Determinants of long-term economic growth redux: A Measurement Error Model Averaging (MEMA) approach

Gernot Doppelhofer, Ole-Petter Moe Hansen and Melvyn Weeks*

February 19, 2019

Abstract

This paper estimates income per capita using recent vintages of the Penn World Tables (PWT) and investigates the robustness of determinants of long-run growth rates in a cross section of countries. We propose a novel *Measurement Error Model Averaging* (MEMA) approach that accounts for measurement error in international income data, model uncertainty over the determinants of long-term growth and allows for outliers in the form of heteroscedastic model errors. Generally, we find that newer vintages of the PWT are more precisely measured than older vintages, and that income in richer countries tend to be more accurate then poorer countries. We find that eighteen variables are robustly related to economic growth using different prior assumptions about the measurement of incomes across PWT vintages and countries. The results are robust to alternative prior assumptions about outliers and the inclusion of growth determinants, as well as inclusion of PWT vintages.

1 Introduction

The central objective of the empirical growth literature is to understand what variables are *robustly* related to economic growth. Extensive attention has been dedicated to ensure that the conclusions are robust to parameter heterogeneity, outliers and model uncertainty (see for example, Durlauf,

*Doppelhofer:Norwegian School of Economics (NHH), Department of Economics, Helleveien 30, 5045 Bergen, Norway, Gernot.Doppelhofer@nhh.no. Hansen: Norwegian School of Economics (NHH), Department of Business and Management Science, Helleveien 30, 5045 Bergen, Norway, Ole-Petter.Hansen@nhh.no. Weeks: University of Cambridge, Faculty of Economics and Clare College, Cambridge CB3 9DD, UK, mw217@cam.ac.uk.
We thank Jesus Crespo-Cuaresma, Herman van Dijk, Chris Papageorgiou, Hashem Pesaran, Xavier Sala-i-Martin, Jan-Egbert Sturm, Andrew Warner, Arnold Zellner and participants of the CESifo, ESOBE and NORMAC conferences, as well seminar participants for helpful comments. All errors are our own.

Johnson and Temple (2008) for a critical survey). Recently, a number of papers have emphasized considerable data uncertainty about the measurement income per capita and economic growth.

The Penn World Tables (PWT), which is the basis for the analysis, publish Purchasing Power Parity (PPP) adjusted income levels for many countries (Kravis, Heston and Summers, 1978). There is a vast literature on the PWT measurement and the underlying International Comparison Program (ICP)[1]. However, the PWT is subject to substantial revisions where each revision is released as a separate *vintage*. Revisions to the PWT are caused by changes in the underlying income and price data, as well as changes in methodology (see for example, Deaton and Heston (2010) and Feenstra et al. (2009)). Recently, Johnson et al. (2013) and Ciccone and Jarociński (2010) have questioned the robustness of important results in the empirical growth literature when conditioning on particular vintages of the PWT and neglecting measurement error.

This paper proposes a novel *Measurement Error Model Averaging* (MEMA) approach that estimates GDP per capita across countries and over time and simultaneously investigates the robustness of determinants of long-run growth. Income is treated as a latent variable, which is observed with classical measurement error. Using a Bayesian measurement error model, we use eight recent vintages of the PWT to identify the posterior distributions of income in 1960 and 1996. Vintage-specific fixed effects capture differences in baseline prices or other methodological differences of measuring income in the PWT. Combining the latent distributions of income per capita with a Bayesian model averaging approach allow us to assess the robustness of determinants of economic growth to measurement error and model uncertainty. We also allow for possible model misspecification and neglected parameter heterogeneity by allowing for heteroscedastic model errors.

The main findings of the paper are as follows: First, we find evidence for systematic differences in the measures of GDP per capita across different vintages of the PWT. Although there are exceptions, we generally find that newer vintages of the PWT are more precisely measured than older vintages. Second, countries differ in the quality of measured levels and growth of income per capita. Richer countries tend to be measured more accurately then poorer countries. However, we find the largest variability in income measurement for middle-to-low income countries, compared to the very poorest countries in the PWT sample. Third, we find that eighteen growth determinants appear robust to measurement error in PWT vintages, model uncertainty *and* heteroscedastic model errors. These include variables measuring initial conditions, such as initial GDP per capita, regional factors controlling for regional differences in economic growth rates, vari-

[1] See Johnson et al. (2013) for a background discussion and the ICP portal website: `http://icp.worldbank.org/icp/GlobalResult.aspx`

ables measuring geographic and climatic conditions, and finally population characteristics and cultural variables.

This paper is related to several strands in the literature.

There is an abundance of papers analysing growth determinants.[2] As shown in Kormendi and Meguire (1985), Grier and Tullock (1989), and Barro (1991) the empirical growth literature have tested alternative models and particular combinations of variables explaining economic growth. The wide variation in results casts doubt on the robustness of growth determinants. Levine and Renelt (1992) use a version of an extreme bounds analysis for growth determinants in a cross-section of countries and found that few (if any) were robust. Sala-i Martin (1997) argues that the test was too extreme and one should rather look at the distribution of model estimates across models. Recent papers address model uncertainty and investigate the robustness of growth determinants using model averaging. Fernandez, Ley and Steel (2001*b*) and Sala-i Martin, Doppelhofer and Miller (2004) came to more optimistic conclusions regarding the robustness of growth determinants and found a number of explanatory variables to be robust to model uncertainty. Durlauf, Johnson and Temple (2008) give a more recent survey the empirical growth literature.

Deaton and Heston (2010) discuss revisions in the PWT, and explain how they are related to changes in factors such price benchmarks, methodology, extrapolation strategies and updates in the underlying data. Johnson et al. (2013) discuss the PWT-revision in general, and find no reason to believe that newer vintages of the PWT are better in terms of measuring growth. An important contribution to the empirical growth literature is Ciccone and Jarociński (2010) showing the sensitivity of results in Sala-i Martin, Doppelhofer and Miller (2004) to different PWT-vintage to measures economic growth. Jarociński (2010) uses Bayesian ridge regressions to estimate growth determinants for different PWT-vintages. This sensitivity of results highlights the need for directing attention to measurement error in growth regressions.

Hausman (2001) and Hyslop and Imbens (2001) discuss the consequences of measurement error in econometric analyses. Carroll et al. (2006, p 1) calls the consequences of measurement error a "triple whammy": Bias in parameter estimates, loss of power and masking of features of the data. Although there is a wide literature on how to model measurement error in a frequentist perspective[3] our approach is more similar to the classical measurement error discussed in Richardson and Gilks (1993).

There are some examples of analyses that combine the PWT-data with measurement error models. Rao, Rambaldi and Doran (2008) proposes a

[2]For a review of *theories* of economic growth, see for example the textbooks by Barro and Sala-i Martin (2004) or Acemoglu (2009).

[3]See e.g. Goldberger (1972), Leamer (1983), Aigner et al. (1984), Black and Smith (2006), Lubotsky and Wittenberg (2006) and Browning and Crossley (2009)

method to construct panels of incomes and prices using also data from national sources. Pinkovsky and Sala-i Martin (2016) highlight measurement error in GDP per capita based on either national accounts data and surveys, and argue that this has important consequences for comparing income levels and economic growth across countries. Finally, Cuaresma et al. (2015) use several PWT-vintages together with a latent variable model to construct consensus measures of income per country. Our paper differs from these papers by simultaneously modelling measurement error of income across countries and over time and simultaneously assessing the robustness of growth determinants.

Temple (2000) argues that growth regressions are hampered by outliers and potential parameter heterogeneity. A natural extension of linear growth regressions accommodates that some observations might differ markedly from most of our data. To accomplish this we use a novel approach based on the Dirichlet distribution (Chigira and Shiba, 2015), as well as more established methods for outlier detection utilising either a binary outlier classification (Hoeting, Raftery and Madigan, 1996) or based upon mixed-normal distributions (Geweke, 1993). Accounting for outliers is important for the robustness of some variables. For example, the importance of *Mining* as growth determinant is essentially driven by one country – Botswana. Furthermore, we find that a normal distribution is ill-suited to capture uncertainty of the growth process. The variance of the growth process is seven times higher in the most compared to the least noisy country. Following Geweke (1993), we find evidence for fat tailed errors of the growth process.

The structure of the paper is as follows. Section 2 presents the measurement error model, model averaging and discusses how we connect these two modules. We estimate the model, and present results in section 3. In section 4 we present and analyse several alternative specifications of the model. Section 5 concludes.

2 The model for measurement error and model uncertainty

This section describes the details of the MEMA-model. Section 2.1 starts with the measurement error model. Section 2.2 discusses model averaging, including the accommodation of heteroscedastic model errors. Finally section 2.3 connects the measurement error and model averaging models.

2.1 Measurement Error

We propose the following model between observed measurements in the PWT and the true levels of income:

$$y_{v,i}^I = a_v + y_i^I + \sigma_{v,i}^I \varepsilon_{v,i}^I \tag{1}$$
$$y_{v,i}^E = a_v + y_i^E + \sigma_{v,i}^E \varepsilon_{v,i}^E \tag{2}$$

$y_{v,i}^I, y_{v,i}^E$ denote, respectively, the *observed* levels of income from the PWT for country i in vintages (v) for initial (I) and end (E) of period GDP; ${y_i}^I$ and ${y_i}^E$ denote the *true* (latent) values of income, and $\varepsilon_{v,i}^I, \varepsilon_{v,i}^E$ are measurement errors unique to each country-vintage. a_v is a vintage-specific level fixed effect, that allows for different PWT-vintages reported in different international US Dollars, but also other effects from the PPP methodology that shifts all measurements in a vintage.[4] To ensure identification, we fix one of the vintage specific fixed effects to zero, such that the level shifters are all defined relative to this fixed vintage. $\sigma_{v,i}^I$ and $\sigma_{v,i}^E$ are parameters that scale the variance of the measurement error for each country-vintage.

We give both the level shifters a_v and the true levels of income y_i^I, y_i^E a uniform prior over a large range. Furthermore, we assume that the measurement errors are independent, standard normal:[5]

$$\begin{aligned} \varepsilon_{v,i}^I &\sim N(0,1) \\ \varepsilon_{v,i}^E &\sim N(0,1) \end{aligned} \tag{3}$$

To close the measurement error model, we need to specify a prior structure for the scale terms $\sigma_{v,i}^I$ and $\sigma_{v,i}^E$ for the measurement errors across vintages and countries. A special feature of the data is that we have repeated measurements of both *countries* (i) and *vintages* (v). It could be the case that measurements in some vintages and some countries are inherently more noisy than others. To allow for this, we separate the scale terms according to the following product:

$$\begin{aligned} {\sigma_{i,v}^I}^2 &= \omega_i^N \omega_v^V \ {\sigma^I}^2 \\ {\sigma_{i,v}^E}^2 &= \omega_i^N \omega_v^V \ {\sigma^E}^2 \end{aligned} \tag{4}$$

[4] We would like to emphasize that even though this parameter is a "fixed effect" with the same value for all income measurements in a given vintage, we still treat the fixed effect as a parameter in a Bayesian sense - i.e. it has both a prior and posterior *distribution.*

[5] By independent, we mean that each error term is independent of all other error terms, which implies that

$$\begin{aligned} Cov(\varepsilon_{j,l}^I, \varepsilon_{h,m}^I) &= 0, \ \forall \ j,l \neq h,m \\ Cov(\varepsilon_{j,l}^E, \varepsilon_{h,m}^E) &= 0, \ \forall \ j,l \neq h,m \\ Cov(\varepsilon_{j,l}^I, \varepsilon_{h,m}^E) &= 0, \ \forall \ j,l \neq h,m \end{aligned}$$

${\sigma^I_{i,v}}^2$ and ${\sigma^E_{i,v}}^2$ are now the variance of measurement error for country i in vintage v for initial and end period income, respectively. ${\sigma^I}^2$ and ${\sigma^E}^2$ are *average* variances of measurement errors for initial and end income across all countries and vintages, and ω^N_i and ω^V_v are the *relative* variance of measurement errors for countries and vintages. This setup implies that the average value of ω^V_v and ω^N_i must both be unity.

A prior that satisfies this condition are referred to as scaled Dirichlet distributions. Starting with the prior on the relative variance of measurement error by countries, we use the following prior:

$$\left(\omega^N_1, ..., \omega^N_N\right)/N \sim Dir\left(\Omega^N_1, ..., \Omega^N_N\right) \tag{5}$$

Where the parameters $\Omega^N_1, ..., \Omega^N_N$ are constants. First, we can note that by setting all constants $\Omega^N_1 ... = \Omega^N_N$ we are taking an *a priori* agnostic approach as to which countries are measured with error. Second, a higher value of these constants imply strengthening the prior. As an example, if we set all $\Omega^N_1, ..., \Omega^N_N$ to the same, high value, we impose a strong belief in that the variance of the measurement error is the same in all countries. Hence, we will force the posterior to be close to the prior. Alternatively, by setting the constants $\Omega^N_1, ..., \Omega^N_N$ to the same low value, we let the data decide where variance of measurement error is higher. Third, we do not have to place an equal value of $\Omega^N_1, ..., \Omega^N_N$. If we have an *a priori* strong belief that some countries have a better data quality than others we can impose this belief through the constants.

Next, we specify the prior for the relative variance of PWT vintages. We specify a similar prior for this as with the relative variance of measurement error by countries, as

$$\begin{aligned}\left(\omega^V_1, ...\omega^V_V\right)/V \sim Dir\left(\Omega^V_1, ...\Omega^V_V\right)\\ \left(\omega^V_1, ...\omega^V_V\right)/V \sim Dir\left(\Omega^V_1, ...\Omega^V_V\right)\end{aligned} \tag{6}$$

This implies that we are completely agnostic as to which PWT vintages contains more measurement error.[6]

Finally, we give an uniform prior for σ^I and σ^E over a large range.[7]

$$\sigma^I \sim U(0, 1000) \tag{7}$$

$$\sigma^E \sim U(0, 1000) \tag{8}$$

This completes the specification of the measurement error model.

[6]For a discussion and relaxation of this prior, see section 4.2.

[7]See Gelman (2006) for a discussion of prior of variance parameters, as well as a brief discussion of the uniform prior on standard deviations.

2.2 Model Averaging

Consider the typical cross-country growth regression of the form:

$$\frac{y_i^E - y_i^I}{T_1 - T_0} = \alpha + \sum_{k=1}^{K} x_{k,i}\beta_k\gamma_k + \sigma_i\varepsilon_i \tag{9}$$

where the left hand side is average growth for country i, where latent initial y_i^I and end period y_i^E income are estimated using the measurement error model outlined in the previous section 2.1. β_k is the coefficient of variable k, σ_i is a scaling parameter and ε_i is an independent, standard normal error term. A particular model is described by the binary parameter γ_k, indicating whether variable k is included in the regression or not. Note that an intercept is always included in the growth regression.

The benchmark model averaging approach assumes that σ_i are identical, i.e. that the errors are *conditionally* homoscedastic across countries. The assumption of homoscedastic model errors may be violated by model misspecification and neglected parameter heterogeneity. We therefore allow for outliers and heteroscedastic model errors. Equation (9) nests all possible linear combinations of growth determinants K. In our setting, this is a fairly large model space. To see this, note that we can use the binary conversion formula

$$M = \sum_{k=1}^{67} \gamma_k 2^{k-1} \tag{10}$$

where M is an integer, denoting one of 2^{67} unique models.

Following the (Bayesian) model averaging literature, the following prior structure is assumed for parameters in each model (see for example Fernandez, Ley and Steel (2001*a*)). The prior slope coefficients β that are included in a given model are normally distributed with mean zero and variance $\sigma^2\mathbf{V}_{0j}$:

$$\beta|\sigma^2, M \sim N(0, \sigma^2\mathbf{V}_{0j}) \tag{11}$$

The prior variance matrix is assumed to be proportional to the sample covariance

$$\mathbf{V}_{0M} = (g_0\mathbf{X}'_M\mathbf{X}_M)^{-1} \tag{12}$$

with factor of proportionality g_0, and $\mathbf{X}_M$ is the matrix of covariates that are included in model M. This *g-prior* was first suggested by Zellner (1986), and is a convenient way to specify the prior variance matrix, in particular in the presence of considerable model uncertainty. Different values of the g-prior parameter g_0 have been proposed in the literature (see Fernandez, Ley and Steel (2001*a*)).[8] To contrast the results in Sala-i Martin, Doppelhofer and

[8] Zeugner and Feldkircher (2009) warn that an overly diffuse prior concentrates estimation on a few models, what they call the 'supermodel effect'. This effect is contributing to the sensitivity of estimates across different samples of the Penn World Tables found by Ciccone and Jarociński (2010).

Miller (2004), this paper follows their assumption that the prior distribution of the slope coefficient β is dominated by the sample information, implying a diffuse prior variance. We therefore set $g_0 = N^{-1}$ as a benchmark.[9]

In the benchmark case, we place a uniform prior on σ over a large, positive range:

$$\sigma \sim U(0, 1000) \tag{13}$$

2.2.1 Model Space Prior

Letting π_k be the independent prior inclusion probability of variable x_k in model M, the prior probability for model M is given by:

$$p(M) = \prod_{k=1}^{K} \pi_k^{\gamma_k}(1-\pi_k)^{1-\gamma_k} \tag{14}$$

Recall the binary indicator variable γ_k measures inclusion (exclusion) of variable x_k.[10] One approach is to assume a completely diffuse or uniform prior across all models, which corresponds to a prior inclusion probability equal to $\pi_k = 1/2$ for all variables. However, with a relatively large number of regressors, a uniform prior implies that the great majority of prior probability is allocated to models with a large number of variables. As an alternative, Sala-i Martin, Doppelhofer and Miller (2004) advocate in their Bayesian Averaging of Classical Estimates (BACE) approach a preference for more parsimonious models with a smaller prior expected model size $\bar{k} = 7$, which seems reasonable given the relatively large number of growth determinants ($K = 67$).[11] We follow the BACE-prior, and place independent Bernoulli priors on the γ_k, with prior inclusion probability 7/67:[12]

$$\gamma_k \sim Bern\left(\frac{7}{67}\right) \tag{15}$$

2.2.2 Modelling Outliers

The empirical growth model can fit poorly for some observations compared to others. This could be caused by a growth process being more variable in some countries than others, or possibly a misspecification of the model where relevant higher order terms are omitted. If this is the case, we would want to prevent these outliers from exerting a strong influence on results.

[9] Appendix 4.3 allows for a hierarchical prior on the hyperparameter g_0.

[10] Mitchell and Beauchamp (1988) first suggested this prior with discrete probability mass or "spike" at zero, representing the prior uncertainty that a regressor should be included. George and McCulloch (1993) propose a Bayesian alternative of using a proper prior distributions with large variance.

[11] O'Hara and Sillanpää (2009) note in their very practical review that "*sparsity has to be forced onto a model; the data themselves may not demand it*" (p 112).

[12] Section 4.3 allows for a hierarchical prior on the prior model size.

To do this we employ a similar method as is commonly used to capture heteroscedasticity in measurement errors. Specifically, we estimate an average model variance (σ^2) as one single parameter, and use a Dirichlet-weighting to estimate the scale of the error for each observation relative to the average.

First, define relative variance of model errors as a Dirichlet of size N:

$$(\omega_1, ..., \omega_N)/N \sim Dir\,(\Omega, ..., \Omega) \tag{16}$$

We interact this with the average variance σ^2, such that the variance for a given observation is

$$\sigma_i^2|\omega_i, \sigma = \omega_i\sigma^2 \tag{17}$$

This setup is very similar to Chigira and Shiba (2015). An alternative would be to specify the model using the mixture-Normal error structure introduced by Geweke (1993).[13] We estimate the model using robust errors following Geweke (1993)Geweke as well as a binary outlier classification method in section 4.

2.3 Measurement Error Model Averaging

We can now combine the measurement error model from section 2.1 with the model averaging 2.2. First, note that the growth equation can be written as

$$y_i^E = \mu_i + \varepsilon_i(T_1 - T_0) \tag{18}$$

where $\mu_i \equiv \left(\alpha + \sum_{k=1}^K x_{k,i}\beta_k\gamma_k\right)(T_1 - T_0) + y_i^I$ is the sum of initial income and economic growth predicted by the regression model. We use equation (18) to substitute for final income in the measurement equation. Considering all V of end-of-period for country i, we can stack these in the following vector:

$$\begin{bmatrix} y_{1,i}^E \\ \vdots \\ y_{V,i}^E \end{bmatrix} = \begin{bmatrix} a_1 + \mu_i + \sigma_i\varepsilon_i(T_1 - T_0) + \sigma_{1,i}^E\varepsilon_{1,i}^E \\ \vdots \\ a_v + \mu_i + \sigma_i\varepsilon_i(T_1 - T_0) + \sigma_{V,i}^E\varepsilon_{V,i}^E \end{bmatrix} \tag{19}$$

This implies that end-of-period measures of GDP per capita of one country have a multivariate normal distribution with a given structure of the covariance matrix:[14]

$$\begin{bmatrix} y_{1,i}^E \\ \vdots \\ y_{V,i}^E \end{bmatrix} \sim N\left(\begin{bmatrix} a_1 + \mu_i \\ \vdots \\ a_V + \mu_i \end{bmatrix}, \begin{bmatrix} \tilde{\sigma}_i^2 + \sigma_{1,i}^{2,E} & \cdots & \tilde{\sigma}_i^2 \\ \vdots & \ddots & \vdots \\ \tilde{\sigma}_i^2 & \cdots & \tilde{\sigma}_i^2 + \sigma_{V,i}^{2,E} \end{bmatrix}\right) \tag{20}$$

[13] See Sims (2010, p20-23) for a discussion of heteroskedasticity robust estimation in a Bayesian setting. Chigira and Shiba (2015) further argue that the Dirichlet-model of heteroskedasticity is superior to the established Geweke (1993) Student-t model of outliers with gamma priors, as it is less informative on the model of heteroskedasticity.

[14] Define $\tilde{\sigma_i}^2 \equiv \sigma_i^2(T_1 - T_0)^2$

Together with the priors for the ME and MA models, as well as the distributional assumptions on initial income, we have now completed the specification of the MEMA-model. The following section report the results we obtain with it.

3 Results

This section presents the results from estimating the MEMA model under three different assumptions. First, we condition on a particular vintage and estimate results by benchmark model averaging, which is a special case of the MEMA model. Second, we allow for measurement errors across countries and PWT vintages using the MEMA model. Third, we allow for outliers using robust model averaging and the MEMA model combined.

3.1 MEMA-model results

Measurements of income per capita and economic growth across different vintages of the PWT exhibit a large degree of uncertainty (see Johnson et al. (2013) or Deaton and Heston (2010)). These papers also warn that there may be systematic mismeasurement across countries, for example that income in poorer countries is likely to be less precisely measured compared to richer countries GDP. We are therefore proposing to address measurement error across PWT-vintages and countries simultaneously.

(Mis)Measurement of incomes

We start by estimating the measurement error (ME) model discussed in section 2.1. We use a flat prior on the relative variances of countries and vintages, where $\Omega_1^V = ... = \Omega_V^V = \Omega_1^N = ... = \Omega_N^N = 1$. This is a fairly uninformative prior, such that the data can pull the relative variances away from the prior. We estimate the true values of initial and end-of-period income per country in 1960 and 1996, respectively.[15]

Figures 1 and 2 show the posterior densities of estimated true initial and end-of-period income. The blue dots indicate median log income, the thick line shows a 68% and the thin line a 95% credible interval, respectively. The figures also show all measurements in all PWT-vintages with black circles. A striking feature of both these figures is that the greatest variability is not for the lowest income countries, but rather for those at the middle-to-low range. Hence, measuring PPP-adjusted income in countries that are close to subsistence might be easier than in countries that have risen somewhat above this low level of income.

[15] The initial value in 1960 and end period in 1996 were chosen for comparison with the literature (see Sala-i Martin, Doppelhofer and Miller (2004), Ciccone and Jarociński (2010)).

[Insert figure 1 about here]

[Insert figure 2 about here]

The measurement error model estimates the relative variances across PWT vintages and countries. This helps us to understand measurement error problems present in this dataset, and make statistical inference and economic implications robust to measurement error.

Table 5 shows the posterior densities of relative variance of measurement error of income per capita for each PWT vintage. In particular, PWT vintage 6.0 has at the mean more than twice the variance compared to the average vintage, whereas recent vintages 8.0 and 8.1 have almost half the variance of the average vintage. Although there are exceptions, we generally find that newer vintages are less noisy than older ones, addressing the question posed by Johnson et al. (2013).

[Insert table 5 about here]

Figure 3 shows relative variances of the measurement error of income per capita per country. There is a vast difference in the variance across countries. Incomes in El Salvador, Zimbabwe and Liberia are at the extremely noisy end of the scale. At the other end of the scale we find France, Belgium and Canada, where there is very little difference of income measurement across different PWT-vintages.

[Insert figure 3 about here]

[Insert table 2 about here]

Growth determinants in a homoscedastic model

We now show estimation results for the growth determinants using the full MEMA-model. The estimated coefficients accounts for both measurement error across PWT vintages and countries and model uncertainty (see section 2.3).

Table 3 shows the posterior inclusion probabilities (PIPs), which represent a summary measure of the importance of a growth determinant. In particular, we can contrast the PIPs shown in the table with the prior inclusion probability, which equals 7/67.

[Insert table 3 about here]

The results in Table 3 show that for thirteen explanatory variables the data are indicating that they are important determinants of economic growth. These variables include variables based on neoclassical (or endogenous) growth models, such as *Initial log GDP per capita*, controlling for initial conditions or determinants of the steady state, *Primary school enrolment in 1960*, controlling for human capital, the *Price of investment goods* or *Life expectancy in 1960.* A second group of variables included regional factors, such as the *East Asian Dummy* and a dummy for *Sub-Saharan Africa.* These variables control for regional differences in economic growth that are present even after controlling for many other plausible growth determinants. A third set of variables measure geographic or climatic conditions, such as the *Fraction of Tropical Area*, *Air Distance*, the overall *Population density in 1960*, as well as *Coastal population density in 1960.* A final group includes population characteristics or cultural variables, such as the population *Fraction Confucian* and *Fraction Muslim.*

The posterior inclusion probabilities in the first four columns shows that the results are quite similar regardless of the exact specification of the variance of the measurement error. As an example, we can note that *Malaria prevalence* is marginally important, and *Fraction Confucian* as an important variable.

3.2 Outlier robust results

An important issue in the context of the empirical growth literature is the possibility of outliers and heteroscedasticity of the model errors. We therefore estimate the MEMA model allowing a more flexible model error structure. The results are shown in the last columns in Tables 4, where the first column is a reproduction of the homoscedastic model from the most flexible MEMA-model of Table 3. The fifth column is the same as the fourth column allowing for measurement error across PWT vintages and countries, except that we also allow for outliers using a binary classification of whether each observation is an outlier. The sixth column uses instead a flexible Dirichlet-weighing of model error variance, with a flat unit priors on $\Omega_1, ..., \Omega_N$.

The results of the MEMA model with and without allowing for outliers adds some interesting features. First, we can note that two additional variables, namely a dummy for *Latin America* and the Malaria Prevalence in 1960, have PIP exceeding the prior inclusion probability in the homoscedastic model in column 1 of Table 4. Interestingly, the PIP associated with these variables *increases* in the last two columns once we allow for heteroscedastic model errors, indicating that outliers might be present in models including these two variables. A few more marginal variables, such as a dummy for *Landlocked countries*, *Openness in the 1960s*, and a *European dummy*, are helped by allowing for outliers. The reverse is true for other variables, the PIPs clearly fall once we allow for heteroscedastic model errors. This implies that variables such as the *Number of years a country is open*, *Political rights*, *Ethnoliguistic fractionalization*, and notably the *Mining share of GDP* are not robust to outliers, indicating that a few extreme observations may be driving these results.

[Insert table 4 about here]

Figure 4 shows the posterior densities of the relative variance of countries' model error. Here, variance of the most noisy country is almost seven times the variance of the least noisy country. Hence, with the Dirichlet weighting the most noisy country - e.g. Botswana - contribute very little to the identification of parameters of in the MA-model.Hence, the posterior inclusion probability of mining, which has a high value in Botswana, drops to 2.7% in the Dirichlet robust model.

[Insert figure 4 about here]

Figure 5 shows model predicted growth from the full MEMA-model with Dirichlet weighing outliers, together with measurements of growth from all PWT-vintages. From this figure we can observe that for countries such as Botswana and Philippines and South Africa the MA-model fits poorly. Botswana is a case special as growth is has been exceptionally high. South Africa and the Philippines are at the other extreme, where performance has been lower than what their initial conditions predict.

[Insert figure 5 about here]

Table 6 shows detailed results for our preferred specification, the full MEMA-model with Dirichlet robust model error. The table reports the mean of each coefficient, conditional on being included and the standard deviation of each coefficient. The table further reports the sign certainty, which is the posterior probability of the sign of the coefficient being equal to the sign of the conditional mean. Finally, the table repeats the posterior inclusion probability for each variable.

[Insert table 6 about here]

The results reported in Table 6 give a clear indication regarding the robustness of growth determinants allowing for measurement error and outliers. Eighteen variables have PIP larger than the prior inclusion probability. Posterior coefficients are relatively precisely estimated with sign certainty exceeding 0.975.[16] For the remaining 49 variables the posterior inclusion probability is below the prior and we are also less sure about the sign of the associated coefficients.

[16] This implies that the sign certainty probability can be interpreted as a test statistic associated with a two-sided confidence interval for a coefficient estimate being zero. The European dummy has sign certainty 0.967 and PIP equal to 0.11 marginally exceeding the prior threshold.

4 Robustness

4.1 Alternative outlier specification

4.1.1 Binary outlier classification

A maintained assumption in the benchmark case is that regression errors are homoscedastic. A useful point of departure to model robustness is to assume that the errors in the growth process can be described by a combination of two normal distributions.

$$p\left(\sigma_i \epsilon_i | \varpi_i, \rho, \sigma\right) = (1 - \varpi_i) N\left(0, \sigma^2\right) + \varpi_i N\left(0, \rho\sigma^2\right) \tag{21}$$

where the mixture is governed by two parameters. The binary parameter ϖ_i identifies whether an observation is an outlier, and the parameter ρ controls the degree of variance-inflation for the outlying observations. Hoeting, Raftery and Madigan (1996) adopt this approach in a study which simultaneously selects regressors and identifies outliers. In the particular application of their paper, the prior probability of an observation being classified as an outlier and ρ are treated as fixed, with the proportion of outliers chosen based upon the size of the dataset. We use the following distributional assumptions:

$$\begin{aligned} \varpi_i &\sim Bernoulli(.1) \\ \rho - 1 &\sim Exp(.1) \end{aligned} \tag{22}$$

This places a 10% prior probability on a given observation being classified as an outlier. The fairly non-informative exponential prior on variance-inflation parameter implies outliers have a far greater variance than non-outliers, with a prior expected value of $E[\rho - 1] = 10$. The variance of an observation in the growth model is therefore

$$\sigma_i^2 | \varpi_i, \sigma, \rho = (1 - \varpi_i)\sigma^2 + \varpi_i \rho \sigma^2 \tag{23}$$

4.1.2 Geweke-robust model errors

Geweke (1993) propose a different parametric approach to the Dirichlet-weighting of country variances in the growth model. In particular, the Geweke-approach use the following setup:

$$\begin{aligned} \nu &\sim exp\left(25^{-1}\right) \\ \nu / \sigma_i &\sim \boldsymbol{\chi}\left(\nu\right) \end{aligned} \tag{24}$$

This is equivalent to interpreting the model errors as draw from a t-distribution with ν degrees of freedom.

4.1.3 Results

Table 7 shows the results using these extensions. The results from the baseline MEMA-model are repeated in the leftmost column. The second column shows PIPs from the extended MEMA-model with random prior model size and Zellner factor g_0. First, we can note that the posterior mean of the hyper-inclusion probability θ is .15, which gives a larger model size compared to baseline MEMA. Furthermore, the posterior mean of the Zellner factor g_0 is .04, slightly higher than the constant provided in baseline MEMA at $N^{-1} \approx .01$. It is therefore not surprising that this extension flags up more variables compared to baseline MEMA, where now an additional six variable have a PIP over 7/67.

[Insert table 7 about here]

The mean of the posterior density if the variance inflation in the binary outlier classification model is 21.43 - i.e. variance of the model error is vastly greater for outliers relative to non-outliers.

The fourth column shows PIPs with Geweke-robust errors. The posterior mean of ν is 2.65, which implies that the growth process has very fat tails. Apart from that, all variables from baseline MEMA still have a PIP over 7/67, in addition to two variables that now have a PIP marginally above the prior.

4.2 Including PWT Vintage 9.0

A new version of the The Penn World Tables has recently been published. As this vintage has not received attention in Pinkovskiy and Sala-i Martin (2016) and Johnson et al. (2013), we do not use this vintage in the main text. However, we examine whether our main results are sensitive the inclusion of this additional vintage. We therefore estimate our preferred model on all PWT vintages from 6.0 to 9.0.

In our preferred specification, we are agnostic as to which PWT-vintage has the least variability of measurement error. However, this approach might be unreasonable. Although not supported by Pinkovskiy and Sala-i Martin (2016), one might expect that newer vintages are better than older vintages. This prior belief could be imposed as a soft constraint, using $\Omega_1^V > ... > \Omega_V^V$, such that the prior expected value of the relative variance is decreasing with newer vintages. As a separate robustness test, we impose the belief that "newer is better" with a hard constraint, using

$$\begin{aligned} \left(\tilde{\omega}_1^V, ...\tilde{\omega}_V^V\right)/V &\sim Dir\left(\Omega_1^V, ...\Omega_V^V\right) \\ \omega_1^V, ..., \omega_V^V &= \textsc{sort}\left(\tilde{\omega}_1^V, ...\tilde{\omega}_V^V\right) \end{aligned} \tag{25}$$

where $\textsc{sort}(\cdot)$ is a function that sorts the elements of $\cdot$ in a descending order. This prior implies that the relative variance of measurement error in vintage

t must always be weakly smaller than the relative variance of measurement error of vintage $t-1$.

4.2.1 Results

[Insert table 8 about here]

4.3 Random model size and g-prior

In the baseline MEMA-model, we assume the prior inclusion probability is given by 7/67. Following Brown, Vannucci and Fearn (1998), we relax this assumption, by replacing the prior distribution of variable inclusion parameters to the following:

$$\begin{aligned} \theta &\sim Beta\left(1, \frac{60}{7}\right) \\ \gamma_k &\sim Bernoulli\left(\theta\right) \end{aligned} \tag{26}$$

This gives a prior expected model size equal to 7/67 as before, but allows for another layer of uncertainty.[17]

Another extension is to allow for more flexibility in the Zellner g-prior. In the benchmark MEMA model, we use $g_0 = N^{-1}$. However, following Liang et al. (1998)[18], this can be treated as a parameter as opposed to a constant with prior given by a Beta distribution

$$\frac{1}{1+g_0} \sim Beta\left(1, N^{-1}\right) \tag{27}$$

which implies that $E\left[\frac{1}{1+g_0}\right] = \frac{1}{1+N^{-1}}$. This is within the range Liang et al. (1998) calls "reasonable".

4.3.1 Results

[Insert table 10 about here]

5 Conclusion

There is considerable uncertainty about the levels and growth rates of income per capita. Although the PWT provides measures of income across countries and over time, there is considerable variation across different vintages of the PWT. The uncertainty about the measures of income spills over to increased uncertainty about the robustness of growth determinants.

[17] See Ley and Steel (2009) for a discussion.

[18] Note that Liang et al. (1998) discuss the distribution of the *inverse* of g_0.

This paper proposes a MEMA approach that models measurement uncertainty together with model uncertainty. Using nine vintages of the PWT to estimate the model, we have found 18 variables robustly related to economic growth from 1960 to 1996. The results are robust to allowing for outliers in the form of heteroscedastic model errors. Furthermore, we have in this paper *quantified* the noisiness of data across both PWT vintages and countries, which extends the qualitative measure of data quality contained in some vintages of the PWT.

We are in this paper trying to remain agnostic in our prior specifications. However, given that we are asking a lot from a very limited amount of data, it is necessary to impose parametric assumptions to ensure a well behaved posterior. The MEMA model can be extended by introducing additional information that can help to identify income and economic growth.

References

Acemoglu, Daron. 2009. *Introduction to Modern Economic Growth.* Princeton University Press.

Aigner, Dennis J., Cheng Hsiao, Arie Kapteyn, and Tom Wansbeek. 1984. “Handbook of Econometrics.” , ed. Z. Griliches and M.D. Intriligator Vol. 2, Chapter Latent Variable Models in Econometrics, 1321–1393. Elsevier.

Barro, Robert J. 1991. “Economic Growth in a Cross Section of Countries.” *Quarterly Journal of Economics*, 106(2): 407–443.

Barro, Robert J., and Xavier Sala-i Martin. 2004. *Economic Growth.* . 2nd ed., MIT Press.

Black, Dan A, and Jeffrey A Smith. 2006. “Estimating the returns to college quality with multiple proxies for quality.” *Journal of labor Economics*, 24(3): 701–728.

Browning, Martin, and Thomas Crossley. 2009. “Are two cheap, noisy measures better than one expensive, accurate one?” *The American Economic Review*, 99(2): 99–103.

Brown, Philip J, Marina Vannucci, and Tom Fearn. 1998. “Multivariate Bayesian variable selection and prediction.” *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 60(3): 627–641.

Carroll, Raymond J, David Ruppert, Leonard A Stefanski, and Ciprian M Crainiceanu. 2006. *Measurement error in nonlinear models: a modern perspective.* CRC press.

Chigira, Hiroaki, and Tsunemasa Shiba. 2015. “Dirichlet Prior for Estimating Unknown Regression Error Heteroskedasticity.” *TERG Discussion Papers*, 2015(341): 1–17.

Ciccone, Antonio, and Marek Jarociński. 2010. “Determinants of economic growth: will data tell?” *American Economic Journal: Macroeconomics*, 2(4): 222–246.

Cuaresma, Jesus Crespo, Martin Feldkircher, Bettina Grűn, Paul Hofmarcher, and Stefan Humer. 2015. “Measuring income with Penn World Table data: Reaching consensus.”

Deaton, Angus, and Alan Heston. 2010. “Undecountries PPPs and PPP-based national accounts.” *American Economic Journal: Macroeconomics*, 2(4): 1–35.

Durlauf, Steven N., Paul A. Johnson, and Jon R.W. Temple. 2008. "Handbook of Economic Growth." , ed. Phillipe Aghion and Steven N. Durlauf, Chapter Growth Econometrics, 555–677. Elsevier: North Holland.

Feenstra, Robert, Alan Heston, Marcel P. Timmer, and Haian Deng. 2009. "Estimating real production and expenditures across countries: A proposal for improving the Penn Wold Tables." *Review of Economics and Staistics*, 91: 201–212.

Fernandez, Carmen, Eduardo Ley, and Mark FJ Steel. 2001*a*. "Benchmark priors for Bayesian model averaging." *Journal of Econometrics*, 100(2): 381–427.

Fernandez, Carmen, Eduardo Ley, and Mark FJ Steel. 2001*b*. "Model uncertainty in cross-country growth regressions." *Journal of applied Econometrics*, 16(5): 563–576.

Gelman, Andrew. 2006. "Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper)." *Bayesian analysis*, 1(3): 515–534.

George, Edward I, and Robert E McCulloch. 1993. "Variable selection via Gibbs sampling." *Journal of the American Statistical Association*, 88(423): 881–889.

Geweke, John. 1993. "Bayesian treatment of the independent student-t linear model." *Journal of Applied Econometrics*, 8(S1): S19–S40.

Goldberger, Arthur S. 1972. "Structural Equation Methods in the Social Sciences." *Econometrica*, 40(6): 979–1001.

Grier, Kevin B., and Gordon Tullock. 1989. "An Empirical Analysis of Cross-National Economic Growth, 1951-80." *Journal of Monetary Economics*, 24(2): 259–276.

Hausman, Jerry. 2001. "Mismeasured Variables in Econometric Analysis: Problems from the Right and Problems from the Left." *Journal of Economic Perspectives*, 15(4): 57–67.

Hoeting, Jennifer, Adrian E Raftery, and David Madigan. 1996. "A method for simultaneous variable selection and outlier identification in linear regression." *Computational Statistics & Data Analysis*, 22(3): 251–270.

Hyslop, Dean R, and Guido W Imbens. 2001. "Bias from classical and other forms of measurement error." *Journal of Business & Economic Statistics*, 19(4): 475–481.

Jarociński, Marek. 2010. "Imposing Parsimony in Cross-Country Growth Regressions." *ECB Working Paper 1234*.

Johnson, Simon, William Larson, Chris Papageorgiou, and Arvind Subramanian. 2013. "Is newer better? Penn World Table revisions and their impact on growth estimates." *Journal of Monetary Economics*, 60(2): 255–274.

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

Kravis, Irving B., Alan Heston, and Robert Summers. 1978. "Real GDP per capita for more then one hundred countries." *Economic Journal*, 88: 215–242.

Leamer, Edward E. 1983. "Handbook of Econometrics." , ed. Z. Griliches and M.D. Intriligator Vol. 1, Chapter Model Choice and Specification Analysis, 285–330. Elsevier.

Levine, Ross E., and David Renelt. 1992. "A Sensitivity Analysis of Cross-Country Growth Regressions." *American Economic Review*, 82(4): 94263.

Ley, Eduardo, and Mark FJ Steel. 2009. "On the effect of prior assumptions in Bayesian model averaging with applications to growth regression." *Journal of applied econometrics*, 24(4): 651–674.

Liang, Feng, Rui Paulo, German Molina, Merlise A Clyde, and Jim O Berger. 1998. "Mixtures of g priors for Bayesian variable selection." *Journal of the American Statistical Association*, 103(481): 410–423.

Lubotsky, Darren, and Martin Wittenberg. 2006. "Interpretation of Regressions with Multiple Proxies." *Review of Economics and Statistics*, 88(3): 549–562.

Mitchell, Toby J, and John J Beauchamp. 1988. "Bayesian variable selection in linear regression." *Journal of the American Statistical Association*, 83(404): 1023–1032.

O'Hara, Robert B, and Mikko J Sillanpää. 2009. "A review of Bayesian variable selection methods: what, how and which." *Bayesian analysis*, 4(1): 85–117.

Pinkovskiy, Maxim, and Xavier Sala-i Martin. 2016. "Newer need not be better: evaluating the Penn World Tables and the World Development Indicators using nighttime lights." National Bureau of Economic Research.

Pinkovsky, Maxim, and Xavier Sala-i Martin. 2016. "Lights, Camera ... Income! Illuminating the National Accounts-Household Surveys Debate." *Quarterly Journal of Economics*, 579–631.

Rao, D.S. Prasada, Alicia Rambaldi, and Howard Doran. 2008. "A Method to Construct World Tables of Purchasing Power Parities and Real Incomes Based on Multiple Benchmarks and Auxiliary Information: Analytical and Empirical Results." *Working Paper WP05/2008, School of Economics, University of Queensland.*, , (WP05/2008).

Richardson, Sylvia, and Walter R Gilks. 1993. "A Bayesian approach to measurement error problems in epidemiology using conditional independence models." *American Journal of Epidemiology*, 138(6): 430–442.

Sala-i Martin, Xavier. 1997. "I just ran 2 million regressions." *American Economic Review (Papers and Proceedings)*, 87(2): 178183.

Sala-i Martin, Xavier, Gernot Doppelhofer, and Ronald I Miller. 2004. "Determinants of long-term growth: A Bayesian averaging of classical estimates (BACE) approach." *American Economic Review*, 813–835.

Sims, Chris. 2010. "Understanding non-bayesians." *Unpublished chapter, Department of Economics, Princeton University.*

Temple, Jonathan. 2000. "Growth regressions and what the textbooks dont tell you." *Bulletin of Economic research*, 52(3): 181–205.

Zellner, Arnold. 1986. "Bayesian estimation and prediction using asymmetric loss functions." *Journal of the American Statistical Association*, 81(394): 446–451.

Zeugner, Stefan, and Martin Feldkircher. 2009. "Benchmark priors revisited: on adaptive shrinkage and the supermodel effect in Bayesian model averaging." *IMF Working Papers*, 1–39.

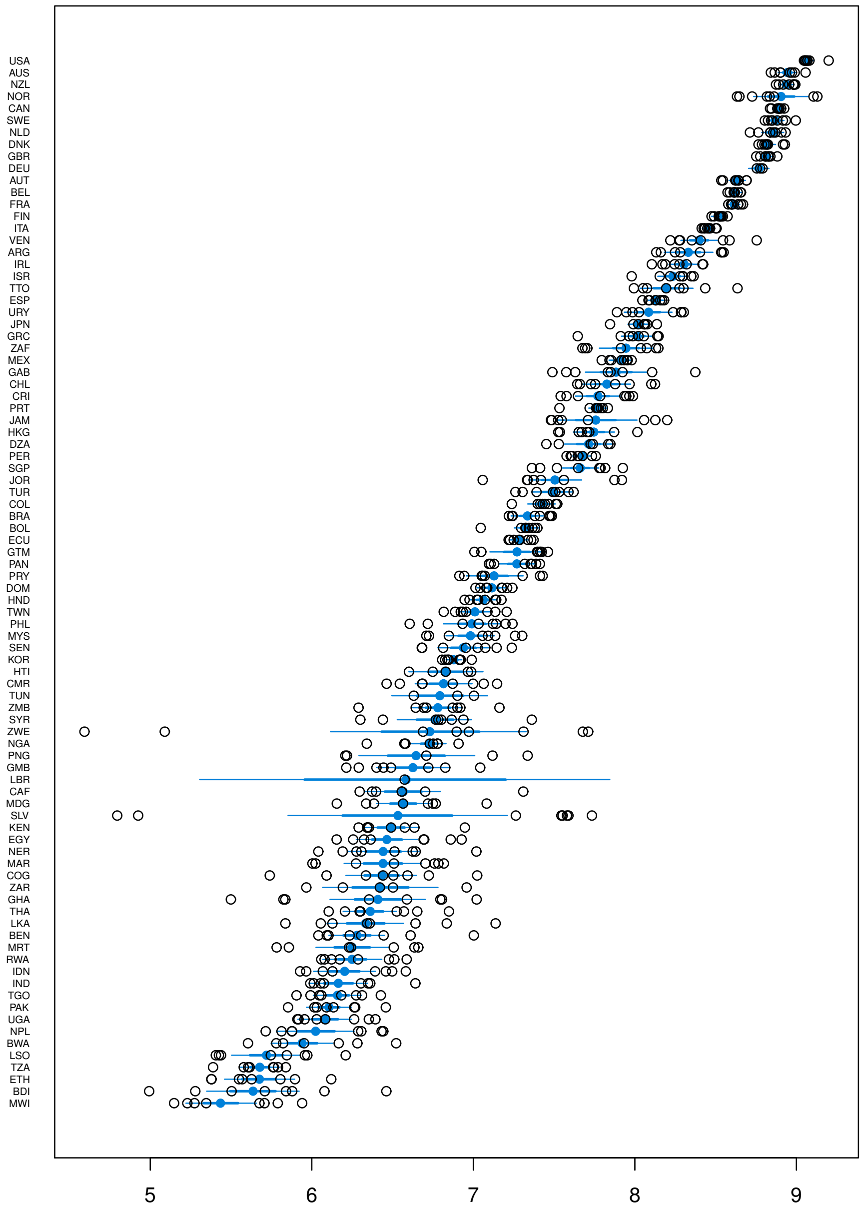

Figure 1: Income income 1960

The figure shows posterior density of PPP adjusted log GDP in 1960, estimated from all vintages 6.0 through 8.1 of the PWT. The circles indicate the level of log GDP the different Penn Vintages. Each vintage is scaled by the mean, estimated vintage level (α), such that all values are reported in "PWT 6.0"-level. The blue dot indicates the median of the estimated log income; the thick line shows a 68% credible interval and the thin line shows a 95% credible interval

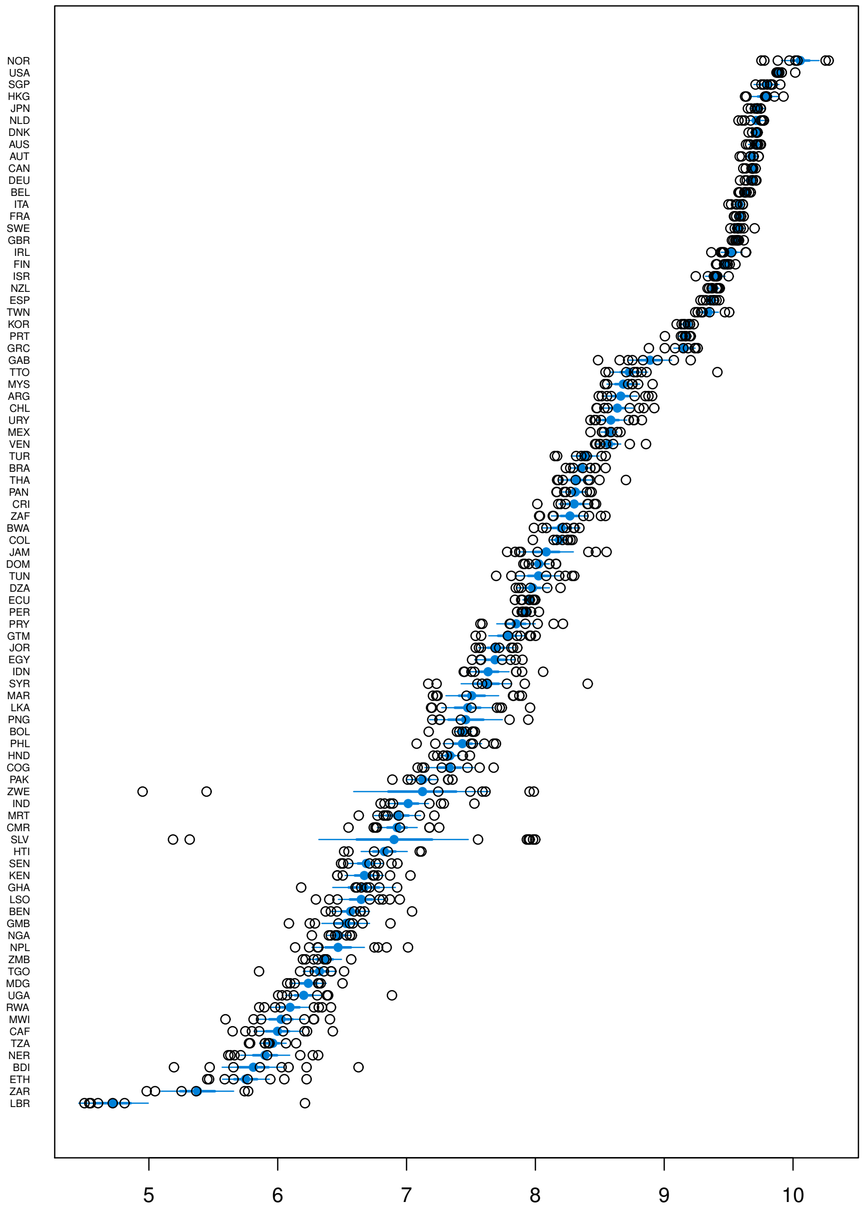

Figure 2: Income income 1996

The figure shows posterior density of PPP adjusted log GDP in 1960, estimated from all vintages 6.0 through 8.1 of the PWT. The circles indicate the level of log GDP the different Penn Vintages. Each vintage is scaled by the mean, estimated vintage level (α), such that all values are reported in "PWT 6.0"-level. The blue dot indicates the median of the estimated log income; the thick line shows a 68% credible interval and the thin line shows a 95% credible interval

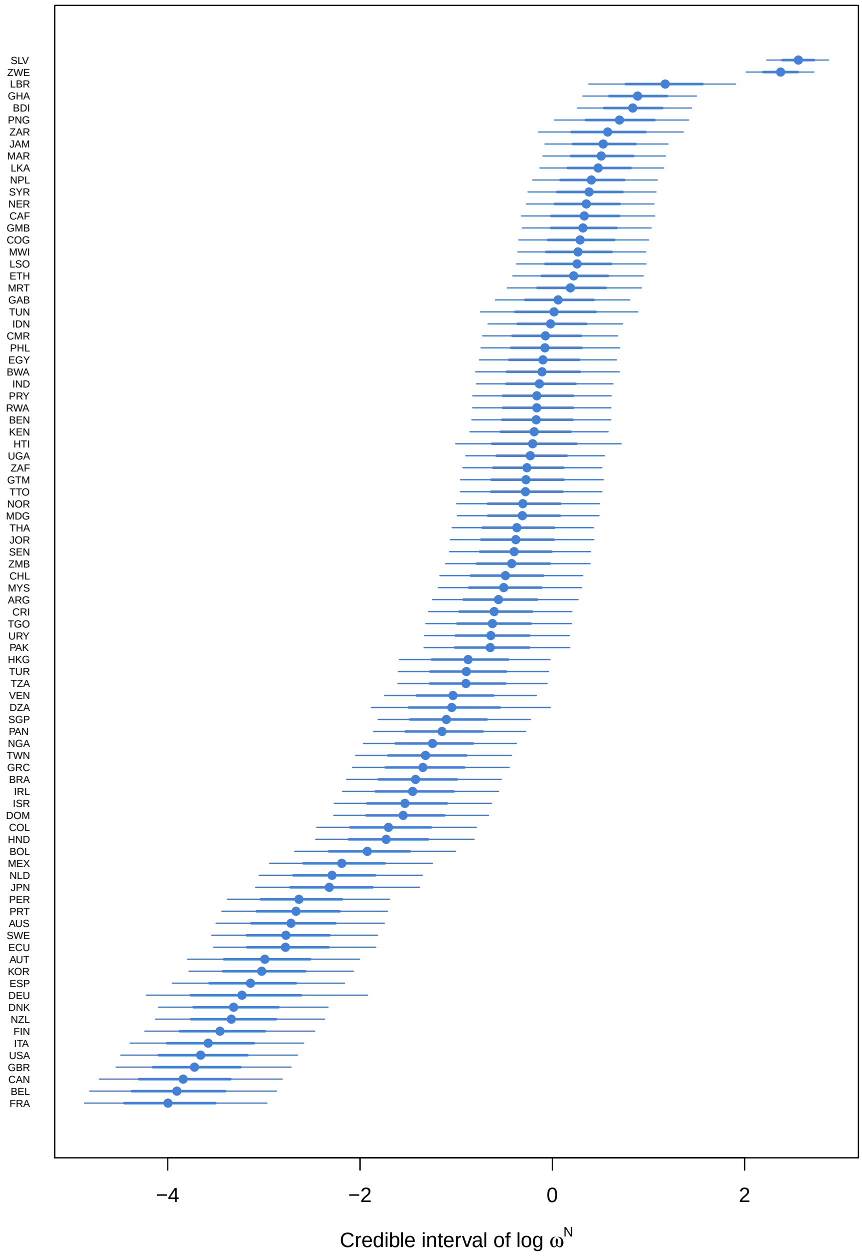

Figure 3: Relative variance of countries' measurement error

The figure shows posterior variance of relative variance of per country measurement error on a log-scale. The dot indicates the median of the distributions; the thick line shows a 68% credible interval and the thin line shows a 95% credible interval.

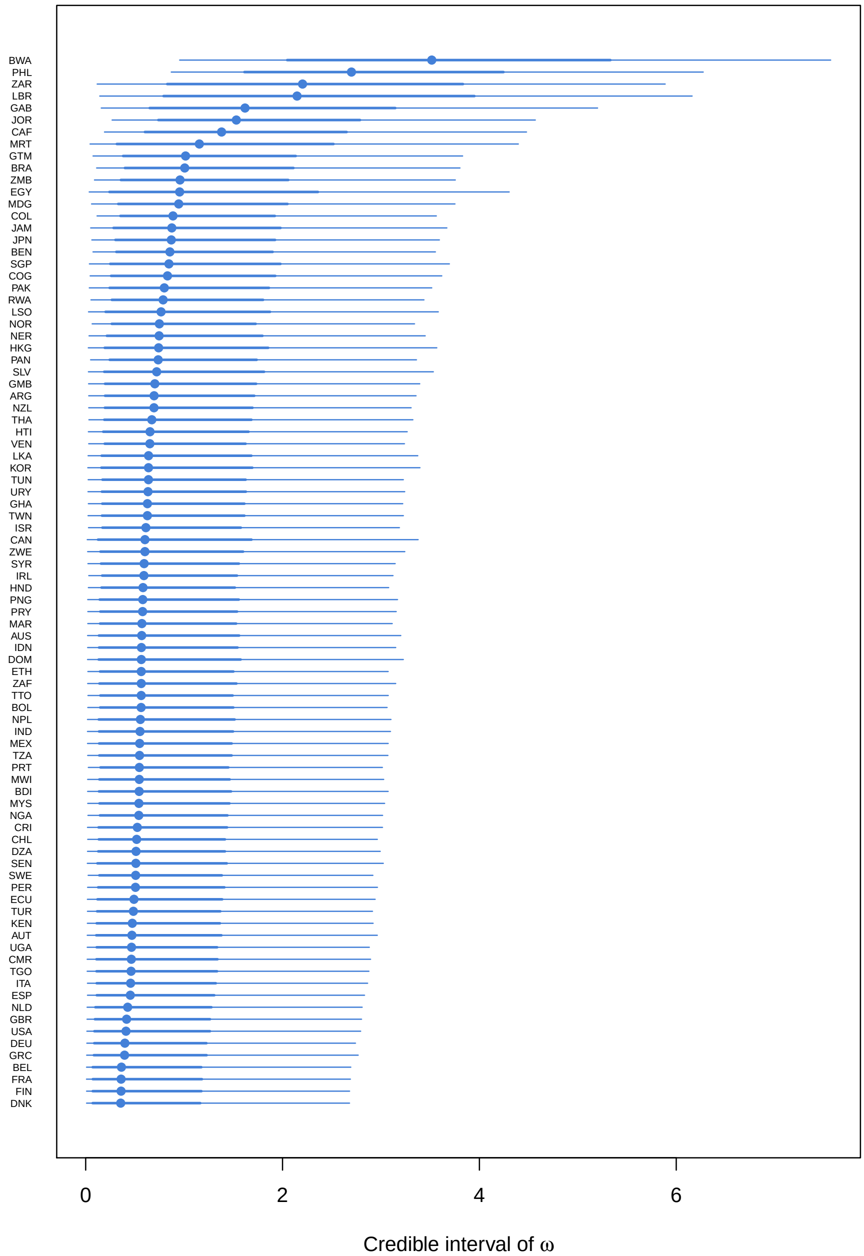

Figure 4: Relative variance of countries' measurement error

The figure shows posterior variance of relative variance of per country model error. The dot indicates the median of the distributions; the thick line shows a 68% credible interval and the thin line shows a 95% credible interval.

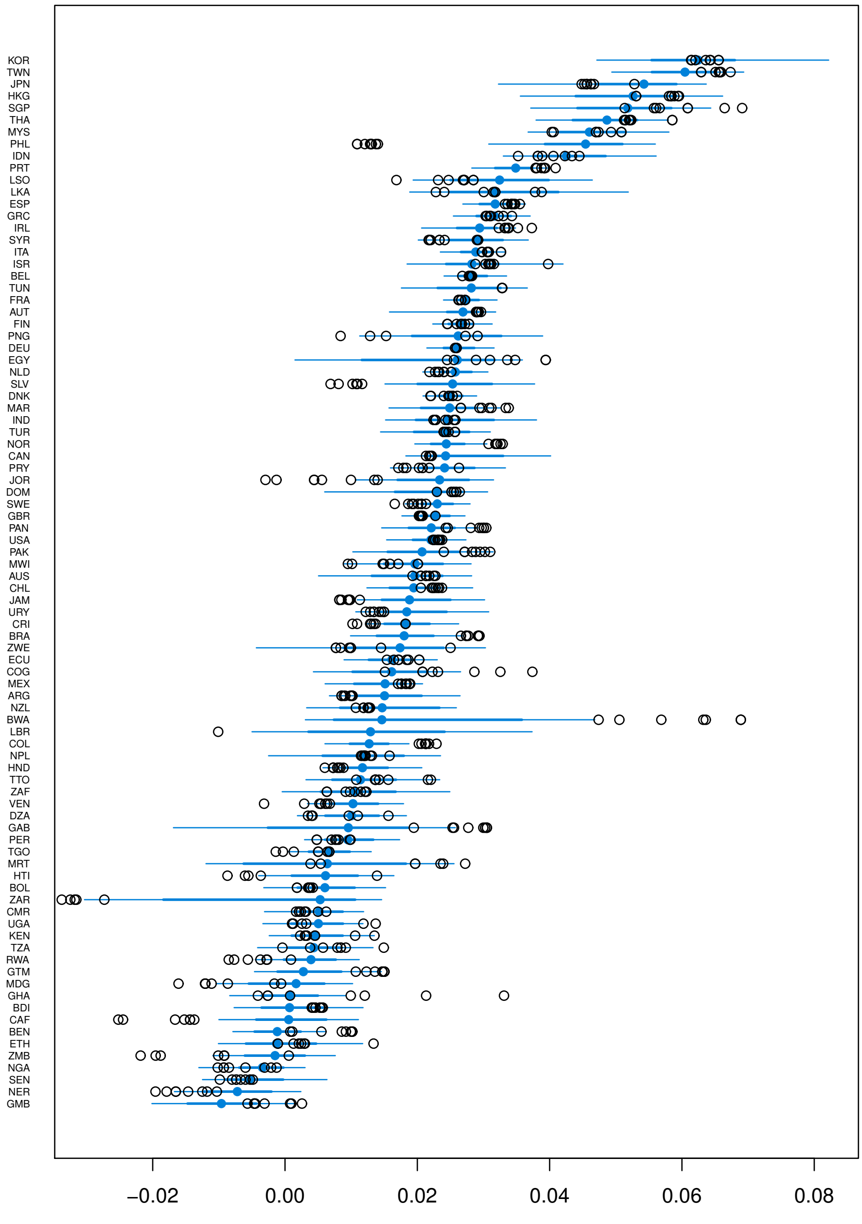

Figure 5: In-sample predicted, average annual growth rate and measures of realised growth rate

The figure shows posterior predicted growth per country. The dot indicates the median of the distributions; the thick line shows a 68% credible interval and the thin line shows a 95% credible interval. Circles indicate growth as measured in each PWT vintage from 6.0 through 8.1.

Table 1: Posterior inclusion probabilities in MA-models with differing data samples

	PWT 6.0	PWT 6.1	PWT 6.2	PWT 6.3	PWT 7.0	PWT 7.1	PWT 8.0	PWT 8.1
Absolute Latitude	3.5	14.9	5.1	9.4	4.5	3.7	10.2	11.9
Air Distance to Big Cities	2.9	1.9	3.8	2.7	2.1	2.4	3.6	3.0
Average Inflation 1960-90	1.9	1.7	1.8	2.0	1.8	1.6	1.9	2.3
British Colony Dummy	2.9	2.8	1.8	2.1	1.7	2.0	1.8	1.8
Capitalism	1.5	2.3	2.2	7.7	1.5	1.4	1.5	1.6
Civil Liberties	3.0	2.7	1.9	1.5	1.7	1.6	3.1	4.2
Colony Dummy	2.9	3.3	7.6	9.3	5.5	5.2	7.1	8.6
Defence Spending Share	2.1	6.1	2.4	2.2	3.3	2.3	1.9	1.9
East Asian Dummy	85.4	79.5	66.9	71.6	87.6	68.4	74.4	67.2
English Speaking Population	2.0	2.6	3.1	4.1	7.4	7.6	6.1	5.5
Ethnolinguistic Fractionalization	9.2	3.0	2.2	3.4	3.6	2.9	7.2	7.4
European Dummy	2.8	2.3	3.7	2.2	2.4	2.9	4.2	6.5
Fertility	3.0	2.7	3.4	2.7	2.7	2.9	2.7	3.0
Fraction Buddhist	9.8	11.4	22.5	3.6	6.4	18.8	28.7	34.0
Fraction Catholic	2.6	2.2	2.7	1.9	1.9	2.4	3.0	3.0
Fraction Confucian	18.4	33.1	48.1	30.1	24.2	44.6	31.3	39.1
Fraction GDP in Mining	9.8	42.6	1.8	38.7	53.9	45.8	79.0	79.4
Fraction Hindu	3.8	3.3	5.2	2.7	2.0	2.1	2.0	2.1
Fraction Muslim	10.1	31.8	19.4	13.8	7.2	13.2	13.1	12.8
Fraction of Tropical Area	52.1	20.6	17.5	8.1	7.9	5.1	44.5	38.5
Fraction Orthodox	1.5	1.4	1.5	1.5	1.4	1.5	1.6	1.6
Fraction Population In Tropics	5.7	15.7	15.8	40.2	69.3	67.0	12.5	15.2
Fraction Population Less than 15	4.3	5.2	2.9	2.8	3.2	3.8	3.8	5.9
Fraction Population Over 65	2.3	4.2	7.0	2.5	2.5	2.2	5.4	9.6
Fraction Protestant	4.3	5.0	7.5	4.7	7.7	10.7	9.5	7.0
Fraction Speaking Foreign Language	6.8	7.3	6.6	13.2	13.0	11.1	9.7	10.9
Fraction Spent in War 1960-90	1.5	1.4	1.4	1.3	1.6	1.4	2.7	2.5
Government Consumption Share	10.4	3.3	1.5	1.8	1.7	1.8	1.5	1.6
Government Share of GDP	5.8	1.9	1.5	1.5	1.6	1.5	1.5	1.6
Higher Education Enrolment	6.7	2.7	5.7	3.9	6.1	5.3	3.9	4.0
Hydrocarbon Deposits	2.1	1.7	2.1	1.9	2.3	2.2	1.8	1.9
Initial Income	55.3	97.0	99.2	100.0	94.6	93.5	71.6	68.5
Interior Density	1.4	1.7	2.2	2.6	1.4	1.5	1.5	1.5
Investment Price	66.2	76.8	1.8	33.3	23.0	26.0	2.3	2.2
Land Area	1.5	1.8	1.7	8.2	2.3	3.2	1.7	1.8
Land Area Near Navigable Water	1.9	1.9	1.6	3.0	1.9	2.4	1.9	1.8
Landlocked Country Dummy	1.8	2.4	2.7	4.0	5.7	6.4	3.3	3.1
Latin American Dummy	13.1	4.8	11.1	2.6	3.2	2.9	12.9	14.3
Life Expectancy	20.7	80.2	75.8	89.3	92.2	90.5	60.5	47.5
Malaria Prevalence	31.5	3.3	6.2	9.3	16.1	12.8	8.4	7.6
Nominal Government Share	3.1	1.6	7.3	1.6	1.8	1.6	2.0	2.1
Oil Producing Country Dummy	1.8	1.6	1.8	1.6	1.5	1.6	1.9	2.0
Openness 1965-74	7.7	25.6	25.9	41.4	6.4	6.7	12.2	15.1
Outward Orientation	2.7	2.4	2.2	2.4	1.9	1.9	1.5	1.5
Political Rights	4.8	1.8	2.1	1.8	1.9	2.2	2.5	2.3
Population Coastal Density	33.7	10.2	4.1	4.1	2.4	2.2	4.3	6.3
Population Density	5.9	11.8	1.6	53.2	44.5	52.7	20.2	20.6
Population Growth Rate 1960-90	1.9	2.5	3.5	1.9	2.2	2.4	2.2	2.2
Population in 1960	2.3	3.2	3.7	3.0	3.8	4.8	2.4	2.3
Primary Exports	5.2	6.3	7.1	5.1	5.0	5.2	9.4	5.1
Primary Schooling Enrollment	71.4	42.9	57.6	32.6	7.5	7.7	27.6	31.9
Public Education Spending Share	2.0	2.3	1.7	1.9	2.4	2.4	5.7	5.7
Public Investment Share	3.6	4.1	1.9	5.2	3.0	2.0	1.9	2.0
Real Exchange Rate Distortions	8.1	3.3	2.2	4.2	5.4	7.0	8.6	10.4
Religion Measure	1.7	1.9	2.3	2.6	2.8	2.8	6.5	5.8
Revolutions and Coups	2.8	1.5	1.5	1.6	1.8	1.9	7.0	8.1
Size of Economy	2.1	22.8	35.0	43.4	5.7	6.3	3.7	4.8
Socialist Dummy	1.8	2.7	1.6	1.7	1.4	1.4	1.6	1.6
Spanish Colony Dummy	13.0	4.3	4.8	2.2	2.8	2.6	6.2	6.2
Square of Inflation 1960-90	1.7	1.5	1.9	1.7	1.7	1.6	1.8	1.9
Sub-Saharan Africa Dummy	12.1	7.8	26.2	6.4	4.8	4.7	28.1	28.5
Terms of Trade Growth in 1960s	2.1	1.8	1.9	1.9	2.2	1.9	2.9	3.2
Terms of Trade Ranking	1.6	1.9	1.6	1.7	1.7	1.8	2.3	2.9
Timing of Independence	1.9	2.4	3.9	5.0	1.9	2.0	2.2	2.3
Tropical Climate Zone	1.7	1.8	1.7	1.6	1.5	1.6	1.9	2.0
War Participation 1960-90	1.5	1.8	1.7	1.4	1.4	1.6	1.5	1.5
Years Open 1950-94	11.5	7.7	9.0	14.6	40.7	48.8	6.4	8.8

The table shows the posterior inclusion probability of each covariate, where we fit the MA-model as explained in section 2.1 on each Penn World Table vintage separately. Columns indicate results for each PWT-vintage. The numbers in the table are reported as percentages.

Table 2: Estimated levels and uncertainty of Income and growth rates

Country	Income 1960	Income 1996	Growth 1960-1996	Country	Income 1960	Income 1996	Growth 1960-1996
ARG	8.332	8.662	0.009	KOR	6.878	9.178	0.064
	(0.076)	(0.064)	(0.003)		(0.025)	(0.022)	(0.001)
AUS	8.944	9.703	0.021	LBR	6.579	4.726	-0.051
	(0.029)	(0.025)	(0.001)		(0.646)	(0.136)	(0.018)
AUT	8.633	9.681	0.029	LKA	6.333	7.474	0.032
	(0.025)	(0.023)	(0.001)		(0.12)	(0.102)	(0.004)
BDI	5.636	5.81	0.005	LSO	5.719	6.647	0.026
	(0.144)	(0.122)	(0.005)		(0.107)	(0.089)	(0.004)
BEL	8.623	9.627	0.028	MAR	6.441	7.506	0.03
	(0.019)	(0.017)	(0.001)		(0.122)	(0.104)	(0.004)
BEN	6.279	6.566	0.008	MDG	6.566	6.237	-0.009
	(0.087)	(0.074)	(0.003)		(0.084)	(0.071)	(0.003)
BOL	7.329	7.438	0.003	MEX	7.912	8.573	0.018
	(0.038)	(0.033)	(0.001)		(0.037)	(0.032)	(0.001)
BRA	7.335	8.355	0.028	MRT	6.25	6.929	0.019
	(0.05)	(0.043)	(0.002)		(0.113)	(0.09)	(0.004)
BWA	5.94	8.2	0.063	MWI	5.436	6.026	0.016
	(0.096)	(0.076)	(0.003)		(0.109)	(0.093)	(0.004)
CAF	6.566	5.998	-0.016	MYS	6.982	8.683	0.047
	(0.118)	(0.094)	(0.004)		(0.076)	(0.064)	(0.003)
CAN	8.889	9.677	0.022	NER	6.442	5.902	-0.015
	(0.019)	(0.017)	(0.001)		(0.112)	(0.096)	(0.004)
CHL	7.826	8.636	0.022	NGA	6.728	6.472	-0.007
	(0.075)	(0.064)	(0.003)		(0.054)	(0.045)	(0.002)
CMR	6.815	6.927	0.003	NLD	8.854	9.714	0.024
	(0.09)	(0.078)	(0.003)		(0.034)	(0.031)	(0.001)
COG	6.433	7.331	0.025	NOR	8.905	10.057	0.032
	(0.111)	(0.095)	(0.004)		(0.086)	(0.073)	(0.003)
COL	7.423	8.186	0.021	NPL	6.023	6.469	0.012
	(0.043)	(0.037)	(0.002)		(0.122)	(0.103)	(0.004)
CRI	7.771	8.301	0.015	NZL	8.939	9.384	0.012
	(0.071)	(0.06)	(0.003)		(0.024)	(0.021)	(0.001)
DEU	8.768	9.67	0.025	PAK	6.106	7.124	0.028
	(0.032)	(0.02)	(0.001)		(0.07)	(0.061)	(0.003)
DNK	8.826	9.71	0.025	PAN	7.269	8.31	0.029
	(0.023)	(0.02)	(0.001)		(0.056)	(0.048)	(0.002)
DOM	7.115	8.027	0.025	PER	7.673	7.921	0.007
	(0.047)	(0.041)	(0.002)		(0.029)	(0.025)	(0.001)
DZA	7.713	7.979	0.007	PHL	6.99	7.434	0.012
	(0.075)	(0.064)	(0.003)		(0.089)	(0.075)	(0.003)
ECU	7.284	7.949	0.018	PNG	6.647	7.46	0.023
	(0.028)	(0.024)	(0.001)		(0.182)	(0.142)	(0.006)
EGY	6.465	7.686	0.034	PRT	7.771	9.161	0.039
	(0.095)	(0.08)	(0.003)		(0.029)	(0.025)	(0.001)
ESP	8.133	9.366	0.034	PRY	7.129	7.85	0.02
	(0.025)	(0.022)	(0.001)		(0.088)	(0.076)	(0.003)
ETH	5.676	5.751	0.002	RWA	6.25	6.095	-0.004
	(0.109)	(0.091)	(0.004)		(0.092)	(0.077)	(0.003)
FIN	8.515	9.472	0.027	SEN	6.939	6.687	-0.007
	(0.021)	(0.019)	(0.001)		(0.08)	(0.068)	(0.003)
FRA	8.615	9.577	0.027	SGP	7.662	9.794	0.059
	(0.018)	(0.017)	(0.001)		(0.058)	(0.049)	(0.002)
GAB	7.884	8.89	0.028	SLV	6.533	6.903	0.01
	(0.098)	(0.084)	(0.004)		(0.345)	(0.297)	(0.012)
GBR	8.815	9.565	0.021	SWE	8.87	9.575	0.02
	(0.02)	(0.018)	(0.001)		(0.027)	(0.024)	(0.001)
GHA	6.41	6.669	0.007	SYR	6.762	7.62	0.024
	(0.149)	(0.123)	(0.005)		(0.116)	(0.099)	(0.004)
GMB	6.626	6.529	-0.003	TGO	6.159	6.324	0.005
	(0.112)	(0.094)	(0.004)		(0.076)	(0.065)	(0.003)
GRC	8.013	9.156	0.032	THA	6.362	8.32	0.054
	(0.049)	(0.041)	(0.002)		(0.084)	(0.07)	(0.003)
GTM	7.27	7.781	0.014	TTO	8.196	8.73	0.015
	(0.084)	(0.072)	(0.003)		(0.083)	(0.072)	(0.003)
HKG	7.746	9.78	0.056	TUN	6.793	8.024	0.034
	(0.065)	(0.054)	(0.002)		(0.15)	(0.086)	(0.005)
HND	7.062	7.337	0.008	TUR	7.491	8.386	0.025
	(0.043)	(0.037)	(0.002)		(0.064)	(0.054)	(0.002)
HTI	6.83	6.827	0	TWN	7.01	9.339	0.065
	(0.115)	(0.091)	(0.004)		(0.049)	(0.042)	(0.002)
IDN	6.203	7.633	0.04	TZA	5.678	5.962	0.008
	(0.097)	(0.083)	(0.003)		(0.062)	(0.053)	(0.002)
IND	6.164	7.013	0.024	UGA	6.081	6.202	0.003
	(0.091)	(0.08)	(0.003)		(0.084)	(0.073)	(0.003)
IRL	8.305	9.522	0.034	URY	8.085	8.585	0.014
	(0.047)	(0.041)	(0.002)		(0.074)	(0.063)	(0.003)
ISR	8.235	9.396	0.032	USA	9.069	9.898	0.023
	(0.046)	(0.039)	(0.002)		(0.019)	(0.018)	(0.001)
ITA	8.454	9.581	0.031	VEN	8.398	8.567	0.005
	(0.02)	(0.018)	(0.001)		(0.058)	(0.049)	(0.002)
JAM	7.758	8.083	0.009	ZAF	7.945	8.268	0.009
	(0.127)	(0.106)	(0.005)		(0.083)	(0.072)	(0.003)
JOR	7.505	7.695	0.005	ZAR	6.425	5.377	-0.029
	(0.083)	(0.069)	(0.003)		(0.181)	(0.141)	(0.006)
JPN	8.023	9.719	0.047	ZMB	6.78	6.363	-0.012
	(0.033)	(0.028)	(0.001)		(0.078)	(0.066)	(0.003)
KEN	6.49	6.672	0.005	ZWE	6.73	7.123	0.011
	(0.087)	(0.076)	(0.003)		(0.309)	(0.268)	(0.011)

The table shows the posterior means of log income in 1960 and 1996, as well as the average, annual growth rate between the two periods. Numbers in parentheses show the standard deviation of the posterior densities. The estimates are made from the ME-model, as explained in section 2.1, where we allow for differing variance of measurement error across

Table 3: Posterior inclusion probabilities in the homoskedastic MEMA-model

ME-variance by country	Fixed, Equal	Fixed, Equal	Random, agnostic	Random, agnostic
ME-variance by vintages	Fixed, Equal	Random, agnostic	Fixed, Equal	Random, agnostic
Initial Income	91.1	95.4	96.9	97.4
Primary Schooling Enrollment	83.3	86.8	88.5	89.4
East Asian Dummy	77.1	76.4	80.0	81.2
Population Density	64.6	66.6	73.9	77.8
Investment Price	43.6	38.8	58.3	58.3
Fraction of Tropical Area	37.8	31.9	44.3	45.8
Primary Exports	36.4	48.6	41.6	41.1
Population Coastal Density	29.0	26.7	38.0	37.3
Air Distance to Big Cities	16.7	12.2	27.6	30.0
Sub-Saharan Africa Dummy	33.7	45.6	29.1	25.3
Life Expectancy	26.0	25.9	26.8	24.6
Fraction Muslim	22.8	22.0	23.6	22.7
Fraction GDP in Mining	22.7	22.3	23.9	22.5
Fraction Population In Tropics	9.3	9.7	17.0	19.9
Fraction Confucian	18.4	14.8	19.4	19.0
Political Rights	14.0	12.4	15.9	15.6
Years Open 1950-94	16.1	12.9	14.1	14.0
Fraction Buddhist	8.6	6.8	12.9	13.9
Ethnolinguistic Fractionalization	11.2	8.4	10.4	10.6
Openness 1965-74	5.0	4.8	8.7	10.5
Real Exchange Rate Distortions	10.0	7.7	10.5	10.0
Latin American Dummy	11.5	13.0	10.0	8.7
Malaria Prevalence	16.5	14.3	10.5	8.5
Fertility	11.0	10.9	8.3	7.5

The table shows the posterior inclusion probability of each covariate, for covariates where the posterior inclusion probability exceeds 7/67% in any model. The columns differ by the prior restrictions on the measurement error variance. The numbers in the table are reported as percentages. Table is sorted by the rightmost column.

Table 4: Posterior inclusion probabilities in the both homoskedastic and outlier robust MEMA-model

Outlier model	Homoscedastic	Dirichlet
ME-variance by country	Random, agnostic	Random, agnostic
ME-variance by vintages	Random, agnostic	Random, agnostic
Initial Income	97.4	99.8
East Asian Dummy	81.2	90.5
Primary Schooling Enrollment	89.4	90.3
Fraction of Tropical Area	45.8	51.1
Air Distance to Big Cities	30.0	43.6
Investment Price	58.3	34.4
Life Expectancy	24.6	34.2
Population Density	77.8	34.0
Sub-Saharan Africa Dummy	25.3	31.8
Population Coastal Density	37.3	26.6
Primary Exports	41.1	20.8
Latin American Dummy	8.7	20.1
Fraction Muslim	22.7	18.1
Malaria Prevalence	8.5	16.6
Fraction Confucian	19.0	14.8
Openness 1965-74	10.5	13.0
Landlocked Country Dummy	5.3	12.4
Outward Orientation	3.9	11.1
Fraction Population In Tropics	19.9	7.8
Years Open 1950-94	14.0	5.9
Political Rights	15.6	5.1
Fraction Buddhist	13.9	4.7
Ethnolinguistic Fractionalization	10.6	3.1
Fraction GDP in Mining	22.5	2.7

The table shows the posterior inclusion probability of each covariate, for covariates where the posterior inclusion probability exceeds 7/67% in any model. The columns differ by the prior restrictions on the measurement error variance. The numbers in the table are reported as percentages. Table is sorted by the rightmost column.

Table 5: Measurement error in the Penn World Tables

PWT 6.0	PWT 6.1	PWT 6.2	PWT 6.3	PWT 7.0	PWT 7.1	PWT 8.0	PWT 8.1
2.11	1.01	1.05	0.9	0.89	0.87	0.59	0.58
(0.22)	(0.14)	(0.14)	(0.12)	(0.13)	(0.13)	(0.11)	(0.1)

The table shows the relative measurement error of eight different Penn World Tables, given by the posterior densities of ω_v as explained in section 2 from the Dirichlet-robust MEMA-model. The first row of numbers shows the posterior mean, and numbers in parentheses gives the standard deviation of the posterior distribution.

Table 6: Expanded results for the robust MEMA-model

	Cond. mean	Std.d.	Sign certainty	PIP
Initial Income	-1.1E-02	2.7E-03	100.0	99.8
East Asian Dummy	2.2E-02	5.5E-03	100.0	90.5
Primary Schooling Enrollment	3.4E-02	8.6E-03	100.0	90.3
Fraction of Tropical Area	-1.3E-02	3.6E-03	99.7	51.1
Air Distance to Big Cities	-1.7E-06	5.5E-07	99.7	43.6
Investment Price	-7.8E-05	2.7E-05	99.5	34.4
Life Expectancy	7.1E-04	2.9E-04	99.6	34.2
Population Density	2.1E-05	7.4E-06	99.3	34.0
Sub-Saharan Africa Dummy	-1.5E-02	6.4E-03	99.5	31.8
Population Coastal Density	6.5E-06	2.3E-06	99.3	26.6
Primary Exports	-1.5E-02	5.8E-03	99.0	20.8
Latin American Dummy	-1.2E-02	5.1E-03	97.5	20.1
Fraction Muslim	1.4E-02	5.7E-03	98.8	18.1
Malaria Prevalence	-1.3E-02	5.7E-03	98.0	16.6
Fraction Confucian	4.7E-02	2.2E-02	98.3	14.8
Openness 1965-74	8.7E-03	4.1E-03	98.2	13.0
Landlocked Country Dummy	-7.2E-03	3.3E-03	98.5	12.4
Outward Orientation	-4.4E-03	2.1E-03	98.2	11.1

The table shows detailed results for the Dirichlet robust MEMA-model, as explained in section 2. The table shows results for covariates with a posterior inclusion probability exceeding 7/67%. The table is sorted by the posterior inclusion probability. The first and second columns shows the posterior mean and standard deviation of the coefficient, conditional on being included in a model. The third column shows the sign certainty - i.e. the cumulative posterior density where the sign of the coefficient is the same as the sign of the posterior mean. The fourth column shows the posterior inclusion probability.

Table 7: Posterior inclusion probabilities in the outlier robust MEMA-model

Outlier model	Dirichlet	Binary	Geweke
ME-variance by country	Random, agnostic	Random, agnostic	Random, agnostic
ME-variance by vintages	Random, agnostic	Random, agnostic	Random, agnostic
Initial Income	99.8	99.9	99.8
East Asian Dummy	90.5	97.9	94.3
Primary Schooling Enrollment	90.3	82.8	88.7
Fraction of Tropical Area	51.1	29.8	41.6
Air Distance to Big Cities	43.6	36.0	43.0
Investment Price	34.4	17.6	23.0
Life Expectancy	34.2	43.6	36.1
Population Density	34.0	17.0	22.3
Sub-Saharan Africa Dummy	31.8	41.4	39.3
Population Coastal Density	26.6	18.5	22.7
Primary Exports	20.8	27.1	24.1
Latin American Dummy	20.1	29.4	24.9
Fraction Muslim	18.1	21.9	21.6
Malaria Prevalence	16.6	37.0	27.6
Fraction Confucian	14.8	12.8	15.9
Openness 1965-74	13.0	12.3	17.0
Landlocked Country Dummy	12.4	38.3	28.8
Outward Orientation	11.1	10.7	13.9
European Dummy	9.6	20.3	16.7
Fraction Speaking Foreign Language	6.9	12.7	10.6

The table shows the posterior inclusion probability of each covariate, for covariates where the posterior inclusion probability exceeds 7/67% in any model. The columns differ by the prior restrictions on the measurement error variance. The numbers in the table are reported as percentages. Table is sorted by the rightmost column.

Table 8: Posterior inclusion probabilities in including PWT 9.0

ME-variance by country	Random, unrestricted	Random, unrestricted	Random, unrestricted
PWT-vintages	6.0 - 8.1	6.0 - 9.0	6.0 - 9.0
Outlier model	Dirichlet	Dirichlet	Dirichlet
Vintage ME-variance	Unrestricted	Newer is better	Unrestricted
Initial Income	99.8	99.6	99.8
East Asian Dummy	90.5	90.1	90.5
Primary Schooling Enrollment	90.3	80.8	87.6
Fraction of Tropical Area	51.1	44.6	49.4
Air Distance to Big Cities	43.6	38.5	42.7
Investment Price	34.4	34.4	35.0
Life Expectancy	34.2	38.9	34.8
Population Density	34.0	28.4	31.4
Sub-Saharan Africa Dummy	31.8	29.4	30.5
Population Coastal Density	26.6	21.5	24.5
Primary Exports	20.8	13.7	17.0
Latin American Dummy	20.1	20.7	20.1
Fraction Muslim	18.1	16.3	18.5
Malaria Prevalence	16.6	27.0	19.1
Fraction Confucian	14.8	14.2	14.4
Openness 1965-74	13.0	11.7	13.1
Landlocked Country Dummy	12.4	27.5	17.0
Outward Orientation	11.1	7.8	9.1
European Dummy	9.6	11.0	10.2

The table shows the posterior inclusion probability of each covariate, for covariates where the posterior inclusion probability exceeds 7/67% in any model. The columns differ by the prior restrictions on the measurement error variance. The numbers in the table are reported as percentages. Table is sorted by the rightmost column.

Table 9: Relative Measurement error in the Penn World Tables

PWT vintage	Unrestricted	Newer is better
PWT 6.0	2.02	2.12
PWT 6.1	0.93	1.12
PWT 6.2	0.95	1.01
PWT 6.3	0.82	0.93
PWT 7.0	0.94	0.88
PWT 7.1	0.92	0.83
PWT 8.0	0.61	0.74
PWT 8.1	0.54	0.70
PWT 9.0	1.28	0.67

The table shows the posterior mean of relative measurement error of nine different Penn World Tables, given by the posterior densities of ω_v as explained in section 4.2

Table 10: Posterior inclusion probabilities with random model size and g-shrinkage

ME-variance by country	Random, agnostic	Random, agnostic
ME-variance by vintages	Random, agnostic	Random, agnostic
Outlier model	Dirichlet	Dirichlet
PWT-vintages	6.0 - 8.1	6.0 - 8.1
Prior model size	Fixed	Random
Prior Zellner g-shrinkage	Fixed	Random
Initial Income	99.8	99.5
East Asian Dummy	90.5	90.5
Primary Schooling Enrollment	90.3	95.0
Fraction of Tropical Area	51.1	58.6
Air Distance to Big Cities	43.6	53.7
Investment Price	34.4	42.3
Life Expectancy	34.2	34.9
Population Density	34.0	55.7
Sub-Saharan Africa Dummy	31.8	29.9
Population Coastal Density	26.6	37.1
Primary Exports	20.8	25.1
Latin American Dummy	20.1	17.1
Fraction Muslim	18.1	23.9
Malaria Prevalence	16.6	13.0
Fraction Confucian	14.8	19.2
Openness 1965-74	13.0	19.0
Landlocked Country Dummy	12.4	17.8
Outward Orientation	11.1	20.6
Real Exchange Rate Distortions	10.0	17.1
European Dummy	9.6	10.6
Fraction Population In Tropics	7.8	11.9
Fraction Speaking Foreign Language	6.9	11.0
Fertility	6.2	10.9
Hydrocarbon Deposits	3.8	12.9

The table shows the posterior inclusion probability of each covariate, for covariates where the posterior inclusion probability exceeds 7/67% in any model. The columns differ by the prior restrictions on the measurement error variance. The numbers in the table are reported as percentages. Table is sorted by the rightmost column.