ELSEVIER

Weak instruments and weak identification, in estimating the effects of education, on democracy

Matteo Bobba [a], Decio Coviello [b],*

[a] *Inter-American Development Bank, United States*
[b] *European University Institute, Italy*

Abstract

Is there any relation between education and democracy? Once we correct for weak instruments and identify education as "weakly exogenous" we find new evidence that education systematically predicts democracy. Our results are robust across model specification, instrumentation strategies, and samples.
© 2007 Elsevier B.V. All rights reserved.

1. Introduction

The relationship between education and democracy has been at the heart of both academic and political debate of the last decade. Recent empirical work has reached mixed results. Glaeser et al. (2004), using OLS, present evidence that countries' level of education predict changes in democracy. Nonetheless, Acemoglu et al. (2005) demonstrate the fragility of those results. Both analyses consider past levels of education exogenous to countries' levels of democracy.

---

* Corresponding author. Department of Economics, European University Institute, Via dei Roccettini, 9 50016 San Domenico di Fiesole (FI) Florence-Italy. Tel.: +39 348 9523474; fax: +39 055 4685 902.
  E-mail addresses: matteob@iadb.org (M. Bobba), decio.coviello@eui.eu (D. Coviello).

In this paper we argue that (1) given the forward-looking nature of investments in human capital, education is endogenous and thus its effects on democracy are weakly identified and (2) given the high persistence in democracy and education, the findings of Acemoglu et al. (2005) are subject to weak identification and weak instruments problems. Indeed, we show that (lagged) levels of education systematically predict changes in democracy by considering a different identification assumption (weak exogeneity) for education and by using additional and more informative moment conditions to instrument all the regressors.

2. Identification assumptions

We consider the following dynamic specification to disentangle the relationship between democracy and education:[1]

$$d_{it} = \alpha\, d_{it-1} + \gamma s_{it-1} + \beta \mathbf{X}_{it-1} + \eta_i + \delta_t + \varepsilon_{it} \tag{1}$$

where d_{it} is the democracy (level) of country i in period t, s_{it-1} is the lagged value of average years of schooling, $\mathbf{X}_{it-1}$ is a vector of (non-contemporaneous) control variables, and η_i and δ_t denote a full set of country effects and time effects. In order to assess the effects of education on democracy, Acemoglu et al. (2005) estimate Eq. (1) in first differences and employ lagged levels of the regressors as instruments for the equation in first differences.[2]

The so-called "Difference" GMM estimator relies upon the following orthogonality conditions:

$$E\big(d_{i,t-s}\varDelta\varepsilon_{it}\big) = 0 \quad \text{for } t = 3, ..., T \text{ and } s \geq 2 \tag{2}$$

where $d_{i,t-s}$ represents the instruments set used in this GMM estimator. In this setting, it is well known that the higher the persistence of the series used as instruments, the lower the correlation between levels and subsequent differences.[3]

It is widely recognized that both democracy and education are highly persistent[4], therefore lagged levels are weak instruments for the differences and it is possible to gain precision in terms of point estimates bias by exploiting some additional moment restrictions.[5] The so-called "System" GMM estimator stacks together the equation in first differences and the equation in levels in a system of equations and employs both lagged levels and differences as internal instruments.

---

[1] This specification is similar to that presented in Glaeser et al. (2004) where the dependent variable is the change in democracy.

[2] The first differences are a standard way to deal with country effects; see Arellano and Bond (1991).

[3] Simulation results show that the Difference GMM may be subject to a large downward finite-sample bias in this case, particularly when T is small. This weak instrument argument can be seen by considering the following transformation of a simple AR(1) process $\Delta d_{it} = (\alpha - 1)d_{it-1} + \varepsilon_{it}$. The higher the persistence of α, the lower $E(\Delta d_{it} d_{it-1})$. See Blundell and Bond (1998).

[4] Our preferred estimates of the univariate AR(1) processes deliver estimates of the autoregressive coefficient for Democracy and Education that are respectively 0.66 and 0.93, thus both very persistent; see Blundell and Bond (1998). Those results are not presented because of space limitations but are available on request.

[5] See Blundell and Bond (1998) for simulation results.

In order to consider the additional moments as valid instruments for Eq. (1), the following additional linear moment conditions must be satisfied:

$$E\big(\Delta d_{i,t-1}(\eta_1 + \varepsilon_{it})\big) = 0 \quad \text{for } t = 4, 5...T. \tag{3}$$

Condition (3) implies that changes in democracy are orthogonal to the country fixed effect. We test the validity of this assumption.

Our identification strategy differs from previous empirical work where past education has always been assumed to be exogenous. We consider this assumption to be strong since education can be viewed as a forward-looking variable: people take into account the expected future level of political development when investing in human capital. With "internal instruments" we can control for a weak form of exogeneity in education (and in other covariates) only by assuming that our explanatory variables can be affected by current and past realizations of democracy but must be uncorrelated with future unpredictable innovations in democracy (the error term). The dataset employed in our analysis is the same used by Acemoglu et al. (2005).[6]

3. Results

In a multivariate panel data framework it is not clear how to test for weak instruments, hence we use the known bias in the Difference GMM by comparing its sample performances with alternative estimators with known properties in dynamic panel data and test whether our preferred estimator, the System GMM, improves the precision of the estimates.[7]

Table 1 reports the main results of estimation of model (1) across various estimators using the baseline sample 1965–2000. Column 1 and 2 show the results of Within Groups and Pooled OLS estimators that provide, respectively, the lower and upper bound for the autoregressive coefficient of democracy.[8]

Columns 3 and 4 employ alternatively one and two step Difference GMM estimators[9]; the estimated autoregressive coefficient is smaller than or equal to the corresponding Within Groups estimate, which should be the lower bound in a short panel like this one. We then conclude that this estimator is likely to be seriously downward biased, and the consequent finding that education is not statistically different from zero (and sometimes negative) under both specifications is not instructive. Columns 5 and 6 report the System GMM estimates, one and two step, respectively, and the results are striking: the estimated autoregressive coefficient lies between the two bounds, and the lagged level of education now has a positive and significant effect on democracy at the 1-percent significance level. The point estimate is 0.099, implying that an additional year of schooling increases the steady-state value of democracy by 18 percentages points.[10] To address the validity of the additional moment conditions (3) we run an

---

[6] Because of space restrictions we do not describe them here and refer the reader to the original paper. Our dataset is a five-year panel with 108 countries spanning the period between 1965 and 2000.

[7] This procedure is known as the Bounding Procedure. For an extended discussion see Bond et al. (2001).

[8] See Bond (2002) for details on the bias of the two bounds.

[9] Standard Wald tests based on two-step efficient generalized method of moments (GMM) estimators are known to have poor finite sample properties in several contexts; see Blundell and Bond (1998). In those contexts the problem worsens when there are weak instruments. We improve our inference in the two-step GMM by implementing the Windmeijer (2005) correction using xtabond2, and we compare the Wald tests with the one-step robust version of the same estimates. See Windmeijer (2005).

[10] The long-run effect is calculated as $(\gamma/1 - \alpha) * 100$.

Table 1
Bounding procedure: Results for baseline model across various estimators

Dependent variables: democracy (t)	Within group	Pooled OLS	Diff-1 GMM	Diff-2 GMM	Sys-1 GMM	Sys-2 GMM	Sys-1 NO-OECD	Sys-2 NO-OECD
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Democracy ($t-1$)	0.387***	0.703***	0.319**	0.399**	0.450***	0.506***	0.422***	0.484***
	(−7.05)	(−19.51)	(−2.06)	(−2.37)	(−4.52)	(−5.1)	(−4.12)	(−4.67)
Education ($t-1$)	−0.005	0.027***	−0.066	−0.024	0.099***	0.062**	0.128***	0.10**
	(−0.25)	(7.17)	(−0.66)	(−0.28)	(3.40)	(2.43)	(3.16)	(2.54)
Constant	0.366***	0.061***			−0.009	0.076	−0.155*	−0.108
	(5.56)	(3.04)			(−0.14)	(1.62)	(−1.83)	(−1.47)
Time effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Hansen			0.49	0.49	0.29	0.29	0.50	0.50
Diff Hansen					0.19	0.19	0.33	0.33
A R(1)			0.00	0.00	0.00	0.00	0.00	0.00
A R (2)			0.69	0.92	0.73	0.73	0.80	0.76
Obs.	775	775	667	667	775	775	601	601
N	108	108	104	104	108	108	86	86

Notes: Dependent variables is the Augmented Freedom House Political Rights Index. Diff-1 GMM and Diff-2 GMM are the one (two) step difference GMM estimation. SYS-1 (-2) GMM are the one (two) step system GMM estimation. Robust standard error are computed for the t-tests reported in "()". For one step estimates are computed Huber–White standard errors, while the two step estimates are Windmeijer corrected. * significant at 10%, ** at 5% and *** at 1%. The values reported for the Hansen test are the p-values for the null hypothesis of instrument validity. The Diff Hansen reports the p-value for the validity of the additional moment restrictions required by the SYS GMM. The values reported for A R (1) and A R (2) are the p-values for first and second order auto correlated disturbances in the first differences equations. Five-year panel, between 1965–2000, as in Acemoglu et al. (2005).

incremental Sargan test for over-identification restriction based on the difference in the Sargan tests between the System and Difference estimator.[11] The corresponding Chi-square statistic does not reject the null of validity of condition (3). A possible concern with this last result is that it may be driven by the presence in our sample of industrial countries in which the level of democracy is very high and persistent across our time span, hence the change in democracy is zero for these countries and condition (3) holds trivially. In columns 7 and 8 we exclude industrial countries from the base sample and estimation results confirm both the validity of condition (3) and the positive and significant effect of education on subsequent democracy.

Note that the estimated coefficient of education might include both direct and indirect effect of education on democracy working through income (or other possible channels); this motivates us to perform various robustness checks of our results by adding standard covariates, as per capita GDP (in logs), investments and population (in logs).[12] Table 2 presents these additional results. Columns 1 to 3 add each covariate and show that none of these variables is significant, while our variable of interest, education, remains significant at the 1-percent level, which is not true for GDP. In order to disentangle the effect of education on democracy, in Column 4 we consider all three covariates simultaneously and

---

[11] See Blundell and Bond (1998).

[12] These are the same variables that appear in Acemoglu et al. (2005) as a robustness check.

Table 2
Robustness checks, additional covariates

Dependent variables: democracy (t)	Sys-2 GMM	Sys-2 GMM	Sys-2 GMM	Sys-2 GMM	Sys-2 GMM
	(1)	(2)	(3)	(4)	(5)
Democracy $(t-1)$	0.562***	0.554***	0.572***	0.556***	−0.444***
	(8.10)	(7.68)	(6.98)	(8.43)	(6.74)
Education $(t-1)$	0.025	0.05***	0.043***	0.029**	0.029**
	(1.42)	(5.31)	(3.37)	(2.29)	(2.29)
Ln (rgdpch) $(t-1)$	0.042			0.035	0.035
	(1.16)			(0.95)	(0.95)
Ln (pop) $(t-1)$		−0.032		−0.013	−0.013
		(−1.33)		(−0.92)	(−0.92)
Ln (ki) $(t-1)$			−0.047	0.154	0.154
			(−0.21)	(0.79)	(0.79)
Constant	−0.156	0.37*	0.022	−0.118	−0.118
	(−0.65)	(1.75)	(0.43)	(−0.41)	(−0.41)
Time effect	Yes	Yes	Yes	Yes	Yes
Hansen	0.27	0.25	0.36	0.99	0.99
Diff Hansen	0.39	0.52	0.8	0.99	0.99
AR(1)	0.00	0.00	0.00	0.00	0.00
AR(2)	0.81	0.75	0.86	0.84	0.84
Obs.	692	763	696	685	685
N	97	106	97	96	96

Notes: Dependent variables is the Augmented Freedom House Political Rights Index. SYS-2 GMM are the two step system GMM estimates. Windmeijer corrected standard error are computed for the *t*-tests reported in "()". * significant at 10%, ** at 5% and *** at 1%. The values reported for the Hansen test are the *p*-values for the null hypothesis of instrument validity. The Diff Hansen reports the *p*-value for the validity of the additional moment restrictions required by the SYS GMM. The values reported for AR(1) and AR(2) are the *p*-values for first and second order autocorrelated disturbances in the first differences equations. Five-year panel, between 1965–2000, as in Acemoglu et al. (2005).

find that the estimated coefficient of education remains highly significant, with a point estimate slightly lower (0.029) than the baseline case of Table 1. We interpret this result as evidence of the primacy of the direct effect of human capital on political institutions and supportive of the interpretation that the indirect effect working from income to democracy is negligible once we take human capital into account.

In order to further stress this argument, we run a regression of the five-year change of our democracy index on past levels of education and three standard covariates. The results are reported in column 5 and are entirely consistent with the findings of Glaeser et al. (2004), but they are now robust to weak instruments and to considering education as weakly exogenous.[13] Levels of schooling appear to be a strong predictor of improving institutional outcomes, while per capita income has no predictive power.[14]

---

[13] See in particular Glaeser et al. (2004), Table 12, panel B.
[14] Due to space limitations we omit comments on all the Sargan tests and AR (1) and AR(2) specification tests which in Table 2 confirm the correct specification of all the estimated models considered.

4. Conclusions

We have revisited the nature of the relationship between political institutions and investment in human capital by taking into account two sources of bias: weak instruments and endogeneity. Using an alternative estimator, a different identification assumption, and different sub-samples we have found evidence of a statistically significant relationship between past levels of education and levels and changes of democracy.

References

Acemoglu, D., et al., 2005. From education to democracy? American Economic Review Papers and Proceedings 95, 44–49.
Arellano, M., Bond, S.R., 1991. Some specification tests for panel data: Monte Carlo evidence and an application to employment equations. Review of Economic Studies 58, 277–298.
Bond, S.R., 2002. Dynamic panel data models: a guide to micro data methods and practice. Portuguese Economic Journal 1, 141–162.
Bond, S.R., Hoeffler, A., Temple, J., 2001. GMM estimation of empirical growth models. CEPR discussion paper, 3048. Centre for Economic Policy Research, London, United Kingdom.
Blundell, R., Bond, S.R., 1998. Initial conditions and moment restrictions in dynamic panel data models. Journal of Econometrics 87, 115–143.
Glaeser, E.L., et al., 2004. Do institutions cause growth? Journal of Economic Growth 9, 271–303.
Windmeijer, F., 2005. A finite sample correction for the variance of linear two-step GMM estimators. Journal of Econometrics 126 (1), 25–51.