

**Online Appendix – Hasan, I., Horvath, R. and J. Mares
(2020) Finance and Wealth Inequality, Journal of International Money and Finance, forthcoming**

Additional robustness checks

Table B1: Dependent variable - average Gini index (wealth) 2010-2016, 73 observations, UIP parameter prior

	PIP	Post Mean	Post SD
Financial institutions efficiency	0.99	-0.36999	0.12386
Value added in agriculture	0.99	-0.56485	0.18154
Access to financial institutions	0.98	-0.44382	0.16204
Financial market development	0.84	0.44193	0.23922
Outward orientation	0.78	0.21853	0.14535
Education index (UN)	0.58	-0.23984	0.24290
Redistribution	0.38	-0.10095	0.15101
Economic freedom index (adjusted)	0.32	-0.10501	0.18144
Net national savings	0.30	0.07686	0.13764
Number of war years	0.23	0.03833	0.08335
Natural resource rents	0.22	0.04549	0.10083
Financial institutions development	0.20	0.10354	0.23661
Net foreign direct investment	0.19	-0.03276	0.08044
Latin America dummy	0.09	0.01404	0.05849
Population density	0.08	-0.01162	0.05108
Average GDP growth	0.08	-0.00950	0.04338
Labor market regulation	0.06	0.00671	0.03585
Population growth	0.06	0.00788	0.04715
Inflation	0.06	0.00568	0.03341
GDP level in 1990	0.06	-0.01404	0.08467
Technological progress	0.05	-0.01188	0.07248
Financial development index (EFW)	0.05	-0.00641	0.04430
Financial markets efficiency	0.05	-0.00499	0.03332
Leftwing orientation	0.05	-0.00400	0.02612
Government expenditures	0.05	0.00463	0.03646
Banking diversification	0.04	-0.00316	0.02370
Value added in industry	0.04	0.00229	0.03279
Life expectancy	0.03	-0.00160	0.03867
Active banking restrictions	0.03	-0.00213	0.02262
Revolutions and coups	0.03	0.00178	0.02012
Financial openness (Chinn-Ito)	0.03	-0.00137	0.02553
Rule of law	0.03	0.00093	0.03789
Civ. liberties and pol. rights	0.03	-0.00131	0.02953
Bank capital regulations	0.03	-0.00131	0.01725
Public education expenditures	0.03	0.00113	0.01817
Business conditions	0.03	-0.00000	0.01732
Labor force participation	0.02	0.00028	0.01376

Table B2: Dependent variable - average Gini index (wealth) 2010-2016, 73 observations, dilution parameter prior

	PIP	Post Mean	Post SD
Financial institutions efficiency	0.93	-0.29559	0.14058
Access to financial institutions	0.88	-0.35265	0.19165
Financial market development	0.85	0.38321	0.21129
Value added in agriculture	0.81	-0.37066	0.23301
Outward orientation	0.66	0.15971	0.14225
Number of war years	0.41	0.06813	0.10412
Net national savings	0.40	0.10489	0.15200
Net foreign direct investment	0.40	-0.06582	0.10158
Education index (UN)	0.33	-0.12682	0.20519
Natural resource rents	0.32	0.06267	0.11045
Redistribution	0.32	-0.08372	0.14239
Latin America dummy	0.25	0.04844	0.10292
Average GDP growth	0.20	-0.02656	0.07126
Value added in industry	0.15	0.03229	0.09069
Financial institutions development	0.14	0.06411	0.17325
Labor market regulation	0.12	0.01228	0.04752
Leftwing orientation	0.11	-0.00800	0.03714
Economic freedom index (adjusted)	0.11	-0.03180	0.10542
Inflation	0.10	0.01006	0.04385
Population density	0.09	-0.00999	0.04676
Banking diversification	0.09	-0.00557	0.03201
Financial development index (EFW)	0.08	-0.01339	0.05852
Bank capital regulations	0.06	-0.00114	0.02308
Labor force participation	0.06	-0.00002	0.02089
Public education expenditures	0.05	0.00208	0.02499
Revolutions and coups	0.05	0.00270	0.02436
Government expenditures	0.04	0.00506	0.03702
Financial markets efficiency	0.04	-0.00350	0.02844
Population growth	0.04	0.00542	0.04010
Active banking restrictions	0.03	-0.00191	0.02272
Financial openness (Chinn-Ito)	0.03	-0.00266	0.02558
Business conditions	0.03	0.00043	0.01735
Civ. liberties and pol. rights	0.01	0.00054	0.01473
Life expectancy	0.00	-0.00069	0.01508
Technological progress	0.00	-0.00099	0.02030
GDP level in 1990	0.00	-0.00102	0.02294
Rule of law	0.00	-0.00013	0.00744

Table B3: Dependent variable - average Gini index (wealth) 2010-2016, 73 observations, relative redistribution measure

	PIP	Post Mean	Post SD
Value added in agriculture	1.00	-0.51152	0.15591
Financial institutions efficiency	0.99	-0.28741	0.11147
Access to financial institutions	0.98	-0.34837	0.15459
Redistribution (rel.)	0.95	-0.27535	0.14043
Outward orientation	0.94	0.23308	0.11250
Financial market depth	0.81	0.34002	0.21938
Education index (UN)	0.72	-0.22528	0.20282
Number of war years	0.59	0.08973	0.10332
Economic freedom index (adjusted)	0.36	-0.08389	0.15606
Labour market regulation	0.32	0.03829	0.07734
Natural resources rents	0.28	0.04065	0.08833
Financial institutions depth	0.28	0.10702	0.21832
Average GDP growth	0.28	-0.03598	0.07976
Rule of law	0.26	0.07442	0.17734
Leftwing orientation	0.22	-0.02359	0.06261
Net foreign direct investment	0.20	-0.02042	0.05956
Net national savings	0.20	0.02747	0.08091
Government expenditures	0.16	0.01994	0.06646
Bank capital regulations	0.11	-0.00826	0.03810
Population density	0.10	-0.00737	0.03893
Civ. liberties and Pol. rights	0.09	-0.00684	0.05487
Business conditions	0.09	-0.00679	0.03889
GDP level in 1990	0.09	-0.00754	0.08113
Public education expenditures	0.09	0.00452	0.03209
Financial openness (Chinn-Ito)	0.08	0.00371	0.04077
Banking diversification	0.08	-0.00453	0.02891
Financial liberalization (EFW)	0.08	-0.00225	0.04299
Active banking restrictions	0.08	-0.00396	0.03201
Latin America dummy	0.07	0.00613	0.08853
Technological progress	0.07	-0.00810	0.06770
Financial markets efficiency	0.06	-0.00027	0.03059
Inflation	0.06	0.00238	0.02706
Labour force participation	0.06	0.00036	0.02077
Life expectancy	0.06	0.00055	0.04579
Population growth	0.06	0.00076	0.03609
Value added in industry	0.05	-0.00063	0.02761
Revolutions and coups	0.05	0.00069	0.02121

Table B4: Dependent variable - average Gini index (wealth) 2010-2016, specific financial indicators as proxies for financial development, 73 observations, dilution parameter prior

	PIP	Post Mean	Post SD
Outward orientation	1.00	0.30288	0.09493
Value added in agriculture	1.00	-0.46969	0.16524
Number of war years	1.00	0.23140	0.09211
Bank branches/1000 inh.	0.99	-0.23286	0.10392
Redistribution	0.96	-0.27204	0.13368
Private credit	0.80	0.26709	0.20234
Average GDP growth	0.72	-0.12719	0.11806
Net interest margin	0.71	0.26047	0.23046
Business conditions	0.63	-0.16526	0.17583
Inflation	0.52	0.08140	0.10963
Education index (UN)	0.43	-0.09997	0.16364
Economic freedom index (adjusted)	0.38	-0.11007	0.18830
Leftwing orientation	0.26	-0.02542	0.06428
Labor market regulation	0.17	0.01351	0.04931
Rule of law	0.17	0.02859	0.11191
Net national savings	0.16	0.01665	0.06290
Natural resource rents	0.16	0.01609	0.06250
Bank Z-score	0.15	0.01193	0.04857
Latin America dummy	0.13	0.01040	0.05422
Banking diversification	0.12	-0.00670	0.03591
Market capitalization	0.11	0.00106	0.04334
Market turnover	0.11	0.00559	0.03372
Civ. liberties and pol. rights	0.11	0.00419	0.05246
Value added in industry	0.11	0.00610	0.04528
Population growth	0.11	0.00659	0.05385
Life expectancy	0.10	-0.00578	0.06521
Technological progress	0.10	0.00530	0.08492
Financial development index (EFW)	0.10	0.00203	0.05079
Net foreign direct investment	0.10	-0.00504	0.03344
GDP level in 1990	0.10	0.00277	0.08595
Financial openness (Chinn-Ito)	0.09	0.00422	0.04314
Public education expenditures	0.09	0.00437	0.03492
Government expenditures	0.09	0.00648	0.04413
Loan-to-deposits	0.09	0.00400	0.03650
Revolutions and coups	0.09	0.00307	0.03130
Active banking restrictions	0.08	0.00076	0.03139
Bank capital regulations	0.08	-0.00113	0.02484
Population density	0.07	0.00112	0.02579
Labor force participation	0.07	-0.00105	0.02323

Grouping of explanatory variables

Table B5: Explanatory Variables Sorted into Groups

GROUP	VARIABLES
Economic	Value added in agriculture Value added in industry Outward orientation Redistribution Net national savings Net foreign direct investment Average GDP growth GDP level in 1990 Inflation Government expenditures Public education expenditures Technological progress Labor force participation
Financial	Financial institutions efficiency Access to financial institutions Financial market development Financial institutions development Financial markets efficiency
Political	Number of war years Leftwing orientation Revolutions and coups Civ. liberties and pol. rights
Institutional	Education index (UN) Economic freedom index (adjusted) Rule of law
Regulatory	Labor market regulation Banking diversification Active banking restrictions Bank capital regulations Financial openness (Chinn-Ito) Business conditions Financial liberalization index (EFW)
Geographical / natural	Natural resource rents Population density Latin America dummy Population growth Life expectancy

Descriptive statistics, correlation matrix, expected effects

Table B6: Descriptive statistics

	Min.	Mean	Max.	Std.dev.
Access to financial institutions	0.02	0.36	0.96	0.26
Active banking restrictions	3.75	7.20	11.00	1.59
Average GDP growth	-0.02	0.02	0.06	0.01
Bank capital regulations	2.00	6.64	10.00	1.61
Banking diversification	0.00	1.32	2.00	0.46
Business conditions	-0.66	-0.28	1.53	0.36
Civ. liberties and Pol. rights	1.00	2.88	5.41	1.42
Economic freedom index (adjusted)	0.48	0.70	0.89	0.10
Education index (UN)	0.27	0.63	0.89	0.15
Financial institutions depth	0.02	0.31	0.86	0.24
Financial institutions efficiency	0.28	0.58	0.76	0.12
Financial liberalization (EFW)	4.01	7.34	9.49	1.52
Financial market depth	0.00	0.22	0.73	0.20
Financial markets efficiency	0.01	0.35	0.95	0.26
Financial openness (Chinn-Ito)	-1.47	0.41	2.39	1.26
GDP level in 1990	6.69	9.00	10.57	1.02
Government expenditures	4.75	16.14	27.48	4.63
Inflation	1.93	46.70	466.21	101.75
Labour force participation	0.00	0.00	0.00	0.00
Labour market regulation	0.46	1.67	2.78	0.51
Latin America dummy	0.00	0.18	1.00	0.39
Leftwing orientation	0.00	8.81	30.00	8.37
Life expectancy	45.51	68.88	78.04	7.86
Natural resources rents	0.00	3.49	31.66	5.30
Net foreign direct investment	0.09	2.95	12.56	2.42
Net national savings	-8.54	8.85	30.00	6.51
Number of war years	0.00	2.38	21.00	4.57
Outward orientation	-0.33	-0.03	0.19	0.08
Population density	2.22	164.99	4547.96	536.87
Population growth	-0.57	1.24	3.62	1.04
Public education expenditures	1.24	4.27	11.18	1.54
Redistribution	-3.40	9.41	22.37	7.07
Revolutions and coups	0.00	2.40	23.00	4.51
Rule of law	-1.23	0.39	1.96	0.95
Technological progress	-1.32	0.37	1.29	0.66
Value added in agriculture	0.41	12.26	45.27	11.79
Value added in industry	16.15	30.71	51.29	6.79

Table B7: Correlation matrix

	GiniWealth	NatRes	PopGrowth	GovExp	NNSavings	EducExp	Infl	VAI	VAA	NetFDI	RuleWB	GDP90	Ygrowth	LifeExp90	LFPart90	PopDens90	RevCoups	WarYears	EcoFreedom	FinLib	CLandPR	OutwardO	LatAm	ChinnIto	LeftWing	ActivRestrict	CapitalReg	DiversIndex	LAMRIG	Tech	EducIndex	FID	FIA	FIE	FMD	FME	BussCond							
NatRes	0.35																																											
PopGrowth	0.24	0.44																																										
GovExp	-0.19	-0.32	-0.43																																									
NNSavings	0.34	0.14	0.42	-0.36																																								
EducExp	-0.03	-0.10	-0.19	0.58	-0.20																																							
Infl	0.22	0.06	-0.06	-0.12	-0.14	-0.11																																						
VAI	0.35	0.24	-0.18	0.00	0.32	0.03	0.20																																					
VAA	-0.08	0.41	0.52	-0.49	-0.02	-0.29	0.04	-0.37																																				
NetFDI	-0.21	-0.12	-0.29	0.29	-0.02	0.14	-0.02	0.13	-0.25																																			
RuleWB	-0.13	-0.42	-0.41	0.48	-0.02	0.26	-0.36	-0.04	-0.64	0.25																																		
GDP90	-0.04	-0.45	-0.66	0.56	-0.18	0.36	-0.16	0.20	-0.85	0.26	0.75																																	
Ygrowth	-0.09	-0.09	0.02	-0.18	0.43	-0.27	-0.26	0.20	-0.08	0.08	0.30	-0.01																																
LifeExp90	-0.07	-0.54	-0.59	0.44	-0.07	0.33	-0.16	0.19	-0.83	0.24	0.67	0.90	0.04																															
LFPart90	-0.17	-0.15	-0.06	0.18	-0.11	0.14	-0.07	-0.05	-0.09	0.11	0.23	0.19	0.07	0.17																														
PopDens90	0.00	-0.13	0.12	-0.21	0.44	-0.19	-0.09	0.00	-0.09	0.44	0.14	0.08	0.27	0.08	0.00																													
RevCoups	0.25	0.32	0.33	-0.45	0.16	-0.15	0.51	0.12	0.19	-0.20	-0.46	-0.39	-0.14	-0.32	-0.13	-0.09																												
WarYears	0.32	0.17	0.30	-0.28	0.25	-0.32	-0.03	-0.06	0.28	-0.31	-0.26	-0.34	0.14	-0.31	-0.16	-0.02	0.02																											
EcoFreedom	-0.20	-0.46	-0.40	0.47	-0.13	0.24	-0.27	-0.05	-0.69	0.39	0.85	0.76	0.23	0.70	0.20	0.21	-0.40	-0.34																										
FinLib	-0.15	-0.35	-0.36	0.39	-0.18	0.31	-0.15	0.02	-0.59	0.34	0.65	0.68	0.04	0.64	0.13	0.10	-0.21	-0.38	0.81																									
CLandPR	0.10	0.43	0.55	-0.44	0.17	-0.33	0.14	-0.05	0.67	-0.06	-0.76	-0.76	-0.04	-0.68	-0.22	0.14	0.23	0.27	-0.67	-0.63																								
OutwardO	0.41	0.50	-0.02	-0.14	0.16	0.06	0.07	0.42	0.03	-0.16	-0.01	0.04	0.01	-0.10	-0.16	-0.22	0.01	-0.06	-0.12	0.53	-0.10	-0.16	0.05	0.05	0.11																			
LatAm	0.28	0.14	0.23	-0.38	0.05	-0.04	0.41	0.20	-0.07	-0.11	-0.34	-0.15	-0.22	0.01	-0.06	-0.12	0.53	-0.10	-0.16	0.05	0.05	0.11																						
ChinnIto	-0.13	-0.29	-0.42	0.35	-0.11	0.26	-0.13	-0.03	-0.54	0.36	0.63	0.65	0.10	0.59	0.04	0.14	-0.24	-0.32	0.74	0.84	-0.59	-0.09	-0.09																					
LeftWing	-0.13	-0.12	-0.18	0.18	-0.17	0.23	-0.06	-0.10	-0.15	-0.05	0.26	0.13	-0.08	0.15	-0.15	-0.13	-0.17	-0.06	0.08	-0.02	-0.24	0.06	0.03	-0.05																				
ActivRestrict	0.00	0.22	0.49	-0.31	0.11	-0.09	-0.02	-0.10	0.42	-0.23	-0.43	-0.53	0.22	-0.45	0.00	-0.02	0.27	0.15	-0.40	-0.37	0.41	-0.11	0.24	-0.50	-0.09																			
CapitalReg	0.02	-0.01	0.23	-0.08	0.14	-0.01	0.00	-0.15	0.23	-0.02	-0.19	-0.28	0.05	-0.30	-0.04	0.08	0.16	0.22	-0.26	-0.19	0.29	-0.09	-0.12	-0.25	-0.15	0.27																		
DiversIndex	-0.04	-0.19	-0.20	0.23	-0.05	0.25	-0.10	0.12	-0.36	0.20	0.27	0.35	-0.06	0.32	-0.05	0.05	-0.24	-0.06	0.29	0.40	-0.25	-0.05	-0.02	0.41	0.01	-0.30	-0.04																	
LAMRIG	-0.08	-0.10	-0.27	0.15	-0.15	0.05	0.20	0.06	-0.05	0.03	-0.20	0.00	-0.21	0.06	0.06	-0.18	0.12	-0.13	-0.17	-0.10	-0.01	-0.05	0.21	-0.07	0.09	-0.03	0.05	-0.12																
Tech	-0.07	-0.44	-0.67	0.60	-0.25	0.40	-0.10	0.17	-0.82	0.33	0.72	0.93	-0.08	0.88	0.16	0.04	-0.37	-0.40	0.73	0.67	-0.68	-0.02	-0.12	0.61	0.17	-0.49	-0.25	0.40	0.02															
EducIndex	-0.11	-0.37	-0.69	0.58	-0.28	0.40	-0.02	0.19	-0.75	0.34	0.70	0.87	-0.03	0.80	0.15	-0.01	-0.31	-0.29	0.72	0.72	-0.73	0.05	-0.14	0.68	0.12	-0.54	-0.32	0.33	-0.01	0.88														
FID	0.08	-0.26	-0.24	0.37	0.06	0.23	-0.24	0.01	-0.64	0.21	0.81	0.71	0.15	0.63	0.08	0.18	-0.32	-0.20	0.74	0.56	-0.61	0.12	-0.22	0.56	0.13	-0.46	-0.25	0.30	-0.29	0.67	0.61													
FIA	-0.20	-0.41	-0.52	0.44	-0.19	0.18	-0.16	-0.03	-0.71	0.16	0.71	0.79	0.09	0.72	0.23	-0.01	-0.30	-0.30	0.66	0.55	-0.73	-0.07	-0.18	0.54	0.13	-0.48	-0.20	0.24	0.06	0.74	0.69	0.73												
FIE	-0.18	-0.19	0.03	0.16	0.41	0.06	-0.48	-0.05	-0.29	0.10	0.52	0.20	0.37	0.27	0.05	0.19	-0.22	-0.13	0.38	0.17	-0.26	-0.02	-0.29	0.23	0.17	-0.19	-0.11	0.13	-0.09	0.16	0.09	0.48	0.29											
FMD	0.19	-0.19	-0.16	0.23	0.19	0.13	-0.27	0.03	-0.55	0.16	0.73	0.64	0.15	0.56	0.03	0.25	-0.26	-0.03	0.63	0.48	-0.50	0.16	-0.28	0.53	0.05	-0.49	-0.14	0.28	-0.30	0.59	0.53	0.91	0.62	0.45										
FME	0.02	-0.35	-0.36	0.27	-0.01	0.13	-0.23	-0.02	-0.39	0.10	0.46	0.52	0.07	0.45	-0.10	0.12	-0.24	-0.01	0.39	0.30	-0.38	0.14	-0.35	0.30	0.14	-0.38	0.07	0.26	0.04	0.50	0.44	0.51	0.47	0.12	0.58									
BussCond	0.16	0.39	0.47	-0.31	0.16	-0.11	0.26	0.18	0.33	-0.25	-0.51	-0.52	-0.22	-0.50	-0.13	-0.16	0.48	0.07	-0.56	-0.36	0.30	0.18	0.45	-0.38	0.03	0.29	0.14	-0.17	0.22	-0.50	-0.47	-0.42	-0.37	-0.19	-0.41	-0.42								
Redist	-0.33	-0.35	-0.64	0.67	-0.42	0.35	-0.28	-0.11	-0.49	0.31	0.63	0.63	-0.06	0.51	0.11	-0.08	-0.46	-0.30	0.61	0.51	-0.61	-0.02	-0.42	0.51	0.22	-0.40	-0.23	0.27	0.12	0.63	0.64	0.50	0.59	0.24	0.38	0.49	-0.47							

Table B8: Expected effects of the explanatory variables

Access to financial institutions	+	/	-	Financial markets efficiency	-	Number of war years	+
Active banking restrictions	+	/	-	Financial openness (Chinn-Ito)	+	Outward orientation	+
Average GDP growth	+	/	-	GDP level in 1990	+	Population density	-
Bank capital regulations	+	/	-	Government expenditures	+	Population growth	-
Banking diversification	+	/	-	Inflation	+	Public education expenditures	-
Business conditions	-			Labour force participation	-	Redistribution	-
Civ. liberties and Pol. rights	+			Labour market regulations	+	Revolutions and coups	+/ -
Economic freedom index (adjusted)	-			Latin America dummy	+	Rule of law	+/ -
Education index (UN)	-			Leftwing orientation	-	Technological progress	+/ -
Financial institutions depth	+	/	-	Life expectancy	-	Value added in agriculture	-
Financial institutions efficiency	-			Natural resources rents	+	Value added in industry	-
Financial liberalization (EFW)	+	/	-	Net foreign direct investment	+		
Financial market depth	+	/	-	Net national savings	+		

Top models by their posterior model probability, group PIPs

Table B9: Top 3 models according to their posterior mode probabilities

Variable	Model 1	Model 2	Model 3
Access to financial institutions	1	1	1
Value added in agriculture	1	1	1
Financial institutions efficiency	1	1	1
Outward orientation	1	1	1
Financial market depth	1	1	1
Education index (UN)	1	1	1
War years	1	1	1
Redistribution	1	0	1
Average GDP growth	0	0	1

Note: 1 marks inclusion of the variable in the model, whereas 0 suggests otherwise. The variables not listed were not included in neither of the models.

Table B10: Group posterior inclusion probabilities

Group	PIP
Financial	1.00
Economic	1.00
Political	0.85
Institutional	0.70
Geographical	0.65
Regulatory	0.34

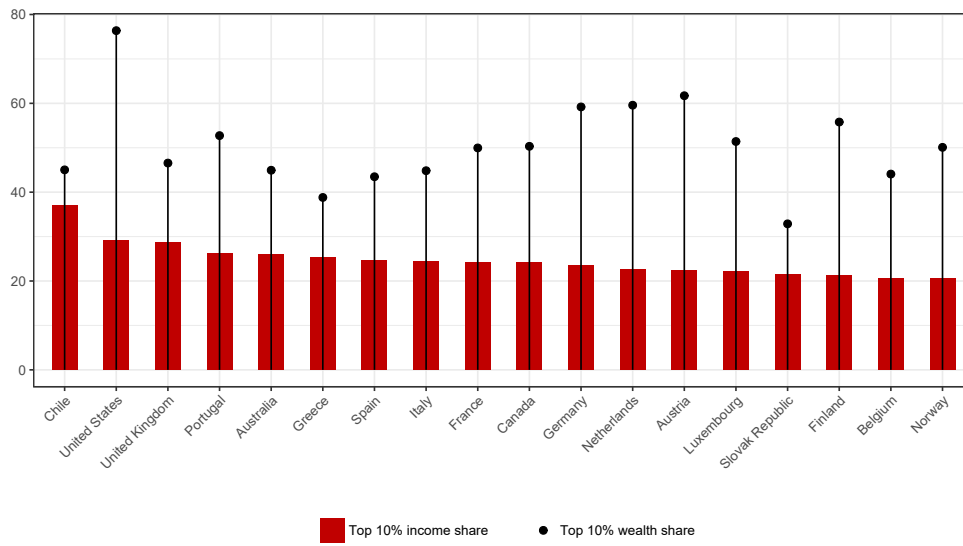
OLS estimates of the restricted models

Table B11: Output of the linear regression specifications, dependent variable GiniWealth

	(1)	(2)
Access to financial institutions	-0.376*** (0.140)	-0.411*** (0.140)
Value added in agriculture	-0.637*** (0.141)	-0.626*** (0.143)
Financial institutions efficiency	-0.356*** (0.100)	-0.377*** (0.100)
Outward orientation	0.319*** (0.086)	0.320*** (0.087)
Financial markets depth	0.470*** (0.124)	0.522*** (0.121)
Education index	-0.388** (0.157)	-0.413** (0.158)
Number of war years	0.146 (0.091)	
Redistribution	-0.213* (0.114)	-0.230* (0.115)
Observations	73	73
R ²	0.574	0.556
Adjusted R ²	0.520	0.509
Residual Std. Error	0.693 (df = 64)	0.701 (df = 65)
F Statistic	10.761*** (df = 8; 64)	11.645*** (df = 7; 65)

Note: *p<0.1; **p<0.05; ***p<0.01. The specification of the model (1) corresponds to the model with the highest posterior model probability, whereas the model (2) contains the regressors with PIP > 0.5 in the baseline BMA estimation.

Figure B1: Top 10% wealth and income shares in OECD countries



Note: Source: Author based on the OECD

Bayesian Model Averaging

First, consider the following linear model:

$$y = \alpha + X\beta + \varepsilon \quad \varepsilon \sim N(0, \sigma^2 I) \quad (\text{B1})$$

where y represents a dependent variable, α is a constant, X is the matrix of explanatory variables, β represents the corresponding coefficients, and ε is a vector of normally distributed IID error terms with variance σ^2 .

BMA takes into consideration all possible combinations of X from equation B1 and takes a weighted average of the estimated coefficients. Even with a modest-sized regression model, the number of combinations rises dramatically, and even with current computers, it is impossible to estimate all regression models. For this reason, a subset of models is considered, and an MCMC sampler is employed (we discuss the sampler in detail below). The substructure of the model is as follows:

$$y = \alpha_i + X_i\beta_i + \varepsilon \quad \varepsilon \sim N(0, \sigma^2 I) \quad (\text{B2})$$

X_i corresponds to a subset of X , and α_i and β_i are the corresponding coefficients. If the number of regressors is K , the total number of models equals 2^K , and $i \in [1, 2^K]$.

Bayes' rule implies that

$$p(\beta|y, X) = \frac{p(y, X|\beta)p(\beta)}{p(y, X)} \quad (\text{B3})$$

where $p(\beta|y, X)$ is the posterior density, $p(y, X|\beta)$ is the marginal likelihood (ML), $p(\beta)$ is the prior density, and $p(y, X)$ is the probability of the data.

The individual regression models are denoted as M_1, \dots, M_i . In the case of K regressors, there are M_1, \dots, M_i regression models, where $i \in [1, 2^K]$. The model is formed using a likelihood function and a prior density, where M_i depends on the parameters β_i , with a posterior probability to be derived in the following manner:

$$p(\beta_i|M_i, y, X) = \frac{p(y|\beta_i, M_i, X)p(\beta_i|M_i)}{p(y|M_i, X)} \quad (\text{B4})$$

Next, we describe the averaging principle of BMA and individual components of equation B3.

Posterior Model Probability

The PMP provides the weights for averaging model parameters across the individual models. The PMP also arises from Bayes' theorem:

$$p(M_i|y, X) = \frac{p(y|M_i, X)p(M_i)}{p(y|X)} \quad (\text{B5})$$

where $p(y|M_i, X)$ is the marginal likelihood (ML) of the model (i.e., the probability of the data given the model M_i), $p(M_i)$ is the prior model probability, and $p(y|X)$ is the integrated likelihood. The term in the denominator is typically disregarded because it is constant across all models under consideration. The PMP then becomes directly proportional to ML and the prior probability. The prior probability $p(M_i \propto 1)$ is typically set to acknowledge that the 'true' model is unknown.

$$p(M_i|y, X) \propto p(y|M_i, X)p(M_i) \quad (\text{B6})$$

We discuss the calculation of ML in detail in section 6. Researchers must set the model prior to reflect the beliefs regarding the data before inspecting them.

Posterior Mean

The parameter point estimates are derived within the Bayesian framework as follows. Zeugner (2011) and Moral-Benito (2012) show that the weighted posterior distribution of any statistic (most notably the β coefficients) is obtained as follows:

$$p(\beta|y, X) = \sum_{i=1}^{2^K} p(\beta_i|M_i, y, X)p(M_i|y, X) \quad (\text{B7})$$

where $p(M_i|y, X)$ is the PMP of the corresponding model M_i from equation B5. The point estimates are obtained by taking expectations:

$$E(\beta|y, X) = \sum_{i=1}^{2^K} E(\beta_i|M_i, y, X)p(M_i|y, X) \quad (\text{B8})$$

$E(\beta|y, X)$ represents the average coefficient, and $E(\beta|M_i, y, X)$ is the estimate of the β_i coefficients from model M_i . The posterior distribution of the coefficients depends on the choice of the prior g . Zeugner (2011) expresses the expected value of the parameter in M_i as follows:

$$E(\beta_i|y, X, g, M_i) = \frac{g}{1+g} \hat{\beta}_i \quad (\text{B9})$$

with $\hat{\beta}_i$ corresponding to the standard OLS estimate.

Posterior Variance

Moral-Benito (2012) provides a formula for the variance corresponding to the expected values of the coefficients derived in the previous subsection:

$$\begin{aligned} \text{Var}(\beta|y, X) &= \sum_{i=1}^{2^K} p(M_i|y, X) \text{Var}(\beta_i|M_i, y, X) \\ &+ \sum_{i=1}^{2^K} p(M_i|y, X) (E(\beta_i|M_i, y, X) - E(\beta|y, X))^2 \end{aligned} \quad (\text{B10})$$

The variance consists of two terms: the weighted average of variance estimates across different models $\text{Var}(\beta_i|M_i, y, X)$ and the weighted variance across different models in the second component $E(\beta_i|M_i, y, X) - E(\beta|y, X)$. $E(\beta|y, X)$ represents the posterior mean from equation B8. As a result, BMA accounts for uncertainty regarding the parameter estimates that arise due to differences across models in addition to the uncertainty of individual models. Zeugner (2011) derives how the value of the prior g affects the posterior variance of the parameters:

$$\text{Cov}(\beta_i|y, X, g, M_i) = \frac{(y - \bar{y})'(y - \bar{y})}{N - 3} \frac{g}{1 + g} \left(1 - \frac{g}{1 + g} R_i^2 \right) (X_i' X_i)^{-1} \quad (\text{B11})$$

where \bar{y} denotes the mean of vector y , N is the sample size, and R_i^2 is the R-squared value corresponding to the model i .

Marginal Likelihood

ML can be calculated using equation B4 for each model M_i . Both sides of the equation must be integrated with respect to β_i . Employing $\int_{\beta} p(\beta_i|M_i, y, X) d\beta_i = 1$, it follows that

$$p(y|M_i, X) = \int_{\beta} p(y|\beta_i, M_i, X) p(\beta_i|M_i, X) d\beta_i \quad (\text{B12})$$

The above equation illustrates the general textbook derivation, but the computation depends on the elicited priors. Zeugner (2011) employs the ‘‘Zellner’s g prior’’ structure, which we also utilize in this paper. The ML for a single model can then be expressed

using the prior as in [Feldkircher and Zeugner \(2009\)](#):

$$p(y|M_i, X, g) = \int_0^\infty \int_\beta p(y|\beta_i, \sigma^2, M_i) p(\beta_i, \sigma^2|g) d\beta d\sigma \quad (\text{B13})$$

Furthermore, [Feldkircher and Zeugner \(2009\)](#) show that ML is in this case simply proportional to

$$p(y|M_i, X, g) \propto (y - \bar{y})'(y - \bar{y})^{-\frac{N-1}{2}} (1 + g)^{-\frac{k_i}{2}} \left(1 - \frac{g}{1+g} R_i^2\right)^{-\frac{N-1}{2}} \quad (\text{B14})$$

In this equation, R_i^2 is the R-squared of model M_i , and k_i is the number of explanatory variables in model i introduced to include a size penalty for the model. N and \bar{y} are the same as in equation [B11](#), i.e., the number of observations and the mean of vector y , respectively.

Posterior Inclusion Probability

The standard BMA framework provides the PIP, which indicates the probability that a particular regressor is included in the “true” model. The PIP is the sum of the PMPs of the models including the variable k :

$$PIP = p(\beta_k \neq 0|y, X) = \sum_{i=1}^{2^K} p(M_i|\beta_k \neq 0, y, X) \quad (\text{B15})$$

MCMC Sampler

One of the limitations of BMA is its computational difficulty when the number of potential regressors K becomes very large. Historically, the computational burden has been the primary factor preventing researchers from employing Bayesian methods. [Zeugner \(2011\)](#) notes that for small models, it is possible to enumerate all variable combinations. However, when $K > 25$, it becomes impossible to evaluate the entire model space within a reasonable time frame. In such cases, BMA utilizes MC³ samplers to approximate the crucial part of the posterior model distribution containing the most likely models. BMA applies the Metropolis-Hastings algorithm, which is outlined in [Zeugner \(2011\)](#) as follows:

At any step i , the sampler is currently at model M_i , having PMP $p(M_i|y, X)$. In the next step $i + 1$, model M_j is proposed to replace M_i . The sampler accepts the new

model M_j with the following probability:

$$p_{i,j} = \min \left(1, \frac{p(M_j|y, X)}{p(M_i|y, X)} \right) \quad (\text{B16})$$

If model M_j is rejected, the next model M_k is suggested and compared with M_i . With an increasing number of iterations, the number of times each model is retained converges to the distribution of posterior model probabilities. Typically, one of the following MC³ samplers is used to construct the models:

- Birth-death sampler - randomly chooses one of the explanatory variables, which is included if it is not already part of the current model M_i or dropped if it is already in M_i .
- Reversible-jump sampler - with 50% probability, the birth-death sampler is used to determine the next candidate model. With 50% probability, the sampler randomly swaps one of the covariates in M_i for a covariate previously excluded from M_i .

Because the sampler can begin with a “poor” model with low PMP, the predefined number of initial draws, the so-called burn-ins, are usually dropped. The quality of the approximation can be evaluated on the basis of the correlation between the PMP derived from an analytical approach and those obtained from the MC³ sampler. It depends on the number of iterations (draws) and the likelihood of the initially selected model. [Zeugner \(2011\)](#) notes that a PMP correlation of approximately 0.9 indicates a “good degree of convergence”. In the event that the correlation is lower, the number of sampler iterations should be increased.