The Reversal of Fortune Thesis Revisited (Research Proposal)

1 Aimed to achieve

Focusing on the *Reversal of Fortune* thesis proposed by AJR (2002), we personally agree with the institution hypothesis which emphasizes the role of institutions of private property in the determination of long-term economic divergence. However, the authors' quantitative methodology in this paper was biasedly implemented and the contribution of *European* settlement is severely overestimated. To better examine the conclusion of "the *British* settlement determines the institutional settings of property protection", we select and discuss three obvious problems appeared in AJR (2002) in order to determine whether the conclusion made in the paper was scientifically reasonable or not.

2 Motivation

What determines the long-term economic divergence and why do some parts of the world which were economically abundant in 1500 tend to be relatively poor now? This question was addressed in AJR (2002)[1] paper, where the authors concluded that the lack of full institutional protections[2] resulted in "Reversal of Fortune".

We agree with this proposition of "institution matters" and are determined to extend the original AJR (2002) by solving three problems. Firstly, the authors' measure of *urbanization in 1500* contains no data from Africa and thus only consists of 41 observations. Without the relevant data obtained from the African continent, the authors' *institution matters* arguments, to some extent, lose its theoretical generosity.

Secondly, authors' measurement of *population density in 1500* does not *accurately* take the total amount of arable lands into consideration. AJR (2002) implements data on arable lands from Mcevedy and Jones(1978)[3]. However, this arable land dataset generates quite curiosity. For example, as Bandyopadhyay and Green(2012) point out, "Mcevedy and Jones(1978) sometimes present data on arable lands inaccurately. In 86 out of 91 observations used by AJR(2002), they list no data on arable land, leading the authors of AJR(2002) to see all land as arable for these observations"[4].

Thirdly, AJR(2002) overestimates the role of European institutions. We believe that what leads to the authors' proposal that institutions of private property determine the long-term economic divergence is the factual economic prosperity among the *Neo-Europe*. Although European settlement does, to some extent, work as a positive force for better property protection, there is no evidence of a significantly general relation between European colonization, institutions of property rights and long-term economic development.

3 Research Design

3.1 Regression Equation

Our paper will adopt the same setting of regression equations just like AJR (2002). The whole regression estimation will be based upon an ordered combination between OLS and IV 2SLS analysis. Again, the puzzle of this paper is to re-estimate the institution hypothesis which claims that the institution of property right determines the long-term economic divergence between 1500 and current days.

Among the regular OLS estimation, our regression equation will be presented as follows:

$$y_{it} = \mu_t + \delta_i + \pi * X_{it} + Z_{it} + \epsilon_{it}$$

[1] Acemoglu, D., Johnson, S., & Robinson, J. A. (2002). Reversal of fortune: Geography and institutions in the making of the modern world income distribution. *The Quarterly journal of economics, 117(4)*, 1231-1294.

[2] In AJR (2002), such an institutional protection is named as *institutions of private property.*

[3] McEvedy, C., & Jones, R. (1978). *Atlas of world population history.* Penguin Books Ltd, Harmondsworth, Middlesex, England.

[4] Bandyopadhyay, S., & Green, E. (2012). The reversal of fortune thesis reconsidered. *Journal of Development Studies, 48(7)*, 817-831.

Here y_{it} is the outcome variable of current economic prosperity in country i at date t. δ_i and μ_t represents the country fixed effect and the time effect, respectively. X_{it} is our focused target, represented by *institution* measurements, which is defined as the measurement of institutions of country i at date t. Z_{it} stands for one combination of the controlled variables, which represents included variables like *urbanization in 1500* and *log population density in 1500*.

Among the IV 2SLS estimation, our combined set of regression equations is presented as follow:

$$\text{First} - \text{stage regression: } X_{it} = \gamma * \text{IV}_{it} + Z_{it} + v_{it}$$
$$\text{Second} - \text{stage regression: } y_{it} = \mu_t + \delta_i + \tilde{\pi} * \hat{X}_{it} + Z_{it} + U_{it}$$

Here IV $_{it}$ represents the presented IV variable in country i at date t. Our main focus is the direction and the statistical significance of coefficient $\tilde{\pi}$.

3.2 Data for Correction and Identification

To correct the measurement bias problem (*Problem 1*), we implement Chandler (1987)'s data on cities with a population of more than 2000 among the continent of Africa and America as an alternative variable measurement.[5] The implementation of this new dataset has two advantages. Firstly, all of this new dataset comes from a single source, which drastically decreases the potential error generated by the merge transformation. Secondly and more importantly, this new dataset allows us to include much detailed information with regard to African pre-modern population.

To correct the measurement inaccuracy problem (*Problem 2*), we implement data from FAO (2000).[6] Austin (2008)[7] suggests that "FAO (2000) is the first-time estimated global dataset for land that is potentially arable for growing any one of the 21 major crops under rain-fed conditions", which makes it possible for us to get access to the pre-modern agricultural societies where modern agriculture technology has not been adopted yet. One advantage of FAO (2000)'s implementation is that this data source allows us to *exclude* the amount of non-arable land, exemplified by the land of deserts, mountains, frozen areas and so on, that was mistakenly calculated in the measurement of arable land in AJR (2002).

In order to deal with the overestimation problem (*Problem 3*), we create a dummy variable *Britain* to represent whether one targeted state receives British colonization in the past history or not. If the answer is Yes, then variable *Britain* is coded as 1 , otherwise it is 0. In this case, we will reorganize the base OLS regression equation as follow:

$$y_{it} = \mu_t + \delta_i + \beta * Britain_{it} * X_{it} + Z_{it} + \epsilon_{it}$$

where $Britain_{it}$ represents whether one state I receives British colonization or not at year t. Our focus is the *direction* of coefficient β.

4 Time Table

We plan to finish the writing process of the research proposal before March 18^{th}. Then we will move on the recalibration of our merged dataset and decide to generate preliminary empirical result before March 26^{th}. Our goal is to finish the first draft of the whole written-up paper before April 1^{st} so that we still have time to polish the quality of this final project.

[5] Chandler, T. (1987). *Four thousand years of urban growth: An historical census.* Mellen.

[6] Bot, A., Nachtergaele, F., & Young, A. (2000). *Land resource potential and constraints at regional and country levels* (No. 90). Food & Agriculture Org.

[7] Austin, G. (2008). The 'reversal of fortune' thesis and the compression of history: perspectives from African and comparative economic history. *Journal of International Development: The Journal of the Development Studies Association, 20(8)*, 996-1027.