The Determinants of Economic Growth

Johann Power

London School of Economics and Political Science

14.01.2022

I. Introduction

Using a novel panel data set, this paper analyses the determinants of economic growth using a fixed effects multiple linear regression model with instrumental variables. Ultimately, contrary to traditional macroeconomic theory, almost no neoclassical growth factors prove to be statistically significant drivers of economic growth.

II. Background

Understanding the determinants of economic growth are important in helping improve standards of living. As Lucas (2000) stated: "How did the world economy of today, with its vast differences in income levels and growth rates, emerge from the world of two centuries ago, in which the richest and the poorest societies had income differing by perhaps a factor of two, and in which no society had ever enjoyed sustained growth in living standards?". There have been many empirical studies on the determinants of economic growth. Romer et al. (1992) provide cross-country evidence to support a labour-augmented Solow model. While Barossi-Filho et al. (2005) find updated evidence to support the predictions of the original Solow model (Solow, 1956). Such papers based on the grounds of a pre-established theoretical model can be more compelling and likely to imply causation. However, they often restrict themselves to analysing only a few determinants of economic growth. Economies constantly evolve and with them growth models have also evolved in their success at explaining economic growth. This paper incorporates additional possible growth determinants alongside the traditional neoclassical variables – notably by breaking down total factor productivity into its physical and social components. Johansson and Tretow (2015) present one such paper which analyses the role of economic freedom (what this paper would classify as a social productivity factor) on economic growth. In addition, many papers apply cross country regressions on all countries irrespective of their level of development. The Clark-Fisher three-sector model (Fisher, 1935) shows how economies structurally change over time and thus so do their main economic growth drivers. We have observed historically how modern day developed nations shifted from agricultural to industrial to service-based economies and so this paper also analyses the role of economic growth factors within groups of countries with similar levels of development.

The biggest problem facing empirical growth economists is the issue of endogeneity. Cross country growth regressions have been rightly criticized for often wrongly inferring causal relationships from their results (Durlauf, 2009). Economic growth factors are complementary and interact with one another, they are almost always not independent variables in a regression model. In a standard OLS regression model, the regression coefficient estimates are usually biased due to simultaneity and reverse causality. Additionally, the data often has measurement errors and there are too many possible growth drivers to consider compared to available data. The standard technique to tackle the endogeneity problem is using instrumental variables. Instruments eliminate bias by acting as a randomiser for the instrumented variable. Instead of randomly assigning the treatment, the instrument randomly generates variations in the treatment variable that we use to estimate the causal effect. For instance, population size has been used as an instrument in many growth regressions, instrumenting for variables such as trade (Spolaore and Wacziarg, 2005) and export sophistication (Hausmann et al. 2007). However, violations in the choice of instrument are often a problem as shown by Bazzi and Clemens (2013). This paper proposes valid instruments based on Cherif, Hasanov and Wang (2018) in using neighbouring countries variables as instruments for endogenous regressors.

III. Data

This paper formalizes output as follows:

$$y = f(k, h, p, s)$$

where:

$y =$ output per capita

$k =$ physical capital per capita

$h =$ human capital per capita

$p =$ physical productivity

$s =$ social productivity

One major difference between this model and other conventional growth models is that total factor productivity (TFP) is split into two components: physical and social productivity. This is done to observe each individual effect and together the datasets used to proxy each productivity factor roughly approximate the theoretically implied TFP level in an economy derived via growth accounting (shown in Appendix I). This paper refers to physical productivity as the technology which contributes to productive efficiency. For example, this includes more efficient machines in a production plant or advanced machine learning capabilities to automate parts of the production process. Whereas social productivity is the social, political or cultural elements which contribute to productive efficiency. For example, this could include the (work) culture, rule of law, role of institutions etc. in a country. Some notable growth determinants which are omitted are land and climate (as there is insufficient variation in land endowments or climate over the 13-year dataset period). Additionally, the impact of a country's climate on economic growth will be accounted for via country fixed effects. In total, this study uses annual data for each growth determinant for 78 countries between 2007 – 2019. GDP per capita is used as a measure of output per capita and is sourced from the World Bank. Gross fixed capital formation, sourced from the World Bank, is used as a measure of physical capital. The World Bank Human Capital Index (HCI) is used as a measure of human capital per capita. Kraay (2018) presents the methodology used in calculating the HCI but essentially it is a measure which assesses how well a country performs in survival, schooling and health. Physical productivity is proxied by the ICT Development Index (IDI) published by the United Nations which gives an approximation of the technology level within a country. The IDI uses 11 internationally agreed indicators, grouped into three sections: access, use and skills, to measure the developments in information and communication technology between countries and over time. This index is not a true measure of all physical productivity factors in an economy as that would be impossible to capture. However, like the other indices used in this paper, it offers a good estimate and makes for useful comparison across countries and time. The Economic Freedom Index (EFI) is used as a measure of social productivity. The EFI measures a country's degree of economic freedom based on four main categories: rule of law, government size, regulatory efficiency and market openness. Any missing data, which overall is low, has been interpolated via linear regression.

IV. Methodology

A. 2SLS Regression Model

Fixed effects multiple linear regression with instrumental variables is used to analyse the impact of each growth determinant. The equation we aim to estimate is:

$$\ln(y_{it}) = \beta_k \ln(k_{it}) + \beta_h \ln(h_{it}) + \beta_p \ln(p_{it}) + \beta_s \ln(s_{it}) + \beta_g (g_{it}) + c_i + d_t + u_{it}$$

Where g_{it} is a control variable representing the average GDP per capita growth rate in a country's neighbouring countries, c_i are country fixed effects and d_t are time fixed effects. The instrument used for each endogenous regressor x is the median value of the x variable of interest in neighbouring land and marine countries.

The first stage regressions are:

$$\ln(k_{it}) = \delta_0 + \delta_1 \ln(z_{k_{it}}) + \delta_2 \ln(z_{h_{it}}) + \delta_3 \ln(z_{p_{it}}) + \delta_4 \ln(z_{s_{it}}) + \beta_g (g_{it}) + c_i + d_t + \epsilon_{k_{it}}$$

$$\ln(h_{it}) = \alpha_0 + \alpha_1 \ln(z_{k_{it}}) + \alpha_2 \ln(z_{h_{it}}) + \alpha_3 \ln(z_{p_{it}}) + \alpha_4 \ln(z_{s_{it}}) + \beta_g (g_{it}) + c_i + d_t + \epsilon_{h_{it}}$$

$$\ln(p_{it}) = \gamma_0 + \gamma_1 \ln(z_{k_{it}}) + \gamma_2 \ln(z_{h_{it}}) + \gamma_3 \ln(z_{p_{it}}) + \gamma_4 \ln(z_{s_{it}}) + \beta_g (g_{it}) + c_i + d_t + \epsilon_{p_{it}}$$

$$\ln(s_{it}) = \pi_0 + \pi_1 \ln(z_{k_{it}}) + \pi_2 \ln(z_{h_{it}}) + \pi_3 \ln(z_{p_{it}}) + \pi_4 \ln(z_{s_{it}}) + \beta_g (g_{it}) + c_i + d_t + \epsilon_{s_{it}}$$

We estimate all first stage equations, create predicted values which are inputted into the second stage and then use OLS estimation to find the IV estimate for each endogenous regressor β_x.

$$\ln(y_{it}) = \beta_k \widehat{\ln(k_{it})} + \beta_h \widehat{\ln(h_{it})} + \beta_p \widehat{\ln(p_{it})} + \beta_s \widehat{\ln(s_{it})} + \beta_g (g_{it}) + c_i + d_t + u_{it}$$

The natural logarithm of each regressor is used as different growth factors vary greatly in magnitude and so this helps reduce heteroskedasticity. Nevertheless, our data still exhibits some heteroskedasticity and autocorrelation. To account for this, clustered standard errors are used which are clustered by country.

B. Validity of Instruments

Our results can be interpreted causally if our instruments are valid. We instrument variables of a country with the average values of these variables in its neighbouring countries. This has the advantage of generating variable-specific instruments and can be applied to a wide range of explanatory variables. An instrument (denoted by z_i) is a valid instrument if it is both *relevant* and *exogenous*. An instrument is relevant if the instrument is correlated with the variable of interest, i.e., $Cov(z_i, x_i) \neq 0$. This is true in our model.

Table 1: Correlation Analysis between Endogenous Variables and Instruments

	Physical Capital	Human Capital	Physical Productivity	Social Productivity
Physical Capital Instrument	0.707			
Human Capital Instrument		0.842		
Physical Productivity Instrument			0.860	
Social Productivity Instrument				0.487

Table 1 shows the strong degree of correlations between endogenous regressors and their instruments so that z_i creates enough variation in $\hat{x_i}$ to allow us to estimate the effect whilst maintaining small standard errors (a low degree of relevance will result in large standard errors and thus statistical insignificance). Intuitively, this is expected as geographic proximity can lead to imitation in education, quality of institutions, trade openness and others.

An instrument is exogenous if the instrument is uncorrelated with the error term of the regression, i.e., $Cov(z_i, u_i) = 0$. The exogeneity assumption can be broken down into two sub-assumptions, the exclusion assumption: which states that the instrument does not directly affect the outcome, i.e., that z_i itself is excluded from u_i, and the randomised assumption: that the instrument is uncorrelated with all unobserved factors in u_i that affect the outcome. A country could be affected by 'spillovers' from its neighbours mostly, but not exclusively, through trade and finance. For example, geographical proximity to fast growing countries could encourage FDI and technological inflows as was the case with the East Asian "Tigers". Thus, spillovers are controlled for by using the average GDP per capita growth rate in neighbouring countries. Additionally, instead of basing instruments on simple averages, the median of variables in neighbouring countries is used. This is because the median neighbour is less likely to be a country's main trading partner. Finally, neighbouring countries typically share common fixed traits like geography and climate which are likely to affect growth which could invalidate our instruments if not accounted for. Fixed effects are included in the model to account for this. Country fixed effects capture time-invariant heterogeneities across countries like climates and cultures. These are time-invariant as over the course of a relatively short panel dataset of 13 years these variables will not change much (for example the impact of climate change will be small). Time fixed effects capture differences in output that vary across time periods but not across individual countries, for instance due to global macroeconomic conditions like the impact of the Covid-19 pandemic. By controlling for spillover effects, common traits and major economic growth determinants we should satisfy the exclusion restriction.

V. Results

This paper categorises countries into three levels of development based on their average ranking in the Economic Complexity Index (ECI) which ranks countries based on how diversified and complex their export basket is. The ECI is a useful measure of development as it more accurately captures the productive potential and advancement of an economy which is the aim of long run economic growth models. Table 2 shows the regression results for each group of countries.

Table 2: IV Fixed Effects Regression Output Table

	Most Developed Countries	**Developed Countries**	**Least Developed Countries**
Log Physical Capital	0.647	-0.009	0.620***
	(1.311)	(1.429)	(0.163)
Log Social Productivity	-3.606	2.439	0.454
	(26.643)	(7.531)	(1.457)
Log Human Capital	8.689	-8.379	-1.540
	(23.544)	(20.969)	(3.026)
Log Physical Productivity	-1.521	2.015	0.294
	(2.819)	(3.956)	(2.792)
Neighbouring Countries GDP per Capita Growth Rate	-0.001	0.002	-0.002
	(0.020)	(0.004)	(0.003)
DF Resid.	211.000	487.000	163.000
nobs	247	546	195

***$p < 0.01$; **$p < 0.05$; *$p < 0.1$

Almost all economic growth determinants across all levels of development are statistically insignificant at the 1%, 5% or 10% significance level. Only physical capital in the least developed countries is statistically significant. At the 1% significance level, we can conclude that a 10% increase in physical capital is associated with (or *causes* if IV assumptions hold), on average, a 6.09% increase in GDP per capita. Whilst infrastructure projects in developing countries have often been rightly criticized whenever they serve little use ('white elephant projects') these results lend support to the view that physical capital investments, particularly when done so correctly, is one of the best means of development in developing countries. Although our economic growth determinants are not statistically significant in the most part individually, they are likely important when collectively combined. Cherif et al. (2018) found similar statistical insignificance of neoclassical growth determinants using an IV FE regression model. However, they show that these growth factors may be important to the extent that they help improve export sophistication which was by far their main determinant of economic growth.

VI. Conclusion

Overall, only physical capital is a statistically significant economic growth determinant and only so in developing countries. For all other growth determinants or when in more developed countries, our data shows that such factors are statistically insignificant, contrary to what neoclassical economic growth theory hypothesizes. One could argue that the instruments used violate exogeneity. However, this paper argues that sufficient controls and amending of weighting procedures used ensure exogeneity holds in our model. Therefore, the main reason why these growth factors are mainly insignificant by themselves is because they are complementary growth engines and perhaps they help promote growth via other channels like the improvement of export sophistication.

References

Barossi-Filho, M., Silva, R. and Diniz, E. (2005). “The Empirics of the Solow Growth Model: Long-Term Evidence”. Journal of Applied Economics, Vol. VIII, No. 1 (May 2005), 31-51.

Bazzi, S. and Clemens, M. (2013). “Blunt Instruments: Avoiding Common Pitfalls in Identifying the Causes of Economic Growth”. American Economic Journal: Macroeconomics 5(2): 152-86.

Cherif, R., Hasanov, F. and Wang, L. (2018). “Sharp Instrument: A Stab at Identifying the Causes of Economic Growth”. IMF Working Paper.

Durlauf, S. (2009). “The Rise and Fall of Cross-Country Growth Regressions”. Duke University Press.

Fisher, A. (1939). "Production, primary, secondary and tertiary". Economic Record.

Hausmann, R., Hwang, J. and Rodrik, D. (2007). “What You Export Matters”. Journal of Economic Growth 12(1): 1-25.

Johansson, E. and Tretow, E. (2015). “Modeling Economic Output and Economic Growth with respect to Economic Freedom”.

Kraay, A. (2018). “Methodology for a World Bank Human Capital Index”. Policy Research Working Papers. The World Bank.

Lucas, R. (2000). "Some Macroeconomics for the 21st Century". Journal of Economic Perspectives, American Economic Association, vol. 14(1), pages 159-168.

Romer, D., Mankiw, G. and Weil, D. (1992). "A Contribution to the Empirics of Economic Growth". The Quarterly Journal of Economics, Oxford University Press, vol. 107(2), pages 407-437.

Solow, R. (1956). “A Contribution to the Theory of Economic Growth”. The Quarterly Journal of Economics, Vol. 70, No. 1 (Feb., 1956), pp. 65-94. MIT Press.

Spolaore, E. and Wacziarg, R. (2005). “Borders and Growth”. Journal of Economic Growth 10(4): 331-386

Appendix I: Theoretical TFP estimate compared to combined TFP value from IDI and EFI

The theoretically implied TFP level in an economy can be derived via growth accounting and compared against the TFP implied value from combining the proxy datasets for physical productivity (IDI) and social productivity (EFI) to assess whether these datasets are good proxies for overall TFP. Using a neoclassical Cobb-Douglas production function, output per capita (y) can be written as a function of TFP (A), physical capital per capita (k) and human capital per capita (h): $y = Ak^{\alpha}h^{1-\alpha}$ where α is the capital share of income and in our instance A is a function of physical and social productivity: $A = f(z_p, z_s) = z_p \times z_s$.

Taking natural logarithms:

$$ln(y) = ln(z_p) + ln(z_s) + \alpha ln(k) + (1-\alpha)ln(h)$$

In period t: $ln(y_t) = ln\left(z_{p_t}\right) + ln\left(z_{s_t}\right) + \alpha ln(k_t) + (1-\alpha)ln(h_t)$

In period t+1: $ln(y_{t+1}) = ln\left(z_{p_{t+1}}\right) + ln\left(z_{s_{t+1}}\right) + \alpha ln(k_{t+1}) + (1-\alpha)ln(h_{t+1})$

Difference between periods:

$$\begin{aligned} &ln(y_{t+1}) - ln(y_t) \\ &\quad = ln\left(z_{p_{t+1}}\right) - ln\left(z_{p_t}\right) + ln\left(z_{s_{t+1}}\right) - ln\left(z_{s_t}\right) + \alpha ln(k_{t+1}) - \alpha ln(k_t) \\ &\quad + (1-\alpha)ln(h_{t+1}) - (1-\alpha)ln(h_t) \end{aligned}$$

Rearrange:

$$\begin{aligned} &ln\left(z_{p_{t+1}}\right) - ln\left(z_{p_t}\right) + ln\left(z_{s_{t+1}}\right) - ln\left(z_{s_t}\right) \\ &\quad = ln(y_{t+1}) - ln(y_t) + \alpha\left(ln(k_t) - ln(k_{t+1})\right) + (1-\alpha)(ln(h_t) - ln(h_{t+1})) \end{aligned}$$

While the left-hand side is realistically unobservable, the right-hand side is easily observable via national accounts and human capital statistics. Therefore, we can compare the theorized TFP values with the actual TFP values calculated from our proxy data sets (IDI and EFI) to judge the extent to which they are a good estimate for TFP. Overall, across all countries, the proxy TFP dataset overestimates the theoretically implied TFP value by only 2.6%. Thus, using the IDI and EFI datasets together seem to be a fairly good proxy of TFP.

Appendix II: List of countries in each group

Most Developed Countries	Developed Countries	Least Developed Countries
Mean ECI ranking ≤ 22	*22 < Mean ECI ranking ≤ 85*	*Mean ECI ranking > 85*
Austria	Albania	Algeria
Belgium	Argentina	Azerbaijan
Czech Republic	Australia	Botswana
Denmark	Bahrain	Burkina Faso
Finland	Brazil	Cameroon
France	Bulgaria	Ecuador
Germany	Canada	Kazakhstan
Hungary	Chile	Madagascar
Ireland	Colombia	Morocco
Italy	Costa Rica	Oman
Japan	Croatia	Paraguay
Luxembourg	Cyprus	Peru
Mexico	Estonia	Senegal
Singapore	Georgia	Uganda
Slovenia	Greece	Zimbabwe
Sweden	Iceland	
Switzerland	Indonesia	
United Kingdom	Israel	
United States	Jordan	
	Latvia	
	Lithuania	
	Malaysia	
	Malta	
	Mauritius	
	Moldova	
	Namibia	
	Netherlands	
	New Zealand	
	Norway	
	Panama	
	Poland	
	Portugal	
	Romania	
	Saudi Arabia	
	South Africa	
	Spain	
	Thailand	
	Tunisia	
	Turkey	
	Ukraine	
	United Arab Emirates	
	Uruguay	