Effect of Unemployment Rate on GDP per Capita

Ali Takahashi

1003737361

Donna Maria Matri

1002693917

April 7, 2019

Contents

1 Introduction

GDP per capita is an important indicator of economic performance and is considered a useful unit to make comparisons of average living standards and economic well being. It is common knowledge that GDP per capita is negatively correlated with the unemployment rate, since GDP is negatively affected for every increase in unemployment rate. However, the extent to which GDP per capita is affected by unemployment rate has been an ongoing question that is yet to be answered extensively.

In this paper, our main aim is to examine the causal effect of unemployment rate on GDP per capita. Unemployment is an important issue in many economies, since high unemployment implies that labour resources are not being used efficiently. However, there are significant country and time effects on GDP per capita, so we have utilized panel data instead of cross sectional data in order to make our results more relevant. We mainly focused on the time period between 2000 - 2015 based on availability and relevance of data, and looked at over 100 countries for our model. We conducted four regression models in this paper; Pooled OLS, First Difference, Fixed Effects, and Random Effects to compare the difference in results between these four models. We will discuss more in detail regarding which estimate is the most suitable based on our model and data set later in this paper.

The rest of the paper is organized as follows: The next section places our analysis in the context of the of the literature. Section 3 describes the data, along with rationale behind why we chose specific variables in our analysis; Section 4 introduces the econometric model with explanations regarding why our method is appropriate; Section 5 presents a discussion of the empirical results, and Section 6 concludes.

2 Relation to Literature

For many years researchers have examined the relationship between GDP per capita and various factors. Naji Meidani and Maryam Zahibi (2011), for example, analyze the relationship between the unemployment rate and GDP per capita for the period 1971 to 2006 by using an auto regressive distributed lag approach (ARDL). The results from the ARDL model revealed that unemployment rate, physical capital, consumer price index, and ratio of government expenditure are statistically significant in determining GDP per capita. They have also determined that the unemployment rate and and the consumer price index are negatively related with GDP per capita, while physical capita and ratio of government expenditure to GDP are positively related with GDP per capita in Iran. We hope to give an overall effect of unemployment rate on GDP per capita worldwide by estimating our model which takes into account multiple countries as opposed to focusing only on Iran.[1]

Daron Acemoglu and Simon Johnson (2006) investigated the effect of life expectancy at birth on economic growth. The results indicated that increase in life expectancy led to a significant increase in population, but they were not able find evidence that the increase in life expectancy led to faster growth of income per capita. It seemed that the increase in life expectancy led to reduced income per capita at first, with the negative effect slowly wearing off over the years. However, it is important to take into account the limitation of their results since their approach exploits the international epidemiological transition around the 1940s, so the results may not be applicable to the present world. We hope to improve on this by focusing on a more recent time period between 2000 - 2015.[2]

Joesph J. MacDougald II (2011) explores how internet use impacts four different measures

1. Ali A. Naji Meidani and Maryam Zabihi, "The Dynamic Effect of Unemployment Rate on Per Capita Real GDP in Iran," *International Journal of Economics and Finance*, 2011,

2. Joseph J. Macdougald, "Internet Use and Economic Development: Evidence and Policy Implications," *University of South Florida*, 2011,

of economic development which includes GDP per capita using several econometric techniques on multi-country panel data. The results suggested that countries benefit differently from increasing Internet use and the magnitude of the effect depends on the income level of the country. The author found that increasing internet use has a significant positive effect on all four measures of economic development in countries that have achieved middle-income status. An increase in the number of internet users by 10% was associated with a 3.2% increase in GDP per capita and an even greater 3.5% in low-income countries.[3]

3 Data Description

3.1 Dataset

We have assembled a panel data set which is strongly balanced that combines information about country, GDP per capita, unemployment rate, and other possible determinants of GDP per capita. Our analysis focuses on the time period 2000 - 2015 for 102 countries and we were able to obtain a comprehensive dataset composed of 1632 observations. One reason we use independently pooled cross sections is to increase the sample size. By pooling random samples from the same country, but at different points in time, we are able to obtain more precise estimators and test statistics with more power. Our variables of interest are unemployment rate, life expectancy at birth, labour force, duration of compulsory years of education, and internet users per 100 people. We notice there are significant differences between developing and developed countries in each of the variables listed in Table 1. For example, the country that possessed the lowest GDP per capita within the dataset is Congo Dem. Rep in the year 2002, and the country that possessed the maximum value is Luxembourg in the year 2007. The results from the summary statistics strongly imply that there are significant country

3. Daron Acemoglu and Simon Johnson, "Disease and Development: The Effect of Life Expectancy on Economic Growth," *National Bureau of Economic Research*, 2006,

and yearly effects affecting GDP per capita. To reflect the fact that each country may have different distributions in different years, we allow the intercept to differ across the periods. This is accomplished by including time dummy variables in our model.

3.2 Dependent Variable

We are interested in measuring the effect of unemployment on GDP per capita, so we have obtained data for 102 countries on GDP per capita from the World Development Indicators database on World Bank. Since GDP per capita in current U.S. dollars (current series) is influenced by the effect of price inflation, we have utilized GDP per capita in constant 2010 U.S. dollars (constant series) which adjusts for the effects of price inflation.

3.3 Independent Variables

3.3.1 Unemployment Rate

We have obtained data on unemployment rate on 102 countries from the International Labour Organization, ILOSTAT database. It is measured as a percentage of the total labour force. The series is part of the ILO estimates and is harmonized to ensure comparability across countries and over time by accounting for differences in country-specific factors.

3.3.2 Life Expectancy at Birth

We have obtained data on life expectancy at birth on 102 countries from the United Nations Population Division , World Population Prospects database. This series indicates the number of years a newborn infant would live if prevailing patterns of mortality at the time of its birth were to stay the same throughout its life.

3.3.3 Labour Force

We have obtained data on labour force on 102 countries from the International Labour Organization, ILOSTAT database. It comprised people aged 15 and older who supply labor for the production of goods and services during a specified period. We have decided to use this series instead of total population since we are more interested in determining the effects labour force has on GDP per capita. Labour force contributes more to overall GDP, and determined that this series would be more relevant for the purpose of this paper. For our model, we will divide labour force by 10,000 for simplicity of units. For example, a 2.2 data value in the data set implies that the total labour force is 22,000.

3.3.4 Compulsory Years of Education

We have obtained data on compulsory years of education on 102 countries from the UNESCO Institute for Statistics. It measures the number of years that children are legally obliged to attend school. We included this series as a variable, since level of education has a significant effect on GDP per capita. Ideally, we would have liked to use data on percentage of net secondary school enrollment; however, due to limited availability of data, we have decided to use this series as a proxy for education.

3.3.5 Internet Users per 100 People

We have obtained data on internet users per 100 people on 102 countries from the Education Statistics database on World Bank. It measures the number of internet users per 100 people. Growth in number of internet users is known to have a positive effect on GDP, so we believe that adding this variable in our regression model would make our results more relevant.

4 Econometric Framework and Model

In this paper, our purpose is to examine the effect of unemployment rate on GDP per capita. Therefore, GDP per capita is the dependent variable with the unemployment rate being the independent variable as the main parameter of interest. To minimize the problem that could result from Omitted Variable Bias, we will include life expectancy at birth, labour force, compulsory years of education, and internet users per 100 people along with unemployment rate in the regression model. Since we are dealing with a panel dataset consisting of 102 countries and 15 years between 2000-2015, we expect there to be significant country and annual effects on our dependent variable.

Based on our analysis and data obtained, we have derived an econometric model given by:

$$y_{it} = \beta_1 x_{1it} + \beta_2 x_{2it} + \beta_3 x_{3it} + \beta_4 x_{4it} + \beta_5 x_{5it} + \beta_k d_t + \alpha_i + \epsilon_{it} \tag{1}$$

where y_{it} is GDP per capita of country i in year t, x_{1it} is unemployment rate, x_{2it} is life expectancy at birth of country, x_{3it} is labour force of country, x_{4it} is compulsory years of education, and x_{5it} is internet users per 100 people. d_t is a time dummy to allow the intercept of the regression to vary over time (we included a dummy variable for each year between 2001-2015 in our model), α_i is the country-specific fixed effect, and ϵ_{it} is the error component. β_1 is our main parameter of interest, and it measures the change in GDP per capita resulting from a percentage increase in unemployment rate.

The key assumption we make in our model is that the idiosyncratic errors are uncorrelated with the explanatory variables in each time period.

$$cov(x_{it}, \epsilon_{it}) = 0 \tag{2}$$

That is, the explanatory variables are strictly exogenous after we take out the country specific effect, α_i.

First, we conduct a Pooled OLS regression based on the econometric model derived above. However, we encounter a problem with this OLS regression. In order for our OLS estimates to be unbiased and consistent, we require the covariance between our independent variables and the error term to be zero.[4] We quickly recognize that the covariance between γ_{it} and our independent variables does not equal zero [5] with a high probability, since α_i has a significant effect on variables such as unemployment rate, life expectancy at birth, and internet users. Different countries have different levels of technology and health care, and those factors would strongly affect an independent variable such as life expectancy.

To take into account the unobserved heterogeneity term, we then conducted the First Differences regression model which is given by:

$$\Delta y_{it} = \beta_1 \Delta x_{1it} + \beta_2 \Delta x_{2it} + \beta_3 \Delta x_{3it} + \beta_4 \Delta x_{4it} + \beta_5 \Delta x_{5it} + \beta_k \Delta d_t + \Delta \epsilon_{it} \tag{3}$$

Since the α_i in our model does not vary across time, we are able to remove the unobserved heterogeneity term which allows us to obtain consistent estimates based on our assumption from (2) [6], and assming no serial correlation of $\Delta\epsilon_{it}$. We observe there to be significant variance in the independent variables across both year t and country i, so we have satisfied all conditions to obtain consistent estimates by running pooled OLS on this first difference system. In this case, β_1 measures the change in variation of GDP per capita for every one unit increase in the change in unemployment rate.

We then conduct both the Fixed Effects and Random Effects regression models. For the fixed effects model, we are able to remove the unobserved heterogeneity term similar to the First Difference model and we are able to obtain consistent estimates. However, the

4. Let's denote the error term as $\gamma_{it} = \alpha_{it} + \epsilon_{it}$
5. $\text{cov}(\gamma_{it},\, x_{it}) = \text{cov}(\alpha_i + \epsilon_{it}, x_{it})$
6. $\text{cov}(\Delta x_{it},\, \Delta\epsilon_{it}) = 0$

Random Effects regression model is based on the assumption that the independent variables are uncorrelated with country-specific effects in our model[7], which we already explained is very unlikely given our dataset and variables. Running the Random Effects regression on our model would give inconsistent estimates which is a concerning issue. To give empirical support to this statement, we also conduct the Hausman Test and found that country-specific effects α_i is indeed correlated with the independent variables x_{it} implying that Fixed Effects would be more suitable over Random Effects.

5 Empirical Results

Table 2 presents the coefficients for the regression models used in our analysis and mentioned in equation (1) and (3). First, we began by running the Pooled OLS estimation model using robust standard errors. We find that the covariances between the response and explanatory variables does not equal zero; indicating that the country-specific fixed effect, α_i, is correlated with the explanatory variables. The Pooled OLS estimation model states that a 1% increase in the unemployment rate results in a 356.7 dollars decrease in GDP per capita. Due to correlation between our regressors and α_i, our β_1 estimate is biased and gives a high standard error.

5.1 First Difference Model

Using the Breusch-Pagan test, we find evidence for heteroskedasticity and to adjust for this, we then conducted the First Differences regression model under the assumption that there are no unobservable time fixed effects. We find that for every unit increase in the change in unemployment rate there is a decrease in the variation of GDP per capita by 22.5.4 dollars and this result is statistically significant at the 0.1% significance level. Additionally, the direction

7. $\text{cov}(\alpha_i, x_{it}) = 0$

of the coefficient supports our initial hypothesis as it shows a negative effect on GDP per capita for every unit change in unemployment. This result has a relatively large economic impact on GDP per capita; however, it has a lower effect in comparison to the Pooled OLS estimation model.

5.2 Fixed Effect and Random Effect Estimation Models

When we ran the Fixed Effect regression model, we checked for heteroskedasticity in the error terms, and found evidence to reject the null hypothesis of homoskedastic standard errors. Therefore, we ran the Fixed Effect with robust standard errors to account for this. The coefficient estimate from the Fixed Effect regression shows a negative correlation between unemployment rate and GDP per capita and its value implies that a single unit increase in the unemployment rate decreases GDP per capita by 160.7 dollars.

The random effect model gives a 163.8 dollars decrease in GDP per capita for every one unit increase in unemployment. This estimate is statistically significant at the 0.1% significance level; however, the assumption that the country-specific fixed effect are uncorrelated with the explanatory variables does not hold in this case and thus, we conclude that this estimate is inconsistent.

We ran the Hausman test, as outlined in the Econometric Framework and Model section and the results of the test were conclusive and in favor of the Fixed Effects model. This means that variation across countries is not random and is correlated with the dependent and independent variables.

Since we choose the Fixed Effect estimate, this rules out the Pooled OLS estimation model. Thus, comparing the First Difference and Fixed Effect models we observe that the First Difference model has a lower standard error for β_1 than the Fixed Effect estimate. This suggests that the predicted values from the First Difference model are closer to the regression line and more accurate. The between R^2 value for the Fixed Effect model implies that around 25% of the variation is due to differences between the countries, whereas the overall R^2 value suggests that our model accounts for 4.03% of the total variation. In the First Difference model, the value of R^2 implies that the model accounts for 11.27% of the variation.

With the Fixed Effect estimation model, all explanatory variables, other than the unemployment rate, are deemed insignificant. The estimates for labour force and internet users; however, are significant at the 1% and 5% level, respectively, according to the First Difference model. We observe differences in the signs of the coefficients across the different models for life expectancy and internet users. Those explanatory variables often depend on the income levels of different countries which might explain the direction of the estimates. Comparing all the estimation models, the First Difference model produces the most significance across the explanatory variables. Acemoglu and Johnson (2007) state that an increase in life expectancy leads to a reduced level of income per capita. In the Fixed Effect model, we find life expectancy has a negative effect on the change in GDP per capita which agrees with the study. Joseph MacDougald (2011) suggests that an increase in the number of internet users increases GDP per capita, and our results reinforced those conclusions. From Tables 3 and 4, it can be seen that time variables are positively correlated with GDP per capita. The coefficients for each year are all statistically significant, and have economically large impacts. This intuitively make sense since GDP per capita has increased with time throughout the world with innovations in technology and other factors.

Although our model gave significant results, it is not immune to limitations. The key assumption this model depends heavily on is that the explanatory variables are strictly exogenous as mentioned in (2). Our regressors would be strictly exogenous if we took into account every possible variable that could affect GDP per capita; however, it is unlikely to completely account for this since we were limited in our selection of regressors due to the limited availability of data. An interesting topic to look at would be the link between population age structure and differences in levels of economic development, since age structure differences account for a significant variation in worldwide per capita GDP. Another limitation we encountered in our model is that due to limited data, we were inclined to use certain proxies as regressors which may not have been most suitable to our model. This is especially seen in our selection of compulsory years of education as a proxy for education. From Table 2, it can be seen that the coefficients from all four regressions are statistically insignificant, and the standard errors are much higher than those of other regressors. We believe that we could improve on this if we are able to find a more suitable variable that could replace this proxy in our model.

6 Conclusion

In this paper, we used panel data and set an econometric framework for estimating the impact of unemployment rate on GDP per capita. Our analysis consisted of a number of explanatory variables including unemployment rate, life expectancy at birth, labour force, years of education and internet users in 102 countries. Across all estimation models, we found that there is a negative correlation between unemployment rate and GDP per capita which reinforces the literature studies. Firstly, we began by running the Pooled OLS regression which produced statistically significant, but biased, estimators for all explanatory variables. To take this into account, we ran the First Difference and Fixed Effects regression models to obtain estimates that are more relevant. The two models allowed us to remove the unobserved heterogenity term which allowed our estimates to be unbiased and inconsistent based on assumptions imposed on our model. We have also conducted the Random Effects regression model; however, this assumes the unobserved heterogeneity term to be uncorrelated with the explanatory variables which is unrealistic for our dataset.

Comparing all the estimation models, we conclude that the First Difference model is the most suitable model for our dataset. Our hypothesis also explored that increasing Internet usage has a positive effect on GDP per capita. We also find that life expectancy at birth has a negative correlation with GDP per capita. Explanatory variables such as life expectancy and internet users can affect this study differently depending on the income level of the country which might explain the difference in the estimates for those variables. For example, an increase in the number of Internet users would have a higher increase in GDP per capita in low-income countries as suggested in MacDougald's paper (2011).

The framework presented here can be used to help assess how much unemployment actually impacts GDP per capita. If unemployment has such an economically large effect on GDP, then we need to find ways to utilize our labour resources more efficiently. An increase in life expectancy impacts this as it raises population, which would decrease the capital-to-labour ratio affecting productivity (MacDougald, 2011). However, with adequate supply of land and capital, this can lead to a positive effect on income and GDP per capita in the long-run, which could reduce the negative effect of unemployment on GDP.

References

Acemoglu, Daron, and Simon Johnson. "Disease and Development: The Effect of Life Expectancy on Economic Growth." *National Bureau of Economic Research*, 2006.

Macdougald, Joseph J. "Internet Use and Economic Development: Evidence and Policy Implications." *University of South Florida*, 2011.

Meidani, Ali A. Naji, and Maryam Zabihi. "The Dynamic Effect of Unemployment Rate on Per Capita Real GDP in Iran." *International Journal of Economics and Finance*, 2011.

Appendix

Table 1: Summary Statistics

	Mean	Std. Dev.	Min	Max
GDP per capita	16,196.14	20,430.10	276.25	111,968.40
Unemployment Rate	7.83	4.95	0.49	29.77
Life Expectancy at Birth	71.53	8.43	45.91	83.79
Labour Force	2,587.32	8,908.79	12.80	78,707.32
Compulsory Years of Education	9.38	2.07	5	15
Internet Users per 100 People	32.54	28.72	0.01	97.33
Observations	1632	1632	1632	1632

Table 2: Regression Results for Independent Variables I

	(1)	(2)	(3)
	OLS	FE	RE
Unemployment Rate	-356.7***	-160.7***	-163.8***
	(53.67)	(43.27)	(42.15)
Life Expectancy at Birth	228.5***	-63.34	129.1
	(41.02)	(60.68)	(69.48)
Labour Force	-0.0935***	-0.275	-0.205
	(0.0144)	(0.220)	(0.117)
Compulsory Years of Education	-589.3***	57.82	72.66
	(175.7)	(104.2)	(104.5)
Internet Users per 100 People	602.4***	24.87	43.56*
	(20.08)	(20.83)	(20.45)
Observations	1631	1631	1631
R^2	0.667	0.0403	0.382

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3: Regression Results for Independent Variables II

	(4)
	FD
Change in Unemployment Rate	-225.4***
	(21.03)
Change in Life Expectancy at Birth	-95.11
	(55.14)
Change in Labour Force	-0.204**
	(0.0748)
Change in Compulsory Years of Education	-71.19
	(81.85)
Change in Internet Users per 100 People	14.01*
	(5.990)
Observations	1528
R^2	0.113

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4: Regression Results for Time Dummies I

	(1) OLS	(2) FE	(3) RE
Year 2001	-1438.2	93.29*	-17.01
	(1662.6)	(46.36)	(43.16)
Year 2002	-3459.1*	238.7*	7.284
	(1642.9)	(104.4)	(92.86)
Year 2003	-4917.5**	408.1*	66.44
	(1606.7)	(170.6)	(153.2)
Year 2004	-6355.9***	773.4**	295.7
	(1622.5)	(248.2)	(214.9)
Year 2005	-7557.0***	1034.7***	453.2
	(1622.6)	(290.3)	(243.9)
Year 2006	-8972.8***	1397.9***	691.4*
	(1650.3)	(345.6)	(282.1)
Year 2007	-10627.0***	1748.9***	918.5**
	(1666.8)	(392.2)	(311.8)
Observations	1631	1631	1631
R^2	0.667	0.0403	0.382

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5: Regression Results for Time Dummies II

	(1) OLS	(2) FE	(3) RE
Year 2008	-12573.3***	1621.5***	659.8
	(1678.0)	(415.5)	(339.6)
Year 2009	-14648.5***	997.1*	-93.46
	(1660.3)	(439.2)	(395.3)
Year 2010	-16270.4***	1306.8*	88.76
	(1685.4)	(513.3)	(454.8)
Year 2011	-17895.3***	1496.7*	139.2
	(1723.5)	(579.9)	(512.4)
Year 2012	-19690.9***	1469.1*	-0.348
	(1768.6)	(649.4)	(580.1)
Year 2013	-21108.4***	1572.3*	-12.68
	(1799.5)	(713.1)	(637.7)
Observations	1631	1631	1631
R^2	0.667	0.0403	0.382

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6: Regression Results for Time Dummies III

	(1)	(2)	(3)
	OLS	FE	RE
yr2014	-22643.1***	1777.1*	77.49
	(1852.1)	(773.8)	(684.5)
yr2015	-24150.8***	2037.4*	246.8
	(1908.2)	(842.5)	(741.3)
Observations	1631	1631	1631
R^2	0.667	0.0403	0.382

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$