Cross-Sectional and Panel Data Econometrics

Topic: **Impact of Population on Economic Growth**

INTRODUCTION

With world population reaching 7.7 billion people and continuously growing at a rate of 1.08%, the existing debates on the implications of population in terms of size, or change in composition and demographic transition, and the quality of life have increased extensively. Additionally, the economy runs due to people and for the people, where population is the source for output production, labour supply, demand for goods, savings and investment, and any other economic factor that we can think of. Due to this, it is highly essential to look into the impact population can possibly has on the economy.

Theoretically, there are three views illustrating the "relationship between population and economic growth, which are *population pessimists, population optimist,* and *population neutralists*". The first view belongs to Malthusian or orthodox school, which elucidated the negative impact of population on economic development. According to Malthus, population of the world keeps increasing while the food production does not, due to which a gap between the demand and supply for food leading to hunger and starvation. Conclusively, with the increase in population there is a slowdown in the economic growth and development, which is referred to as the Malthusian trap of population growth and subsistence. "A more contemporary theory highlighting this negative impact is given by Solow, who explains that in the transitional to the steady state, higher population growth will lead to a decline in the income per capita and thus highlighting the negative effect" (Klasen, 2007). Similarly, "Mankiw, Roemer, and Weil (1992) found out that an increase in the population growth rate of 10% would reduce per capita income in the steady state by 5%". This is aggravated when human capital is considered a factor in production, because the pressure builds up on savings in order to equip more people with physical and human capital. On the other hand, "optimistic view advocates that rapid population growth allows economic development because of the application of 'economies of scale' which is able to promote technological and institutional development and innovation". Further, "optimists believe that human capital is in-fact the main weapon in economic growth and this human capital comes from nothing but the growth in population". Lastly, the neutralists contend that population growth has "no significant effects, either negative or positive, on the economic development" and is rather growing in isolation.

According to the above mentioned three views, it can be concluded that "the relationship between the population growth and economic growth is a widely debated one, but the differences in the causality indicates the lack of empirical evidence". Therefore, population- growth relationship remains one of the oldest problems in economics. (Synopsis, Parnika Singh)

LITERATURE REVIEW

The relationship between "*population and economic growth*" has been debated extensively, and thus provides plethora of literature review. Some models exclusively study a particular demography while some are similar on the basis of the variables they study or the methodology they use. While most of the analysis follows time series analysis, few also presents their evidence based on panel regression analysis.

Gordon R. Stavig (1979) in his paper, *"The Impact of Population Growth on the Economy of Countries"* concluded that "rapid population increase had a negative impact on changes in many crucial economic indicators during the 1955-71 periods for the 94 countries that were studied". He not only explained that "population growth reduces GDP per capita, but there is a negative impact of population growth on private consumption and thus reducing the savings rate". Moreover, he highlighted that in most of the developing countries, only a smaller part of the population is actively participating in the labor force, and therefore with an increase in the population this ratio reduces even further. In totality, population growth leads to the slowdown of the economy due to various factors explained above.

In another paper by Klasen (2007), *"The Impact of Population Growth on Economic Growth and Poverty Reduction in Uganda",* the paper "examined the link between population and per capita economic growth, and poverty, using the case of Uganda". He concluded that, "by combining macro and micro-econometric approach and using a panel data, they found both theoretical and empirical evidence suggesting that the high population growth puts a considerable break on per capita growth prospects and contributes significantly to low achievement in poverty reduction". Similarly, Hamza (2015) in his paper, *"Panel Data Analysis of Population Growth and it's implication on Economic Growth of Developing Countries",* also came to the conclusion that, "the effect of population growth on per capita GDP growth is linear and continuous". He also highlighted the "causality between the birth rates and death rates with the economic growth for both the short and the long runs".

On the other hand, Koduru (2016), in his paper, *"Effect of Population Growth on Economic Development in India"* highlights that "there is in-fact a positive relationship between the population growth and economic development for India, where in their analysis for every unit increase in population, the GDP grows by 3.38 units". (Synopsis, Parnika Singh)

OBJECTIVE

The objective of this paper is to capture the relationship between population growth and economic growth, thus contributing to the population-economic growth nexus. We study this relationship "by a cross-country analysis of 58 developing countries (from Latin America, Africa and Asia), for the time period 2000-2015 using panel data analysis". The regression analysis helps us conclude if the developing countries fall under the pessimistic, optimistic or neutral view of population-economic growth. (Synopsis, Parnika Singh)

RESEARCH DESIGN

Methodology

The nature of the data is a mixture of both cross sectional and a time series data, and therefore we analyze the data by doing a panel regression analysis, where we run pooled regression, fixed effect method as well as random effect method. "This is done to capture the dynamic behavior and efficient estimation of the parameters. For choosing between pooled regressions, fixed effect or random effect models on the basis of the results from Breusch Pagan and Hausman Test". (Cameron, Trivedi)

Variables Description

For our analysis, we take GDP, as the dependent variable and a proxy of economic growth, while independent variables consist of birth rate, death rate, and net population migration as proxy variables for population growth. We also take, unemployment rate and savings rate as other variables that can have causality with economic growth.

- *Dependent Variable*

  Gross Domestic Product: GDP is defined as, "Total market value of all finished goods and services produced in a year, as well as investments, government spending, and exports minus imports" (World Bank), and hence provides a good proxy for economic growth. There are a large number of variables that can possibly affect the GDP level in a country, however we focus on the effect of population growth on GDP.

- *Independent Variables*

  Birth Rate: Birth rate is the "average number of births per 1000 persons in a year" (World Bank). We take this variable because number of births affects the population growth in any country, hence it serves as a proxy variable for population growth. It is assumed that birthrate can have a positive, negative or no effect on the economic growth indicating the existence of optimistic, pessimistic or neutral position on the relationship between economic growth and population growth.

  Death Rate: Death rate is the "average number of deaths per 1000 persons in a year" (World Bank). We take this variable as a proxy for population growth and can have either positive, negative or no effect on economic growth.

  Net Population Migration Rate: It is defined as the "rate at which the number of immigrants (people coming into an area) and the number of emigrants (people leaving an area) throughout the year" (World Bank). We take this variable as a proxy for population growth and can have either positive, negative or no effect on economic growth.

  Unemployment Rate: "The share of the labor force that is jobless, expressed as a percentage" (World Bank), is defined as the unemployment rate. It is assumed that unemployment rate has a negative impact on the economic growth, because higher unemployment rate means low labour in the economy which results in low output, as a result of which the GDP is low. Hence higher unemployment rate translates to lower economic growth.

  Savings Rate: "The savings rate is the percentage of disposable personal income that a person or group of people save rather than spend on consumption" (World Bank). This income is then used for investments which therefore increases the GDP. Hence higher savings rate will lead to higher GDP.

Data Source

The data has been collected for 58 developing countries from Asia, Africa and Latin America region for the time period 2000-2015. The reason for the short time period is due to the unavailability of the data for net population migration rate. Despite having more than 150 developing nations, we used 58 countries because of the large missing values of migration rate, unemployment rate and savings rate. The data has been collected from the Open Data Bank of the World Bank.

Model

The model based on the above mentioned dependent and independent variables in the most basic form can be expressed as

$$GDP_{it} = \alpha + \beta_1 BR_{it} + \beta_2 DR_{it} + \beta_3 POP_{it} + \beta_4 S_{it} + \beta_5 UNEM_{it} + U_{it}$$

Where GDP_{it} = GDP

α = intercept

BR_{it} = Birth rate in thousand

DR_{it} = Death rate in thousand

POP_{it} =Net Migration Rate

S_{it} = Savings Rate

$UNEM_{it}$ =Unemployment rate

U_{it} =Panel error term

i = Countries

t = time period

Before carrying out the exploratory and confirmatory analysis, we first set up the correct functional form regression equation, by using the command "gladder" in Stata. This can also be checked by summary of the variables, where the variable is not normally distributed. We transform GDP its logarithmic form, which thus yields the following model.

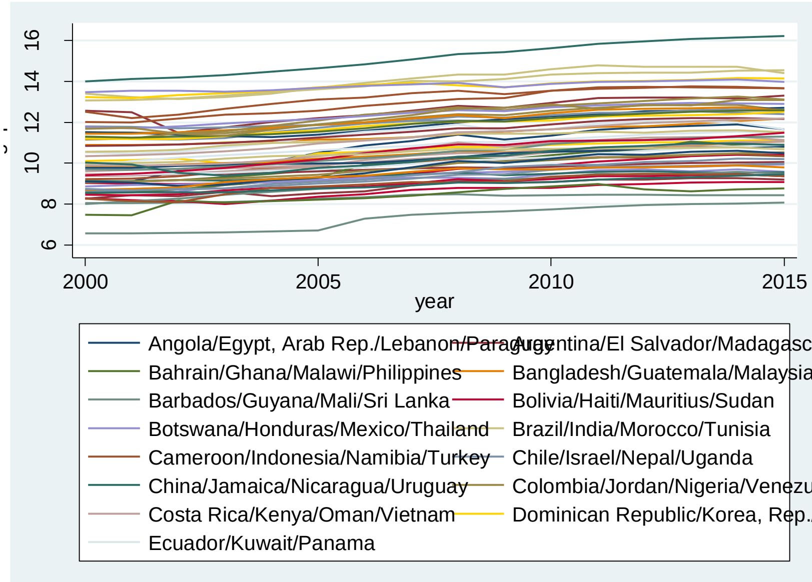
$$LGDP_{it} = \alpha_i + \beta_1 BR_{it} + \beta_2 DR_{it} + \beta_3 POP_{it} + \beta_4 S_{it} + \beta_5 UNEM_{it} + U_{it}$$

Where, LGDP is the logarithm of GDP.

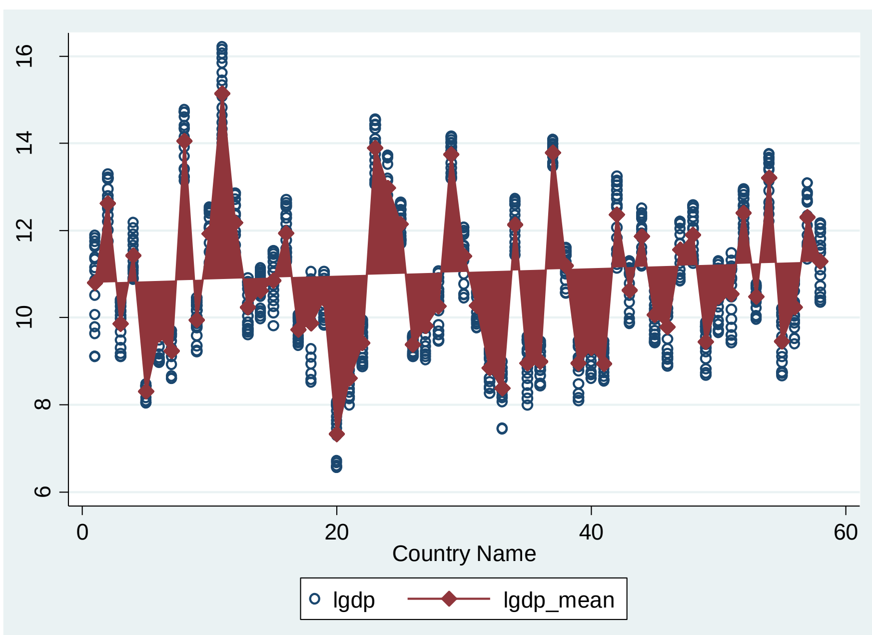
<u>EXPLORATORY ANALYSIS</u>

Gross Domestic Product Trends

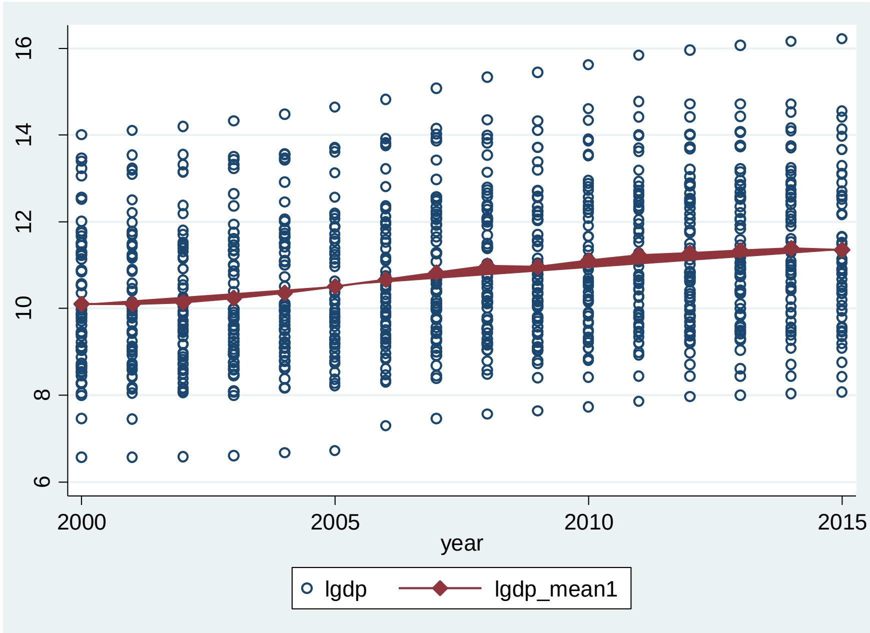
The graph for the GDP for 58 countries over the period of 16 years indicates that there has been an increase in the level of GDP in each countries. Some countries on one hand showed a higher rate of economic growth, which is highlighted by the steeper curve, while other countries grew at a slower pace.



A feature of panel data is that "it captures the unobserved heterogeneity across both cross sectional units and time" (Gujarati). The graph below indicates vast variation in the GDP across countries where some countries have higher economic growth than others.

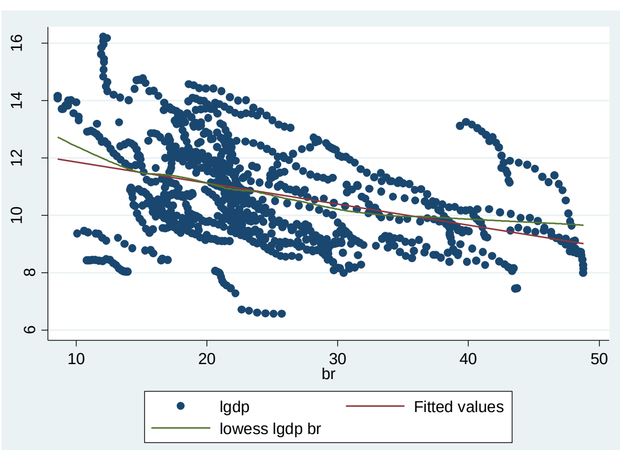


Similarly, the graph below shows how GDP has changed over time, and the circles represent the countries for the particular time period. The graph indicates that there has been an increase in the GDP over time, which further suggests that there is time fixed effects in the model.

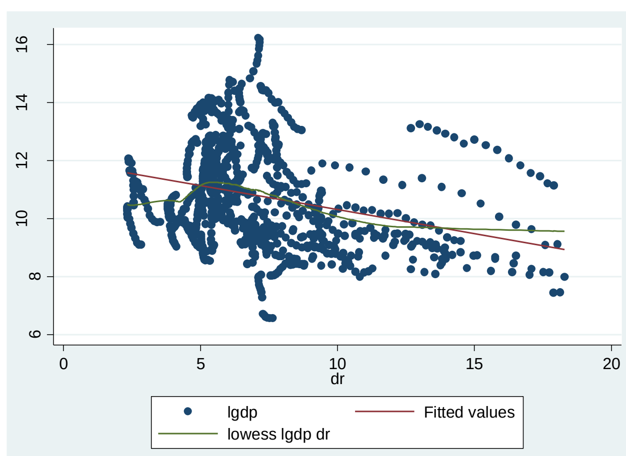


Dependent and Independent Variables

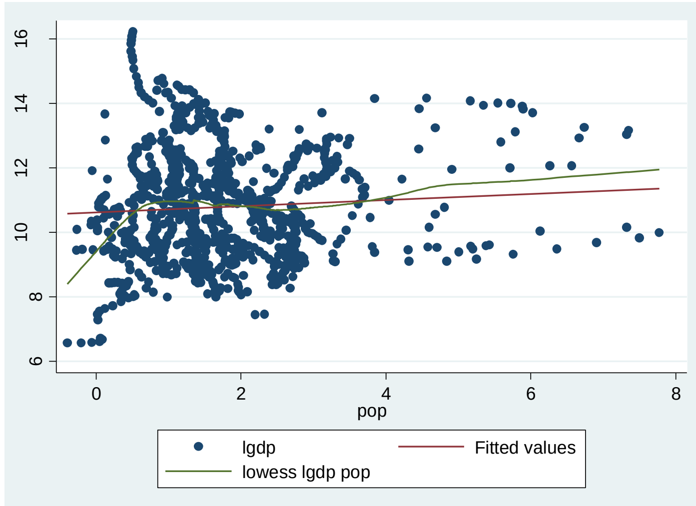
We check how population growth proxy variables are related to economic growth, i.e. GDP, by using scatter plots, Lowess Curve ("locally weighted scatterplot smoothing"), and fitted lines. The graphs show that there is a "negative relationship between Birth Rate and GDP, indicating a pessimistic opinion of population growth". Additionally , "with the increase in Death Rate, the GDP rises, and an increase in migration increases the economic growth slightly, indicating the presence of optimistic opinion".



Relation between GDP and Birth Rate



Relationship between GDP and Death Rate



Relationship between GDP and Net Migration Rate

Similar results are obtained from correlation matrix between the variables as shown below. The correlation matrix depicts that net migration rate and savings rate have a positive relationship with GDP, with correlation 0.0674 and 0.4063 respectively. While birth rate, death rate and unemployment rate are negatively correlated with economic growth, with correlation -0.3925, -0.2888 and -0.2700 respectively. The correlation between net migration and GDP is quite low indicating a weak relation between the two.

```
. corr lgdp unem dr br s pop
(obs=928)
```

	lgdp	unem	dr	br	s	pop
lgdp	1.0000					
unem	-0.2700	1.0000				
dr	-0.2888	0.3013	1.0000			
br	-0.3925	0.1212	0.6803	1.0000		
s	0.4063	-0.2128	-0.1455	-0.1815	1.0000	
pop	0.0674	-0.2307	-0.0501	0.0380	0.1816	1.0000

Further the summary of the variables and data set by running *xtsum* is given by the following, “where it gives mean, standard deviation, minimum and maximum values, and tells the total number of observation, observations in cross section and in time series, for overall, between (cross country) and within (time series) for our dependent and independent variables after necessary transformation”. (Appendix)

REGRESSION ANALYSIS

Under confirmatory analysis, we estimate the model by running regression and interpret the output, thus explaining the "total variation in dependent variable due to the explanatory variables and the residuals" (Gujarati). In panel data analysis, we can run either "*pooled OLS model, fixed effect least squares dummy variable model, fixed effect model or random effect model* depending on the sample".
Proceeding to the regression analysis, we first convert the data from wide form to long form, because "the number of cross sectional units are higher than the time periods and then declare the data to be balanced panel data, where there are same numbers of observation for every unit". (Cameron, Trivedi)

For declaration we run *xtset* command, which shows that the panel data is balanced and that there are no missing observation.

```
. xtset country year
       panel variable:  country (strongly balanced)
        time variable:  year, 2000 to 2015
                delta:  1 unit
```

After declaring the data to be panel data, we then make run regressions for *Pooled OLS Model, Fixed Effect Regression and Random Effect Regression* and showing which model is best suited for the study. We further run diagnostic tests for various problems that can take place in the model and try correcting them.

Pooled OLS Regression

In pooled OLS regression, "we pool together all the observation, but assume that the regression coefficients are the same for all the countries", which means that in pooled OLS the heterogeneity between different countries has been camouflage. The model also assumes that "the explanatory variables are strictly exogenous and the error terms are normally distributed". (Gujarati)

In Stata, the pooled OLS regression command is *regression dependent variable independent variables* which yields the following results:

```
regress lgdp pop br dr s unem

      Source |       SS           df       MS      Number of obs   =       928
-------------+----------------------------------   F(5, 922)       =     77.95
       Model |  818.710828         5  163.742166   Prob > F        =    0.0000
    Residual |  1936.87659       922  2.10073383   R-squared       =    0.2971
-------------+----------------------------------   Adj R-squared   =    0.2933
       Total |  2755.58742       927  2.97258621   Root MSE        =    1.4494

------------------------------------------------------------------------------
        lgdp |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         pop |  -.0221677   .0403618    -0.55   0.583    -.1013793    .0570439
          br |  -.0640522   .0072058    -8.89   0.000    -.0781938   -.0499106
          dr |   .0244713   .0226528     1.08   0.280    -.0199857    .0689282
           s |   .0514956   .0047324    10.88   0.000      .042208    .0607832
        unem |  -.0643344   .0109438    -5.88   0.000    -.0858121   -.0428567
       _cons |   11.51166   .2134372    53.93   0.000     11.09279    11.93054
```

We can conclude the following from the results:

- The model is statistically significant we can reject the null hypothesis that the model is insignificant because Prob>F is close to 0.00
- The coefficient values of most of the independent variable is statistically significant, while coefficients of net migration and death rate are statistically insignificant.
- R-square of the model is 29.71% which is fairly good, and highlights that the model can explain 30% variation in dependent variable due to the independent variables.

However, "the problem with this method is that it does not take into account the heterogeneity across the cross sectional units, i.e. countries", which further results in the problem of autocorrelation in the model. In such cases, we use *cluster-robust* standard errors that cluster on individuals for getting better results. (Cameron, Trivedi)

```
. regress lgdp pop br dr s unem, vce (cluster country)

Linear regression                               Number of obs     =        928
                                                F(5, 57)          =       7.34
                                                Prob > F          =     0.0000
                                                R-squared         =     0.2971
                                                Root MSE          =     1.4494

                                 (Std. Err. adjusted for 58 clusters in country)
------------------------------------------------------------------------------
             |               Robust
        lgdp |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         pop |  -.0221677   .1494194    -0.15   0.883    -.3213749    .2770396
          br |  -.0640522   .0252547    -2.54   0.014    -.1146238   -.0134806
          dr |   .0244713   .0745225     0.33   0.744    -.1247575      .1737
           s |   .0514956   .0160181     3.21   0.002     .0194198    .0835714
        unem |  -.0643344   .0330969    -1.94   0.057    -.1306098    .0019409
       _cons |   11.51166   .8133588    14.15   0.000     9.882941    13.14039
```

The results above highlights that only coefficient of birth rate, savings rate and constant are significant because there p-value is lower than 5% and thus we can reject the null hypothesis. However, due to high p-values of other coefficients, we cannot reject the null hypothesis and thus they are insignificant.

To further improve the results by correcting autocorrelation using *xtreg, pa corr(ar1) vce(robust)* command, which gives us the following results

```
. xtreg lgdp pop br dr s unem, pa corr(ar1) vce(robust)

Iteration 1: tolerance = .15897797
Iteration 2: tolerance = .03518812
Iteration 3: tolerance = .00476096
Iteration 4: tolerance = .00066141
Iteration 5: tolerance = .00009219
Iteration 6: tolerance = .00001286
Iteration 7: tolerance = 1.793e-06
Iteration 8: tolerance = 2.500e-07

GEE population-averaged model               Number of obs      =        928
Group and time vars:         country year   Number of groups   =         58
```

```
Link:                      identity      Obs per group:
Family:                    Gaussian                     min =         16
Correlation:                  AR(1)                     avg =       16.0
                                                        max =         16
                                         Wald chi2(5)       =     155.70
Scale parameter:           2.644177      Prob > chi2        =     0.0000

                                (Std. Err. adjusted for clustering on country)
```

lgdp	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
pop	.0054403	.0177259	0.31	0.759	-.0293019	.0401825
br	-.0985822	.0135968	-7.25	0.000	-.1252315	-.0719329
dr	-.1049865	.0263534	-3.98	0.000	-.1566383	-.0533347
s	.0039491	.0020106	1.96	0.050	8.42e-06	.0078899
unem	-.0364538	.0095216	-3.83	0.000	-.0551158	-.0177918
_cons	14.09406	.3822653	36.87	0.000	13.34483	14.84328

In this regression, the coefficient changes considerably from pooled OLS and the cluster robust standard errors are small which leads to improved efficiency due to better modelling. We can also see that in the above model only coefficient of net migration is statistically insignificant, while other coefficients are significant at 5% significance level. The estimated correlation matrix is stored in e(R) which is slightly different than the autocorrelation in the pooled OLS model. (Appendix)

However, there are other methods for treating panel data which are proved to be better because they are able to capture the heterogeneity in the model. Also by the exploratory analysis, we saw that there is heterogeneity in cross-sectional units which thus cannot be treated as unobserved.

Fixed Effect Model

Another method of panel regression is to "eliminate the fixed effect by expressing the values of the dependent and explanatory variables for each country as deviation from their respective mean values". "The resulting values are called *mean corrected values* which are then used for running the regression". This method is called *Fixed Effect Regression Model*, which differs from pooled OLS regression in the sense that pooled OLS does not take into account the heterogeneity across cross-sectional units but fixed effect incorporates heterogeneity, "but by eliminating heterogeneity by differencing sample observations around their sample means rather than using dummy variables".

Another feature is that "the method controls for all time-invariant differences between the individuals, so estimated coefficients cannot be biased because of omitted time-invariant characteristics like religion, gender, race, etc. However because of this, the model cannot be used to investigate time-invariant causes of the dependent variables". (Gujarti)

In Stata, the regression command for fixed effect is *xtreg, fe* (having default standard error)

```
. xtreg lgdp unem br dr pop s, fe

Fixed-effects (within) regression               Number of obs     =        928
Group variable: country                         Number of groups  =         58

R-sq:                                           Obs per group:
     within  = 0.6210                                         min =         16
     between = 0.1578                                         avg =       16.0
     overall = 0.1771                                         max =         16

                                                F(5,865)          =     283.52
corr(u_i, Xb)  = -0.5808                        Prob > F          =     0.0000

------------------------------------------------------------------------------
        lgdp |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        unem |  -.0650792   .0064808   -10.04   0.000    -.0777992   -.0523593
          br |  -.1761411   .0069502   -25.34   0.000    -.1897823   -.1624998
          dr |   -.061458    .011353    -5.41   0.000    -.0837407   -.0391753
         pop |   .0165432    .017976     0.92   0.358    -.0187385     .051825

           s |  -.0022284   .0022427    -0.99   0.321    -.0066301    .0021733
       _cons |   16.04795   .1631821    98.34   0.000     15.72767    16.36823
-------------+----------------------------------------------------------------
     sigma_u |    1.89874
     sigma_e |  .31996252
         rho |  .97238745   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F test that all u_i=0: F(57, 865) = 316.74                   Prob > F = 0.0000
```

Some of the observation from the above results are

- The model is statistically significant we reject the null hypothesis of insignificant model because Prob>F is close to 0.00
- 'rho' represents the intra-class correlation and thus suggest that 97% variation is due to different cross panels.
- All the coefficients, except the coefficient for savings and net population migration, are statistically significant "because two-tail p-value is less than 0.05, thus we cannot reject the null hypothesis (coefficient is insignificant) at 5% significance level".
- The correlation between the error term and the regressors in the fixed effect model is equal to -0.5808
- The R-square within the model is 0.6210, which suggests that 62% of the variation within the panel units has been explained by the model.

It has been highlighted that the fixed effect model with default standard error assumes iid error terms. However, by using *cluster-robust* command we can relax this assumption. These results are shown below, which highlights that due to only within variation of the data, the standard errors are triple the standard error of pooled OLS. While R-square, rho and coefficient values remain same, in the *cluster-robust* command, the standard errors have changed and net migration and savings remain statistically insignificant because of the high p-value.

```
. xtreg lgdp unem br dr pop s, fe vce(cluster country)

Fixed-effects (within) regression               Number of obs     =        928
Group variable: country                         Number of groups  =         58

R-sq:                                           Obs per group:
     within  = 0.6210                                         min =         16
     between = 0.1578                                         avg =       16.0
     overall = 0.1771                                         max =         16

                                                F(5,57)           =      46.29
corr(u_i, Xb)  = -0.5808                        Prob > F          =     0.0000

                                (Std. Err. adjusted for 58 clusters in country)
------------------------------------------------------------------------------
             |               Robust
        lgdp |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        unem |  -.0650792   .0131929    -4.93   0.000    -.0914975    -.038661
             |
          br |  -.1761411   .0163357   -10.78   0.000    -.2088528   -.1434293
          dr |   -.061458   .0300058    -2.05   0.045    -.1215435   -.0013725
         pop |   .0165432   .0418841     0.39   0.694    -.0673282    .1004147
           s |  -.0022284   .0057744    -0.39   0.701    -.0137914    .0093346
       _cons |   16.04795   .3877539    41.39   0.000     15.27148    16.82441
-------------+----------------------------------------------------------------
     sigma_u |    1.89874
     sigma_e |  .31996252
         rho |  .97238745   (fraction of variance due to u_i)
```

Random Effect Model

In random effect model, it is assumed that "the intercept of an individual unit is a random drawing from a much larger population with a constant mean value, thus the individual intercept is expressed as a deviation from this constant mean". Normally, random effect model is chosen "when the intercept of the cross-sectional unit is uncorrelated to the regressors, or when we have to introduce time-invariant variables". (Gujarati)

In Stata, we can estimate random effect model by using *xtreg, re* command, which gives the following results with default standard errors

```
. xtreg lgdp unem br dr pop s, re

Random-effects GLS regression                   Number of obs     =        928
Group variable: country                         Number of groups  =         58

R-sq:                                           Obs per group:
     within  = 0.6206                                         min =         16
     between = 0.1605                                         avg =       16.0
     overall = 0.1800                                         max =         16

                                                Wald chi2(5)      =    1339.66
corr(u_i, X)   = 0 (assumed)                    Prob > chi2       =     0.0000
```

lgdp	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
unem	-.0674754	.0065378	-10.32	0.000	-.0802892	-.0546616
br	-.1658578	.0068295	-24.29	0.000	-.1792434	-.1524722
dr	-.0616651	.0115159	-5.35	0.000	-.0842359	-.0390943
pop	.0165992	.0182573	0.91	0.363	-.0191845	.0523829
s	-.0014424	.0022765	-0.63	0.526	-.0059042	.0030194
_cons	15.79569	.2509849	62.93	0.000	15.30377	16.28761
sigma_u	1.4335601					
sigma_e	.31996252					
rho	.95254813	(fraction of variance due to u_i)				

Some of the observations from the above results are:

- The model is statistically significant we reject the null hypothesis of insignificant model because Prob>F is close to 0.00
- 'rho' represents the intra-class correlation and thus suggest that 95% variation is due to different cross panels.
- All the coefficients, except the coefficient for savings and net migration, are statistically significant because "two-tail p-value is less than 0.05, thus we cannot reject the null hypothesis (coefficient is insignificant) at 5% significance level".
- The correlation between the error term and the regressors in the fixed effect model is assumed to be 0.
- The R-square within the model is 0.6206, which suggests that 62% of the variation within the panel units has been explained by the model. Other R-squares are 16% between estimators and 18% overall.

However, the random effect estimator assumption of equi-correlated errors is very strong, and therefore in order to relax the assumption the *vce(cluster id)* command is recommended to use in order to obtain cluster-robust standard errors. The results are as follows

```
. xtreg lgdp unem br dr pop s, re vce(cluster country)

Random-effects GLS regression                   Number of obs     =        928
Group variable: country                         Number of groups  =         58

R-sq:                                           Obs per group:
     within  = 0.6206                                         min =         16
     between = 0.1605                                         avg =       16.0
     overall = 0.1800                                         max =         16

                                                Wald chi2(5)      =     252.17
corr(u_i, X)   = 0 (assumed)                    Prob > chi2       =     0.0000

                             (Std. Err. adjusted for 58 clusters in country)
```

lgdp	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
unem	-.0674754	.012848	-5.25	0.000	-.0926571	-.0422937
br	-.1658578	.0152163	-10.90	0.000	-.1956813	-.1360343
dr	-.0616651	.0288953	-2.13	0.033	-.1182988	-.0050314
pop	.0165992	.044687	0.37	0.710	-.0709856	.104184
s	-.0014424	.0057495	-0.25	0.802	-.0127112	.0098263
_cons	15.79569	.4136359	38.19	0.000	14.98498	16.6064
sigma_u	1.4335601					
sigma_e	.31996252					
rho	.95254813	(fraction of variance due to u_i)				

<u>Estimates Table</u>

Variable	OLS_rob	fe_rob	re_rob
pop	-0.0222	0.0165	0.0166
br	-0.0641*	-0.1761***	-0.1659***
dr	0.0245	-0.0615*	-0.0617*
s	0.0515**	-0.0022	-0.0014
unem	-0.0643	-0.0651***	-0.0675***
_cons	11.5117***	16.0479***	15.7957***
N	928	928	928
r2	0.2971	0.6210	
r2_o		0.1771	0.1800
r2_b		0.1578	0.1605
r2_w		0.6210	0.6206
sigma_u		1.8987	1.4336
sigma_e		0.3200	0.3200
rho		0.9724	0.9525

legend: * p<0.05; ** p<0.01; *** p<0.001

The table compares all the results from the pooled robust, fixed effect robust, and random effect robust. In the table we can see the variables which are statistically significant at 5% in each model. We can also see that the coefficient values of fixed effect and that of random effects values are close each other, and while fixed effect has higher R-square than pooled OLS.

Correct Panel Regression Model

The decision between the "fixed effect model and the random effect model is based on whether unobserved factor is correlated with regressors or whether it is uncorrelated, consequently fixed effect model is preferable when unobserved factor is correlated with regressors". This can be decided by using the Hausman Test, "where the null hypothesis is that the fixed effect and random effect estimators do not differ substantially", and the test follows chi-square distribution. "According to the test, if the null hypothesis is rejected, then we can conclude that fixed effect model is appropriate to use".

In Stata, we run command *hausman fixed random.* However it is better to run *hausman fixed random, sigmamore* command because "it specifies that both covariance matricies are based on the (same) estimated disturbance variance from the efficient estimator". The results from the Hausman test are as follows, (Gujarati)

```
. hausman fixed random, sigmamore

                 ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |     fixed        random       Difference          S.E.
-------------+----------------------------------------------------------------
        unem |   -.0650792    -.0674754        .0023962        .0010333
          br |   -.1761411    -.1658578       -.0102833         .001935
          dr |    -.061458    -.0616651        .0002072        .0013517
         pop |    .0165432     .0165992        -.000056        .0019304
           s |   -.0022284    -.0014424        -.000786        .0002529
------------------------------------------------------------------------------
                           b = consistent under Ho and Ha; obtained from xtreg
            B = inconsistent under Ha, efficient under Ho; obtained from xtreg

    Test:  Ho:  difference in coefficients not systematic

                  chi2(5) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                          =       42.87
                Prob>chi2 =      0.0000
```

The results depicts that we reject the null hypothesis at 1% or 5% level because Prob>chi2 is very close to 0.00. Hence the "Hausman test indicates that fixed effect model is more appropriate than the random effect model". This is primarily because the geographic differences in the countries, i.e. the cross-sectional units, due to which random effect model is an inappropriate one.

Additionally, we can also run "*Breusch-Pagan Lagrange Multiplier* test which indicates the appropriateness of the random effect model". The null hypothesis of the test is that "variance across entities is zero, which means that there is no significant difference across units". By rejecting the null hypothesis we accept the appropriateness of the random effects model. It also helps to "decide between random effect model and simple pooled OLS regression model" (Princeton).

In Stata, we run *xttest0* command for running the LM test. The results are as follows,

```
. xttest0

Breusch and Pagan Lagrangian multiplier test for random effects

        lgdp[country,t] = Xb + u[country] + e[country,t]

        Estimated results:
                         |       Var     sd = sqrt(Var)
                ---------+-----------------------------
                    lgdp |   2.972586       1.724119
                       e |    .102376       .3199625
                       u |   2.055095        1.43356

        Test:   Var(u) = 0
                             chibar2(01) =  5516.39
                          Prob > chibar2 =   0.0000
```

Since "Prob>chibar2 is close to 0.00, this means that at 1% or 5% significance level the null hypothesis of $\sigma\mu^2 = 0$ is rejected". Therefore, according to the LM test random effect model is appropriate and should be used instead of fixed effect or pooled OLS regression model. However, in this study we give higher weightage to the Hausman Test and thus conclude that it is most appropriate to estimate fixed effect model.

DIAGNOSTIC TESTS

Diagnostic tests are undertaken to see whether the model follows certain assumptions and if not then we can correct these to get accurate results.

Time-Fixed Effect

To check whether time effects are needed while running fixed effect model use the command "*testparm*, which jointly tests if dummies for all years are equal to 0, and if they are then there is no time effect." After running regression using years as dummy using the command *xi: xtreg y x i.year, fe*, and then testing time effects we get

```
. testparm _Iyear*

 ( 1)  _Iyear_2001 = 0
 ( 2)  _Iyear_2002 = 0
 ( 3)  _Iyear_2003 = 0
 ( 4)  _Iyear_2004 = 0
 ( 5)  _Iyear_2005 = 0
 ( 6)  _Iyear_2006 = 0
 ( 7)  _Iyear_2007 = 0
 ( 8)  _Iyear_2008 = 0
 ( 9)  _Iyear_2009 = 0
 (10)  _Iyear_2010 = 0
 (11)  _Iyear_2011 = 0
 (12)  _Iyear_2012 = 0
 (13)  _Iyear_2013 = 0
 (14)  _Iyear_2014 = 0
 (15)  _Iyear_2015 = 0

       F( 15,   907) =    3.99
            Prob > F =    0.0000
```

Results suggests that since “Prob>F is very close to 0.00, therefore we can reject the null hypothesis at 1% or 5% level of significance, which means that the coefficients for all years are not equal to zero jointly”, and therefore we have to incorporate time dummies for accurate regression results.

Similarly, we can test for fixed effects by testing for the dummies for each country. “The null hypothesis is that jointly the dummies for countries are equal to zero and if not, we have to incorporate fixed effects” (Princeton). After running regression using countries as dummy using the command *xi: xtreg y x i.country, fe*, and then testing fixed effects we get

```
( 1)  _Icountry_2 = 0  (21)  _Icountry_22 = 0 (41)  _Icountry_42 = 0
( 2)  _Icountry_3 = 0  (22)  _Icountry_23 = 0 (42)  _Icountry_43 = 0
( 3)  _Icountry_4 = 0  (23)  _Icountry_24 = 0 (43)  _Icountry_44 = 0
( 4)  _Icountry_5 = 0  (24)  _Icountry_25 = 0 (44)  _Icountry_45 = 0
( 5)  _Icountry_6 = 0  (25)  _Icountry_26 = 0 (45)  _Icountry_46 = 0
( 6)  _Icountry_7 = 0  (26)  _Icountry_27 = 0 (46)  _Icountry_47 = 0
( 7)  _Icountry_8 = 0  (27)  _Icountry_28 = 0 (47)  _Icountry_48 = 0
( 8)  _Icountry_9 = 0  (28)  _Icountry_29 = 0 (48)  _Icountry_49 = 0
( 9)  _Icountry_10 = 0 (29)  _Icountry_30 = 0 (49)  _Icountry_50 = 0
(10)  _Icountry_11 = 0 (30)  _Icountry_31 = 0 (50)  _Icountry_51 = 0
(11)  _Icountry_12 = 0 (31)  _Icountry_32 = 0 (51)  _Icountry_52 = 0
(12)  _Icountry_13 = 0 (32)  _Icountry_33 = 0 (52)  _Icountry_53 = 0
(13)  _Icountry_14 = 0 (33)  _Icountry_34 = 0 (53)  _Icountry_54 = 0
(14)  _Icountry_15 = 0 (34)  _Icountry_35 = 0 (54)  _Icountry_55 = 0
(15)  _Icountry_16 = 0 (35)  _Icountry_36 = 0 (55)  _Icountry_56 = 0
(16)  _Icountry_17 = 0 (36)  _Icountry_37 = 0 (56)  _Icountry_57 = 0
(17)  _Icountry_18 = 0 (37)  _Icountry_38 = 0 (57)  _Icountry_58 = 0
(18)  _Icountry_19 = 0 (38)  _Icountry_39 = 0
(19)  _Icountry_20 = 0 (39)  _Icountry_40 = 0       F( 57,   865) =  316.74
(20)  _Icountry_21 = 0 (40)  _Icountry_41 = 0            Prob > F =    0.0000
```

The above results show that “we reject the null hypothesis at 1% or 5% significance level because the p-value is very close to 0.00”.

Further we can also test for “random effects using *Breusch-Pagan Lagrange Multiplier test* which helps us decide between random effect regression and pooled OLS regression (The test has been discussed above where random effect model seems to be more appropriate then OLS regression model)”. The regression results incorporating fixed effects and time effects separately are shown in the Appendix.

Heteroscedasticity

The model suffers from heteroscedasticity “when the error terms do not have constant variance, and then its presence leads to biased and inefficient standard errors which further affects the coefficients and R-square”. Additionally, hypothesis testing is nor reliable in the presence of heteroscedasticity and the coefficients have wrong signs.

In Stata, we use the command *xttest3* which “runs modified *Wald Test* for group wise heteroscedasticity”. The null hypothesis is “that there is no heteroscedasticity and the variances are constance”. The results from *xttest3* are as follows,

```
. xttest3

Modified Wald test for groupwise heteroskedasticity
in fixed effect regression model

H0: sigma(i)^2 = sigma^2 for all i

chi2 (58)  =    66156.23
Prob>chi2 =      0.0000
```

The results from the above test suggest that "at 1% or 5% significance level we reject the null hypothesis of constant variance". This implies that the model is suffering from the problem of heteroscedasticity. The problem can be solved by either using robust standard errors while running regression by using command *VCE(robust)*, or we can also use *xtscc* command which gives even better results "as it takes care of heteroscedasticity as well as autocorrelation and makes standard errors more robust". (Stata Journal)

Serial Correlation

When the observations depends on its lagged values then the model suffers from serial correlation. It has been suggested that the serial correlation normally occurs with "data with long time series and is generally not a problem in micro panel". The problem caused standard errors of coefficients to be very small as a result of which the R-square value is very high.

In Stata, the command for running test for serial correlation is *xtserial, "*which runs Wooldrige test for autocorrelation in the panel data. Here the null hypothesis is that there is no serial correlation". The results from the Wooldrige test are as follows,

```
. xtserial lgdp unem br dr pop s

Wooldridge test for autocorrelation in panel data
H0: no first order autocorrelation
    F(  1,      57) =    267.230
           Prob > F =      0.0000
```

The results from the Wooldrige test suggests "that since the p-value is very low therefore we can reject the null hypothesis that there is no autocorrelation, and conclusively the model above is suffering from serial correlation problem".

It has been highlighted that "if there is structraul correlation in the idiosyncratic error term, clustering at the panel level will produce consistent estimates of the standard erorrs and therefore the estimators produced will be more efficient than before" (Princeton). Since the countries are heterogeneous and have their own cycles, it can be suggested that the model suffers from AR(1).

Normality

It is important to check if the error terms are normally distributed because "it helps to validate specification test, forecasting and other inferences procedures" (Gujarati). In Stata, we use the command *xtsktest* which tests for skeweness and kurtosis for each error term. The results are as follows,

```
. xtsktest
(running _xtsktest_calculations on estimation sample)

Bootstrap replications (50)
----+--- 1 ---+--- 2 ---+--- 3 ---+--- 4 ---+--- 5
..................................................    50

Tests for skewness and kurtosis                 Number of obs     =        928
                                                Replications      =         50

                              (Replications based on 58 clusters in country)
------------------------------------------------------------------------------
             |   Observed   Bootstrap                         Normal-based
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
  Skewness_e |   -.002335    .0202437    -0.12   0.908     -.042012     .037342
  Kurtosis_e |   .0656705    .0656342     1.00   0.317    -.0629701    .1943112
  Skewness_u |   .5068996    .5618262     0.90   0.367    -.5942596    1.608059
  Kurtosis_u |  -2.082135    1.427148    -1.46   0.145    -4.879293    .7150232
------------------------------------------------------------------------------
Joint test for Normality on e:        chi2(2) =   1.01    Prob > chi2 = 0.6022
Joint test for Normality on u:        chi2(2) =   2.94    Prob > chi2 = 0.2296
```

The results from *xtsktest* show that the joint test "for normality on e and on u cannot be rejected at 1% or 5% level of significance because the p-value is very high". This means that error terms are symmetric and are normally distributed.

Cross-Section Dependence/ Contemporaneous Correlation

For testing cross sectional dependence in panel data we use *Pesaran's CD test,* which tells if "error term of two cross sectional units in particular time period is correlated with each other". The null hypothesis of *Pesaran's CD test* is "the estimated residuals are not contemporaneously correlated against the alternative that of cross sectional dependence or presence of contemporaneous correlation among residuals". Cross-sectional dependence leads to biased and inconsistent results therefore it is suggested to test using the command *xtcsd, pesaran ab* in Stata. The results show are as follows, (Krishna Ram)

```
. xtcsd, pesaran abs


Pesaran's test of cross sectional independence =    35.921, Pr = 0.0000

Average absolute value of the off-diagonal elements =     0.543
```

The above results suggests "that at 1% or 5% level of significance we can reject the null hypothesis and assert that there is the presence of cross-sectional dependence".

Selection of Stata Command for producing appropriate Robust Standard Error

The model described above suffers from time-effect, autocorrelation, heteroscedasticity and cross-sectional dependence. The best command in order to produce robust standard errors which are then able to tackle these disturbances is *xtscc* where we make time dummies. *Xtscc* command runs regression with Driscoll-Kraay standard errors which corrects for the disturbances, We do not use *xtpcse* command despite it correcting the problems because it is best for running with small time periods. Hence we use *xtscc* for interpretation and estimation.

REGRESSION RESULTS

After checking the most appropriate panel model from pooled OLS regression, fixed effect model and random effect model, and then checking for various disturbances in the model, we came to the conclusion of using fixed effect model with time dummies which is regressed using the command *xtscc* in Stata for correcting the disturbances, i.e. " running regression with Driscoll-Kraay standard errors".

The regression equation is

$$GDP_{it} = \alpha + \beta_1 BR_{it} + \beta_2 DR_{it} + \beta_3 POP_{it} + \beta_4 S_{it} + \beta_5 UNEM_{it} + \delta_t T_t + U_{it}$$

Where T_t= Time Dummies (where benchmark is 2000)

Running the regression gives us the following results,

```
. xtscc lgdp unem br dr pop s i.year, fe

Regression with Driscoll-Kraay standard errors   Number of obs     =       928
Method: Fixed-effects regression                 Number of groups  =        58
Group variable (i): country                      F( 20,     15)    =   5112.48
maximum lag: 2                                   Prob > F          =    0.0000
                                                 within R-squared  =    0.8856
```

lgdp	Coef.	Drisc/Kraay Std. Err.	t	P>|t|	[95% Conf.	Interval]
unem	-.0277797	.0043311	-6.41	0.000	-.0370112	-.0185482
br	.0126965	.0055546	2.29	0.037	.000857	.0245359
dr	-.0296472	.0188302	-1.57	0.136	-.0697829	.0104885
pop	-.0004299	.0130908	-0.03	0.974	-.0283324	.0274725
s	.0026715	.001306	2.05	0.059	-.0001121	.0054551
year						
2000	0	(empty)				
2001	.0109602	.0028303	3.87	0.002	.0049277	.0169928
2002	.0536673	.0038302	14.01	0.000	.0455034	.0618313
2003	.1446262	.0043035	33.61	0.000	.1354536	.1537988
2004	.2664343	.0045468	58.60	0.000	.256743	.2761256
2005	.4004769	.0059261	67.58	0.000	.3878456	.4131082
2006	.5371449	.0075762	70.90	0.000	.5209965	.5532933
2007	.6852543	.0090701	75.55	0.000	.6659219	.7045867
2008	.8405752	.0104257	80.63	0.000	.8183534	.862797
2009	.844324	.0118415	71.30	0.000	.8190844	.8695636
2010	.9740817	.0126147	77.22	0.000	.9471941	1.000969
2011	1.095084	.0141808	77.22	0.000	1.064858	1.12531
2012	1.145569	.0153354	74.70	0.000	1.112882	1.178255
2013	1.209047	.0168457	71.77	0.000	1.173141	1.244953
2014	1.251193	.0173021	72.31	0.000	1.214315	1.288072
2015	1.229534	.0183435	67.03	0.000	1.190436	1.268633
_cons	10.15811	.1082775	93.82	0.000	9.927326	10.3889

Results from the regression output:

- The model is overall statistically significant because "the p-value of the F-test is very close to 0.00 which means that we can reject the null hypothesis of insignificant model".
- The within R-square values of the model is equal to 0.8856, which implies that the model is able to explain approximately 88% "of the variations within the panel units has been explained by the model. It shows that the model has a fairly high goodness of fit".

Interpretation of regression coefficients

- **Birth Rate:** The coefficient of death rate has statistically significant effect on GDP, since the p-value for two-tailed t test is less than 0.05, i.e. 5% significance level which means that we can reject the null hypothesis of insignificant coefficient. Coefficient is interpreted as "for a given country, as birth rate increases across time by one unit, GDP increases by 0.0126%, keeping all other variables as constant". This suggests "a positive relationship between birth rate and economic growth, which implies that according to birth rate population growth and economic growth has an optimistic relationship, because higher birth rate means higher population level which gives higher GDP".

- **Death Rate:** The coefficient of birth rate has statistically significant effect on GDP, but only at a very high significance level like 14%. Coefficient is interpreted as "for a given country, as death rate increases across time by one unit, GDP decreases by 0.0296%, keeping all other variables as constant". This suggests that there is "a negative relationship between birth rate and economic growth, which implies that according to death rate population growth and economic growth has an optimistic relationship, because higher death rate means lower population level which gives declining GDP".

- **Net Migration Rate:** The regression output suggests that coefficient for net migration rate is statistically unsignificant because "the p-value is very high and as a result of which we cannot reject the null hypothesis of insignificant coefficient".

- **Unemployment Rate:** The coefficient of unemployment rate is statistically significant because the p-value is close to 0.00. Coefficient is interpreted as "for a given country, as unemployment rate increases across time by one unit, GDP decreases by 0.0278%, keeping all other variables as constant". This suggests that there is "a negative relationship between unemployment rate and GDP as highlighted by the economic theory".

- **Savings Rate:** The coefficient of savings rate is statistically significant at 6% or 10% significance level because the p-value for the two-tail test is close to 0.05. Coefficient is interpreted as "for a given country, as savings rate increases across time by one unit, GDP increases by 0.0026%, keeping all the other variables as constant". This suggests that there is "a positive relationship between savings rate and GDP as highlighted by the economic theory".

- **Time Dummies:** The time dummy variable coefficients, where the benchmark category is year 2000, are statistically significant as the p-values of the two-tail test is close to 0.00, "which indicates that there is time effect in the above model". This has also been tested by *testparm* command in Stata.

CONCLUSION

The main objective of the study was “to determine the relationship between population growth and economic development and decide whether the pessimistic, optimistic or neutral school of thought is correct”, by running panel regression for 58 countries for time period 2000-2015. Based on the Hausman test and other tests testing for the disturbances in the model, we came to the conclusion that fixed effect time dummy model with robust standard errors (obtained by running command *xtscc* in Stata) was the most appropriate model for estimation.

According to the coefficients of birth rate and death rate, we can conclude that since in each case, with the increase in the population level the GDP increases. “Hence birth rate and death rate as a proxy for population growth and GDP as the proxy for economic growth suggests that there is in-fact a positive relationship between population growth and economic growth”. Conclusively, the optimistic school of though is correct for the data set we have taken. However, the level of immigration growth did not have a statistically significant coefficient and hence is unable to have an affect on the economic growth of the data set. Aditionally, when exploratory analysis was undertaken then “birthrate had a negative effect on economic growth, however the results from regression analysis suggested the opposite”. This might be due to the distrurbances in the data collected.

Furthermore, this case is not always true that population growth will always lead to positive economic growth as highlighted in the literature review where “population growth had negative impact on economic growth, therefore the scope for policy recommendation remains limited”.

REFERENCES

A. Colin Camera, Pravin K. Treivedi 2008, *Microeconometrics Using Stata*

Bhanu Phani Krishna Koduru (March 2016): Effect *of population growth on economic development in India*, Birla Institute of Technology and Science Pilani.

Damodar N. Guajarati 2004, *Basic Econometrics*

Gordon R. Stavig, The *Impact of Population on the Economy of countries*, Economic Development and Cultural Change, Vol. 27, No. 4 (Jul,. 1979), pp. 735-750, The University of Chicago Press.

H. Joseph Newton, Nicholas J. Cox 2003, *Testing for serial correlatio in linear panel-data models*, The Stata Journal

Klasen, Stephan; Lawson, David (2007) : *The impact of population growth on economic growth and poverty reduction in Uganda*, Diskussionsbeiträge, No. 133, Georg-August-Universität Göttingen, Volkswirtschaftliches Seminar, Göttingen.

Krihna Ram, *Panel Data Regression using STATA*

Labaran Hamaza, *Panel Data analysis of population growth and its implication on economic growth of development countries*, April 2015, Department of economics, SRM University, Chennai, India.

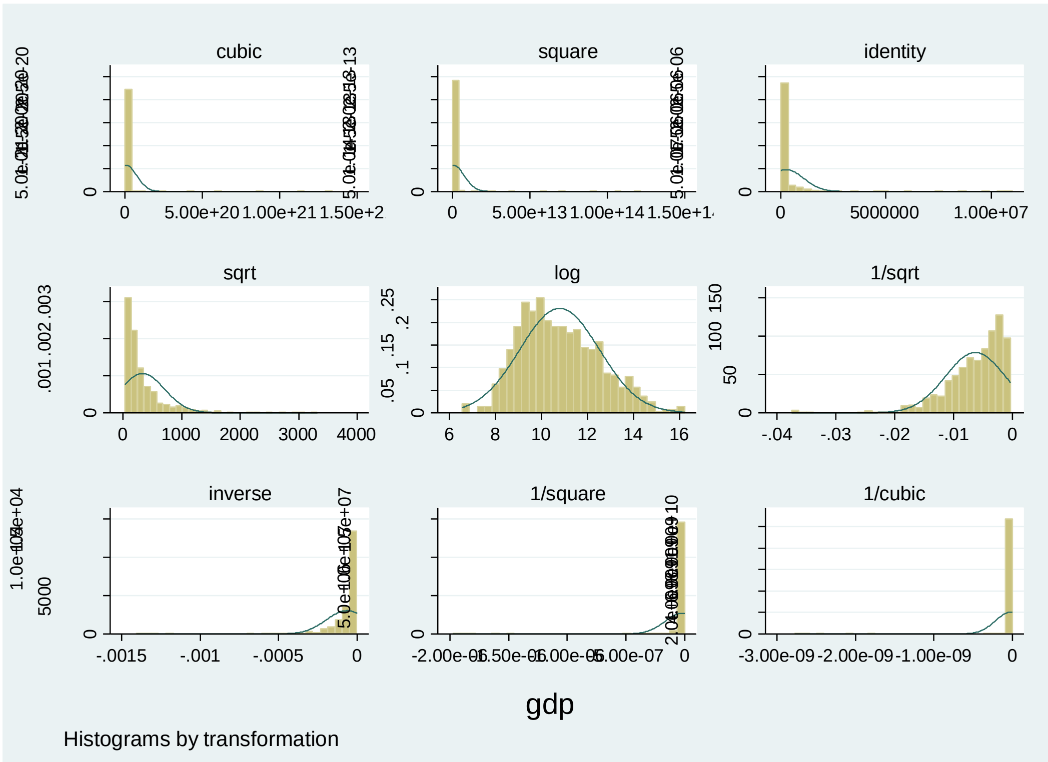
Mankiw, G. D. Roemer, and P.Weil 1992, *A Contribution to the Empirics of Economic Growth*, Quarterly Journal of Economics.

Princeton University, *2007: PANEL DATA ANALYSIS FIXED AND RANDOM EFFECTS USING STATA"*

World Bank indicators, https://data.worldbank.org/

APPENDIX

Gladder for GDP



Autocorreltaion Matrix

```
. matrix list e(R)

symmetric e(R)[16,16]
```

	c1	c2	c3	c4	c5	c6	c7	c8	c9	c10	c11	c12	c13	c14	c15	c16
r1	1															
r2	.99892341	1														
r3	.99784798	.99892341	1													
r4	.99677371	.99784798	.99892341	1												
r5	.99570059	.99677371	.99784798	.99892341	1											
r6	.99462863	.99570059	.99677371	.99784798	.99892341	1										
r7	.99355783	.99462863	.99570059	.99677371	.99784798	.99892341	1									
r8	.99248817	.99355783	.99462863	.99570059	.99677371	.99784798	.99892341	1								
r9	.99141967	.99248817	.99355783	.99462863	.99570059	.99677371	.99784798	.99892341	1							
r10	.99035232	.99141967	.99248817	.99355783	.99462863	.99570059	.99677371	.99784798	.99892341	1						
r11	.98928612	.99035232	.99141967	.99248817	.99355783	.99462863	.99570059	.99677371	.99784798	.99892341	1					
r12	.98822106	.98928612	.99035232	.99141967	.99248817	.99355783	.99462863	.99570059	.99677371	.99784798	.99892341	1				
r13	.98715715	.98822106	.98928612	.99035232	.99141967	.99248817	.99355783	.99462863	.99570059	.99677371	.99784798	.99892341	1			
r14	.98609439	.98715715	.98822106	.98928612	.99035232	.99141967	.99248817	.99355783	.99462863	.99570059	.99677371	.99784798	.99892341	1		
r15	.98503277	.98609439	.98715715	.98822106	.98928612	.99035232	.99141967	.99248817	.99355783	.99462863	.99570059	.99677371	.99784798	.99892341	1	
r16	.9839723	.98503277	.98609439	.98715715	.98822106	.98928612	.99035232	.99141967	.99248817	.99355783	.99462863	.99570059	.99677371	.99784798	.99892341	1

Summary

```
. xtsum
```

Variable		Mean	Std. Dev.	Min	Max	Observations	
countr~e	overall	.	.	.	.	N =	0
	between		.	.	.	n =	0
	within		.	.	.	T =	.
year	overall	2007.5	4.612258	2000	2015	N =	928
	between		0	2007.5	2007.5	n =	58
	within		4.612258	2000	2015	T =	16
gdp	overall	250553	835927.2	712.1676	1.10e+07	N =	928
	between		698023.7	1791.065	4944797	n =	58
	within		468424.3	-3482897	6321298	T =	16
unem	overall	7.203394	4.769474	.398	23.925	N =	928
	between		4.483986	1.049188	20.4875	n =	58
	within		1.722524	-.4506057	16.19539	T =	16
s	overall	22.8951	10.56075	-16.35903	64.71603	N =	928
	between		9.525794	3.742388	49.07809	n =	58
	within		4.717732	-5.661052	43.78008	T =	16
dr	overall	7.124363	3.010265	2.324	18.292	N =	928
	between		2.834777	2.4645	15.29188	n =	58
	within		1.075067	1.471301	12.20749	T =	16
br	overall	24.59316	9.209542	8.6	48.785	N =	928
	between		9.11361	9.7375	46.9595	n =	58
	within		1.76118	18.98547	31.18079	T =	16
pop	overall	1.754227	1.23037	-.3954859	7.776022	N =	928
	between		1.090363	.1194919	4.799171	n =	58
	within		.5866504	-1.891114	5.117084	T =	16
country	overall	29.5	16.7497	1	58	N =	928
	between		16.88688	1	58	n =	58
	within		0	29.5	29.5	T =	16
lgdp	overall	10.78744	1.724119	6.568313	16.21482	N =	928
	between		1.662902	7.331446	15.1457	n =	58
	within		.5020832	9.083579	11.97069	T =	16

Regression incorporating time dummies

```
. xtreg lgdp unem br dr pop s i.year, fe

Fixed-effects (within) regression               Number of obs     =        928
Group variable: country                         Number of groups  =         58

R-sq:                                           Obs per group:
     within  = 0.8856                                         min =         16
     between = 0.0182                                         avg =       16.0
     overall = 0.0907                                         max =         16

                                                F(20,850)         =     329.11
corr(u_i, Xb)  = 0.0118                         Prob > F          =     0.0000

------------------------------------------------------------------------------
        lgdp |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        unem |  -.0277797   .0037207    -7.47   0.000    -.0350826   -.0204768
          br |   .0126965    .005942     2.14   0.033     .0010338    .0243592
          dr |  -.0296472   .0063377    -4.68   0.000    -.0420866   -.0172078
         pop |  -.0004299   .0099819    -0.04   0.966     -.020022    .0191621
           s |   .0026715   .0013204     2.02   0.043     .0000799    .0052631
             |
        year |
       2001  |   .0109602    .033118     0.33   0.741    -.0540423    .0759628
       2002  |   .0536673    .033436     1.61   0.109    -.0119594     .119294
       2003  |   .1446262   .0338107     4.28   0.000     .0782639    .2109885
       2004  |   .2664343    .034285     7.77   0.000     .1991412    .3337274
       2005  |   .4004769   .0349136    11.47   0.000       .33195    .4690038
       2006  |   .5371449   .0357429    15.03   0.000     .4669903    .6072995
       2007  |   .6852543   .0364826    18.78   0.000     .6136478    .7568607
       2008  |   .8405752    .037202    22.59   0.000     .7675566    .9135937
       2009  |    .844324   .0378316    22.32   0.000     .7700697    .9185783
       2010  |   .9740817   .0386226    25.22   0.000     .8982748    1.049889
       2011  |   1.095084   .0395926    27.66   0.000     1.017373    1.172795
       2012  |   1.145569   .0405929    28.22   0.000     1.065895    1.225243
       2013  |   1.209047   .0418089    28.92   0.000     1.126986    1.291108
       2014  |   1.251193   .0428695    29.19   0.000     1.167051    1.335336
       2015  |   1.229534   .0441983    27.82   0.000     1.142784    1.316285
             |
       _cons |   10.15811   .1707096    59.51   0.000     9.823052    10.49318
-------------+----------------------------------------------------------------
     sigma_u |  1.6488141
     sigma_e |  .17731896
         rho |  .98856666   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F test that all u_i=0: F(57, 850) = 998.93                   Prob > F = 0.0000
```

Regression incorporating cross-sectional dummies

```
. xi:reg lgdp unem dr br s pop i.country
i.country         _Icountry_1-58      (naturally coded; _Icountry_1 omitted)
```

Source	SS	df	MS
Model	2667.03217	62	43.0166478
Residual	88.55525	865	.102376012
Total	2755.58742	927	2.97258621

Number of obs	=	928
F(62, 865)	=	420.18
Prob > F	=	0.0000
R-squared	=	0.9679
Adj R-squared	=	0.9656
Root MSE	=	.31996

lgdp	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
unem	-.0650792	.0064808	-10.04	0.000	-.0777992	-.0523593
dr	-.061458	.011353	-5.41	0.000	-.0837407	-.0391753
br	-.1761411	.0069502	-25.34	0.000	-.1897823	-.1624998
s	-.0022284	.0022427	-0.99	0.321	-.0066301	.0021733
pop	.0165432	.017976	0.92	0.358	-.0187385	.051825
_Icountry_2	-3.676743	.2128048	-17.28	0.000	-4.094417	-3.259069
_Icountry_3	-7.523413	.2109436	-35.67	0.000	-7.937434	-7.109391
_Icountry_4	-4.620468	.1846505	-25.02	0.000	-4.982883	-4.258052
_Icountry_5	-9.085591	.254528	-35.70	0.000	-9.585156	-8.586026
_Icountry_6	-5.790308	.1761694	-32.87	0.000	-6.136077	-5.444538
_Icountry_7	-4.626345	.168878	-27.39	0.000	-4.957804	-4.294886
_Icountry_8	-2.832532	.2200986	-12.87	0.000	-3.264522	-2.400542
_Icountry_9	-2.736718	.1388996	-19.70	0.000	-3.009337	-2.464098
_Icountry_10	-5.260486	.2263766	-23.24	0.000	-5.704798	-4.816175
_Icountry_11	-2.620838	.2386701	-10.98	0.000	-3.089278	-2.152398
_Icountry_12	-4.216512	.2131524	-19.78	0.000	-4.634869	-3.798156
_Icountry_13	-6.868212	.2205451	-31.14	0.000	-7.301078	-6.435345
_Icountry_14	-5.380276	.19555	-27.51	0.000	-5.764084	-4.996468
_Icountry_15	-5.33254	.1885178	-28.29	0.000	-5.702545	-4.962534
_Icountry_16	-3.067102	.1722266	-17.81	0.000	-3.405133	-2.729071
_Icountry_17	-6.533676	.2043356	-31.98	0.000	-6.934727	-6.132624
_Icountry_18	-4.086782	.1504353	-27.17	0.000	-4.382043	-3.791521
_Icountry_19	-4.528619	.1692443	-26.76	0.000	-4.860797	-4.196442
_Icountry_20	-8.256779	.2008548	-41.11	0.000	-8.650999	-7.86256
_Icountry_21	-5.544153	.1620243	-34.22	0.000	-5.86216	-5.226146
_Icountry_22	-5.926012	.172257	-34.40	0.000	-6.264103	-5.587922
_Icountry_23	-2.176503	.1884526	-11.55	0.000	-2.54638	-1.806625
_Icountry_24	-3.208734	.1918044	-16.73	0.000	-3.58519	-2.832277
_Icountry_25	-3.843747	.196531	-19.56	0.000	-4.229481	-3.458014
_Icountry_26	-6.85578	.2107325	-32.53	0.000	-7.269387	-6.442173
_Icountry_27	-4.69812	.1689299	-27.81	0.000	-5.029681	-4.36656
_Icountry_28	-2.886922	.1494752	-19.31	0.000	-3.180299	-2.593546
_Icountry_29	-4.732589	.2483554	-19.06	0.000	-5.220039	-4.245139
_Icountry_30	-5.496661	.1998749	-27.50	0.000	-5.888958	-5.104364
_Icountry_31	-6.716741	.2214226	-30.33	0.000	-7.151329	-6.282152
_Icountry_32	-4.75138	.1523411	-31.19	0.000	-5.050381	-4.452379
_Icountry_33	-3.888922	.1415126	-27.48	0.000	-4.16667	-3.611174
_Icountry_34	-4.879573	.2032035	-24.01	0.000	-5.278402	-4.480743
_Icountry_35	-2.108	.1420691	-14.84	0.000	-2.38684	-1.829159
_Icountry_36	-8.369428	.2349962	-35.62	0.000	-8.830657	-7.908199
_Icountry_37	-2.659024	.1971791	-13.49	0.000	-3.046029	-2.272018
_Icountry_38	-4.744441	.1936179	-24.50	0.000	-5.124457	-4.364425
_Icountry_39	-4.396034	.1725112	-25.48	0.000	-4.734624	-4.057445
_Icountry_40	-6.438187	.1797167	-35.82	0.000	-6.790919	-6.085455
_Icountry_41	-6.808949	.1824499	-37.32	0.000	-7.167045	-6.450853
_Icountry_42	.1590572	.1405613	1.13	0.258	-.116824	.4349383
_Icountry_43	-5.745626	.1896128	-30.30	0.000	-6.117781	-5.373471
_Icountry_44	-2.808813	.1626408	-17.27	0.000	-3.12803	-2.489596
_Icountry_45	-6.318985	.1933979	-32.67	0.000	-6.698569	-5.939401
_Icountry_46	-6.078009	.186407	-32.61	0.000	-6.443872	-5.712146
_Icountry_47	-4.835554	.1951112	-24.78	0.000	-5.2185	-4.452607
_Icountry_48	-3.604059	.1699121	-21.21	0.000	-3.937548	-3.270571
_Icountry_49	-3.502084	.1467591	-23.86	0.000	-3.790129	-3.214038
_Icountry_50	-6.275717	.2098913	-29.90	0.000	-6.687672	-5.863761
_Icountry_51	-2.263005	.1444646	-15.66	0.000	-2.546547	-1.979463
_Icountry_52	-5.623011	.2386085	-23.57	0.000	-6.091331	-5.154692
_Icountry_53	-5.827943	.2225883	-26.18	0.000	-6.264819	-5.391067
_Icountry_54	-3.20217	.2050755	-15.61	0.000	-3.604674	-2.799666
_Icountry_55	-2.365267	.1489668	-15.88	0.000	-2.657646	-2.072889
_Icountry_56	-6.693486	.2333887	-28.68	0.000	-7.15156	-6.235412
_Icountry_57	-3.62325	.1932562	-18.75	0.000	-4.002556	-3.243944
_Icountry_58	-5.917042	.2131743	-27.76	0.000	-6.335442	-5.498643
_cons	20.74965	.3091871	67.11	0.000	20.14281	21.3565