
GROWTH AND INEQUALITY

2ND PART OF ASSIGNMENT: REGRESSIONS

Helen Kermiser, 23rd January 2015

1. INTRODUCTION

With this work, I want to check whether inequality has an effect on GDP growth and if yes, if it is positive or negative. I do so by trying to replicate what Cingano has done in 2014 in his paper for the OECD "Trends in Income Inequality and its Impact on Economic Growth".

I use the Gini coefficient as the inequality measure. The general idea of the Gini coefficient is to show the difference between the actual cumulative income distribution and perfect equality. If the Gini coefficient would be 0, everybody would get the same income; if the Gini was 1, one person would get all the income. The OECD offers three Gini measures: Gini (disposable income, post taxes and transfers), Gini (market income, before taxes and transfers) and Gini (gross income, before taxes). I use the Gini coefficient measured in terms of disposable income, i.e. the after-tax Gini index. It considers the effect of taxes and social spending.

The OECD offers one growth rate in its economic data set which is the growth rate of the crude GDP in US \$ at constant prices with a constant purchasing power parity (for the reference year 2010). This most often mentioned growth indicator controls for inflation but not for the size of the population. This can lead to a distorted picture since the population of countries is constantly changing - and not necessarily getting bigger. For example, countries such as Japan are well-known for their shrinking population. Consequently, I introduced the growth rate of the GDP per capita to control for both inflation and demographics. In contrast, Cingano uses 5 year intervals of the logarithm of real GDP per capita to account for growth.

2. CORRELATION

As an introduction, I start with the correlation between GDP growth and the Gini coefficient. Three growth of GDP measures have been discussed in the guided group.

```
> cor(data.01$Real_GDP_cap_growth, data.01$GINI_disp_income,
use="pairwise.complete.obs")
[1] -0.004260325
> cor(log(data.01$Real_GDP_per_capita), data.01$GINI_disp_income,
use="pairwise.complete.obs")
[1] -0.3036656
> cor(data.01$GDP_growth, data.01$GINI_disp_income, use="pairwise.complete.obs")
[1] 0.0602686
```

If we compare the results of a simple correlation analysis for the three indicators we find that they yield deviating algebraic signs. A higher Gini coefficient is positively correlated with the standard measure for GDP growth but negatively with the GDP per capita growth and log of GDP per capita. However, the absolute value is 14 times bigger for the standard GDP growth measure than for the GDP per capita growth. The latter is very small at only -0.004260325.

The correlation of log of GDP per capita and the Gini is the highest at -0.3036656 but the log of GDP cannot be used as a measure of GDP growth like this but only if taken as a difference to the log of GDP per capita at another time. Cingano does so for a time interval of 5 years. If we use this growth rate and correlate it with the 5 year lag of the Gini we get -0.01005944, i.e. a negative correlation, too.

```
> cor(log(data.01$Real_GDP_per_capita)-log(lag(data.01$Real_GDP_per_capita,5)),
lag(data.01$GINI_disp_income,5), use="pairwise.complete.obs")
[1] -0.01005944
```

3. SIMPLE LINEAR REGRESSION

I start with a simple linear regression following in Cingano's footsteps - at first only with the Gini as an independent variable and secondly, with more regressors, i.e. investments and schooling.

3.1. WITH GINI INEQUALITY INDICATOR AS AN INDEPENDENT VARIABLE

At first, I compute a simple linear model with the cumulative GDP growth over five years as the dependent variable and the 5 year lagged Gini, i.e. the Gini in the beginning of the interval, as the independent variable.

```
> lm_Gini_log_GDP_01b <- lm((log(data.01$Real_GDP_per_capita)-
log(lag(data.01$Real_GDP_per_capita,5))) ~ (lag(data.01$GINI_disp_income,5)))
> summary(lm_Gini_log_GDP_01b)
```

Call:

```
lm(formula = (log(data.01$Real_GDP_per_capita) -
log(lag(data.01$Real_GDP_per_capita,
5))) ~ (lag(data.01$GINI_disp_income, 5)))
```

Residuals:

```
      Min       1Q   Median       3Q      Max
-0.35651 -0.05195  0.00196  0.05546  0.23488
```

Coefficients:

```
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    0.06727    0.02490   2.701  0.00726 **
lag(data.01$GINI_disp_income, 5) -0.01491    0.08170  -0.182  0.85533
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 0.08103 on 329 degrees of freedom
(1199 observations deleted due to missingness)
```

```
Multiple R-squared:  0.0001012, Adjusted R-squared:  -0.002938
F-statistic: 0.0333 on 1 and 329 DF,  p-value: 0.8553
```

This simple model finds a coefficient of -0.01491 for the Gini. If it was a statistically significant, that would mean that a 0.01 higher Gini, i.e. more inequality, decreases cumulative growth of five years by 0.01491 percentage points. However, the p-value is very high at 0.8553. This means that the result is not significant and very likely to have occurred by chance. Also, the t-value is only -0.182 which also makes it very likely that the effect of the Gini occurred by random chance. In addition, R-squared is only 0.0001012 which means that the model explains barely 0.1 percent of the variability of the response data around its mean. Consequently, we need more regressors to estimate GDP growth.

3.2. WITH MORE REGRESSORS (SCHOOLING, INVESTMENT)

With regard to GDP growth, it makes sense to include regressors that allow us to control for human and physical capital. I follow Cingano and add investment and schooling as regressors (with a five year lag). Schooling can be seen as a proxy for different human capital endowments. Our schooling variable includes the average years of schooling of working age 15 to 64. For investment, we use the gross fixed capital formation which shows us how much of value added is invested rather than consumed. If we did not consider investment separately, high inequality and high growth in the same country would always support the neoclassical claim that high inequality makes rich people richer and - even though their marginal propensity to consume is lower - this increased savings which automatically become investments. In addition, I add - just like Cingano - the log of GDP per capita for the first period (i.e. initial income) of the time interval to control for convergence.

```
> lm_Gini_log_GDP_02a <- lm((log(data.01$Real_GDP_per_capita)-
log(lag(data.01$Real_GDP_per_capita,5))) ~ log(lag(data.01$Real_GDP_per_capita,5))
+ (lag(data.01$GINI_disp_income,5)) + lag(data.01$GFCF_GDP,5) +
lag(data.01$Schooling,5))
> summary(lm_Gini_log_GDP_02a)
```

```
Call:
lm(formula = (log(data.01$Real_GDP_per_capita) -
log(lag(data.01$Real_GDP_per_capita,
5))) ~ log(lag(data.01$Real_GDP_per_capita, 5)) +
(lag(data.01$GINI_disp_income,
5)) + lag(data.01$GFCF_GDP, 5) + lag(data.01$Schooling, 5))
```

```
Residuals:
    Min       1Q   Median       3Q      Max
-0.138460 -0.033632 -0.001531  0.034846  0.159262
```

```
Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)      1.160824   0.274125   4.235 7.85e-05 ***
log(lag(data.01$Real_GDP_per_capita, 5)) -0.088939   0.027287  -3.259  0.00183 **
lag(data.01$GINI_disp_income, 5) -0.290948   0.148497  -1.959  0.05465 .
lag(data.01$GFCF_GDP, 5) -0.667575   0.234779  -2.843  0.00606 **
lag(data.01$Schooling, 5)  0.005680   0.006419   0.885  0.37971
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 0.05877 on 61 degrees of freedom
(1464 observations deleted due to missingness)
Multiple R-squared:  0.2743,    Adjusted R-squared:  0.2267
F-statistic: 5.764 on 4 and 61 DF,  p-value: 0.0005306
```

The p-value is 0.0005306 and thereby much less than 0.05 which means we can reject the null hypothesis and that there is a significant relationship between the variables in the linear regression model of the data set. The R-squared of 0.2743 indicates that the model explains more than 27 percent of the variability of the response data around its mean. In addition, the residuals are reasonably good because their median -0.001531 is close to 0 and the two sides are relatively balanced.

This linear model with more regressors also finds a negative coefficient for the Gini. If it was a statistically significant, a 0.01 increase in the Gini would decrease cumulative five year growth by nearly 0.3 percentage points. With a t-value of -1.959 and a p-value of 0.05465 the Gini coefficient is nearly in the 95 percent confidence interval and therefore significant with a probability of nearly 95 percent. We have thus good reason to believe that the correlation between the Gini and growth is real.

However, we find even more significant effects for the effect of the initial income with a p-value of 0.00183 and a t-value of -3.259, and the effect of investment with a p-value of 0.00606 and a t-value of -2.843. Consequently, a higher initial income seems to yield lower growth (albeit only a little lower) and higher investments in the first period yield lower growth over the following years. The latter does not seem logical. One would rather expect a positive effect of investments in the first period. However, the schooling variable seems to be positive but is insignificant with a p-value of 0.37971. All in all, this linear model with more regressors has considerably more significance than the first linear model with only one regressor.

Exclusion of insignificant independent variables

If we exclude Schooling from the linear regression model, it becomes even more significant with a p-value smaller than 2.2e-16. The residuals are also reasonably good and the R-squared indicates that the model explains more than 24 percent of the variability of the response data around its mean. Also, the difference between R-squared and R-squared adjusted is reduced considerably.

```
> lm_Gini_log_GDP_02b <- lm((log(data.01$Real_GDP_per_capita) -
log(lag(data.01$Real_GDP_per_capita,5)))) ~ log(lag(data.01$Real_GDP_per_capita,5))
+ (lag(data.01$GINI_disp_income,5)) + lag(data.01$GFCF_GDP,5))
> summary(lm_Gini_log_GDP_02b)
```

Call:

```
lm(formula = (log(data.01$Real_GDP_per_capita) -
log(lag(data.01$Real_GDP_per_capita,
5)))) ~ log(lag(data.01$Real_GDP_per_capita, 5)) +
(lag(data.01$GINI_disp_income,
5)) + lag(data.01$GFCF_GDP, 5))
```

Residuals:

```
      Min       1Q   Median       3Q      Max
-0.33643 -0.03857  0.00126  0.03890  0.21374
```

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	1.25269	0.13238	9.463	< 2e-16 ***
log(lag(data.01\$Real_GDP_per_capita, 5))	-0.09379	0.01176	-7.975	2.59e-14 ***
lag(data.01\$GINI_disp_income, 5)	-0.25163	0.07506	-3.352	0.000895 ***

```

lag(data.01$GFCF_GDP, 5)                -0.66991    0.10289  -6.511 2.81e-10 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.07071 on 327 degrees of freedom
(1199 observations deleted due to missingness)
Multiple R-squared:  0.2434,    Adjusted R-squared:  0.2364
F-statistic: 35.06 on 3 and 327 DF,  p-value: < 2.2e-16

```

With this modification, the remaining three independent variables become so significant that we can conclude that they do not only seem to have an effect but have an effect in this model. While it can be explained theoretically that a higher initial GDP and a higher initial Gini have negative effects on growth, this is very hard for investment which retains its negative coefficient, too.

3.3. DIFFERENT TIME TREATMENTS

One interesting point in econometrics is how to treat time. So far, we have ignored it. However, we could add a time trend variable or a time dummy variable as a regressor. I will use the linear model without schooling to do so.

3.3.1. TIME DUMMY

A dummy variable gives us a year-specific effect that affects all our units. This is most appropriate if we suspect that there are specific effects to one year, and want to model them. For example, this could be a global shock. We include such a time fixed effect by including `factor(data.01$Year)` in our model.

```

> lm_Gini_log_GDP_02c <- lm((log(data.01$Real_GDP_per_capita) -
log(lag(data.01$Real_GDP_per_capita, 5))) ~ log(lag(data.01$Real_GDP_per_capita, 5))
+ (lag(data.01$GINI_disp_income, 5)) + lag(data.01$GFCF_GDP, 5) +
factor(data.01$Year))
> summary(lm_Gini_log_GDP_02c)

```

```

Call:
lm(formula = (log(data.01$Real_GDP_per_capita) -
log(lag(data.01$Real_GDP_per_capita,
5))) ~ log(lag(data.01$Real_GDP_per_capita, 5)) +
(lag(data.01$GINI_disp_income,
5)) + lag(data.01$GFCF_GDP, 5) + factor(data.01$Year))

```

```

Residuals:
    Min       1Q   Median       3Q      Max
-0.294148 -0.035881  0.000042  0.031881  0.222671

```

```

Coefficients:
(Intercept)                Estimate Std. Error t value Pr(>|t|)
log(lag(data.01$Real_GDP_per_capita, 5)) -0.076990  0.011680  -6.591 2.03e-10 ***
lag(data.01$GINI_disp_income, 5)         -0.156096  0.074570  -2.093 0.03719 *
lag(data.01$GFCF_GDP, 5)                 -0.313507  0.109324  -2.868 0.00444 **
factor(data.01$Year)1980                 -0.100926  0.063098  -1.600 0.11079
factor(data.01$Year)1981                 -0.052491  0.076516  -0.686 0.49325
factor(data.01$Year)1982                 -0.125395  0.062691  -2.000 0.04640 *
factor(data.01$Year)1983                 -0.131479  0.076592  -1.717 0.08711 .
factor(data.01$Year)1984                 -0.092763  0.062560  -1.483 0.13921
factor(data.01$Year)1985                 -0.057177  0.062569  -0.914 0.36157
factor(data.01$Year)1986                 -0.053631  0.062540  -0.858 0.39184
factor(data.01$Year)1987                  0.002747  0.062530   0.044 0.96499
factor(data.01$Year)1988                 -0.001986  0.057279  -0.035 0.97236

```

```

factor (data.01$Year) 1989      -0.070718    0.054041   -1.309    0.19170
factor (data.01$Year) 1990      -0.045004    0.049705   -0.905    0.36599
factor (data.01$Year) 1991      -0.062040    0.050521   -1.228    0.22044
factor (data.01$Year) 1992      -0.124109    0.052399   -2.369    0.01851 *
factor (data.01$Year) 1993      -0.144008    0.054542   -2.640    0.00873 **
factor (data.01$Year) 1994      -0.124473    0.054524   -2.283    0.02316 *
factor (data.01$Year) 1995      -0.091608    0.049648   -1.845    0.06603 .
factor (data.01$Year) 1996      -0.092708    0.050560   -1.834    0.06773 .
factor (data.01$Year) 1997      -0.058748    0.050562   -1.162    0.24622
factor (data.01$Year) 1998      -0.024996    0.051380   -0.486    0.62698
factor (data.01$Year) 1999      -0.036582    0.048375   -0.756    0.45013
factor (data.01$Year) 2000      -0.008330    0.047521   -0.175    0.86097
factor (data.01$Year) 2001       0.005845    0.049239    0.119    0.90559
factor (data.01$Year) 2002      -0.015762    0.050446   -0.312    0.75491
factor (data.01$Year) 2003      -0.032595    0.050459   -0.646    0.51880
factor (data.01$Year) 2004      -0.019690    0.049121   -0.401    0.68882
factor (data.01$Year) 2005      -0.054084    0.046965   -1.152    0.25044
factor (data.01$Year) 2006      -0.028200    0.050616   -0.557    0.57787
factor (data.01$Year) 2007      -0.007060    0.050439   -0.140    0.88878
factor (data.01$Year) 2008      -0.043567    0.049113   -0.887    0.37577
factor (data.01$Year) 2009      -0.083044    0.046233   -1.796    0.07350 .
factor (data.01$Year) 2010      -0.087127    0.046531   -1.872    0.06214 .
factor (data.01$Year) 2011      -0.097376    0.046613   -2.089    0.03757 *
factor (data.01$Year) 2012      -0.130309    0.046740   -2.788    0.00565 **
factor (data.01$Year) 2013      -0.133344    0.046305   -2.880    0.00428 **
factor (data.01$Year) 2014      -0.076808    0.045852   -1.675    0.09498 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

```

Residual standard error: 0.06233 on 292 degrees of freedom
(1199 observations deleted due to missingness)
Multiple R-squared:  0.4749,    Adjusted R-squared:  0.4066
F-statistic: 6.951 on 38 and 292 DF,  p-value: < 2.2e-16

```

If we include time dummies into our model the p-value stays very low at $< 2.2e-16$ while the R-squared and adjusted R-squared improve considerably. The sign of our initial coefficients stays the same but two of them (the Gini and investments) lose in significance but stay significant within a 0.05 confidence interval. There are many year dummies without a significant effect at common confidence intervals while the years 1982, 1983, 1992 to 1996 and 2009 to 2014 seem to have an effect within a 0.05 confidence interval. However, most of these effects are not very sizeable.

5.2. TIME TREND

If the effect if time is not specific to a particular given year it can also build up across years. Such a trend means that moving from one year to the other (i.e. increasing time by one unit) yields effects on our outcome variable. In our case, this could be technological progress.

Originally, Year is integer but I want numeric for my calculation. So I make it numeric using `as.numeric`. Our first year is 1970, so we have to calculate `Year - 1969` to give the value 1 to the first year.

```

> Year <- as.numeric(data.01$Year)
> time_trend <- Year - 1969
> lm_Gini_log_GDP_02d <- lm((log(data.01$Real_GDP_per_capita)-
log(lag(data.01$Real_GDP_per_capita,5))) ~ log(lag(data.01$Real_GDP_per_capita,5))
+ (lag(data.01$GINI_disp_income,5)) + lag(data.01$GFCF_GDP,5) + time_trend)
> summary(lm_Gini_log_GDP_02d)

```

```

Call:
lm(formula = (log(data.01$Real_GDP_per_capita) -
log(lag(data.01$Real_GDP_per_capita,
5))) ~ log(lag(data.01$Real_GDP_per_capita, 5)) +
(lag(data.01$GINI_disp_income,
5)) + lag(data.01$GFCF_GDP, 5) + time_trend)

Residuals:
    Min       1Q   Median       3Q      Max
-0.33026 -0.03761 -0.00157  0.04175  0.20420

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)      -1.0875570   1.1061161   -0.983   0.3262
log(lag(data.01$Real_GDP_per_capita, 5)) -0.0834212   0.0126686  -6.585 1.82e-10 ***
lag(data.01$GINI_disp_income, 5)      -0.1969520   0.0789416  -2.495   0.0131 *
lag(data.01$GFCF_GDP, 5)              -0.5613564   0.1143162  -4.911 1.44e-06 ***
time_trend        -0.0011343   0.0005323  -2.131   0.0338 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.07033 on 326 degrees of freedom
(1199 observations deleted due to missingness)
Multiple R-squared:  0.2538,    Adjusted R-squared:  0.2446
F-statistic: 27.72 on 4 and 326 DF,  p-value: < 2.2e-16

```

If we add a time trend as an independent variable our intercept becomes negative and insignificant while the other regressors keep their algebraic sign while getting a smaller effect and staying significant (with Gini, however, becoming less significant). The time trend itself has a significant negative effect on growth which means that growth decreased over time since the 1970s. The effect is very small though at -0.0011343. The R-squared of the model increases only slightly while the p-value stays very small.

4. PANEL DATA MODEL

A regular ordinary least square regression (which is what I did in the first part) does not consider heterogeneity across groups or time. So in the second part, I work with panel data to control for these two things.

4.1. WITH GINI INEQUALITY INDICATOR AS AN INDEPENDENT VARIABLE

Again, firstly, I compute the linear model with the cumulative GDP growth over five years as the dependent variable and the 5 year lagged Gini, i.e. the Gini in the beginning of the interval, as the only independent variable.

```

> data.01$five_yr_GDP <- (log(data.01$Real_GDP_per_capita) -
log(lag(data.01$Real_GDP_per_capita,5)))
> View(data.01)
> lm_Gini_log_GDP_03a <- plm(five_yr_GDP ~ (lag(GINI_disp_income,5)), data=data.01,
index=c("Country", "Year"), model="within")
> summary(lm_Gini_log_GDP_03a)
Oneway (individual) effect Within Model

Call:
plm(formula = five_yr_GDP ~ (lag(GINI_disp_income, 5)), data = data.01,
     model = "within", index = c("Country", "Year"))

```

Unbalanced Panel: n=34, T=1-34, N=331

Residuals :
Min. 1st Qu. Median 3rd Qu. Max.
-0.28800 -0.03970 0.00329 0.04000 0.22800

Coefficients :
Estimate Std. Error t-value Pr(>|t|)
lag(GINI_disp_income, 5) -0.21608 0.23799 -0.9079 0.3647

Total Sum of Squares: 1.5512
Residual Sum of Squares: 1.5469
R-Squared: 0.0027772
Adj. R-Squared: 0.0024835
F-statistic: 0.824326 on 1 and 296 DF, p-value: 0.36466

This model finds a coefficient of -0.21608 for the Gini. If it was a statistically significant, that would mean that a 0.01 higher Gini, i.e. more inequality, decreases cumulative growth of five years by 0.21608 percentage points. However, the p-value is high with 0.36466. This means that the result is not significant and very likely to have occurred by chance. Also, the t-value is only -- 0.9079 and R-squared only 0.0027772 which means that the model explains barely 0.2 percent of the variability of the response data around its mean. Consequently, the model can be rejected.

4.2. WITH MORE REGRESSORS (SCHOOLING, INVESTMENT)

If we include more regressors into the model, the algebraic sign of Gini's coefficient changes to positive though it stays insignificant. In contrast, Schooling's coefficient becomes negative while staying insignificant. The estimated effect of initial GDP per capita stays negative (as in the simple linear model) but insignificant at most used confidence levels. With a total number of 66 observations we are actually missing data for a lot of periods. The general p-value of the model is very low at 0.00013131 with the coefficient of investments being significant (in a 0.001 confidence interval) and negative. The difference between R-squared and R-squared adjusted is very high which does not speak in favour of the model. We should look for a better one.

```
> lm_Gini_log_GDP_04f <- plm(five_yr_GDP ~ log(lag(Real_GDP_per_capita,5)) +  
(lag(GINI_disp_income,5)) + lag(GFCF_GDP,5) + lag(Schooling, 5), data=data.01,  
index=c("Country", "Year"), model="within")  
> summary(lm_Gini_log_GDP_04f)  
Oneway (individual) effect Within Model
```

```
Call:  
plm(formula = five_yr_GDP ~ log(lag(Real_GDP_per_capita, 5)) +  
(lag(GINI_disp_income, 5)) + lag(GFCF_GDP, 5) + lag(Schooling,  
5), data = data.01, model = "within", index = c("Country",  
"Year"))
```

Unbalanced Panel: n=24, T=1-6, N=66

Residuals :
Min. 1st Qu. Median 3rd Qu. Max.
-0.0648 -0.0316 0.0000 0.0245 0.1220

Coefficients :
Estimate Std. Error t-value Pr(>|t|)
log(lag(Real_GDP_per_capita, 5)) -0.1704580 0.1145811 -1.4877 0.145092
lag(GINI_disp_income, 5) 0.8730258 0.5719158 1.5265 0.135168
lag(GFCF_GDP, 5) -1.7710749 0.5653891 -3.1325 0.003331 **
lag(Schooling, 5) -0.0017653 0.0175880 -0.1004 0.920578

```

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Total Sum of Squares:    0.1949
Residual Sum of Squares: 0.10819
R-Squared:              0.44492
Adj. R-Squared:         0.25617
F-statistic: 7.61478 on 4 and 38 DF, p-value: 0.00013131

```

4.3. WITH OTHER EFFECTS

If we add the effect "twoways" to include the fixed effects of country (as above) and time, the model does not get considerably better with illogical signs of the coefficients and now the initial GDP per capita as the only significant independent variable. The difference between R-squared and R-squared adjusted is, again, big. The model does not seem convincing.

```

> lm_Gini_log_GDP_04h <- plm(five_yr_GDP ~ log(lag(Real_GDP_per_capita,5)) +
(lag(GINI_disp_income,5)) + lag(GFCF_GDP,5) + lag(Schooling, 5), data=data.01,
index=c("Country", "Year"), model="within", effect = "twoways")
> summary(lm_Gini_log_GDP_04h)
Twoways effects Within Model

```

```

Call:
plm(formula = five_yr_GDP ~ log(lag(Real_GDP_per_capita, 5)) +
      (lag(GINI_disp_income, 5)) + lag(GFCF_GDP, 5) + lag(Schooling,
      5), data = data.01, effect = "twoways", model = "within",
      index = c("Country", "Year"))

```

Unbalanced Panel: n=24, T=1-6, N=66

```

Residuals :
  Min. 1st Qu.  Median 3rd Qu.    Max.
-0.0843 -0.0113  0.0000  0.0125  0.0881

```

```

Coefficients :
              Estimate Std. Error t-value Pr(>|t|)
log(lag(Real_GDP_per_capita, 5)) -0.671187  0.155260 -4.3230 0.0001402 ***
lag(GINI_disp_income, 5)         0.585183  0.435581  1.3435 0.1885753
lag(GFCF_GDP, 5)                 -0.614486  0.477684 -1.2864 0.2075381
lag(Schooling, 5)                -0.013108  0.014883 -0.8808 0.3850101

```

```

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

```

Total Sum of Squares:    0.099999
Residual Sum of Squares: 0.046817
R-Squared:              0.53182
Adj. R-Squared:         0.25785
F-statistic: 9.08754 on 4 and 32 DF, p-value: 5.0665e-05

```

4.4. GMM INSTEAD OF PLM

However, Cingano does not use least square estimators but GMM.

5. CONCLUSION

With regard to the linear model with more regressors we can say that inequality seems to have a significant albeit small negative effect on growth. However, I find a more sizeable negative effect of investments in the first period on cumulative five year growth. This seems at odds with the

theory. It might help to review other investment measures. Also, it should be interesting to include measures of top or bottom inequality into the model. The treatment of time can help to make our model more sound. For example, we see that time dummies have a considerable effect on R-squared in the linear model.

While I can get to a similar conclusion to Cingano (2014) with regard to inequality and growth using the linear model, I did not succeed in doing so with a panel model. The models I have computed lack in significance and I did not succeed with the GMM.