The Impact of Data Revisions on the Robustness of Growth Determinants - A Note on 'Determinants of Economic Growth. Will Data Tell?'*

Martin Feldkircher†
Oesterreichische Nationalbank

Stefan Zeugner‡
ECARES, Université libre de Bruxelles

22 February 2012

Abstract

Ciccone and Jarociński (2010) show that inference in Bayesian Model Averaging (BMA) can be highly sensitive to small data perturbations. In particular they demonstrate that the importance attributed to potential growth determinants varies tremendously over different revisions of international income data. They conclude that 'agnostic' priors appear too sensitive for this strand of growth empirics. In response, we show that the found instability owes much to a specific BMA set-up: First, comparing the *same* countries over data revisions improves robustness. Second, much of the remaining variation can be reduced by applying an evenly 'agnostic', but flexible prior.

Keywords: Bayesian model averaging, Growth determinants, Zellner's g prior, Model uncertainty.
JEL Classifications: C11, C15, E01, O47.

*The opinions in this paper are those of the authors and do not necessarily coincide with those of the Oesterreichische Nationalbank. We would like to thank Antonio Ciccone, Jesús Crespo Cuaresma, Gernot Doppelhofer, Marek Jarociński, Aart Kraay, Tomáš Slačík and two anonymous referees for helpful comments.

†Oesterreichische Nationalbank, Otto-Wagner-Platz 3, 1090 Vienna, Austria. E-mail: martin.feldkircher@oenb.at.

‡Corresponding author – address: ULB, CP 114, 50 Avenue Roosevelt, 1050 Brussels, Belgium.
E-mail: stefan.zeugner@ulb.ac.be. Vox: +32 2 2963534. URL: http://www.zeugner.eu

1 Introduction

Ciccone and Jarociński (2010) present an intriguing paper that points to severe weaknesses of Bayesian Model Averaging (BMA). In particular, they criticize the appealing use of non-informative ('agnostic') priors as being plagued by robustness problems. Based on the ubiquitous growth dataset by Sala-i-Martin et al. (2004), the authors analyze the impact of data revisions employing three different vintages of international income data provided by the Penn World Tables (PWT). In this manner, Ciccone and Jarociński (2010) show that small data perturbations can lead to striking differences in posterior inclusion probabilities (PIP) – i.e. the importance of a covariate in explaining the data. The variable 'investment price', for instance, seems very important for growth with a PIP of 0.98 based on PWT 6.1 data, but it exhibits a mere 0.02 under PWT 6.2. Further variables with a worrisome degree of PIP variation reported in Ciccone and Jarociński (2010) are the fraction of Confucian population, population density in 1960, population density in coastal areas in 1960, the fertility rate in 1960, a dummy for east Asia, a dummy for African countries and the fraction of tropical area per country. The conclusions made in Ciccone and Jarociński (2010, p.222) are thus rather pessimistic in that *"margins of error in international income estimates are too large for agnostic growth empirics."*

In a replication exercise we confirm the view that results from 'agnostic' BMA have to be interpreted with caution. In the following sections, however, we point to two caveats to the above findings: First, a considerable share of the PIP variation found by the authors is due to changing sample size and composition over PWT revisions. Secondly, the remaining variation can be attributed to the use of a 'default' prior framework that embodies overly confident prior beliefs – this second caveat constitutes the focus of this note. In response to Ciccone and Jarociński (2010), we show that conditioning on the same country set through all revisions and employing an alternative, flexible prior greatly reduces PIP instability.

2 Robustness under Data Revisions

Ciccone and Jarociński (2010) use three different PWT revisions in order to represent income, and cross-check the results with World Bank data. The conditional convergence growth regression employed is of the following form:

$$\Delta y^j = \alpha + \gamma y_0^j + \vec{\beta}_s X_s + \varepsilon, \tag{1}$$

with Δy^j denoting the average annual growth of income per capita over the period from 1960 to 1996 for N countries, α the intercept term and $\mathbf{X}_s = (\mathbf{x}_1 \dots \mathbf{x}_s)$ a matrix whose columns are stacked data for s explanatory variables and ε an error term. Initial income (y_0^j) and income growth (Δy^j) are the only variables that change with revisions. The remaining potential growth determinants grouped in $\mathbf{X}_s$ are the ones originally put forward in Sala-i-Martin et al. (2004) and employed in Ciccone and Jarociński (2010). These 66 variables comprise measures for factor accumulation and convergence (as implied by the Solow growth model), human capital variables, variables measuring political stability and socio-geographical variables. The estimation is carried out for each of the three considered PWT revisions, indexed by $j \in \{\text{PWT 6.0}, \text{PWT 6.1 and PWT 6.2}\}$. In what follows, we denote by σ^2 the variance, N the total number of observations, R_s^2 the OLS R-squared of model s and K the total number of available covariates.

Turning to the econometric framework, Ciccone and Jarociński (2010) use two approaches from the model averaging literature: The 'BACE' methodology proposed by Sala-i-Martin et al. (2004) as well as the popular 'benchmark' BMA type employed by Fernández et al. (2001). Since both yield broadly similar results in the empirical application by Ciccone and Jarociński (2010, p.223), we follow the purely Bayesian approach akin to Fernández et al. (2001). This framework relies on Zellner's g prior for the coefficient vector $\vec{\theta}_s|\sigma^2 \sim N\left(\vec{0}_s, \sigma^2 g[X_s' X_s]^{-1}\right)$ with $\vec{\theta}_s \subset (\beta_s, \gamma)$.

3 The Impact of Changing Samples – Or How much Variation can be Attributed to Africa?

A special feature of PWT data is that with each vintage the coverage of countries is likely to change. For the data employed by Ciccone and Jarociński (2010), the number of countries ranges from 88 (PWT 6.0) to 79 (PWT 6.2). Ciccone and Jarociński (2010, p.230) argue in favor of keeping sample sizes to their possible maximum, as it might be uncertain which countries may be included in future revisions. While this might have its virtues from the perspective of a forward-looking policy-maker, it has the potential to blur the conclusions on the robustness of the method (BMA).

A closer look at the PWT samples reveals that the countries dropped or added between revisions are mostly African.[1] In terms of growth, these were either very successful or unsuccessful countries with respect to their regional peers. As a consequence, and recently demonstrated in Masanjala and Papageorgiou (2008) and Crespo Cuaresma (2010), growth determinants in Africa systematically differ in various instances and it is difficult to consider this particular set of countries as a randomly chosen subsample of the data.

Varying Sample	$g = K^2$	$g = N$	hyper-g
Overall Max / Min Ratio	4.8580	2.0421	1.5224
PWT 6.0 vs. PWT 6.1	2.1179	1.4123	1.2114
PWT 6.0 vs. PWT 6.2	2.5888	1.6562	1.2778
PWT 6.1 vs. PWT 6.2	3.9466	1.6634	1.3518
Common Sample	$g = K^2*$	$g = N*$	hyper-$g*$
Overall Max / Min Ratio	2.4217	1.7344	1.4013
PWT 6.0 vs. PWT 6.1	1.9675	1.5866	1.3260
PWT 6.0 vs. PWT 6.2	1.8357	1.5434	1.3085
PWT 6.1 vs. PWT 6.2	1.5880	1.2470	1.1451

Table 1: Average PIP Max/Min ratios: for each revision pair, the figures above display the mean of the ratio maximum vs. minimum PIP per variable. The asterisk denotes the use of the common country set.

In order to identify the sources of PIP variability, we first replicate[2] the results of Ciccone and Jarociński (2010), employing the changing PWT samples and same prior set-up: a uniform prior[3] on the model space, and the 'benchmark' coefficient prior of Fernández et al. (2001), that is $g = \max(K^2, N)$. As a measure of PIP instability we calculate the max/min ratio of the posterior inclusion probability over the three data vintages per variable and report its average (over covariates) as the *overall max/min ratio*. The results are summarized in the first column of Table 1.

The first column of the top panel ('varying sample') shows by far the greatest overall max/min ratio

[1] Precisely, PWT 6.1 includes the countries of PWT 6.2 plus Botswana, Central African Republic, Mauritania, DR Congo (Zaire) and Papua New Guinea. PWT 6.0 includes the PWT 6.1 countries plus Liberia, Tunisia, West Germany and Haiti.

[2] All computations were carried out with the `R` package `BMS`. The data and detailed instructions for replication are available at `http://bms.zeugner.eu/datarev.php`.

[3] Note that we have omitted the model prior from the discussion since its impact seems to be rather limited as compared to the importance of the g-prior for this application (Ciccone and Jarociński, 2010, p.226).

(and thus PIP variation), implying that the modeling strategy of Ciccone and Jarociński (2010) is indeed the one that is most severely plagued by instability. Furthermore, note that the amount of PIP variation reported in the first column is outstanding for all PWT vintages. That is, the strong variation is not driven by a single PWT vintage, but is a characteristic of the empirical framework employed by Ciccone and Jarociński (2010).

In order to quantify the role played by the changing sample, we replicate the estimations in Ciccone and Jarociński (2010) for the three data revisions but condition on the *same* set of countries. The first column, bottom panel ('common sample') illustrates that conditioning on the same observations per PWT vintage reduces PIP instability by a half. This suggests that the additional countries from PWT 6.0 and PWT 6.1 can be regarded as innovations (outliers) with the potential to change results.

4 The use of Default Priors – Or is fixing Prior Beliefs always advisable?

As alluded to before, the BMA framework used by Ciccone and Jarociński (2010) calls for eliciting the key hyperparameter Zellner's g. This parameter g reflects the strength of the researcher's prior guess on slope coefficients. Small values correspond to stronger beliefs that the regression slopes are zero (i.e. the prior is tightened).[4] By construction, the parameter g directly affects the posterior model probability (PMP) – the weight attributed to model M_s – and thus final inference:

$$p(M_s|y^j, X) \propto \left(1 - \frac{g}{1+g}\right)^{\frac{k_s}{2}} \left(1 - \frac{g}{1+g} R_s^2\right)^{-\frac{N-1}{2}} \tag{2}$$

The choice of g can be crucial for the results: Consider a relatively large g (implying a large shrinkage factor $\frac{g}{1+g}$). As is evident from equation (2), this will not only favor parsimonious models but also amplify any – potentially very small – differences in R_s^2. The resulting distribution of posterior model probabilities will therefore be highly concentrated on the few parsimonious models with the very highest R_s^2.

Figure 1 illustrates this effect by plotting cumulative PMPs based on different settings for g. The chart demonstrates that for PWT data, larger g attributes more weight to the first-best model relative to the remaining ones. If the data is dominated by noise, this *supermodel effect* (Feldkircher and Zeugner, 2009) will skew posterior mass to concentrate on a few 'supermodels'. Consequently it will skew the distribution of PIPs and thus amplify variations that may be due to noise. In contrast, employing a smaller g will limit such PIP variations, as exemplified by the second column of Table 1 (which uses $g = N$).

A remedy for the supermodel effect: Several 'default' mechanisms have been proposed to elicit g, the most prominent being the 'benchmark prior' by Fernández et al. (2001), who recommend $g = \max(N, K^2)$. Still, any of these fixed mechanisms risks to set g too small or too large with respect to the noise component in the data.

In response, Feldkircher and Zeugner (2009) propose to forgo fixed g-priors outright and to place a prior distribution on the parameter instead:[5] We follow the fairly general *hyper-g* prior approach

[4] Note that in principle, it would be possible to elicit individual priors for the slopes, but we follow Ciccone and Jarociński (2010) and the bulk of the literature in centering all coefficient priors at zero.

[5] Note that there are also other flexible prior frameworks, such as the Empirical Bayes (EB) approach. For the sake of brevity, we have omitted the EB results in this note, since they are very similar to those under the hyper-g prior (Feldkircher and Zeugner, 2009, Appendix).

put forward by Liang et al. (2008), which belongs to the family of mixtures of g-priors. With a (hyper-)prior on g, the prior on the coefficient vector can be interpreted as a mixture of normal distributions with fatter tails (Ley and Steel, 2011). Technically, employing the hyper-g prior boils down to placing a Beta prior on the shrinkage factor $g/(1+g) \sim Beta(1, \frac{a}{2} - 1)$. Choosing the hyperparameter a accordingly allows for formulating prior beliefs on g that match popular fixed-g settings.[6]

As opposed to fixing beliefs a priori, this hierarchical approach updates the prior beliefs according to the data. In that sense employing the hyper-g prior is less prone to misalignments of data and beliefs. Furthermore, the hyper-g prior is more flexible as it allows for *model-specific* g_s values (and shrinkage factors) that adjust to data quality:[7] If the data is characterized by minor noise, then posterior mass will concentrate on the true model(s) – even more than under fixed settings with large g. Conversely, if noise dominates the data, then posterior statistics will decrease g and PMPs under the hyper-g prior will be distributed more evenly. Even in such a case, BMA under a large fixed g would always come up with clear-cut results (a single model and a few covariates that obtain overwhelming support), not taking into account the small degree of data quality. The hyper-g framework, in contrast, will then point to inconclusiveness mirrored in evenly spread PMPs and PIPs.

The supermodel effect and PWT revisions: In their application, Ciccone and Jarociński (2010) follow the 'benchmark' recommendation by Fernández et al. (2001) and set $g = \max(N, K^2) = (66+1)^2$. Note that this renders the shrinkage factor very close to unity ($g/(1+g) \approx 0.9998$). We can thus expect PMPs to become highly concentrated, and small differences in R_s^2 to be translated into large differences in PMPs and PIPs. This large shrinkage factor treats the data as very informative with respect to the prior. In fact, however, the quality of the PWT data sets turns out to be rather poor. Using flexible g priors such as the hyper-g prior implies far smaller average shrinkage factors of around 0.95 (Table 2). This suggests that in order to avoid the supermodel effect, a fixed-g framework on PWT data should rather elicit $g \approx 19$. That said, any fixed g will always lack the flexibility of the hyper-g prior in adjusting to data quality.

Varying Sample	$g = K^2$	$g = N$	hyper-g
PWT 6.0	0.9998	0.9888	0.9103
PWT 6.1	0.9998	0.9882	0.9293
PWT 6.2	0.9998	0.9875	0.9454
Common Sample	$g = K^2*$	$g = N*$	hyper-$g*$
PWT 6.0	0.9998	0.9875	0.9361
PWT 6.1	0.9998	0.9875	0.9476
PWT 6.2	0.9998	0.9875	0.9454

Table 2: Average shrinkage factors for three PWT revisions under the benchmark case $g = 67^2$, $g = 79$ and the hyper-g prior (with prior expected shrinkage factor $\frac{g}{1+g} = 67^2/(1+67^2)$). The asterisk denotes the use of the common country set. All results are based on a Metropolis-Hastings model sampler as in Fernández et al. (2001), with 80 million iterations after 20 million burn-ins.

In order to examine the role of the hyper-g prior in the context of growth determinants and PWT revisions, we provide the full results in Table 5. The hyper-g prior results greatly reduce

[6] For our purpose, we defined the parameter a such that prior expected shrinkage matches the $g/(1+g)$ used in Ciccone and Jarociński (2010). Experimenting with other parameters a produces results that are very close to the ones reported here.

[7] In this note we refer to data quality by the degree of variation of the dependent variable explained by the data at hand.

the variability of PIPs over revisions for the 8 covariates mentioned before. In particular, the investment price and population density variables as well as the tropical area dummy do not appear to matter for growth under either revision while the African dummy seems robust. As a further observation one might stress that under the hyper-g prior, the PIPs are much larger on average – this results from the data inducing lower posterior shrinkage factors and therefore emphasizing less parsimonious models than under the 'benchmark' priors. The smaller model size penalty results into much larger posterior model size (≈ 25 vs. $7-9$ under the Benchmark case). As a consequence, comparing *absolute* PIPs across different priors is prone to misleading conclusions since the sum of PIPs is by construction equal to the respective posterior model size – which in turn strongly depends on the value of g.[8]

Consider Figure 1 to see how the hyper-g prior induces data-dependent shrinkage. The top panel plots the cumulative PMP based on the PWT 6.1 vintage (varying sample composition). As expected, the benchmark prior ($g = K^2$) results in the by far most concentrated PMP distribution, whereas the hyper-g prior spreads PMPs most evenly. The Figure also shows that the shrinkage factor induced by the fixed $g = N$ setting is still too high to match the PMP distribution of the hyper-g prior. The bottom panel of Figure 1 provides the same plot for the sample of common countries, which was already shown to induce less PIP variation. Comparing the two figures reveals that the PMP distribution under neither fixed g prior responds substantially to the change in samples. The only line that adjusts considerably is that of the hyper-g prior. This can be best seen by considering the differences in PMP distributions of the hyper-g prior and the fixed UIP ($g = N$) prior: while this difference is pronounced under the varying sample framework, it is considerably smaller for the common sample data. The smaller noise component inherent in the common sample induces the hyper-g prior to put more weight to the data relative to the prior and thus to skew PMP mass. The fixed prior settings, in contrast, do not allow for adjusting the PMP distribution to a change in the noise component.

To assess the degree of PIP variation for the different prior set-ups, we consider the overall max/min figures provided in Table 1. For illustration, compare first the results for the original framework with varying sample size (Table 1, top panel). A decrease in g (such as $g = N$) already lowers PIP variation substantially, which demonstrates the supermodel effect. However, the most remarkable reduction is achieved by the hyper-g prior, with a drop in variation of close to 70% compared to the benchmark framework (Table 1, top panel, columns 1 and 3). The results for the common samples (Table 1, bottom panel) yield a similar picture: The hyper-g prior greatly reduces PIP instability in comparison to the other fixed prior settings. With respect to the prior set-up used in Ciccone and Jarociński (2010) (Table 1, bottom panel, column 1), PIP variation is reduced by close to 50% – from already decreased levels due to conditioning on the same countries. Finally, Table 1 reveals that the minimal overall max/min ratio is achieved by employing the hyper-g prior coupled with using identical countries over PWT vintages. That said, note that PIP variation is always smaller with the hyper-g framework then under the popular fixed g settings. Thus, even when – for data availability reasons – sample composition changes with revisions, we strongly recommend the use of a hyper-g prior.

There is one important caveat, though: the quite low shrinkage factor induced by hyper-g not only decreases PIP variation over data revisions, but also over covariates for a given revision. Consequently, this implies that there are less covariates that could be identified as 'considerably more important' than others. As has been noted above, this is a direct result from posterior mass being spread out more evenly over models due to an important noise component in the data. Yet this trait may be desirable, as noisy data should not lead to the striking conclusions from the 'agnostic' approach criticized by Ciccone and Jarociński (2010).

[8] For instance, when $g \to \infty$, all PIPs will tend to zero, thus their absolute differences will vanish.

5 Conclusion

This note addresses an important issue raised by Ciccone and Jarociński (2010): inference of (agnostic) BMA applied to growth empirics is not robust under small perturbations of international income data. Our response demonstrates that such instability is partly due to the overconfident 'default' g-prior framework employed by the authors. Instead, we propose to rely on the hyper-g prior. While 'agnostic' in the sense of Ciccone and Jarociński (2010), it adjusts to data quality and induces smaller shrinkage factors according to the data's considerable noise component. This in turn renders BMA results considerably more stable over different revisions of PWT growth data. However, reflecting poor data quality, results under hyper-g discriminate far less among covariates. The empirical findings under a flexible prior may thus be characterized as 'robust ambiguity', limiting statements about the importance of growth determinants to a quite small subset of covariates.

References

Ciccone A, Jarociński M. 2010. Determinants of Economic Growth: Will Data Tell? *American Economic Journal: Macroeconomics* **2**: 222–46.

Crespo Cuaresma J. 2010. How different is Africa? A comment on Masanjala and Papageorgiou (2008). *Journal of Applied Econometrics* **26**: 1048–1050.

Feldkircher M, Zeugner S. 2009. Benchmark Priors Revisited: On Adaptive Shrinkage and the Supermodel Effect in Bayesian Model Averaging. *IMF Working Paper* **WP/09/202**.

Fernández C, Ley E, Steel MF. 2001. Benchmark Priors for Bayesian Model Averaging. *Journal of Econometrics* **100**: 381–427.

Ley E, Steel MF. 2011. Mixtures of g-priors for bayesian model averaging with economic applications. Policy Research Working Paper 5732, World Bank.

Liang F, Paulo R, Molina G, Clyde MA, Berger JO. 2008. Mixtures of g Priors for Bayesian Variable Selection. *Journal of the American Statistical Association* **103**: 410–423.

Masanjala WH, Papageorgiou C. 2008. Rough and Lonely Road to Prosperity: A Reexamination of the Sources of Growth in Africa Using Bayesian Model Averaging. *Journal of Applied Econometrics* **23**: 671–682.

Sala-i-Martin X, Doppelhofer G, Miller RI. 2004. Determinants of Long-Term Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach. *American Economic Review* **94**: 813–835.

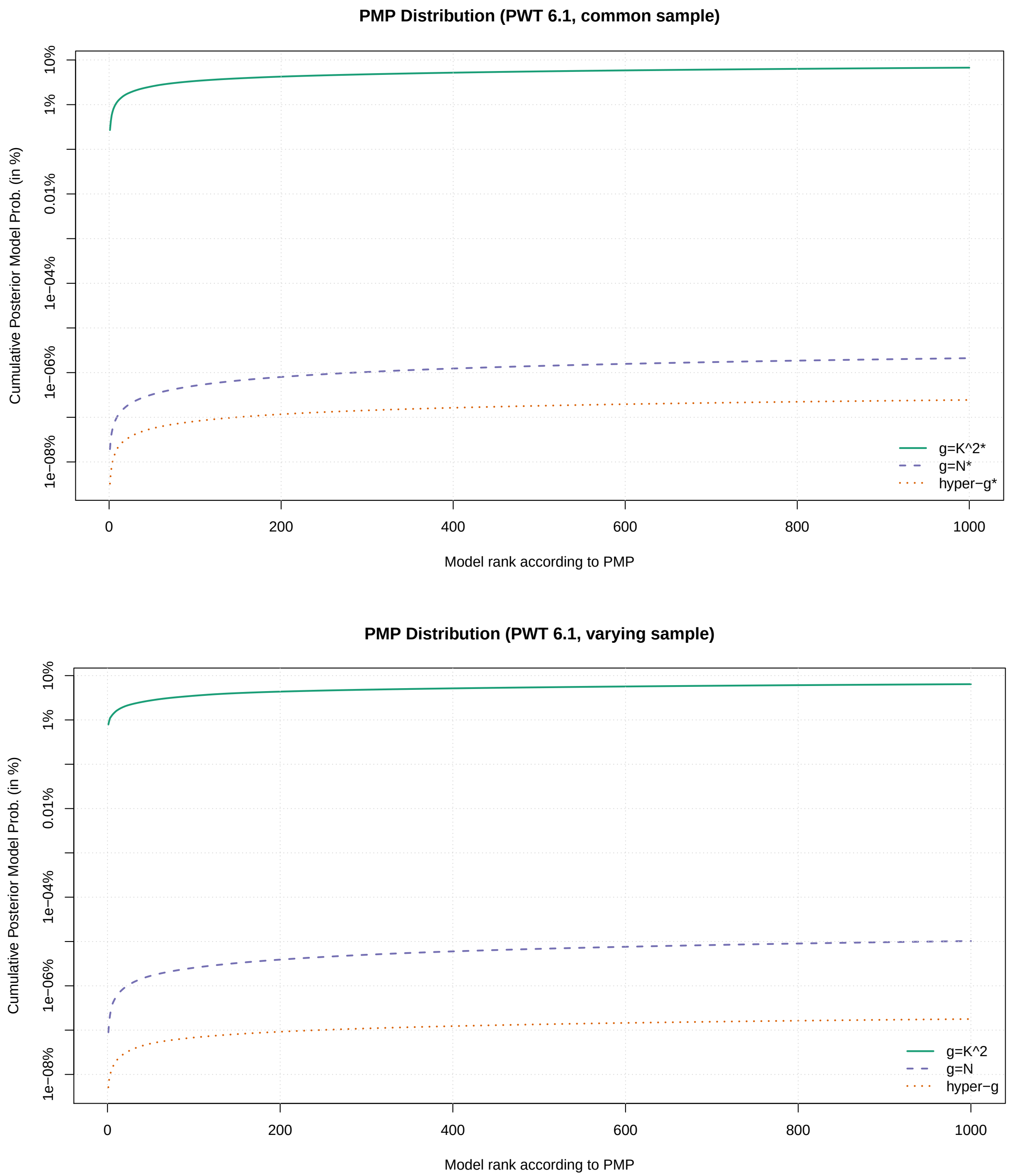

Figure 1: Cumulative posterior model probabilities for the best 1000 models (by PMP) under different settings for g. The lower panel displays results for the original PWT 6.1 data with 88 countries. The upper panel shows results for the PWT 6.1 sample with observations restricted to the 'common sample' countries that appear in later PWT revisions (79 observations).

	$g = K^2$			$g = K^{2*}$			hyper-g*		
PWT revisions	6.0	6.1	6.2	6.0	6.1	6.2	6.0	6.1	6.2
GDP in 1960 (log)	0.73	1.00	1.00	0.81	1.00	1.00	0.87	1.00	1.00
Absolute Latitude	0.04	0.03	0.03	0.06	0.02	0.03	0.38	0.26	0.29
Air Distance to Big Cities	0.05	0.43	0.05	0.07	0.09	0.05	0.34	0.29	0.30
Ethnolinguistic Fractionalization	0.12	0.03	0.04	0.02	0.05	0.04	0.24	0.34	0.28
British Colony Dummy	0.03	0.03	0.02	0.02	0.02	0.02	0.25	0.26	0.23
Fraction Buddhist	0.13	0.13	0.30	0.12	0.17	0.30	0.39	0.59	0.48
Fraction Catholic	0.04	0.02	0.07	0.13	0.06	0.07	0.58	0.39	0.45
Civil Liberties	0.03	0.02	0.02	0.07	0.04	0.02	0.40	0.30	0.24
Colony Dummy	0.04	0.09	0.02	0.12	0.05	0.02	0.36	0.25	0.24
Fraction Confucius	**0.24**	**0.17**	**0.84**	**0.55**	**0.61**	**0.84**	**0.77**	**0.93**	**0.96**
Population Density 1960	**0.11**	**0.71**	**0.02**	**0.02**	**0.04**	**0.02**	**0.29**	**0.25**	**0.24**
Population Density Coastal in 1960s	**0.44**	**0.76**	**0.11**	**0.42**	**0.37**	**0.11**	**0.34**	**0.31**	**0.38**
Interior Density	0.02	0.02	0.03	0.03	0.06	0.03	0.28	0.37	0.26
Population Growth Rate 1960-90	0.02	0.03	0.03	0.03	0.03	0.03	0.30	0.23	0.28
East Asian Dummy	**0.79**	**0.75**	**0.33**	**0.54**	**0.46**	**0.33**	**0.55**	**0.37**	**0.32**
Capitalism	0.02	0.02	0.02	0.02	0.03	0.02	0.24	0.29	0.28
English Speaking Population	0.02	0.02	0.02	0.02	0.02	0.02	0.26	0.28	0.25
European Dummy	0.03	0.04	0.07	0.07	0.08	0.07	0.46	0.50	0.36
Fertility in 1960s	**0.04**	**0.15**	**0.90**	**0.23**	**0.65**	**0.90**	**0.65**	**0.73**	**0.80**
Defense Spending Share	0.02	0.02	0.04	0.03	0.02	0.04	0.31	0.30	0.30
Public Education Spending Share in GDP in 1960s	0.02	0.02	0.02	0.02	0.03	0.02	0.25	0.26	0.25
Public Investment Share	0.07	0.06	0.02	0.03	0.02	0.02	0.33	0.24	0.28
Nominal Government GDP Share 1960s	0.05	0.02	0.29	0.28	0.11	0.29	0.84	0.46	0.74
Government Share of GDP in 1960s	0.08	0.04	0.05	0.09	0.13	0.05	0.31	0.53	0.34
Gov. Consumption Share 1960s	0.12	0.06	0.03	0.05	0.06	0.03	0.29	0.37	0.29
Higher Education 1960	0.07	0.02	0.03	0.14	0.03	0.03	0.60	0.25	0.27
Religion Measure	0.02	0.03	0.02	0.02	0.02	0.02	0.30	0.27	0.24
Fraction Hindus	0.05	0.02	0.03	0.03	0.03	0.03	0.37	0.40	0.33
Investment Price	**0.81**	**0.98**	**0.02**	**0.03**	**0.07**	**0.02**	**0.26**	**0.28**	**0.22**
Latin American Dummy	0.18	0.08	0.34	0.34	0.28	0.34	0.39	0.33	0.34
Land Area	0.02	0.02	0.02	0.04	0.05	0.02	0.49	0.32	0.29
Landlocked Country Dummy	0.02	0.09	0.03	0.03	0.05	0.03	0.31	0.32	0.24
Hydrocarbon Deposits in 1993	0.03	0.13	0.11	0.03	0.22	0.11	0.35	0.73	0.65
Life Expectancy in 1960	0.23	0.26	0.03	0.03	0.03	0.03	0.25	0.23	0.28
Fraction of Land Area Near Navigable Water	0.02	0.05	0.02	0.03	0.04	0.02	0.28	0.26	0.24
Malaria Prevalence in 1960s	0.23	0.03	0.03	0.05	0.03	0.03	0.30	0.26	0.30
Fraction GDP in Mining	0.16	0.28	0.02	0.04	0.02	0.02	0.31	0.22	0.23
Fraction Muslim	0.13	0.22	0.43	0.06	0.49	0.42	0.50	0.88	0.84
Timing of Independence	0.02	0.09	0.14	0.03	0.09	0.14	0.30	0.70	0.85
Oil Producing Country Dummy	0.02	0.02	0.02	0.02	0.02	0.02	0.23	0.22	0.23
Openness measure 1965-74	0.08	0.07	0.16	0.13	0.31	0.16	0.66	0.68	0.62
Fraction Orthodox	0.02	0.02	0.03	0.02	0.02	0.03	0.28	0.35	0.37
Fraction Speaking Foreign Language	0.10	0.04	0.07	0.04	0.03	0.07	0.28	0.26	0.34
Primary Schooling in 1960	0.81	0.99	1.00	0.98	1.00	1.00	0.99	1.00	1.00
Average Inflation 1960-90	0.02	0.02	0.06	0.02	0.03	0.06	0.24	0.25	0.24
Square of Inflation 1960-90	0.02	0.02	0.04	0.02	0.02	0.04	0.24	0.23	0.23
Political Rights	0.07	0.24	0.02	0.06	0.04	0.02	0.43	0.44	0.42
Fraction Population Less than 15	0.04	0.03	0.04	0.03	0.04	0.03	0.42	0.33	0.32
Population in 1960	0.03	0.02	0.02	0.02	0.02	0.02	0.34	0.22	0.23
Fraction Population Over 65	0.03	0.06	0.05	0.05	0.07	0.05	0.48	0.41	0.39
Primary Exports 1970	0.06	0.23	0.28	0.13	0.48	0.28	0.44	0.84	0.76
Fraction Protestants	0.05	0.02	0.07	0.14	0.06	0.07	0.53	0.37	0.44
Real Exchange Rate Distortions	0.10	0.05	0.02	0.11	0.03	0.02	0.60	0.23	0.21
Revolutions and Coups	0.04	0.03	0.03	0.03	0.02	0.03	0.27	0.32	0.34
African Dummy	**0.19**	**0.21**	**0.85**	**0.86**	**0.83**	**0.85**	**0.83**	**0.72**	**0.78**
Outward Orientation	0.04	0.04	0.02	0.02	0.02	0.02	0.26	0.22	0.25
Size of Economy	0.02	0.03	0.04	0.02	0.07	0.04	0.31	0.28	0.28
Socialist Dummy	0.02	0.03	0.03	0.03	0.06	0.03	0.24	0.28	0.22
Spanish Colony	0.12	0.02	0.07	0.28	0.04	0.07	0.41	0.23	0.31
Terms of Trade Growth in 1960s	0.02	0.02	0.02	0.02	0.02	0.02	0.23	0.30	0.28
Terms of Trade Ranking	0.02	0.02	0.02	0.02	0.02	0.02	0.28	0.24	0.22
Fraction of Tropical Area	**0.56**	**0.67**	**0.05**	**0.32**	**0.23**	**0.05**	**0.49**	**0.30**	**0.25**
Fraction Population In Tropics	0.06	0.16	0.04	0.02	0.02	0.04	0.26	0.23	0.28
Fraction Spent in War 1960-90	0.02	0.02	0.02	0.02	0.02	0.02	0.25	0.22	0.26
War Participation 1960-90	0.02	0.02	0.03	0.02	0.02	0.03	0.25	0.22	0.26
Years Open 1950-94	0.12	0.06	0.09	0.12	0.08	0.09	0.31	0.24	0.25
Tropical Climate Zone	0.02	0.03	0.04	0.02	0.04	0.04	0.23	0.36	0.44
Number of Observations	88	84	79	79	79	79	79	79	79

Table 3: Posterior Inclusion Probabilities. Left panel corresponds to data set with varying number of countries, center panel to data set with common countries over the three revisions. Right panel displays results of hyper-g BMA over common countries. The asterisk denotes the use of the common country set. Results are based on 80 million posterior draws after a burn-in phase of 20 million draws.