (weighted)[2] | CDF Non-Normal (Not Weighted)[3] |
|---|---|---|---|---|---|---|---|---|
| | | | | Focus Variables | | | | |
| GX | -0.024 | 1.542 | 0.999 | 0.256 | 0.038 | 1.000 | 1.000 | 1.000 |
| LAAM | -2.726 | 0.447 | 0.958 | -1.087 | 0.344 | 0.999 | 0.995 | 0.995 |
| GM | -0.876 | 0.974 | 0.918 | 0.265 | 0.047 | 1.000 | 0.990 | 0.990 |
| SED | -4.539 | 7.407 | 0.864 | 3.046 | 1.177 | 0.995 | 0.983 | 0.984 |
| AFRICA | -2.653 | 0.747 | 0.753 | -0.928 | 0.377 | 0.993 | 0.981 | 0.979 |
| SEC | -4.286 | 7.560 | 0.836 | 2.958 | 1.272 | 0.990 | 0.962 | 0.965 |
| BMP | -0.035 | 0.027 | 0.609 | -0.006 | 0.003 | 0.961 | 0.952 | 0.951 |
| BMS | -0.019 | 0.018 | 0.450 | -0.003 | 0.002 | 0.936 | 0.941 | 0.943 |
| GPO | -1.237 | 0.572 | 0.586 | -0.412 | 0.200 | 0.981 | 0.938 | 0.944 |
| CIVL | -0.629 | 0.336 | 0.349 | -0.207 | 0.121 | 0.956 | 0.930 | 0.924 |
| GPI | -0.046 | 0.024 | 0.239 | -0.012 | 0.007 | 0.943 | 0.924 | 0.926 |
| XSG | -1.575 | 0.349 | 0.267 | 0.019 | 0.054 | 0.638 | 0.919 | 0.926 |
| PI6089 | -0.139 | 0.035 | 0.222 | -0.005 | 0.004 | 0.869 | 0.913 | 0.916 |
| OECD | -2.038 | 2.624 | 0.358 | 0.780 | 0.518 | 0.934 | 0.895 | 0.897 |
| STPI | -0.008 | 0.028 | 0.055 | -0.001 | 0.001 | 0.723 | 0.885 | 0.886 |
| STDC | -0.022 | 0.012 | 0.109 | -0.003 | 0.002 | 0.901 | 0.883 | 0.882 |
| X | -10.209 | 6.972 | 0.145 | 0.967 | 1.018 | 0.829 | 0.867 | 0.859 |
| REVC | -3.090 | 2.216 | 0.127 | -0.803 | 0.653 | 0.891 | 0.860 | 0.853 |
| DCPYI | -2.524 | 4.452 | 0.076 | 1.318 | 1.110 | 0.883 | 0.860 | 0.861 |
| PRJ | -4.227 | 2.132 | 0.145 | -1.069 | 0.924 | 0.877 | 0.846 | 0.855 |
| M | -4.909 | 8.757 | 0.040 | 0.968 | 0.988 | 0.837 | 0.830 | 0.821 |
| LIT60 | -4.426 | 4.912 | 0.125 | 1.036 | 1.068 | 0.834 | 0.819 | 0.826 |
| MSG | -1.008 | 0.887 | 0.181 | -0.021 | 0.065 | 0.626 | 0.811 | 0.819 |
| POP70 | -0.005 | 0.010 | 0.000 | 0.002 | 0.002 | 0.814 | 0.797 | 0.794 |
| SOC | -1.708 | 2.821 | 0.002 | -0.225 | 0.442 | 0.695 | 0.757 | 0.758 |
| OIL | -2.806 | 1.916 | 0.068 | -0.417 | 0.592 | 0.760 | 0.746 | 0.760 |
| YRSCH | -0.130 | 0.185 | 0.001 | 0.031 | 0.047 | 0.744 | 0.744 | 0.735 |
| GDC | -0.029 | 0.074 | 0.004 | 0.001 | 0.007 | 0.569 | 0.706 | 0.708 |
| GOV | -11.733 | 15.686 | 0.001 | 1.276 | 3.224 | 0.654 | 0.687 | 0.695 |
| MIX | -0.973 | 0.865 | 0.000 | -0.077 | 0.279 | 0.608 | 0.658 | 0.659 |
| SCOUT | -0.766 | 1.190 | 0.001 | 0.087 | 0.285 | 0.620 | 0.657 | 0.664 |
| GSG | -0.312 | 0.197 | 0.005 | -0.023 | 0.067 | 0.634 | 0.649 | 0.664 |
| | | | | Free Variables | | | | |
| INV | -1.236 | 19.593 | 0.989 | 10.660 | 2.693 | 1.000 | 0.999 | 0.999 |
| PRI | -0.965 | 5.060 | 0.948 | 2.246 | 0.628 | 1.000 | 0.994 | 0.994 |
| RGDP60 | -0.825 | 0.182 | 0.624 | -0.221 | 0.105 | 0.982 | 0.938 | 0.937 |

Notes: See Sala-i-Martin (1997a,b) for a general discussion of the methods and the formulae used in this table. Heavy line divides variables into "robust" (above) and "non-robust" (below the line) on a 95-percent criterion.

[1]"CDF normal is the proportion of the CDF of the estimated coefficient assuming that the distribution of the estimator is normal.

[2]CDF non-normal (weighted) does not assume normality but weights the estimated CDFs using the integrated likelihood for each regression.

[3]CDF non-normal (non-weighted) does not assume normality and does not weight the CDFs.

**Table 4. The Importance of the Determinants of Growth Rates Based on the General-to-Specific Search of the Levine and Renelt (1992) Data**

| Variable [1] | | Characteristics of the Data [2] | | | Coefficient [1] | Effect on Growth Rate Attributable to Variable Evaluated at Its: [3] | | |
|---|---|---|---|---|---|---|---|---|
| Mnemonic | Name | Minimum | Median | Maximum | | Median [4] | Favorable Extreme (Marginal Effect) [5] | Unfavorable Extreme (Marginal Effect) [6] |
| M | Import Share of GDP | 1.02 | 6.19 | 19.8 | -0.665 | -4.12 | 3.44 | -9.07 |
| RGDP60 | Real Per Capita GDP in 1960 | 0.208 | 1.04 | 7.38 | -0.0482 | -0.050 | 0.04 | -0.306 |
| XSG | Growth of Export Share of GDP | -3.37 | 2.01 | 24.3 | 0.0416 | 0.0835 | 0.926 | -0.224 |
| X | Export Share of GDP | 0.0517 | 0.242 | 1.30 | 1.24 | 0.300 | 1.31 | -0.236 |
| GM | Growth of Imports | 1.02 | 6.19 | 19.8 | 0.892 | 5.53 | 12.2 | -4.62 |

Notes: Variables are listed in ascending order of the effect on the growth rate attributable to each variable evaluated at its median value.

[1]Variables and coefficient values are those that were statistically significantly different from zero at the 5-percent confidence level in the general-to-specific search reported in Table 3. Data characteristics and coefficient values reported to three significant digits.

[2]The characteristics of the data are computed using all available countries in the Levine and Renelt (1992) data set.

[3]Values are reported to two decimal places. Because of rounding, reported numbers may not correspond to the calculations described in notes 4-6.

[4]Equals Coefficient x Median

[5]The additional contribution to the growth rate of GDP from the variable evaluated at the extreme value (minimum or maximum, depending on the sign of the coefficient) above the contribution evaluated at the median.

[6]The additional loss to the growth rate of GDP from the variable evaluated at the extreme value (minimum or maximum, depending on the sign of the coefficient) below the contribution evaluated at the median.

**Table 5. Regressions of the Growth Rate of GDP Per Capita (GR) Using Sala-i-Martin (1997a, b) Data**

| Variable | | Specification Search Method | | | | | |
|---|---|---|---|---|---|---|---|
| | | General-to-Specific | | | Modified Extreme Bounds | | |
| Mnemonic | Name | Coefficient | Standard Error | *p*-value | Coefficient | Standard Error | *p*-value |
| | Constant | 5.04 | 1.10 | 0.000 | 8.01 | 1.86 | 0.000 |
| JEW | Fraction Jewish[1] | 0.035 | 0.00984 | 0.000 | | | |
| STDC | Standard Deviation of Domestic Credit | -0.013 | 0.00297 | 0.000 | | | |
| HYR | Average Years of Higher School | 0.492 | 0.7808 | 0.5317 | | | |
| STPI | Standard Deviation of Inflation | 0.00103 | 0.000459 | 0.0304 | | | |
| BMS | Standard Deviation of Black Market Premium for Exchange Rates | -0.00144 | 0.00157 | 0.366 | -0.000582 | 0.00103 | 0.576 |
| BUDDHA | Fraction Buddhist[1] | 0.0113 | 0.00503 | 0.0287 | 0.00960 | 0.00791 | 0.231 |
| CONFUC | Fraction Confucist[1] | 0.0546 | 0.00536 | 0.000 | 0.0513 | 0.0113 | 0.000 |
| EQINV | Equipment Investment (Fraction of GDP)[1] | 0.102 | 0.0351 | 0.0056 | 0.102 | 0.0305 | 0.0016 |
| GDPSH60L | Log(Per Capita GDP in 1960) | -1.26 | 0.227 | 0.000 | -1.68 | 0.2884 | 0.000 |
| LIFEE060 | Life Expectancy in 1960 | 0.492 | 0.7808 | 0.532 | 0.086 | 0.0228 | 0.04 |
| MUSLIM | Fraction Muslim[1] | 0.013815 | 0.002975 | 0.000 | 0.016206 | 0.5089 | 0.25 |
| NONEQINV | Nonequipment Investment (Fraction of GDP)[1] | 0.074379 | 0.023072 | 0.0023 | 0.048022 | 0.0324 | 0.144 |
| PROT | Fraction Protestant[1] | -0.00919 | 0.002012 | 0.000 | -0.00785 | 0.00420 | 0.0674 |
| RERD | Exchange Rate Distortion | -0.00882 | 0.00325 | 0.0094 | -0.00281 | 0.00364 | 0.4435 |
| REVCOUP | Revolution and Coups | 1.54 | 0.691 | 0.0302 | 0.4768 | 0.551 | 0.391 |
| RULELAW | Rule of Law (Index) | 01.22 | 0.3729 | 0.002 | 0.7243 | 0.512 | 0.164 |
| WARDUM | Wars (=1; =0 if none) | -0.412 | 0.2047 | 0.0499 | -0.269 | 0.217 | 0.222 |
| ABSLATIT | Absolute Latitude (Degrees) | | | | -0.00582 | 0.00909 | 0.525 |
| CATH | Fraction Roman Catholic[1] | | | | 0.000898 | 0.00460 | 0.846 |
| CIVLIBB | Civil Liberties Index (1 (high) – 7 (low)) | | | | -0.126 | 0.212 | 0.555 |
| ECORG | Degree of Capitalism (1 (high) – 5 (low)) | | | | 0.172 | 0.105 | 0.106 |
| LAAM | Latin American Country (=1; =0 otherwise) | | | | -0.00241 | 0.4502 | 0.996 |
| MINING | Mining (Fraction of GDP)[1] | | | | 0.0408 | 0.0143 | 0.0063 |
| P60 | Primary School Enrollment in 1960[1] | | | | 0.00785 | 0.0112 | 0.489 |
| PRIEXP70 | Primary Export as Fraction of Exports in 1970[1] | | | | -0.908 | 0.6343 | 0.159 |
| PRIGHTSB | Index of Political Rights (1 (high) – 7 (low)) | | | | 0.00747 | 0.183 | 0.968 |
| SAFRICA | Subsaharan African Country (=1, =0 otherwise) | | | | -0.888 | 0.4127 | 0.0365 |
| YRSOPEN | Years as an Open Economy | | | | 0.6504 | 0.4177 | 0.126 |

Notes: Shaded areas identify variables common to both search procedures.
[1]Original units of the data (natural fractions) converted to percentage points.

Table continued next page

## Table 5 continued

**Summary Statistics**

| | | |
|---|---|---|
| $\bar{R}^2$ | 0.89 | 0.86 |
| Standard Error of Regression | 0.548 | 0.690 |
| Sum of Squared Residuals | 0.138 | 0.228 |
| Number of Observations | 64 | 73 |
| Mean of the Dependent Variable [1] | 2.02 | 2.06 |
| Standard Deviation of the Dependent Variable [1] | 1.62 | 1.81 |
| Log Likelihood | 252 | 275 |
| Adjusted Schwarz Criterion [1] | -6.43 | -5.78 |
| F-statistic for All Regressors | 29.7 | 18.8 |

[1] $ASC = -2L/T + k\log(T_{min})/T_{min}$, where $T_{min}$ is the fewest number of observations among the regressions being compared. In this case, regressions using the variables from the general-to-specific and modified extreme bounds searches are compared and $T_{min} = 64$.

**Table 6. The Importance of the Determinants of Growth Rates**
**Based on the General-to-Specific Search of the Sala-i-Martin (1997a, b)Data**

| Variable [2] | | Characteristics of the Data [3] | | | Coefficient [2] | Effect on Growth Rate Attributable to Variable Evaluated at Its: [4] | | |
|---|---|---|---|---|---|---|---|---|
| Mnemonic | Name | Minimum | Median | Maximum | | Median [5] | Favorable Extreme (Marginal Effect) [6] | Unfavorable Extreme (Marginal Effect) [7] |
| GDPSH60L | Log(Per Capita GDP in 1960) | 5.52 | 7.20 | 9.19 | -1.26 | -9.05 | 2.12 | -2.49 |
| RERD | Exchange Rate Distortion | 51.0 | 116 | 277 | -0.00882 | -1.0231 | 0.573 | -1.42 |
| STDC | Standard Deviation of Domestic Credit | 2.73 | 14.7 | 590 | -0.013 | -0.184 | 0.149 | -7.19 |
| PROT | Fraction Protestant [1] | 0 | 5 | 98 | -0.00919 | -0.0460 | 0.0460 | -0.855 |
| JEW | Fraction Jewish [1] | 0 | 0 | 82 | 0.035 | 0.000 | 2.88 | 0.000 |
| BUDDHA | Fraction Buddhist [1] | 0 | 0 | 95 | 0.0113 | 0.000 | 1.08 | 0.000 |
| CONFUC | Fraction Confucist [1] | 0 | 0 | 60 | 0.0546 | 0.000 | 3.28 | 0.000 |
| WARDUM | Wars (=1; =0 if none) | 0.0 | 0.0 | 1.0 | -0.412 | 0.000 | 0.000 | -0.412 |
| STPI | Standard Deviation of Inflation | 1.76 | 8.23 | 2131 | 0.00103 | 0.00848 | 2.19 | -0.00666 |
| MUSLIM | Fraction Muslim [1] | 0 | 20 | 100 | 0.013815 | 0.0276 | 1.35 | -0.0276 |
| REVCOUP | Revolution and Coups | 0.000 | 0.120 | 1.190 | 1.54 | 0.185 | 1.65 | -0.185 |
| EQINV | Equipment Investment (Fraction of GDP) [1] | 0.000 | 2.91 | 14.800 | 0.102 | 0.296 | 1.22 | -0.294 |
| RULELAW | Rule of Law (Index) | 0.0 | 0.5 | 1.0 | 1.22 | 0.611 | 0.611 | -0.611 |

Notes: Variables are listed in ascending order of the effect on the growth rate attributable to each variable evaluated at its median value.
[1] Original units of the data (natural fractions) converted to percentage points.
[2] Variables and coefficient values are those that were statistically significantly different from zero at the 5-percent confidence level in the general-to-specific search reported in Table 3. Data characteristics and coefficient values reported to three significant digits.
[3] The characteristics of the data are computed using all available countries in the Sala-i-Martin (1997a,b) data set.
[4] Values are reported to two decimal places. Because of rounding, reported numbers may not correspond to the calculations described in notes 5-7.
[5] Equals Coefficient x Median
[6] The additional contribution to the growth rate of GDP per capita from the variable evaluated at the extreme value (minimum or maximum, depending on the sign of the coefficient) above the contribution evaluated at the median.
[7] The additional loss to the growth rate of GDP per capita from the variable evaluated at the extreme value (minimum or maximum, depending on the sign of the coefficient) below the contribution evaluated at the median.