THE WILLIAM DAVIDSON INSTITUTE
AT THE UNIVERSITY OF MICHIGAN

# Infrastructure and Growth: Empirical Evidence

*By: Balazs Egert, Tomasz Kozluk and Douglas Sutherland*

William Davidson Institute Working Paper Number 957
April 2009

# Infrastructure and growth: Empirical evidence

Balázs Égert[‡] Tomasz Koźluk[†] Douglas Sutherland[§]

Abstract

Investment in network infrastructure can boost long-term economic growth in OECD countries. Moreover, infrastructure investment can have a positive effect on growth that goes beyond the effect of the capital stock because of economies of scale, the existence of network externalities competition enhancing effects. This paper analyses the empirical relationship between infrastructure and economic growth. Time-series results reveal a positive impact of infrastructure investment on growth. They also show that this effect varies across countries and sectors and over time. In some cases, these results reveal evidence of possible over-investment. Bayesian model averaging of cross-section growth regressions confirms that infrastructure investment in telecommunications and the electricity sectors has a robust positive effect on long-term growth (but not in railways and road networks). Furthermore, this effect is highly nonlinear as the impact is stronger if the physical stock is lower.

*JEL Classification: E22, O11, O40*

*Keywords*: investment, infrastructure, network industry, economic growth, cointegration, Bayesian model averaging

---

[‡] OECD, Economics Department; CESifo; University of Paris X-Nanterre and the William Davidson Institute. Email: balazs.egert@oecd.org

[†] OECD, Economics Department, Tomasz.kozluk@oecd.org

[§] OECD, Economics Department, douglas.sutherland@oecd.org

The authors are indebted to Jørgen Elmeskov, Giuseppe Nicoletti, Jean-Luc Schneider, Fabio Schiantarelli and participants of an OECD internal seminar for useful comments, as well as Irene Sinha for editorial assistance. The opinions expressed in this paper are those of the author and are not necessarily shared by the OECD.

# 1. Introduction

A wide debate on the influence of infrastructure on output levels and growth has led to attempts to quantify this effect and to ask about optimal levels of investment in infrastructure, particularly over the past two decades. While there is a wide consensus that some basic level of infrastructure is necessary for development, the ranges of estimates of the effects of infrastructure have varied widely.

Infrastructure has often been seen as increasing productivity and attracting business activity by lowering transport and production costs and facilitating market access. Some of these effects cancel out on the aggregate level, given that infrastructure has to be paid for. However, even in the absence of the "free input" effect, the effects of infrastructure on output on the aggregate level may still differ from the effects of total capital for a number of reasons. First, there may be significant economies of scale that differ from other investments. Second, network externalities may characterise infrastructure investments, through connecting both regions and countries. Third, infrastructure may have a competition enhancing effect, allowing for improved market access, such as through lowering transport costs. However, the causal link between infrastructure and growth may operate in the opposite direction, as countries with high levels of output will also be able to fund higher infrastructure investments, which may be desirable for social reasons. Moreover infrastructure investment will, to some extent, reflect expectations of future capacity utilization.

The focus in this paper is on physical capital stocks in network sectors: transport (roads, motorways and railways) and non-transport (electricity, telecommunications). All these sectors can be expected to have network externalities and large economies of scale. Their expansion can be fostering competition in other segments by facilitating market access through lowering the costs of transport and communication.

After briefly reviewing previous research, this paper applies a simple exogenous growth model to capture the effect of physical infrastructure levels on GDP per capita in an annual panel of OECD countries since the 1960s. The data are then described and the estimations reported. This is followed by the estimation of cross-section growth regressions for multi-annual periods. This part of the paper describes the main methods of model selection, before presenting results for linear and non-linear models.

The main conclusions are as follows. First, regarding measures of infrastructure, data quality limits the scope of the empirical work. National account data for investment or capital stocks in infrastructure sectors are available in long series for only a handful of countries and still contain methodological differences. Measures of infrastructure quality are even more scarce and of poorer comparability. Overall, the most robust available measures of infrastructure for a sample of OECD countries over time are physical indicators. It should be noted that much of the literature seems to confuse infrastructure with public capital stocks or public investment, which, due to corporatisation, privatisation and market liberalisation are increasingly unreliable measures of infrastructure.

Regarding annual time-series growth regressions, we find that the contributions of infrastructure to long-run output levels and growth are not homogenous across countries. Results indicate that the expansion of infrastructure could be both more or less productive with respect to other capital expenditure. Furthermore, the result that more does not always mean better (in terms of GDP per capita) seems to be robust across different specifications including control variables such as human capital, trade openness and tax revenues. Finally, the validity of the full sample estimates holds for more recent years, and in most countries the effect does not seem to change.

For multi-annual cross-section growth regressions, the evidence based on Bayesian averaging of classical estimates suggests greater provision of broad measures of infrastructure is associated with higher subsequent growth rates. The results also suggest that the link is non-linear, with a potentially higher impact of additional infrastructure in countries with initially lower levels of provision.

## 2. Previous research

In much of the literature, especially aggregate-level studies, public capital is often regarded as a synonym for (public) infrastructure.[1] Therefore, the effects of public investment (general government gross fixed capital formation in Sturm et al. 1999) or estimates of public capital (Kamps, 2005), are often assumed to be the effects of infrastructure on growth or output levels. However, this assumption is increasingly problematic. First, due to corporatisation and the privatisation of firms in network industries in many of the OECD countries, together with a liberalisation of entry into these sectors, much of the physical capital and investments are no longer classified as government. In sectors such as telecommunications, electricity or rail, most entities are not included in the definition of general government.[2] Furthermore, a growing share of government fixed capital formation will often include investment in schools, hospitals and government buildings.

---

[1] Romp and De Haan (2007) survey the effect of public investment and infrastructure on growth.

[2] In national accounts (SNA93/ESA55), general government excludes state-owned corporations, quasi-corporations and public utility firms.

In the economic literature, a number of channels through which infrastructure can affect aggregate GDP levels and growth have been identified. A standard approach, following Aschauer (1989), incorporates infrastructure into the production function as a third input with capital and labour.[3] Infrastructure is treated separately, due to those features that distinguish it as an input and most of the specifications allow the use of physical stocks. Alternatively, infrastructure can be treated as a total factor productivity augmenting input: by lowering the costs of production (*e.g.* through the costs of transport and communication) it increases the technological index.[4]

From the empirical point of view, two mainstream approaches can be distinguished. First, a large number of papers adopt variations of a production function approach and estimate either a simultaneous equation model with a production and an investment function or a closed-form solution to a growth model most commonly based on a Cobb-Douglas or trans-log production function. Examples include:

- *Cross-country:* Calderon and Severn (2002) estimate the effect of various types of physical infrastructure on growth and inequality for over 100 countries[5] or Esfahani and Ramirez (2002) who develop and estimate a structural model of infrastructure and growth for 75 countries.[6]
- *National:* for both one country or for a group or panel of countries, see for instance Ford and Poret (1991), for a study on OECD countries,[7]
- *Regional:* La Ferrara and Marcellino (2000) study the effect of infrastructure on productivity in Italian regions[8] while Stephan (2000) attempts to assess the effect of transportation infrastructure stock in French and German regions.

---

[3] Aschauer's (1989) results sparked a discussion on the effects of public capital on output and productivity. His results, showing that the productivity of public investment can be much higher than that of private investment, caused strong controversy (being criticised from many methodological points of view, ranging from inappropriate estimation techniques to the specific characteristics of the period analysed).

[4] Sturm *et al.* (1998) showed that an estimated Cobb-Douglas production function cannot distinguish between these two specifications.

[5] The authors focus on Sub-Saharan Africa and experiment with both quantity and quality variables for infrastructure (electricity, road and telecommunications). They include a number of control variables in their equation to account for education, trade, financial development, institutional quality etc. Using the system GMM of Arellano and Bover (1995) they find strong positive effects of infrastructure on growth.

[6] The authors find benefits from infrastructure investment and performance in infrastructure sectors but show that achieving better outcomes (in terms of growth) requires institutional and organisational improvements.

[7] The study of 11 OECD economies uses two alternative definitions of public capital, of which the "broad" definition includes structures in transport, communications and electricity to test its influence on total factor productivity. The results show mixed experiences for OECD countries.

- *Industry:* Shanks and Barnes (2008) estimate the effect of road and communications infrastructure on industry-level multi-factor productivity growth for Australia.

The alternative mainstream approach is estimating a cost-function approach based on the idea that firms optimise by maximising profits given a price of output and a cost function which includes infrastructure as one of the inputs. From a firm level this input is often assumed as fixed, externally given (*i.e.* exogenous) and "free", but a firm decides on the amount of input it uses resulting in an aggregate demand for infrastructure – thus determining an environment in which it operates. Infrastructure is assumed to have a cost-reducing effect for firms. The cost-function approach can also be adopted at various levels of aggregation:

- *National and cross-country*: Demetriades and Mamuneas (2000) generally find positive effects of public capital on output supply and input demand for 12 OECD countries.[9]
- *Regional*: La Ferrara and Marcellino (2000) who tend to find signs of overinvestment in Central Italy.
- *Industry:* Moreno *et al*. (2003) find wide heterogeneity of the cost elasticity of infrastructure capital across industries and regions of Spain for the 1980s.

One strand of the empirical literature attempts to tackle the issue of geographical network externalities. For example, Fernald (1999) finds high productivity of additional roads in the United States until the interstate highway system was completed, but cannot reject the effect being different from zero once the network was broadly completed. This allows benefits from the abundance of infrastructure in neighbouring regions or countries, or where firm suffer from the fact that others use infrastructure.

Finally, a pool of studies uses so-called vector auto-regressions or vector error-correction models (VAR and VECM, respectively). These are somewhat similar to a production function approach,

---

8. Among the approaches taken by the authors is the estimation of a production function approach where public capital, acting as a proxy for infrastructure, has a productivity augmenting effect on output (measured by regional value-added in the manufacturing sector). They find mixed effects across four Italian macro-regions though note that the use of a common perpetual inventory method to construct the public capital variable may drive part of the effect as different regions may differ in the efficiency of spending the money devoted to investment.

9. The results obtained using the inter-temporal optimisation framework for the manufacturing sectors of the 12 OECD countries point to a general under-supply of public infrastructure in the countries, though large heterogeneity can be observed, both across time and countries.

using the same variables, but allow more flexibility in modelling the cross-relationships between the variables. For example, Flores de Frutos *et al.* (1998), find permanent effects of shocks to public infrastructure on output, employment, private and public capital in Spain. However, most of these models suffer from the problem of short series, tend to run out of degrees of freedom quickly and the assumptions necessary to identify shocks are often unconvincing. On the positive side, these approaches are better able to deal dealing with reverse causality, which is one of the main problems of estimating the effect of infrastructure on growth.

*The problem of causality*

1. In general, causality is difficult to establish convincingly in growth regressions. In principle, the saving rate is exogenous in the benchmark Solow model, implying that the effect of any type of investment (be it in total capital, infrastructure or human capital) on growth should be unidirectional. In practice, however, this assumption is too restrictive as the saving rate may be influenced by the growth rate. In empirical settings, a number of approaches have been used to address the problem of causality:

- *Instrumental variables*: The first notable attempt in this setting was by Aschauer (1989) who used lagged investment as an instrument, which is questionable. Finding a convincing instrument that is correlated with infrastructure variables but not with output has proved extremely difficult.
- *Granger causality:* Another approach is to test for Granger causality. However this approach requires either the estimation of a multivariate system which would require a long data set or restricting to bi-variate causality tests where the omitted variables would pose a problem for the estimated coefficients. In general, Granger causality must be used with caution, as it does not reflect the pure concept of causality, rather the usefulness for predictions, and despite finding a variable Granger-causes another the variables may be driven by a third process.[10]

---

[10] Simple bivariate Granger causality tests were conducted on the stationary first differences of per capita GDP and infrastructure variables, both in single series and in a panel setting. With 4 lags, the tests lack power, and can only reject the lack of Granger causality (in both directions) for energy and telecommunications in a panel setting. The panel setting imposes common cross country coefficients, which as will be shown later is not necessarily correct. The single series tests do not provide strong results rejecting (at 10%) the lack of Granger causality from infrastructure to GDP for only a handful of countries, mostly in the case of electricity and motorways (the results are reported in the Appendix Tables A.5 and A.6).

- *Simultaneous equations*: Other approaches propose simultaneous equation estimation derived from simple structural models. These in turn require assumptions on the channels through which output influences infrastructure investments (see Esfahani and Ramirez 2003) in particular on the political decision-making process of public investment (*e.g.* Cadot *et al.* 1999, 2006).

- *Impulse response functions:* Another way of going about causality, applied using public capital estimates are impulse response functions, as in Kamps (2005). They require estimation of a VAR system with all the variables in the model, together with their lags and tend to quickly run out of degrees of freedom. Thus with relatively short time series the standard errors are likely to be large, which is most probably why the author of the aforementioned paper reports 68% confidence intervals, while not accounting for the large uncertainty regarding the construction of one of the independent variables (public capital stock) and still fails to find much significant effects.

- *Using stocks rather than flows*: Other authors (see Arnold *et al.* 2007) argue that the use of stocks (usually human capital) instead of flows reduces the problem of reverse causality. The reasoning is based on the fact that the feedback from output to stocks, which contain an accumulated investment over many years, will be smaller than in the case of investment. Although the argument should in principle apply to infrastructure (physical capital which is generally composed of assets with a relatively long life span), it only reduces the problem and does not eliminate it convincingly enough to disregard the issue of causality in long-run relationships.

- *Assuming it away:* Canning and Bennathan (2000) argue that the use of panel data sets rids of the issue of causality if one assumes a common long run production function relationship across countries while allowing for heterogeneous investment relationships and pools the data across countries.

- *Exploiting co-integrating relations:* Finally, Canning and Pedroni (2008) have attempted to tackle the issue by estimating small country-specific error-correction models for GDP per capita and infrastructure per capita. However, the co-integrating relationship between the two variables is underpinned by strong assumptions, such as the instantaneous depreciation of infrastructure, in order to overcome the issue of omitted variables and reduce the problem to two variables. Other details of the specification, such as conceptual problems of being, on average, above or below an estimated equilibrium, or the possibility of the inversion of the results towards the end of the sample show that the problems with establishing causality persist.

As causality is not convincingly established in any of these approaches in simple aggregate growth regressions, it is often dealt with at industry or company level where one can more easily assume that infrastructure is exogenous, or in cross section regressions such as those reported in Section 6.

## 3. Basic Model

The model underlying the empirical estimations is based on a simple exogenous growth framework proposed by Mankiw, Romer and Weil (1992), MRW hereafter. The model is based on a human capital augmented production function where human capital is treated as an 'additional' factor of production to capital, population and technology:

$$Y(t) = K(t)^{\alpha} H(t)^{\beta} (A(t)L(t))^{1-\alpha-\beta} \tag{1}$$

where: Y, K, H, A and L represent GDP; total capital; human capital; the level of technology and the labour force, respectively. With $\alpha + \beta < 1$ the production function $\boldsymbol{\alpha + \beta < 1}$ exhibits decreasing returns to all capital. Capital accumulation functions (where lower cases indicate variables per effective unit of labour e.g. *y=Y/AL*) are given by:

$$\begin{aligned} \dot{k}(t) &= (s^{K})y(t) - (n+g+\delta)k(t) \\ \dot{h}(t) &= (s^{H})y(t) - (n+g+\delta)h(t) \end{aligned} \tag{2}$$

Given the production function and capital accumulation functions, MRW derive their basic specification:[11]

$$\ln(\frac{Y_t}{L_t}) = \ln(A_0) + gt + \frac{\alpha}{1-\alpha}\ln(s_t^{K}) + \frac{\beta}{1-\alpha}\ln(h_t\,) - \frac{\alpha}{1-\alpha}\ln(n_t + g + \delta)$$

This is the basis for the approach taken in this paper, where as a first step infrastructure stock (inf) is assumed to be a factor of production. Appropriately, an equivalent equation can be derived:

---

[11] This is equation 12 in the MRW paper, which uses the saving rate and stock of human capital. A second version of the basic model includes the savings rates into both types of capital. As data on investment into capital stock in infrastructure sectors is practically unavailable for a broad set of OECD countries and for a long time series, and the comparability of the available series is poor, this paper focuses on the specification with one of the variables taken as stock. By construction the estimated coefficients differ slightly but in principle not in sign or significance.

$$\ln(\frac{Y_t}{L_t}) = \ln(A_0) + gt + \frac{\alpha}{1-\alpha}\ln(s_t^K) + \frac{\beta}{1-\alpha}\ln(\inf_t) - \frac{\alpha}{1-\alpha}\ln(n_t + g + \delta) \qquad (3)$$

This equation is then used to estimate the long run levels relationship, with human capital also an important control variable.[12]

One of the primary aims of this exercise is to single out infrastructure as a separate, additional factor of production, as has been done with human capital in MRW. However, a difference relative to human capital is that infrastructure capital is generally included in measures of total capital stock. However, there are reasons to believe that the effect of infrastructure capital can be different than that of an average unit of capital. For example, the infrastructure stock often exhibits features of a natural monopoly, tends to have public good characteristics, network effects and spillovers into other sectors. Furthermore, investments are often large and their life-cycle long and will differ in terms of financing (private versus public). In this context, when physical infrastructure is included together with total capital stock or total capital investment the coefficient can be crudely interpreted as the additional effect of infrastructure, which is the different effect on GDP relative to other types of capital.

Interpreting the additional effect of infrastructure is difficult. If both stocks were measured using the same units the increase in capital would include a proportional increase in physical infrastructure stock and the total effect would be α+ β. Since the equation is estimated with infrastructure stocks measured in physical units, this is only a crude approximation of the effect. In this light, direct comparison is problematic, especially in the specification using investment rates, as the coefficients are only proportional to α and β and with relatively wide estimated confidence intervals. Nevertheless, the sign and significance of the coefficient on infrastructure can be interpreted as an indication of the additional effect that investment in infrastructure would have – a positive significant coefficient may indicate that expanding infrastructure would be associated with higher output while a negative one may indicate inappropriate investment – in which case investment in other types of capital may be more productive.

The framework is an exogenous growth model, and hence by construction an effect of infrastructure on long-run GDP levels and on short run GDP growth can exist, while the long-run growth rates at the steady state are exogenously determined by technological progress.[13] Moreover, the effect of

---

[12] Equation 3 could also be derived to include the level of human capital, though the notation becomes more cumbersome.

[13] This may not hold in all circumstances (see Bond *et al.* 2004). For instance, Kocherlakota and Yi (1997) propose an interesting approach to testing growth exogeneity versus endogeneity by including (lags of) public investment together with tax revenues on the right hand side of the equation and testing for joint

infrastructure on GDP levels and short run growth cannot be assessed by looking solely at the long-run coefficient. In order for an equilibrium correction mechanism to exist (and thus to justify the two-step approach) the short-run coefficient on the ECM should be negative.

## 4. Data

One limitation of the exercise is that the quality of the data is poor. National Accounts data on investment or capital stock volumes in specific sectors are available only for a handful of OECD countries, and usually for only a very short time and of uncertain comparability. For example, infrastructure capital stock estimates rely on assumptions about depreciation and scrapping rates, which are often poorly observed. Physical measures are available for a wider coverage of countries and for much longer time periods. However, physical capital stocks have a number of shortcomings. Most of the available data do not contain any information on differences in cost and quality. For example, the costs of setting up the infrastructure can vary markedly (an additional kilometre of road or rail track would be more expensive if requiring a bridge or a tunnel), while the quality of infrastructure may also vary (well maintained stocks may yield more benefits than poorly maintained ones). Additionally, even the physical capital stock data encounters the problem of the lack of a uniform methodology across countries. To the extent that this does not change across time, the problem would be overcome by the inclusion of country-specific fixed effects in the regression.

The variables are in natural logarithms, and the broadest approach is based on the year sample 1960-2005 for 24 OECD countries (the Czech Republic, Hungary, Poland, Slovakia, Germany and Luxembourg are excluded due to problems with the data). In practice, the panel is unbalanced (with a minimum of 16 observations per country) and the inclusion of countries into the individual specifications is determined by the availability of the individual variables.

In general, the infrastructure variables used in the model are in total physical stocks per capita (in line with the basic model specification), *i.e. roads*, *motorways* and *railtracks* are in kilometres of length per capita, *electricity* is defined as total plant generating capacity and *telephone mainlines* as number of fixed lines (both per capita). At first glance some types of infrastructure may make sense when measured in terms of density (for example, the length of roads infrastructure may seem more meaningful with respect to land area in some cases). However, as the estimation uses fixed effects, this is no different from including the raw level of infrastructure (dividing by land area is equivalent to dividing by any other constant). These values are used as an additional robustness check. In the

---

coefficient significance. While tax revenues seem a potentially attractive counterpart to public investment, finding an equivalent variable for infrastructure is problematic.

case of telephone mainlines, the series exhibit fast growth in the first 30-40 years of the sample with a sharp slowdown and even fall in most countries in the last decade. This is probably related to technological change, such as the introduction of mobile phones and lines with greater bandwidth. In order to try to account for this we included the variable of total (mobile and fixed line) subscribers. However, this is not a pure infrastructure variable, and a rise in this variable could actually reflect an increase in use (possibly leading to congestion) rather than an increase in actual infrastructure.

The dependent variable is measured as GDP per capita (of 15-64 population) in 2000 PPP terms. The other variables include saving or actually investment rate (measured using total gross fixed capital formation relative to GDP), private gross fixed capital formation relative to GDP, total physical capital stock to GDP, human capital measured as the average number of years of schooling for the adult population, general government tax revenue relative to GDP and trade openness as imports plus exports relative to GDP.

### *Time-series properties*

The time series properties of the variables were examined. The results for both single series and panel unit root test were:

- In the single series tests, the null hypothesis that they contain a unit root cannot be rejected for the majority of the series, both when including and when not including a deterministic trend (see Appendix Table A.1 for excerpts of the results). The hypothesis that their first differences are non-stationary can be rejected, implying that most of the series (with or without a deterministic trend) are I(1). Some of the results suggest that infrastructure variables could be regarded as I(0).
- A set of panel unit root tests were run using the Levin, Lin and Chu (2002) test which imposes the fairly strong assumption of a common unit root process across the series and the Im, Pesaran and Shin (2003) test which relaxes this restriction. The latter must be treated with caution, as the null hypothesis is formulated as all series contain a unit root while the alternative hypothesis is that there is a non-zero share of stationary series in the panel. Thus, if the panel includes both stationary and non-stationary series the test will tend to reject the null.[14] Among the tested series this may be the case for GDP per capita,

---

[14] For instance, when testing for a unit root in GDP per capita levels the null for a panel of all countries can be rejected, but if any two of the three countries which seemed to have stationary series according to the ADF single series tests are excluded, the null can no longer be rejected even at 90% confidence levels. In case of the investment rate to GDP the situation is similar, therefore for these two variables the ADF tests on single (country) series are reported in the Appendix, showing that for most countries the assumption that

investment rates or some of the infrastructure series for which the evidence in single series test was mixed.[15] An excerpt of the results is presented in Appendix Table A.2.

Clearly the comparatively small sample of OECD countries limits the strength of the conclusions.[16] However, one can fairly confidently assume that log GDP is an integrated variable when looking at results from wider panels used in other studies (Canning, 1999; Canning and Pedroni, 2008).

The next step involves testing for cointegration in the various specifications (reported in Appendix Table A.3). For the majority of proposed specifications in levels which represent long-run equations, the null hypothesis of no cointegration can be rejected, regardless of whether we assume a homogenous panel (as in Kao, 1999) or a heterogenous panel (as in Pedroni, 1999). The inclusion of a deterministic trend, along the line of the MRW model, further improves the rejection rates in heterogenous panels.

## 5. Annual time-series estimations

The time-series properties of the data determine the empirical estimation approach. Despite some of the relevant tests being inconclusive, one cannot exclude the possibility of there being a co-integrating relationship among variables which contain a unit root. As a result, the estimation uses the so-called two-step Engle-Granger (1987) approach in a heterogeneous panel setting to address this possibility. In this approach, the long run levels equation is estimated, first, as follows:[17]

$$\ln(y_{it}) = b_{0i} + b_{1i}\ln(s_{it}^{K}) + b_{2i}\ln(\inf_{it}) + b_{3i}n_{it} + b_{4i}t \qquad (4)$$

The estimation is done using Dynamic OLS, which has favourable asymptotic and finite-sample properties (see Stock and Watson, 1993 or Mark and Sul, 2003) in estimating a long run relationship. Dynamic OLS includes lags and leads (in this case 2 and 1 respectively, due to the relatively short time span of the sample) and contemporaneous values of the first differenced variables in order to

---

variables are integrated of order one is reasonable. Both panel and single series tests point to the stationarity of the series of motorways per capita however.

[15] In principle the possibility that some of the variables are indeed stationary will not invalidate the results, as the methodology used allows for the presence of such variables. However, if the case was reversed, *i.e.* stationarity was wrongfully assumed and the possibility of unit roots ignored during the estimation, one would encounter problems of spurious correlations and invalid inference.

[16] The single series tests are troubled by the problem of low power in a relatively short sample.

[17] In the panel setup, due to the fact that the yearly population growth rate can be negative, the levels value has been used instead of the logarithm, thus the coefficient should be interpreted as a semi-elasticity. In the original cross-section setting of MRW the population growth rate is a non-negative multi-year average, thus poses no such problem.

yield asymptotically efficient estimates of the coefficients. The residuals from the first step are then plugged into the short run first difference equation with a lag of one period:

$$\begin{aligned}\Delta \ln(y_{it}) &= B_{0i} + \sum\nolimits_{s=0}^{T_1} B_{1it-s} \Delta \ln(s_{it-s}^{K}) + \sum\nolimits_{s=0}^{T_2} B_{2it-s} \Delta \ln(\inf_{it-s}) \\ &+ \sum\nolimits_{s=0}^{T_3} B_{3it-s} \Delta n_{it-s} + \lambda_i ECM_{it-1}\end{aligned} \tag{5}$$

where the lag structure is determined by the data using the AIC criterion. This second step is estimated using OLS (due to the fact that it should contain only stationary variables).

The main focus of the paper is on the long run coefficients, though certain properties of the short run cannot be ignored, *e.g.* the negative coefficient on the ECM term which will guarantee reversion to the equilibrium relationship. Most generally, the model is estimated for a heterogenous panel, with country-specific coefficients and for a homogenous panel with common coefficients for all countries. The intermediate approach has also been explored – applying the Pooled Mean Group estimator where the long run coefficients are homogenous (*i.e.* common across countries) while the short-run dynamics are not restricted to be the same across countries. However, coefficient homogeneity tests (both Wald Chi-squared and Hausman's tests) strongly reject the null of coefficient homogeneity.[18] This indicates large individual coefficient heterogeneity and in the case of the PMG estimator, this points to the inconsistency of the pooled estimates. Therefore, for the estimations the mean-coefficients for every variable (which are the averages of the individual coefficients and not the pooled mean group coefficients) and the individual coefficients for the infrastructure variable have been reported.

The basic regression formulation used for interpretation includes the total economy investment rate, physical infrastructure stocks per capita, population growth and a deterministic trend. As a first robustness check, the regressions are run using total capital stock instead of the investment rate. While the coefficient sizes can be affected (due to different derivation of the model) the signs should not be affected if the relationship is robust. Additionally private investment is used, and control variables are inserted one by one. The estimations use each infrastructure stock variable separately. The inclusion of the physical stock variables one-by-one (instead of all at a time) is determined by the amount of data available. This raises the possibility of an omitted variable bias. However, the coefficients on the other variables seem relatively robust to this approach and the inclusion of too many infrastructure variables causes the errors to increase dramatically due to multicollinearity.

---

[18] In the Wald Chi-squared test for the equality of coefficients across countries is rejected for all the types of infrastructure at 99% confidence levels.

One of the common problems of growth regressions is the abundance of potential explanatory variables (Temple, 2000, see also Section 6). There may be many factors explaining growth and the available data does not allow the inclusion of all of them in one specification. Furthermore, including even a sub-set may be undesirable due to the fact that many variables are likely to be collinear. In a simple attempt to extend the basic specification, the robustness of the results is checked by including, one-by-one, potential additional explanatory variables: first of all human capital (proxied by average years of schooling of the working age population) which is one of the standard explanatory variables. Next, the level of tax burdens (proxied by tax receipts to GDP) and trade openness (proxied by the sum of imports and export relative to GDP) are added. As a further robustness check, an additional set of estimations was run using private investment instead of total investment, as there is likely to be less of a problem of double counting infrastructure in the investment measure. However, as infrastructure investment throughout our sample may have been increasingly done by private companies, as a consequence of corporatisation, privatisation and the liberalisation of entry, the results have to be treated with caution.

The lack of comparable cross-country series is a serious limitation to this type of analysis. First, while physical stocks have advantages over estimates of capital stock and investment series, they nonetheless fail to account for quality and are a crude measure of the abundance of infrastructure. Second, as usual in the simple growth regression approaches, causality can work both ways, partly to the extent that infrastructure investments reflect expectations of future growth and higher output levels. Therefore, the results need to be considered with some caution.

## 5.1. Results

The results for the general specification are presented in Tables 1 and 2, while the robustness analysis follows. The estimated mean group coefficients on investment are significant and in line with the estimated effects of capital obtained by MRW and subsequent studies. The estimated coefficients on investment range from 0.39 to 0.53 in the specifications without human capital and are smaller when human capital is included ranging from 0.3 to 0.42, which is similar to the MRW finding.[19] The coefficient on population growth is not directly comparable with MRW but is similarly insignificant. The time trend, which accounts mainly for technological progress is significantly positive, but becomes insignificant once human capital is included. Human capital is usually not significant, contrary to many previous findings, but is significant when using different

---

[19] These estimated coefficients reflect $\alpha /(1- \alpha)$ from equation (3) and can be used to obtain output elasticities with respect to capital $\alpha$ which are in the range of 0.28-0.35 with human capital excluded and 0.23 to 0.30 when it is included.

estimation techniques.[20] As mentioned the mean group coefficient on infrastructure capital, is in most cases not significant (with the exception of electricity generating capacity, where the estimated mean group coefficient is 0.17, significant at 95%).[21]

---

[20] For illustrative purposes, Appendix Table A.6 reports estimation results for a long-run panel with a homogenous coefficient assumption, where indeed the coefficient on human capital tends to be significant more often, in line with previous estimates. Estimates obtained via Dynamic OLS tend to have wider confidence intervals than the Pooled Mean Group estimates and the use of mean group coefficients results in a further widening of the confidence intervals. Moreover, human capital to some extent may be collinear with the infrastructure variables as the variable resembles a deterministic trend in many countries, which causes problems with identification.

[21] This may be in part due to the limited sample, which consists of a subset of OECD countries, all of which are relatively high-income countries and are relatively abundant in all the examined types of infrastructure. Moreover the limited time span of the sample plays a large role

**Table 1. Time-series estimation results: transport infrastructure**

| | Roads | | Rail | | Motorways | |
|---|---|---|---|---|---|---|
| | Coefficient | Coefficient | Coefficient | Coefficient | Coefficient | Coefficient |
| | | | ***Long run (mean group)*** | | | |
| Investment | 0.46 ** | 0.3 * | 0.53 *** | 0.39 ** | 0.42 *** | 0.4 *** |
| Population growth | 0.032 | 0.019 | 0.013 | -0.007 | 0.019 | -0.005 |
| Human capital | | 0.18 | | 0.08 | | -0.03 |
| trend | 0.02 *** | -0.03 | 0.03 *** | -0.08 | 0.02 *** | -0.11 |
| | | | *Country specific coefficients for infrastructure* | | | |
| Australia | 0.17 | 0.07 | 0.46 *** | 0.50 *** | | |
| Austria | -0.13 | 0.07 | 2.27 *** | 1.04 *** | 0.30 *** | 0.17 *** |
| Belgium | 0.27 | 0.12 | -1.01 *** | -0.39 ** | 0.18 *** | 0.12 |
| Canada | | | 0.45 | 3.02 | | |
| Denmark | 1.19 * | -0.75 | -0.20 *** | -0.11 | 0.15 *** | 0.10 |
| Finland | 1.66 | -0.32 | 0.29 | -0.48 | 0.01 | 0.00 |
| France | -0.81 *** | -0.52 *** | -2.52 *** | 2.21 ** | 0.14 *** | 0.09 |
| Greece | -0.09 *** | -0.09 ** | 2.22 *** | 0.93 *** | | |
| Iceland | -1.45 *** | | | | | |
| Ireland | -2.29 *** | 0.83 | 2.02 *** | 0.03 | 0.00 | 0.00 |
| Italy | -0.28 *** | -0.04 | -0.94 *** | -0.45 | 0.17 *** | 0.06 |
| Japan | 0.64 | 1.43 | 2.46 *** | 0.28 | 0.17 *** | 0.13 *** |
| Korea | 0.17 | | 1.06 *** | | | |
| Mexico | 0.17 * | | | | | |
| Netherlands | -0.45 * | -0.75 *** | -0.15 | -0.91 *** | 0.12 ** | 1.00 *** |
| New Zealand | 1.85 *** | 2.51 *** | 0.95 *** | 1.45 *** | -0.34 *** | 0.05 |
| Norway | 0.75 * | 1.21 | -1.37 * | -0.13 | | |
| Portugal | 0.30 *** | -0.04 | 0.09 | -0.44 *** | -0.16 *** | 0.00 |
| Spain | -0.43 * | -0.48 ** | -1.28 *** | -1.95 *** | 0.17 *** | 0.16 *** |
| Sweden | -0.14 | -0.35 | -0.22 | -0.21 | 0.23 *** | 0.16 |
| Switzerland | -0.55 * | -0.59 | -3.65 ** | 0.70 | 0.08 | 0.11 |
| Turkey | -0.13 | | -0.83 | | | |
| United Kingdom | 0.92 ** | 1.20 *** | 0.30 ** | 0.80 *** | -0.02 | -0.12 |
| United States | 1.86 | 2.00 | -0.07 | 1.31 *** | -0.10 | -0.47 |
| | | | ***Short run (mean group)*** | | | |
| Error correction term (-1) | -0.26 | -0.39 | -0.25 | -0.53 * | -0.4 | -0.56 |
| Adjusted R-squared long run | 0.994 | 0.995 | 0.993 | 0.995 | 0.995 | 0.996 |
| Adjusted R-squared short run | 0.4 | 0.42 | 0.4 | 0.45 | 0.46 | 0.47 |
| F-test | 5.18 | 4.34 | 5.38 | 5.5 | 5.67 | 5.96 |
| Durbin Watson statistic | 1.47 | 1.68 | 1.55 | 1.74 | 1.75 | 1.82 |
| Number of observations | 849 | 615 | 845 | 666 | 600 | 529 |

Note: The top panel gives the mean-group coefficients for the long run as well as the country-specific long-run coefficients for the electricity variable; the intermediate panel gives the coefficients for the short-run error correction term; the bottom panel gives regression diagnostics; ***, **, * denote the 1%, 5% and 10% level of significance, respectively; heterogenous coefficients were used as the Wald test on homogenous coefficients was rejected for each regressor variable individually and for all regressors jointly.

The coefficient of the infrastructure stock should be interpreted as the effect in addition to the effect of just adding to the productive capital stock. In this sense, a positive (negative) coefficient implies that the impact on output would be higher (lower).

**Table 2. Time series estimation results: non-transport infrastructure**

| | Electricity | | Telephone mainlines | | Telephone subscriptions | |
|---|---|---|---|---|---|---|
| | Coefficient | Coefficient | Coefficient | Coefficient | Coefficient | Coefficient |
| | | | ***Long run (mean-group)*** | | | |
| Investment | 0.39 *** | 0.39 *** | 0.39 *** | 0.42 *** | 0.45 *** | 0.34 ** |
| Population growth | 0.004 | -0.006 | 0.021 | 0.003 | 0.024 | 0.009 |
| Human capital | | 0.08 | | 0.15 | | 0.68 |
| trend | 0.02 *** | -0.11 | 0.02 *** | -0.13 | 0.02 *** | -0.07 |
| | | | *Country specific coefficients for infrastructure* | | | |
| Australia | -0.04 | -0.23 ** | -0.46 *** | -0.43 *** | 0.26 | 0.41 *** |
| Austria | 0.40 *** | 0.24 *** | 0.39 *** | 0.21 *** | 0.71 *** | 0.18 |
| Belgium | 0.54 *** | 0.22 *** | 0.37 ** | -0.08 | -0.70 *** | -0.24 * |
| Canada | 0.04 | -0.08 | 0.02 | 0.01 | 0.02 | 0.56 |
| Denmark | 0.26 *** | 0.36 ** | 0.21 * | -0.14 | 0.25 | 0.22 |
| Finland | 0.00 | -0.04 | -0.03 | -0.03 | 1.03 | 0.98 |
| France | 0.31 *** | 0.15 ** | 0.10 *** | 0.01 | -0.26 *** | -0.11 * |
| Greece | 0.31 *** | 0.38 *** | 0.32 *** | 0.34 *** | 0.16 *** | 0.28 *** |
| Iceland | 0.25 *** | | -0.60 *** | | 0.29 *** | |
| Ireland | -0.40 *** | -0.40 *** | -0.56 *** | -0.19 | -1.19 *** | -0.05 |
| Italy | 1.15 *** | 1.13 *** | 0.42 *** | 0.32 *** | -0.14 *** | -0.11 |
| Japan | 0.54 *** | 0.40 ** | 0.33 *** | 0.12 | -0.25 *** | -0.13 *** |
| Korea | -0.23 *** | | 0.02 | | 0.03 | |
| Mexico | 0.58 *** | | 0.68 *** | | 0.87 *** | |
| Netherlands | 0.25 *** | 0.21 | -0.12 * | 0.00 | -0.31 *** | -0.75 *** |
| New Zealand | -0.28 *** | -0.29 ** | -0.80 *** | -1.06 *** | 0.18 *** | 0.11 *** |
| Norway | 0.14 *** | 0.34 | 0.10 | 0.13 ** | -0.19 | -0.34 * |
| Portugal | 0.26 *** | -0.04 | 0.31 *** | 0.07 | -0.30 *** | -0.26 *** |
| Spain | 0.35 *** | 0.37 *** | 0.19 | 0.64 *** | -0.57 *** | -0.75 *** |
| Sweden | 0.03 | -0.01 | -0.11 | -0.02 | -0.01 | 0.14 |
| Switzerland | 0.08 | -0.16 | 0.13 | -0.32 | -0.04 | 0.03 |
| Turkey | 0.26 *** | | 0.08 | | 0.28 ** | |
| United Kingdom | 0.09 | 0.49 *** | -0.21 *** | -0.29 *** | -0.39 *** | 0.64 *** |
| United States | -0.08 * | -0.18 | 0.55 * | 0.24 | 0.31 ** | 0.47 |
| | | | ***Short run (mean group)*** | | | |
| Error correction term (-1) | -0.24 | -0.41 | -0.24 | -0.49 | -0.35 | -0.58 |
| | | | | | | |
| Adjusted R-squared long run | 0.996 | 0.996 | 0.996 | 0.996 | 0.997 | 0.998 |
| Adjusted R-squared short run | 0.43 | 0.41 | 0.45 | 0.45 | 0.42 | 0.45 |
| F-test | 4.85 | 5.1 | 6.58 | 5.74 | 5.53 | 3.75 |
| Durbin Watson statistic | 1.55 | 1.69 | 1.63 | 1.7 | 1.5 | 1.64 |
| Number of observations | 961 | 700 | 958 | 697 | 912 | 669 |

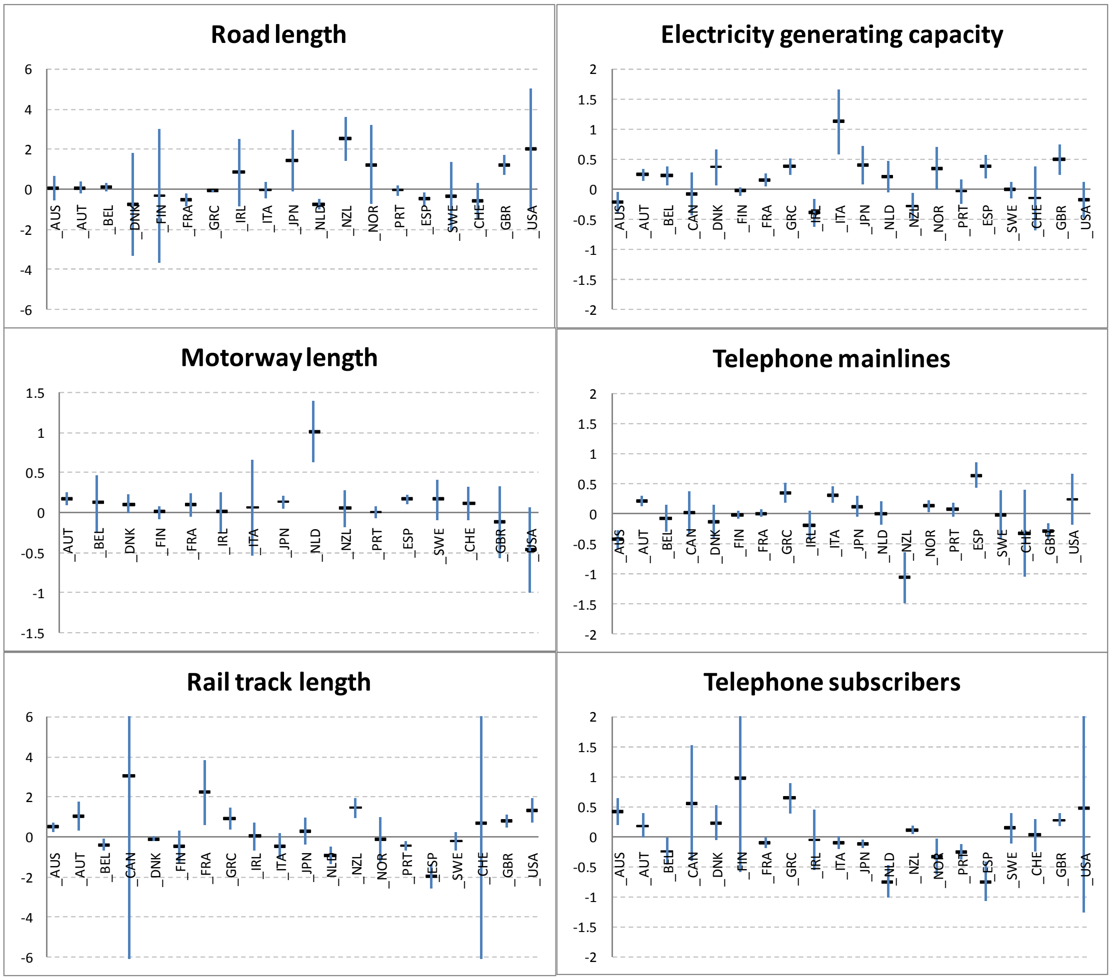
Note: The top panel gives the mean-group coefficients for the long run as well as the country-specific long-run coefficients for the electricity variable; the intermediate panel gives the coefficients for the short-run error correction term; the bottom panel gives regression diagnostics; ***, **, * denote the 1%, 5% and 10% level of significance, respectively; heterogenous coefficients were used as the Wald test on homogenous coefficients was rejected for each regressor variable individually and for all regressors jointly.

The coefficient of the infrastructure stock should be interpreted as the effect in addition to the effect of just adding to the productive capital stock. In this sense, a positive (negative) coefficient implies that the impact on output would be higher (lower).

While there does not appear to be a common effect of infrastructure on output and growth there are significant country-specific effects. Figure 1 displays the coefficients on the infrastructure stock variable for each sector graphed together with 90% confidence intervals (which graphically demonstrates why the assumption of a homogenous effect was rejected). They are estimated for the whole sample using the basic specification, total capital investment and including a human capital stock variable. The different specifications used in the robustness check are presented in Figure 2. Shares of positive significant coefficients are presented above the axis, while shares of negative significant coefficients are presented below the axis.[22]

**Figure 1. Infrastructure coefficient estimates from growth regressions**

Coefficient estimate and 90% confidence intervals



Road length

Electricity generating capacity

Motorway length

Telephone mainlines

Rail track length

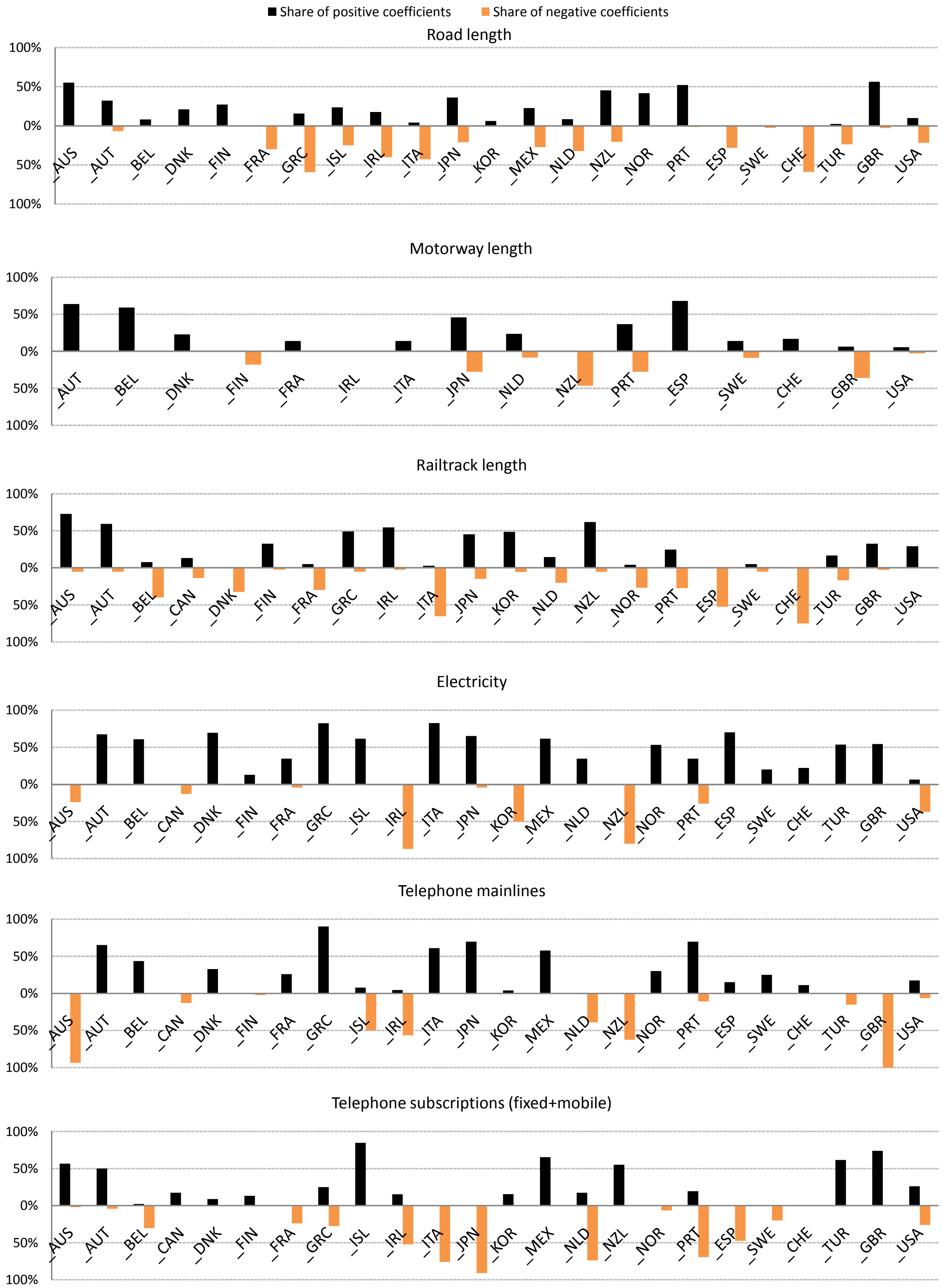
Telephone subscribers

---

[22] Due to different availability of variables for each country the shares are not calculated over the same amount of regressions. They are calculated over all available specifications, which vary from 20 to 48.

**Figure 2. Robustness check for time-series growth regressions**

share of significant coefficients



### *Road*

In case of the transport sector, a strong positive influence of length of roads per capita on GDP per capita levels and short term growth can be identified for the United Kingdom and New Zealand, being relatively robust. The alternative specifications show a possible similar effect for Australia and Portugal. If length of motorways is used, the effect is positive in Austria, Spain, Japan and the Netherlands. Interestingly, only in the first two is the effect robust to the inclusion of controls, while the effect for Belgium is generally positive, albeit not in the basic specification.

A negative effect of roads can be seen for France, Greece, Spain and the Netherlands, of which the latter two, as mentioned had shown a positive effect of motorways. It can be noted that in none of the countries is a robust negative effect of motorways found, though the United Kingdom and Japan show some weak signs of potential overinvestment (assuming the relationship is non linear). In general positive estimates for the coefficients on infrastructure can be crudely interpreted as investment in kilometres of roads being more productive than other types of investment, as roads are already included in total capital. Therefore, the results suggest that throughout our sample in the United Kingdom and New Zealand an increase in road length was associated with higher output and growth than other capital investment while it was the contrary in Greece, France, Spain and the Netherlands, where other types of investment have shown to be more productive.

In a number of countries the coefficients on roads tend to be negative or insignificant while the respective coefficients on motorways are positive. This may reflect the fact that the road networks were relatively well developed through most part of the sample, and additional roads were associated with fewer benefits in terms of output levels than the expansion of the motorway networks. Moreover, the objectives behind building roads may to a higher extent reflect social considerations rather than the large scale, fast transport routes, or international connections which are more often done *via* motorways.

### *Rail track*

With respect to rail tracks, positive significant effects on output levels and growth can be found in Australia, Austria, Ireland (in most specifications), Greece, New Zealand, the United Kingdom, France, and the United States, suggesting that in these countries investment in rail tracks was associated with higher output levels. Only in the case of the first four this seems robust to different specifications. A negative significant effect is found in the Benelux countries, and on the Iberian Peninsula, which may indicate some type of over-investment. In a large number of specifications this seems also true for Switzerland and Italy. It must be noted that in many OECD countries the length of rail tracks per capita was actually decreasing throughout the sample, thereby associated with an increase in output per capita in the two groups of countries.

### *Electricity generation*

In case of electricity generating capacity, the effects are positive and significant in Austria, Belgium, Denmark, Spain, Greece, United Kingdom, France, Iceland, Japan, Mexico, Turkey and the Netherlands, while being negative in Australia, Ireland and New Zealand. With the exception of Australia, these findings are robust, and the mean-group coefficient is positive and significant.

### *Telecommunications*

The increase in telephone mainlines per capita was associated with higher output levels and growth in Austria, Spain, Greece, Italy, Norway, Mexico and Iceland. On the other hand, it was associated with a slowdown in growth and lower levels of output for Australia, United Kingdom, Ireland and New Zealand, where the relationship is very robust. However, as noted above, technological developments may obscure the relationship between telecommunications infrastructure and growth.[23]A simple but rough way of accounting for this is to look at estimates for the variable of telephone subscribers (fixed and mobile), for which the effect turns positive for Australia, United Kingdom and New Zealand becoming insignificant for Ireland.

As the estimates are drawn from a period which covers the 1960s till 2005, the negative or positive contributions of infrastructure stock to output levels and growth fail to give an indication of whether recent investment conforms to the experience over the longer period. In other words, an indication of overinvestment (that is an estimated negative contribution around the deterministic trend) or underinvestment (a positive contribution) may be strongly driven by events in the early part of the sample. In order to account for this, at least partly, specifications allowing for non-linear effects at the end of the period have been estimated. This has been done in line with the robustness tests (*i.e.* across different specifications) with an additional inclusion of an interaction variable which is equal to the infrastructure stock in respectively the last 11, 7 and 5 years of the sample, and equal to zero before.[24]

---

[23] For example, the negative correlation may be related to the increase in mainlines slowing and then falling while growth was accelerating.

[24] Three types of time intervals are chosen in order to avoid the influence of cyclical factors on the estimates – they encompass commonly assumed lengths of business cycle duration (*i.e.* 5 to 11 years).

Table 3 shows that in case of most countries no significant separate effect at the end of the sample can be identified. However, in a number of countries the effect became stronger, suggesting for example that further increases in electricity generation capacity can be related to a decrease in output in Australia and Austria, similarly to motorways in Austria, New Zealand and Switzerland and rail tracks in Ireland and the Netherlands, whereas increases in road capacity may be associated with an increase in output in Greece, Ireland and the United Kingdom and additional electricity generation capacity in Portugal may support growth.

**Table 3. End of sample effects of infrastructure variables**

| | | positive | negative |
|---|---|---|---|
| Roads | p.c. | - | - |
| | *density* | *GRC(0), IRL(0), GBR(0)* | *NZL(-)* |
| Motorways | p.c. | - | AUT(+), NZL(0), CHE(0) |
| | *density* | - | - |
| Rail | p.c. | - | IRL(0), NLD(-) |
| | *density* | *GRC(0)* | *IRL(0)* |
| Electricity | p.c. | PRT(0) | AUS(-), AUT(+) |
| | *density* | *PRT(0)* | *AUS(-), AUT(+)* |
| Telephone mainlines | p.c. | - | AUT(+), ISL(0), IRL(0) |
| | *density* | *IRL(-)* | - |
| Telephone subscribers | p.c. | - | AUT(+), GRC(+) |
| | *density* | - | *AUT(+), GRC(+)* |

Reported only if significant in more than two thirds of the specifications (different types of capital, different control variables, end of sample 11, 7 and 5 years). – indicates that no countries met this criterion.

The sign of the effect estimated over the entire sample is shown in brackets (0 if insignificant).

Much of the empirical work focusing on a broader set of countries (see for example Canning and Bennathan, 2000, Canning and Pedroni, 1999) also fails to capture a common effect of infrastructure, even though dealing with a larger sample and inevitably more variation in GDP and infrastructure levels. Hulten (1996) fails to find any significant common effect of infrastructure stocks (telephone mainlines, rail, paved roads and electricity generation capacity) on growth in a cross-section study derived from the MRW approach. Nevertheless the results suggest that quality and efficiency of use may matter. Moreover, non-linear influences of infrastructure may also be at work – however, most of these effects can only be analysed at regional, industry or company level.

The negative estimated coefficients on infrastructure confirm that in the case of OECD countries more infrastructure is not always "better".[25] The negative estimated coefficients may be signalling over-investment as increases in infrastructure are associated with lower increases in GDP.[26] Over-investment in terms of quantity or perhaps high costs (admittedly not captured here directly) associated with poor investment decisions, with inadequate location or the high costs of expanding already developed networks can actually negatively influence GDP per capita levels. The finding is relatively robust to different specifications and underlines the importance of proper cost-benefit analysis of infrastructure investments as well as the importance of the fact that with limited resources infrastructure investments may not be the most productive.

# 6. Cross-country growth regressions

The cross-country approach uses multi-year averages with a limited number of observations in the panel. This approach requires a different interpretation from the annual panel data approach: cross-sectional data help explain why economic growth differs across countries over long periods of time. The same argument applies to the role of infrastructure investment. This section explains cross-country heterogeneity in economic growth that is attributable to differences in the level of physical infrastructure (rather than the effect due to the evolution of physical infrastructure within a given country).

## 6.1 Variable selection and data issues

The starting point of the analysis is to identify the set of explanatory variables that will be used in the growth regressions from the large number of possible drivers of economic growth.[27] In practice, very similar data to those used in the annual panels are used in this analysis. Furthermore, the variables used are limited to those that are proven to be robust in previous research (Sala-i-Martin *et al.* 2004; Doppelhofer and Weeks, 2008). In some cases these variables are modified to capture differences

---

[25] In this context "better" is defined as higher GDP per capita and higher GDP growth. While this seems a straightforward measure of well being, there may be others as well. Fore example, social considerations for the building of infrastructure associated with universal service.

[26] A negative coefficient can also signal indivisibility in investment and small increases in infrastructure do not tend to influence output positively, while a major investment may yield substantial network effects and push the economy into a new equilibrium. This issue is explored in a slightly different setting in Section 6.

[27] The previous literature has used a large number of variables that can be excluded from the analysis of recent OECD performance, such as dummy variables capturing wars, political instability, the spread of malaria, Sub-Saharan Africa, colonial past or the main religion.

across OECD countries (such as human capital and openness variables, see below)[28] and variables that are not available for large cross-sections of countries, but are available for OECD economies, are included, (such as the OECD indicators of product market regulation). The possible explanatory variables considered are:

- Log initial GDP per capita level
- Human capital measured in terms of educational achievements in the population over 15 years
- The growth rate of investment prices (which is often used instead of investment rate in cross-country growth regressions)
- Government investment as a share of GDP
- Life expectancy (at birth)
- Openness measured in terms of goods and services relative to GDP
- Growth rate of labour force
- Growth rate of the consumer price index
- The degree of regulation (measured by the OECD indicators of product market regulation)

The measures of infrastructure include the physical measures used in the previous section (railtrack, motorways, line subscriptions and electricity generation capacity). In some cases, a few modifications are made to take into account variations across countries.

- For transport networks, adequacy can depend on population density and geographical features. In this light, the railway and motorway variables are constructed to take into account per capita provision while accounting for geographical network density with a correction to account for areas where one would expect network density to be lower.[29]

---

[28] Primary or secondary school education and the number of years of openness used in the literature have a large cross-sectional variation for wide country cross-sections but exhibit fairly low cross-country heterogeneity for OECD countries.

[29] The railway and motorway variables are constructed as the total length of the physical network per square km, per population of working age (15 to 64 years of age), multiplied by the share of forests in the country's total area. The forest correction factor accounts for the fact that areas covered by forest may be inhabited.

- The telecommunication indicator measures the number of fixed line and mobile phone subscriptions per capita and thus takes account of the decline in fixed line subscription and the increasing popularity of mobile phones observed from the mid-1990s. However, line subscriptions is not strictly a physical measure of telecommunications networks
- The physical indicators for energy include overall energy generation and electricity capacity as well as consumption which may give an indication of the quality and the size of the transmission and distribution networks. In this light, there are four alternative variables:
  - Energy consumption per capita that considers not only domestic generation but also energy imports;
  - Energy generation per capita;
  - Electricity generation;
  - Total electricity generation capacity.

Given that the analysis uses a limited number of cross-sections observations, the measures of infrastructure were constructed to exploit the variation of these variables across countries. The principal components technique was used to construct one or two principal components that combine linearly the variance of four series (rail, motorways, telecom and energy). Four sets of principal components were computed using the four different measures of energy and electricity consumption or generation. The first two principal components account for around 80% of the total variance of the data. The first component is dominated by the measures of railway and motorway networks, whereas the second component largely reflects variance in the energy and telecommunications variables (Appendix Table A.7). The combination coefficients or factor loadings show that an increase in the first and second principal component coincides with a rise in the underlying measures of physical infrastructure.

### *Sample and country coverage*

2. Country coverage covers all the OECD with the exception of Luxembourg. For the cross-sections, 10-year and 8-year averages are constructed. The sample with 10-year average include at most three observations per country (1977-1986, 1987-1996, 1997-2006), whilst maximum four observations per country are available for the sample with 8-year averages (1975-1982, 1983-1990, 1991-1998, 1999-2006). The time span of the different variables is the same for any given country.

## 6.2 Model uncertainty and model averaging

With a number of candidate variables there is uncertainty about which variables should be included in the empirical growth model. The main techniques that have been used to account for this model uncertainty include extreme bounds analysis, classical model averaging and Bayesian model averaging. These approaches can be roughly categorised as testing the robustness of the candidate variable to the inclusion of other variables or examining how the inclusion of the candidate variable improves the explanatory power of growth models. The following sections briefly describe these techniques and present results.

### *6.2.1 Extreme bounds analyses (EBA)*

The extreme bounds approach of Levin and Renelt (1992) seeks to identify the sensitivity of the sign of a given variable to the inclusion of a number of other potential regressors. In this approach, the growth rate of per capita income ($Y$) is regressed on the variable of interest ($I$), a set of control variables ($\overline{C}$) that are always included and a set of other explanatory variables drawn from a larger set of potential explanatory variables ($\overline{Z}$):[30]

$$Y = \alpha \cdot I + \sum_{i=1}^{m=3} \beta_i \cdot C_i + \sum_{j=1}^{n=3} \gamma_j \cdot Z_j \text{ , with } Z_1, Z_2, Z_3 \subset \overline{Z} \qquad (6)$$

The lower bound is obtained as the lowest value of $\alpha$ in equation (6) minus two standard errors, and the highest values is calculated as the highest value of $\alpha$ plus two standard errors. If the lower and upper bounds are on the same side (either both negative or both positive), the variable of interest is labelled as being robust to the inclusion of $\overline{C}$ and $\overline{Z}$. However, if a single coefficient estimate of $\alpha$ becomes statistically insignificant at the 5% level or switches sign, the lower and upper bounds will have the opposite sign, and the variable of interest will be called fragile, conditional on $\overline{C}$ and $\overline{Z}$.

In our empirical analysis, we slightly modify the procedure described above. The estimated equations always include one principal component capturing cross country differences in physical infrastructure. Five additional explanatory variables are selected from the pool of possible explanatory variables $\overline{Z}$. For each principal component, all possible 5-variable combinations from

---

[30] In the original application, Levin and Renelt (1992) used three fixed control variables (m=3) and up to 3 variables from the set of potential regressors $\overline{Z}$. With $I$ and $\overline{C}$ being always included, regression (1) is estimated for all possible combinations of $\overline{Z}$ up to n=3..

the 9 possible (non-infrastructure) explanatory variables were estimated. The lower and upper bounds are then calculated in order to analyse whether the measures of physical infrastructure are fragile or robust to the inclusion of different explanatory variables. In line with previous applications to cross-section growth equations, the extreme bounds analysis shows that none of the infrastructures measures is robust to the inclusion of the controls as the lower and upper bounds always have the opposite sign (Table 4).[31] Indeed, Sala -i-Martin (1997) demonstrated that none of the conventional explanatory variables used in empirical growth equations would pass the test of the extreme bounds analysis.

**Table 4. Extreme-bound Analysis results**

| | 8-year averages | | 10-year averages | |
|---|---|---|---|---|
| | low bound | high bound | low bound | high bound |
| ***First - mainly transport - principal component*** | | | | |
| energy use (PC11) | -5.40 | 3.72 | -6.72 | 9.96 |
| energy generation (PC12) | -4.93 | 2.61 | -6.01 | 6.44 |
| electricity generation (PC13) | -4.26 | 3.43 | -5.07 | 8.46 |
| electricity generation capacity (PC14) | -4.84 | 3.74 | -5.80 | 9.92 |
| ***Second - mainly energy and telecommunications - principal component*** | | | | |
| energy use (PC21) | -2.75 | 3.76 | -8.99 | 10.47 |
| energy generation (PC22) | -1.44 | 1.16 | -1.20 | 3.14 |
| electricity generation (PC23) | -2.97 | 3.24 | -4.39 | 8.63 |
| electricity generation capacity (PC24) | -2.54 | 3.05 | -2.49 | 7.09 |

Note: The table reports the lower and upper coefficient estimates for the different variants of the principal components.

### *6.2.2 Model averaging à la Sala-i-Martin (1997)*

Sala-i-Martin (1997) proposed an alternative approach for assessing variable robustness that constructs averages of the individual coefficient estimates and standard errors. Relying on equation (1), this approach computes weighted averages of the mean and variance of $\alpha$. The weights are calculated as the likelihood of a given model related to the summed likelihood of all estimated models. However, as some mis-specified models, such as those subject to endogeneity, will fit the data better and may thus out-perform other models, Sala-i-Martin (1997) also used unweighted averages. Furthermore, significance is assessed assuming that the estimated coefficients are either normally or not normally distributed.

---

[31] This finding is not sensitive to possible outliers. When the countries are dropped from the sample one at a time, the same results are obtained.

The results of the model averaging are presented in Table 5, which shows the weighted and unweighted means, and the corresponding p-values under the assumptions of both normality and non-normality of the coefficient estimates. The results do not support the inclusion of the principal component measures of infrastructure when 8-year averages are used. For the 10-year averages, however, a positive relationship is found in some cases between the per capita growth rate and the second principal components (pc21, pc23 and pc24). These results are not robust (particularly for pc21 and pc23) given that they are sensitive to the sample used and no relationship could be established using likelihood-based weights while assuming a normal distribution.[32]

**Table 5. Model averaging : à la Sala-i-Martin (1997)**

Panel A. 8 year averages

| | | coefficient | p-value | | coefficient | p-value | |
|---|---|---|---|---|---|---|---|
| | sample | Weighted | normal | non-normal | Unweighted | normal | non-normal |
| ***First - mainly transport - principal component*** | | | | | | | |
| energy use (PC11) | all | -1.297 | 0.238 | 0.212 | -1.212 | 0.269 | 0.242 |
| energy generation (PC12) | all | -0.95 | 0.359 | 0.373 | -0.902 | 0.377 | 0.393 |
| electricity generation (PC13) | all | -0.919 | 0.352 | 0.307 | -0.839 | 0.395 | 0.348 |
| electricity generation capacity (PC14) | all | -1.153 | 0.283 | 0.25 | -1.068 | 0.32 | 0.285 |
| ***Second - mainly energy and telecommunications - principal component*** | | | | | | | |
| energy use (PC21) | all | 0.449 | 0.603 | 0.461 | 0.581 | 0.491 | 0.352 |
| energy generation (PC22) | all | -0.105 | 0.745 | 0.713 | -0.08 | 0.805 | 0.771 |
| electricity generation (PC23) | all | 0.113 | 0.883 | 0.82 | 0.202 | 0.791 | 0.726 |
| electricity generation capacity (PC24) | all | 0.195 | 0.785 | 0.721 | 0.277 | 0.695 | 0.628 |

Panel B. 10 year averages

| | | coefficient | p-value | | coefficient | p-value | |
|---|---|---|---|---|---|---|---|
| | sample | Weighted | normal | non-normal | Unweighted | normal | non-normal |
| ***First - mainly transport - principal component*** | | | | | | | |
| energy use (PC11) | all | 0.688 | 0.738 | 0.744 | 0.885 | 0.663 | 0.661 |
| energy generation (PC12) | all | -0.263 | 0.88 | 0.896 | -0.203 | 0.905 | 0.923 |
| electricity generation (PC13) | all | 1.132 | 0.493 | 0.494 | 1.341 | 0.412 | 0.405 |
| electricity generation capacity (PC14) | all | 1.168 | 0.547 | 0.552 | 1.386 | 0.47 | 0.467 |
| ***Second - mainly energy and telecommunications - principal component*** | | | | | | | |
| energy use (PC21) | excl _BEL | 1.404 | 0.505 | 0.216 | 2.274 | 0.249 | **0.049** |
| | excl _FRA | 1.393 | 0.503 | 0.199 | 2.299 | 0.237 | **0.038** |
| | excl _IRL | 1.603 | 0.437 | 0.172 | 2.496 | 0.195 | **0.036** |
| | excl _KOR | 2.067 | 0.29 | 0.136 | 2.663 | 0.157 | **0.049** |
| | excl _MEX | 2.895 | 0.124 | 0.042 | 3.617 | **0.041** | **0.007** |
| | excl _NZL | 1.465 | 0.466 | 0.161 | 2.407 | 0.196 | **0.025** |
| | excl _NOR | 1.463 | 0.541 | 0.199 | 2.343 | 0.293 | **0.046** |
| | excl _GBR | 1.421 | 0.519 | 0.214 | 2.317 | 0.259 | **0.045** |
| energy generation (PC22) | all | 0.373 | 0.5 | 0.54 | 0.411 | 0.459 | 0.486 |
| electricity generation (PC23) | all | 1.827 | 0.249 | 0.235 | 2.129 | 0.176 | 0.162 |
| | excl _NOR | 2.711 | 0.208 | 0.1 | 3.305 | 0.102 | **0.036** |
| electricity generation capacity (PC24) | all | 2.045 | 0.128 | 0.072 | 2.435 | 0.056 | **0.025** |

Note: Figures in bold indicate statistical significance at the 5% level.

---

[32] If outliers are accounted for, the unweighted p-values with a non-normal distribution (the ratio of the coefficient estimate over the standard errors) are lower than 0.05 and the unweighted coefficients exhibit a positive sign.

### *6.2.3 Bayesian averaging of classical estimates (BACE)*

A more recent approach to addressing model uncertainty is to assess whether the inclusion of a candidate variable improves the fit of the model (Sala-i-Martin *et al*. 2004). This approach estimates all possible combinations of the (K) candidate explanatory variables, which is given by $2^K$, or some subset of models.[33] Given the relatively low number of potential explanatory variables used here, all possible models are estimated.

The Bayesian averaging of classical estimates (BACE) determines the posterior probability attributed to each single model $M_j$ that includes the variable of interest and conditioned on the underlying dataset ($P(M_j|y)$).

$$P(M_j|y) = \frac{P(M_j)T^{-k_j/2}SSE_j^{-T/2}}{\sum_{i=1}^{2^K} P(M_i)T^{-k_i/2}SSE_i^{-T/2}} \qquad (7)$$

where SSE is the sum of squared residuals, T is the number of observations, k denotes the number of explanatory variables included in the specific model and K is the number of all explanatory variables considered. Expression (7) gives the contribution of a given model to explaining the dependent variable as compared to the other models. Expression (7) is then summed up for the models that contain the variable of interest to obtain the posterior inclusion probability of this variable. If the posterior inclusion probability is higher than the prior inclusion probability, one can conclude that the candidate variable should be included in the estimated models.[34]

The posterior mean and the square root of the variance (standard error) conditional on inclusion can be used to obtain t-statistics and to determine the significance of the individual variables upon inclusion. The posterior mean conditional on inclusion ($E(\beta|y)$) is the average of the individual OLS estimates weighted by $P(M_j|y)$. As the unconditional posterior mean considers all regressions (even those without the variable of interest), the unconditional posterior mean of any given variable

---

[33] If the number of models to be estimated is too large, techniques such as Markov-Chain Monte-Carlo, stochastic search variable selection, or random sampling are alternative approaches to estimating all possible models.

[34] Sala-i-Martin *et al.* (2004) compare the posterior inclusion probability to a prior inclusion probability for their 67 explanatory variables in 7 variable models. The prior inclusion probability is then 7/67=0.1044.

can be derived as the product of the conditional posterior mean and the posterior inclusion probability. The posterior variance of $\beta$ ($Var(\beta|y)$) can be calculated as follows:

$$Var(\beta|y)=\sum_{j=1}^{2^K}P(M_j|y)Var(\beta|y,M_j)+\sum_{j=1}^{2^K}P(M_j|y)(\hat{\beta}_j-E(\beta|y))^2 \tag{8}$$

In addition, White's heteroscedasticity-corrected standard errors are used in all estimations, which include not only the full sample but also sub samples which exclude one country at a time. This jack-knifing of the sample makes it possible to evaluate the impact of individual countries on the robustness of the results and to identify potential outliers. All explanatory variables were used with a one period lag (8 to 10 years) in order to minimise potential problems with endogeneity that may potentially affect most of the explanatory variables.[35]

The random sampling procedures employed in previous applications of model averaging have often experienced difficulties by duplicating the estimation of particular models because they fails to distinguish between identical models.[36] When the recurrence of different orderings of the same variables is not controlled for, good models including more variables receive a considerably higher weight than similarly performing models including fewer variables. The approach adopted here eliminates the bias towards larger models by ensuring that each model is only estimated once.

### *Nonlinear extensions*

An extension to the basic approach is to assess possible nonlinear links between infrastructure and economic growth. This possibility can be addressed using threshold models proposed by Hansen (1999). These take the following form:

$$Y=\begin{cases}\alpha+\sum_{j=1}^{n-1}\beta_j\cdot X_j+\varphi_1\cdot X_n+\varepsilon & if \quad T\le\rho\\ \alpha+\sum_{j=1}^{n-1}\beta_j\cdot X_j+\varphi_2\cdot X_n+\varepsilon & if \quad T>\rho\end{cases} \tag{9a}$$

where $\rho$ is the threshold variable and T denotes the threshold value that separates the two regimes. This type of model can be easily extended to three or even more regimes such as in equation (4b):

---

[35] The (first difference and system) GMM estimator would necessitate at least three observations per country. There are some countries in the sample that have less than three observations for the 8- and 10-year averages. Therefore, 5-year averages were used for the GMM estimator for the specification including all variables comparing with fixed effect OLS estimates to check whether or not endogeniety is a real concern.

[36] Such as, Y=a+bX1+cX2+dX3 or Y=a+ bX2+cX1+ dX3 or Y=a+ bX3+cX2+dX1

$$Y = \begin{cases} \alpha + \sum_{j=1}^{n-1} \beta_j \cdot X_j + \varphi_1 \cdot X_n + \varepsilon & if \quad T_1 \leq \rho \\ \alpha + \sum_{j=1}^{n-1} \beta_j \cdot X_j + \varphi_2 \cdot X_n + \varepsilon & if \quad T_2 \geq \rho > T_1 \\ \alpha + \sum_{j=1}^{n-1} \beta_j \cdot X_j + \varphi_3 \cdot X_n + \varepsilon & if \quad \rho > T_2 \end{cases} \tag{9b}$$

Equation (4a) integrated into a BACE framework is termed Bayesian Averaging of Thresholds (BAT) (Crespo-Cuaresma and Doppelhofer, 2007). The BAT technique relies on a random sampling procedure for the variables of interest as well as a random sampling of the threshold variable in order to estimate the non-linear model instead of the linear model. The BAT approach reveals the inclusion probability for the nonlinear explanatory variable and the corresponding threshold values of the threshold variable, but it does not test whether $\varphi = \gamma$ - in equation (9) - and whether the nonlinear variant of the model is superior to the linear version.[37]

The approach used here to analyse non-linearity tests explicitly whether the linear variant of the model can be rejected in favour of the nonlinear variant. This is done in two approaches:

- The first approach selects the variables that pass the inclusion test from the linear model averaging, estimates the OLS model, and analyses nonlinearity within this model.
- The second approach, more in the spirit of model averaging, estimates all the possible combinations of the candidate explanatory variables. For each combination, the linear, two-regime and three-regime model for the nonlinear variable are estimated. An advantage of this methodology is that only a single linear or non-linear model is selected for a given set of explanatory variables.

The selection between linear and nonlinear models is done as follows. The first step estimates the linear model and the two-regime model. A grid search with steps of 1% of the distribution is carried out to find the value of the threshold variable (principal component measure of infrastructure) that minimizes the sum of squared residuals of the estimated two-regime model.[381] Hansen (1999) shows

---

37. Note that the prior inclusion probability for the nonlinear and threshold variables are considerably lower in the BAT modelling framework than in the standard linear modelling framework because the prior inclusion probability accounts not only for the number of potential variables but also for the grid search of the threshold variable. The prior inclusion probability decreases with the number of steps used for the grid search.

[38] While steps of 1% in the grid search seem to be sufficient given the sample size of about 80 observations, steps of 0.1% were also used. The estimated threshold that separates the two regimes did not change. The

that the null hypothesis of $\varphi_1 = \varphi_2$ from equations (9a) can be tested using a likelihood ratio test. Given that the likelihood ratio test statistic does not follow a standard asymptotic distribution as the threshold value is not identified under the null hypothesis, the distribution of the test statistic is obtained through bootstrapping (Hansen, 1999). The bootstrap procedure consists in the following steps:

1. The linear and nonlinear models that minimise the sum of squared residuals are estimated and the likelihood ratio test computed.
2. Repeated random draws with the probability of 1/n at each draw are made from the residuals of the alternative model to construct the bootstrapped residual.
3. The bootstrapped dependent variable is obtained using the bootstrapped residuals of the alternative model (two-regime model) and the coefficient estimates of the benchmark model (linear model).
4. The models that are tested against each other (linear versus two-regime, and two-regime *vs.* three-regime models) are re-estimated and the likelihood ratio tests re-calculated.
5. Steps 2 to 4 are repeated 1 000 and 500 times for respectively for the simple non-linear model and the non-linear models imbedded in model averaging, respectively. The likelihood ratio tests obtained on the basis of the bootstrapped sample are then saved.
6. The likelihood ratio test from the original sample is compared with the upper 90%, 95% or 99% of the distribution of the bootstrapped likelihood ratio tests. If the likelihood ratio test statistic rejects the null hypothesis of the linear model against the two-regime model (on the basis of the bootstrapped critical values), whether there are three different regimes rather than only two regimes is also analysed. A three-regime model is estimated based on two threshold values of the threshold variable that minimise the sum of squared residuals across the estimated models.[39] The bootstrap procedure described above is applied to the two-regime and three-regime models.

---

grid search starts at 25% of the distribution and stops at 75% to ensure that at least 20 observations fall in one particular regime.

[39] The threshold from the two-regime model is held fixed and a grid search is used to identify the second threshold. We impose the restriction that the two thresholds should be separated at least by 25% of our sample observations. Once the second threshold is identified, a backward grid search is done to identify the first threshold as suggested by Hansen (1999).

### *6.3.4 Empirical results: the case of linearity*

The full results (presented in Appendix Tables A.8 and A.9) reveal that important drivers of GDP per capita growth include, as expected, initial per capita income as well as openness, life expectancy and human capital. Government investment (a proxy for the tax burden on the economy) as well as investment price inflation as expected relate negatively with economic growth. These results are broadly in line with earlier findings (OECD, 2003; Sala-i-Martin *et al*. 2004).[40]

The results concerning the infrastructure variables of model averaging for the full sample are presented in Table 6. The posterior inclusion probability of the first principal component, which mainly measures railway and motorways infrastructure, is almost always higher than the prior inclusion probability of 0.5 (50%). In these models, the mean of the principal component is unstable and never statistically significant. The second principal component, which measures mainly energy and telecommunications infrastructure, is generally found to have high estimated inclusion probabilities, always exceeding 0.5 (50%) regardless of the averaging method. Furthermore, the mean is significant.

**Table 6. Full model averaging results**

linearity, full sample

| | 8-year averages | | | 10-year averages | | |
|---|---|---|---|---|---|---|
| | Inclusion probability | mean | standard error | Inclusion probability | mean | standard error |
| ***First - mainly transport - principal component*** | | | | | | |
| energy use (PC11) | 1 * | -0.02 | 1.15 | 0.5 * | 0.7 | 1.45 |
| energy generation (PC12) | 1 * | -0.29 | 1.02 | 0.36 | 0.04 | 0.54 |
| electricity generation (PC13) | 1 * | 0.23 | 1 | 0.73 * | 1.35 | 1.8 |
| electricity generation capacity (PC14) | 1 * | 0.12 | 1.12 | 0.68 * | 1.31 | 2.08 |
| ***Second - mainly energy and telecommunications - principal component*** | | | | | | |
| energy use (PC21) | 1 * | 1.69 | 0.61 | 1 * | 5.73 | 1.12 * |
| energy generation (PC22) | 1 * | 0.15 | 0.34 | 0.72 * | 0.44 | 0.37 |
| electricity generation (PC23) | 1 * | 0.96 | 0.68 | 1 * | 2.91 | 1.19 * |
| electricity generation capacity (PC24) | 1 * | 0.96 | 0.59 | 1 * | 3.07 | 0.83 * |

Note: inclusion probability is the posterior inclusion probability, the mean and standard error are the posteriors conditional on inclusion. * indicates that the estimated inclusion probability is higher than 0.50, and that the mean/s.e. ratio is higher than 2.

---

[40] Given the possibility of endogeneity in the relationship between infrastructure and growth, the (first difference) GMM estimator was used for the specification including all variables and using 5 year averages. These results are compared with fixed effect estimates. The results suggest that for this sample there is not a major problem of endogenity for the explanatory variables. The results are reported in the Appendix (Table A.10).

In order to check the robustness of the results for the infrastructure variables, the model averaging was also conducted for sub-samples that dropped one country at a time (The results are presented in Figures 3 and 4, see also Appendix Figures A.1 and A.2).

- *The mainly transportation (railways and motorways) principal component*: Figure 4 shows the inclusion probability is almost always above 0.5 and most frequently 1 when 8 year averages are used, but is often below 0.5 for some versions of the principal components when 10 year averages are used.[41]In addition, the distribution of the estimated means suggests that these results are sensitive to the inclusion of particular countries given that changes in the sample can switch the sign of the posterior mean. Overall, the relation between the first principal component and economic growth is not very robust.
- *The mainly energy and telecommunications principal component*: Figure 5 shows that the inclusion probability is nearly always 1 for almost all the variants of the second principal component. Furthermore, the estimated means of the infrastructure variables always have a positive sign both for the whole sample and all sub-samples.[42] Overall, variants of the second principal components (with one exception) appear to have a strong positive relationship to economic growth.

---

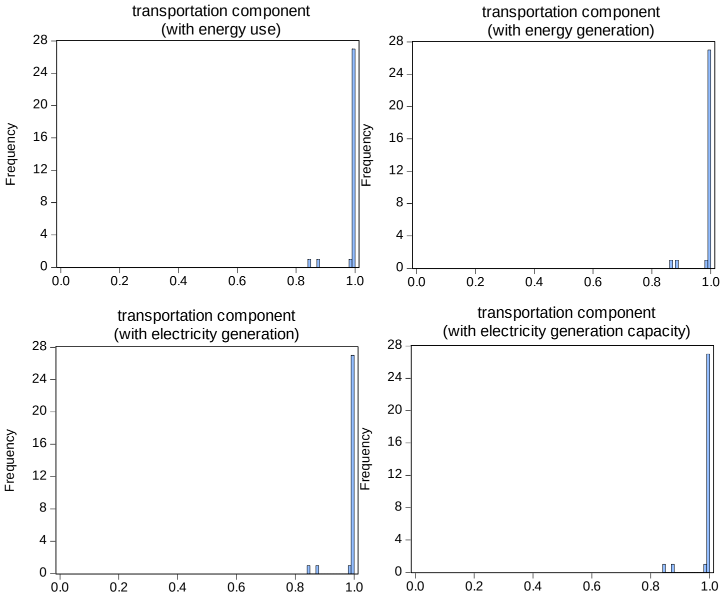
41. This is particularly the case when New Zealand and Germany are excluded from the sample. When Portugal is dropped from the sample, the inclusion probability for the only principal component that falls below 50% in the full sample jumps to around 80%.

[42]. Problems arise only for the second variant of the second principal component for 10-year averages when Germany, New Zealand and Portugal are excluded from the sample. In these cases, the estimated inclusion probabilities drop slightly below 50%.

**Figure 3. Inclusion probabilities and posterior means; transport**

8 year averages

Panel A. Distribution of inclusion probability

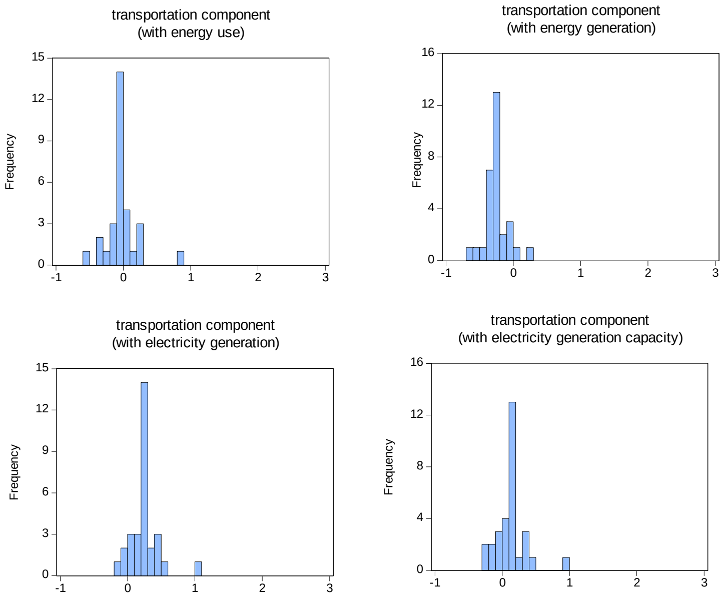

transportation component (with energy use)

transportation component (with energy generation)

transportation component (with electricity generation)

transportation component (with electricity generation capacity)

Panel B. Distribution of posterior mean conditional on inclusion


transportation component (with energy use)

transportation component (with energy generation)

transportation component (with electricity generation)

transportation component (with electricity generation capacity)

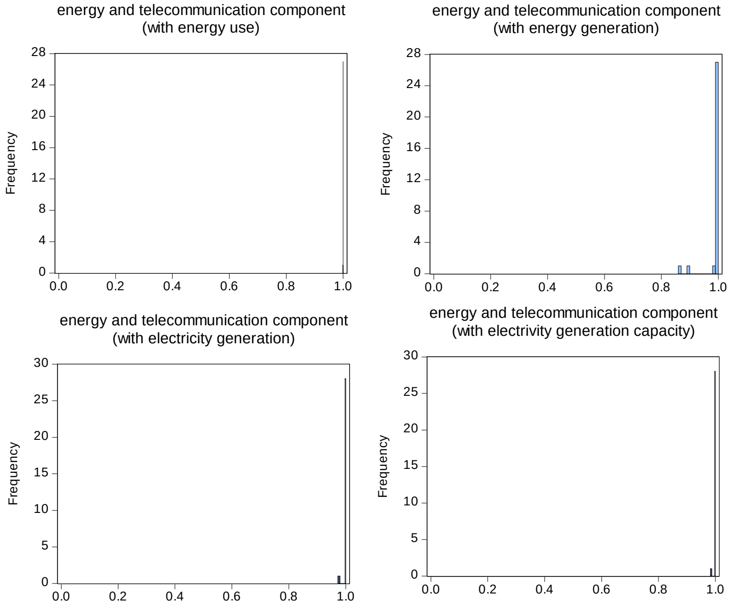
Note: Panel A is a histogram of the distribution of the estimates inclusion probabilities for all estimates (full sample and sub-samples that drop a single country from the sample). The vertical axis shows the number of models for which the values of the inclusion probability are shown on the horizontal axis are estimated.

Panel B is a histogram of the distribution of the posterior mean conditional on inclusion for all estimates (full sample and sub-samples that drop a single country from the sample). The vertical axis shows the number of models for which the values of the estimated mean are shown on the horizontal axis are estimated.

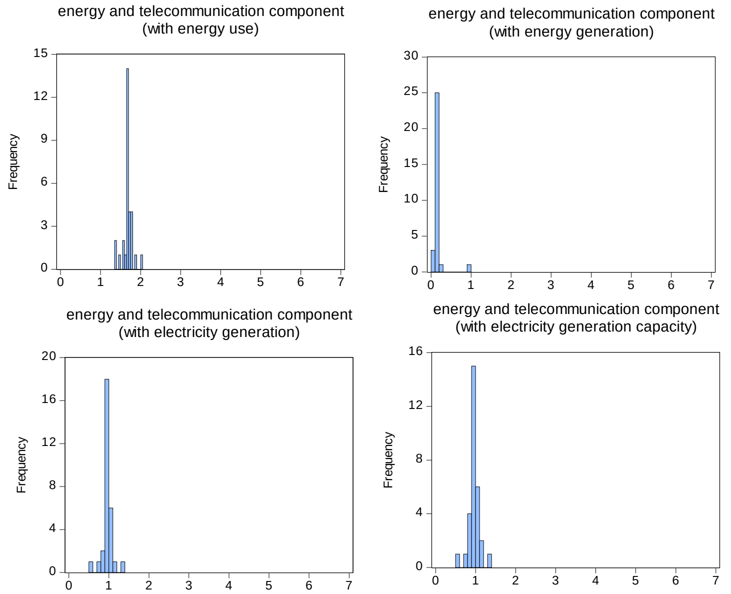
**Figure 4. Inclusion probabilities and posterior means; energy and telecoms**

8 year averages

Panel A. Distribution of inclusion probability



Panel B. Distribution of posterior mean conditional on inclusion



Note: Panel A is a histogram of the distribution of the estimates inclusion probabilities for all estimates (full sample and sub-samples that drop a single country from the sample). The vertical axis shows the number of models for which the values of the inclusion probability are shown on the horizontal axis are estimated.

Panel B is a histogram of the distribution of the posterior mean conditional on inclusion for all estimates (full sample and sub-samples that drop a single country from the sample). The vertical axis shows the number of models for which the values of the estimated mean are shown on the horizontal axis are estimated.

### 6.3.5 *Empirical results: the case of nonlinearity*

*Non-linearity in a simple framework*

To test for the non-linearity of the impact of the different measures of infrastructure on growth, the first approach uses the variables with posterior probabilities exceeding prior probabilities from the linear full model averaging.[43] The estimation results reported in Appendix Table A.11 provide little robust empirical evidence in favour of non-linearity for the various measures of the first principal component capturing railway and motorway networks and the coefficients are also generally found to be statistically insignificant.

The results for the second principal component, reflecting energy and telecommunication, provide stronger support for nonlinear effects of infrastructure on growth (Table 7, Annex Table A.12). The two-regime model is selected against the linear and the three-regime models for three of the four versions of the second principal component when 8-year averages or 10-year averages are used. The coefficients in the lower and upper regimes are statistically significant and have a positive sign for 4 (2) versions of the principal component for the 8-year (10-year) averages. Furthermore, the coefficients in the lower regime are considerable larger than those in the upper regime. This implies that an increase in physical infrastructure in energy and telecommunication has a considerable stronger impact on economic growth if the level of physical infrastructure is lower.

**Table 7. The nonlinear relationship between infrastructure and growth**

8-year averages

| | **2nd principal component (PC21) With energy use** | | | **2nd principal component (PC22) With energy generation** | | |
|---|---|---|---|---|---|---|
| | Linear model | 2-regime Model | 3-regime model | Linear model | 2-regime Model | 3-regime model |
| Low | 1.376** | 2.770*** | 2.437** | 0.140 | 0.892 | 1.321 |
| High/mid | | 1.578*** | 0.960 | | 0.227 | 0.624 |
| high | | | 1.666*** | | | 0.228 |
| p-value (bootstrap) | | 0.004 | 0.447 | | 0.209 | 0.549 |
| | **2nd principal component (PC23) With electricity generation** | | | **2nd principal component (PC24) With electricity generation capacity** | | |
| | Linear model | 2-regime Model | 3-regime model | Linear model | 2-regime model | 3-regime model |
| Low | 0.770 | 2.831** | 1.724* | 0.747 | 2.470*** | 2.176** |
| High/mid | | 1.365** | 1.011 | | 1.184** | 0.752 |
| high | | | 1.361** | | | 1.209** |
| p-value (bootstrap) | | | 0.077 | 0.721 | | 0.013 |

---

[43] The approach estimates the linear, two-regime and three-regime models using the Likelihood Ratio (LR) test statistic with the boostrapped critical values to test for nonlinearity.

Note: *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively. P-values lower than 0.1, 0.05 and 0.01 indicate that the null hypothesis of the linear model (2-regime model) can be rejected against the alternative hypothesis of the 2-regime model (3-regime model).

### *Non-linearity in a Bayesian model averaging framework*

The second approach to analysing nonlinearity incorporates the nonlinear models in the model averaging framework.[44] The results for the first principal component are mixed, with positive and negative effects of infrastructure on growth estimated for different regimes and for the various infrastructure measures (see Appendix Tables A.13-A.16). The results for the second principal component provided somewhat stronger evidence of nonlinear effects in particular when 8-year averages are used (Table 8). The two-regime model is selected for three of four variants of the principal component. The coefficient estimates of the infrastructure measures are always considerably larger in the lower regime than in the higher regime.

**Table 8. Results of model averaging – nonlinearity of the infrastructure variable**

8-year averages, 2nd principal component

| | **PC21**<br>**With energy use** | | | | **PC22**<br>**With energy generation** | | | |
|---|---|---|---|---|---|---|---|---|
| | Inclusion probability | | mean | s.e. | Inclusion probability | | mean | s.e. |
| Linear | 0.000 | | 0.000 | 0.000 | **0.586** | * | 0.085 | 0.241 |
| 2-regime | **1.000** | * | | | 0.412 | | | |
| lower regime | | | 2.633 | 4.501 | | | 0.643 | 0.557 |
| higher regime | | | 1.524 | 1.511 | | | 0.119 | 0.168 |
| 3-regime | 0.000 | | | | 0.000 | | | |
| low regime | | | 0.000 | 0.000 | | | 0.000 | 0.000 |
| middle regime | | | 0.000 | 0.000 | | | 0.000 | 0.000 |
| high regime | | | 0.000 | 0.000 | | | 0.000 | 0.000 |

| | **PC23**<br>**With electricity generation** | | | | **PC24**<br>**With electricity generation capacity** | | | |
|---|---|---|---|---|---|---|---|---|
| | Inclusion probability | | mean | s.e. | Inclusion probability | | mean | s.e. |
| Linear | 0.002 | | 0.002 | 0.002 | 0.000 | | 0.000 | 0.000 |
| 2-regime | **0.998** | * | | | **1.000** | * | | |
| lower regime | | | 2.705 | 3.750 | | | 2.596 | 4.088 |
| higher regime | | | 1.312 | 1.111 | | | 1.236 | 1.041 |
| 3-regime | 0.000 | | | | 0.000 | | | |
| low regime | | | 0.000 | 0.000 | | | 0.000 | 0.000 |
| middle regime | | | 0.000 | 0.000 | | | 0.000 | 0.000 |
| high regime | | | 0.000 | 0.000 | | | 0.000 | 0.000 |

Note: * indicates that the estimated inclusion probability is higher than 0.50

---

[44] The selection of the linear and nonlinear models is carried out at the 5% level of significance.

# 7. Conclusions

This paper analysed the relation between infrastructure investment and economic growth. We emphasised issues related to data quality limits. For instance, national account data for investment or capital stocks in infrastructure sectors are available in long series for only a handful of countries and still contain methodological differences. Measures of infrastructure quality are even more scarce and of poorer comparability. Overall, the most robust available measures of infrastructure for a sample of OECD countries over time are physical indicators. We pointed out that much of the literature seems to confuse infrastructure with public capital stocks or public investment, which, due to corporatisation, privatisation and market liberalisation are increasingly unreliable measures of infrastructure.

Keeping this in mind, our empirical results based on annual time-series growth regressions indicate that the contributions of infrastructure to long-run output levels and growth are not homogenous across countries and that the expansion of infrastructure could be both more or less productive with respect to other capital expenditure. Furthermore, the result that more does not always mean better (in terms of GDP per capita) seems to be robust across different specifications including control variables such as human capital, trade openness and tax revenues. Importantly, the validity of the full sample estimates holds for more recent years, and in most countries the effect does not seem to change.

We identified a robust positive and highly nonlinear link between infrastructure and economic growth using low-frequency multi-annual average using Bayesian averaging of classical estimates. The cross-section growth regressions suggests greater provision of broad measures of infrastructure is associated with higher subsequent growth rates and that the link is non-linear, with a potentially higher impact of additional infrastructure in countries with initially lower levels of provision.

# REFERENCES

Arellano, M. and O. Bover (1995), “Another Look at the Instrumental Variable Estimation of Error-Components Models”, *Journal of Econometrics,* 68(1), pp. 29-51.

Arnold, J., A. Bassanini, and S. Scarpetta (2007), “Solow or Lucas?: Testing Growth Models Using Panel Data from OECD Countries", *OECD Economics Department Working Papers* No. 592.

Aschauer, D.A. (1989), “Is Public Expenditure Productive?”, *Journal of Monetary Economics,* 23, pp. 177–200.

Bond, S., A. Leblecioglu and F. Schiantarelli (2004), "Capital Accumulation and Growth: A New Look at the Empirical Evidence", *Boston College Working Paper* No. 591.

Cadot, O., L.H. Röller and A. Stephan (1999), “A Political Economy Model of Infrastructure Allocation: An Empirical Assessment”, *CEPR Discussion Paper* No. 2336.

Cadot, O., L.H. Röller and A. Stephan (2006), “Contribution to Productivity or Pork-barrel? The Two Faces of Infrastructure Investment”, *Journal of Public Economics,* 90(6-7), pp. 1133–1153.

Calderón, C. and L. Servén (2002), “The Output Cost of Latin America's Infrastructure Gap”, *Central Bank of Chile Working Paper* No. 186.

Canning, D. (1999), “The Contribution of Infrastructure to Aggregate Output”, *World Bank Working Paper* No. 2246.

Canning, D. and E. Bennathan (2000) “The Social Rate of Return on Infrastructure Investments”, *World Bank Working Paper* No. 2390.

Canning, D. and P. Pedroni (2008), “Infrastructure, Long-run Economic Growth and Causality Tests for Cointegrated Panels”, *Manchester School*, 76(5), pp. 504-527.

Crespo-Cuaresma, J. and G. Doppelhofer (2007), “Nonlinearities in Cross-Country Growth Regressions: A Bayesian Averaging of Thresholds (BAT) Approach”, *Journal of Macroeconomics*, 29(3), 541-554.

Demetriades, P.O. and T.P. Mamuneas (2000), “Intertemporal Output and Employment Effects of Public Infrastructure Capital: Evidence from 12 OECD Economies”, *Economic Journal,* 110(465), pp. 687–712.

Doppelhofer, G. and M. Weeks (2008), “Robust model averaging”, paper presented at the CESifo Area Conference Macro, Money and International Finance, 14-15 March, 2008.

Engle, R.F. and C.W.J. Granger (1987), “Cointegration and Error-Correction: Representation, Estimation, and Testing”, *Econometrica* 55(3), pp. 251-276.

Esfahani, H. and M.T. Ramíres (2003), “Institutions, Infrastructure and Economic Growth.” *Journal of Development Economics* 70, pp. 443–477.

Fernald, J. (1999), “Assessing the Link between Public Capital and Productivity”, *American Economic Review* 89, pp. 619–638.

La Ferrara, E.L. and M. Marcellino (2000), “TFP, Costs, and Public Infrastructure: An Equivocal Relationship”, *IGIER Working Paper* No. 176.

Flores de Frutos, R., M. Garcia-Diez,and R. Perez-Amaral (1998), “Public Capital and Economic Growth: An Analysis of the Spanish Economy”, *Applied Economics*, 30, pp. 985–994.

Ford, R. and P. Poret (1991), “Infrastructure and Private-Sector Productivity”, *OECD Economic Studies* 17, pp. 63–89.

Hulten, C.R. (1996), “Infrastructure Capital and Economic Growth, How Well You Use It May Be More Important Than How Much You Have”, *NBER Working Paper* No. 5847.

Im, K.S., H.M. Pesaran and Y. Shin (2003), “Testing for Unit Roots in Heterogeneous Panels”, *Journal of Econometrics* 115(1), pp. 53-74.

Kamps, C. (2005), “The Dynamic Effects of Public Capital: VAR Evidence for 22 OECD Countries”, *International Tax and Public Finance*, 12(4), pp. 533-558.

Kao, C. (1999), "Spurious Regression and Residual-Based Tests for Cointegration in Panel Data", *Journal of Econometrics*, 90(1), pp. 1-44.

Kocherlakota, N.R. and K.-M. Yi (1997), "Is There Endogenous Long-Run Growth? Evidence from the United States and the United Kingdom", *Journal of Money, Credit and Banking*, 29(2), pp. 235-62.

Levin, A., C.-F. Lin and C.S.J. Chu (2002), “Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties”, *Journal of Econometrics* 108(1), pp. 1-24.

Mankiw, G.N., D. Romer and D.N. Weil (1992), “A Contribution to the Empirics of Economic Growth”, *The Quarterly Journal of Economics*, 107(2), pp. 407-437.

Mark, N.C. and D. Sul (2003), “Cointegration Vector Estimation by Panel DOLS and Long-run Money Demand”, *Oxford Bulletin of Economics & Statistics,* 65(5), pp. 655-680.

Moreno, R., E. López-Bazo and M. Artís (2003), “On the Effectiveness of Private and Public Capital”, *Applied Economics*, 35, pp. 727–740.

Pedroni, P. (1999), “Critical Values for Cointegration Tests in Heterogeneous Panels with Multiple Regressors”, *Oxford Bulletin of Economics and Statistics*, 61(4), pp. 653-670.

Pedroni, P. (2004), “Panel Cointegration; Asymptotic and Finite Sample Properties of Pooled Time Series Tests with an Application to the Purchasing Power Parity Hypothesis”, *Econometric Theory*, 20(3), pp. 597-325.

Romp, W. and J. de Haan  (2007), “Public Capital and Economic Growth: A Critical Survey”, *Perspektiven der Wirtschaftspolitik*,8(S1), pp. 6-52.

Sala-i-Martin, X. (1997), “I Just Ran Four Million Regressions”, *American Economic Review*, 97(2), pp. 178-83.

Sala-i-Martin, X., G. Dopperhofer and R. Miller (2004), “Determinants of Long-Run Growth: A Bayesian Averaging of Classical Estimates (BACE) approach”, *American Economic Review*, 94(4), pp. 813-835.

Stephan, A. (2000), “Regional Infrastructure Policy and Its Impact on Productivity: A Comparison of Germany and France”, *Applied Economics Quarterly,* 46, pp. 327–356.

Stock, J.H. and M.W. Watson (1993), “A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems”, Econometrica, 61(4), pp. 783–820.

Sturm, J.E., J. Jacobs and P. Groote (1999), “Output Effects of Infrastructure Investment in the Netherlands, 1853–1913,” *Journal of Macroeconomics* 21, pp. 355–380.

Sturm, J.E., G.H. Kuper and J. de Haan (1998), “Modelling Government Investment and Economic Growth on a Macro Level: A Review”, in: S. Brakman, H. van Ees and S.K. Kuipers (eds.), *Market Behaviour and Macroeconomic Modelling*. Macmillan Press, London.

Temple, J. (2000), "Growth Regressions and What the Textbooks Don't Tell You", *Bulletin of Economic Research*, 52(3), pp. 181-205.

Levine, R. and D. Renelt (1992), A Sensitivity Analysis of Cross-Country Growth Regressions, American Economic Review, 82(4), pp. 942-963.

Hansen, B. (1999), "Threshold Effects in Non-Dynamic Panels: Estimation, Testing and Inference", *Journal of Econometrics*, 93, pp. 345-368.

# Appendix: Supplementary tables and Figures

## [Table A.1. Single series ADF unit root tests for selected variables.]

| maximum lags =2 | ADF constant, no trend | | | ADF constant, trend | | | ADF constant, first diff. | | |
|---|---|---|---|---|---|---|---|---|---|
| AIC criterion for selection | stat. | p-val. | obs | stat. | p-val. | obs | stat. | p-val. | obs |
| *GDP per capita (PPP)* | | | | | | | | | |
| Australia | -0.49 | 0.88 | 44 | -2.60 | 0.28 | 45 | -5.20 | 0.00 *** | 44 |
| Austria | -4.36 | 0.00 *** | 46 | -2.52 | 0.32 | 46 | -1.88 | 0.34 | 43 |
| Belgium | -4.02 | 0.00 *** | 46 | -2.72 | 0.23 | 46 | -3.04 | 0.04 ** | 44 |
| Canada | -1.37 | 0.59 | 45 | -4.01 | 0.02 ** | 45 | -4.91 | 0.00 *** | 45 |
| Denmark | -1.49 | 0.53 | 40 | -4.19 | 0.01 ** | 39 | -5.18 | 0.00 *** | 39 |
| Finland | -0.88 | 0.79 | 44 | -2.36 | 0.39 | 44 | -4.63 | 0.00 *** | 44 |
| France | -2.46 | 0.13 | 42 | -2.72 | 0.23 | 42 | -3.52 | 0.01 *** | 42 |
| Greece | -2.40 | 0.15 | 45 | -2.21 | 0.47 | 45 | -4.51 | 0.00 *** | 45 |
| Iceland | -1.66 | 0.44 | 44 | -2.17 | 0.50 | 44 | -4.87 | 0.00 *** | 44 |
| Ireland | 0.65 | 0.99 | 45 | -1.59 | 0.78 | 45 | -4.38 | 0.00 *** | 45 |
| Italy | -4.95 | 0.00 *** | 46 | -2.45 | 0.35 | 46 | -4.76 | 0.00 *** | 45 |
| Japan | -3.16 | 0.03 ** | 45 | -2.02 | 0.58 | 45 | -1.90 | 0.33 | 43 |
| Korea | -0.40 | 0.90 | 36 | -1.85 | 0.66 | 36 | -5.54 | 0.00 *** | 35 |
| Mexico | -3.55 | 0.01 ** | 45 | -2.68 | 0.25 | 45 | -4.06 | 0.00 *** | 45 |
| Netherlands | -1.25 | 0.65 | 45 | -2.44 | 0.35 | 45 | -3.73 | 0.01 *** | 45 |
| New Zealand | -0.32 | 0.91 | 46 | -1.92 | 0.63 | 46 | -6.53 | 0.00 *** | 45 |
| Norway | -1.79 | 0.38 | 45 | -1.20 | 0.90 | 45 | -4.24 | 0.00 *** | 45 |
| Portugal | -2.50 | 0.12 | 45 | -1.95 | 0.61 | 45 | -4.02 | 0.00 *** | 45 |
| Spain | -1.48 | 0.54 | 45 | -2.47 | 0.34 | 45 | -3.53 | 0.01 *** | 45 |
| Sweden | -1.56 | 0.50 | 44 | -2.77 | 0.22 | 45 | -4.17 | 0.00 *** | 45 |
| Switzerland | -1.73 | 0.41 | 39 | -4.00 | 0.02 ** | 40 | -4.25 | 0.00 *** | 39 |
| Turkey | -1.24 | 0.65 | 46 | -2.48 | 0.34 | 46 | -7.28 | 0.00 *** | 45 |
| United Kingdom | -0.88 | 0.79 | 44 | -3.05 | 0.13 | 45 | -5.29 | 0.00 *** | 44 |
| United States | -0.80 | 0.81 | 44 | -3.82 | 0.02 ** | 45 | -5.03 | 0.00 *** | 44 |
| *Total investment to GDP* | | | | | | | | | |
| Australia | -1.70 | 0.43 | 46 | -2.59 | 0.29 | 45 | -5.49 | 0.00 *** | 44 |
| Austria | -1.18 | 0.68 | 46 | -2.94 | 0.16 | 46 | -6.50 | 0.00 *** | 45 |
| Belgium | -1.94 | 0.31 | 45 | -2.24 | 0.46 | 45 | -5.04 | 0.00 *** | 45 |
| Canada | -2.49 | 0.12 | 45 | -3.38 | 0.07 * | 45 | -4.98 | 0.00 *** | 43 |
| Denmark | -1.94 | 0.31 | 45 | -1.99 | 0.59 | 45 | -5.48 | 0.00 *** | 45 |
| Finland | -2.01 | 0.28 | 45 | -3.33 | 0.07 * | 45 | -4.48 | 0.00 *** | 43 |
| France | -1.64 | 0.46 | 45 | -2.45 | 0.35 | 45 | -4.21 | 0.00 *** | 45 |
| Greece | -2.25 | 0.19 | 46 | -2.31 | 0.42 | 46 | -6.65 | 0.00 *** | 45 |
| Iceland | -1.46 | 0.55 | 45 | -1.38 | 0.85 | 45 | -6.08 | 0.00 *** | 45 |
| Ireland | -2.00 | 0.28 | 45 | -1.98 | 0.60 | 45 | -4.56 | 0.00 *** | 45 |
| Italy | -1.75 | 0.40 | 44 | -2.22 | 0.47 | 44 | -5.68 | 0.00 *** | 44 |
| Japan | -1.22 | 0.66 | 45 | -2.46 | 0.35 | 45 | -4.39 | 0.00 *** | 44 |
| Korea | -2.64 | 0.09 * | 34 | -1.68 | 0.74 | 34 | -6.17 | 0.00 *** | 34 |
| Mexico | -3.27 | 0.02 ** | 45 | -3.24 | 0.09 * | 45 | -6.23 | 0.00 *** | 45 |
| Netherlands | -1.01 | 0.74 | 46 | -2.65 | 0.26 | 44 | -5.85 | 0.00 *** | 45 |
| New Zealand | -3.00 | 0.04 | 45 | -2.99 | 0.14 | 45 | -5.80 | 0.00 *** | 45 |
| Norway | -0.91 | 0.77 | 44 | -3.04 | 0.13 | 45 | -5.10 | 0.00 *** | 44 |
| Portugal | -3.06 | 0.04 ** | 45 | -3.20 | 0.10 * | 45 | -4.88 | 0.00 *** | 44 |
| Spain | -2.06 | 0.26 | 45 | -2.07 | 0.55 | 45 | -3.90 | 0.00 *** | 45 |
| Sweden | -1.86 | 0.35 | 45 | -3.74 | 0.03 ** | 45 | -4.21 | 0.00 *** | 44 |
| Switzerland | -2.22 | 0.20 | 45 | -4.36 | 0.01 * | 45 | -4.41 | 0.00 *** | 44 |
| Turkey | -3.35 | 0.02 ** | 34 | -3.36 | 0.07 * | 34 | -6.60 | 0.00 *** | 33 |
| United Kingdom | -2.84 | 0.06 ** | 45 | -3.34 | 0.07 * | 45 | -4.73 | 0.00 *** | 45 |
| United States | -2.45 | 0.13 | 44 | -2.45 | 0.35 | 44 | -5.67 | 0.00 *** | 44 |

*, **, *** denote that the null hypothesis of a unit root in the series can be rejected at 90%, 95% and 99% confidence levels.

## [Table A.2. Panel unit root test results for the main variables]

|  | Levin, Lin and Chu (2002) |  |  |  |  |  | Im, Pesaran and Shin (2003) |  |  |  |  |  |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| variable | constants, trends |  |  | first differences, constants |  |  | constants, trends |  |  | first differences, constants |  |  |
| maxlags =2, AIC | stat. | p-val. | obs. | stat. | p-val. | obs. | stat. | p-val. | obs. | stat. | p-val. | obs. |
| GDP per capita (PPP) | -4.39 | 0 *** | 1024 | -15.57 | 0 *** | 1045 | -2.54 | 0.01 ** | 1060 | -16.68 | 0 *** | 1045 |
| Total Investment/GDP | -1.05 | 0.15 | 1012 | -21.23 | 0 *** | 1011 | -3.51 | 0 *** | 1023 | -20.28 | 0 *** | 1011 |
| Population growth | 0.51 | 0.7 | 1027 | -3.81 | 0 *** | 1015 | 2.66 | 1 | 1027 | -6.91 | 0 *** | 1015 |
| Electricity (per capita) | -3.9 | 0 *** | 1022 | -9.72 | 0 *** | 989 | 2.69 | 1 | 1022 | -10.35 | 0 *** | 989 |
| Telephone mainln. (per cap.) | 4.54 | 1 | 1006 | 0.13 | 0.55 | 990 | 13.14 | 1 | 1006 | -1.31 | 0.09 * | 990 |
| Telephone subs. (per cap.) | -2.02 | 0.02 ** | 946 | -4.64 | 0 *** | 939 | 2.26 | 0.99 | 946 | -5.94 | 0 *** | 939 |
| Roads (per capita) | -0.39 | 0.35 | 905 | -24.58 | 0 *** | 887 | 1.16 | 0.88 | 905 | -22.82 | 0 *** | 887 |
| Rail (per capita) | -1.02 | 0.15 | 910 | -16.19 | 0 *** | 893 | 1.34 | 0.91 | 910 | -16.55 | 0 *** | 893 |
| Motorway (per capita) | -12.95 | 0 *** | 664 | -19 | 0 *** | 643 | -11.81 | 0 *** | 664 | -14.02 | 0 *** | 643 |
| Human Capital | -2.7 | 0 *** | 672 | -9.5 | 0 *** | 664 | -0.55 | 0.29 | 672 | -8.83 | 0 *** | 664 |
| Private Investment/GDP | -2.13 | 0.02 ** | 865 | -17.73 | 0 *** | 850 | -3.18 | 0 *** | 865 | -17.49 | 0 *** | 850 |
| Capital Stock | -3.68 | 0 *** | 695 | -12.86 | 0 *** | 687 | -2.22 | 0.01 | 695 | -8.42 | 0 *** | 687 |
| Tax revenue/GDP | -0.02 | 0.49 | 529 | -10.08 | 0 *** | 503 | 0.44 | 0.67 | 529 | -11.58 | 0 *** | 503 |
| Trade Openness | -0.83 | 0.2 | 1026 | -25.94 | 0 *** | 1009 | -1.69 | 0.05 * | 1026 | -24.03 | 0 *** | 1009 |
| Electricity | -4.36 | 0 *** | 1020 | -8.86 | 0 *** | 988 | 2.24 | 0.99 | 1020 | -9.52 | 0 *** | 988 |
| Telephone mainlines | 5.13 | 1 | 1008 | 0.89 | 0.81 | 988 | 13.79 | 1 | 1008 | -0.18 | 0.43 | 988 |
| Telephone subscriptions | -2.14 | 0.02 ** | 946 | -4.19 | 0 *** | 938 | 2.42 | 0.99 | 946 | -5.35 | 0 *** | 938 |
| Roads | -2.79 | 0 *** | 911 | -25.23 | 0 *** | 887 | 0.13 | 0.55 | 911 | -23.94 | 0 *** | 887 |
| Rail | -1.01 | 0.16 | 914 | -19.55 | 0 *** | 895 | 2.4 | 0.99 | 914 | -19.69 | 0 *** | 895 |
| Motorway | -14.43 | 0 *** | 666 | -18.48 | 0 *** | 642 | -13.08 | 0 *** | 666 | -13.7 | 0 *** | 642 |

*, **, *** denote that the null hypothesis of a unit root in the series can be rejected at 90%, 95% and 99% confidence levels.

## [Table A.3. Cointegration tests of the basic equations]

|  | Pedroni (1999) |  | Kao (1999) |  |
|---|---|---|---|---|
| Lag selection = AIC, maximum = 2 | panel-v |  | ADF |  |
| country specific constant & trend | stat. | p-val. | stat. | p-val. |
| Y, Invest_total, n, Roads | 6.67 | 0.00 | -2.54 | 0.01 |
| Y, Invest_total, n, Rail | 7.89 | 0.00 | -2.75 | 0.00 |
| Y, Invest_total, n, Motorways | 12.02 | 0.00 | -6.22 | 0.00 |
| Y, Invest_total, n, Electricity | 6.74 | 0.00 | -2.73 | 0.00 |
| Y, Invest_total, n, Tele_mainlines | 3.54 | 0.00 | -2.67 | 0.00 |
| Y, Invest_total, n, Tele_subscriptions | 6.04 | 0.00 | -3.96 | 0.00 |
|  |  |  |  |  |
| with human capital: |  |  |  |  |
| Y, Invest_total, n, Roads, HumanCapital | 4.40 | 0.00 | -1.86 | 0.03 |
| Y, Invest_total, n, Rail, HumanCapital | 5.00 | 0.00 | -1.93 | 0.03 |
| Y, Invest_total, n, Motorways, HumanCapital | 4.16 | 0.00 | -4.07 | 0.00 |
| Y, Invest_total, n, Electricity, HumanCapital | 4.03 | 0.00 | -1.76 | 0.04 |
| Y, Invest_total, n, Tele_mainlines, HumanCapital | 4.25 | 0.00 | -1.50 | 0.07 |
| Y, Invest_total, n, Tele_subscriptions, HumanCapital | 3.13 | 0.00 | -3.07 | 0.00 |

The null hypothesis in both tests is no cointegration. Pedroni (1999) test allows individual (heterogenous) cointegrating relationships; Kao (1999) assumes a common (homogenous) relationship.

### [Table A.4. Bi-variate Granger causality tests for single series]

| single series<br>number of lags = 4<br>first differences, constant | Null hypothesis<br>Infrastructure does not cause output<br>p-value | <br>Outpud does not cause infrastructure<br>p-value |
|---|---|---|
| *Roads* | | |
| Denmark | 0.40 | 0.03 ** |
| Japan | 0.02 ** | 0.97 |
| Netherlands | 0.08 * | 0.21 |
| *Motorways* | | |
| France | 0.01 *** | 0.16 |
| Italy | 0.02 ** | 0.36 |
| Sweden | 0.02 ** | 0.87 |
| *Rail* | | |
| Australia | 0.06 * | 0.81 |
| France | 0.13 | 0.04 ** |
| *Electricity* | | |
| Greece | 0.02 ** | 0.06 * |
| Iceland | 0.07 * | 0.33 |
| Italy | 0.07 * | 0.11 |
| Japan | 0.44 | 0.01 ** |
| Mexico | 0.35 | 0.02 ** |
| Norway | 0.09 * | 0.30 |
| Spain | 0.05 ** | 0.50 |
| *Telephone mainlines* | | |
| Japan | 0.04 ** | 0.07 * |
| Mexico | 0.82 | 0.03 ** |
| Portugal | 0.04 ** | 0.23 |
| Spain | 0.48 | 0.01 *** |
| *Telephone subscriptions* | | |
| Korea | 0.30 | 0.02 ** |
| Mexico | 0.02 ** | 0.59 |
| Netherlands | 0.06 * | 0.66 |

1. Only reported if the lack one of Grange causality in at least one direction is rejected.
2. *, **, *** denote that the null hypothesis can be rejected at 90%, 95% and 99% confidence levels.

### [Table A.5. Panel bi-variate Granger causality test results]

| homogenous panel<br>with country fixed effects<br>number of lags = 4<br>first differences | Null hypothesis<br>Infrastructure does not cause output<br>p-value | <br>Output does not cause infrastructure<br>p-value |
|---|---|---|
| Road | 0.71 | 0.09 * |
| Motorways | 0.18 | 0.22 |
| Rail | 0.36 | 0.07 ** |
| Electricity | 0.00 *** | 0.00 *** |
| Telephone mainlines | 0.00 *** | 0.00 *** |
| Telephone subscriptions | 0.15 | 0.02 * |

3. The test assumes the homogeneity of coefficients.
4. *, **, *** denote that the null hypothesis can be rejected at 90%, 95% and 99% confidence levels respectively

## [Table A.6. Panel coefficient estimates with homogenous coefficient assumptions]

| | Road | Motorway | Rail |
|---|---|---|---|
| Investment | 0.15 *** | 0.3 *** | 0.16 *** |
| Population growth | 0.03 *** | -0.01 | 0.04 *** |
| Human Capital | 0.68 *** | 0.15 | 0.04 |
| Infrastructure | -0.19 *** | 0.09 *** | 0.05 |
| Trend | 0.02 *** | 0.02 *** | 0.02 *** |
| Obs. | 557 | 484 | 605 |

| | Electricity | Telephone main. | Telephone subs. |
|---|---|---|---|
| Investment | 0.15 *** | 0.18 *** | 0.16 *** |
| Population growth | 0.04 *** | 0.04 *** | 0.04 *** |
| Human Capital | 0.14 | 0.11 | 0.28 *** |
| Infrastructure | 0.09 *** | 0.12 *** | -0.03 ** |
| Trend | 0.02 *** | 0.02 *** | 0.02 *** |
| Obs. | 637 | 634 | 618 |

5. Long run coefficients reported. Panel includes country specific fixed effects. Dynamic OLS used.

6. *, **, *** denote that the significance at 90%, 95% and 99% confidence levels respectively.

## [Table A.7.. Principal components: Factor loadings ]

| | Energy 1<br>WDI – energy use | | Energy 2<br>WDI – energy consumption | | Energy 3<br>IEA – electricity generation | | Energy 4<br>IEA – electricity capacity | |
|---|---|---|---|---|---|---|---|---|
| | **1st PC** | **2nd PC** | **1st PC** | **2nd PC** | **1st PC** | **2nd PC** | **1st PC** | **2nd PC** |
| **Cumulative variance** | 0.44 | 0.79 | 0.45 | 0.75 | 0.45 | 0.79 | 0.44 | 0.80 |
| | **loadings** | | **loadings** | | **loadings** | | **loadings** | |
| 1 (energy) | 0.07 | 0.71 | -0.19 | 0.69 | 0.24 | 0.67 | 0.09 | 0.71 |
| 2 (motorw) | 0.70 | -0.11 | 0.69 | 0.06 | 0.66 | -0.25 | 0.69 | -0.15 |
| 3 (rail) | 0.67 | -0.21 | 0.68 | -0.04 | 0.64 | -0.31 | 0.67 | -0.21 |
| 4 (telecom) | 0.26 | 0.66 | 0.16 | 0.72 | 0.32 | 0.62 | 0.27 | 0.65 |
| **Cumulative variance** | 0.45 | 0.80 | 0.45 | 0.76 | 0.46 | 0.79 | 0.45 | 0.80 |
| | **loadings** | | **loadings** | | **loadings** | | **loadings** | |
| 1 (energy) | 0.04 | 0.72 | -0.19 | 0.70 | 0.23 | 0.68 | 0.09 | 0.72 |
| 2 (motorw) | 0.70 | -0.09 | 0.69 | 0.05 | 0.66 | -0.26 | 0.69 | -0.15 |
| 3 (rail) | 0.67 | -0.19 | 0.68 | -0.04 | 0.64 | -0.29 | 0.67 | -0.21 |
| 4 (telecom) | 0.25 | 0.66 | 0.17 | 0.72 | 0.32 | 0.62 | 0.27 | 0.65 |

## Table A.8. Full model averaging, 1$^{st}$ principal component

linearity, full sample

### PANEL A: 5-year averages

| | pc11 | | | pc12 | | | pc13 | | | pc14 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Incl. prob. | mean | s.e. | Incl. prob. | mean | s.e. | Incl. prob. | mean | s.e. | Incl. prob. | mean | s.e. |
| capita(-1) | **0.88** | -3.41 | 3.97 | **0.96** | -4.05 | 3.37 | **0.78** | -2.86 | 4.46 | **0.86** | -3.28 | 4.13 |
| life_exp(-1) | 0.37 | 0.10 | 0.09 | **0.50** | 0.15 | 0.12 | 0.35 | 0.09 | 0.09 | 0.37 | 0.10 | 0.09 |
| cpi(-1) | 0.11 | 0.00 | 0.00 | 0.11 | 0.00 | 0.00 | 0.11 | 0.00 | 0.00 | 0.11 | 0.00 | 0.00 |
| reg(-1) | **0.51** | -0.28 | 0.21 | 0.36 | -0.17 | 0.15 | **0.60** | -0.35 | 0.24 | **0.56** | -0.31 | 0.22 |
| d_lf(-1) | 0.15 | 0.07 | 0.10 | 0.13 | 0.05 | 0.08 | 0.17 | 0.09 | 0.12 | 0.16 | 0.08 | 0.11 |
| open(-1) | **1.00** | 0.10 | 0.02 | **1.00** | 0.10 | 0.03 | **1.00** | 0.10 | 0.02 | **1.00** | 0.10 | 0.02 |
| inv_gov(-1) | 0.13 | -0.01 | 0.03 | 0.12 | -0.01 | 0.02 | 0.14 | -0.02 | 0.03 | 0.13 | -0.02 | 0.03 |
| inv_price(-1) | **0.92** | -0.05 | 0.03 | **0.94** | -0.05 | 0.03 | **0.93** | -0.05 | 0.03 | **0.93** | -0.05 | 0.03 |
| edu(-1) | 0.39 | 0.19 | 0.17 | 0.34 | 0.16 | 0.15 | 0.37 | 0.18 | 0.17 | 0.38 | 0.19 | 0.17 |
| infrastr(-1) | **1.00** | -1.30 | 1.10 | **1.00** | -1.48 | 0.81 | **1.00** | -1.52 | 1.05 | **1.00** | -1.39 | 1.11 |

### PANEL B: 8-year averages

| | pc11 | | | pc12 | | | pc13 | | | pc14 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Incl. prob. | mean | s.e. | Incl. prob. | mean | s.e. | Incl. prob. | mean | s.e. | Incl. prob. | mean | s.e. |
| capita(-1) | **1.00** | -7.03 | 2.56 | **1.00** | -7.09 | 2.37 | **1.00** | -7.14 | 2.61 | **1.00** | -7.07 | 2.58 |
| life_exp(-1) | **0.56** | 0.17 | 0.10 | **0.61** | 0.19 | 0.11 | **0.52** | 0.15 | 0.10 | **0.54** | 0.16 | 0.10 |
| cpi(-1) | 0.43 | 0.01 | 0.02 | 0.43 | 0.01 | 0.02 | 0.42 | 0.01 | 0.02 | 0.43 | 0.01 | 0.02 |
| reg(-1) | 0.15 | -0.02 | 0.03 | 0.15 | -0.02 | 0.03 | 0.14 | -0.02 | 0.03 | 0.15 | -0.02 | 0.03 |
| d_lf(-1) | 0.12 | 0.00 | 0.06 | 0.12 | 0.00 | 0.06 | 0.12 | -0.01 | 0.07 | 0.12 | -0.01 | 0.07 |
| open(-1) | **0.74** | 0.03 | 0.02 | **0.77** | 0.03 | 0.02 | **0.73** | 0.03 | 0.02 | **0.73** | 0.03 | 0.02 |
| inv_gov(-1) | **0.54** | -0.14 | 0.08 | 0.48 | -0.12 | 0.08 | **0.57** | -0.15 | 0.09 | **0.55** | -0.14 | 0.09 |
| inv_price(-1) | **0.66** | -0.04 | 0.02 | **0.66** | -0.04 | 0.02 | **0.65** | -0.03 | 0.02 | **0.66** | -0.04 | 0.02 |
| edu(-1) | **1.00** | 0.93 | 0.29 | **0.99** | 0.90 | 0.30 | **1.00** | 0.95 | 0.28 | **1.00** | 0.94 | 0.28 |
| infrastr(-1) | **1.00** | -0.02 | 1.15 | **1.00** | -0.29 | 1.02 | **1.00** | 0.23 | 1.00 | **1.00** | 0.12 | 1.12 |

### PANEL C: 10-year averages

| | pc11 | | | pc12 | | | pc13 | | | pc14 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Incl. prob. | mean | s.e. | Incl. prob. | mean | s.e. | Incl. prob. | mean | s.e. | Incl. prob. | mean | s.e. |
| capita(-1) | **1.00** | -11.20 | 2.97 | **1.00** | -11.05 | 3.05 | **1.00** | -11.26 | 2.67 | **1.00** | -11.28 | 2.86 |
| life_exp(-1) | **1.00** | 0.67 | 0.21 | **1.00** | 0.70 | 0.22 | **1.00** | 0.64 | 0.18 | **1.00** | 0.64 | 0.19 |
| cpi(-1) | **0.85** | -0.10 | 0.04 | **0.86** | -0.10 | 0.04 | **0.84** | -0.09 | 0.04 | **0.84** | -0.09 | 0.04 |
| reg(-1) | **0.90** | -0.57 | 0.31 | **0.93** | -0.62 | 0.32 | **0.75** | -0.42 | 0.29 | **0.83** | -0.50 | 0.30 |
| d_lf(-1) | 0.32 | 0.31 | 0.57 | 0.32 | 0.32 | 0.60 | 0.29 | 0.26 | 0.49 | 0.32 | 0.30 | 0.55 |
| open(-1) | 0.27 | 0.01 | 0.01 | 0.28 | 0.01 | 0.01 | 0.26 | 0.01 | 0.01 | 0.26 | 0.01 | 0.01 |
| inv_gov(-1) | 0.32 | 0.08 | 0.09 | 0.36 | 0.09 | 0.10 | 0.24 | 0.05 | 0.06 | 0.27 | 0.06 | 0.07 |
| inv_price(-1) | 0.33 | 0.00 | 0.04 | 0.33 | 0.01 | 0.04 | 0.32 | 0.00 | 0.03 | 0.33 | 0.00 | 0.04 |
| edu(-1) | 0.41 | 0.17 | 0.19 | 0.36 | 0.14 | 0.17 | 0.39 | 0.15 | 0.17 | 0.46 | 0.20 | 0.21 |
| infrastr(-1) | **0.50** | 0.70 | 1.45 | 0.36 | 0.04 | 0.54 | **0.73** | 1.35 | 1.80 | **0.68** | 1.31 | 2.08 |

Notes: PC11: based on principal components using energy consumption (World Bank), PC12: based on principal components using energy production (World Bank), PC13: based on principal components using electricity generation (IEA), PC14: based on principal components using total electricity generating capacity (IEA).

inclusion probability is the posterior inclusion probability, mean is the posterior mean conditional on inclusion, s.e. is the posterior standard error conditional on inclusion. Bold figures for the posterior inclusion probability indicate that a given variable passes the prior inclusion probability of 1/2..

## Table A.9. Full model averaging, 2nd principal component, linearity, full sample

**PANEL A: 5-year averages**

| | pc21 | | | pc22 | | | pc23 | | | pc24 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Incl. prob. | mean | s.e. | Incl. prob. | mean | s.e. | Incl. prob. | mean | s.e. | Incl. prob. | mean | s.e. |
| capita(-1) | **0.96** | -4.44 | 2.94 | **0.99** | -4.57 | 2.83 | **0.91** | -3.76 | 3.65 | **0.94** | -3.94 | 3.36 |
| life_exp(-1) | 0.25 | 0.06 | 0.06 | 0.35 | 0.09 | 0.08 | 0.32 | 0.08 | 0.08 | 0.32 | 0.08 | 0.08 |
| cpi(-1) | 0.11 | 0.00 | 0.00 | 0.11 | 0.00 | 0.00 | 0.11 | 0.00 | 0.00 | 0.11 | 0.00 | 0.00 |
| reg(-1) | 0.35 | -0.17 | 0.15 | 0.29 | -0.13 | 0.12 | **0.55** | -0.33 | 0.24 | **0.52** | -0.31 | 0.24 |
| d_lf(-1) | 0.09 | 0.02 | 0.05 | 0.09 | 0.01 | 0.04 | **0.10** | 0.02 | 0.05 | 0.10 | 0.02 | 0.05 |
| open(-1) | **1.00** | 0.10 | 0.02 | **1.00** | 0.10 | 0.02 | **1.00** | 0.09 | 0.02 | **1.00** | 0.09 | 0.02 |
| inv_gov(-1) | 0.12 | -0.01 | 0.02 | 0.12 | -0.01 | 0.02 | 0.13 | -0.01 | 0.03 | 0.13 | -0.01 | 0.03 |
| inv_price(-1) | **0.87** | -0.04 | 0.03 | **0.92** | -0.05 | 0.03 | **0.93** | -0.05 | 0.03 | **0.92** | -0.05 | 0.03 |
| edu(-1) | 0.42 | 0.21 | 0.19 | 0.41 | 0.21 | 0.19 | 0.47 | 0.26 | 0.22 | 0.47 | 0.25 | 0.21 |
| infrastr(-1) | **1.00** | 0.50 | 1.13 | **1.00** | 0.35 | 0.28 | **1.00** | -0.38 | 0.93 | **1.00** | -0.22 | 0.94 |

**PANEL B: 8-year averages**

| | pc21 | | | pc22 | | | pc23 | | | pc24 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Incl. prob. | mean | s.e. | Incl. prob. | mean | s.e. | Incl. prob. | mean | s.e. | Incl. prob. | mean | s.e. |
| capita(-1) | **1.00** | -8.62 | 1.62 | **1.00** | -7.33 | 2.68 | **1.00** | -7.82 | 2.20 | **1.00** | -7.83 | 2.11 |
| life_exp(-1) | 0.20 | 0.03 | 0.04 | **0.59** | 0.18 | 0.11 | 0.42 | 0.10 | 0.08 | 0.38 | 0.09 | 0.07 |
| cpi(-1) | 0.25 | 0.00 | 0.01 | 0.42 | 0.00 | 0.02 | 0.35 | 0.00 | 0.01 | 0.33 | 0.00 | 0.01 |
| reg(-1) | 0.15 | 0.02 | 0.03 | 0.13 | -0.01 | 0.03 | 0.11 | 0.00 | 0.02 | 0.12 | 0.01 | 0.03 |
| d_lf(-1) | 0.18 | -0.06 | 0.11 | 0.12 | -0.01 | 0.07 | 0.13 | -0.02 | 0.07 | 0.13 | -0.03 | 0.08 |
| open(-1) | **0.96** | 0.05 | 0.02 | **0.79** | 0.04 | 0.02 | **0.87** | 0.04 | 0.02 | **0.84** | 0.04 | 0.02 |
| inv_gov(-1) | **0.57** | -0.12 | 0.07 | **0.49** | -0.12 | 0.08 | **0.52** | -0.12 | 0.07 | **0.54** | -0.12 | 0.08 |
| inv_price(-1) | 0.31 | -0.01 | 0.01 | **0.64** | -0.03 | 0.02 | 0.48 | -0.02 | 0.01 | 0.45 | -0.02 | 0.01 |
| edu(-1) | **1.00** | 0.91 | 0.28 | **1.00** | 0.92 | 0.28 | **1.00** | 0.92 | 0.28 | **1.00** | 0.94 | 0.28 |
| infrastr(-1) | **1.00** | 1.69 | 0.61 | **1.00** | 0.15 | 0.34 | **1.00** | 0.96 | 0.68 | **1.00** | 0.96 | 0.59 |

**PANEL C: 10-year averages**

| | pc21 | | | pc22 | | | pc23 | | | pc24 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Incl. prob. | mean | s.e. | Incl. prob. | mean | s.e. | Incl. prob. | mean | s.e. | Incl. prob. | mean | s.e. |
| **10-year averages** | | | | | | | | | | | | |
| life_exp(-1) | 0.36 | 0.08 | 0.09 | **1.00** | 0.68 | 0.19 | **0.99** | 0.51 | 0.18 | **0.99** | 0.43 | 0.17 |
| cpi(-1) | 0.26 | -0.01 | 0.02 | **0.88** | -0.10 | 0.04 | **0.91** | -0.09 | 0.04 | **0.88** | -0.07 | 0.04 |
| reg(-1) | 0.14 | -0.01 | 0.04 | **0.68** | -0.38 | 0.29 | 0.15 | -0.02 | 0.05 | 0.13 | 0.01 | 0.04 |
| d_lf(-1) | 0.42 | -0.35 | 0.51 | 0.25 | 0.20 | 0.42 | 0.16 | 0.06 | 0.17 | 0.13 | 0.01 | 0.11 |
| open(-1) | **1.00** | 0.11 | 0.03 | 0.33 | 0.01 | 0.01 | **0.72** | 0.04 | 0.03 | **0.85** | 0.04 | 0.03 |
| inv_gov(-1) | 0.17 | -0.01 | 0.03 | 0.25 | 0.05 | 0.07 | 0.13 | 0.00 | 0.02 | 0.13 | 0.00 | 0.02 |
| inv_price(-1) | 0.20 | 0.01 | 0.02 | 0.30 | 0.01 | 0.04 | 0.35 | 0.02 | 0.04 | 0.36 | 0.01 | 0.04 |
| edu(-1) | 0.36 | 0.11 | 0.12 | 0.23 | 0.06 | 0.10 | 0.14 | 0.01 | 0.04 | 0.24 | 0.05 | 0.07 |
| infrastr(-1) | **1.00** | 5.73 | 1.12 | **0.72** | 0.44 | 0.37 | **1.00** | 2.91 | 1.19 | **1.00** | 3.07 | 0.83 |

Notes: PC21: based on principal components using energy consumption (World Bank), PC22: based on principal components using energy production (World Bank), PC23: based on principal components using electricity generation (IEA), PC24: based on principal components using total electricity generating capacity (IEA).

inclusion probability is the posterior inclusion probability, mean is the posterior mean conditional on inclusion, s.e. is the posterior standard error conditional on inclusion. Bold figures for the posterior inclusion probability indicate that a given variable passes the prior inclusion probability of 1/2.

## Table A.10. Fixed effect OLS and GMM estimations,

5-year averages

**Fixed effect OLS**

| | 1st principal component | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | **PC11** | | **PC12** | | **PC13** | | **PC14** | |
| | Coef | p-value | Coef | p-value | Coef | p-value | Coef | p-value |
| Cappp | -5.41 | 0.00 | -5.94 | 0.00 | -5.03 | 0.00 | -5.29 | 0.00 |
| life_exp | 0.29 | 0.14 | 0.30 | 0.12 | 0.30 | 0.11 | 0.30 | 0.12 |
| cpi | 0.01 | 0.22 | 0.00 | 0.27 | 0.01 | 0.19 | 0.01 | 0.21 |
| reg | -0.53 | 0.02 | -0.43 | 0.05 | -0.59 | 0.01 | -0.56 | 0.01 |
| dpop1564 | 0.57 | 0.15 | 0.48 | 0.20 | 0.61 | 0.12 | 0.60 | 0.14 |
| opengs | 0.10 | 0.00 | 0.10 | 0.00 | 0.10 | 0.00 | 0.10 | 0.00 |
| ig_gdp | 0.06 | 0.70 | 0.07 | 0.64 | 0.06 | 0.69 | 0.06 | 0.69 |
| i_price | -0.06 | 0.10 | -0.06 | 0.10 | -0.06 | 0.09 | -0.06 | 0.09 |
| edu2 | 0.46 | 0.03 | 0.43 | 0.05 | 0.47 | 0.02 | 0.46 | 0.03 |
| **infrastr** | -1.72 | 0.06 | -1.58 | 0.04 | -1.91 | 0.03 | -1.87 | 0.04 |

| | 2nd principal component | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | **PC21** | | **PC22** | | **PC23** | | **PC24** | |
| | Coef | p-value | Coef | p-value | Coef | p-value | Coef | p-value |
| Cappp | -5.80 | 0.00 | -6.20 | 0.00 | -5.47 | 0.00 | -5.53 | 0.00 |
| life_exp | 0.25 | 0.24 | 0.23 | 0.28 | 0.26 | 0.21 | 0.27 | 0.18 |
| cpi | 0.01 | 0.28 | 0.01 | 0.28 | 0.01 | 0.23 | 0.01 | 0.22 |
| reg | -0.51 | 0.03 | -0.43 | 0.07 | -0.61 | 0.01 | -0.64 | 0.01 |
| dpop1564 | 0.36 | 0.35 | 0.31 | 0.42 | 0.38 | 0.34 | 0.40 | 0.30 |
| opengs | 0.09 | 0.00 | 0.09 | 0.00 | 0.09 | 0.00 | 0.09 | 0.00 |
| ig_gdp | 0.04 | 0.81 | 0.04 | 0.80 | 0.04 | 0.79 | 0.05 | 0.77 |
| i_price | -0.06 | 0.13 | -0.06 | 0.11 | -0.06 | 0.11 | -0.06 | 0.11 |
| edu2 | 0.51 | 0.03 | 0.49 | 0.04 | 0.53 | 0.02 | 0.53 | 0.02 |
| **infrastr** | -0.35 | 0.65 | 0.12 | 0.60 | -0.86 | 0.23 | -0.87 | 0.18 |

**First difference GMM**

| | 1st principal component | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | **PC11** | | **PC12** | | **PC13** | | **PC14** | |
| | Coef | p-value | Coef | p-value | Coef | p-value | Coef | p-value |
| cappp | -5.76 | 0.00 | -6.34 | 0.00 | -5.52 | 0.00 | -5.71 | 0.00 |
| life_exp | 0.30 | 0.13 | 0.33 | 0.12 | 0.33 | 0.11 | 0.32 | 0.11 |
| cpi | 0.01 | 0.21 | 0.00 | 0.28 | 0.01 | 0.18 | 0.01 | 0.21 |
| reg | -0.58 | 0.01 | -0.45 | 0.03 | -0.65 | 0.00 | -0.62 | 0.00 |
| dpop1564 | 0.57 | 0.12 | 0.48 | 0.19 | 0.59 | 0.10 | 0.59 | 0.11 |
| opengs | 0.12 | 0.00 | 0.11 | 0.00 | 0.12 | 0.00 | 0.12 | 0.00 |
| ig_gdp | 0.10 | 0.52 | 0.11 | 0.48 | 0.11 | 0.44 | 0.10 | 0.48 |
| i_price | -0.06 | 0.08 | -0.06 | 0.08 | -0.06 | 0.08 | -0.06 | 0.08 |
| edu2 | 0.58 | 0.02 | 0.52 | 0.03 | 0.60 | 0.02 | 0.58 | 0.02 |
| **infrastr** | -2.22 | 0.03 | -1.90 | 0.04 | -2.37 | 0.01 | -2.32 | 0.02 |
| Sargan | 0.54 | | 0.49 | | 0.61 | | 0.57 | |

| | 2nd principal component | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | **PC21** | | **PC22** | | **PC23** | | **PC24** | |
| | Coef | p-value | Coef | p-value | Coef | p-value | Coef | p-value |
| cappp | -6.67 | 0.00 | -6.77 | 0.00 | -5.82 | 0.00 | -6.07 | 0.00 |
| life_exp | 0.27 | 0.21 | 0.28 | 0.23 | 0.25 | 0.26 | 0.28 | 0.20 |
| cpi | 0.00 | 0.38 | 0.00 | 0.30 | 0.01 | 0.22 | 0.01 | 0.24 |
| reg | -0.46 | 0.03 | -0.44 | 0.03 | -0.57 | 0.01 | -0.58 | 0.02 |
| dpop1564 | 0.35 | 0.37 | 0.37 | 0.31 | 0.39 | 0.29 | 0.40 | 0.27 |
| opengs | 0.10 | 0.00 | 0.10 | 0.00 | 0.09 | 0.00 | 0.10 | 0.00 |
| ig_gdp | 0.10 | 0.54 | 0.11 | 0.54 | 0.07 | 0.65 | 0.09 | 0.57 |
| i_price | -0.06 | 0.14 | -0.05 | 0.11 | -0.06 | 0.10 | -0.06 | 0.11 |
| edu2 | 0.68 | 0.00 | 0.68 | 0.00 | 0.72 | 0.00 | 0.69 | 0.00 |
| **infrastr** | -0.08 | 0.93 | -0.01 | 0.99 | -0.71 | 0.30 | -0.63 | 0.37 |
| Sargan | 0.44 | | 0.43 | | 0.44 | | 0.45 | |

Notes: PC11: based on principal components using energy consumption (World Bank), PC12: based on principal components using energy production (World Bank), PC13: based on principal components using electricity generation (IEA), PC14: based on principal components using total electricity generating capacity (IEA). PC21: based on principal components using energy consumption (World Bank), PC22: based on principal components using energy production (World Bank), PC23: based on principal components using electricity generation (IEA), PC24: based on principal components using total electricity generating capacity (IEA). P-values are shown for the Arellano, Sargan and Hansen tests.

## Table A.11. The nonlinear relationship between growth and infrastructure

(1st principal component)

### PANEL A: 8-year averages

| | PC11 | | | | | | PC12 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Linear model | | 2-regime model | | 3-regime model | | Linear model | | 2-regime Model | | 3-regime model | |
| Capita | -7.258 | *** | -7.287 | *** | -3.574 | *** | -7.310 | *** | -7.171 | *** | -4.835 | *** |
| life_exp | 0.193 | | 0.091 | | 0.393 | | 0.211 | | 0.162 | | 1.172 | |
| opengs | 0.037 | | 0.027 | | 0.954 | | 0.040 | | 0.030 | | 1.137 | |
| ig_gdp | -0.163 | | -0.258 | * | -1.692 | * | -0.149 | | -0.222 | | -1.323 | |
| i_price | -0.032 | *** | -0.026 | *** | -3.099 | *** | -0.032 | *** | -0.032 | *** | -2.861 | *** |
| edu2 | 0.916 | *** | 0.990 | *** | 4.327 | *** | 0.888 | *** | 0.869 | *** | 3.776 | *** |
| Infrastr | | | | | | | | | | | | |
| Low | -0.152 | | 0.904 | | -0.247 | | -0.411 | | 0.433 | | -0.295 | |
| High/mid | | | 0.027 | | -1.506 | | | | -0.654 | | -1.697 | |
| high | | | | | 0.377 | | | | | | 0.004 | |
| t1 | | | 0.320 | | 0.32 | | | | 0.410 | | 0.405 | |
| t2 | | | | | 0.725 | | | | | | 0.735 | |
| p-value (bootstrap) | | | 0.052 | | 0.317 | | | | 0.136 | | 0.294 | |

| | PC13 | | | | | | PC14 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Linear model | | 2-regime model | | 3-regime model | | Linear model | | 2-regime Model | | 3-regime model | |
| Capita | -7.391 | *** | -7.307 | *** | -3.706 | *** | -7.309 | *** | -6.612 | *** | -3.704 | *** |
| life_exp | 0.182 | | 0.270 | | 1.540 | | 0.188 | | 0.348 | ** | 2.242 | ** |
| opengs | 0.037 | | 0.033 | | 1.006 | | 0.037 | | 0.046 | * | 1.781 | * |
| ig_gdp | -0.167 | | -0.118 | | -1.017 | | -0.165 | | -0.014 | | -0.064 | |
| i_price | -0.032 | *** | -0.032 | *** | -3.155 | *** | -0.032 | *** | -0.038 | *** | -3.038 | *** |
| edu2 | 0.931 | *** | 0.928 | *** | 3.301 | *** | 0.923 | *** | 0.965 | *** | 3.753 | *** |
| Infrastr | | | | | | | | | | | | |
| Low | 0.105 | | -0.728 | | 0.328 | | -0.031 | | -3.578 | | -1.184 | |
| High/mid | | | -0.087 | | 1.384 | | | | -0.756 | | 0.831 | |
| high | | | | | -0.260 | | | | | | -1.296 | |
| t1 | | | 0.285 | | 0.325 | | | | 0.435 | | 0.435 | |
| t2 | | | | | 0.715 | | | | | | 0.685 | |
| p-value (bootstrap) | | | 0.253 | | 0.253 | | | | 0.035 | | 0.216 | |

### PANEL A: 10-year averages

| | PC11 | | | | | | PC12 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Linear model | | 2-regime model | | 3-regime model | | Linear model | | 2-regime Model | | 3-regime model | |
| Capita | -10.766 | *** | -7.155 | ** | -2.498 | ** | -10.294 | *** | -10.590 | *** | -5.331 | *** |
| life_exp | 0.597 | *** | 0.599 | *** | 3.743 | *** | 0.608 | *** | 0.659 | *** | 3.977 | *** |
| opengs | -0.089 | *** | -0.109 | *** | -8.131 | *** | -0.094 | *** | -0.094 | *** | -7.220 | *** |
| ig_gdp | -0.565 | * | -0.269 | | -0.908 | | -0.590 | * | -0.610 | ** | -2.132 | ** |
| i_price | 0.023 | | 0.033 | | 1.096 | | 0.030 | | 0.012 | | 0.114 | |
| Infrastr | | | | | | | | | | | | |
| Low | 0.904 | | -3.876 | | -1.251 | | -0.488 | | -1.302 | | -0.456 | |
| High/mid | | | -0.099 | | -1.277 | | | | 0.202 | | 1.031 | |
| high | | | | | 0.036 | | | | | | -0.276 | |
| t1 | | | 0.425 | | 0.425 | | | | 0.445 | | 0.445 | |
| t2 | | | | | 0.745 | | | | | | 0.745 | |
| p-value (bootstrap) | | | 0.015 | | 0.636 | | | | 0.01 | | 0.640 | |

| | PC13 | | | | | | PC14 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Linear model | | 2-regime model | | 3-regime model | | Linear model | | 2-regime Model | | 3-regime model | |
| Capita | -11.068 | *** | -8.977 | *** | -4.849 | *** | -10.932 | *** | -9.316 | *** | -4.773 | *** |
| life_exp | 0.581 | *** | 0.620 | *** | 4.013 | *** | 0.591 | *** | 0.599 | *** | 3.997 | *** |
| opengs | -0.086 | *** | -0.088 | *** | -5.044 | *** | -0.087 | *** | -0.112 | *** | -8.130 | *** |
| ig_gdp | -0.484 | * | -0.448 | | -1.472 | | -0.524 | * | -0.501 | * | -1.586 | |
| i_price | 0.023 | | 0.009 | | 0.044 | | 0.021 | | 0.048 | | 1.472 | |
| edu2 | | | | | | | | | | | | |
| Infrastr | | | | | | | | | | | | |
| Low | 1.483 | | -0.711 | | 0.101 | | 1.361 | | -2.825 | | -0.972 | |
| High/mid | | | 0.875 | | 2.028 | | | | 0.891 | | 0.580 | |
| high | | | | | 0.829 | | | | | | 0.933 | |
| t1 | | | 0.305 | | 0.310 | | | | 0.495 | | 0.495 | |
| t2 | | | | | 0.695 | | | | | | 0.745 | |
| p-value (bootstrap) | | | 0.014 | | 0.497 | | | | 0.016 | | 0.879 | |

Notes: low, high/mid, high are the lower regimes, the upper regimes (2-regime model)/the middle-regime (3-regime model), and the upper regime for the 3-regime model. T1 and t2 are the two thresholds. PC11: based on principal components using energy consumption (World Bank), PC12: based on principal components using energy production (World Bank), PC13: based on principal components using electricity generation (IEA), PC14: based on principal components using total electricity generating capacity (IEA).

## Table A.12. The nonlinear relationship between growth and infrastructure

(2nd principal component)

**PANEL A: 8-year averages**

| | PC21 | | | | | | PC22 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Linear model | | 2-regime model | | 3-regime model | | Linear model | | 2-regime Model | | 3-regime model | |
| Capita | -8.584 | *** | -8.983 | *** | -5.713 | *** | -7.583 | *** | -8.399 | *** | -4.098 | *** |
| life_exp | 0.060 | | -0.152 | | -0.893 | | 0.192 | | 0.164 | | 0.805 | |
| opengs | 0.046 | * | 0.046 | ** | 1.912 | * | 0.038 | | 0.036 | | 1.374 | |
| ig_gdp | -0.178 | | -0.369 | *** | -2.984 | *** | -0.154 | | -0.162 | | -1.211 | |
| i_price | -0.021 | * | -0.023 | ** | -1.962 | * | -0.031 | *** | -0.031 | *** | -2.608 | ** |
| edu2 | 0.906 | *** | 0.905 | *** | 3.747 | *** | 0.911 | *** | 0.871 | *** | 3.676 | *** |
| Infrastr | | | | | | | | | | | | |
| Low | 1.376 | ** | 2.770 | *** | 2.437 | ** | 0.140 | | 0.892 | | 1.321 | |
| High/mid | | | 1.578 | *** | 0.960 | | | | 0.227 | | 0.624 | |
| high | | | | | 1.666 | *** | | | | | 0.228 | |
| t1 | | | 0.255 | | 0.250 | | | | 0.300 | | 0.305 | |
| t2 | | | | | 0.53 | | | | | | 0.730 | |
| p-value (bootstrap) | | | 0.004 | | 0.447 | | | | 0.209 | | 0.549 | |

| | PC23 | | | | | | PC24 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Linear model | | 2-regime Model | | 3-regime model | | Linear model | | 2-regime Model | | 3-regime Model | |
| Capita | -8.055 | *** | -8.688 | *** | -4.852 | *** | -8.066 | *** | -9.152 | *** | -5.343 | *** |
| life_exp | 0.142 | | -0.093 | | -0.389 | | 0.131 | | -0.037 | | -0.115 | |
| opengs | 0.040 | | 0.038 | | 1.570 | | 0.039 | | 0.027 | | 1.035 | |
| ig_gdp | -0.156 | | -0.342 | ** | -2.368 | ** | -0.157 | | -0.298 | ** | -2.545 | ** |
| i_price | -0.026 | ** | -0.027 | ** | -2.610 | ** | -0.025 | ** | -0.026 | ** | -2.488 | ** |
| edu2 | 0.907 | *** | 0.882 | *** | 3.577 | *** | 0.927 | *** | 0.946 | *** | 4.001 | *** |
| Infrastr | | | | | | | | | | | | |
| Low | 0.770 | | 2.831 | ** | 1.724 | * | 0.747 | | 2.470 | *** | 2.176 | ** |
| High/mid | | | 1.365 | ** | 1.011 | | | | 1.184 | ** | 0.752 | |
| high | | | | | 1.361 | ** | | | | | 1.209 | ** |
| t1 | | | 0.265 | | 0.260 | | | | 0.290 | | 0.295 | |
| t2 | | | | | 0.525 | | | | | | 0.57 | |
| p-value (bootstrap) | | | 0.077 | | 0.721 | | | | 0.013 | | 0.504 | |

**PANEL A: 10-year averages**

| | PC21 | | | | | | PC22 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Linear model | | 2-regime Model | | 3-regime model | | Linear model | | 2-regime Model | | 3-regime model | |
| Capita | -12.730 | *** | -11.263 | *** | -5.654 | *** | -10.793 | *** | -11.852 | *** | -5.701 | *** |
| life_exp | 0.275 | | 0.410 | * | 2.255 | ** | 0.590 | *** | 0.434 | *** | 3.155 | *** |
| opengs | -0.033 | | -0.041 | | -1.701 | | -0.090 | *** | -0.063 | *** | -2.823 | *** |
| ig_gdp | -0.148 | | -0.108 | | -0.392 | | -0.422 | | -0.210 | | -1.243 | |
| i_price | 0.081 | * | 0.063 | | 1.512 | | 0.035 | | 0.053 | | 1.287 | |
| Infrastr | | | | | | | | | | | | |
| Low | 4.380 | ** | 2.230 | | 1.119 | | 0.588 | | 3.459 | ** | 1.912 | * |
| High/mid | | | 4.004 | ** | 4.129 | *** | | | 0.834 | *** | 4.249 | *** |
| high | | | | | 3.311 | ** | | | | | 0.678 | *** |
| t1 | | | 0.405 | | 0.405 | | | | 0.650 | | 0.655 | |
| t2 | | | | | 0.745 | | | | | | 0.340 | |
| p-value (bootstrap) | | | 0.026 | | 0.452 | | | | 0.051 | | 0.008 | |

| | PC23 | | | | | | PC24 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Linear model | | 2-regime model | | 3-regime Model | | Linear model | | 2-regime Model | | 3-regime model | |
| Capita | -11.554 | *** | -12.259 | *** | -5.719 | *** | -11.583 | *** | -12.408 | *** | -6.289 | *** |
| life_exp | 0.454 | *** | 0.439 | *** | 3.342 | *** | 0.401 | *** | 0.384 | *** | 2.505 | ** |
| opengs | -0.070 | *** | -0.067 | *** | -4.111 | *** | -0.055 | *** | -0.048 | ** | -2.267 | ** |
| ig_gdp | -0.108 | | -0.232 | | -0.670 | | 0.082 | | 0.019 | | -0.064 | |
| i_price | 0.051 | | 0.037 | | 0.968 | | 0.052 | | 0.042 | | 1.021 | |
| Infrastr | | | | | | | | | | | | |
| Low | 2.823 | ** | 3.718 | *** | 1.804 | * | 3.201 | *** | 3.925 | *** | 2.123 | ** |
| High/mid | | | 2.765 | ** | 1.644 | | | | 3.329 | *** | 2.954 | |
| high | | | | | 2.650 | ** | | | | | 3.267 | *** |
| t1 | | | 0.255 | | 0.430 | | | | 0.295 | | 0.300 | |
| t2 | | | | | 0.715 | | | | | | 0.640 | |
| p-value (bootstrap) | | | 0.048 | | 0.116 | | | | 0.186 | | 0.768 | |

Notes: low, high/mid, high are the lower regimes, the upper regimes (2-regime model)/the middle-regime (3-regime model), and the upper regime for the 3-regime model. T1 and t2 are the two thresholds. PC21: based on principal components using energy consumption (World Bank), PC22: based on principal components using energy production (World Bank), PC23: based on principal components using electricity generation (IEA), PC24: based on principal components using total electricity generating capacity (IEA).

## Table A.13. Results of model averaging – nonlinearity of the infrastructure variable

### ($1^{st}$ principal component) – 8 year averages

| | Inclus. Prob. | mean | s.e | Inclus. Prob. | mean | s.e | Inclus. Prob. | mean | s.e | Inclus. Prob. | mean | s.e. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Cappp | **1.000** | -6.750 | 1.993 | **1.000** | -6.150 | 1.273 | **1.000** | -7.412 | 2.286 | **1.000** | -6.241 | 2.534 |
| life_exp | 0.488 | 0.166 | 0.095 | 0.045 | 0.012 | 0.010 | 0.464 | 0.200 | 0.111 | **0.687** | 0.229 | 0.112 |
| Cpi | 0.404 | 0.009 | 0.017 | 0.424 | -0.003 | 0.012 | 0.447 | -0.005 | 0.011 | **0.533** | 0.044 | 0.038 |
| Reg | 0.112 | 0.000 | 0.021 | 0.140 | -0.006 | 0.029 | 0.140 | -0.019 | 0.028 | 0.111 | -0.004 | 0.021 |
| dpop1564 | 0.123 | -0.010 | 0.062 | 0.163 | -0.046 | 0.080 | 0.103 | -0.003 | 0.061 | 0.110 | -0.001 | 0.056 |
| Opengs | **0.596** | 0.025 | 0.016 | 0.215 | 0.007 | 0.006 | 0.317 | 0.012 | 0.009 | **0.837** | 0.041 | 0.022 |
| ig_gdp | **0.546** | -0.186 | 0.094 | **0.966** | -0.359 | 0.117 | **0.564** | -0.194 | 0.093 | 0.298 | -0.071 | 0.055 |
| i_price | **0.642** | -0.037 | 0.021 | **0.626** | -0.027 | 0.015 | 0.454 | -0.018 | 0.012 | **0.788** | -0.081 | 0.044 |
| edu2 | **1.000** | 0.975 | 0.260 | **0.998** | 0.913 | 0.288 | **0.953** | 0.870 | 0.309 | **0.997** | 0.993 | 0.254 |
| **Infrastr** | | **Pc11** | | | **Pc12** | | | **Pc13** | | | **Pc14** | |
| Linear | 0.000 | 0.000 | 0.000 | 0.032 | -0.013 | 0.055 | 0.145 | 0.024 | 0.261 | 0.001 | 0.000 | 0.002 |
| 2-regime | | | | | | | | | | | | |
| lower regime | **0.636** | 0.363 | 1.416 | **0.966** | 1.116 | 0.941 | **0.809** | 0.539 | 1.577 | **0.996** | -2.709 | 22.868 |
| higher regime | | 0.110 | 1.157 | | -0.147 | 1.715 | | 0.245 | 1.231 | | -0.554 | 3.428 |
| 3-regime | | | | | | | | | | | | |
| low regime | 0.363 | -1.224 | 5.619 | 0.002 | -0.002 | 0.007 | 0.045 | 0.008 | 0.174 | 0.003 | -0.006 | 0.022 |
| middle regime | | -0.403 | 1.221 | | 0.002 | 0.005 | | -0.033 | 0.089 | | -0.004 | 0.007 |
| high regime | | 0.094 | 1.232 | | -0.006 | 0.008 | | 0.068 | 0.173 | | 0.002 | 0.016 |

Notes: PC11: based on principal components using energy consumption (World Bank), PC12: based on principal components using energy production (World Bank), PC13: based on principal components using electricity generation (IEA), PC14: based on principal components using total electricity generating capacity (IEA).

## Table A.14. Results of model averaging – nonlinearity of the infrastructure variable

### ($2^{nd}$ principal component) – 8 year averages

| | Inclusion probability | mean | s.e. | Inclusion probability | mean | s.e. | Inclusion probability | mean | s.e. | Inclusion probability | mean | s.e. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| cappp | **1.000** | -9.480 | 1.350 | **1.000** | -7.463 | 2.255 | **1.000** | -8.947 | 1.589 | **1.000** | -9.337 | 1.481 |
| life_exp | 0.231 | -0.042 | 0.049 | 0.356 | 0.105 | 0.075 | 0.112 | -0.008 | 0.024 | 0.116 | -0.001 | 0.022 |
| cpi | 0.448 | -0.010 | 0.008 | 0.348 | 0.002 | 0.013 | 0.441 | -0.009 | 0.009 | 0.473 | -0.011 | 0.008 |
| reg | 0.119 | -0.011 | 0.024 | 0.106 | -0.011 | 0.022 | 0.124 | -0.015 | 0.028 | 0.105 | -0.003 | 0.023 |
| dpop1564 | 0.183 | -0.064 | 0.101 | 0.219 | -0.072 | 0.138 | 0.104 | -0.007 | 0.047 | 0.153 | -0.042 | 0.084 |
| opengs | **0.935** | 0.041 | 0.020 | **0.812** | 0.041 | 0.022 | **0.705** | 0.026 | 0.017 | 0.375 | 0.010 | 0.009 |
| ig_gdp | **0.979** | -0.308 | 0.108 | 0.382 | -0.096 | 0.063 | **0.986** | -0.317 | 0.110 | **0.975** | -0.317 | 0.110 |
| i_price | 0.405 | -0.008 | 0.009 | 0.439 | -0.022 | 0.015 | 0.472 | -0.013 | 0.010 | 0.453 | -0.011 | 0.009 |
| edu2 | **1.000** | 0.853 | 0.239 | **0.994** | 0.941 | 0.297 | **0.998** | 0.834 | 0.254 | **1.000** | 0.892 | 0.246 |
| **infrastr** | | **Pc21** | | | **Pc22** | | | **Pc23** | | | **Pc24** | |
| linear | 0.000 | 0.000 | 0.000 | **0.586** | 0.085 | 0.241 | 0.002 | 0.002 | 0.002 | 0.000 | 0.000 | 0.000 |
| 2-regime | | | | | | | | | | | | |
| lower regime | **1.000** | 2.633 | 4.501 | 0.412 | 0.643 | 0.557 | **0.998** | 2.705 | 3.750 | **1.000** | 2.596 | 4.088 |
| higher regime | | 1.524 | 1.511 | | 0.119 | 0.168 | | 1.312 | 1.111 | | 1.236 | 1.041 |
| 3-regime | | | | | | | | | | | | |
| low regime | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| middle regime | | 0.000 | 0.000 | | 0.000 | 0.000 | | 0.000 | 0.000 | | 0.000 | 0.000 |
| high regime | | 0.000 | 0.000 | | 0.000 | 0.000 | | 0.000 | 0.000 | | 0.000 | 0.000 |

Notes: PC21: based on principal components using energy consumption (World Bank), PC22: based on principal components using energy production (World Bank), PC23: based on principal components using electricity generation (IEA), PC24: based on principal components using total electricity generating capacity (IEA).

## Table A.15. Results of model averaging – nonlinearity of the infrastructure variable

(1st principal component) – 10-year averages

| | Inclus. Prob. | mean | s.e | Inclus. Prob. | mean | s.e | Inclus. Prob. | mean | s.e | Inclus. Prob. | mean | s.e. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| cappp | **0.956** | -5.337 | 4.277 | **1.000** | -11.358 | 6.505 | **1.000** | -10.609 | 2.359 | **1.000** | -7.848 | 3.666 |
| life_exp | **0.955** | 0.481 | 0.175 | **0.986** | 0.652 | 0.165 | **1.000** | 0.698 | 0.201 | **0.854** | 0.370 | 0.163 |
| cpi | 0.255 | 0.001 | 0.021 | **0.757** | -0.072 | 0.014 | **0.811** | -0.091 | 0.023 | 0.456 | 0.055 | 0.041 |
| reg | **0.963** | -0.687 | 0.253 | **0.715** | -0.552 | 0.350 | **0.999** | -0.998 | 0.352 | **1.000** | -0.990 | 0.179 |
| dpop1564 | 0.285 | 0.187 | 0.300 | 0.052 | 0.017 | 0.067 | 0.249 | 0.234 | 0.338 | 0.233 | 0.112 | 0.169 |
| opengs | **1.000** | 0.097 | 0.021 | **0.596** | 0.025 | 0.018 | **0.829** | 0.060 | 0.028 | **1.000** | 0.125 | 0.024 |
| ig_gdp | **0.997** | 0.608 | 0.164 | **0.688** | 0.251 | 0.185 | **0.994** | 0.482 | 0.206 | **1.000** | 0.537 | 0.135 |
| i_price | **0.933** | -0.176 | 0.039 | 0.285 | -0.032 | 0.013 | 0.276 | -0.030 | 0.020 | **0.999** | -0.240 | 0.060 |
| edu2 | **1.000** | 0.913 | 0.385 | **0.728** | 0.441 | 0.449 | **0.998** | 0.748 | 0.372 | **0.999** | 0.654 | 0.387 |
| **infrastr** | | **Pc11** | | | **Pc12** | | | **Pc13** | | | **Pc14** | |
| linear | 0.000 | 0.000 | 0.000 | 0.001 | 0.001 | 0.004 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| 2-regime | | | | | | | | | | | | |
| lower regime | 0.000 | -0.001 | 0.001 | 0.063 | -0.079 | 0.173 | 0.011 | -0.013 | 0.049 | 0.001 | -0.008 | 0.007 |
| higher regime | | 0.000 | 0.000 | | 0.028 | 0.201 | | 0.009 | 0.038 | | 0.000 | 0.004 |
| 3-regime | | | | | | | | | | | | |
| low regime | **1.000** | -12.035 | 226.455 | **0.929** | -3.227 | 22.627 | **0.989** | -3.556 | 30.319 | **0.999** | -8.647 | 121.087 |
| middle regime | | -0.278 | 4.218 | | 1.706 | 1.860 | | 0.817 | 1.735 | | 0.622 | 1.858 |
| high regime | | -6.462 | 83.365 | | -4.534 | 41.751 | | -1.226 | 7.814 | | -1.826 | 12.581 |

Notes: PC11: based on principal components using energy consumption (World Bank), PC12: based on principal components using energy production (World Bank), PC13: based on principal components using electricity generation (IEA), PC14: based on principal components using total electricity generating capacity (IEA).

## Table A.16. results of model averaging – nonlinearity of the infrastructure variable

(2nd principal component) – 10-year averages

| | Inclusion probability | mean | s.e. | Inclusion probability | mean | s.e. | Inclusion probability | mean | s.e. | Inclusion probability | mean | s.e. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| cappp | **1.000** | -12.357 | 3.057 | **1.000** | -17.769 | 2.125 | **1.000** | -13.805 | 3.238 | **1.000** | -13.074 | 2.574 |
| life_exp | **0.990** | 0.502 | 0.263 | **1.000** | 1.298 | 0.226 | **0.998** | 0.626 | 0.183 | **0.973** | 0.474 | 0.198 |
| cpi | **0.868** | -0.111 | 0.080 | **1.000** | -0.443 | 0.074 | **0.972** | -0.117 | 0.055 | **0.916** | -0.083 | 0.050 |
| reg | 0.139 | -0.008 | 0.031 | **0.999** | -0.980 | 0.272 | 0.266 | -0.093 | 0.131 | 0.128 | -0.006 | 0.043 |
| dpop1564 | 0.176 | -0.034 | 0.165 | **0.988** | 1.672 | 0.687 | 0.152 | 0.091 | 0.209 | 0.195 | 0.101 | 0.248 |
| opengs | **0.956** | 0.063 | 0.034 | **0.999** | 0.063 | 0.022 | 0.120 | 0.005 | 0.005 | 0.366 | 0.019 | 0.012 |
| ig_gdp | 0.142 | -0.009 | 0.026 | **0.998** | 0.423 | 0.163 | 0.225 | 0.011 | 0.045 | 0.158 | -0.004 | 0.028 |
| i_price | **0.684** | 0.093 | 0.091 | **1.000** | 0.440 | 0.085 | 0.411 | 0.051 | 0.054 | 0.426 | 0.042 | 0.047 |
| edu2 | 0.208 | 0.039 | 0.068 | **0.999** | 0.743 | 0.313 | 0.125 | 0.023 | 0.039 | 0.405 | 0.108 | 0.113 |
| **infrastr** | | **Pc21** | | | **Pc22** | | | **Pc23** | | | **Pc24** | |
| linear | 0.002 | 0.013 | 0.004 | 0.000 | 0.000 | 0.000 | 0.022 | 0.066 | 0.054 | 0.095 | 0.294 | 0.181 |
| 2-regime | | | | | | | | | | | | |
| lower regime | **0.996** | 2.764 | 2.845 | 0.000 | 0.000 | 0.000 | 0.043 | 0.164 | 0.135 | **0.898** | 3.585 | 8.343 |
| higher regime | | 4.567 | 12.354 | | 0.000 | 0.000 | | 0.115 | 0.124 | | 2.910 | 5.477 |
| 3-regime | | | | | | | | | | | | |
| low regime | 0.002 | 0.002 | 0.009 | **1.000** | 3.462 | 6.032 | **0.935** | 3.775 | 5.737 | 0.007 | 0.027 | 0.020 |
| middle regime | | 0.004 | 0.007 | | 0.354 | 0.248 | | 2.375 | 2.946 | | 0.021 | 0.012 |
| high regime | | 0.000 | 0.009 | | 6.647 | 28.199 | | 1.254 | 1.429 | | 0.016 | 0.026 |

Notes: PC21: based on principal components using energy consumption (World Bank), PC22: based on principal components using energy production (World Bank), PC23: based on principal components using electricity generation (IEA), PC24: based on principal components using total electricity generating capacity (IEA).

Figure A.1.. Distribution of inclusion probability, 1st principal component

**10-year averages**

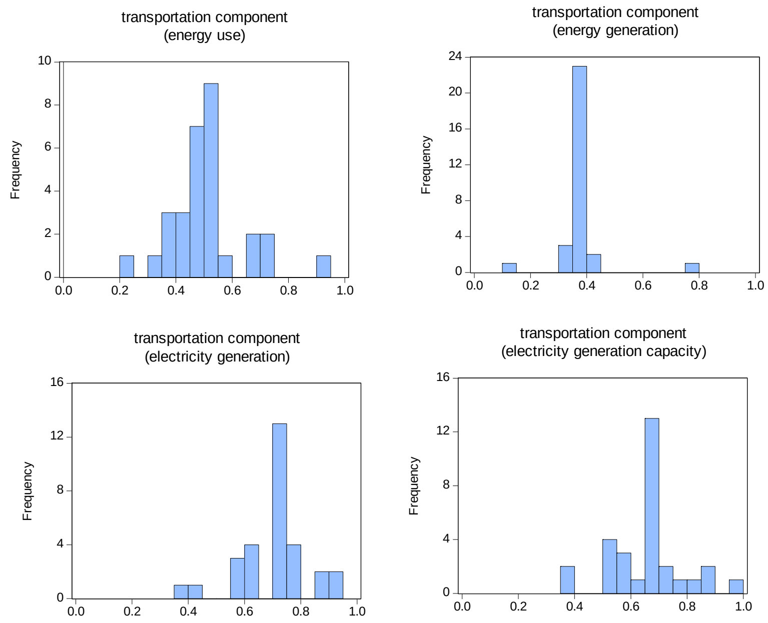
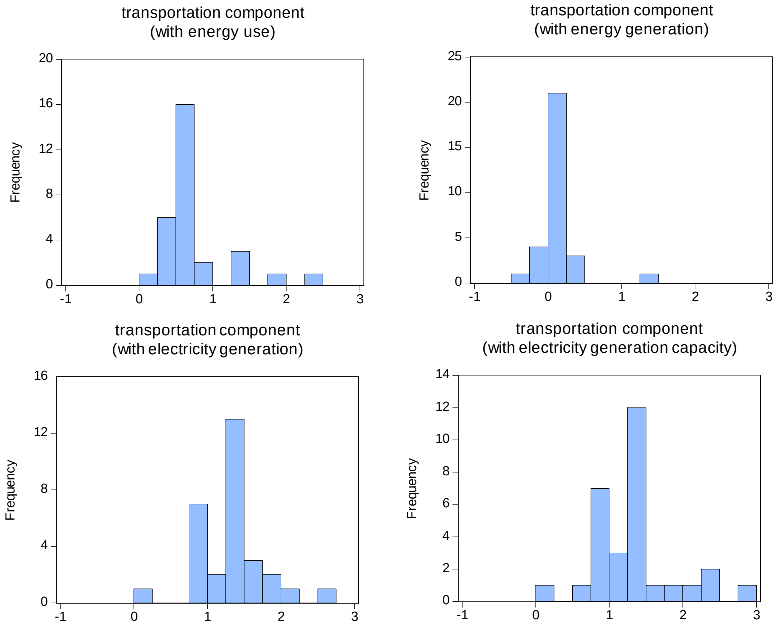


Figure 1B. Distribution of posterior mean conditional on inclusion, 1st principal component

**10-year averages**



Note: The distribution of the inclusion probability shows the distribution of the models (full sample + subsamples excluding one country at a time) in terms of inclusion probabilities. The vertical axis shows the number of models for which the values of the inclusion probability are shown on the horizontal axis are estimated.

Note: The distribution of the posterior mean conditional on inclusion shows the distribution of the models (full sample + subsamples excluding one country at a time) in terms of the estimated means on variants of the principal component. The vertical axis shows the number of models for which the values of the estimated mean are shown on the horizontal axis are estimated.

Figure A.2. Distribution of inclusion probability, 2nd principal component

**10-year averages**

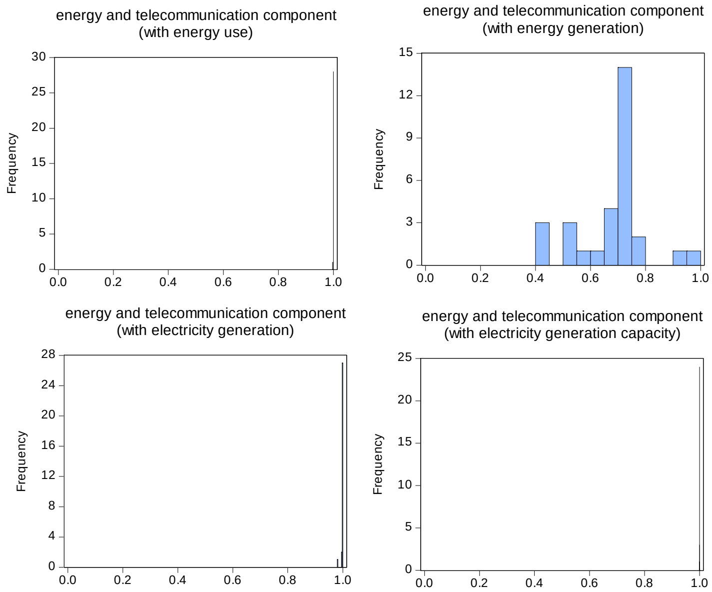
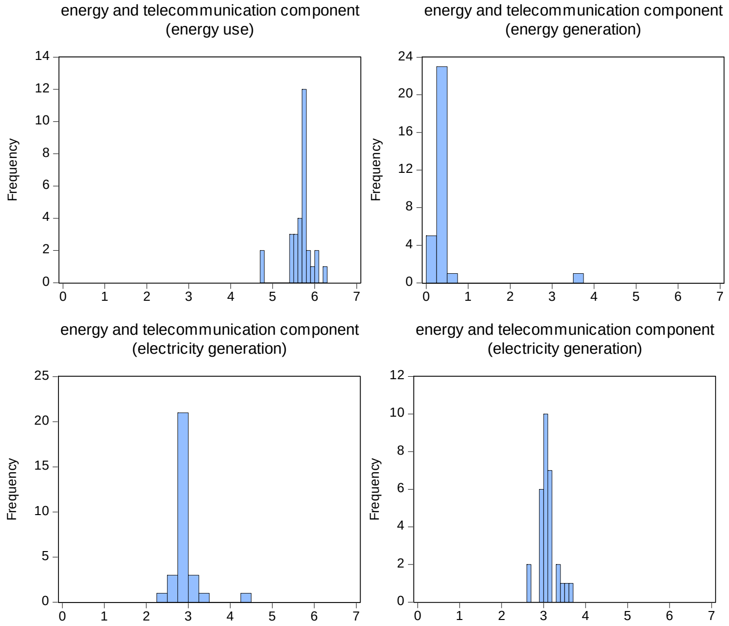


Figure 2B. Distribution of posterior mean conditional on inclusion, 2nd principal component



Note: The distribution of the inclusion probability shows the distribution of the models (full sample + subsamples excluding one country at a time) in terms of inclusion probabilities. The vertical axis shows the number of models for which the values of the inclusion probability are shown on the horizontal axis are estimated.

Note: The distribution of the posterior mean conditional on inclusion shows the distribution of the models (full sample + subsamples excluding one country at a time) in terms of the estimated means on variants of the principal component. The vertical axis shows the number of models for which the values of the estimated mean are shown on the horizontal axis are estimated.

# DAVIDSON INSTITUTE WORKING PAPER SERIES - Most Recent Papers

The entire Working Paper Series may be downloaded free of charge at: www.wdi.umich.edu

CURRENT AS OF 4/17/09

| Publication | Authors | Date |
|---|---|---|
| *No. 957: Infrastructure and growth: Empirical evidence* | Balazs Egert, Tomasz Kozluk and Douglas Sutherland | April 2009 |
| *No. 956: Infrastructure investment in network industries: The role of incentive regulation and regulatory independence* | Balazs Egert | April 2009 |
| *No. 955: The impact of monetary and commodity fundamentals, macro news and central bank communication on the exchange rate: Evidence from South Africa* | Balazs Egert | April 2009 |
| *No. 954: "Family" ownership, tunneling and earnings management: A review of the literature* | Sumon Bhaumik and Andros Gregoriou | March 2009 |
| *No. 953: Impact of Foreign Direct Investments on Industrial Productivity: A Subnational Study of India* | Krishna C. Vadlamannati | March 2009 |
| *No. 952: Determinants of Exchange Rate Practices in the MENA Countries: Some Further Empirical Results* | Sfia, M. Daly and Mouley Sami | Jan 2009 |
| *No. 951: An Analytic Approach To Selecting A Nonprofit* | Andrés Ramírez and Hakan Saraoglu | Jan 2009 |
| *No. 950: Anthropometry of Love Height and Gender Asymmetries in Interethnic Marriages* | Michèle Belot and Jan Fidrmuc | Jan 2009 |
| *No. 949: IS THE COST OF LIVING IN RUSSIA REALLY THAT LOW?* | Konstantin Gluschenko | Dec 2008 |
| *No. 948: Banking Market Liberalization and Bank Performance: the Role of Entry Modes* | Ngoc-Anh Vo Thi | Jan 2009 |
| *No. 947: The Monetary Union: The Decade Ahead. The Case of Non-Member States* | Daniel Daianu and Laurian Lungu | Jan 2009 |
| *No. 946: Currency Substitution: A Case Of Kazakhstan (2000:1-2007:12)* | Mesut Yilmaz, Yessengali Oskenbayev & Kanat Abdulla | Jan 2009 |
| *No. 945: Determinants Of Pollution Abatement And Control Expenditure: Evidence From Romania* | Guglielmo Caporale, Christophe Rault, Robert Sova & Anamaria Sova | Jan 2009 |
| *No. 944: Bootstrap panel Granger-causality between government spending and revenue in the EU* | Antonio Afonso and Christophe Rault | Jan 2009 |
| *No. 943: Regional Determinants of FDI Distribution in Poland* | Agnieszka Chidlow and Stephen Young | Nov 2008 |
| *No. 942: Structural Reform and Firm Exports* | Alvaro Cuervo-Cazurra & Luis Alfonso Dau | Sept 2008 |
| *No. 941: Exploring The Relationship Between Military Spending & Human Rights Performance In South Asia* | Krishna Chaitanya, Vadlamannati and K K Shakya Lahiru Pathmalal | Oct 2008 |
| *No. 940: Structural Reform And Firm Profitability In Developing Countries* | Alvaro Cuervo-Cazurra & Luis Alfonso Dau | July 2008 |
| *No. 939: Does Timing Of Elections Instigate Riots? A Subnational Study Of 16 Indian States, 1958 – 2004* | Krishna Chaitanya Vadlamannati | Oct 2008 |
| *No. 938: Price Setting And Market Structure: An Empirical Analysis Of Micro Data* | Fabrizio Coricelli & Roman Horvath | Sept 2008 |
| *No. 937: Inflation Differentials in EU New Member States: An Empirical Evidence* | Roman Horvath & Kamila Koprnicka | Oct 2008 |
| *No. 936: Market Reforms and Growth in Post-socialist Economies: Evidence from Panel Cointegration and Equilibrium Correction Model* | Igor Pelipas and Alexander Chubrik | Sept 2008 |
| *No. 935: Derivatives trading and the volume-volatility link in the Indian Stock Market* | S. Bhaumiky, M. Karanasosy and A. Kartsaklas | April 2008 |
| *No. 934: Yuan Real Exchange Rate Undervaluation, 1997-2006. How Much, How Often? Not Much, Not Often* | Jeff Chen, Wende Deng & David Kemme | Aug 2008 |