Institutions and growth revisited: OLS, 2SLS, G2SLS random effects IV regression and panel fixed (within) IV regression with cross-country data

Dushko Josheski and Risto Fotov and Darko Lazarov and Cane Koteski

2. October 2011

Institutions and Growth revisited: OLS, 2SLS, G2SLS Random effects IV regression and Panel Fixed (within) IV regression with cross-country data

Dushko Josheski (dushkojosheski@gmail.com)

Darko Lazarov (darko.lazarov@ugd.edu.mk)

Risto Fotov (risto.fotov@ugd.edu.mk)

Cane Koteski (cane.koteski@ugd.edu.mk)

Abstract

This paper revisits the Institutions and growth models. Econometric techniques have been applied on cross-country data, just to confirm the *apriori* knowledge that Institutions effect on growth is positive and highly statistically significant. This evidence was confirmed by all four models. OLS proved as a better technique for our data than 2SLS, this simply because overidentification test showed that instrument cannot be considered exogenous, also Hausman test showed that OLS is better than 2SLS at 1% and 5% levels of significance. G2SLS estimator and Fixed effects panel estimators just confirmed the results from the OLS and 2SLS. As a proxy variable for institutions we used Rule of law variable, also as instruments were used revolutions and Freedom house rating as well as War casualties variables. Also as conclusion here Trade is insignificant in influence to GDP growth compared with quality of institutions.

Key words: Institutions, Growth, 2SLS, OLS, G2SLS Random effects IV regression and Panel Fixed (within) IV regression, cross-country data, Hausman test, Overidentification test

Literature review of Institution and growth

The growth theory tries to explain the dynamic of growth process and the enormous differences of income per capita and economic performance among countries. From historical perspective, some group of countries have accomplished very high rate of growth and economic performance compared with other countries which face with economic problems (slowly dynamic of growth process). There are many explanations about this fact, basically, three theories analyze the factors which determinate cross-country differences in income levels and growth rate. First, the neoclassical theory of economic growth, based on work of Solow (1956), Lucas (1988), and others, focuses on the inputs of physical and human capital as a main resource of growth process, and late, Romer (1990) focus on technology advances through R&D activities (activities that create new ideas in economy) as a engine of growth. Second, the geographic/location theory explain that the geographic location of country (access to market) and the climate condition are very important for income level and economic performance. The theoretical and empirical research present the strong causality between the geographic location and the income level, the geographic/location theory explain only the income level differences among countries. In other side, the most important question for economist is the engine of growth, and in this direction the growth theory tries to explain the factors which determent the rate of growth. Third, the institutional approach emphasizes the importance of creating an institutional environment and institutions that support and encourage the main foundation of market economy (e.g. protection of property rights, rule of law, enforcement of contracts, and voluntary exchange of market-determined price. Institutions refer to rules, regulations, laws and policies that affect economic incentives such as incentives to invest in technology, physical capital and human capital. In this regard, the good institution framework is necessary for high level investment. Investors do not prefer to risk their capital when the protection of property rights is poorly, there are weak in rule of law and enforcement of contracts, and other illegal activities in market foundation economy. The theoretical explanations for growth that we introduced above are not inconsistent each other and all might play important role, but institutions are the major fundamental cause of economic growth and cross-country differences in economic performance.

The research of our paper focuses on the causality relationship between institutions and growth, and analyzes how quality of institutions influences growth rate. The empirical investigate show the more strong direction of causality of institutional quality to growth than the influence of growth to quality institutions. The explanation of this result is the fact that poor counties have more incentive to improve the quality of their institutions to achieve higher growth rate, rather than develop counties with high growth do not need to improve the institutional environment because that countries already have reached high-quality institutions.

Theoretical model of institutions, capital and economic growth

To develop the growth model with institutions, we start our analysis with aggregate production function which describes how the inputs (physical and human capital, labor and technology) are combined to produce output.[1]

$$Y_t = A_t K_t^{\alpha} H_t^{\beta} L_t^{1-\alpha-\beta} \text{ }^{2} \tag{1}$$

where Y is output, the parameter A represent the level of technology in economy, K is physical capital, H is human capital, and L is labor. We should make distinction between human capital and labor. The labor force is amount of people who are able to work, in the other side, human capital is the knowledge, skills and abilities of people who are or who may be involved in production process.

The equation of production function can write in per capita form:

$$\frac{Y_t}{L_t} = \frac{K_t^{\alpha}}{L_t} \frac{H_t^{\beta}}{L_t} \frac{A_t L_t^{1-\alpha-\beta}}{L_t} \tag{2}$$

$$y_t = A_t k_t^{\alpha} h^{\beta} \tag{3}$$

Traditional macroeconomic growth models do not include the influence of institutional quality as a factor of economic growth. These models implicitly assume an underlying set of good institutions. The fact that institutions have important role in growth process, the economists try to implement the institutional quality in growth models.

---

[1] The production function is characterize with constant return, $\alpha + \beta \leq 1$.

[2] The equation (1) we can write in this terms: $Y_t = K_t^{\alpha} H_t^{\beta} (A_t L_t^{1-\alpha-\beta})$.

$$A_t = A_0 k_t^{\delta_1 (In - In^*)} h_t^{\delta_2 (In - In^*)} \quad (4)$$

where A_0 represents the basic level of technology, In^* represents the best quality institutions, these ideal institutions are assumed in the traditional growth model, and In is the country's current level of institutional quality. The mathematical statement $(In - In^*)$ measures the degree to which the country's institutions fall short of the best conditions. The traditional growth model assume that economies function close to best-quality institutions, $In = In^*$, thus, these growth model reduce the influence of quality institutions.

Substituting the equation (3) into equation of production function per worker, we get:

$$y_t = A_0 k_t^{\delta_1 (In - In^*)} h_t^{\delta_2 (In - In^*)} k_t^{\alpha} h_t^{\beta} \quad (5)$$

Rewriting this equation we get:

$$y_t = A_0 k_t^{\alpha + \delta_1 (In - In^*)} h_t^{\beta + \delta_2 (In - In^*)} \quad (6)$$

To study the dynamic of output per capita, we will use a simple *mathematical trick* that economists often used in the study of growth.[3] The mathematical trick is to "take logs and then derivatives".

If we take logs of equation (6), we obtain:

$$\log y_t = \log A_0 + \left[\alpha + \delta_1 (In - In^*)\right] \log k_t + \left[\beta + \delta_2 (In - In^*)\right] \log h_t \quad (6)$$

Derivatives regarding time t, we obtain following form:

[3] Mathematical notes: The theory of growth uses some properties of natural logarithms. One of that properties is: The statement regarding the timing of the logarithms of a variable, gives the growth rate of that variable: If $y(t) = \log x(t)$, than, $\frac{dy}{dt} = \frac{dy}{dx}\frac{dx}{dt} = \frac{1}{x}\Delta x = \frac{\Delta x}{x}$.

$$\frac{d\log y_t}{dt} = \frac{d\log A_0}{dt} + \left[\alpha + \delta_1(In - In^*)\right]\frac{d\log k_t}{dt} + \left[\beta + \delta_2(In - In^*)\right]\frac{d\log h_t}{dt} \qquad (7)$$

As we can see, the equation (8), show the growth rate of output per capita:

$$\frac{\Delta y_t}{y_t} = \frac{\Delta A_0}{A_0} + \left[\alpha + \delta_1(In - In^*)\right]\frac{\Delta k_t}{k_t} + \left[\beta + \delta_2(In - In^*)\right]\frac{\Delta h_t}{h_t} \text{[4]} \qquad (8)$$

Rewriting equation (8) we get following form of growth rate of output per capita:

$$\frac{\Delta y_t}{y_t} = \frac{\Delta A_0}{A_0} + \left[(\alpha - \delta_1 In^*) + \delta_1 In\right]\frac{\Delta k_t}{k_t} + \left[(\beta - \delta_2 In^*) + \delta_2 In\right]\frac{\Delta h_t}{h_t} \qquad (9)$$

If we assume that: $\varphi_1 = (\alpha - \delta_1 In^*)$; $\varphi_2 = (\beta - \delta_2 In^*)$ and $\alpha_0 = \Delta A_0$, and adding an error term ε_t, we get final equation of growth rate of output per capita:

$$\frac{\Delta y_t}{y_y} = \alpha_0 + \varphi_1\frac{\Delta k_t}{k_t} + \delta_1 In\frac{\Delta k_t}{k_t} + \varphi_2\frac{\Delta h_t}{h_t} + \delta_2 In\frac{\Delta h_t}{h_t} + \varepsilon_t \qquad (10)$$

The final basic equation that we got in our theoretical model can use to test the impact of institution on the growth by the influence of institution's quality on the productivity of physical and human capital. In addition, we explain the coefficient estimates for $\varphi_1, \varphi_2, \delta_1, \delta_2$. The coefficient φ_1 and φ_2 measure the return to physical and human capital investments (the productivity of capital investments) in a country with the worst possible institutional quality, while coefficient δ_1 and δ_2 showing an increasing return to these capital investments as the country's institutional quality improves to the ideal level for economy based of market foundations.

Measuring problems with institutional quality and their influence of growth

In our theoretical model of institutions, capital and growth we can see that some parameters are relatively easy to measure, for example, K is amount of physical capital and H

[4] Where symbol, Δ, denotes changes of parameters.

is human capita that measure by years of schooling. On the other hand, institutions are not easily to quantifiable and this makes problem to measure the influence of institutions to economic growth. Economists try to solve the problem with measuring the quality of institutions by including some instrumental variables.

First, we will define the range of institutions and put some variables to measure different aspects of institutional environment. Institutions are the rule of game and it encompasses different type of social arrangements, laws, regulation, enforcement of property rights and so on. This definition of institutions is very widely and we can learn relatively little by emphasizing the importance of such a broad set of institutions. It is therefore important to try to understand what types of institutions are more important for economic growth. This is very useful for our empirical analysis of institutions and economic growth. There three type of institutions: political, financial and economic institutions. The quality of each of these type of institutions are measured through different variables. For example, the main variables for political institutions are: political rights and civil liberties that contain the political freedom index, rule of law that contain rule of law index, control of corruption and corruption freedom that contain index of corruption and other variables. On the other hand, the main variables of economic institutions are: protection of property rights, regulation and business freedom index that refer to trade freedom, freedom in doing business, financial freedom, investment freedom, and quality of regulation system.

The investigation of relative roles of different types of institutions is very important because as we can see above different type of institution have different influence of growth and economic performance. The economic institutions have the major role for growth, and in this regard when economist testified the relationship between institutions and growth, have to measure variables that cause quality of economic institutions more that quality of political institutions.

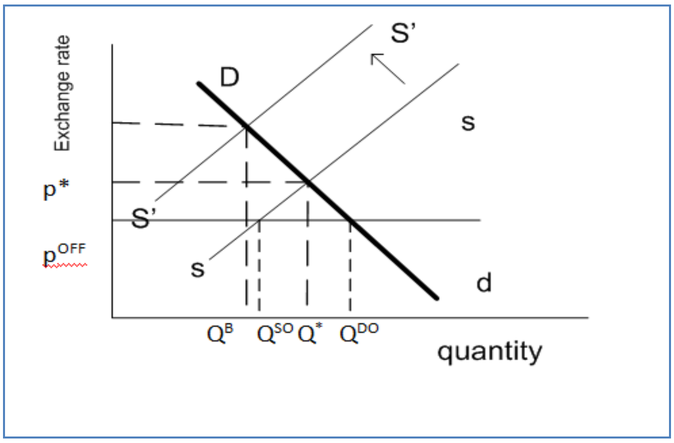
Data and the methodology

Data are from 212 groups of countries and geographic regions. These cross-country data were used in more than one study, including those from Dollar and Kraay (2003). In our study we are going to test the influence of institutions on average GDP growth per capita at PPP. The other variables are:

Rulellaw-law and order rating, we use this variable as proxy for quality of institutions, this variables is expected to be positively correlated with the average growth of GDP per capita.

Wardead-war casualties, **frehouserating**-freedom house rating, **cima_v**-contractintensive money (measure of property rights), **revolution**-revolutions, these variables are proxies for rulellaw. These variables are being used as instruments for rule of law variable and are proxies for quality of institutions.

gdppercap~a-average GDP per capita growth at PPP. This variable is variable of interest in our study. Dependent variable is being expressed in per capita terms and PPP conversion factor for more comparable result has been added. This variable is expressed in log terms.
govconshar~p-government consumption as share of GDP. This variable is expected to be positively correlated with average GDP per capita growth variable. This variable is expressed in log terms.
fdiinflow_~p-FDI inflows as percentage to GDP.
linvestmen~p-log of investment as fraction to GDP
lnbmp-this variable is log of (1+black market premium). Black market premium refers to the amount in excess of the official exchange rate that must be paid to purchase foreign exchange on an illegal ("black") market. Black market premium when the official rate is not market clearing is presented on the next graph. The premium typically arises when a country fixes the value of its exchange rate in relation to another currency irrespective of the rate that would prevail in the commercial market. It is akin to the authorities' fixing a price for a commodity at a non-market-clearing level.



In figure 1, schedule DD reflects demand for foreign exchange, while schedule SS reflects the supply. Under normal circumstances DD will be downward sloping, meaning that demand for foreign exchange will be greater as the price (in units of domestic currency) declines. Similarly, SS will slope upward, since additional foreign currency will be supplied to the market only as the price (in units of local currency per unit of foreign currency) increases. Provided normal economic conditions prevail, the market can be expected to clear at price P*, where the supply and demand schedules intersect. At this price, quantity Q* of foreign exchange will be bought and sold. When a nation fixes its exchange rate at a nonmarket-clearing rate, the normalmarket mechanism is disrupted. At the official exchange rate, POFF, demand for foreign exchange, QDO, exceeds the available supply, QSO. Those wishing to purchase foreign exchange cannot obtain it at the official price in the commercial market. If they seek to obtain foreign exchange from a private source, rather than using the queuing mechanism established by the authorities, they will need to pay more than the official price.The margin will reflect the scarcity value of the foreign exchange, plus a premium to compensate sellers for participating in an illegal (''black'') market. This risk can be depicted by a leftward (upward) shift in the supply curve to S0S0, making the market-clearing exchange rate, PB, likely to exceed the clearing rate in a legal market. The difference between the clearing rate in the illegal market, PB, and the official exchange rate, POFF, is the black market premium. This variable it is expected to be negatively correlated wioth the average growth of GDP per capita.

Instrumental variables (2SLS) versus OLS

An **Instrumental Variable** is a variable that is correlated with *X* but uncorrelated with *e*.

If Z_i is an instrumental variable:

1. $E(Z_i X_i) \neq 0$
2. $E(Z_i e_i) = 0$

The econometrician can use an instrumental variable *Z* to estimate the effect on *Y* of only that part of *X* that is correlated with *Z.* Because *Z* is uncorrelated with *e*, any part of *X* that is correlated with *Z* must also be uncorrelated with *e*. An instrumental variable lets the econometrician find a part of X that behaves as though it had been randomly assigned. When the economist is worried about measurement error, a good choice of instrument is simply a different measure of the same variable. The new measure may have its own errors, but these errors are unlikely to be correlated with the mistakes in the first measure, or with any other component of *e* (Murray, 2006). Instrumental variables are NOT the explanator of interest. We do not simply use instrumental variables as proxies for the explanator of interest.
Instead, we use IV's as a tool to tease out the "random" (or at least uncorrelated) component of *X*. Let's construct a consistent IV estimator for the case of measurement error.

1. $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \quad E(\varepsilon_i) = 0$

2. $Var(\varepsilon_i) = \sigma_\varepsilon^2 < \infty \quad Cov(\varepsilon_i, \varepsilon_j) = 0, i \neq j$

3. $E(X_i, \varepsilon_i) = 0, \quad \frac{1}{n}\sum(x_i^2) \to \sigma_X^2 < \infty$

4. $M_i = X_i + v_i \quad E(v_i) = 0$

5. $Var(v_i) = \sigma_v^2 \; Cov(v_i, v_j) = 0, i \neq j$
6. $Cov(v_i, X_i) = 0 \; Cov(Z_i, X_i) \neq 0$
7. $Cov(Z_i, \varepsilon_i) = 0$

If X_i were uncorrelated with e_i , we would want to weight more heavily observations with a high x_i value. We know that Z_i is correlated with the "clean" part of X_i , so now we want to weight more heavily observations with a high z_i value. Here we ask question what is expectation for IV?

$$E(\hat{\beta}_i^{IV}) = E\left(\frac{\sum z_i Y_i}{\sum z_i x_i}\right) = E\left(\frac{\sum z_i(\beta_0 + \beta_1 X_i + \varepsilon_i)}{\sum z_i x_i}\right)$$
$$= \beta_1 E\left(\frac{\sum z_i X_i}{\sum z_i x_i}\right) + \left(\frac{\sum z_i \varepsilon_i}{\sum z_i x_i}\right)$$
$$= \beta_1 + \sum E\left(\frac{\sum z_i \varepsilon_i}{\sum z_i x_i}\right)$$

Because $Cov(X_i, \varepsilon_i) \neq 0$, the bias term cannot be eliminated IV is biased in the same direction as the bias in OLS.

A variable Z_i can instrument for a particular troublesome explanator, X_{Ri}, if:

$Cov(Z_i, X_{Ri}) \neq 0$

$Cov(Z_i, e_i) = 0$

Z_i must be correlated with the troublesome variable for which it instruments, but need not be correlated with all of the troublesome variables. To estimate a multiple regression consistently, we need at least one instrumental variable for each troublesome explanator. When we have just enough instruments for consistent estimation, we say the regression equation is **exactly identified**. When we have more than enough instruments, the regression equation is **over identified**. When we do not have enough instruments, the equation is **under identified** (and inconsistent). An **Instrumental Variable** is a variable that is correlated with *X* but uncorrelated with *e*.

If Z_i is an instrumental variable:

$E(Z_i X_i) \neq 0$

$E(Z_i e_i) = 0$

If X_i were uncorrelated with e_i , we would want to weight more heavily observations with a high x_i value.We know that Z_i is correlated with the "clean" part of X_i , so now we want to weight more heavily observations with a high z_i value.

Beta estimator is

$$\hat{\beta}^{IV} = \frac{\sum z_i Y_i}{\sum z_i x_i}$$

When the regression is under identified, then we do not have a consistent estimator.
When the regression is exactly identified, then we simply use Instrumental Variables Least Squares. When the regression is over identified, we have more instruments than we need. The methods we learned last time are only suitable for the exactly identified case. When the regression equation is over identified, we have more instruments than we need. We could simply discard the additional instruments, but then we throw out valuable information. Ignoring valid instruments is inefficient. Standard OLS estimator is BLUE best linear unbiased estimator, to test whether OLS coefficients or 2SLS coefficients are better we are going to perform Hausman test. The Hausman specification test performs test of significance of one estimator versus alternative estimator

Panel Fixed effects IV model versus Random effects IV model

Potential unobserved heterogeneity is a form of omitted variables bias."Unobserved heterogeneity" refers to omitted variables that are fixed for an individual (at least over a long period of time). With cross-sectional data, there is no particular reason to differentiate between omitted variables that are fixed over time and omitted variables that are changing. However, when an omitted variable is fixed over time; panel data offers another tool for eliminating the bias. **Panel Data** is data in which we observe repeated cross-sections of the same individuals. Examples:

- Annual unemployment rates of each state over several years
- Quarterly sales of individual stores over several quarters
- Wages for the same worker, working at several different jobs

By far the leading type of panel data is repeated cross-sections over time. The key feature of panel data is that we observe the same individual in more than one condition. Omitted variables that are fixed will take on the same values each time we observe the same individual. The Fixed Effects Estimator basic idea is to estimate a separate intercept for each individual.

$$Y_{it} = \beta_0 + \beta_1 X_{1it} + \beta_1 X_{1it} + v_i + \mu_{it}$$

$$-Y_{it} = \beta_0 + \beta_1 X_{1it} + \beta_1 X_{1it} + v_i + \mu_{it,}$$

$$(Y_{it} - Y_{it}) = 0 + \beta_1 \left(X_{it} - X_{it,}\right) + 0 + 0 + \mu_{it} - \mu_{it}$$

When we difference, the heterogeneity term v_i drops out. (In the distinct intercepts model, the b_{0i} would drop out). By assumption, the m_{it} are uncorrelated with the X_{it} OLS would be a consistent estimator of b_1.

When unobserved heterogeneity is uncorrelated with explanators, panel data techniques are not needed to produce a consistent estimator. However, we do need to correct for serial correlation between observations of the same individual. When $E(X_{it}, v_i) = 0,$, panel data does not offer special benefits. We use Random Effects to overcome the serial correlation of panel data. The key idea of random effects:

- Estimate s_v^2 and s_m^2
- Use these estimates to construct efficient weights of panel data observations

Once we have estimates of s_v^2 and s_m^2, we can re-weight the observations optimally.

These calculations are complicated, but most computer packages can implement them.

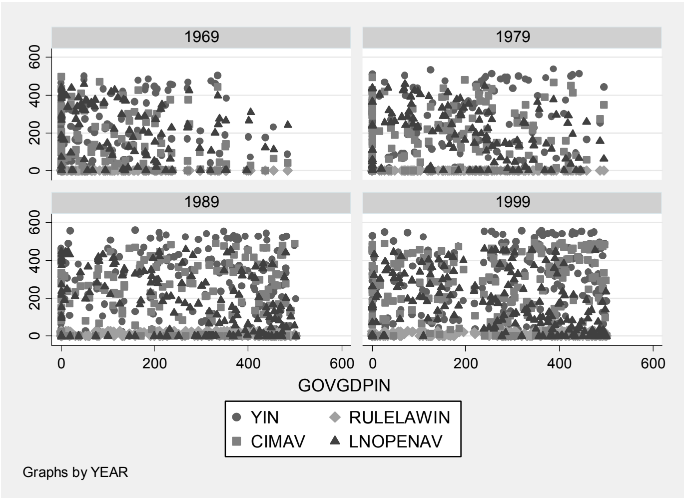
Descriptive statistics of the model

Descriptive statistics of the model is given in the following table

Variable	Obs	Mean	Std. Dev.	Min	Max
lgdppercap~a	848	191.1038	184.5586	1	560
rulellaw	848	5.643868	9.014775	1	31
lavertrade	848	125.4929	150.5476	1	460
govconshar~p	848	150.888	166.4599	1	502

lnbmp |	848	110.2618	132.5916	1	420
linvestmen~p |	848	3.576252	2.632151	0	6.326149
fdiinflow_~p |	848	125.5778	148.9089	1	458
cima_v |	848	145.7642	163.9984	1	496
wardead |	848	12.44458	28.41316	1	133
revolution |	848	4.548349	5.94604	1	30
frehousera~g |	848	11.05896	13.34896	1	37

In our sample we use decadal data. Sample contains 4 observations for each of 212 groups in the panel, contains data from 1969-1979,1979-1989, and 1989-1999. Moving of the variables through four decades is shown on the next graphs.



Where YIN here is annual average growth of GDP pre capita in PPP terms variable. Cimav are contract intensive money. Contract Intensive Money (CIM) = (M2 - money outside the banking system)/M2 where M2= Money + Quasi money. Proportion of money supply held by the banking system, sometimes interpreted as a proxy for the rule of law or an indicator of the credibility of financial institutions.LNOPENAV is natural logarithm of the average trade openness of the country, i.e. Average trade. RULELAWIN is the rule of law variable it law and order rating variable.

2SLS VS OLS [5]

2SLS regression is modeled as follows:

[5] See Appendix 1 2SLS regression

$$\ln(GDPpercapita) = \beta_0 + \beta_1 institutions + \beta_2 Trade + \beta_3 controls + u_i$$

Dependent variable log of GDP per capita in PPP terms.

Instrumental variables (2SLS) regression	Variables	Coefficients	p-value P>|t|
rulellaw	Rule of law proxy for quality of institutions	11.45504	0.005
lavertrade	Log of average trade	-0.0905889	0.071
lnbmp	Log of black market premium	-0.1623014	0.000
linvestmen~p	Log of investment as a fraction to GDP	31.56	0.000
govconshar~p	Government consumption as a share to GDP	0.1011464	0.114
fdiinflow_~p	FDI inflows as proportion to GDP	0.126112	0.003
_cons	Constant term	11.75178	0.285

Instrumented: rulellaw

Instruments: lavertrade lnbmp linvestmentgdp govconsharegdp fdiinflow_gdp frehouserating revolution cima_v

From the above Table we can see that the rule of law is highly positively correlated with growth, coefficient is 11.45, p-value is 0.005, meaning that the coefficient is statistically significant at all conventional levels. This is expected positive sign from the theory. Coefficient on the logarithm of average trade is small of size (-0.09), but is statistically significant up to 7% level of significance. Growth is positively correlated with average trade, but trade compared with other explanatory variables here has negative sign, meaning that compared to the institutions is growth deteriorating. Logarithm of black market premium exerts negative sign, which is expected from the *apriori* knowledge. Black market is non-regulated market that doesn't pay taxes to the country in which exists coefficient is -0.16, and is significant at all conventional levels. Private investment and government consumption as a fraction to GDP are expectedly positively correlated with growth with coefficients of 31.56 and 0.11 respectively. And Investment as a fraction to GDP is significant at all conventional levels, while government consumption is almost significant at 10% level of significance. FDI are positively correlated with growth as it is expected from the theory with a sign 0.12. Here instruments for Rule of law are contract intensive money, war casualties and revolutions. OLS regression is presented in a Table [6]

Dependent variable log of GDP per capita in PPP terms.			
Ordinary least squares regression	Variables	Coefficients	p-value P>|t|
rulellaw	Rule of law proxy for quality of institutions	5.024089	0.000
lavertrade	Log of average trade	-0.0384768	0.268
lnbmp	Log of black market premium	-0.1948633	0.000
linvestmen~p	Log of investment as a fraction to GDP	33.33	0.000
govconshar~p	Government consumption as a share to GDP	0.1868692	0.000
fdiinflow_~p	FDI inflows as proportion to GDP	0.1501029	0.000
_cons	Constant term	22.83623	0.003

Ramsey Reset test using powers of the fitted values of the dependent variable

$F(3, 838) = 1.78$

$Prob > F = 0.1490$

From the above Table only the coefficient of trade is negative and insignificant at all conventional levels. Rule of law as a proxy for institutional quality is again as expected positively correlated with growth, coefficient of 5.02 and highly significant at all levels of significance. Black market premium is negative -0.19 and is significant at all conventional levels. Investment as fraction to GDP, government consumption as a share to GDP and FDI inflows as a fraction to GDP are positively correlated with growth. Coefficients respectively are: 33.33,0.18 and 0.15 and are significant at all conventional levels. Ramsey Reset test showed that the model does not suffer from omitted variables bias. If we reject the null hypothesis of no omitted variables , probability of making Type I error is 15%.

[6] See Appendix 2 OLS regression

Hausman test

This command computes the Hausman test statistic. The null hypothesis is that the OLS estimator is consistent. If accepted, we probably would prefer to use OLS instead of 2SLS. The option constant is necessary to tell Stata to include the constant term in the comparison of both estimates. The sigmamore option tells Stata to use the same estimate of the variance of the error term for both models. This is desirable here since the error term has the same interpretation in both models. The df(1) option tells Stata that the null distribution has one degree of freedom. Stata was able to figure this out when I left this option out, even though the Hausman test is comparing values of two 5- element (not one-element) vectors. It probably knew this by finding only one non-zero eigenvalue of the 5-by-5 covariance matrix estimate that it calls (V_b-V_B) in the output. It's safer to impose the d.f. in the hausman command as above.

```
                 ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |     ivreg          .           Difference          S.E.

    rulellaw |    11.45504     5.024089         6.43095        3.736097
  lavertrade |   -.0905889    -.0384768       -.0521121        .0302748
       lnbmp |   -.1623014    -.1948633         .032562        .0189171
linvestmen~p |       31.56     33.32564       -1.765634        1.025755
govconshar~p |    .1011464     .1868692       -.0857229        .0498012
fdiinflow_~p |     .126112     .1501029       -.0239909        .0139376
       _cons |    11.75178     22.83623       -11.08445        6.439575
```

b = consistent under Ho and Ha; obtained from ivreg

B = inconsistent under Ha, efficient under Ho; obtained from regress

Test: Ho: difference in coefficients not systematic

chi2(1) = (b-B)'[(V_b-V_B)^(-1)](b-B)

= 2.96

Prob>chi2 = 0.0852

(V_b-V_B is not positive definite)

From the above result from Hausman test, we can see that OLS is acceptable at 1% and 5% level of significance, but not at 10% .Otherwise 2SLS squares would be more preferable.

Over identification test[7]

Next are presented results from the overidentification test.

scalar list x^2 pval

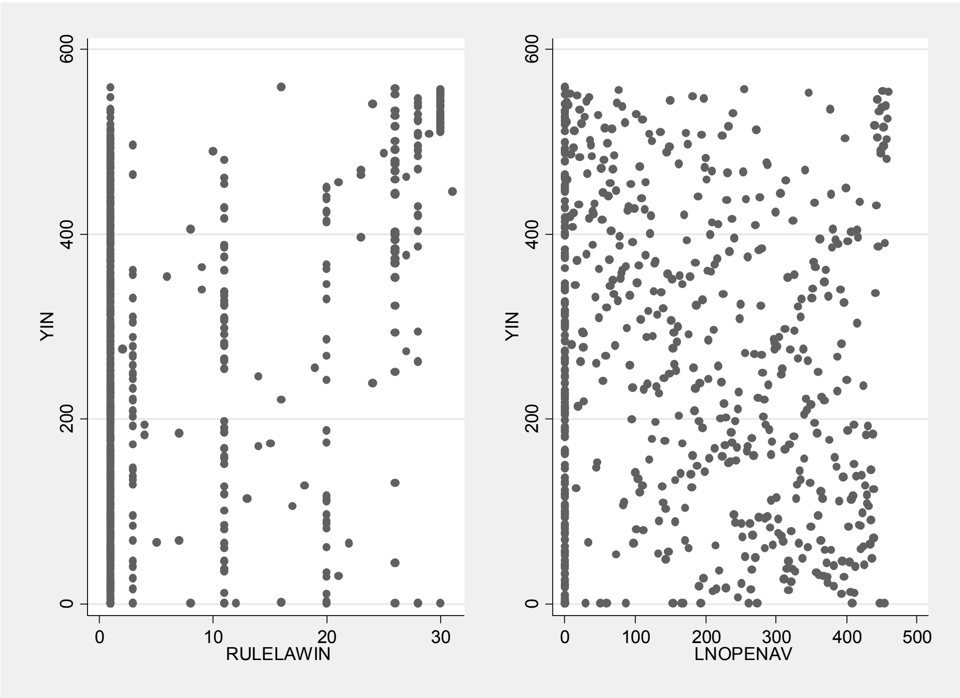
x^2 = 474.82519

[7] See Hausman test in Appendix 3

pval =	0

So at all conventional levels of significance we can drop hypothesis that instruments are exogenous. We can drop one or two of them but we can't be sure if that solves the problem.

So in conclusion about this part we can say that OLS won the battle and is better estimator than OLS , since it has better results in Hausman test and 2SLS did not show good overidentification test. From the below scatters it is evident that Rule of law variable and openness variable are positively correlated with growth.



G2SLS random-effects (RE) model

IV estimation can also be combined with panel data models in a straight forward manner Recall, that under the assumption of unobserved heterogeneity we removed the unobserved heterogeneity by either first differencing or fixed effects. This left us back in the world of OLS. However, one of the demeaned or first-differenced repressors could still be correlated with the error term, suggesting that IV could be helpful. *Ctry* variable i.e. country is panel IIS , ID variable. [8]

[8] See Appendix 4 G2SLS random-effects (RE) model

Dependent variable log of GDP per capita in PPP terms.			
Instrumental variables (G2SLS) regression Random effects model	Variables	Coefficients	p-value P>|t|
rulellaw	Rule of law proxy for quality of institutions	1.622535	0.000
lavertrade	Log of average trade	-0.0008549	0.981
linvestmen~p	Log of investment as a fraction to GDP	0.3291961	0.000
govconshar~p	Government consumption as a share to GDP	0.1058485	0.011
_cons	Constant term	65.90368	0.000
Group variable :ctry			
Instrumented: rulellaw			
Instruments: lavertrade investmentgdp govconsharegdp frehouserating wardead revolution cima_v			

From the above regression we can see that rulellaw variable which is being used as proxy for quality of institutions, is positively correlated with growth of GDP per capita variable at PPP terms, coefficient is 1.6 and p-value is 0.000. Coefficient on Trade is highly insignificant, pvalue is 0.981. Investment and government consumption are positively and statistically significant with coefficients 0.32 and 0.11 respectively.
As conclusion Trade is insignificant to growth compared with institutions.

Fixed effects regression (within)IV model[9]

In the next Table is presented Fixed effects panel regression IV model with panel ID variable ctry.

Dependent variable log of GDP per capita in PPP terms.			
Fixed effects regression (within)IV	Variables	Coefficients	p-value P>|t|

[9] See Appendix 5 Fixed effects regression (within)IV model

model			
rulellaw	Rule of law proxy for quality of institutions	1.579087	0.000
lavertrade	Log of average trade	-0.020254	0.640
linvestmen~p	Log of investment as a fraction to GDP	0.2575612	0.000
govconshar~p	Government consumption as a share to GDP	0.0961099	0.024
_cons	Constant term	84.53991	0.000
Group variable :ctry			

In conclusion institutions and investment as fraction to GDP and government consumption as share to GDP are positively and statistically significantly correlated.

Appendix 2SLS regression

```
Instrumental variables (2SLS) regression

      Source |       SS       df       MS              Number of obs =     848
-------------+------------------------------           F(  6,   841) =  124.71
       Model |  13000377.3     6  2166729.55           Prob > F      =  0.0000
    Residual |  15850017.6   841  18846.6321           R-squared     =  0.4506
-------------+------------------------------           Adj R-squared =  0.4467
       Total |  28850394.9   847  34061.8593           Root MSE      =  137.28

------------------------------------------------------------------------------
lgdppercap~a |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    rulellaw |   11.45504   4.102134     2.79   0.005     3.403417    19.50666
  lavertrade |  -.0905889   .0500865    -1.81   0.071    -.1888982    .0077204
       lnbmp |  -.1623014   .0445351    -3.64   0.000    -.2497144   -.0748884
linvestmen~p |      31.56   2.686769    11.75   0.000     26.28644    36.83356
govconshar~p |   .1011464   .0639289     1.58   0.114    -.0243325    .2266253
fdiinflow_~p |    .126112   .0420451     3.00   0.003     .0435863    .2086377
       _cons |   11.75178   10.98684     1.07   0.285    -9.813075    33.31663
------------------------------------------------------------------------------
Instrumented:  rulellaw
Instruments:   lavertrade lnbmp linvestmentgdp govconsharegdp fdiinflow_gdp
               frehouserating revolution cima_v
------------------------------------------------------------------------------
```

Appendix 2 OLS regression

```
      Source |       SS       df       MS              Number of obs =     848
-------------+------------------------------           F(  6,   841) =  161.59
       Model |  15449333.1     6  2574888.86           Prob > F      =  0.0000
    Residual |  13401061.7   841  15934.6751           R-squared     =  0.5355
-------------+------------------------------           Adj R-squared =  0.5322
       Total |  28850394.9   847  34061.8593           Root MSE      =  126.23

------------------------------------------------------------------------------
lgdppercap~a |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    rulellaw |   5.024089   .5187478     9.69   0.000     4.005897    6.042282
  lavertrade |  -.0384768   .0347058    -1.11   0.268    -.1065969    .0296433
       lnbmp |  -.1948633    .036319    -5.37   0.000    -.2661499   -.1235767
linvestmen~p |   33.32564   2.247488    14.83   0.000     28.91429    37.73698
govconshar~p |   .1868692   .0312295     5.98   0.000     .1255722    .2481662
fdiinflow_~p |   .1501029    .036061     4.16   0.000     .0793227     .220883
       _cons |   22.83623   7.784074     2.93   0.003     7.557735    38.11472
------------------------------------------------------------------------------
Ramsey RESET test using powers of the fitted values of lgdppercapita
       Ho:  model has no omitted variables
                 F(3, 838) =      1.78
                  Prob > F =      0.1490
```

Appendix 3 Hausman test

quietly reg ivresid   ruleoflaw lavertrade investmentgdp govconsharegdp

. predict explresid,xb

. matrix accum rssmat = explresid,noconstant
(obs=848)

. matrix accum rssmat = explresid,noconstant
(obs=848)

. matrix accum tssmat = ivresid,noconstant
(obs=847)

. scalar nobs=e(N)

. scalar x2=nobs*rssmat[1,1]/tssmat[1,1]

. scalar pval=1-chi2(1,x2)

. scalar list x2 pval
     x2 = 474.82519
   pval =   0

Appendix 4 G2SLS random effects IV regression

```
G2SLS random-effects IV regression              Number of obs      =       848
Group variable: ctry                            Number of groups   =       212

R-sq:  within  = 0.3022                         Obs per group: min =         4
       between = 0.6248                                        avg =       4.0
       overall = 0.4837                                        max =         4

                                                Wald chi2(4)       =    437.92
corr(u_i, X)       = 0 (assumed)                Prob > chi2        =    0.0000

------------------------------------------------------------------------------
lgdppercap~a |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   ruleoflaw |   1.622535    .257857     6.29   0.000     1.117144    2.127925
  lavertrade |  -.0008549   .0366775    -0.02   0.981    -.0727415    .0710317
investment~p |   .3291961   .0285336    11.54   0.000     .2732712     .385121
govconshar~p |   .1058485   .0417191     2.54   0.011     .0240807    .1876164
       _cons |   65.90368   11.21311     5.88   0.000     43.92639    87.88097
-------------+----------------------------------------------------------------
     sigma_u |  128.00592
     sigma_e |  91.331967
         rho |  .66265566   (fraction of variance due to u_i)
------------------------------------------------------------------------------
Instrumented:   ruleoflaw
Instruments:    lavertrade investmentgdp govconsharegdp frehouserating wardead revolution
                cima_v
------------------------------------------------------------------------------
```

Appendix 5 Panel Fixed effect IV regression

```
Fixed-effects (within) IV regression         Number of obs      =         848
Group variable: ctry                         Number of groups   =         212

R-sq:  within  = 0.1198                      Obs per group: min =           4
       between = 0.6100                                     avg =         4.0
       overall = 0.4553                                     max =           4

                                             Wald chi2(4)       =     3974.14
corr(u_i, Xb)  = 0.2832                      Prob > chi2        =      0.0000

------------------------------------------------------------------------------
lgdppercap~a |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   ruleoflaw |   1.579087   .2395886     6.59   0.000     1.109502    2.048672
  lavertrade |   -.020254   .0432842    -0.47   0.640    -.1050896    .0645816
investment~p |   .2575612   .0336336     7.66   0.000     .1916405    .3234819
govconshar~p |   .0961099   .0425786     2.26   0.024     .0126573    .1795625
       _cons |   84.53991   8.688616     9.73   0.000     67.51053    101.5693
-------------+----------------------------------------------------------------
     sigma_u |   111.5128
     sigma_e |  91.331967
         rho |  .59851397   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F  test that all u_i=0:     F(211,632) =     4.94         Prob > F    = 0.0000
------------------------------------------------------------------------------
```

References

1. Acemoglu, Daron, Simon Johnson, and James A. Robinson, "The Colonial Origins of Comparative Development: An Empirical Investigation." *American EconomicReview* 91, no. 5 (December 2001), pp. 1369-140
2. Barro, Robert J., "Democracy and Growth." *Journal of Economic Growth* 1, no. 1 (1996), pp. 1-27.
3. Dollar,D, Kraay,A,(2003), Institutions, Trade, and Growth Revisiting the Evidence,The World bank working paper
4. Gwartney, James, and Robert Lawson, *Economic Freedom of the World Annual Report 2003.* Vancouver, B.C.: Fraser Institute, 2003.
5. Green, William. 1993. Econometric Analysis. Second Edition. Macmillan. Canada.
6. Gujarati, Damodar. 2003. Basic Econometrics. Fourth Edition. McGraw Hill. USA.
7. Hall, Robert E., and Charles I. Jones, "Why Do Some Countries Produce So Much More Output Per Worker than Others?" *Quarterly Journal of Economics* 114, no. 1(February 1999), pp. 83-116.
8. Murray, P.,Michael(2006), Econometrics A modern introduction, Pearson ,Addison - Wesley
9. Scully, Gerald W., "The Institutional Framework and Economic Development." *Journal of Political Economy* 96, no. 3 (June 1988), pp. 652-662.
10. Wooldridge, Jeffery. 2006. Introductory Econometrics. A modern approach. Third Edition. Thomson. Canada.