Assignment 4 – Panel Data
Advanced Econometrics 2

Stepan Svoboda - 11762616, Jan Hynek - 11748494

March 5, 2018

1 Report

1.1 Introduction

In this assignment we are asked to extend results from papers by Levine, Loayza and Beck (2000) and Aghion, Howitt and Mayer-Foulkes (2005). We are asked to use both standard panel data methods, i.e. least squares, and also alternative GMM estimation technique. Firstly we perform an analysis using the standard fixed effects model where we tackle the omitted variable bias of each country having a specific profile and specific growth background. We are explaining the growth using the financial development and several control variables. Secondly we investigated the possibility of the growth and financial development being endogenous by employing the GMM framework and using the lagged variables of our endogenous variables as instruments.

1.2 Methodology and motivation

We have approached each estimation in two ways, firstly we estimated the model with variables of interest, i.e. the financial development, growth, interaction term and initial GDP, secondly we added the control variables. We have followed in the first step the AHM paper and used only variable 'average private credit' as our proxy for financial development. The control variable that we chose were mainly in line with what AHM chose with some minor changes, we added the average secondary schooling in total population on top of the overall schooling. The reason for this deviation is that we believe that combination of these factors can better capture the overall development of the country since we are controlling for a possibility of a population subgroup being highly educated with another population having no education whatsoever. Including the secondary education level should capture and in tandem with the overall schooling level nicely explain the overall development which should be highly correlated with growth.

    We estimated the fixed effects model using the standard within model. In this part we assumed that endogeneity is only caused by omitted variable bias, which we have removed by the demeaning and that strong exogeneity applies. This a strong assumption but if we are assuming that there is no simultaneity it is not a big stretch to assume that the covariates we picked could be not only weakly exogenous but also strongly exogenous. This assumption may be possibly violated but we believe it is a reasonable assumption.

    After this we decided to explore the possibility of simultaneity and estimated the same model using GMM with lags of our endogenous variables as instruments. Our extension of this model is to include lagged value of growth as an explanatory variable and making the model a dynamic panel data model. We used the first-differences version based on the famous Arellano and Bond estimator. The reasoning behind is that we believe growth is sticky in some regards, e.g. if there is a boom many firms and individuals believe it will last forever and act as if they expect this growth will be forever. By including two lags of growth and taking as or instruments further lags, i.e. 3, 4 and lags of our endogenous variables,

we have created a model that has very convenient properties and seems to function well based on our results.

1.3   Results

The summary of our results is in the only table. The first two columns describe the first regression that used the within estimator only on our variables of interest. All of our estimates are significant (not all at 5%, but at least at 10%). The standard errors are obtained by using panel-robust inference to serial correlation and heteroskedasticity. In the second case we added several control variables and obtained slightly higher standard errors, as expected, and slightly shifted estimates of our coefficients. Those became either more positive or negative. Several of our control variables are also significant, including the secondary schooling in population, which in our opinion justifies is inclusion.

Table 1: Estimated coefficients and their standard errors

	rowname	β_{core}	SE_{core}	β_{full}	SE_{full}	β_{GMM}	SE_{GMM}	β_{GMM}^{full}	SE_{GMM}^{full}
1	privo	0.1555	0.0595	0.1724	0.0630	-0.1305	0.1515	-0.1358	0.1571
2	log(initial)	-0.0441	0.0065	-0.0543	0.0081	-0.0670	0.0308	-0.0592	0.0323
3	privo:log(initial)	-0.0151	0.0061	-0.0175	0.0064	0.0165	0.0176	0.0160	0.0172
4	school			-0.0034	0.0029			-0.0054	0.0058
5	gov			-0.1306	0.0675			-0.0555	0.0936
6	log1p(pi)			-0.0282	0.0084			-0.0408	0.0101
7	log1p(bmp)			-0.0045	0.0064			-0.0076	0.0076
8	sec			0.0085	0.0039			0.0023	0.0092
9	lag(growth, 1)					-0.4601	0.1912	-0.4173	0.1999
10	lag(growth, 2)					-0.0037	0.0923	-0.0415	0.1043

The second part of our results includes the GMM estimation. We decided unlike AHM and LLB to not use a cross-section analysis but to try and explain the growth by a dynamic panel model. We then assumed two endogenous variables, growth and private credit, and were explaining today's growth by the same variables as in the 'core' FE regression plus two lags of growth. This regression led to only the initial state of GDP and the first lag of growth being significant with both financial development term and the interaction term being insignificant. In the same regression with control variables added the initial level of GDP lost its significance at 5% level but still was significant at 10% level, first lag stayed significant and from the control variables inflation became significant. As our instruments we used third, fourth and fifth lag of growth and of private credit since two lags of growth were included in the model.

The second part of our results, which assumes endogeneity significantly shifted some of our estimates and made our standard errors much larger. The higher standard errors are not a surprise but unfortunately it led to loss of significance in quite a few of our variables.

1.4   Conclusion

We had two different approaches and in both we had a core specification of only variables of interest and a full specification of model with control variables included.

When we assumed no simultaneity we got nice and overall significant results but those were quite likely biased since we ignored a possible source of endogeneity. After including this possibility in our model we lost statistical significance in several of our variables but those that stayed significant throughout the analysis kept their sign and order of magnitude, which is encouraging. These were the logarithm of initial GDP and the first lag of growth.