Impact of Infrastructure Investment on Developed and Developing Economies

By Xuehui Han, Jiaqi Su and Jang Ping Thia*

June 2020

Abstract

This paper uses two longitudinal datasets—one with more limited coverage from the Organisation for Economic Co-operation and Development and another constructed using general government gross fixed capital formation—to test for the relative effects of infrastructure versus non-infrastructure investment on output per worker, between developed and developing economies. The paper presents evidence that increasing infrastructure per worker has a larger relative impact on developing economies. This also implies that the share of gross capital formation devoted to infrastructure should be higher in developing economies.

Keywords: Gross fixed capital formation, infrastructure, output per worker, growth

*Email: jangping.thia@aiib.org, xuehui.han@aiib.org, jiaqi.su@aiib.org.

1 Introduction

Infrastructure, always known to be an important factor in economic development, has attracted much renewed attention from policy-makers in recent years [(International Monetary Fund, 2014) (Asian Development Bank, 2017) (Fay & Rozenburg, Beyond the Gap, 2019) (Fay, Lee, Mastruzzi, Han, & Cho, 2019)]. This is the result of a confluence of a few important agenda, such as the need for sustainable infrastructure to meet climate change challenges, China's impressive infrastructure development and its subsequent efforts to internationalize this model through the promotion of the Belt and Road Initiative (BRI), and Sustainable Development Goals (SDGs) for which infrastructure development is a key part.

The effect of poor infrastructure is pernicious, impeding economies in many ways – reducing access to markets and opportunities, raising costs of amenities, increasing risks and uncertainty for businesses and people. Stagnating investments also coincide with slower productivity growth, including in developing economies (World Bank, 2020). The "fading promise of convergence" will mean that it would be harder for the long tail of countries developing economies to raise incomes towards middle-income levels.

China initiated the Asian Infrastructure Investment Bank (AIIB), which began operations in 2016, and as of 2019, has attracted 100 members (including many non-Asian developing countries). It has approved more than USD10 billion of loans and other types of investments. The BRIC (Brazil, Russia, India, China) group launched the New Development Bank (NDB), also in 2016, to promote more infrastructure investments in this group. The G20 set up the Global Infrastructure Hub (GIH) in 2014 as a coordinating center to facilitate more investments into infrastructure for both developed and developing economies. Added to these efforts are also numerous regional and national initiatives to this effect.

At the heart of this is the belief that infrastructure is critical for economic development, and that developing economies would need to invest significantly to address infrastructure needs, reduce infrastructure gap, in order to develop their economies.

1.1 Investment in East Asia and High Economic Growth

There are examples in recent history that support this view. Several Asian economies sustained a high level of infrastructure spending during their respective phases of fast economic development (Figure 1). Infrastructure investment as a share of GDP exceeded 10 percent in some years, and in some cases well above the global average. China remains exceptional in sustaining infrastructure investment of above 15 percent of GDP since the mid-1990s.[1] The scatter plot between GDP growth and infrastructure spending also shows a positive correlation (Figure 2).

[1] Even if there is overestimation of China's expenditure on infrastructure, and with some discounting from the statistics, it would still constitute a very exceptional rate of investment in infrastructure.

Figure 1: Infrastructure Investment as a % of GDP
in Select Asian Economies (1960-2017) Compared to World Average

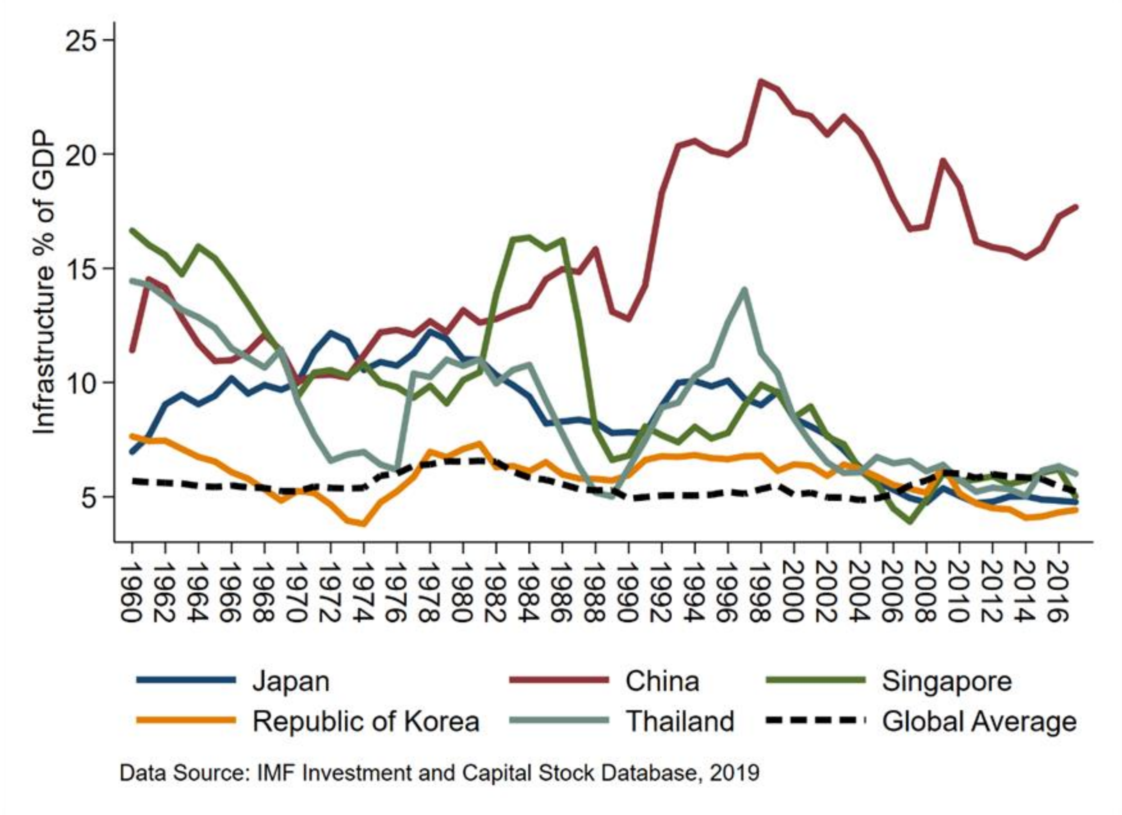
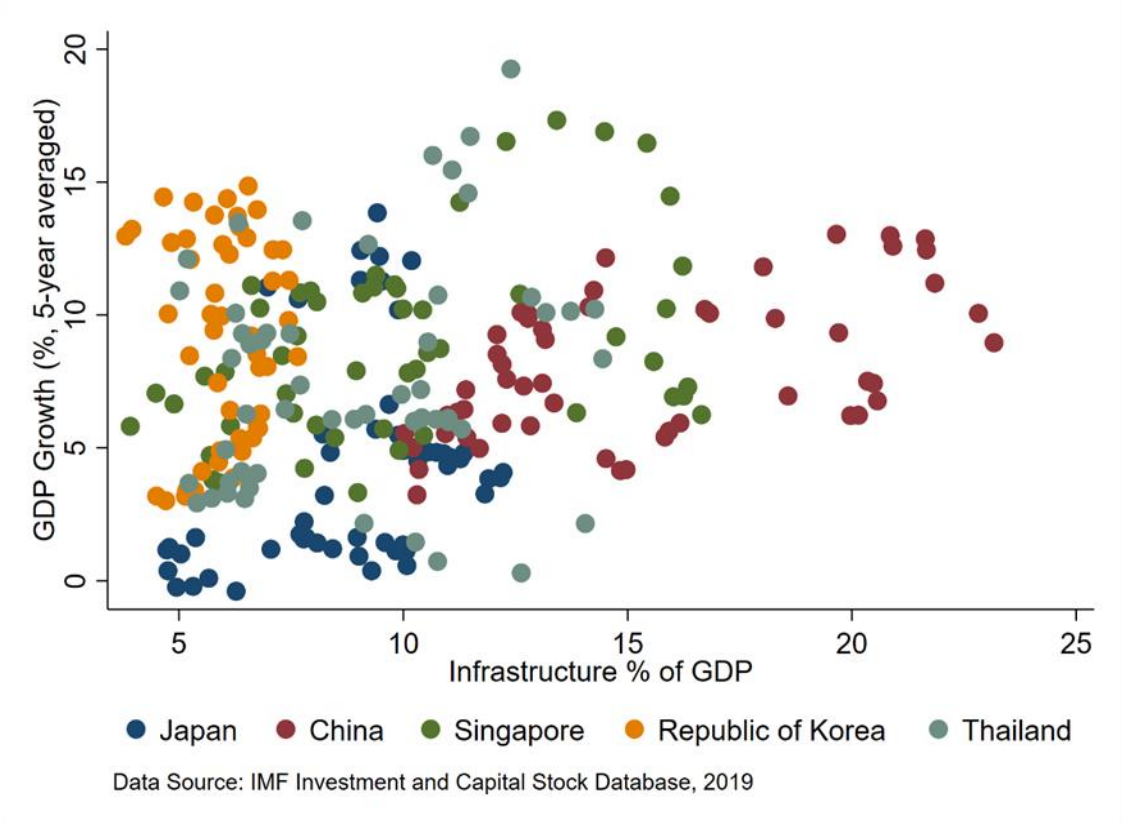


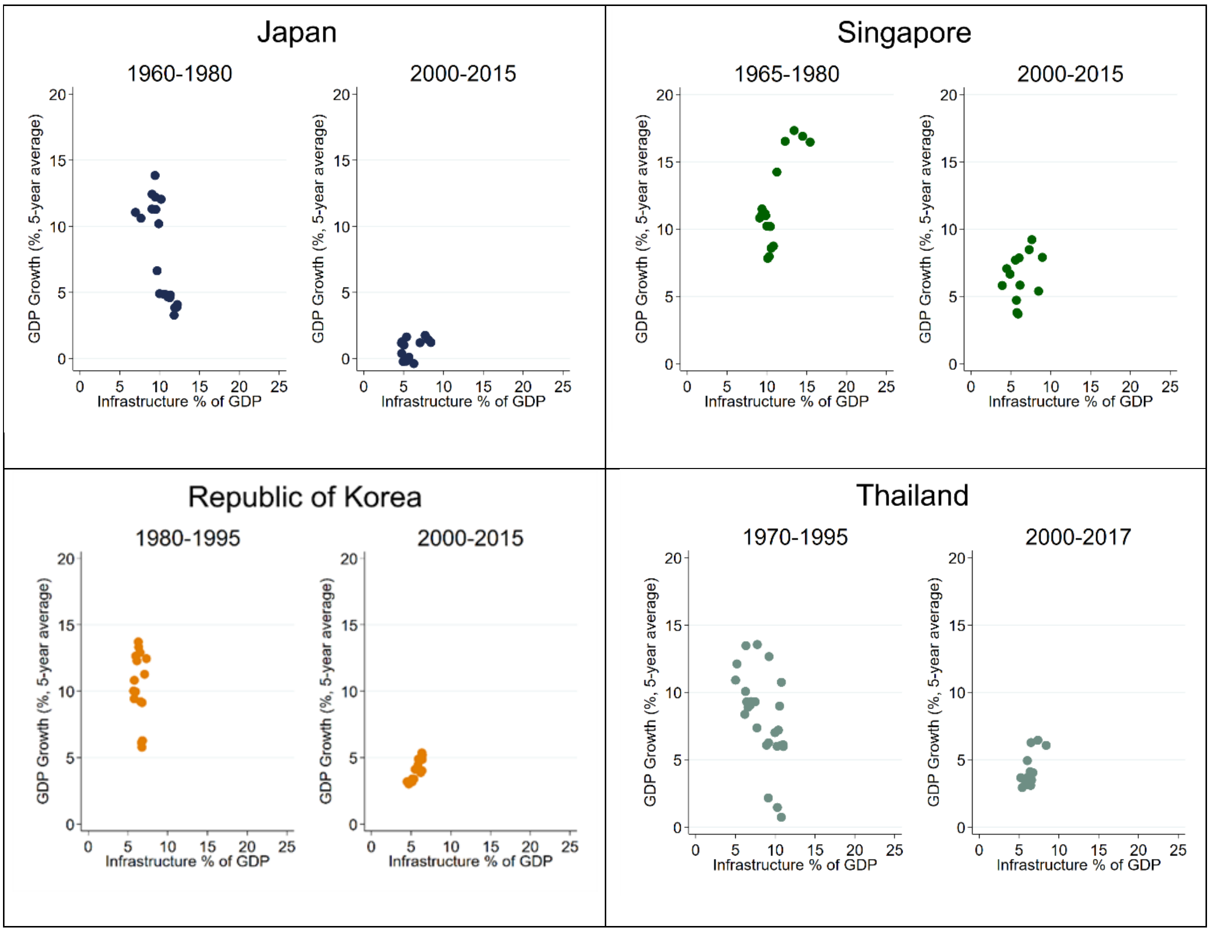
Figure 2: Scatter Plot of GDP Growth and Infrastructure Investment as % of GDP



For each of the select Asian economies, we further divide the scatter plots into high growth and slower growth periods to bring out the stylized facts (Figure 3). For Japan, Singapore, Korea and Thailand, it can be seen that the earlier high growth periods (left panels) were also accompanied by higher shares of GDP devoted to infrastructure. Infrastructure investment then slowed down and became smaller shares of GDP as economies reached maturity with lower growth rates (right panels).

These stylized facts thus point to high infrastructure investment correlating with high economic growth. This positive correlation, at least for these Asian economies, reinforces the view that infrastructure investment is critical to economic development.

Figure 3: Comparing High versus Low Growth Periods
and Infrastructure Investment as % of GDP for Select Asian Economies



Data Source: IMF Investment and Capital Stock Database

The central question we pose in this paper is whether developing economies should invest more in infrastructure. This goes beyond just relying on country-specific stylized facts (which are nonetheless important) so that one can arrive at more general conclusions.

2 Literature Review

What is the empirical evidence that developing economies should invest more in infrastructure? More precisely, we ask if there is empirical support that developing economies should invest a larger share of their GDP in infrastructure?

An intuitive answer would be affirmative. After all, much has been said about the low quality of infrastructure and the infrastructure gap in many developing economies. Yet a closer examination of the question suggests that the answer is not immediately obvious.

Take a classical growth theory in the form of the Solow growth model. The optimal rate of capital accumulation—the so-called "Golden Rule"—gives no indication that developing economies should invest a higher share of GDP on infrastructure.[2] The lower wages in developing economies relative to capital costs may even suggest less capital should be invested (ostensibly, this also applies to the infrastructure where the capital cost is high). There is even a strand of literature that suggests that returns on capital in developing economies are not high given the lack of human capital or poor institutional quality [(Lucas, 1990); (Alfaro, Kalemli-Ozcan, & Volosovych, 2008)].

There has been substantial research into infrastructure and a consensus on its positive impact. At the most macro level, it is accepted that infrastructure raises economies' supply-side capacity [(Aschauer, 1990) (Calderon & Serven, 2004) (Allcott, Collard-Wexler, & D O'Connell, 2016)] and can also be part of any counter-cyclical policy measures [(International Monetary Fund, 2014)].

At a more granular level, researchers have found that transport infrastructure contributed to improving market access, trade, poverty reduction and welfare [(Nuno & Venables, 2001) (Michaels, 2008) (Donaldson & Hornbeck, 2016) (Normaz Wana & Jamilah Mohd, 2015) (Aggarwal, 2018)], often leading to agglomeration [(Martin & Rogers, 1995) (Kline & Moretti, 2014)], and accompanied by changing comparative advantage and trade patterns [(Yeaple & Golub, 2007)]. Infrastructure generates greater competition [(Aghion & Schankerman, 1999)] and also raises firms' productivity [(Allcott, Collard-Wexler, & D O'Connell, 2016)]. Infrastructure investment is also found to generate further investments at the neighborhood level, thereby improving social outcomes [(McIntosh, Algeria, Ordonez, & Zenteno, 2018)]. The effects of infrastructure are generally positive, often quite large, and also varied at different levels of economic activity.

Yet, many of these do have one or more of the following features. Firstly, many studies are commendably well designed – including the use of natural experiments to account for confounding factors – they often only deal with the local effects of infrastructure, which may be the result of the transfer of activity from one region to another rather than gains in the overall economy. Secondly, some of the studies are limited to only a single or a small number of regional economies. Thirdly, there is sometimes a lack of distinction between public infrastructure expenditure between developed and developing economies. Given the different stages of development, public infrastructure can have significantly different impacts between the two groups.

An exception to the generally positive findings on infrastructure investment is (Devarajan, Swaroop, & Zou, 1996). The authors find that after controlling for total government expenditure, public investments in capital goods has a negative and significant coefficient on

[2] The Golden Rule merely states that countries should choose an optimal saving rate and accumulate capital to the point where the marginal return to capital is the same as offsetting factors, namely the depreciation rate of capital plus population growth.

per capita GDP growth rate for developing economies. This will be discussed in greater details in the following sections.

2.1 Contributions of This Study

This paper complements the existing literature by address the key question—whether elasticity of growth to infrastructure investment in developing economies is higher, based on an available set of data on infrastructure and non-infrastructure spending.

As a preface to the remainder of the paper, this research affirms the significantly positive impact of infrastructure. The paper also brings two key contributions to the literature. First, we put together a longitudinal dataset that contains consistent (though imperfect) measures of infrastructure investment together with other macroeconomic variables, and this allows for the relative impact of infrastructure versus non-infrastructure gross fixed capital formation (GFCF) to be estimated.

There has not been a standardized approach to measuring infrastructure investment, but there is an emerging consensus that the best approximate would be through the general government GFCF data provided by IMF, augmented by some private sector investment data from World Bank's Private Participation in Infrastructure (PPI) dataset [(Asian Development Bank, 2017) (Fay, Lee, Mastruzzi, Han, & Cho, 2019)]. Thus far, only snapshots of the data are provided. In this research, the longitudinal dataset is constructed with the methodology that is in line with the consensus. Note that in separate regressions, we also make use of an OECD dataset which contains more direct reporting of infrastructure spending but for a smaller set of economies. The characteristics of both datasets will be elaborated later in the paper, and the regression results from both datasets will be provided for robustness checks.

The second key contribution of this paper is to document the differing impacts (or elasticities) infrastructure has on developed versus developing economies. Our paper shows that for developing economies, infrastructure GFCF investment (per worker) has a stronger relative effect on output (per worker) compared to other forms of GFCF. To our best knowledge, this is the first time such effects are documented. This result is simple and yet important in what it implies. It suggests that developing economies need to direct a larger share of GFCF toward infrastructure (and by extension, have a higher infrastructure spending to GDP ratio in general) compared to developed economies. While acknowledging country heterogeneity, we provide a ballpark estimate on how many developing economies should invest in the infrastructure.

In terms of approach, this paper is similar to (Calderon & Serven, 2004) in using longitudinal datasets to reach a conclusion—except that this paper uses aggregate GFCF spending (of two separate datasets) rather than measured physical quantities of infrastructure types and quality of infrastructure. This paper also draws on (Esfahani & Ramirez, 2003) in using departure from steady states as a way to pin down the structure of the regression to be estimated.[3] In our estimation, we also arrived at broadly reasonable estimates of Total Factor Productivity (TFP) growth for developed and developing economies, giving some added confidence on the robustness of the result.

The rest of the paper is organized as follows. Section 3 will provide a description of the two datasets used for the empirical work, and in particular, how we construct the infrastructure GFCF data in the second dataset. We will also provide the general trends and summary

[3] Different from (Esfahani & Ramirez, 2003), this paper does not seek to model the influences on infrastructure capital accumulation. As will be explained later, we thus avoid directly having to model savings rate or the institutional factors that affect infrastructure accumulation. Secondly, we draw on much larger longitudinal datasets in arriving at our conclusions. This also allows us to partition the data into developing versus developed economies and highlight the key differences.

statistics of GFCF investment in this section. Section 4 will provide the estimation framework. The regression results are also presented in this section. Section 5 will provide a discussion of the results and attending policy implications. Section 6 concludes.

3 Data and Trends

In this paper, there are two sets of data to measure infrastructure - OECD Data and General Government GFCF Dataset, and one set of data to measure growth - output per worker.

3.1 OECD Data

The first data source for this piece of research comes from the Organisation for Economic Co-operation and Development [(OECD, Investment (GFCF) (indicator), 2019). (OECD, Investment by asset (indicator), 2019)]. This dataset covers 144 economies and 58 years, from 1960 to 2017. However, the dataset is unbalanced, especially with the early years when there are very few countries with data reported.

The key advantage is that OECD provides data on GFCF investments by asset types. GFCF asset types include six groups: dwellings (excluding land); other buildings and structures (roads, bridges, airfields, dams, etc.); transport equipment (ships, trains, aircraft, etc.); other machinery and equipment (ICT equipment, office machinery and hardware, as well as weapons systems etc.); cultivated assets (managed forests, livestock raised for milk production, etc.) and intellectual property products (intangible fixed assets such as R&D, mineral exploration, software and databases, and literary and artistic originals, etc.).

Here, “other buildings are structures” is taken as the paper’s measure of infrastructure GFCF, which is a subset of an economy’s total GFCF. OECD countries provide their data according to the system of national accounts (United Nations., 2008).[4] The definition of other buildings includes public non-residential properties such as warehouses, industrial and commercial buildings, hospitals, schools, etc., while other structures include flood protection, highways and roads, harbors, pipelines, communication, and power lines. In general, based on data definition, the coverage is larger than infrastructure per se (that is, potentially resulting in upward bias for infrastructure investment) but this is mitigated by the fact that there would also likely be some omissions as well (e.g., investment in ICT equipment, or transport equipment such as rolling stocks, which provide infrastructure like merit goods, are not included). Note that housing, typically the largest building asset class, is being excluded. Hence, this piece of data is taken as a good enough indicator for infrastructure spending.[5]

For the analysis, the paper separates countries into developing and developed economies for some of the regressions in order to highlight the difference. Developed economies are countries with a 10-year average (2008-2017) GDP per capita of above USD25,000. The rest are classified as developing economies.[6]

We first report OECD data on total GFCF and infrastructure GFCF as a percentage of economies for developed and developing economies. The mean percentages are provided from 1970 onward. But as countries exhibit heterogeneity, the P25 and P75 ratios are also

[4] “Dwellings” includes “Other buildings and structures” for Chile.

[5] Note that for regression purposes, it would matter less if any bias in reporting is steady over time—resulting in no impact in coefficients. If data is recorded with idiosyncratic errors, the subsequent regressions could be biased downward, but can be taken as the lower bound estimate of impact of such infrastructure.

[6] Developed economies in the OECD dataset are Australia, Austria, Belgium, Canada, Switzerland, Germany, Denmark, Spain, Finland, France, United Kingdom, Iceland, Ireland, Israel, Italy, Japan, Luxembourg, Netherlands, Norway, New Zealand, Sweden and the United States. Developing economies in the OECD dataset are Brazil, Chile, Colombia, Costa Rica, Czech Republic, Estonia, Greece, Hungary, Korea, Lithuania, Latvia, Mexico, Poland, Portugal, Slovak Republic, Slovenia and South Africa.

provided. From Figure 4, it can be seen that developed economies, on average, invest around 21-26 percent of GDP over the sample period (with slightly lower rate in recent years). This is not too dissimilar for developing economies with 17-26 percent, but with developing economies exhibiting greater heterogeneity as seen in Figure 5.

Figure 4: Total GFCF as a Percentage of GDP for Developed Economies (OECD Dataset)

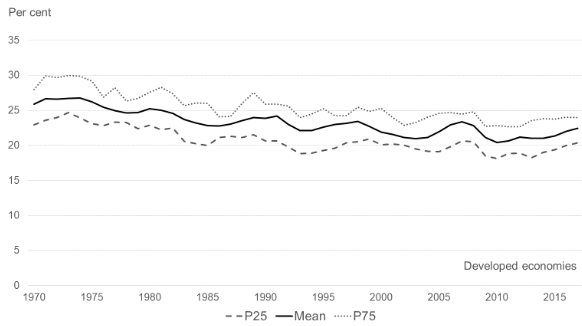


Figure 5: Total GFCF as a Percentage of GDP for Developing Economies (OECD Dataset)

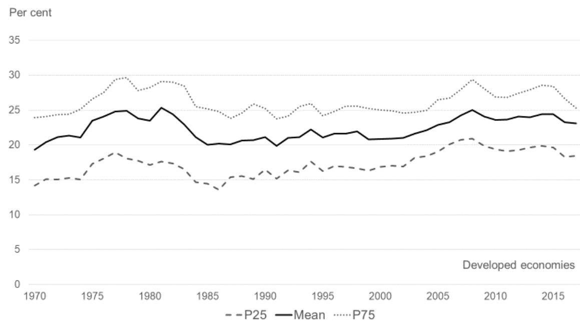


Figure 6: Infrastructure GFCF as a Percentage of GDP for Developed Economies (OECD Dataset)

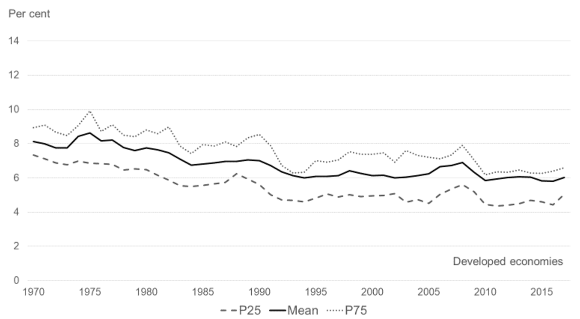
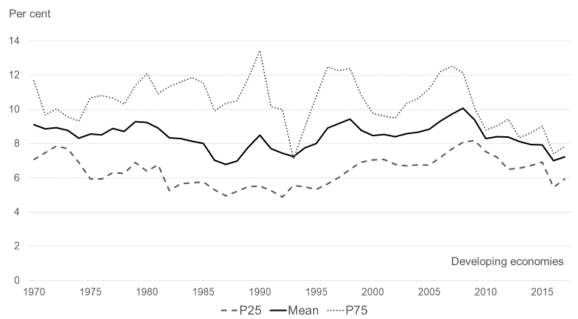


Figure 7: Infrastructure GFCF as a Percentage of GDP for Developing Economies (OECD Dataset)



Developed economies invest around 4.7-6.5 percent of GDP on infrastructure GFCF in the past 10 years based on P25 and P75 ratios, as seen in Figure 6. It is also clear that there is a small general downward trend from the 1970s, and this is broadly in line with the general downward trend in overall GFCF.

Developing economies invested around 7-9 percent of GDP in infrastructure in the past 10 years, as seen in Figure 7. A few other observations are also evident. First, infrastructure GFCF to GDP ratio is quite volatile, when compared to developing economies GFCF investment ratio. Second, developing countries exhibit more heterogeneity in investment ratios, as indicated by the wider band between P25 and P75 countries. Third, it is also clear that developing countries' infrastructure investment rate was higher before the financial crisis in the late 2000s, reaching an average of 10 percent of GDP. The decline in infrastructure GFCF ratio in the recent decade is quite across the board, resulting in a tighter but lower P25 to P75 band of infrastructure investment to GDP ratios.

3.2 General Government GFCF Dataset (or GFCF-GG Dataset)

A key disadvantage of the OECD longitudinal dataset is its relatively sparse coverage of developing economies. To overcome this and to test the robustness of findings, the research constructs a longitudinal dataset using IMF data and the increasingly accepted method of estimating infrastructure spending—which is based on GFCF of General Government [see

(Asian Development Bank, 2017) (Fay, Lee, Mastruzzi, Han, & Cho, 2019)]. The definition of infrastructure is

$$Infrastructure \ = GFCF_{GG} + PPP_{total}$$

Where $\mathrm{GFCF_{GG}}$ is the GFCF of general government (from IMF investment and capital stock dataset, 2019) and $\mathrm{PPP_{total}}$ is the public-private partnership (PPP) investment (also from IMF investment and capital stock dataset, 2019).[7] In this dataset, both $\mathrm{GFCF_{GG}}$ and $\mathrm{PPP_{total}}$ are measured in constant 2011 international dollars and expressed as the percentage of GDP.[8]

$\mathrm{GFCF_{GG}}$ captures the gross fixed capital formation by central, state, and local governments. It measures public involvement in infrastructure. The other variable $\mathrm{PPP_{total}}$ tracks the total commitment of projects where there is public-private-partnership. (Fay, Lee, Mastruzzi, Han, & Cho, 2019) evaluated the existing estimation methods of infrastructure and highlighted three advantages—ready accessibility, wide country coverage, and long-time series. With the update of the Investment and Capital Stock Database in August 2019 extending the period of GFCF-GG data from 2015 to 2017, this balanced panel dataset covers 189 economies and 58 years from 1960 to 2017.

However, the caveat is that this definition is an approximation rather than a precise measurement of infrastructure. $\mathrm{GFCF_{GG}}$ is initially designed to measure public investment in the national account system. It includes a certain portion of non-infrastructure related investments, e.g., residential housing and social investments, such as health, education, national defense, and mining. On the flip side, it excludes State-Owned-Enterprise (SOE) activities, which could be the main undertaker of infrastructure investment in certain countries.

As for $\mathrm{PPP_{total}}$, the information comes from two main sources: The World Bank Private Participation in Infrastructure Database for low- and middle-income economies and the European PPP Expertise Center data at the European Investment Bank for high-income European economies.[9] Instead of estimating the exact percentage of private investment, $\mathrm{PPP_{total}}$ takes the total commitment value of each project into the calculation by spreading the value of project commitments over five years then add together the investment by each country and every year. The caveat of this method is that it excludes the fully privatized projects (error of omission) and that given its public-private nature, potentially lead to double counting when added to $\mathrm{GFCF_{GG}}$ (error of inclusion).

Therefore, there is a need to acknowledge the possibility of over or underestimation on infrastructure by choosing this definition. But the above-mentioned factors seem to suggest that some errors (upside or downside bias) would cancel out. This underpins the emerging consensus that it is a good proxy measure for infrastructure investment.[10] The trends for developed and developing economies using this measure are provided from Figure 8 to Figure 11.

[7] Information on public investment comes from three sources: the OECD Analytical Database (2019 version) for OECD countries, and a combination of the National Accounts of the Penn World Tables (PWT, version 9.1) and the IMF World Economic Outlook (WEO, April 2019 vintage) for non-OECD countries..
[8] In some of the regressions later, we use the shares reported, multiplied these to relevant variables reported in USD, and essentially converted GFCF or infrastructure figures into USD for regressions. This additional regressions in USD are meant to provide robustness checks.
[9] World Bank Private Participation in Infrastructure Database: https://ppi.worldbank.org/en/ppi [Last accessed Oct. 24, 2019]. European PPP Expertise Center: https://www.eib.org/epec/ [Last accessed Oct. 24, 2019].
[10] Improving the precision of infrastructure measurement is a nontrivial task and a subject for further research.

Figure 8: Total GFCF as a Percentage of GDP for Developed Economies (GFCF-GG Dataset)

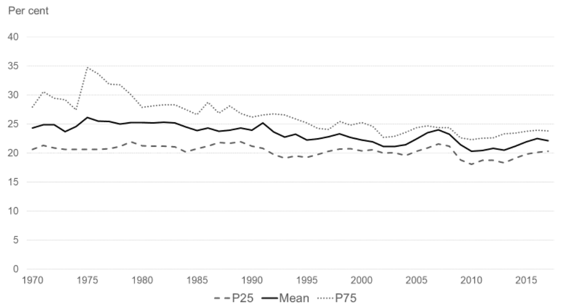


Figure 9: Total GFCF as a Percentage of GDP for Developing Economies (GFCF-GG Dataset)

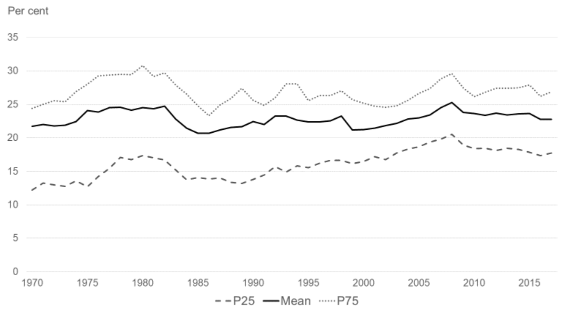


Figure 10: Infrastructure GFCF as a Percentage of GDP for Developed Economies (GFCF-GG Dataset)

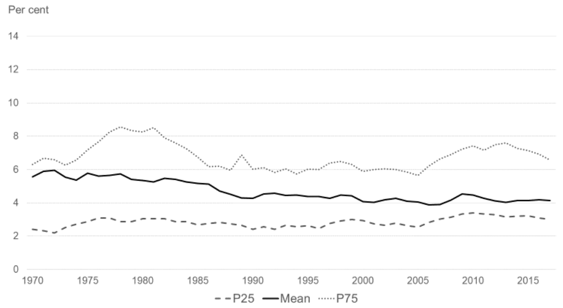
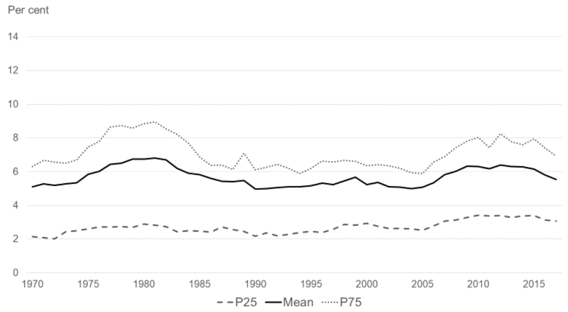


Figure 11: Infrastructure GFCF as a Percentage of GDP for Developing Economies (GFCF-GG Dataset)



From the GFCF-GG dataset, developed economies' GFCF is around 20-24 percent of GDP (based on P25 to P75 values), and infrastructure is 3.9-4.6 percent of GDP, over the sample period. Consistent with the OECD dataset, these show a slightly declining trend.[11] For developing economies, GFCF is 15-26 percent and infrastructure 2.6-7.0 percent of GDP, respectively. Investment in developing economies is more volatile, with greater heterogeneity.

3.2 Other variables and summary table

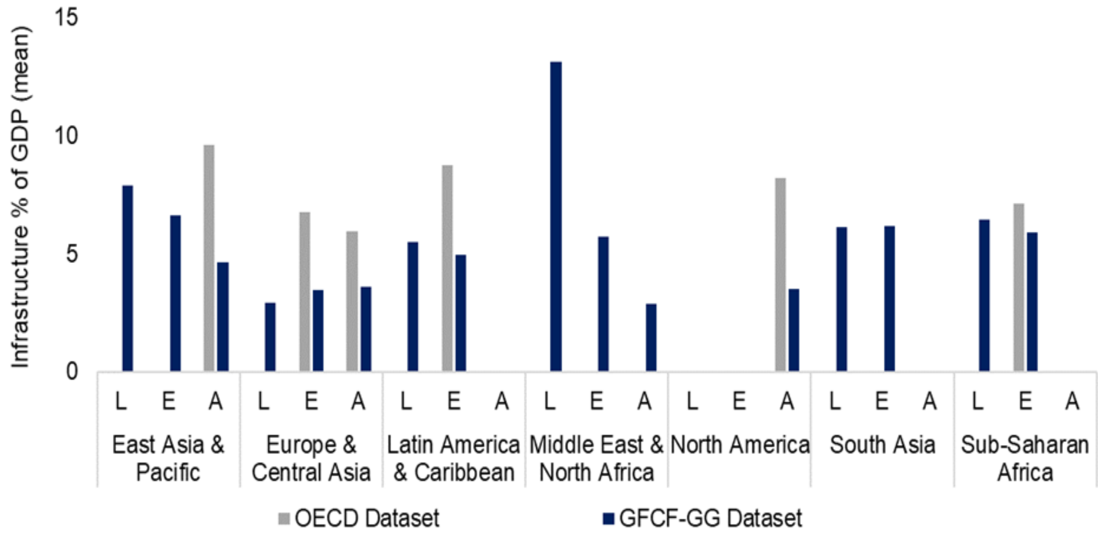
In addition to the infrastructure-related data, four more variables were added for the analysis in this paper, which are gross domestic production, population, employment rate and gross fixed capital formation. The definition and source of these variables are listed in Appendix A. With these variables, the paper is able to derive output per worker, infra GFCF per worker, non-infrastructure GFCF per worker, etc., which are then used for the regressions in the next section.

To show the difference between these two datasets, Figure 12 below breaks down the infrastructure investment rate in 2017 by income group and region. Within the OECD dataset, the average infrastructure investment rate in 2017 is 8 percent across regions, with advanced economies in East Asia being the highest and advanced economies in Europe and Central Asia scoring the lowest. As for the GFCF-GG dataset, it shows that advanced economies generally invested less than the low-income economies within the same region, which is to be expected. In addition, there was a large disparity of investment levels between income groups in the Middle East and North Africa region. The advanced economies in this region invested

[11] Infrastructure spending reported from OECD and GFCF-GG dataset are slightly different. The former is gathered through more detailed components of national accounts reported by OECD countries, while the latter is proxied mostly using general government GFCF.

the lowest – at 2.9% of GDP, while the low-income economies, invested the highest - at 13% of GDP. East Asia and Pacific region invest high comparatively across all the income groups.

Figure 12: Overview of the Infrastructure Investment Rate by Region and Income Group in 2017



Note: L= Low Income Economies; E = Emerging Economies; A = Advanced Economies
Data Source: OECD Database and IMF Investment and Capital Stock Database, 2019

4 Model Framework and Estimation

4.1 Regression Model

Following (Esfahani & Ramirez, 2003), the growth model is

$$Y = K^{\alpha} N^{\beta} (QL)^{1-\alpha-\beta}$$

Where Y is the F K and N are infrastructure and non-infrastructure capital stock respectively. L is labor, and Q is the TFP in a labor augmenting technology. In per capita terms, this becomes

Equation 1

$$y = k^{\alpha} n^{\beta} Q^{1-\alpha-\beta}$$

where y, k, n are expressed in per worker terms. Expressed as growth terms, the equation becomes

Equation 2

$$\gamma_y = \alpha\gamma_k + \beta\gamma_n + (1-\alpha-\beta)\gamma_Q$$

Where γ_y is the growth rate of output per worker (the same analogs hold for other variables with γ). In a balanced growth path, all endogenous variables grow at the same steady-state rate q^* which is underlying TFP growth rate

$$\bar{\gamma}_Q = \bar{\gamma}_y = \bar{\gamma}_k = \bar{\gamma}_n = q^*$$

This allows Equation *2* to be written as

Equation 3

$$\gamma_y = q^* + \alpha(\gamma_k - q^*) + \beta(\gamma_n - q^*) + (1 - \alpha - \beta)\left(\gamma_Q - q^*\right)$$

Observed economic growth consists of an underlying steady-state rate q^* and the components related to any shocks to infrastructure and non-infrastructure growth, and some TFP shocks $\gamma_Q - q^*$, which we treat as the error term. This is the key identification equation for subsequent regressions. For the subsequent regressions, we need effective measures for the regressors. By definition

$$\gamma_k \equiv \frac{\Delta K}{K} - \frac{\Delta L}{L}$$

Note that capital accumulation follows the standard process

$$\frac{\Delta K}{K} = \frac{GFCF_K}{K} - \delta$$

where $\Delta K = GFCF_K - \delta K$ is the net flow of investment into infrastructure, with $GFCF_K$ as new capital formation and δ accounting for the rate of depreciation of existing capital stock. This implies that

Equation 4

$$\gamma_k - q^* = \left(\frac{\Delta K}{K} - q^*\right) - \frac{\Delta L}{L} \approx \left(\frac{GFCF_K}{K} - q^*\right) - \frac{\Delta L}{L}$$

where the approximation ignores the effect of depreciation. In other words, the change in infrastructure stock can be proxied by the gross fixed capital formation in infrastructure.[12]

Extending this further, the paper uses the first log difference of GFCF per worker as the proxy for $\gamma_k - q^*$. The argument is as follows. Annual GFCF in the economy captures the increase in capital stock. This can also be easily converted into per worker terms, dividing by the number of workers. Nevertheless, annual GFCF investment per worker does not tell us whether the rate of investment is above or below steady-state, which is what is required from Equation 4.

On the other hand, a positive log first difference in GFCF (in per worker terms) will correlate to a positive shock to per worker capital stock growth away from steady-state. Conversely, a negative first log difference in GFCF will correlate with a negative shock to per worker capital stock growth.[13] The term $\gamma_k - q^*$ can thus be better measured by first log difference in per worker infrastructure GFCF. Similarly, $\gamma_N - q^*$ is measured by first log difference in per worker non-infrastructure GFCF. This completes the characterization of the regression setup and allows us to estimate Equation 3.

In all regressions, changes in output per worker (log difference) are the independent variable. In regression 1 (shorthand R1), this is regressed against infra GFCF investment per worker and non-infrastructure GFCF per worker in log difference, using random effects. In R2, fixed effects are used. R3 and R4 also employ fixed effect estimation but with different subsamples

[12] This is an approximation. Suppose the depreciation rate is small, the GFCF (which is a flow) in each year would be closely matched to the increase in capital stock. Most capital stock series are constructed using rolling annual GFCF figures (net increase after accounting for depreciation).

[13] Note that in the scenario of the negative first difference, capital per worker could still be rising, but it will be rising at a rate that is below steady state.

of developed and developing economies, respectively. Regression R5 has all samples, while R6 and R7 use subsamples of developed and developing countries, respectively.

From R5 to R7, instrumental variables are used for address potential endogeneity. Instruments are the second to fourth lags (L2 to L4) of explanatory variables – namely infra GFCF per worker and non-infra GFCF per worker (in log level terms). The use of lagged regressors as instruments provides consistent (though likely inefficient) estimates in panel regressions. The number of lagged instruments will not in principle affect consistency, so long as they provide sufficient information on the instrumented variable. In the research, a shorter lag (L2 to L3) is also tested and they provide broadly similar coefficients.[14]

4.2 Regression Results

The results using the OECD dataset are presented in Table 1, with the types of estimation methods, sample coverage, fully described. The respective regressions from R1 to R7 regressions are repeated on the GFCF-GG dataset (from regressions R8 to R14), and results are presented in Table 2.

Table 1: Regression Results of Output Per Worker on Infrastructure and Non-Infrastructure GFCF (OECD data in USD)[15]

	R1 Random effect	R2 Fixed effect	R3 Fixed effect	R4 Fixed effect	R5 Fixed effect	R6 Fixed effect	R7 Fixed effect
Constant	0.016 0.036 10.680	0.016 0.001 13.960	0.011 0.002 6.750	0.023 0.002 14.020	0.015 0.002 9.340	0.008 0.002 3.190	0.024 0.002 11.620
Infra GFCF per worker (log difference)	0.241 0.018 13.300	0.240 0.018 13.460	0.270 0.030 9.050	0.211 0.019 11.330	0.247 0.057 4.380	0.155 0.032 4.800	0.346 0.083 4.170
Non infra GFCF per worker (log difference)	0.389 0.036 10.680	0.385 0.037 10.510	0.399 0.055 7.290	0.375 0.049 7.700	0.446 0.081 5.490	0.600 0.078 7.690	0.301 0.086 3.490
Observations	918	918	542	376	801	476	325
Country groups	39	39	22	17	39	22	17
R-square within	0.731	0.731	0.761	0.708	0.738	0.729	0.708
R-square between	0.910	0.910	0.887	0.947	0.814	0.670	0.893
R-square overall	0.736	0.736	0.762	0.716	0.745	0.734	0.721
P	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Samples	All	All	Developed	Developing	All	Developed	Developing
Instruments	No	No	No	No	L2 to L4 infra GFCF and non-infra GFCF per worker	L2 to L4 infra GFCF and non-infra GFCF per worker	L2 to L4 infra GFCF and non-infra GFCF per worker

The standard errors and t-tests are provided on the second and third rows of corresponding coefficients.

[14] Can be made available on request.
[15] Note: Infrastructure is abbreviated as "infra" in the result table.

Table 1 displays the result using the OECD dataset, with 39 countries having sufficient data for the regression. Focusing on R6 and R7 (i.e., between developed and developing economies, accounting for endogeneity), the contribution of infra per worker is higher in the developing economies group (0.346) compared to developed economies (0.155).

Table 2: Regression Results of Output Per Worker on Infrastructure and Non-Infrastructure GFCF Using GFCF-GG Data (International Currency)

	R8	**R9**	**R10**	**R11**	**R12**	**R13**	**R14**
	Random Effect	Fixed Effect	Fixed Effect	Fixed Effect	Fixed Effect	Fixed Effect	Fixed Effect
Constant	0.019	0.018	0.011	0.020	0.018	0.009	0.020
	0.008	0.000	0.000	0.000	0.001	0.001	0.002
	5.760	47.330	59.650	43.880	13.790	17.460	12.540
Infra GFCF per worker (log difference)	0.049	0.047	0.032	0.047	0.113	0.138	0.112
	0.008	0.008	0.013	0.008	0.043	0.057	0.045
	5.860	5.740	2.540	5.580	2.620	2.440	2.460
Non infra GFCF per worker (log difference)	0.043	0.043	0.087	0.041	0.039	0.130	0.033
	0.008	0.007	0.009	0.008	0.017	0.024	0.018
	5.760	5.730	9.990	5.520	2.260	5.500	1.780
Observations	3753	3753	666	3107	3270	585	2685
Country groups	157	157	27	130	156	27	129
R-square within	0.146	0.146	0.115	0.152	.	.	.
R-square between	0.248	0.246	0.532	0.229	0.345	0.257	0.313
R-square overall	0.151	0.151	0.122	0.155	0.107	0.076	0.107
P	0.000	0.000	0.000	0.000	0.008	0.000	0.025
Samples	All	All	Developed	Developing	All	Developed	Developing
Instruments	No	No	No	No	L2 to L4 infra GFCF and non-infra GFCF per worker	L2 to L4 infra GFCF and non-infra GFCF per worker	L2 to L4 infra GFCF and non-infra GFCF per worker

The standard errors and t-tests are provided on the second and third rows of corresponding coefficients.

Regressions R15 and R16 use the same formulation as R13 and R14, respectively (and also R6 and R7), but are weighted by the employment size of each country. The reason is as follows. In the GFCF-GG dataset with 157 economies, 32 of which are with GDP smaller than USD10 billion, and 49 of which are with GDP smaller than USD20 billion. Given the smallness of their GDP, the recorded macro data can be more idiosyncratic.[16] Rather than an arbitrary cut off to exclude some small economies, regressions R13 and R14 correct this by giving

[16] Small economies have some idiosyncratic characteristics. For example, some recorded very large current account deficits—Maldives, Liberia, Mongolia, Nicaragua, Sao Tome and Principe, Guinea-Bissau, Montenegro, Mozambique, Congo, had all registered current account deficits in excess of 40 percent of GDP. Furthermore, Mali, Zambia, Equatorial Guinea, Bhutan, Mongolia, Gabon had in some years recorded government gross capital formation in excess of 50 percent of GDP. Per capita income growth can also be erratic, with Liberia, Equatorial Guinea, Azerbaijan, Maldives and others registering more than 20-percent growth in output per worker (even in PPP terms) in certain years.

larger economies (regardless of income levels) more weight. Finally, R17 and R18 repeat R15 and R16 but with per capita worker output and GFCF recorded in current USD (instead of international dollars). The results are presented in Table 3. Since these regressions (R15 to R20) are scaled up by employment in each country, it would be inaccurate to over-interpret the standard errors and t-statistics.

4.3 Low-Middle Income Subset

For the regression for developing economies, there is a legitimate concern on whether the results hold for a subset of developing economies with low to middle income. While the development needs of lower-income countries are large, it is also highly plausible that institutional factors in these countries would constrain the effectiveness of infrastructure (though this would most likely constrain the effectiveness of other forms of capital as well).

Nevertheless, for robustness checks, we tested versions of regressions R16 (in international dollars) and R18 (in USD) but reduced the sample points (arbitrarily) to those economies' 10-year average per capita GDP of between USD1,000 and USD5,000. We chose a lower cutoff of USD1,000 to exclude economies with a very low level of development, with quite a few that are affected by specific factors such as civil strife or natural disasters during the sample period.[17] The respective results are presented in the R19 and R20. Note that in R19, the infrastructure coefficient is positive, while non-infrastructure GFCF is not.

[17] Countries below the cut off of per capita GDP USD1,000 are Niger, Malawi, Madagascar, Central African Republic, Ethiopia, Mozambique, Sierra Leone, Togo, Gambia, Burkina Faso, Liberia, Guinea-Bissau, Uganda, Rwanda, Nepal, Guinea, Haiti, Mali, Benin, Tanzania, Tajikistan, Chad. There are 49 countries between USD1,000 and USD5,000, including those with sizable populations such as India, Bangladesh, Indonesia, Pakistan, Nigeria, Vietnam, Philippines, Egypt, etc., giving a fairly representative sample (note that China is above USD5,000 and not included in this subset).

Table 3: Employment Weighted Regression Results of Output Per Worker on Infrastructure and Non-Infrastructure GFCF Using GFCF-GG Data (International Currency and USD)

	R15	R16	R17	R18	R19	R20
Infra GFCF per worker (log difference)	0.039 0.000 192	0.043 0.000 495	0.301 0.000 1000	0.305 0.000 2466	0.040 0.000 556	0.208 0.000 1598
Non infra GFCF per worker (log difference)	0.165 0.000 905	0.064 0.000 445	0.570 0.000 1777	0.482 0.000 2001	-0.023 0.000 -143	0.357 0.000 1218
Observations Centered R-square P	11158147 0.115 0.000	73859840 0.219 0.000	11158147 0.753 0.000	73596134 0.491 0.000	31206081 -0.070 0.000	31051122 0.601 0.000
Samples	Developed	Developing	Developed	Developing	Economies per capita between USD 1000 and 5000	Economies per capita between USD 1000 and 5000
Country groups	27	128	27	128	49	49
Recorded in	International	International	USD	USD	International	USD
Regression	Fixed Effect	Fixed Effect	Fixed Effect	Fixed Effect	Fixed Effect	Fixed Effect
Instruments	L2 to L4 infra GFCF and non-infra GFCF per worker	L2 to L4 infra GFCF and non-infra GFCF per worker	L2 to L4 infra GFCF and non-infra GFCF per worker	L2 to L4 infra GFCF and non-infra GFCF per worker	L2 to L4 infra GFCF and non-infra GFCF per worker	L2 to L4 infra GFCF and non-infra GFCF per worker

While the standard errors and t-tests are provided on the second and third rows of corresponding coefficients, the regressions in this table are employment weighted - standard errors and t-tests should be interpretated cautiously

5 Discussion

5.1 Summary of Key Results

We begin the discussion by highlighting a few results. First, the combined coefficients $\alpha + \beta$ (for infrastructure and non-infrastructure) are larger for developed economies. This can be seen by comparing R6 and R7 of OECD data, R13 and R14 of GFCF-GG data, as well as R15 and R16 (and also R17 and R18) of the weighted regressions. From R15 and R16, the total coefficients are around 0.20 and 0.11 for developed and developing economies, respectively (or 0.27 and 0.15 if the unweighted regressions of R13 and R14 are used as comparisons). Capital it seems is indeed more and almost twice as productive in developed economies.[18]

Second, because the regressions are expressed in growth terms, there is a natural interpretation of the constant term—which is the Total Factor Productivity (TFP) growth—as seen in Equation 3. From R13 and R14, TFP growth, on average, is 0.9 percent per annum for developed economies and 2.0 percent per annum for developing economies. The OECD regressions (R6 and R7) show 0.8 and 2.4 percent per annum respectively. The paper

[18] As an aside, we also notice that regressions where output and GFCF are recorded in USD tend to give very large coefficients (see R17 and R18), and also OECD regressions (R6 and R7). We attribute this to the effect of currency affecting both LHS variable (output per worker) and also RHS variable (GFCF) thereby biasing the elasticities upward. To be clear, the upward bias should affect both infrastructure and non-infrastructure coefficients.

believes that these are reasonable estimates that provide some confidence that the regressions are robust.

Third, which is the key and new insight presented in this paper, is that the infrastructure coefficient is relatively larger in developing economies. Comparing R15 and R16, infrastructure is 19 percent of the total coefficient for developed economies but a higher 40 percent for developing economies. For R13 and R14, infrastructure is around 52 percent of total coefficient for developed economies, but a higher 77 percent of total coefficient for developing economies.[19] In other words, across a range of comparable regressions, infrastructure growth in developing economies has a higher relative impact on per capita GDP growth compared to non-infrastructure capital.

5.2 Implications on Infrastructure and Non-Infrastructure GFCF Investment

Taking first-order conditions from Equation 1 and equating them to the cost of capital r_K and r_N, one arrives at the following standard equations:

$$\frac{dy}{dk} = \alpha k^{\alpha-1} n^{\beta} Q^{1-\alpha-\beta} = r_K$$

$$\frac{dy}{dn} = \beta k^{\alpha} n^{\beta-1} Q^{1-\alpha-\beta} = r_N$$

and the optimal capital ratios

$$\frac{K}{N} = \frac{r_N}{r_K}\frac{\alpha}{\beta}$$

If there aren't significant differences between the costs of capital for infrastructure and non-infrastructure or other distortions in the economy, the $\frac{K}{N}$ ratio in the economy should be the ratios of the elasticities $\frac{\alpha}{\beta}$ in the production function. Similarly, the $\frac{\alpha}{\alpha+\beta}$ ratio would represent the fraction of GFCF that should be invested in infrastructure, on average. The key statistics of various datasets and regression ratios are presented in Table 4.

[19] Similar pattern holds for regressions with variables recorded in USD. R6 and R7 (21 percent for developed, 53 percent for developing); R17 and R18 (35 percent for developed, 39 percent for developing). Note that even though regressions with variables in USD may result in upward bias for the coefficients (see footnote 18), the relative impact should not be affected.

Table 4: Key Statistics from Various Datasets and Regressions

	Developed			Developing		
Dataset	OECD	GFCF-GG	GFCF-GG	OECD	GFCF-GG	GFCF-GG
Regression	R6	R15	R17	R7	R16	R18
Recorded currency	USD	International	USD	USD	International	USD
α/(α+β)	0.205	0.193	0.346	0.535	0.402	0.387
GFCF to GDP (P75)	25.5	23.8	23.8	26.2	25.5	25.5
GFCF to GDP (mean)	23.0	22.2	22.2	22.3	21.7	21.7
GFCF to GDP (P25)	20.8	20.3	20.3	17.3	15.5	15.5
Regression projected infra to GDP (P75)	5.2	4.6	8.2	14.0	10.2	9.9
Regression projected infra to GDP (mean)	4.7	4.3	7.7	11.9	8.7	8.4
Regression projected infra to GDP (P25)	4.3	3.9	7.0	9.2	6.2	6.0
Actual infra to GDP (P75)	7.6	6.6	6.6	10.4	7.0	7.0
Actual infra to GDP (mean)	6.8	4.1	4.1	8.4	5.6	5.6
Actual infra to GDP (P25)	5.5	3.0	3.0	6.4	2.6	2.6

5.2.1 Ballpark Estimates

Our preferred benchmark regressions are R15 and R16. These are carried out using the more comprehensive GFCF-GG dataset, in international dollars (to avoid effects caused by currency fluctuations), and employment weighted to avoid the idiosyncrasies of smaller economies.

Based on a $\frac{\alpha}{\alpha+\beta}$ ratio of 0.193 and a P25-P75 range of GFCF being 20.3-23.8 percent of GDP, regression estimates suggest developed economies should invest 3.9-4.6 percent of GDP for infrastructure. Broadly speaking, developed economies appear to invest in line with what regression estimates suggest.

On the other hand, developing economies should invest around 6-10 percent of GDP on infrastructure, based on a $\frac{\alpha}{\alpha+\beta}$ ratio of 0.4 and a P25 – P75 GFCF of 15.5 to 25.5 percent of GDP. Actual infrastructure spending of 2.6 to 7.0 percent of GDP appears to fall below what is suggested via regression estimates.

It is important to note that the paper is not making a claim that developing economies should invest in more capital. One could argue that the total GFCF ratio in developing economies should be higher (but it is certainly not the case based on data). How much should an economy invest is a function of many other parameters, such as its intertemporal preferences, the real interest rate it faces, access to capital markets, and of course, various institutional quality affecting the productivity of capital.

As mentioned, there are natural skepticism around more capital investments by governments in developing economies. (Devarajan, Swaroop, & Zou, 1996) find that public sector expenditure on capital goods has a negative effect on per capita GDP growth in developing economies, after controlling for total public expenditure in the economy. The authors find that current government expenditure has a more beneficial effect on per capita GDP growth, and suggest that government misallocates resources to capital spending.

In this paper, the test is not on per capita GDP but on output per worker. Given that developing economies have higher population growth compared to developed economies, output per worker is a more accurate measure of economic convergence (as opposed to per capita GDP, which in developing economies will be slowed down by population growth). The research here also accounts for the relative difference between infrastructure and non-infrastructure GFCF, while acknowledging that total GFCF returns appear to be larger in developed economies. Here is also where this paper finds (Esfahani & Ramirez, 2003) framework attractive, with a proper functional form and highlighting growth effects of employment, infrastructure and non-infrastructure investments on output per worker growth.[20]

The claim here is that developing economies are not directing as much capital into infrastructure investments as relative growth elasticities suggest. The evidence here arises directly from the estimation of the growth Equation 3, which comes from the production function in Equation 1. This is a general framework, and the paper does not rely on strong assumptions or very specific growth models to generate the results. We have also shown a range of regressions, using two separate datasets, to provide greater confidence in the results.

In the end, the results tell a simple story—namely that the returns to infrastructure are relatively higher in developing economies, compared to other forms of capital investment. Infrastructure provides basic services to various economic and social activities. One can argue that infrastructure has to be in place so that other forms of capital (human or physical) can become productive. It is in this context that for developing economies, it becomes more important to divert more resources into infrastructure. On the other hand, in developed economies, returns to infrastructure are relatively smaller compared to other forms of capital.[21]

5.3 Population Growth

In this model, the variables are normalized to per worker units. The paper otherwise does not model demographic changes formally. Hence, we would state the following point briefly. Developing economies have higher population growth. Potentially, this implies a larger group of young population that has yet to enter the labor force (i.e., not yet productive) but would nonetheless require infrastructure services. This is another argument why developing countries need to invest more in infrastructure (as a share of GDP) compared to developed economies.

5.4 Limitations of This Study

There are several important limitations to this study. Firstly, we have not accounted for the effects of various factors, such as institutional quality, human capital, etc., that would have an impact on growth as well as the efficiency of infrastructure. In principle, it would not be difficult to include these variables in the dataset, but endogeneity and identification of causality would have complicated the analysis significantly. It is not the intent of the paper to wade into these debates. In defense, our position is that such factors would likely affect infrastructure and non-infrastructure returns equally, and perhaps not affect the key result in this paper too much.

Secondly, the regression results reflect the average effects of infrastructure. There is considerable heterogeneity across countries. While we suggest as a ballpark estimate that

[20] Population growth or employment effects were not considered in the (Devarajan, Swaroop, & Zou, 1996) study.
[21] (Fernald, 1999) for example assessed that repeat investment in roads would not give the same productivity boost compared to initial investment. See also (Holtz-Eakin & Schwartz, 1995), which found relatively low elasticity of infrastructure capital in the US.

developing economies should invest around 6-10 percent of GDP in infrastructure, we recognize there will be considerable differences from country to country. In this regard, development agencies would typically provide country-specific diagnostics and assess the infrastructure needs of each economy separately.

Finally, while we provide some evidence that developing economies are not investing sufficiently in infrastructure, the research has not studied the underlying root causes. Why developing economies do not invest as much in infrastructure as regression elasticity would remain unanswered. Nonetheless, there are well-known hypotheses amongst the development community—such as the higher up-front capital, longer payback period, and the divergence between social and private returns—that set infrastructure apart from other forms of private capital and explain the relatively low levels of investments, especially in developing countries. This is related to the previous point on the country-specific context. While we make the general recommendation that developing economies should invest more in infrastructure, the "how" will have to depend on in-depth country studies and would necessarily differ from country to country.

6 Conclusion

By conventional wisdom, developing economies face a large infrastructure gap and should invest more in infrastructure. Nevertheless, there has not been much macroeconomic research that provides evidence for this. In fact, developing economies' infrastructure needs are sometimes based on economic growth projections—that is, with causality running from GDP to infrastructure. This paper's key contribution is to provide evidence of the relative larger growth impact (elasticity) of infrastructure in developing economies. In doing so, and while acknowledging that country circumstances differ and that there are limitations of the study, the paper provides some support that developing economies should optimally invest a larger share, around6-10 percent of GDP in infrastructure, which is higher than corresponding percentages for developed economies.

References

Aggarwal, S. (2018). Do Rural Roads Creat Pathways Out of Poverty? Evidence from India. *Journal of Development Economics*, 375 - 395.

Aghion, P., & Schankerman, M. (1999). Competition, Entry and Social Returns to Infrastructure in Transition Economies. *Economics of Transition*, 79 - 101.

Alfaro, L., Kalemli-Ozcan, S., & Volosovych, V. (2008). Why Doesn't Capital Flow from Rich to Poor Countries? An Empirical Investigation. *Review of Economics and Statistics*, 347-368.

Allcott, H., Collard-Wexler, A., & D O'Connell, S. (2016). How Do Electricity Shortages Affect Industry? Evidence from India. *American Economic Review*, 586-624.

Aschauer, D. (1990). Why is Infrastructure Important? *Federal Reserve Bank of Boston*, 21-68.

Asian Development Bank. (2017). *Meeting Asia's Infrastructure Needs.*

Aynaoui, K., Agenor, P., & Bayraktar, N. (2008). Roads Out Of Poverty? Assessing the Links between Aid, Public Investment, Growth and Poverty Reduction. *Journal of Development Economics*, 277 - 295.

Calderon, C., & Serven, L. (2004). *The Effects of Infrastructure on Growth and Income Distribution.* World Bank Policy Research Paper 3400.

Devarajan, S., Swaroop, V., & Zou, H. (1996). The Composition of Public Expenditure and Economic Growth. *Journal of Monetary Economics*, 313 - 344.

Donaldson, D., & Hornbeck, R. (2016). Railroads and American Economic Growth: A "Market Access" Approach. *The Quarterly Journal Of Economics 131(2)*, 799-858.

Esfahani, H., & Ramirez, M. (2003). Institutions, Infrastructure and Economic Growth. *Journal of Development Economics*, 443 - 447.

Fay, M., & Rozenburg, J. (2019). *Beyond the Gap.* World Bank Group.

Fay, M., Lee, H., Mastruzzi, M., Han, S., & Cho, M. (2019). *Hitting the Trillion Mark: A Look at How Much Countries Are Spending on Infrastructure.* World Bank.

Fernald, J. (1999). Roads to Prosperity? Assessing the Link Between Public Capital and Productivity. *American Economic Review*, 619 - 638.

Holtz-Eakin, D., & Schwartz, A. (1995). Infrastructure in a Structural Model of Economic Growth. *Regional Science and Urban Studies*, 131 - 151.

International Monetary Fund. (2014). *World Economic Outlook October 2014.*

Kline, P., & Moretti, E. (2014). Local Economic Development, Agglomeration, and the Big Push: 100 Years of Evidence from the Tenessee Valley Authority. *Quarterly Journal of Economics*, 275 - 331 .

Lucas, R. E. (1990). Why Doesn't Capital Flow from Rich to Poor Countries. *American Economic Review*, 92-96.

Martin, P., & Rogers, C. (1995). Industrial Location and Public Infrastructure. *Journal of International Economics 39*, 335 - 351.

McIntosh, C., Algeria, T., Ordonez, G., & Zenteno, R. (2018). The Neighbourhood Impact of Infrastructure Invesment: Evidence from Urban Mexico. *American Economic Journal: Applied Economics*, 263 - 286.

Michaels, G. (2008). The Effect of Trade on the Demand for Skills - Evidence from the Interstate Highway System. *Review of Economics and Statistics 90(4)*, 683-701.

Normaz Wana, I., & Jamilah Mohd, M. (2015). The Impact of Infrastructure on Trade and Economic Growth in Selected Economies in Asia. *Asian Development Bank Institute Working Paper No.553*.

Nuno, L., & Venables, A. (2001). Infrastructure, Geographical Disadvantages, Transport Costs, and Trade. *The World Bank Economic Review 15(3)*, 451-379.

OECD. (2019). *Investment (GFCF) (indicator).*

OECD. (2019). *Investment by asset (indicator).*

Straub, S. (2008). Infrastructure and Development: A Critical Appraisal of the Macro Level Literature. *World Bank Policy Research Working Paper 4590*.

United Nations., E. C.-o. (2008). *System of National Accounts.*

World Bank. (2020). *Global Economic Prospects: Slow Growth, Policy Challenges.*

Yeaple, S., & Golub, S. (2007). International Productivity Differences, Infrastructure and Comparative Advantages. *Review of International Economics*, 223 - 242.

Appendix A:

List of Variables and Estimated Values

OECD Dataset
(annual data, 144 economies, 1960-2017)

Variable	**Definition**	**Unit**	**Source**
GDP_usd	Gross Domestic Product	Billions, Current US Dollar	Work Bank World Development Indicators
Population	Population, both sexes	Billions	Work Bank World Development Indicators
Employment Rate	Employment to population ratio, Age 15+, both sexes (modeled ILO estimate)	Percentage	Work Bank World Development Indicators
Worker	Employment to population ratio multiplies total population	Billions	Author Calculation
Output per Worker	GDP divided by the number of Worker	Current US Dollar	Author Calculation
GFCF_GDP	Gross fixed capital formation as percentage of GDP	Percentage	World Bank World Development Indicators
GFCF_usd	Gross fixed capital formation (GFCF) is defined as the acquisition of produced assets (including purchases of second-hand assets), including the production of such assets by producers for their own use, minus disposals (SNA 2008)	Billions, current US Dollar	OECD
GFCF-GG_usd	One of the asset types of gross fixed capital formation, named as “other buildings are structures”	Billion, current US Dollar	OECD
Infrastructure_usd	Same as “GFCF-GG_usd”	Billions, current US Dollar	Author Calculation
Non-Infrastructure_usd	Gross fixed capital formation minus Infrastructure investment	Billions, current US Dollar	Author Calculation
Infrastructure Output per worker	Infrastructure investment divided by the number of workers	Current US Dollar	Author Calculation
Non infrastructure Output per worker	Non-infrastructure investment divided by the number of workers	Current US Dollar	Author Calculation

IMF GFCF-GG Dataset
(annual data, 189 economies, 1960-2017)

Variable	Definition	Unit	Source
GDP_rppp	Gross domestic product	Billions, constant 2011 international dollars	IMF Investment and Capital Stock Database
Population	Population, both sexes	Billions	Work Bank World Development Indicators
Employment Rate	Employment to population ratio, Age 15+, both sexes (modeled ILO estimate)	Percentage	Work Bank World Development Indicators
Worker	Employment to population ratio multiplies total population	Billions	Author Calculation
Output per Worker	GDP divided by the number of Worker	Constant 2011 international dollars	Author Calculation
GFCF_GDP	Gross fixed capital formation as percentage of GDP	Percentage	World Bank World Development Indicators
GFCF_rppp	Gross fixed capital formation as percentage of GDP multiplies gross domestic product	Billions, constant 2011 international dollars	Author Calculation
GFCF-GG_rppp	General government investment (gross fixed capital formation)	Billions, constant 2011 international dollars	IMF Investment and Capital Stock Database
PPP, total	Public-private partnership (PPP) investment, total commitment	Billions, constant 2011 international dollars	IMF Investment and Capital Stock Database
Infrastructure_rppp	General government investment plus Public-private partnership investment	Billions, constant 2011 international dollars	Author Calculation
Non-Infrastructure_rppp	Gross fixed capital formation minus Infrastructure	Billions, constant 2011 international dollars	Author Calculation
Infrastructure Output per Worker	Infrastructure investment divided by the number of workers	Constant 2011 international dollars	Author Calculation
Non-Infrastructure Output per Worker	Non-infrastructure investment divided by the number of workers	Constant 2011 international dollars	Author Calculation

Appendix B: Geographic Coverage between OECD and GFCF-GG Dataset

Africa

GFCF-GG Only	Both GFCF-GG&OECD	OECD Only
Algeria	South Africa	
Angola		
Benin		
Botswana		
Burkina Faso		
Burundi		
Cabo Verde		
Cameroon		
Central African Republic		
Chad		
Comoros		
Congo, Dem. Rep.		
Congo, Rep.		
Cote d'Ivoire		
Djibouti		
Egypt, Arab Rep.		
Equatorial Guinea		
Eswatini		
Ethiopia		
Gabon		
Gambia, The		
Ghana		
Guinea		
Guinea-Bissau		
Kenya		
Lesotho		
Madagascar		
Malawi		
Mali		
Mauritania		
Mauritius		
Morocco		
Mozambique		
Namibia		
Niger		
Nigeria		
Rwanda		
Sao Tome and Principe		
Senegal		
Sierra Leone		
Sudan		
Tanzania		
Togo		
Tunisia		
Uganda		
Zambia		
Zimbabwe		

Oceania

GFCF-GG Only	Both GFCF-GG&OECD	OECD Only
Fiji	New Zealand	Australia

Americas

GFCF-GG Only	Both GFCF-GG&OECD	OECD Only
Argentina	Brazil	
Bahamas, The	Canada	
Barbados	Chile	
Belize	Colombia	
Bolivia	Costa Rica	
Dominican Republic	Mexico	
Ecuador	United States	
El Salvador		
Guatemala		
Haiti		
Honduras		
Nicaragua		
Panama		
Paraguay		
Peru		
St. Lucia		
St. Vincent and the Grenadines		
Suriname		
Uruguay		
Venezuela, RB		

Europe

GFCF-GG Only	Both GFCF-GG&OECD	OECD Only
Albania	Austria	
Belarus	Belgium	
Bosnia and Herzegovina	Czech Republic	
Bulgaria	Denmark	
Croatia	Estonia	
Malta	Finland	
Moldova	France	
Montenegro	Germany	
North Macedonia	Greece	
Romania	Hungary	
Russian Federation	Iceland	
Serbia	Ireland	
Ukraine	Italy	
	Latvia	
	Lithuania	
	Luxembourg	
	Netherlands	
	Norway	
	Poland	
	Portugal	
	Slovak Republic	
	Slovenia	
	Spain	
	Sweden	
	Switzerland	
	United Kingdom	

Asia

GFCF-GG Only	Both GFCF-GG&OECD	OECD Only
Armenia	Israel	
Azerbaijan	Japan	
Bahrain	Korea, Rep.	
Bangladesh		
Bhutan		
Cambodia		
China		
Cyprus		
Georgia		
Hong Kong SAR, China		
India		
Indonesia		
Iran, Islamic Rep.		
Iraq		
Jordan		
Kazakhstan		
Kuwait		
Lebanon		
Macao SAR, China		
Malaysia		
Maldives		
Mongolia		
Myanmar		
Nepal		
Oman		
Pakistan		
Philippines		
Saudi Arabia		
Singapore		
Sri Lanka		
Syrian Arab Republic		
Tajikistan		
Thailand		
Turkey		
United Arab Emirates		
Uzbekistan		
Vietnam		
Yemen, Rep.		