Determinants of Economic Growth: Will Data Tell?

by

Antonio Ciccone and Marek Jarocinski*

October 2007

Abstract

Many factors inhibiting and facilitating economic growth have been suggested. Will international income data tell which matter when all are treated symmetrically a priori? We find that growth determinants emerging from agnostic Bayesian model averaging and classical model selection procedures are sensitive to income differences across datasets. For example, many of the 1975-1996 growth determinants according to World Bank income data turn out to be irrelevant when using Penn World Table data instead (the WB adjusts for purchasing power using a slightly different methodology). And each revision of the 1960-1996 PWT income data brings substantial changes regarding growth determinants. We show that research based on stronger priors about potential growth determinants is more robust to imperfect international income data.

Keywords: growth regressions, robust growth determinants
JEL Codes: E01, O47

---

* ICREA-Universitat Pompeu Fabra; and European Central Bank. Emails: antonio.ciccone@upf.edu and marek.jarocinski@ecb.int. We are grateful to Manuel Arellano, Tim Cogley, and Yuri Dikhanov for very useful comments. We also thank Eduardo Ley and Mark Steel for kindly making their programs available. The opinions expressed herein are those of the authors and do not necessarily represent those of the European Central Bank. Most of the results of this paper were generated using Ox. Ciccone gratefully acknowledges research support from the CREI and CREA research institutes.

1. Introduction

Why does income grow faster in some countries than in others? Most of the empirical work addressing this question focuses on a few explanatory variables to deal with the statistical challenges raised by the limited number of countries (see, for example, Kormendi and Meguire, 1985; Grier and Tullock, 1989; Barro, 1991; for reviews of the literature see Barro and Sala-i-Martin, 2003; Durlauf, Johnson, and Temple, 2005). Selected variables are motivated by their importance for theory or policy. But as researchers disagree on what is most important, there is usually only partial overlap among the variables considered in different empirical works.

It is therefore natural to try and see which of the explanatory variables suggested in the literature emerge as growth determinants when all are treated symmetrically a priori. The idea is to find out in which direction the data guides an agnostic. Following Sala-i-Martin (1997), this is the task tackled by Fernandez, Ley, and Steel (2001a), Sala-i-Martin, Doppelhofer, and Miller (2004), and Hendry and Krolzig (2004).[1] As many potential explanatory variables have been suggested, these agnostic empirical approaches inevitably need to start out with a long list of variables. We show that, as a result, the growth determinants emerging from these approaches turn out to be sensitive to seemingly minor variations in international income estimates across datasets. This is because strong conclusions are drawn from small differences in the R^2 across growth regressions. Small changes in relative model fit—due to Penn World Table income data revisions or methodological differences between the PWT and the World Bank income data for example—can therefore lead to substantial changes regarding growth determinants.

Given the difficulties faced when estimating international incomes, differences across datasets are not surprising. A major challenge is the limited coverage and quality of the

[1] The starting point of the literature on the sensitivity of cross-country growth empirics is Levine and Renelt (1992).

underlying price benchmark and national income data.[2] The Penn World Table income data—the dataset used in most cross-country empirical work—have therefore undergone periodic revisions to eliminate errors, incorporate improved national income data, or account for new price benchmarks.[3] For example, the 1960-1996 income data in the PWT 6.1 corrects the now-retired PWT 6.0 estimates for the same period.[4] Relative to previous revisions, changes seem minor.[5] Still, the two datasets yield rather different growth determinants using agnostic empirical approaches. Consider, for example, the Bayesian averaging of classical estimates (BACE) approach of Sala-i-Martin, Doppelhofer, and Miller (2004). SDM use this approach to obtain the 1960-1996 growth determinants with PWT 6.0 data. When we update their results with the corrected income data in PWT 6.1 it turns out that the two versions of the PWT disagree on 13 of 25 determinants of 1960-1996 growth that emerge using one of them. A further update using the latest PWT 6.2 data for 1960-1996 yields disagreement on 13 of 23 variables. A Monte Carlo study confirms that agnostic empirical approaches are sensitive to small income data revisions by PWT standards.

Disaccord concerning the 1960-1996 growth determinants across PWT income data revisions affects factors that have featured prominently in the literature. For instance, one BACE finding with PWT 6.0 income data that has been used for policy analysis is the negative effect of malaria prevalence in the 1960s on 1960-1996 growth (e.g. Sachs, 2005). But with PWT 6.1 or PWT 6.2 data for the same period, malaria prevalence in the 1960s becomes an insignificant growth factor. Another important policy issue is whether greater trade openness raises the rate of economic growth (e.g. Sachs and Warner, 1995; Rodrik and Rodriguez, 2001). But, according to the BACE criterion used by SDM, the number of years an economy has been open is a positive growth determinant with PWT 6.0 income data but not with PWT 6.1 or PWT 6.2 data. Prominent potential growth determinants that are significant for 1960-1996 growth according to both PWT 6.0 and PWT 6.1 income data

[2] Many income estimates are therefore obtained by extrapolation and margins of error for income levels and growth rates are often estimated to be large (e.g. Summers and Heston, 1991; Heston, 1994).

[3] PWT revisions are sometimes accompanied by a change in the base year for prices, which alone can change relative income estimates (e.g. Dowrick and Quiggin, 1997)

[4] See the PWT website page http://pwt.econ.upenn.edu/whatsnew.htm and our Web Appendix B for details.

[5] For example, the correlation between 1960-1996 income per capita growth rates in the two databases is 0.98. This is high compared to the 0.88 correlation of PWT 6.0 income per capita growth for 1960-1985 with the corresponding data from PWT 5.6.

but irrelevant when using PWT 6.2 data are life expectancy, the abundance of mining resources, the relative price of investment goods, and location in the tropics (for earlier results on these growth determinants and their policy implications see, for example, Barro, 1991; DeLong and Summers, 1991; Jones, 1994; Sachs and Warner, 1995; Congressional Budget Office, 2003). Fertility, on the other hand, is an interesting example going in the opposite direction. While insignificant when using PWT 6.0 income data, high fertility reduces 1960-1996 growth according to PWT 6.1 and 6.2 (for earlier results on the link between fertility and growth see, for instance, Barro, 1991, 1998; Barro and Lee, 1994). As the PWT income data will almost certainly undergo further revisions, it is important to be aware that past revisions have led to substantial changes in growth determinants.

Incomplete or erroneous price benchmark and national income data are not the only reasons why international income statistics are imperfect. As is well known, there is no best method for obtaining internationally comparable real income data. Different methods have advantages and disadvantages (e.g. Neary, 2004). For example, the PWT data adjusts for cross-country price differences using the Geary-Khamis method. The main advantage is that results for each country are additive across different levels of aggregation. This fits the PWT approach as aggregate income is obtained by summing over expenditures categories. The main disadvantage of the GK method is that country comparisons are based on a reference price vector (so-called international prices) which ends up being more characteristic for rich than poor countries (e.g. Kravis, Heston, and Summers, 1982; Neary, 2004; Dowrick, 2005). The World Bank's international income data uses the Elteto-Koves-Szulc method. This method does not rely on a single set of international prices but is instead based on bilateral income comparisons.[6]

Despite the difference in underlying methodological choices, the PWT and World Bank international income data are closely related.[7] For example, the correlation between 1975-1996 growth rates according to PWT 6.2 and the World Development Indicators is 96% and the correlation between 1975 income per capita levels is 97% (the WDI international income data starts in 1975). Still, all-inclusive agnostic empirical analysis often yields substantial disagreement on the determinants of 1975-1996 growth between the two

[6] The EKS method is also used by the Statistical Agency of the European Union and the OECD. The PWT and World Bank data also differ in the way they estimate incomes in non-benchmark countries and the extent to which they try to ensure consistency across price benchmarks.
[7] This is not surprising as they use the same national income and price benchmark data.

datasets. For instance, the PWT 6.2 and WDI income data yield disagreement on 8 of 15 variables when using SDM's Bayesian criterion; and the model selection approach of Hendry and Krolzig (2004) disagrees on 7 out of 10 variables.[8] It is important to be aware that the methodological choices underlying international income estimates have such strong effects on the conclusions of agnostic empirical analysis.

A potential strength of cross-country growth regressions focusing on few explanatory variables is that conclusions may be more robust across data revisions and alternative datasets. This is because, with fewer correlated variables, the conditioning of the ordinary least-squares regression problem can be expected to improve (see Belsley, Kuh, and Welsch, 1980, for example). To see whether this is the case for our application, we compare results of ordinary least-squares regressions across PWT income data revisions and also between PWT and WDI income data. Our results show that stronger priors about potential growth determinants translate into results that are less sensitive to data revisions and discrepancies across datasets. For example, when we consider cross-country regressions with 7 explanatory variables, a model size often found in the literature, we find that PWT and WDI results coincide on average on 4 of 5 variables that are statistically significant according to one of the datasets. Stronger priors about potential growth determinants also reduce the sensitivity of Bayesian model averaging and classical model selection procedures to imperfections in the income data. For instance, when we implement SDM's BACE approach starting with lists of only 18 candidate explanatory variables (instead of the 67 candidate variables SDM start out with), we find that PWT 6.1 and PWT 6.2 income data agree on average on 2 of 3 determinants of 1960-1996.

The agnostic empirical approaches considered here do not account for possible feedback from economic growth to the suggested explanatory variables (this could be done in principle, see Tsangarides, 2005, for example;[9] alternatively, the analysis could be restricted to potential growth determinants that can be considered exogenous). As a result, *robust growth correlates* might be a better name for what the literature refers to as growth determinants (or, sometimes, *robust* growth determinants). In any case, our objective here

[8] Comparing the results of Fernandez, Ley, and Steel's (2001a) approach across datasets requires some additional definitions as variables are not classified according to whether or not they are (*robust*) *growth determinants*.
[9] Tsangarides (2005) derives approximate formulas. An exact analysis could in principle follow Kleibergen and van Dijk (1998) and Kleibergen and Paap (2002). In practice, finding valid instruments is difficult, especially with multiple endogenous right-hand-side variables.

is to examine whether different income datasets point researchers looking for robust correlations in the same direction when many explanatory variables are treated symmetrically a priori. Whether these correlations reflect causal effects is an important but separate issue.

The remainder of the paper is structured as follows. Section 2 presents the three agnostic empirical approaches we consider and examines their sensitivity to income data using a small-scale Monte Carlo study. In Section 3, we use the agnostic approaches to obtain the 1960-1996 growth determinants according to the latest income estimates from the PWT 6.2. Our findings are compared with results using the 1960-1996 income data in PWT 6.1 and PWT 6.0. We also compare the 1975-1996 growth determinants according to the PWT 6.2 and WDI income data. Section 4 examines the effect of reducing the number of candidate explanatory variables on the sensitivity of results. Section 5 concludes. Appendix A describes our Monte Carlo study, and Appendix B points to three possible avenues for reducing the sensitivity of Bayesian model averaging to imperfections in the income data. More details on the data and our findings can be found in a Web Appendix.[10]

2. Agnostic Approaches to Growth Determinants

Consider the problem of identifying the determinants of economic growth across countries. If the number of countries (C) were large relative to the number of explanatory variables (K), we could find the statistically significant explanatory variables by regressing the growth rate of countries on all candidate variables. With C close to K, this approach tends to yield estimates that are too noisy to be of interest (the approach is infeasible when $C<K$). This has led to the use of two alternative approaches, *Bayesian model averaging* and *general-to-specific model selection*. We briefly discuss these approaches and illustrate their sensitivity using PWT data revisions.

A. Bayesian Approaches

Bayesian model averaging (BMA). Bayesian methods frame the problem of identifying the determinants of economic growth in terms of uncertainty about the true set of explanatory variables (*model uncertainty*). The Bayesian approach to model uncertainty is

[10] See http://www.antoniociccone.eu.

to first attach prior probabilities to alternative sets of explanatory variables and then update these probabilities using data.

To develop the Bayesian approach to growth determinants more formally it is useful to collect all candidate explanatory variables in a vector x_c, $c = 1,...,C$. The 2^K subsets of x_c are denoted by x_{jc}, $j = 1,...,2^K$, and called *models*. The cross-country growth regressions considered are of the form

$$y_c = a + x_{jc} b_j + e_{jc} \tag{1}$$

where y_c is the growth rate of per capita GDP in country *c*; *a* is the constant term; b_j is the effect of the explanatory variables in model *j* on growth; and e_{jc} is a Gaussian error term. The ingredients of BMA are: priors for models (p_j); priors for all parameters (*a*, b_j and the variance of the error term); and the likelihood function of the data for each model *j*. A key intermediate statistic is the likelihood of model *j* integrated with respect to the parameters using their prior distributions (the marginal likelihood of model *j*, $l_y(M_j)$). Bayesian approaches use Bayes' theorem to translate the density of the data conditional on the model (the marginal likelihood) into a posterior probability of the model conditional on the observed data,

$$p(M_j \mid y) = \frac{l_y(M_j) p_j}{\sum_{h \in \text{all models}} l_y(M_h) p_h}. \tag{2}$$

Detailed discussions of BMA can be found in Leamer (1978) and Hoeting et al. (1999) for example.

Agnostic BMA (BACE and BMA with benchmark priors). The idea of the agnostic Bayesian approaches to growth determinants of Fernandez, Ley, and Steel (2001b) and Sala-i-Martin, Doppelhofer, and Miller (2004) is to limit the subjectivity of Bayesian analysis. This is done by treating all candidate explanatory variables symmetrically a priori and by using coefficient priors that have a negligible effect on the posterior distribution of model coefficients (*loose priors*). FLS assume equal prior probabilities for all models, irrespective of model size and composition. SDM assume equal priors for models of the same size but favor models of a predetermined size (7 in their preferred specification). SDM specify non-informative priors (*improper priors*) for model coefficients that make posterior distributions equal to classical sampling distributions of ordinary least-squares coefficients.[11] This is why they refer to their approach as Bayesian averaging of classical estimates (BACE). FLS use priors proposed in Fernandez, Ley and Steel (2001a) (*benchmark priors*) that are designed to have a negligible effect on the posterior distribution of model coefficients.

SDM's choices of priors for model coefficients yield the following (approximate) marginal likelihood of the data

$$l_y(M_j) \propto C^{-\frac{k_j}{2}} SSE_j^{-\frac{C}{2}}, \tag{3}$$

where k_j is the number of candidate explanatory variables included in model j and SSE_j is the sum of squared ordinary least-squares residuals associated with the model. Hence, posterior probabilities of models are increasing in model fit and decreasing in the number of candidate variables included in the model. The marginal likelihood of the data in FLS (in their equation (8)) is

$$l_y(M_j) \propto \left(\frac{g}{g+1}\right)^{\frac{k_j}{2}} \left(\frac{1}{g+1} SSE_j + \frac{g}{g+1}(y-\bar{y})'(y-\bar{y})\right)^{-\frac{C-1}{2}}, \tag{4}$$

where $g = 1/\max\left\{C, K^2\right\}$, y is a vector collecting growth rates for all countries, and $\bar{y}$ is the average growth rate in the sample multiplied by a vector of ones.

BMA allows computation of a *posterior inclusion probability* for each of the K candidate explanatory variables. The posterior inclusion probability for a given candidate explanatory variable is obtained by summing posterior probabilities of all models including the variable. Other outputs of BMA are *posterior means conditional on inclusion* and *unconditional posterior means* of all coefficients. Posterior means conditional on inclusion are computed by averaging the coefficients of each variable over all models that include it, using weights equal to the posterior probabilities of the models. Unconditional means are computed similarly, but averaging over all models, taking the coefficient value of zero when the variable is not included.

Potential effects of data imperfections. Agnostic Bayesian approaches to growth determinants put much weight on the sum of squared errors when assigning posterior

[11] Such priors are not a proper distribution but emerge as a limit of a certain sequence of increasingly loose priors (Leamer, 1978).

inclusion probabilities to candidate explanatory variables. This can be seen immediately from the SDM marginal likelihood in (3) where the sum of squared errors (*SSE*) is raised to the power $-C/2$, which in SDM's case equals −44 as they have data on 88 countries. It is also true for (4), as the benchmark priors of FLS assume a very small value for g. As a result, posterior inclusion probabilities are sensitive to small changes in the sum of squared errors.

To get an initial sense of the magnitude of this effect, suppose that we want to determine the posterior inclusion probabilities of the 67 candidate explanatory variables considered by SDM (the data are available for 88 countries). To simplify, let us limit attention to models of a predetermined size. In this case, substituting (3) into (2) and summing across all models containing a given variable, yields that the posterior inclusion probability of a candidate explanatory variable v relative to w is

$$\text{(6)} \qquad \frac{\text{Posterior probability variable } v}{\text{Posterior probability variable } w} = \frac{\sum_{j \in \text{all models including variable } v} SSE_j^{-\frac{C}{2}}}{\sum_{j \in \text{all models including variable } w} SSE_j^{-\frac{C}{2}}}.$$

Now suppose that a data revision leads to a fall in the sum of squared errors generated by candidate variable v of 1.5% in all models. With data on 88 countries, this implies that the posterior inclusion probability of this variable almost doubles relative to other variables.

How strongly do actual data revisions affect posterior inclusion probabilities? To illustrate the effect we assume a predetermined model size of 1 and determine the posterior inclusion probability of the 67 variables in the SDM dataset with both PWT 6.1 and PWT 6.2 income data for the 1960-1996 period. In the top-left panel of Figure 1 we plot the R^2 of all 67 models including just the constant term and each of the variables (sorted by decreasing R^2) using PWT 6.1 data. In the top-right panel we display the corresponding posterior probabilities (computed as in SDM) for the first 16 variables. The comparison of these two panels illustrates how small differences in R^2 translate into large differences in inclusion probabilities. The best variable (which turns out to be the number of years countries have been open to international trade) has an R^2 of 0.33 and gets a posterior inclusion probability of 0.84. The sixth variable has an R^2 of 0.25 and an inclusion probability of around 0.006 (beyond the sixth variable inclusion probabilities are negligible). The bottom-left and bottom-right panels display the R^2 and inclusion probability for variables in the same order as in the top-left panel but using PWT 6.2 data.

It can be seen that changing the dataset perturbs the inclusion probabilities drastically when compared to changes in R^2. For example, number of years open, which was best with PWT 6.1, is almost irrelevant with PWT 6.2 (the inclusion probability of 0.03). The relevance of the second variable, the dummy for East Asian countries, jumps from 0.06 to 0.97.

When the predetermined model size is greater than 1, the posterior inclusion probability of a variable will be the sum of posterior inclusion probabilities across models containing the variable. What if data revisions were to lead to changes in the sum of squared errors that are unsystematic across models containing this variable? Will such data imperfections average out and therefore have small effects on posterior inclusion probabilities of variables? It turns out that they may not average out in theory and practice. To see this, note that when C is large, then $\sum_{j \in \text{all models including variable } v} SSE_j^{-\frac{C}{2}}$ is dominated by the sum of squared errors of the best fitting model (the model with the lowest SSE). In this case, we can therefore approximate the relative posterior inclusion probabilities in (6) by an expression that only involves the best fitting models for each variable

$$\frac{\text{Posterior probability variable } v}{\text{Posterior probability variable } w} \cong \frac{\max\left\{SSE_j^{-\frac{C}{2}} : j \in \text{all models including variable } v\right\}}{\max\left\{SSE_j^{-\frac{C}{2}} : j \in \text{all models including variable } w\right\}}.$$

Small unsystematic changes in the sum of squared errors across models can therefore have large effects on posterior inclusion probabilities of variables.

To illustrate the effect of the best fitting models on posterior inclusion probabilities for actual data revisions, we return to the example where we try to determine the posterior inclusion probabilities of the 67 candidate explanatory variables for economic growth considered by SDM with both PWT 6.1 and PWT 6.2 data. But now we take the predetermined model size to be 3. In this case, each variable is part of 66*65/2= 2145 models, which means that the posterior inclusion probability of a variable is the sum of posterior inclusion probabilities across 2145 models. A quick way to see whether this sum is sensitive to the sum of squared errors produced by the best fitting model is to examine rank correlations. We therefore rank variables according to their posterior inclusion probabilities (from highest to lowest probability) and also according to the median sum of squared errors across models that include the variable (from lowest to highest sum of squared errors). The simple correlation coefficient between these ranks is 71%. We then repeat the exercise but rank variables according to the sum of squared errors at the 5th percentile across models that include the variable (instead of the median). Now the correlation is 86%. When we rank variables according to the lowest sum of squared errors across models including the variable, the correlation is almost perfect (96%).

A useful alternative perspective on the sensitivity of posterior inclusion probabilities to the sum of squared errors of the best-fitting model can be obtained in two steps. We first determine the sum of squared errors of models including a given candidate explanatory variable at the 1st, 5th, 25th, and 50th percentiles of the distribution. Then we regress the log posterior inclusion probability of all variables on the log sum of squared errors at these percentiles as well as the log sum of squared errors of the best-fitting model. Pooling the results obtained using the PWT 6.1 and PWT 6.2 income datasets (which results in 2*67 observations) yields

$$\ln \hat{p}^{X} = \text{DatasetDummies} \underset{(0.95)}{-26.87} \ln SSE_{\min}^{X} \underset{(1.14)}{-12.645} \ln SSE_{1Pct}^{X}$$

$$\underset{(1.61)}{+1.53} \ln SSE_{5Pct}^{X} \underset{(1.68)}{-0.67} \ln SSE_{25Pct}^{X} \underset{(1.07)}{-0.13} \ln SSE_{50Pct}^{X}$$

where $\ln \hat{p}^{X}$ is the predicted log posterior inclusion probability of candidate variable X (the regression R^2 is 99%); $SSE_{\min}^{X}$ denotes the minimum sum of squared errors across all 2145 models that include *X*; and SSE_{zPct}^{X} denotes the sum of squared errors at the *z*th percentile (numbers in parentheses are standard errors). Note that only the minimum sum of squared errors and the sum of squared errors at the 1st percentile are (highly) statistically significant. Hence, the sum of squared errors of the best fitting model has a statistically significant effect on posterior inclusion probabilities, even when the sum of squared errors at the 1st, 5th, 25th, and 50th percentiles are controlled for. Moreover, the effect is large. A 1% fall in the minimum sum of squared errors is associated with an increase in the posterior inclusion probabilities of almost 27%. Small changes in the sum of squared errors of a few models can therefore have large effects on posterior inclusion probabilities.

A small-scale Monte Carlo study. To examine the sensitivity of agnostic Bayesian procedures as applied in the literature, we perform a small-scale Monte Carlo study. We first generate 30 artificial datasets by randomly perturbating PWT 6.1 1960-1996 annualized growth rates. The distribution from which we draw the perturbations is calibrated using the difference between PWT 6.1 and 6.0 income growth rates (we use these two versions to be conservative, as changes between PWT 6.1 and PWT 6.2 growth rates are somewhat larger). The variance of the perturbations is calibrated to be decreasing in income per capita of countries. To make the exercise as close as possible to a minor income data revision by PWT standards, we draw from the calibrated distribution until we have generated 30 growth perturbations whose correlation with PWT 6.1 growth is between 0.975 and 0.979 (the interval in centered on 0.977, the correlation between PWT 6.1 and 6.0 growth rates). For comparison, the correlation between PWT 6.1 and PWT 6.2 growth rates is 0.937. The construction of the perturbed growth rates is explained in more detail in Appendix A.

A few statistics on our 30 randomly generated growth series corroborate that perturbations relative to PWT 6.1 growth are in fact conservative relative to the difference between growth rates in PWT 6.1 and PWT 6.2. The average mean squared difference between the 30 growth series and the PWT 6.1 data is 0.0041, the same as the mean squared difference of growth rates between PWT 6.0 and PWT 6.1, but smaller than the mean squared difference between PWT 6.1 and 6.2 (0.0059). The mean squared differences of individual artificial datasets with PWT 6.1 growth rates range from 0.0038 to 0.0044. Hence, the mean square difference of each of the 30 growth series with PWT 6.1 is smaller than the mean squared difference between PWT 6.2 and PWT 6.1. The average maximum absolute perturbation of an individual country across the 30 artificial datasets is 0.013 (1.3 percentage points of annualized growth). The maximum individual perturbation across all 30 artificial datasets is 0.018, which is smaller than the maximum absolute difference in growth rates of the PWT 6.0-6.1 revision or the PWT 6.1-6.2 revision.

We then apply SDM's BACE procedure to the 66 variables of the SDM dataset, plus the 1960 income level from PWT 6.1, with the dependent variable taken to be 1960-1996 income growth according to the PWT 6.1 dataset and each of the 30 growth perturbations. This gives 31 sets of BACE results. The variation between them is quite substantial. Comparing pairwise actual PWT 6.1 results with each of the perturbated datasets, we find that the average ratio of the greater to smaller posterior inclusion probability is 1.88 on average across all variables. SDM's robustness criterion is that the posterior inclusion probability of a variable exceeds the prior inclusion probability. Only four variables satisfy this criterion in all perturbed datasets, out of 35 variables which satisfy it in at least one of the dataset. As there are a total of 67 candidate explanatory variables, this implies that more than half of the candidates turn out to be robust at least once and 11% of the variables robust at least once are always robust. In other words, close to half of the candidate variables (31=35-4) emerge as growth determinants for some growth perturbation but not for another. When we repeat the same exercise applying the BMA with benchmark priors of FLS, we find similar results. For example, when comparing pairwise actual PWT 6.1 results with each of the perturbated datasets, the average ratio of the greater to smaller inclusion probability is 1.77.

Our findings above are consistent with what has been noted by other researchers. Results of Ley and Steel (2007) show that the single best model often dominates the BMA results for growth regressions. They also perform a Monte Carlo experiment where they generate artificial samples by randomly dropping 15% of observations, and find posterior inclusion probabilities of some variables fluctuating between zero and almost certainty. In other contexts, Pesaran and Zaffaroni (2006) and Garrat et al. (2007) also find that the Bayesian average of models often turns out to be dominated by the best model, and that model weights react very strongly to small changes in the data.[12]

B. General-to-Specific Approach

Hendry and Krolzig (2004) identify determinants of economic growth using a general-to-specific strategy, as implemented in their PcGets model-selection computer package (see Hendry and Krolzig, 2001; Campos, Ericsson and Hendry, 2005; Hendry and Krolzig, 2005). The algorithm tells relevant from irrelevant variables by performing a series of econometric tests. It tests significance of individual variables and their groups, as well as the correct specification of the resulting models. By following all possible reduction paths, the algorithm ensures that results do not depend on which insignificant variable is removed first. The output of the PcGets algorithm is the final, specific model, which includes only the variables that have a statistically significant effect on the dependent variable. The

[12] In Appendix B, we point to three possible ways of reducing the sensitivity of Bayesian model averaging to data imperfections: (i) shrinkage priors; (ii) the incorporation of priors about measurement error; (iii) Zellner's (2002) adjustment for data quality.

coefficients are estimated with ordinary least-squares, and standard t-ratios and R^2 are reported.

The general-to-specific strategy ends up picking a best model based on sequences of t-tests and F-tests. Because these tests are functions of the *SSE*, the best model will most likely be different across income data revisions as well as alternative income datasets. It is however difficult to assess theoretically whether such differences may be large. To get a sense of the sensitivity of the general-to-specific strategy, we therefore return to our small Monte-Carlo setup and apply the PcGets algorithm to the SDM dataset paired with the 30 growth perturbations. In this case, 28 candidate explanatory variables are selected at least once and none is selected always. To put it differently, approximately 40% of the candidate explanatory variables end up being part of the best model for some growth perturbation but not another.

3. Income Data and Growth Determinants

We start by analyzing the determinants of economic growth for the 1960-1996 period using the latest Penn World Table income data (PWT 6.2). To assess the sensitivity of agnostic all-inclusive empirical analysis to income data revisions, we then compare PWT 6.2 results with those of earlier PWT income data for 1960-1996 (PWT 6.0 and 6.1). As potential determinants we use the dataset of 67 variables, compiled by Sala-i-Martin, Doppelhofer, and Miller (2004) (see Web Appendix Tables A1a-b).

We also examine how much growth determinants depend on the different methodological choices underlying the PWT and the World Development Indicators income data. This analysis is for the 1975-1996 period, as the WDI purchasing power parity income estimates are only available since 1975. As potential determinants we use the same 67 variables, with values updated wherever necessary (see Web Appendix Table A1c).

A. Determinants of 1960-1996 Growth: The Effect of PWT Revisions

BACE and BMA with benchmark priors of FLS. In Table 1, we list all 67 candidate explanatory variables from the SDM dataset and their BACE posterior inclusion probabilities according to PWT 6.2, PWT 6.1 and PWT 6.0 income data. Following SDM, (*robust*) *growth determinants* are defined as variables with a posterior inclusion probability higher than the prior inclusion probability (which, with the prior model size of 7 and 67 candidate variables, equals 7/67=0.104). These posterior inclusion probabilities are shown in boldface. Unconditional mean effects of the main variables are listed in Table 2.

The original SDM BACE exercise was performed with PWT 6.0 income data on a sample of 88 countries. When we implement BACE with PWT 6.1 and PWT 6.2 data, we use the SDM data for all variables except 1960-1996 per capita growth rates and (log) 1960 incomes, which are taken from PWT 6.1 and PWT 6.2 respectively. The PWT 6.1 income data are available for 84 of the countries in the SDM sample and the PWT 6.2 data for 79 countries.[13] Ultimately, we want to use the history of PWT revisions to learn about how much the set of growth determinants using the latest income data might change with future revisions. We therefore compute BACE results on the largest possible samples for which all the necessary data are available, as it is impossible to know which income estimates will be dropped because of their unreliability in the future.[14]

Implementing BACE with PWT 6.2 1960-1996 income data yields 14 growth determinants. To get an idea of the effects of data revisions, note that there are 23 growth determinants for 1960-1996 according to PWT 6.2 or PWT 6.1. And the two versions of the PWT disagree on 13 of these variables. This disagreement is not driven by small changes in posterior inclusion probabilities (around the robustness threshold). For example, the investment price variable (which has played an important role in the growth literature, see, for example, De Long and Summers, 1991; Jones, 1994) is the variable with the third highest posterior inclusion probability (97%) according to PWT 6.1 income data, but practically irrelevant in PWT 6.2 (the posterior inclusion probability is 2%). The inclusion probability of the variable capturing location in the tropics (fraction tropical area) drops from above 70% to 5%. A similar drop is experienced by population density in 1960 and the population density of coastal areas in the 1960s. Air distance to big cities is another geographic country characteristic whose relevance for growth diminishes with the PWT 6.2 dataset. Life expectancy in 1960, the fraction of GDP produced in the mining sector, and political rights experience smaller, but still important decreases in their posterior inclusion probabilities (from around 25% to below 3%; for earlier results on the role of life expectancy and political rights for economic growth see Limongi and Przeworski, 1993;

[13] See Web Appendices A and B for details on the samples.

[14] BACE results are sensitive to PWT income data revisions even when one considers the largest common sample, see the next footnote.

Barro, 1991, 1996). Nominal government expenditures, on the other hand, obtain a posterior inclusion probability of 26% according to PWT 6.2 but are irrelevant according to PWT 6.1. Other variables with high posterior inclusion probabilities (above 83%) with PWT 6.2 data but low posterior inclusion probabilities (below 18%) according to PWT 6.1 are location in Africa, the fraction of the population that adheres to Confucianism, and fertility.

There is even greater disagreement regarding the determinants of 1960-1996 growth when we compare the results with PWT 6.2 income data to those using PWT 6.0. Examples of variables that go from irrelevance in PWT 6.0 to robustness in PWT 6.2 are fertility and primary export dependence (for earlier results see Sachs and Warner, 1995). Examples going the other way are the variables measuring the degree of ethnolinguistic fractionalization of the population, which was borderline with PWT 6.0 (for more on this variable, see Easterly and Levine, 1997; Alesina et al., 2003; Alesina and La Ferrara, 2005), and malaria prevalence. When we look across all 3 revisions of the PWT income data, we find that they disagree on 20 of 28 variables that are classified as growth determinants according to one of the versions.

Table 1 also reports statistics on differences in BACE posterior inclusion probabilities across PWT revisions. The statistic reported is the ratio of the larger to the smaller inclusion probability (MAX/MIN) between datasets. The two bottom rows of the table contain the average MAX/MIN value across all variables and across variables selected using one of the two datasets. Across variables that are robust according to one of the datasets compared, the average MAX/MIN value for the PWT 6.2-PWT 6.1 comparison is 7.97. The average MAX/MIN value across all 67 candidate variables is 4.33. Hence, on average, the larger posterior inclusion probability exceeds the smaller posterior inclusion probability by a factor greater than four when we compare PWT 6.2 and PWT 6.1. For the PWT 6.2-PWT 6.0 comparison, the average MAX/MIN value is 3.18 across variables that are robust according to one of the datasets and 2.26 across all variables.[15]

BMA with benchmark priors of FLS also yields results that are sensitive to PWT revisions. For example, when we compare PWT 6.1 and PWT 6.2 results, the average MAX/MIN posterior inclusion probability across all 67 candidate variables is 3.97, quite

[15] When we calculate the same statistics for the largest common samples of PWT versions 6.0, 6.1, and 6.2 (79 countries), we get an average MAX/MIN PIP value of 2.48 for the PWT 6.2-PWT 6.0 comparison and an average MAX/MIN PIP value of 1.82 for the PWT 6.2-PWT 6.1 comparison.

similar to what we obtained with BACE. The comparison between PWT 6.0 and PWT 6.2 results yields a MAX/MIN value across all 67 candidate variables of 3.3 (BACE yielded 5.52). Detailed results can be found in Web Appendix Table C1.[16]

General-to-specific approach. In Table 3, we summarize the results using the general-to-specific approach. Several of the variables of the final, specific empirical model using PWT 6.2 income data are also among the growth determinants emerging from the two Bayesian approaches. It can also be seen that the general-to-specific approach is rather sensitive to data revisions, as PWT 6.2 and PWT 6.1 comparison yields disagreement on 8 of 11 1960-1996 growth factors. The three versions of the PWT agree on only 1 growth determinant (primary schooling).

B. Determinants of 1975-1996 Growth: PWT versus WDI

The correlation between 1975-1996 growth rates for the 112 countries in both the PWT 6.2 and the World Development Indicators dataset is 96.2% (WDI purchasing power parity incomes estimates are only available from 1975). Limiting the analysis to the 87 countries in the SDM sample for which there are WDI income data, the correlation between growth rates is 95.4% and the correlation between 1975 income per capita levels is 96.2%. Still, as reported in Table 4, the Bayesian criterion of SDM yields disagreement on 8 of 15 variables that are classified as growth determinants using one of the two datasets.[17] For example, political freedom emerges as a positive growth determinant (unconditional mean effects of the main variables are reported in Table 5) with WDI but not PWT 6.2 income data. The same is true for three different indicators of openness (trade openness, years open, and exchange rate distortions). Averaging the MAX/MIN posterior inclusion probability ratios across variables that are robust using one of the datasets yields 2.34.

[16] An important difference between BACE as implemented by SDM and BMA of FLS is the prior model size. FLS set it to the number of explanatory variables over 2 (33.5 as we are using a set of 67 potential explanatory variables), while SDM chose 7. The larger prior model size of FLS turns out to reduce the sensitivity of results to data revisions. When we implement BMA of FLS with a prior model size of 7, the average MAX/MIN value across all 67 candidate variables is 10.84 comparing PWT 6.1 and PWT 6.2; comparing PWT 6.0 and PWT 6.2, the average MAX/MIN inclusion probability is 7.93. See Web Appendix Table C3.

[17] To examine the determinants of 1975-1996 growth rates we updated the SDM dataset, replacing the variables which were measured in the 1960s with the analogous values for the 1970s. See the details in Web Appendix Table A1c.

Using the PWT 6.1 income data, the discrepancy with the WDI is even more striking. Averaging the MAX/MIN posterior inclusion probability ratios across variables that are robust using one of the datasets yields 4.81. Disagreement extends to 12 of 20 variables that are classified as growth determinants using one of the two datasets. Results are similar when we use FLS-BMA. In this case, the MAX/MIN posterior inclusion probability ratios average to 1.62 when comparing WDI and PWT 6.2, and 2.55 when comparing WDI and PWT 6.1, see Web Appendix Table C5.[18]

Application of the general-to-specific approach to the PWT 6.2 and WDI international income data yields disagreement on 7 of 10 explanatory variables that are in the final, specific model using one of the two datasets. For example, the East Asian country dummy enters the final model according to the WDI income data (positively) but not the PWT 6.2 data. On the other hand, the PWT 6.2 income data indicates that religion may matter for growth, while the WDI final model does not contain a single indicator of religion.

4. Implications of More Informative Priors

One reason why most of the existing empirical work focuses on cross-country regressions with few explanatory variables is that such models can be expected to be more robust to data imperfections. With fewer correlated variables, the conditioning of the ordinary least-squares problem can be expected to improve (on data matrix conditioning and the sensitivity of OLS estimator to small changes in data, see Belsley, Kuh, and Welsch, 1980, for example). This point can be illustrated using the PWT 6.2 revision of the PWT 6.0 growth data for 1960-1996. When we run a least-squares regression of 1960-1996 economic growth using PWT 6.2 data on all 67 explanatory variables of SDM, we find 14 variables with an absolute t-statistic greater than 2 (which corresponds to statistical significance at the 95% level approximately).[19] When we repeat the analysis with 1960-1996 growth data from PWT 6.0, only 1 variable is significant and it is not among the significant variables using PWT 6.2. Hence, researchers using PWT 6.0 income data would have had to totally revise their conclusions about growth determinants with the arrival of PWT 6.2. Now consider cutting the number of explanatory variables to 18. To avoid

[18] As before, FLS-BMA with prior model size 7 yields somewhat larger disagreement, see Web Appendix Table C6.

[19] This threshold t-value is arbitrary of course. Other conventional values yield similar results.

findings being driven by a particular list of explanatory variables, rather than there being relatively few of them, we randomly extract 500 lists of 18 variables and compare average statistics across lists.[20] In this case, we find that 54% of the variables that are statistically significant at the 95% confidence level using one of the datasets are significant and enter with the same sign in both. If we reduce the number of explanatory variables to 10, the overlap increases to almost 2/3 (for the PWT 6.2-WDI comparison, overlap is 82% in this case). Clearly, stronger priors about potential growth determinants lead to results that are less fragile with respect to PWT income data revisions.

Will stronger priors about candidate explanatory variables also reduce the sensitivity of Bayesian empirical analysis to income differences across datasets? The two top panels of Table 6A summarize the results starting with 500 randomly selected lists of 18 and 10 candidate variables respectively. For comparison, we also give results for the full list of 67 candidate variables. When starting with 67 candidates, SDM's Bayesian averaging of classical estimates yields that PWT 6.0 and PWT 6.2 disagree on 59% of the 1960-1996 growth determinants (top panel). The average MAX/MIN ratio of posterior inclusion probabilities across variables (3.18) also indicates substantial disaccord. When we start with lists of 18 variables, disagreement between the two versions of the PWT falls. Now they disagree on 42% of variables on average across the 500 lists, and the average MAX/MIN ratio of posterior inclusion probabilities is 1.84 when we average across lists. Disagreement falls to 35% and the average MAX/MIN ratio to 1.37 when we start with lists of 10 variables. The average MAX/MIN ratio of posterior inclusion probabilities obtained with FLS's Bayesian model averaging also falls as candidate lists become shorter (middle panel). Moreover, the results in Table 6A show that the general-to-specific model selection approach also becomes less sensitive to PWT income revisions when fewer variables are considered a priori (bottom panel). Findings are similar for the PWT 6.2-WDI comparison in Table 6B. Hence, just as in the case of ordinary least-squares analysis, reducing the number of potential growth determinants yields results that are less sensitive to income differences across datasets.

[20] Extracting more than 500 lists did not change our findings.

5. Conclusions

It is easy to see, and should not be surprising, that the available international income data are imperfect. One only needs to examine how 1960-1996 growth rates have been changing with each revision of the Penn World Table. But what does this imply for empirical work on the determinants of economic growth? We find that each revision of the 1960-1996 income data in the PWT leads to substantive changes regarding growth determinants with agnostic empirical analysis. A case in point is the latest revision (PWT 6.2). Using Sala-i-Martin, Doppelhofer, and Miller's (2004) Bayesian averaging of classical estimates approach, PWT 6.2 and the previous version (PWT 6.1) disagree on 13 of 23 growth determinants for the 1960-1996 period that emerge with one of the two datasets. Other agnostic approaches we consider yield similar results. The explanatory variables to which an agnostic should pay attention according to some versions of the 1960-1996 income data, but not others, are related to the debates on trade openness, religion, geography, demography, health, etc. A Monte Carlo study confirms that agnostic empirical approaches are sensitive to small income data revisions by PWT standards.

Agnostic empirical analysis also results in only limited coincidence regarding growth determinants when we use international income estimates obtained with alternative methodologies. For instance, the latest PWT and World Bank international income data yield disagreement on 8 of 15 growth determinants for the 1975-1996 period with the Bayesian averaging of classical estimates approach.

Our findings suggest that the available income data may be too imperfect for agnostic empirical analysis. At the same time, we find that the sensitivity of growth determinants to income differences across data revisions and datasets falls considerably when priors about potential growth determinants become stronger. That is, the data appears good enough to differentiate among a limited number of hypotheses. Empirical models of the typical size in the literature, for example, tend to point to the same growth determinants using different versions of the PWT or the World Bank income data. Researchers who want to continue giving equal a priori weight to all potential growth determinants in the literature should consider shrinkage priors, explicitly incorporating priors about measurement error in the income data, or implementing Zellner's (2002) adjustment for data quality.

References

Alesina, A., Devleeschauwer, A., Easterly, W., Kurlat, S., and Wacziarg, R. (2003). Fractionalization. Journal of Economic Growth, 8(2):155–94.

Alesina, A. and La Ferrara, E. (2005). Ethnic diversity and economic performance. Journal of Economic Literature, 43:721–61.

Barro, R. J. (1991). Economic growth in a cross section of countries. Quarterly Journal of Economics, 106(2):407–43.

Barro, R. J. (1996). Democracy and growth. Journal of Economic Growth, 1(1):1–27.

Barro, R. J. (1998). Determinants of Economic Growth: A Cross-country Empirical Study. Lionel Robbins Lectures. MIT Press, first edition.

Barro, R. J. and Lee, J.-W. (1994). Sources of economic growth. Carnegie-Rochester Conference Series on Public Policy, 40:1–46.

Barro, R. J. and Lee, J.-W. (2000). International data on educational attainment: Updates and implications. Technical report, Harvard University.

Barro, R. J. and Sala-i-Martin, X. (2003). Economic Growth. MIT Press, second edition.

Belsley, D. A., Kuh, E., and Welsch, R. E. (1980). Regression Diagnostics: Identifying Influential Data and Sources of Collinearity. John Wiley and Sons.

Campos, J., Ericsson, N. R., and Hendry, D. (2005). Editors' introduction. In Campos, J., Ericsson, N. R., and Hendry, D., editors, General to Specific Modelling. Edward Elgar Publishing, Cheltenham.

Congressional Budget Office (2003). The economics of climate change: A primer. http://www.cbo.gov.

De Long, J. B. and Summers, L. H. (1991). Equipment investment and economic growth. Quarterly Journal of Economics, 106(2):445–502.

Doan, T., Litterman, R., and Sims, C. (1984). Forecasting and conditional projections using realistic prior distributions. Econometric Reviews, 3(1):1–100.

Dowrick, S. and Quiggin, J. (1997). Convergence in GDP and living standards: A revealed preference approach. American Economic Review, 67(1):41–64.

Dowrick, S. (2005). The Penn World Table: A review. Australian Economic Review, 38(2):223–228.

Durlauf, S. N., Johnson, P. A., and Temple, J. R. W. (2005). Growth econometrics. In Aghion, P. and Durlauf, S. N., editors, Handbook of Economic Growth. North-Holland.

Easterly, W. and Levine, R. E. (1997). Africa's growth tragedy: Policies and ethnic divisions. Quarterly Journal of Economics, 112(4):1203–50.

Fernandez, C., Ley, E., and Steel, M. F. (2001a). Benchmark priors for Bayesian model averaging. Journal of Econometrics, 100:381–427.

Fernandez, C., Ley, E., and Steel, M. F. (2001b). Model uncertainty in cross- country growth regressions. Journal of Applied Econometrics, 16(5):563–76.

Garratt, A., Lee, K., Mise, E., and Shields, K. (2007). Real time representation of the UK output gap in the presence of model uncertainty. Technical report, Birkbeck College.

Grier, K. B. and Tullock, G. (1989). An empirical analysis of cross-national economic growth. Journal of Monetary Economics, 24(2):259–76.

Hendry, D. F. and Krolzig, H.-M. (2001). Automatic Econometric Model Selection using PcGets. Timberlake Consultants Press, London.

Hendry, D. F. and Krolzig, H.-M. (2004). We ran one regression. Oxford Bulletin of Economics and Statistics, 66:799–810.

Hendry, D. F. and Krolzig, H.-M. (2005). The properties of automatic Gets modelling. Economic Journal, 115:C32–C61.

Heston, A. (1994). National accounts: A brief review of some problems in using national accounts data in level of output comparisons and growth studies. Journal of Development Economics, 44:29–52.

Heston, A., Summers, R., and Aten, B. (2001,2002,2006). Penn World Table version 6.0,6.1,6.2. Data Web site: http://pwt.econ.upenn.edu/. Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania.

Hoerl, A. E. and Kennard, R. W. (1970). Ridge regression: biased estimation for nonorthogonal problems. Technometrics, 12(1):55–67.

Hoeting, J. A., Madigan, D., Raftery, A. E., and Volinsky, C. T. (1999). Bayesian model averaging: A tutorial. Statistical Science, 14(4):382–417.

Jones, C. I. (1994). Economic growth and the relative price of capital. Journal of Monetary Economics, 34(3):359–382.

Kleibergen, F. and Paap, R. (2002). Priors, posteriors and Bayes' factors for a Bayesian analysis of cointegration. Journal of Econometrics, 111(2):223–249.

Kleibergen, F. and van Dijk, H. K. (1998). Bayesian simultaneous equations analysis using reduced rank structures. Econometric Theory, 14:701–743.

Kormendi, R. C. and Meguire, P. (1985). Macroeconomic determinants of growth: Cross-country evidence. Journal of Monetary Economics, 16(2):141–63.

Kravis, I. B., Heston, A., and Summers, R. (1982). World Product and Income: International Comparisons of Real Gross Products. Baltimore: The Johns Hopkins University Press.

Leamer, E. E. (1978). Specification Searches. New York: John Wiley and Sons.

Levine, R. E. and Renelt, D. (1992). A sensitivity analysis of cross-country growth regressions. American Economic Review, 82(4):942–63.

Ley, E. and Steel, M. F. (2007). On the effect of prior assumptions in Bayesian model averaging with applications to growth regression. MPRA Paper 3214, University Library of Munich, Germany.

Limongi, F. and Przeworski, A. (1993). Political regimes and economic growth. Journal Of Economic Perspectives, 7(3).

Neary, J. P. (2004). Rationalizing the Penn World Table: True multilateral indices for international comparisons of real income. American Economic Review, 94(5):1411–1428.

Pesaran, M. H. and Zaffaroni, P. (2006). Model averaging and value-at-risk based evaluation of large multi asset volatility models for risk management. Technical report.

Rodriguez, F. and Rodrik, D. (2001). Trade policy and economic growth: A sceptic's guide to the cross-national evidence. In Bernanke, B. S. and Rogoff, K., editors, NBER Macroeconomics Annual 2000. Cambridge, MA: MIT Press.

Sachs, J. D. (2005). Can extreme poverty be eliminated? Scientific American.

Sachs, J. D. and Warner, A. M. (1995). Economic reform and the process of economic integration. Brookings Papers on Economic Activity, (1):1–95.

Sala-i-Martin, X. (1997). I just ran 2 million regressions. The American Economic Review, 87(2):178–83. Papers and Proceedings.

Sala-i-Martin, X., Doppelhofer, G., and Miller, R. I. (2004). Determinants of long-term growth: A Bayesian averaging of classical estimates (BACE) approach. The American Economic Review, 94(4):813–835.

Summers, R. and Heston, A. (1991). The Penn World Table (Mark 5): An expanded set of international comparisons, 1950-1988. Quarterly Journal of Economics, pages 327–368.

Tsangarides, C. G. (2005). Growth empirics under model uncertainty: Is Africa different? IMF Working Paper 05/18, International Monetary Fund.

Zellner, A. (2002). Information processing and Bayesian analysis. Journal of Econometrics, 107:41–50.

Appendix

Appendix A: Design of Monte Carlo Study

We generate 30 perturbated 1960-96 growth series starting from PWT 6.1 GDP per capita growth. The perturbations are drawn from distributions that are calibrated to the difference between the PWT 6.0 and PWT 6.1 income data (which is smaller than the difference between PWT 6.1 and PWT 6.2 income data). The variance of these perturbations is taken to be decreasing in income per capita of a country. This reflects the observed heteroscedasticity of the measurement error; the income of richer countries is more exactly measured than that of poorer countries. In particular, we take the variance of the perturbations to be the fitted value from a regression of the squared differences between PWT 6.1 and 6.0 growth rates on a constant and PWT 6.1 log income per capita in 1960 (see Table A1 below for the results). Fitted values of the 17 richest (in 1960) countries are negative, so we replace them by 0, i.e. we do not perturb their growth rates. We draw from this distribution until we have generated 30 growth perturbations whose correlation with PWT 6.1 growth is between 0.975 and 0.979 (the interval in centered on 0.977, the correlation between PWT 6.1 and 6.0 growth rates). Summary statistics about perturbed data are reported in Table A2 below.

Appendix Table A1. Ordinary least-squares regression of squares of revisions of income data (growth and levels) on the level of income in 1960

Constant	0.000160
	(0.000048)
$\log y_{1960}^{PWT\,6.1}$	-0.000019
	(0.000006)
R^2	0.10
Number of observations	84

Notes: Standard errors are in parenthesis.

Appendix Table A2. Perturbed growth rate series compared to PWT 6.1 1960-1996 growth rates

	Correlation with PWT 6.1 1960-1996 growth rates	R^2 of regression on constant and PWT 6.1 1960-1996 growth rates
Min	0.975	0.950
Average	0.977	0.954
Max	0.979	0.959

Appendix B: Could Bayesian Model Averaging be Made More Robust?

Our results in Section 2 suggest that the effect of small changes in the sum of squared errors on Bayesian posterior odds ratios is implausibly strong when there are doubts regarding data quality. We now take the Bayesian model averaging specification of Fernandez, Ley and Steel (2001a) as the baseline and discuss some departures that make odds ratios less extreme functions of sum of squared errors.

Priors on model coefficients. The agnostic Bayesian approaches we have discussed assume very loose priors for model coefficients. In the benchmark prior of FLS for example, the prior variance is proportional to g^{-1} and g is taken to be very small. In the expression in (4), the term in the second bracket is a weighted average of the sum of squared errors of a least-squares regression using the explanatory variables in model j (SSE_j) and the sum of squared errors when setting all model coefficients equal to their prior means (as prior means are zero, this term is equal to $(y-\bar{y})'(y-\bar{y})$). In this framework, making a prior for coefficients more informative by reducing its prior variance results in a larger value of the parameter g in (4). This implies that a greater weight is put on $(y-\bar{y})'(y-\bar{y})$, which does not vary across models. (The Bayesian updating completely ignores least-squares coefficients in the limit where prior variances are zero.) A given change in SSE_j therefore leads to a smaller change in the posterior probability. Such more informative priors, called shrinkage priors, are an established way to improve reliability of estimation with short samples (see, for example, Doan, Litterman, and Sims, 1984; Hoerl and Kennard, 1970).

Explicit modeling of the measurement error in the growth data. Another possibility is to specify an informative prior on the variance of e_{jc} in (1), which can be thought of as capturing measurement error in the growth data. Assuming a known variance σ^2 implies that the odds ratio for models *j* and *k* (assuming they are equally likely a priori) is

$$\frac{l\left(M_j\right)}{l\left(M_s\right)} = \left(\frac{g}{1+g}\right)^{\frac{k_j - k_s}{2}} \exp\left(-\frac{1}{2}\frac{\frac{1}{1+g}\left(SSE_j - SSE_s\right)}{\sigma^2}\right).$$

(In their baseline setup, FLS assume a non-informative prior for σ^2.) Hence, the greater the variance, the smaller the percentage change in odds ratios in response to a change in the difference in fit between models ($SSE_j - SSE_s$). A similar effect can be obtained when the variance is not known but an appropriate informative prior distribution is specified for it.

Using *quality adjusted* likelihood. Zellner (2002) proposes to account for low quality sample information by using a *quality adjusted* likelihood, which is obtained as the original likelihood raised to a power *a*, $0 < a < 1$ (for further references on this approach see Zellner). The usual case of a fully reliable sample corresponds to a=1. Lower values of a make the density more spread out, which captures that low quality data carries less information. Introducing Zellner's quality adjustment in the FLS setup, we obtain a marginal likelihood of the form,

$$l_y(M_j) \propto \left(\frac{g}{g+a}\right)^{\frac{k_j}{2}} \left(\frac{a}{g+a} SSE_j + \frac{g}{g+a}(y-\bar{y})'(y-\bar{y})\right)^{-\frac{Ca-1}{2}}.$$

Accounting for low-quality data makes odds ratios react less to changes in the sum of squared errors because it results in the second bracket being taken to a lower power in absolute terms and reduces the weight on model fit.

Tables and Figures

Table 1. Determinants of 1960-1996 income growth with the BACE approach: Posterior inclusion probabilities using income data from Penn World Table versions 6.2, 6.1, and 6.0

	PWT6.2	PWT6.1	PWT6.0	PWT6.2/6.1	PWT6.2/6.0
	Posterior inclusion probabilities			MAX/MIN	MAX/MIN
GDP in 1960 (log)	**1.00**	**1.00**	**0.69**	1.00	1.46
Primary Schooling in 1960	**1.00**	**0.99**	**0.79**	1.01	1.25
African Dummy	**0.86**	**0.18**	**0.16**	4.74	1.17
Fraction Confucius	**0.83**	**0.13**	**0.21**	6.25	1.56
Fraction Muslim	**0.40**	**0.18**	**0.11**	2.16	1.60
East Asian Dummy	**0.33**	**0.78**	**0.82**	2.32	1.06
Fraction Buddhist	**0.28**	**0.11**	**0.12**	2.57	1.06
Population Density Coastal in 1960s	**0.11**	**0.79**	**0.43**	7.46	1.84
Fertility in 1960s	**0.91**	**0.12**	0.03	7.75	3.84
Latin American Dummy	**0.35**	0.07	**0.15**	4.92	2.08
Primary Exports 1970	**0.27**	**0.20**	0.05	1.35	3.86
Fraction of Tropical Area	0.05	**0.71**	**0.56**	13.59	1.26
Life Expectancy in 1960	0.03	**0.25**	**0.22**	7.92	1.16
Investment Price	0.02	**0.97**	**0.78**	47.55	1.25
Fraction GDP in Mining	0.02	**0.24**	**0.13**	14.23	1.88
Nominal Government GDP Share 1960s	**0.26**	0.02	0.04	15.59	2.15
Openness Measure 1965-74	**0.15**	0.06	0.07	2.38	1.17
Timing of Independence	**0.12**	0.06	0.02	1.79	3.42
Hydrocarbon Deposits in 1993	0.10	**0.11**	0.03	1.11	4.07
Years Open 1950-94	0.08	0.05	**0.11**	1.58	2.06
Spanish Colony	0.07	0.02	**0.13**	3.30	6.33
Air Distance to Big Cities	0.04	**0.45**	0.04	10.33	11.89
Ethnolinguistic Fractionalization	0.04	0.03	**0.11**	1.24	3.68
Fraction Population in Tropics	0.03	**0.15**	0.06	4.39	2.56
Gov. Consumption Share 1960s	0.03	0.05	**0.11**	1.80	2.28
Malaria Prevalence in 1960s	0.02	0.02	**0.25**	1.02	10.67
Political Rights	0.02	**0.26**	0.06	12.78	3.97
Population Density 1960	0.02	**0.73**	0.09	40.90	8.42
Fraction Protestants	0.07	0.02	0.05	3.87	2.60
Fraction Speaking Foreign Language	0.06	0.04	0.08	1.60	1.96
Fraction Catholic	0.06	0.02	0.03	3.12	1.62
European Dummy	0.06	0.03	0.03	1.74	1.18
Average Inflation 1960-90	0.05	0.02	0.02	2.45	1.08
Government Share of GDP in 1960s	0.05	0.04	0.06	1.12	1.52
Fraction Population Over 65	0.05	0.05	0.02	1.08	2.18
Square of Inflation 1960-90	0.04	0.02	0.02	2.54	1.14
Size of Economy	0.03	0.02	0.02	1.46	1.14
Tropical Climate Zone	0.03	0.03	0.02	1.13	1.87
Defence Spending Share	0.03	0.02	0.02	1.65	1.03
Fraction Population Less than 15	0.03	0.03	0.04	1.27	1.63
Landlocked Country Dummy	0.03	0.08	0.02	2.66	4.32
Revolutions and Coups	0.03	0.03	0.03	1.14	1.09
Population Growth Rate 1960-90	0.03	0.03	0.02	1.00	1.61
Higher Education 1960	0.03	0.02	0.06	1.85	4.11
Absolute Latitude	0.03	0.03	0.03	1.08	1.06
Fraction Orthodox	0.03	0.01	0.01	1.91	1.07

Fraction Hindus	0.03	0.02	0.04	1.31	2.22
Interior Density	0.03	0.02	0.01	1.51	1.25
War Participation 1960-90	0.02	0.01	0.01	1.61	1.02
Socialist Dummy	0.02	0.03	0.02	1.24	1.61
Colony Dummy	0.02	0.08	0.03	3.79	2.61
Public Investment Share	0.02	0.04	0.05	2.24	1.16
Oil Producing Country Dummy	0.02	0.02	0.02	1.22	1.17
Capitalism	0.02	0.01	0.02	1.31	1.27
Land Area	0.02	0.02	0.02	1.15	1.08
Real Exchange Rate Distortions	0.02	0.04	0.08	2.34	1.97
British Colony Dummy	0.02	0.03	0.03	1.40	1.08
Population in 1960	0.02	0.02	0.02	1.15	1.57
Public Education Spending Share in GDP in 1960s	0.02	0.02	0.02	1.08	1.32
Fraction of Land Area Near Navigable Water	0.02	0.05	0.02	2.81	2.52
Religion Measure	0.02	0.03	0.02	1.94	1.68
Fraction Spent in War 1960-90	0.02	0.01	0.01	1.29	1.14
Civil Liberties	0.02	0.02	0.03	1.30	1.39
Terms of Trade Growth in 1960s	0.02	0.02	0.02	1.29	1.06
English Speaking Population	0.02	0.02	0.02	1.02	1.24
Terms of Trade Ranking	0.02	0.02	0.02	1.08	1.06
Outward Orientation	0.01	0.03	0.03	2.34	1.12
Average for variables robust at least once				7.97	3.18
Average for all variables				4.33	2.26

Notes: Variables come from the Sala-i-Martin, Doppelhofer, and Miller (2004) dataset (see Tables A1a-b in the Web Appendix). Posterior inclusion probabilities higher than the prior inclusion probabilities (here: 7/67) are in boldface.

Table 2. Determinants of 1960-1996 income growth with the BACE approach: Standardized unconditional posterior mean effects on 1960-1996 growth using income data from Penn World Table versions 6.2, 6.1, and 6.0

	PWT6.2	PWT6.1	PWT6.0
	Unconditional mean effects		
GDP in 1960 (log)	-1.28	-1.38	-0.53
Primary Schooling in 1960	0.91	1.01	0.62
African Dummy	-0.56	-0.10	-0.10
Fraction Confucius	0.35	0.05	0.09
Fraction Muslim	0.16	0.08	0.04
East Asian Dummy	0.13	0.38	0.57
Fraction Buddhist	0.08	0.03	0.04
Population Density Coastal in 1960s	0.02	0.39	0.19
Fertility in 1960s	-0.68	-0.07	-0.01
Latin American Dummy	-0.14	-0.03	-0.08
Primary Exports 1970	-0.10	-0.10	-0.02
Fraction of Tropical Area	-0.01	-0.50	-0.39
Life Expectancy in 1960	0.01	0.19	0.21
Investment Price	0.00	-0.45	-0.35
Fraction GDP in Mining	0.00	0.09	0.04
Nominal Government GDP Share 1960s	-0.07	0.00	-0.01
Openness Measure 1965-74	0.03	0.02	0.02
Timing of Independence	-0.04	-0.02	0.00
Hydrocarbon Deposits in 1993	0.02	0.03	0.00
Years Open 1950-94	0.02	0.02	0.04
Spanish Colony	-0.02	0.00	-0.05
Air Distance to Big Cities	-0.01	-0.18	-0.01
Ethnolinguistic Fractionalization	0.01	0.00	-0.04
Fraction Population in Tropics	-0.01	-0.08	-0.02
Gov. Consumption Share 1960s	0.00	-0.01	-0.03
Malaria Prevalence in 1960s	0.00	0.00	-0.17
Political Rights	0.00	-0.11	-0.02
Population Density 1960	0.00	0.31	0.02
Fraction Protestants	-0.02	0.00	-0.02
Fraction Speaking Foreign Language	0.01	0.01	0.02
Fraction Catholic	-0.02	0.00	-0.01
European Dummy	0.01	0.01	0.00
Average Inflation 1960-90	-0.01	0.00	0.00
Government Share of GDP in 1960s	-0.01	-0.01	-0.02
Fraction Population Over 65	0.01	0.02	0.00

Notes: Coefficients of standardized variables are obtained as the coefficients of the original variables multiplied by their in-sample standard deviation times 100. Therefore, they show the effect of a one standard deviation change in the variable in terms of percentage points of average annual growth rate over 1960-1996. For example, according to the first entry, countries with a one standard deviation higher initial income had, on average and controlling for other variables, 1.282% lower annual growth rates over 1960-1996. For brevity, only the coefficients of the first 35 variables in Table 1 are reported.

Table 3. Determinants of 1960-1996 income growth with the PcGets approach: Final models selected by PcGets using income data from Penn World Table versions 6.2, 6.1, and 6.0

RESULTS WITH PWT 6.0 INCOME DATA

	Coefficient	t-Statistics
Fraction Buddhist	0.505	3.77
Fraction Confucius	0.615	4.63
Investment Price	-0.536	-4.03
Primary Schooling in 1960	0.745	5.51

RESULTS WITH PWT 6.1 INCOME DATA

	Coefficient	t-Statistic
Fraction Buddhist	0.439	3.31
Investment Price	-0.600	-4.50
Primary Schooling in 1960	1.013	5.58
Primary Exports 1970	-0.722	-4.87
GDP in 1960 (log)	-0.893	-4.67

RESULTS WITH PWT 6.2 INCOME DATA

	Coefficient	t-Statistic
Fraction Confucius	0.579	5.98
Fertility in 1960s	-0.805	-5.00
Defence Spending Share	-0.276	-2.71
Hydrocarbon Deposits in 1993	0.293	3.15
Fraction Muslim	0.679	6.18
Timing of Independence	-0.430	-3.94
Primary Schooling in 1960	1.383	10.29
Primary Exports 1970	-0.455	-3.41
GDP in 1960 (log)	-1.726	-10.07

Notes: These are s*pecific models* obtained using *conservative strategy* with PcGets version 1.02. The settings used allow to replicate the results of Hendry and Krolzig (2004). Coefficients are standardized as in Table 2.

Table 4. Determinants of 1975-1996 income growth with the BACE approach. Posterior inclusion probabilities (PIP) with PWT and WDI income data

	PWT6.2 PIP	WDI common sample	PWT6.1 PIP	WDI common sample	PWT6.2/WDI MAX/MIN	PWT6.1/WDI MAX/MIN
East Asian Dummy	**0.98**	**0.99**	**0.52**	**0.99**	1.01	1.91
GDP in 1975 (log)	**0.88**	**1.00**	**0.50**	**1.00**	1.13	2.01
Life Expectancy in 1975	**0.86**	**0.97**	**0.41**	**0.99**	1.12	2.41
Fraction of Tropical Area	**0.73**	**0.68**	**0.20**	**0.72**	1.07	3.61
Fraction GDP in Mining	**0.27**	**0.15**	**0.48**	**0.21**	1.82	2.29
Absolute Latitude	**0.21**	**0.31**	**0.21**	**0.27**	1.48	1.26
Investment Price	**0.97**	**0.21**	**0.27**	0.05	4.66	5.39
Real Exchange Rate Distortions	0.08	**0.32**	**0.33**	**0.35**	4.12	1.04
Fraction Confucius	**0.12**	0.06	**0.41**	0.05	2.07	8.42
Political Rights	0.09	**0.24**	0.02	**0.19**	2.55	7.75
Openness Measure 1965-74	0.09	**0.17**	0.10	**0.25**	1.82	2.59
Years Open 1950-94	0.07	**0.30**	0.08	**0.24**	4.44	2.98
Government Share of GDP in 1970s	0.03	0.02	**0.29**	**0.17**	1.42	1.65
Gov. Consumption Share 1970s	0.02	**0.15**	0.09	**0.19**	6.79	2.21
British Colony Dummy	0.08	**0.11**	0.10	0.07	1.34	1.41
Fraction Population in Tropics	0.06	0.03	**0.20**	0.04	1.76	5.47
African Dummy	0.06	0.02	**0.54**	0.03	2.53	17.58
Fraction Buddhist	0.05	0.02	**0.29**	0.02	2.30	14.85
Fraction Speaking Foreign Language	0.04	**0.11**	0.02	0.09	2.62	4.16
Latin American Dummy	0.03	0.02	**0.15**	0.02	1.41	6.39
Nominal Government GDP Share 1970s	0.02	0.02	**0.30**	0.05	1.28	5.91
Defence Spending Share	0.02	0.04	**0.23**	0.05	2.24	4.62
Primary Schooling in 1975	0.08	0.03	0.02	0.03	2.88	1.51
Public Investment Share	0.07	0.02	0.05	0.02	2.98	2.07
Population Density Coastal in 1960s	0.07	0.08	0.06	0.07	1.16	1.09
Population Density 1975	0.07	0.07	0.06	0.06	1.12	1.02
Malaria Prevalence in 1960s	0.05	0.09	0.02	0.10	1.93	4.74
Terms of Trade Ranking	0.05	0.06	0.09	0.05	1.25	1.76
Ethnolinguistic Fractionalization	0.04	0.05	0.10	0.04	1.16	2.36
Revolutions and Coups	0.04	0.02	0.03	0.03	1.60	1.13
Higher Education 1975	0.04	0.03	0.03	0.03	1.04	1.18
Population in 1975	0.03	0.04	0.04	0.05	1.15	1.07
Fraction Muslim	0.03	0.05	0.09	0.04	1.44	2.31
Fraction Hindus	0.03	0.03	0.05	0.03	1.15	1.62
Fraction Orthodox	0.03	0.02	0.03	0.02	1.40	1.43
Capitalism	0.03	0.02	0.02	0.02	1.14	1.35
Civil Liberties	0.02	0.03	0.02	0.03	1.07	1.34
Land Area	0.02	0.02	0.02	0.02	1.02	1.33
Religion Measure	0.02	0.02	0.02	0.02	1.19	1.29
Average Inflation 1960-90	0.02	0.01	0.03	0.02	1.41	2.13
Fraction Population Less than 15	0.02	0.02	0.08	0.02	1.19	4.61
Fraction Spent in War 1960-90	0.02	0.01	0.02	0.01	1.47	1.35
Timing of Independence	0.02	0.02	0.02	0.03	1.10	1.17
Fraction Protestants	0.02	0.03	0.05	0.03	1.57	1.91
Square of Inflation 1960-90	0.02	0.01	0.02	0.01	1.27	1.75

Size of Economy	0.02	0.02	0.03	0.02	1.05	1.46
European Dummy	0.02	0.02	0.04	0.02	1.16	2.03
Interior Density	0.02	0.01	0.02	0.01	1.23	1.08
Fraction Population Over 65	0.02	0.02	0.05	0.02	1.18	2.36
Spanish Colony	0.02	0.02	0.10	0.02	1.03	5.39
Primary Exports 1970	0.02	0.02	0.02	0.02	1.07	1.51
Fraction of Land Area Near Navigable Water	0.02	0.02	0.02	0.02	1.01	1.04
Fertility in 1960s	0.02	0.02	0.03	0.02	1.01	2.12
Fraction Catholic	0.02	0.02	0.04	0.02	1.04	2.60
Colony Dummy	0.02	0.02	0.02	0.02	1.10	1.27
Air Distance to Big Cities	0.02	0.02	0.03	0.02	1.10	1.47
Hydrocarbon Deposits in 1993	0.02	0.01	0.02	0.01	1.10	1.57
Population Growth Rate 1960-90	0.02	0.02	0.03	0.02	1.09	1.60
Oil Producing Country Dummy	0.02	0.02	0.02	0.02	1.04	1.16
Terms of Trade Growth in 1960s	0.02	0.01	0.02	0.02	1.05	1.14
Outward Orientation	0.01	0.03	0.02	0.03	2.11	1.65
Public Education Spending Share in GDP in 1970s	0.01	0.02	0.02	0.02	1.29	1.24
Tropical Climate Zone	0.01	0.03	0.02	0.04	2.30	2.06
Landlocked Country Dummy	0.01	0.01	0.02	0.02	1.00	1.25
English Speaking Population	0.01	0.02	0.02	0.02	1.12	1.43
War Participation 1960-90	0.01	0.01	0.02	0.01	1.05	1.16
Socialist Dummy	0.01	0.02	0.03	0.02	1.15	1.95
Average for variables robust at least once					2.32	4.81
Average for all variables					1.64	2.78

Notes: The variables are based on Sala-i-Martin, Doppelhofer, and Miller (2004), but wherever applicable, variables were updated from the 1960s to the 1970s. See the Web Appendix for the list of the updated variables and data sources.

Table 5. Determinants of 1975-1996 income growth with the BACE approach. Standardized unconditional posterior mean effects on 1960-1996 growth with PWT and WDI data

	PWT6.2	WDI (PWT6.2 sample)	PWT6.1	WDI (PWT6.1 sample)
East Asian Dummy	0.86	0.80	0.38	0.81
GDP in 1975 (log)	-1.11	-1.57	-0.66	-1.63
Life Expectancy in 1975	1.14	1.33	0.54	1.41
Fraction of Tropical Area	-0.50	-0.56	-0.13	-0.62
Fraction GDP in Mining	0.10	0.05	0.25	0.07
Absolute Latitude	0.14	0.26	0.14	0.21
Investment Price	-0.51	-0.07	-0.12	-0.01
Real Exchange Rate Distortions	-0.03	-0.14	-0.18	-0.15
Fraction Confucius	0.04	0.02	0.21	0.01
Political Rights	-0.04	-0.14	0.01	-0.10
Openness Measure 1965-74	0.03	0.06	0.04	0.10
Years Open 1950-94	0.02	0.16	0.04	0.12
Government Share of GDP in 1970s	-0.01	0.00	-0.13	-0.06
Gov. Consumption Share 1970s	0.00	-0.05	-0.04	-0.07
British Colony Dummy	0.02	0.03	0.04	0.02
Fraction Population in Tropics	-0.03	-0.01	-0.14	-0.01
African Dummy	-0.03	-0.01	-0.52	-0.01
Fraction Buddhist	0.02	0.00	0.15	0.00
Fraction Speaking Foreign Language	0.01	0.03	0.00	0.03
Latin American Dummy	-0.01	0.00	-0.08	0.00
Nominal Government GDP Share 1970s	0.00	0.00	-0.14	-0.01
Defence Spending Share	0.00	0.01	0.11	0.01
Primary Schooling in 1975	-0.03	-0.01	0.00	-0.01
Public Investment Share	-0.02	0.00	-0.01	0.00
Population Density Coastal in 1960s	0.02	0.03	0.02	0.02
Population Density 1975	0.02	0.02	0.02	0.02
Malaria Prevalence in 1960s	-0.02	-0.04	0.00	-0.05
Terms of Trade Ranking	-0.01	-0.02	-0.04	-0.01
Ethnolinguistic Fractionalization	-0.01	-0.02	-0.05	-0.01
Revolutions and Coups	-0.01	0.00	-0.01	0.00
Higher Education 1975	-0.01	-0.01	-0.01	-0.01
Population in 1975	0.01	0.01	0.01	0.01
Fraction Muslim	0.01	0.01	0.04	0.01
Fraction Hindus	0.01	0.00	0.01	0.00
Fraction Orthodox	0.00	0.00	-0.01	0.00

Notes: Coefficients of standardized variables are obtained as the coefficients of the original variables multiplied by their in-sample standard deviation times 100. Therefore, they show the effect of a one standard deviation change in the variable in terms of percentage points of average annual growth rate over 1960-1996. For brevity, only the coefficients of the first 35 variables of Table 3 are reported.

Table 6A. Differences between PWT6.2 and PWT6.0, with different numbers of candidate variables

Number of lists	Candidates per list	*Share always*	PIP avg(MAX/MIN)
BACE			
1	67	0.41	3.18
500	18	0.58	1.84
500	10	0.65	1.37
FLS			
1	67	-	3.30
500	18	-	2.13
500	10	-	1.62
PcGets			
1	67	0.18	-
500	18	0.53	-
500	10	0.60	-

Table 6B. Differences between PWT6.2 and WDI, with different numbers of candidate variables

Number of lists	Candidates per list	*Share always*	PIP avg(MAX/MIN)
BACE			
1	67	0.47	1.64
500	18	0.79	1.32
500	10	0.84	1.16
FLS			
1	67	-	1.62
500	18	-	1.34
500	10	-	1.24
PcGets			
1	67	0.30	-
500	18	0.68	-
500	10	0.79	-

Notes: *Share always* denotes the share of variables selected with at least one of the datasets that are selected with both datasets. PIP avg(max/min) denotes the ratio of the higher to the smaller posterior inclusion probability obtained for each variable with one of the datasets, averaged over variables, and (where applicable) over 500 lists of variables.

Figure 1. R^2s and posterior probabilities in one-variable models, with PWT6.1 and PWT6.2 data