Concept-based Bayesian model averaging and growth empirics:

Supplementary document*

April 18, 2014

Jan R. Magnus
Department of Econometrics and Operations Research,
Vrije Universiteit Amsterdam

Wendun Wang
Econometric Institute, Erasmus University Rotterdam

*E-mail addresses: jan@janmagnus.nl (Jan Magnus), wangwendun@gmail.com (Wendun Wang).

Contents

1 Introduction

This document provides some supplementary material and additional results relating to Magnus and Wang (2013). It contains the within-group correlation of five groups in which we re-sign a variable, the BACE results (Sala-i-Martin et al., 2004) which we use to compare our results with, and the full results using data-dependent prior probabilities. It provides the procedure and detailed results of sensitivity analysis with respect to the prior π and to the grouping. It also contains a growth empirics study using another data set to test the robustness of the endogenous growth model.

2 Scaling

We present five groups in which we re-sign a variable in Table 1, so that variables within one group are positively correlated.

TABLE 1

Note that the original variables in each group are highly and negatively correlated. Therefore, averaging estimates without scaling cancels the effect of these variables.

3 Results of BACE and HWALS-F1

The original results of BACE in Sala-i-Martin et al. (2004) are posterior mean and standard deviation estimates, conditional on inclusion along with posterior conditional probabilities. Since model uncertainty is not fully taken into account in the posterior standard deviations conditional on inclusion, the precision of the estimates is misleading as explained in Magnus et al. (2010). To 'fairly' compare the estimates

produced by BACE with those of WALS and HWALS, we compute the unconditional ('true') moments of BACE, based on Equations (8) and (14) in Sala-i-Martin et al. (2004).

TABLE 2

The unconditional posterior mean can be computed by multiplying the conditional mean times the posterior inclusion probability, and the unconditional variance can be calculated as

$$\sigma^2_{uncond} = \left(\sigma^2_{cond} + \beta^2_{cond}\right) \times (\textit{posterior inclusion prob}) - \beta^2_{uncond}.$$

Both the conditional and the unconditional estimates of BACE are given in Table 2, where the variables are ordered in the same way as in Magnus and Wang (2013).

Results of HWALS-F1 are also provided in Table 2 as a direct comparison between HWALS and BMA because *all* explanatory variables are allowed to be either included or excluded.

4 Complete results of data-dependent prior

In Section 2.3 of the paper we discussed two updating algorithms based on data-dependent priors: one-step updating and two-step updating. This section provides complete results of HWALS-F8 using two updating methods.

TABLE 3 and 4

In Table 3 and 4 we present the updated priors and the new HWALS-F8 estimates for all groups. The robustness of the updated probabilities and resulting estimates confirms that model specification only has a marginal effect in the updating procedure.

5 Robustness of data-dependent prior with respective to misspecifications

We check whether misspecification is negligible in updating the prior by randomly including some additional controls in the regressions. The results of focus variables are presented in Table 5 (the effects on auxiliary variables are generally smaller). We find that both the updated priors and the HWALS estimates are hardly affected, which confirms the merit of the updating procedure.

TABLE 5

6 Sensitivity with respect to the prior π

In Section 3.3 of the paper we distinguished between four levels of belief regarding the specification of the prior probabilities π. In Section 5 of the paper we present the empirical results based on the first two levels (default equal priors and data-dependent priors). In this section we investigate the effects of the belief of π on the estimates and standard deviations. In other words, we ask how sensitive the empirical results and the conclusions are with respect to π. We discuss the two levels in turn: equal priors and ordered priors. The analysis of equal priors corresponds to the first case where researchers have no information, and we study how the change in the prior of one variable affects the results, keeping priors of other variables equal; The other case of order priors corresponds to the third and fourth cases where researchers have unequal information and are able to order the variables in one group. The way of assigning priors and the effect of priors are very different from Ley and Steel (2009) or Eicher et al. (2011) who found sensitive results with respect to priors, because the priors in our case are assigned to the variables in one group, unlike the priors of the models or parameters. Therefore, we expect the effect of priors on our results is also

different from the findings in Ley and Steel (2009) and Eicher et al. (2011).

6.1 Equal priors

Suppose that in one group, say group l, the m_l variables do not all have the same prior probability $1/m_l$, but that one of the variables, say variable j, has a different probability π_l^j, while the remaining variables have equal probabilities

$$\pi_l^i = \frac{1 - \pi_l^j}{m_l - 1} \qquad (i = 1, \ldots, j-1, j+1, \ldots, m_l). \tag{1}$$

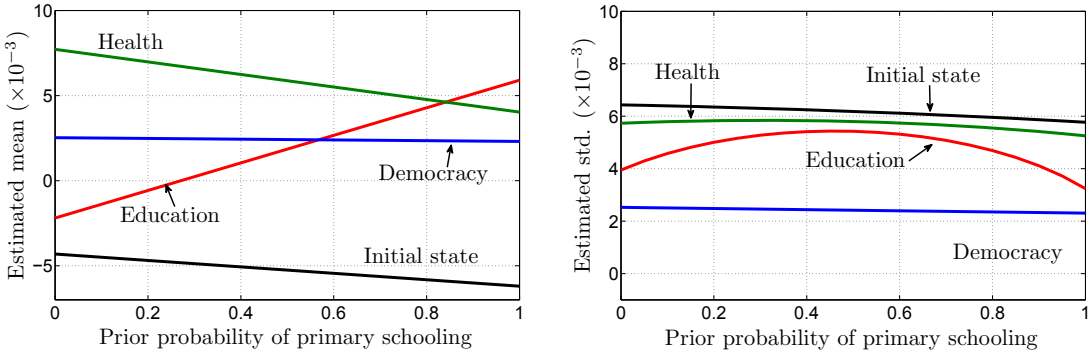
This assumption can be made for any of the type I groups. For our sensitivity experiment we choose one focus variable 'education' and one auxiliary variable 'democracy'. We choose education, because this group contains nine variables with relatively large deviations, and the effect of education on economic growth depends on which variable is used; see Table 3 of the paper. Democracy is of interest because it has strong policy implications (Barro, 1999) and its effect on economic growth is controversial. Within education we choose the variable 'primary schooling' as the one whose prior probability is different, because its role in explaining economic growth appears to differ from other education variables; within democracy we choose the variable 'political rights'.

In Figure 1 we report the sensitivity of four groups: education, democracy, health, and initial state. These groups are chosen because they are proximate determinants that are typically regarded as the most important growth theories. Also, by including education and democracy, we can investigate the direct effect (effect on the group itself) and the indirect effect (effect on other groups) of changing the prior probability. In the figure, the prior probability of primary schooling (top panel) and political rights (bottom panel) varies between 0 and 1. Note that the estimated means of all variables are linear with the varying prior probability, because π_l^i in Equation (1) is a linear function of π_l^j.

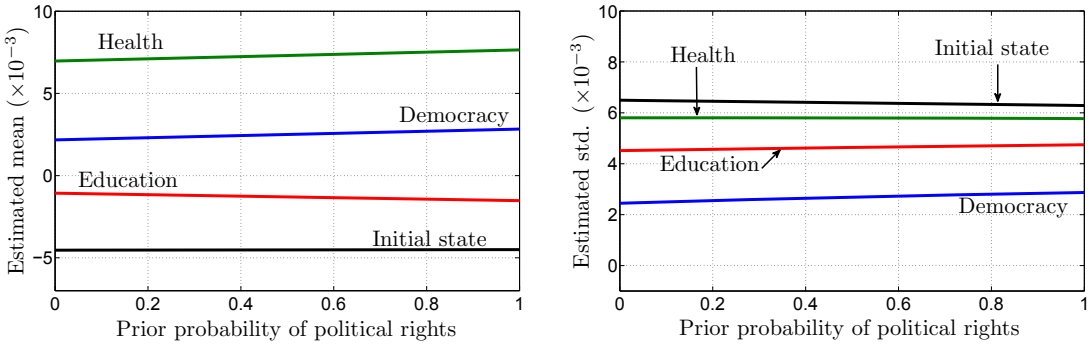
Changing the prior probability of primary schooling has a serious direct effect on the estimated mean and standard deviation of education. The estimated mean of

Figure 1: Sensitivity with respect to π: unequal priors

Unequal prior in group 'education'



Unequal prior in group 'democracy'



education is negative when the prior probability of primary schooling is less than 0.2, but becomes positive when it is larger than 0.4. This is due to the fact that primary schooling has a strong positive effect on growth while the effect of other education variables is weak.

The estimated standard deviation of education is a concave function of the prior probability. It is obvious that when primary schooling has weight 1, we obtain the minimal standard deviation because there is no variation between variables. It is less obvious that when primary schooling has a small weight, we also obtain small standard deviations. The reason lies in the fact that primary schooling differs much

from other education variables, so that a small weight leads to a small cross term bb' in Equation (12) of the paper.

The estimates in the other three groups (indirect effect) are less sensitive than those of education. The estimated means never change sign. Among the three groups, health appears to be the most sensitive to the change in prior. The estimated standard deviations of the three groups are all insensitive to the change in prior.

In the bottom panel we change the prior on political rights in the group democracy. The effects are very small. Even the estimate of democracy itself (the direct effect) is not sensitive.

6.2 Ordered priors

Next suppose that we can order the priors in group l so that the priors of the variables $x_l^1, \ldots, x_l^{m_l}$ are known to satisfy $\pi_l^1 > \cdots > \pi_l^{m_l}$. In particular, assume that $\pi_l^{i+1} = r\pi_l^i$ for some $0 < r < 1$. Then,

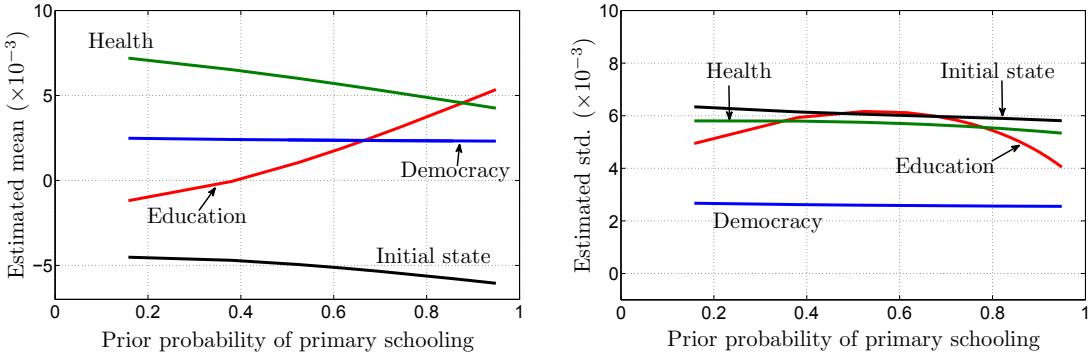
$$\pi_l^i = \frac{(1-r)r^{i-1}}{1-r^{m_l}} \qquad (i = 1, \ldots, m_l). \tag{2}$$

The smaller is r the more weight is placed on the important variables. Equation (2) allows the prior probability of the most important measurement, π_l^1, to change over the interval $(1/m_l, 1)$.

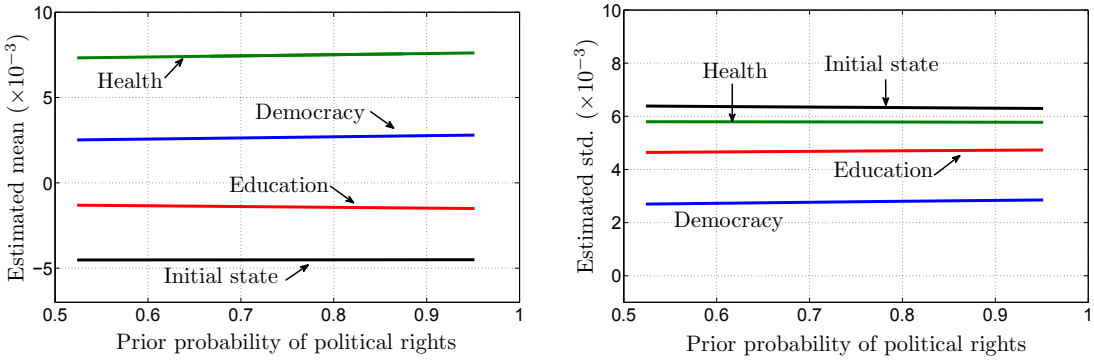
Figure 2 presents some representative examples when the priors are ordered. To perform this experiment we set $r = 1/2$ and we need a predetermined ordering of the variables. In the group education we select primary schooling as the most important variable and we order the other variables randomly. Unlike the previous case, neither the estimated mean nor the standard deviation is a linear function of the prior probability. Still, the main results are essentially the same as before. In particular, as the prior probability of primary schooling increases, the estimated mean of education changes from negative to positive, and the estimated standard deviation of education is a concave function. The health effect is weakened as the

Figure 2: Sensitivity with respect to π: ordered priors

Ordered priors in group 'education'



Ordered priors in group 'democracy'



prior probability of primary schooling increases, while initial state and democracy are insensitive to the probability change.

In the group democracy we select political rights as the most important variable. In this group there are only two variables, so that $\pi^1 = 1/(1+r)$ and $\pi^2 = r/(1+r)$. Changing the probability hardly changes the estimated means and standard deviations for any of the groups.

We repeated these experiments for all other type I groups, both for the unequal prior case and for the ordered priors case. Based on these experiments we draw three conclusions regarding the sensitivity with respect to the prior probability. First, the

effects of proximate determinants on economic growth is robust to the choice of prior probability, except for the group education. Second, the indirect effects of the prior probability are very small, while the direct effect varies across groups. The direct effect is large for those groups whose variables vary greatly, such as education. But it is small for those groups whose variables are highly correlated, such as health, inflation, and scale effect. Third, the standard deviations are quite robust.

7 Sensitivity with respect to grouping

It is not always easy to decide which variable belongs in which group. In this section we investigate the sensitivity of estimates and standard deviations with respect to grouping. We consider six scenarios. First, we consider separating GDP per capita in 1960 and the initial size of the economy. This is motivated by the neoclassical growth model where initial GDP per capita has a structural role and thus should always be included. This is scenario S1. Second, we question whether the variable 'public spending on education' belongs in the group 'education'. This variable has a low correlation with other education variables, so perhaps it should be placed in a separate group (scenario S2). Third, one might make a case for placing this variable in the group 'relative government size' (scenario S3). Besides separating public education spending, one could also consider the possibility that enrollment rates and attainment levels (school years) have a different effect on growth, because the former is a flow measure while the latter captures the stock of human capital. Thus, in scenario S4, we separate enrollment rates (variables 5–7 and 8), school years (variables 9–13), and public education spending. In scenario S5, we allow that lower level (primary and secondary) and higher education may have different effects because the first is related with basic literacy necessary for simpler activities while the latter provides advanced capability useful in some innovative industry. Finally, we consider separating latitude from tropic effect group since it could also measure proximity to major economic hubs (scenario S6).

TABLE 6

Table 6 presents the results of focus variables under alternative groupings. We see that S1 leads to a much larger estimated coefficient ($b = -0.0098$) of the initial level of income and a smaller variance ($V = 0.0053$), making initial income as one of the most important determinants explaining cross-country growth differences. This also provides strong evidence of convergence. The new grouping also has an impact on other focus variables, but this impact is not large. For example, education effect on steady-state growth is still weakly negative, while the effects of health, ethnolinguistic fractionalization, and religion are strengthened to different extents. Results of other focus variables and auxiliary variables are marginally affected except that estimated coefficients of democracy and scale effects reduced by around 28% and 32%, respectively, and that of trade statistic is doubled (standard deviation unchanged). Separating public education spending makes education group slightly more negative but with larger variance given by column S2, while assigning public education spending in 'relative government size' group (column S3) has a weaker effect on education than S2. Column S4 shows that the effects of both the flow measure and the stock measure of education are weak. Distinguishing different levels of education (S5) suggests that primary and secondary education has a very weak effect on growth, while the effect higher education is even negative. These results confirm large variation of education variables as well as their distinct effects on growth. In general we see most education variables are weakly related with growth except for primary schooling. Finally, we see from column S6 that it does not make much difference distinguishing between absolute latitude and other variables in the tropic effect group.

We also look at the average relative change between the HWALS-F8 estimates and standard deviations from the new grouping and the original grouping. If we consider all variables, then some changes in grouping have a relatively significant effect on the estimates, but the standard deviations are not much affected. This

applies in particular to the case where public spending in education is separated from group 'education', and we find 30% change in estimates and 0.01% change in standard deviations. Grouping public spending into relative government size affects the results only marginally, with 6% change in estimates and 0.6% standard deviation change. Distinguishing between lower level and higher education (S4) and flow and stock (S5) lead to moderate changes in both estimates and standard deviations, and separating absolute latitude also has a weak effect. The average relative change is much less if we only compare the focus groups.

8 Growth models: Alternative data set

This section applies hierarchical model averaging to a small set of growth data. There are three reasons for doing this. First, we test the robustness of the endogenous growth model with the distinction between focus and auxiliary regressors. Second, we compare our results with Magnus et al.'s (2010) Model 2 that uses a similar data set. Finally, the small data set allows us to study *all* combinations of measurements, so that we can have more information on the distribution of our estimates, not just the first two moments.

This data set introduces two variables that are not listed in Tables 1 and 2 of our main paper, namely

75 Equipment investment

76 Rule of law index,

both taken from Sala-i-Martin (1997). We follow a similar specification as in Model 2 of Magnus et al. (2010), in which nine focus variables (labeled 'F' in Table 7) and three auxiliary regressors were considered. The auxiliary variables are: political rights (25), fraction GDP in mining (59), and population growth rate (65).

We note two deviations from the specification in Magnus et al. (2010). First, we take into account different alternative measurements of four type I groups, namely

education, health, democracy, and tropics effect. Second, since malaria prevalence belongs to the group 'health', it is not estimated as a separate auxiliary variable as in Magnus et al. (2010).

TABLE 7

The estimates based on the small data set are given in Table 7. The WALS results in the table are different in magnitude from Magnus et al. (2010) (with the same signs) due to the scaling of the regressors and also the different number of observations (We have 6 countries less than Magnus et al. (2010) since we include additional alternative measurements.) For type I groups, the WALS estimates correspond to variables while the estimates in HWALS correspond to groups.

We first compare the signs, and then the precisions as in the large data set. Three groups have different signs: education, tropics effect, and democracy. The counterintuitive sign of education produced by HWALS is mainly due to large variation of nine measurements as discussed in our main paper. For the group tropics effect, HWALS produces a positive estimate but very insignificant. The standard deviation is more than four times the mean, so that the sign of the mean estimate is very uncertain. The large standard deviation is mainly caused by the fact that the variables 'tropic climate zone' and 'absolute latitude' are insignificant and not robust. If we fix this group to 'fraction of tropical area' and re-estimate using HWALS, *ceteris paribus*, we obtain a strongly negative tropics effect. As for the group democracy, HWALS reports a negative effect, but insignificant as well. This is in line with most studies on the association between growth and democracy, as discussed in more detail in our paper.

Next, we comment on the precision of the estimates. Unlike in the large data set, estimates from HWALS, in this case, have larger standard deviations than estimates from WALS. This is because HWALS standard deviation explicitly takes into account additional uncertainty on the choice of measurement that is not considered

in WALS. It is also because there are not many type I groups in this small data set, and thus not many highly correlated variables are included in the WALS regressions. Therefore, multicollinearity is not a serious problem here.

References

R. J. Barro. Determinants of democracy. *The Journal of Political Economy*, 107: S158–S183, 1999.

T. S. Eicher, C. Papageorgiou, and A. E. Raftery. Default priors and predictive performance in Bayesian model averaging, with application to growth determinants. *Journal of Applied Econometrics*, 26:30–55, 2011.

E. Ley and M. F. Steel. On the effect of prior assumptions in bayesian model averaging with applications to growth regression this article was published online on 30 march 2009. an error was subsequently identified. this not. *Journal of Applied Econometrics*, 24:651–674, 2009.

J. R. Magnus and W. Wang. Concept-based Bayesian model averaging and growth empirics. Discussion paper, 2013.

J. R. Magnus, O. Powell, and P. Prüfer. A comparison of two model averaging techniques with an application to growth empirics. *Journal of Econometrics*, 154: 139–153, 2010.

X. Sala-i-Martin. I just ran two million regressions. *American Economic Review*, 87: 178–183, 1997.

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

9 Tables

Table 1: Within-group correlations

g	Group	v	Variable	Correlation
(1)	Demographic characteristics	1*	Fraction population over 65	
		2	Fraction population under 15	−0.91
(2)	Economy system	3	Capitalism	−0.58
		4*	Socialist dummy	
(5)	Health	19	Life expectancy in 1960	−0.73
		20*	Malaria prevalence in 1960s	
(8)	Democracy	25	Political rights	−0.83
		26*	Civil liberties	
(11)	Tropics effect	31	Fraction of tropical area	−0.89
		32	Tropical climate zone	−0.60
		33*	Absolute latitude	

* Adjusted variable.

Table 2: BACE results and HWALS-F1

Variable	BACE results		HWALS-F1
	Conditional posterior estimates	Unconditional posterior estimates	
Education			−0.0013 (0.0038)
5 Primary schooling	0.0269 (0.0080)	0.0214 (0.0130)	
6 Secondary schooling			
7 Higher education	−0.0697 (0.0418)	−0.0043 (0.0196)	
8 Public edu. spending	0.1295 (0.1728)	0.0027 (0.0312)	
9 Primary school yrs			
10 Secondary school yrs			
11 Higher education yrs			
12 Ave. school yrs			
13 Ave. school yrs × logGDP			
Health			0.0045 (0.0044)
19 Life expectancy	0.0008 (0.0004)	0.0002 (0.0004)	
20 Malaria prevalence	−0.0157 (0.0062)	−0.0040 (0.0075)	
Initial state			−0.0030 (0.0046)
23 GDP in 1960 (log)	−0.0085 (0.0029)	−0.0058 (0.0046)	
24 Size of economy	−0.0005 (0.0014)	−0.0000 (0.0002)	
Tropics effect			−0.0029 (0.0032)
31 Frac. of tropical area	−0.0148 (0.0042)	−0.0083 (0.0080)	
32 Tropical climate zone	−0.0021 (0.0066)	0.0000 (0.0009)	
33 Absolute latitude	0.0001 (0.0002)	0.0000 (0.0000)	
Ethnicity and language			
36 Ethnolinguistic frac.	−0.0113 (0.0058)	−0.0012 (0.0039)	−0.0024 (0.0025)
37 English-speaking pop.	−0.0037 (0.0071)	−0.0001 (0.0011)	
38 Frac. foreign language	0.0070 (0.0040)	0.0006 (0.0022)	
Religion			
39 Fraction Confucian	0.0544 (0.0224)	0.0112 (0.0242)	
40 Fraction Muslim	0.0126 (0.0063)	0.0014 (0.0045)	
41 Fraction Buddhist	0.0217 (0.0107)	0.0023 (0.0076)	
42 Fraction Protestant	−0.0119 (0.0093)	−0.0005 (0.0032)	
43 Fraction Hindu	0.0176 (0.0126)	0.0008 (0.0045)	
44 Fraction Catholic	−0.0084 (0.0085)	−0.0003 (0.0022)	
45 Fraction Orthodox	0.0057 (0.0136)	0.0001 (0.0018)	
46 Religious intensity	−0.0047 (0.0072)	−0.0001 (0.0012)	−0.0009 (0.0017)
Price distortion			
70 Investment price	−0.0001 (0.0000)	−0.0001 (0.0000)	−0.0029 (0.0015)

Table 2: BACE results and HWALS-F1 (Continued)

Variable	BACE results		HWALS-F1
	Conditional posterior estimates	Unconditional posterior estimates	
Demographic characteristics			0.0026 (0.0047)
1 Frac. pop. over 65	0.0194 (0.1195)	0.0004 (0.0180)	
2 Frac. pop. under 15	0.0450 (0.0411)	0.0018 (0.0122)	
Economy system			−0.0010 (0.0016)
3 Capitalism	−0.0002 (0.0011)	0.0000 (0.0001)	
4 Socialist dummy	0.0040 (0.0050)	0.0001 (0.0009)	
Relative government size			−0.0004 (0.0021)
14 Public investment share	−0.0615 (0.0430)	−0.0030 (0.0162)	
15 Public consumption share (excl. education and defense)			
16 Gov. consumption share	−0.0442 (0.0254)	−0.0046 (0.0158)	
17 Gov. share of GDP	−0.0349 (0.0294)	−0.0022 (0.0112)	
18 Nominal gov. GDP share	−0.0336 (0.0274)	−0.0012 (0.0081)	
Inflation			0.0005 (0.0022)
21 Average inflation	−0.0001 (0.0001)	−0.0000 (0.0000)	
22 Square of inflation	0.0000 (0.0000)	0.0000 (0.0000)	
Democracy			0.0025 (0.0027)
25 Political rights	−0.0018 (0.0102)	−0.0001 (0.0005)	
26 Civil liberties	−0.0072 (0.0071)	−0.0002 (0.0017)	
Scale effect			0.0028 (0.0027)
27 Land area	0.0000 (0.0000)	0.0000 (0.0000)	
28 Population	0.0000 (0.0000)	0.0000 (0.0000)	
Trade policy indices			0.0010 (0.0025)
29 Outward orientation	−0.0033 (0.0027)	−0.0001 (0.0007)	
30 Years open	0.0122 (0.0063)	0.0015 (0.0045)	
War			−0.0001 (0.0017)
34 Frac. spent in war	−0.0014 (0.0092)	−0.0000 (0.0012)	
35 War participation	−0.0007 (0.0030)	−0.0000 (0.0004)	
Trade statistics			
47 Openness measure	0.0089 (0.0052)	0.0007 (0.0028)	0.0007 (0.0029)
48 Primary exports	−0.0113 (0.0075)	−0.0006 (0.0031)	
Terms of trade			
49 Terms of trade ranking	−0.0037 (0.0096)	−0.0001 (0.0013)	0.0004 (0.0027)
50 Terms of trade growth	0.0326 (0.0467)	0.0007 (0.0082)	0.0036 (0.0024)
Regional effect			
51 East Asian dummy	0.0218 (0.0061)	0.0179 (0.0100)	0.0062 (0.0028)
52 African dummy	−0.0147 (0.0069)	−0.0023 (0.0060)	−0.0033 (0.0035)
53 European dummy	−0.0023 (0.0105)	−0.0001 (0.0019)	0.0016 (0.0045)
54 Latin-American dummy	−0.0128 (0.0058)	−0.0019 (0.0051)	−0.0012 (0.0045)
55 Colony dummy	−0.0050 (0.0047)	−0.0001 (0.0012)	−0.0040 (0.0035)
56 British colony	0.0037 (0.0036)	0.0001 (0.0008)	0.0028 (0.0027)
57 Spanish colony	−0.0107 (0.0050)	−0.0013 (0.0039)	0.0015 (0.0033)

Table 2: BACE results and HWALS-F1 (Continued)

Natural resource			
58 Hydrocarbon deposits	0.0003 (0.0004)	0.0000 (0.0001)	0.0001 (0.0019)
59 Frac. GDP in mining	0.0388 (0.0193)	0.0048 (0.0145)	−0.0014 (0.0019)
60 Oil country dummy	0.0048 (0.0071)	0.0001 (0.0012)	−0.0018 (0.0023)
Population			
61 Population density coastal	0.0000 (0.0000)	0.0000 (0.0000)	0.0010 (0.0029)
62 Interior density	0.0000 (0.0000)	0.0000 (0.0000)	−0.0011 (0.0017)
63 Fraction pop. in tropics	−0.0107 (0.0068)	−0.0006 (0.0030)	0.0014 (0.0032)
64 Population density	0.0000 (0.0000)	0.0000 (0.0000)	−0.0016 (0.0021)
65 Population growth rate	0.0208 (0.3078)	0.0004 (0.0425)	0.0013 (0.0053)
66 Fertility	−0.0075 (0.0101)	−0.0002 (0.0022)	−0.0031 (0.0061)
Geography (excl. tropics effect)			
67 Frac. land area near water	−0.0026 (0.0059)	0.0000 (0.0009)	0.0018 (0.0032)
68 Landlocked country dummy	−0.0021 (0.0042)	0.0000 (0.0007)	0.0003 (0.0018)
69 Air distance to big cities	0.0000 (0.0000)	0.0000 (0.0000)	0.0010 (0.0025)
Real exchange rate			
71 Real exchange rate dist.	−0.0001 (0.0000)	0.0000 (0.0000)	−0.0024 (0.0020)
Defense			
72 Defense spending share	0.0453 (0.0768)	0.0010 (0.0129)	−0.0003 (0.0017)
Political instability			
73 Revolutions and coups	−0.0071 (0.0061)	−0.0002 (0.0016)	−0.0005 (0.0019)
Independence			
74 Timing of independence	0.0011 (0.0021)	0.0000 (0.0003)	0.0006 (0.0025)

Table 3: HWALS estimates using data-dependent priors: Focus variables

	One-step updating		Two-step updating	
Variable	HWALS-F8	updated π	HWALS-F8	updated π
Education	0.0051 (0.0034)		0.0050 (0.0034)	
5 Primary schooling		0.9784		0.9784
6 Secondary schooling		0.0033		0.0033
7 Higher education		0.0024		0.0024
8 Public edu. spending		0.0027		0.0027
9 Primary school yrs		0.0025		0.0025
10 Secondary school yrs		0.0031		0.0031
11 Higher education yrs		0.0019		0.0019
12 Ave. school yrs		0.0024		0.0024
13 Ave. school yrs × logGDP		0.0033		0.0033
Health	0.0062 (0.0059)		0.0065 (0.0060)	
19 Life expectancy		0.8142		0.8142
20 Malaria prevalence		0.1858		0.1858
Initial state	−0.0084 (0.0057)		−0.0082 (0.0059)	
23 GDP in 1960 (log)		0.6923		0.6923
24 Size of economy		0.3077		0.3077
Tropics effect	−0.0041 (0.0034)		−0.0040 (0.0034)	
31 Frac. of tropical area		0.5488		0.5488
32 Tropical climate zone		0.1489		0.1489
33 Absolute latitude		0.3023		0.3023
Ethnicity and language				
36 Ethnolinguistic frac.	−0.0022 (0.0026)	–	−0.0023 (0.0028)	–
37 English-speaking pop.		–		–
38 Frac. foreign language		–		–
Religion				
39 Fraction Confucian		–		–
40 Fraction Muslim		–		–
41 Fraction Buddhist		–		–
42 Fraction Protestant		–		–
43 Fraction Hindu		–		–
44 Fraction Catholic		–		–
45 Fraction Orthodox		–		–
46 Religious intensity	−0.0022 (0.0018)	–	−0.0022 (0.0018)	–
Price distortion				
70 Investment price	−0.0046 (0.0015)	–	−0.0045 (0.0016)	–

Table 4: HWALS estimates using data-dependent priors: Auxiliary variables

	One-step updating		Two-step updating	
Variable	HWALS-F8	updated π	HWALS-F8	updated π
Demographic characteristics	0.0026 (0.0044)		0.0019 (0.0041)	
1 Frac. pop. over 65		0.7202		0.3282
2 Frac. pop. under 15		0.2798		0.6718
Economy system	−0.0007 (0.0015)		−0.0007 (0.0014)	
3 Capitalism		0.5227		0.5516
4 Socialist dummy		0.4773		0.4484
Relative government size	0.0005 (0.0020)		0.0004 (0.0020)	
14 Public investment share		0.1555		0.1647
15 Public consumption share (excl. education and defense)		0.1710		0.0220
16 Gov. consumption share		0.4317		0.5210
17 Gov. share of GDP		0.0827		0.0440
18 Nominal gov. GDP share		0.1590		0.2484
Inflation	0.0004 (0.0019)		0.0007 (0.0018)	
21 Average inflation		0.5005		0.5372
22 Square of inflation		0.4995		0.4628
Democracy	0.0015 (0.0024)		0.0017 (0.0024)	
25 Political rights		0.3017		0.2217
26 Civil liberties		0.6983		0.7783
Scale effect	0.0018 (0.0024)		0.0012 (0.0019)	
27 Land area		0.5293		0.1411
28 Population		0.4707		0.8589
Trade policy indices	0.0009 (0.0026)		0.0011 (0.0027)	
29 Outward orientation		0.0002		0.0005
30 Years open		0.9998		0.9995
War	−0.0003 (0.0015)		−0.0005 (0.0015)	
34 Frac. spent in war		0.5932		0.6160
35 War participation		0.4068		0.3840
Trade statistics				
47 Openness measure	−0.0003 (0.0025)	–	−1.50e-5(0.0025)	–
48 Primary exports		–		–
Terms of trade				
49 Terms of trade ranking	0.0002 (0.0024)	–	−0.0003 (0.0023)	–
50 Terms of trade growth	0.0021 (0.0022)	–	0.0019 (0.0022)	–
Regional effect				
51 East Asian dummy	0.0046 (0.0025)	–	0.0046 (0.0025)	–
52 African dummy	−0.0020 (0.0032)	–	−0.0018 (0.0032)	–
53 European dummy	0.0009 (0.0040)	–	0.0014 (0.0040)	–
54 Latin-American dummy	−0.0002 (0.0042)	–	0.0004 (0.0042)	–
55 Colony dummy	−0.0040 (0.0031)	–	−0.0036 (0.0030)	–
56 British colony	0.0022 (0.0026)	–	0.0021 (0.0025)	–
57 Spanish colony	0.0007 (0.0031)	–	0.0002 (0.0029)	–

Table 4: HWALS estimates using data-dependent priors: Auxiliary variables (Continued)

	One-step updating		Two-step updating	
Variable	HWALS-F8	updated π	HWALS-F8	updated π
Natural resource				
58 Hydrocarbon deposits	0.0005 (0.0017)	–	0.0005 (0.0017)	–
59 Frac. GDP in mining	−0.0012 (0.0017)	–	−0.0012 (0.0017)	–
60 Oil country dummy	−0.0004 (0.0021)	–	−0.0003 (0.0020)	–
Population				
61 Population density coastal	0.0026 (0.0025)	–	0.0024 (0.0026)	–
62 Interior density	−0.0008 (0.0015)	–	−0.0007 (0.0015)	–
63 Fraction pop. in tropics	0.0009 (0.0028)	–	0.0007 (0.0028)	–
64 Population density	−0.0009 (0.0018)	–	−0.0007 (0.0017)	–
65 Population growth rate	0.0003 (0.0047)	–	0.0012 (0.0047)	–
66 Fertility	−0.0006 (0.0052)	–	−0.0004 (0.0052)	–
Geography (excl. tropics effect)				
67 Frac. land area near water	0.0001 (0.0030)	–	−0.0004 (0.0027)	–
68 Landlocked country dummy	−0.0003 (0.0016)	–	−0.0005 (0.0016)	–
69 Air distance to big cities	−0.0001 (0.0023)	–	−0.0003 (0.0022)	–
Real exchange rate				
71 Real exchange rate dist.	−0.0019 (0.0019)	–	−0.0019 (0.0019)	–
Defense				
72 Defense spending share	−0.0007 (0.0016)	–	−0.0008 (0.0016)	–
Political instability				
73 Revolutions and coups	−0.0003 (0.0017)	–	−0.0005 (0.0017)	–
Independence				
74 Timing of independence	0.0010 (0.0023)	–	0.0010 (0.0022)	–

Table 5: Robustness of HWALS estimates using data-dependent priors: Focus variables

Variable	(1)	(2)	(3)
Education	0.0051 (0.0034)	0.0058 (0.0033)	0.0051 (0.0032)
Health	0.0062 (0.0059)	0.0047 (0.0053)	0.0074 (0.0054)
Initial state	−0.0084 (0.0057)	−0.0072 (0.0058)	−0.0110 (0.0038)
Tropics effect	−0.0041 (0.0034)	−0.0028 (0.0034)	−0.0040 (0.0032)
Ethnicity and language	−0.0022 (0.0026)	−0.0022 (0.0026)	−0.0023 (0.0024)
Religion	−0.0022 (0.0018)	−0.0021 (0.0018)	−0.0022 (0.0017)
Price distortion	−0.0046 (0.0015)	−0.0043 (0.0016)	−0.0047 (0.0015)

Notes: Column (1) gives the estimates with no auxiliary variables in the updating procedure (the same as reported in the paper); Column (2) gives the estimates from the same updating procedure except that two randomly selected controls are added when updating the prior; Column (3) gives the estimates with four randomly selected controls included.

Table 6: Sensitivity analysis on grouping: Focus variables

	S1	S2	S3	S4	S5	S6
Education	−0.0002 (0.0043)	−0.0021 (0.0050)	−0.0017 (0.0048)			−0.0015 (0.0048)
Edu.stock				−0.0020 (0.0066)		
Edu.flow				−0.0011 (0.0042)		
Lower edu.					0.0000 (0.0049)	
Higher edu.					−0.0044 (0.0032)	
Health	0.0088 (0.0064)	0.0077 (0.0060)	0.0074 (0.0058)	0.0077 (0.0060)	0.0082 (0.0062)	0.0071 (0.0058)
Initial state	−0.0098 (0.0053)	−0.0046 (0.0065)	−0.0045 (0.0064)	−0.0043 (0.0062)	−0.0042 (0.0063)	−0.0048 (0.0064)
Tropics effect	−0.0030 (0.0033)	−0.0029 (0.0034)	−0.0031 (0.0034)	−0.0032 (0.0034)	−0.0031 (0.0034)	−0.0028 (0.0041)
Ethnicity and lang.	−0.0035 (0.0028)	−0.0028 (0.0028)	−0.0030 (0.0028)	−0.0024 (0.0028)	−0.0024 (0.0028)	−0.0030 (0.0028)
Religion	−0.0017 (0.0019)	−0.0017 (0.0019)	−0.0015 (0.0019)	−0.0018 (0.0019)	−0.0017 (0.0020)	−0.0016 (0.0020)
Price distortion	−0.0041 (0.0017)	−0.0042 (0.0017)	−0.0041 (0.0017)	−0.0044 (0.0017)	−0.0045 (0.0017)	−0.0040 (0.0018)

Notes: Public education spending is treated as auxiliary in S2–S5, and other education-related groups are focus.

Table 7: Results for small data set

Group/Variable	WALS	HWALS
Constant (F)	0.0211 (0.0013)	0.0211 (0.0014)
Type I groups/variables		
Education (F)		−0.0009 (0.0038)
5 Primary schooling (F)	0.0039 (0.0026)	
Health (F)		0.0060 (0.0045)
19 Life expectancy (F)	0.0065 (0.0041)	
20 Malaria prevalence	0.0022 (0.0018)	
Initial state (F)		−0.0062 (0.0068)
23 GDP in 1960 (log) (F)	−0.0149 (0.0030)	
Democracy		0.0008 (0.0027)
25 Political rights	−0.0014 (0.0018)	
Tropics effect (F)		0.0005 (0.0024)
31 Frac. of tropical area (F)	−0.0017 (0.0019)	
Type II variables		
36 Ethnolinguistic frac. (F)	−0.0020 (0.0018)	−0.0026 (0.0022)
39 Frac. Confucian (F)	0.0049 (0.0015)	0.0060 (0.0019)
59 Frac. GDP in mining	−0.0003 (0.0013)	−0.0006 (0.0015)
65 Population growth rate	0.0015 (0.0021)	0.0013 (0.0025)
75 Equipment investment (F)	0.0041 (0.0020)	0.0053 (0.0024)
76 Rule of law index (F)	0.0074 (0.0024)	0.0073 (0.0031)