Infrastructure, Institutions and Economic Growth: A Time-Series Study of the South African Economy

Robert J. Garlick

A thesis submitted for the fulfillment of the degree of Master of Commerce (Economics)

Supervised by Prof. Johan Fedderke

Plagiarism Declaration

- I know that plagiarism is wrong. Plagiarism is to use another's work and pretend that it is one's own.
- I have used the Harvard convention for citation and referencing. Each contribution to and quotation in this dissertation from the work(s) of other people has been cited and referenced.
- This dissertation is my own work.
- I acknowledge that copying someone else's assignment or essay, or part thereof, is wrong, and declare that this is my own work.

Abstract

A growing empirical has analysed the historical relationship between infrastructure and output in South Africa, finding a broadly positive effect that operates largely *via* the marginal productivity of private capital. We extend this literature by investigating the relationship between infrastructure, output and institutional quality: protection of property rights, political fractionation and political and economic risk. We develop a model in the spirit of Barro (1990), which predicts a nonlinear relationship between infrastructure and output (positive at low levels of infrastructure and subsequently negative) and a positive effect of institutional capital on both output and the response of output to changes in infrastructure stock. We test this model using both univariate and multivariate cointegration models, finding support for the predicted nonlinear relationship and mixed evidence with respect to all other predictions. We conclude that the omission of institutional measures may have biased prior empirical analyses but that we are unable to generate more robust findings with the short available time series.

Contents

1 Introduction

The Accelerated and Shared Growth Initiative for South Africa (ASGI-SA) has identified infrastructure spending in South Africa as a key driver of sustained economic growth and job creation, stating that:

> "Backlogs in infrastructure and investment, and in some cases market structures that do not encourage competition, make the price of moving goods and conveying services over distance higher than it should be. Deficiencies in logistics are keenly felt in a country of South Africa's size, with considerable concentration of production inland, and which is some distance from the major industrial markets (South Africa, 2007a)."

Reflecting this emphasis, the National Treasury has allocated R416 billion to spending on infrastructure development and maintenance, broadly defined, in the current three-year budget cycle (South Africa, 2007b). This follows a period from 1976 to 2002 in which annual infrastructure investment fell from 8.1% to 2.6% of GDP, with per capita per annum expenditure falling from R1268 to R356 (Fedderke and Bogetić, 2006a).

The relationship between infrastructure and economic growth is thus of crucial importance to South Africa and to sub-Saharan Africa more broadly. This relationship has been explored in a number of recent papers using time-series and panel data on the South African economy (Perkins, *et al*, 2005; Fedderke, *et al*, 2006; Fedderke and Bogetić, 2006 and Kularatne, 2006). At the same time, another strand of the empirical literature on South African growth has examined the relationship between growth and institutions such as property rights and political risk (Fielding, 1997; Mariotti, 2002; Kularatne, 2002 and Fedderke, 2004).

These two strands of the literature have, however, proceeded largely in isolation in South Africa. In contrast, the international academic and policy literature has begun to place significant emphasis on institutional dimensions of the relationship between infrastructure spending and output: both Guasch (2004) and Kessides (2004) emphasize that institutions play a crucial role in determining the type of infrastructure projects undertaken and the ability of the private sector to take advantage of those projects. Only Fedderke, *et al* (2006), and Kularatne (2006) briefly explore this three-way interaction in the South African context.

This paper aims to determine the implications for the infrastructure-output relationship of taking explicit account of the institutional factors affecting both infrastructure and output. Analysis proceeds through the development of a theoretical model in the spirit of Barro (1990) and the testing of this model against time-series data on the South African economy between 1960 and 2006.

Section 2 very briefly reviews the most relevant international and South African literature. Section 3 develops the theoretical model of the interaction of infrastructure, institutions and economic growth and section 4 introduces the data used to test these models. Section 5 reports the estimation results, with subsections 5.2 and 5.3 reporting the univariate and multivariate estimates, respectively. Section 6 concludes. The appendix provides more details on construction of the institutional variables.

2 Literature Review

A comprehensive review of the existing literature on infrastructure and output falls beyond the scope of this paper, so comment is limited to a brief, approximately chronological discussion of the development of key international and domestic findings. Emphasis is placed on understanding the way in which this literature might be limited by methodological considerations and specifically by its omission of the potential role of institutions in shaping the infrastructure-output relationship. For a comprehensive discussion of the South African infrastructure-output literature see Fedderke and Garlick (2007).

Empirical investigation of the relationship between infrastructure and output dates back to the late 1980s (Aschauer, 1989a; Aschauer, 1989b; Aschauer, 1989c; Munnell, 1990), with most studies focusing only on the developed world and finding significant, positive growth elasticities of infrastructure spending. The South African literature followed suit, finding growth elasticities of infrastructure spending between 0.17 and 0.33 (Abedian and Van Seventer, 1995; Coetzee and Le Roux, 1998; Development Bank of Southern Africa, 1998). These initial findings were, however, strongly criticised in subsequent literature on numerous fronts:

- applying least-squares estimators to non-stationary time series;
- assuming a linear relationship between infrastructure and output;
- ignoring the possibility of an indirect relationship between infrastructure and output *via*, for example, changes in the marginal productivity of capital or labour;
- failing to control for other determinants of economic growth;
- employing highly aggregated data which ignored region-, sector- and infrastructure-specific characteristics; and
- failing to control for the possibility of simultaneity between infrastructure and output.

These problems are reported in more depth in Fedderke and Bogetić (2006), Fedderke and Garlick (2007) and Gramlich (1994). The latter deals with the international literature and the former two with the South African literature.

Subsequent research has attempted to address specifically these early problems and has produced a significantly more nuanced picture of the relationship between infrastructure and output. Perkins, *et al*, (2005) broadened the

South African literature on two levels. Firstly, the paper used both financial (investment during a period or accumulated capital stock) and physical (kilometres of road, number of fixed-line telephones, number of passengers departing from an airport during a period, etc.) measures of infrastructure. While both are "noisy" measures of actual investment - the former approach implicitly assumes a uniform "infrastructure output" for each unit of currency invested and the latter approach measures only the quantity of infrastructure, not the quality - this was nonetheless an improvement on the early literature's sole focus on financial measures. Secondly, they specifically examined the direction of association between infrastructure and output. Using a PSS ARDL estimation, they find that aggregate infrastructure expenditure, aggregate infrastructure capital stock and total road length drive gross domestic product, while other physical measures of output are either driven by or ambiguously related to GDP.

This finding is replicated by Fedderke, *et al*, (2006) who also use a Johansen VECM structure to estimate the long-run relationship between infrastructure and output. They conclude that there is a significant positive relationship but that it is indirect, operating through the relationship between infrastructure stock and fixed capital stock (elasticity 1.37) and between fixed capital stock and GDP (elasticity 0.06). Their findings are, however, highly sensitive to the use of different physical measures of infrastructure stock.

These papers all focused exclusively on the relationship between output and *physical infrastucture*, which includes transport, communications and power generation infrastructure. The relationship between output and *social infrastructure*, which includes health, education and sanitation infrastructure,[1] has received little attention in the South African literature. A notable exception is Kularatne (2006), which explores the interaction between social infrastructure, physical infrastructure, private investment and gross value added.

[1]This distinction is widely used in the literature, both empirical and theoretical. The distinction is not always, however, entirely clear: water supply and sanitation infrastructure, for example, may be classified as either economic or social infrastructure.

Using a PSS ARDL model, he finds that social infrastructure drives all of the other three variables but that the direction of association between the other variables is ambiguous. A VECM structure estimates a value-added elasticity of 0.06 with respect to social infrastructure investment and of 2.5 with respect to private investment (which in turn has an elasticity of 0.02 with respect to physical infrastructure investment).

Several other papers examine the indirect relationship between infrastructure and output through the role of infrastructure in driving other determinants of economic growth. Fedderke and Bogetić (2006) focus on the relationship between infrastructure and both labour and total factor productivity, using instrumental variables to take cognisance of the possibility of simultaneity between infrastructure and output. They find that aggregate infrastructure investment and stock impact on labour productivity with elasticities 0.2 and 0.19, respectively, and impact total factor productivity with elasticties of 0.04 and approximately zero. These results are qualitatively robust to replacing financial measures of infrastructure with physical measures, though the elasticities are substantially larger for measures of the extensiveness of the road and rail networks and substantially smaller for other forms of physical infrastructure.

Alves and Edwards (2005) and Edwards and Johnny (2006) examine the relationship between infrastructure and export performance. PSS ARDL analysis suggests that infrastructure stock drives export performance, rather than *vice versa* and firm-level surveys indicate that infrastructure capacity positively impacts on firms' decisions to enter the export market.

The bulk of these results, however, do not take cognisance of the potential for institutional variables to affect the relationship between infrastructure and output. Fedderke, *et al* (2006), augment their baseline model to include measures of property rights and political instability and find that the direct relationship between infrastructure stock and GDP becomes positive and significant with this control (elasticity 0.4 - 0.5, depending on the specification).

The indirect relationship between infrastructure and output *via* investment is largely unaffected. These results are, however, not robust to changes in the model specification. Kularatne (2006), controlling for political instability, finds that only the indirect relationship between infrastructure and gross value added *via* investment is significant (investment elasticity of 0.02 with respect to infrastructure and GVA elasticity of 0.08 with respect to investment) for physical infrastructure but finds that social infrastructure impacts directly on GVA (elasticity 0.06).

The international literature is surprisingly sparse in this area: many papers acknowledge the high *a priori* probability that the infrastructure-output relationship is affected by institutions but there are few empirical tests of this hypothesis. Esfahani and Ramírez (2002), one exception to this pattern, experiment with the inclusion of indicators of democracy and ethnolinguistic fractionation in a cross-country growth regression and find that the effect of these variables on the estimated relationship between output and telecommunications and power generation capital is ambiguous.

3 Theory

3.1 A simple growth model

In line with several previous studies of infrastructure-output relationship in South Africa, we develop a growth model in the spirit of Barro (1990), which readily lends itself to the inclusion of multiple forms of capital. We assume a closed economy with no technological progress and no labour-leisure choice (i.e. inelastic labour supply). We assume a production function

$$y = F(\vec{K}) \tag{1}$$

where $\vec{K}$ is a vector of all forms of capital and the production function is increasing and concave in each individual argument (i.e. $F_i > 0$ and $F_{ii} < 0$) but has constant returns to scale in all arguments. Both y and $\vec{K}$ are in per capita terms. We assume that $\vec{K}$ includes only one reproducible form of capital, private physical capital, which we denote as k. (Though it may include other, non-reproducible forms of capital.) We define y as output net of depreciation of private physical capital so that k is governed by the equation of motion

$$\dot{k} = F(\vec{K}) - c. \tag{2}$$

Household preferences are given by

$$U = \int_0^\infty e^{-\rho t} \left(\frac{c^{1-\theta} - 1}{1 - \theta} \right) dt \tag{3}$$

and we normalise the number of households in the economy to unity.

Assuming that the representative household is a utility maximiser, it faces a problem of optimal control with the present-value Hamiltonian equation

$$H = e^{-\rho t} \cdot \frac{c^{1-\theta} - 1}{1 - \theta} + \lambda \cdot \left[F(\vec{K}) - c \right].$$

Treating c as a control variable and $\vec{K}$ as a vector of state variables, Pontryagin's Maximum Principle yields the first-order conditions

$$e^{-\rho t} \cdot c^{-\theta} = \lambda$$

and

$$-\dot{\lambda} = \lambda \cdot F_k(\vec{K})$$

As the second differential equation is potentially non-linear in some elements of $\vec{K}$, we cannot directly obtain a closed form solution to this system of equations. Instead, we log-transform both sides of the first equation and then differentiate with respect to time, yielding

$$-\rho - \theta \cdot \left(\frac{\dot{c}}{c} \right) = \frac{\dot{\lambda}}{\lambda}.$$

Substituting this into the second equation yields

$$-\rho - \theta \cdot \left(\frac{\dot{c}}{c} \right) = -F_k(\vec{K})$$

so the Euler equation for consumption is

$$\frac{\dot{c}}{c} = \frac{1}{\theta} \cdot \left[F_k(\vec{K}) - \rho \right]. \tag{4}$$

As F has constant returns to scale in all its arguments, F_k is homogenous of degree zero (using Euler's formula) and so $F_k(\vec{K})$ is a constant. Thus, the growth rate of consumption is constant and we denote it by η_1.[2]

If $\vec{K}$ is a one-dimensional vector (i.e. private physical capital is the only form of capital in the economy), then the production function can be written as

$$y = Ak$$

and log-transforming both sides of this equation and differentiating with respect to time yields

$$\frac{\dot{y}}{y} = \frac{\dot{A}}{A} + \frac{\dot{k}}{k}.$$

Under the assumption of no technological progress, $\dot{A} = 0$ and so the growth rates of output and capital are equal. Given that the growth rate of consumption is η_1, both output and capital must then also grow at the constant rate η_1. Thus, the economy is always in a steady state with constant growth of output, capital and consumption, all in per capita terms. Given an initial capital stock, we can use equations (1) and (2) to determine the complete time paths of all three variables.

3.2 A model of infrastructure and growth

We now adapt the model above to include a role for government expenditure. We assume that government purchases output from the private sector to provide services that are qualitatively different from those provided by

[2]In order to ensure positive but finite consumption growth we must impose the additional conditions that $F_k(\vec{K}) > \rho > (1 - \theta) \cdot F_k(\vec{K})$.

the private sector.[3] Given the focus of this paper, we assume that government expenditure is devoted entirely to infrastructure spending and that the private sector provides no infrastructure services (the latter assumption is broadly consistent with South Africa's historical experience). In order to continue to use variables in per-capita terms, we must assume that there are no congestion effects on government expenditure.

Production technology is now

$$y = F(g, k)$$

where g is infrastructure expenditure per capita. To clarify our exposition, we assume that production technology is Cobb-Douglas

$$y = A \cdot g^{\alpha} \cdot k^{1-\alpha}$$

though the analysis holds for any constant elasticity of substitution production technology (which includes linear and Leontief technology). We further assume that infrastructure expenditure is funded by a lump-sum income tax $\tau = \frac{g}{y}$ which is constant through time, so that g and y grow at the same rate.

Assuming that the representative household's preferences are still given by equation (3), we can repeat the optimal control analysis above with the new production function and obtain the revised Euler equation for consumption

$$\frac{\dot{c}}{c} = \frac{1}{\theta} \cdot \left[(1-\tau) \cdot (1-\alpha) \cdot A \cdot \left(\frac{g}{k}\right)^{\alpha} - \rho\right]. \qquad (5)$$

With the assumption that $\tau = \frac{g}{y}$ is constant, the growth rate of consumption is again constant in this specification and we denote it by η_2. Log-transforming and time-differentiating the production function yields

$$\frac{\dot{y}}{y} = \alpha \cdot \frac{\dot{g}}{g} + (1-\alpha) \cdot \frac{\dot{k}}{k}$$

[3]The conclusions of the model are identical if we allow government to own and produce capital, provided that it has the same rate of physical depreciation as private capital. However, this complicates the analysis of the model and so we limit our exposition to the simpler case in which government owns no capital.

and so, given the assumption that output and infrastructure expenditure grow at the same rate, output and capital stock also grow at the same rate, which must be constant and equal to η_2. This augmented model is thus also in steady state at all times, though the growth rate of the per capita variables may differ from the steady state of the basic model.

Using the fact that $\frac{g}{k} = \frac{g}{y} \cdot A \cdot \left(\frac{g}{k}\right)^{\alpha}$, we can rewrite equation (5) as

$$\eta_2 = \frac{\dot{c}}{c} = \frac{1}{\theta} \cdot \left[\left(1 - \frac{g}{y}\right) \cdot (1-\alpha) \cdot \left(\frac{g}{y}\right)^{-1} \cdot \left(\frac{g}{k}\right) - \rho\right].$$

Differentiating this with respect to (g/y) and simplifying (recalling that $\alpha = \frac{\partial y}{\partial g} \cdot \frac{g}{y}$) yields

$$\frac{d\eta_2}{d(g/y)} = \frac{1}{\theta} \cdot A \cdot \left(\frac{g}{k}\right)^{\alpha} \cdot \left(\frac{\partial y}{\partial g} - 1\right). \tag{6}$$

Thus, an increase in infrastructure expenditure relative to output raises the growth rate of per capita consumption if and only if $\frac{\partial y}{\partial g} > 1$. The non-linear nature of this relationship arises from the fact that infrastructure expenditure affects output through two channels: it raises the marginal productivity of capital (as $\frac{\partial y}{\partial k} = (1-\alpha) \cdot A \cdot \left(\frac{g}{k}\right)^{\alpha}$, which is increasing in g) and reduces disposable income by increasing the tax rate τ.

As $\frac{\partial y}{\partial g} = \alpha \cdot A \cdot \left(\frac{k}{g}\right)^{1-\alpha}$, $\frac{\partial y}{\partial g}$ is clearly falling in g and so the capital productivity effect outweighs the tax effect for low values of g relative to y and *vice versa*. We thus expect increases in $\frac{g}{y}$ to be initially growth-enhancing and then, past a certain threshold, growth-retarding.

3.3 A model of institutions, infrastructure and growth

The final step in our theoretical analysis is the adaptation of the infrastructure-augmented growth model above to include an institutional dimension. We specify a revised production function, maintaining the assumption of Cobb-Douglas technology,

$$y = A \cdot g^{\beta} \cdot i^{\gamma} \cdot k^{1-\beta-\gamma} \tag{7}$$

where i is a measure of institutional capital, incorporating elements such as the protection of property rights and the accountability of the executive government. Using this production function,

$$\frac{\partial y}{\partial g} = \beta \cdot A \cdot \left(\frac{k}{g}\right)^{1-\beta} \cdot \left(\frac{i}{k}\right)^{\gamma}$$

and

$$\frac{\partial y}{\partial k} = (1 - \beta - \gamma) \cdot A \cdot \left(\frac{g}{k}\right)^{\beta} \cdot \left(\frac{i}{k}\right)^{\gamma}$$

so the marginal productivity of both private capital and infrastructure spending is increasing in the level of institutional capital. This specification captures the argument that more secure property rights and lower political instability make capital more productive. Similarly, greater political accountability raises the probability of infrastructure spending being concentrated on projects that are genuinely useful to the private sector (instead of the "prestige" projects common to less democratic polities), raising the marginal productivity of government infrastructure spending.

Maintaining the household preferences specified in equation (3) and repeating the optimal control analysis, we obtain the Euler equation

$$\frac{\dot{c}}{c} = \frac{1}{\theta} \cdot \left[(1-\tau) \cdot (1 - \beta - \gamma) \cdot A \cdot \left(\frac{g}{k}\right)^{\beta} \left(\frac{i}{k}\right)^{\gamma} - \rho\right]. \tag{8}$$

Assuming that τ is constant (i.e. g and y grow at the same rate) and that i and y grow at the same rate, the growth rate of consumption is again constant and we denote it by η_3. Furthermore, y, k, g and i also all grow at the same constant rate, η_3. This growth rate is clearly increasing in the level of institutional capital.

Using the fact that $\frac{g}{k} = \frac{g}{y} \cdot A \cdot \left(\frac{g}{k}\right)^{\beta} \left(\frac{i}{k}\right)^{\gamma}$, we can rewrite equation (8) as

$$\begin{aligned} \eta_3 = \frac{\dot{c}}{c} &= \frac{1}{\theta} \cdot \left[\left(1 - \frac{g}{y}\right) \cdot (1 - \beta - \gamma) \cdot \left(\frac{g}{y}\right)^{-1} \cdot \left(\frac{g}{k}\right) - \rho\right] \\ &= \frac{1}{\theta} \cdot \left[\left(\left(\frac{g}{y}\right)^{-1} - 1\right) \cdot (1 - \beta - \gamma) \cdot \left(\frac{g}{k}\right) - \rho\right]. \end{aligned}$$

Differentiating this with respect to (g/y) and simplifying (recalling that $\beta = \frac{\partial y}{\partial g} \cdot \frac{g}{y}$ and $\frac{g}{k} = \frac{g}{y} \cdot A \cdot \left(\frac{g}{k}\right)^{\beta} \cdot \left(\frac{i}{k}\right)^{\gamma}$) yields

$$\frac{d\eta_3}{d(g/y)} = \frac{1}{\theta} \cdot A \cdot \left(\frac{g}{k}\right)^{\beta} \cdot \left(\frac{i}{k}\right)^{\gamma} \left[\frac{\partial y}{\partial g} - (1-\gamma)\right]. \tag{9}$$

This condition permits the same interpretation as that in the model without institutional capital: rising government expenditure enchances growth for low values of g relative to y (corresponding to high values of $\frac{\partial y}{\partial g}$) but retards growth past some level.

The assumption of equal growth rates for i and y may be critiqued as overly restrictive, particularly given that this assumption is not standard in institutional growth models. However, its sole role in the model is for expositional simplicity. If we allow the value of $\frac{i}{y}$ to vary, the interpretation of η_3 and its response to a change in $\frac{g}{y}$ are unchanged, though the magnitude of such responses will vary with $\frac{i}{y}$.

3.4 Empirical specification

The model above provides a number of empirical predictions that can be tested in a regression framework. Consider the model

$$y = f(g, i, k, X) \tag{10}$$

where X is a (potentially empty) vector of appropriate control variables. The model above predicts that f_k and f_i are strictly positive and the sign of f_g varies according to the level of $\frac{g}{y}$: it is initially positive and then becomes negative once $\frac{g}{y}$ rises past a certain point.

This non-linearity can be explored in the multivariate regression framework

$$y = f(g, i, k, X) \tag{11}$$

$$k = h(g, i, y, X) \tag{12}$$

in which the model predicts $f_g < 0$ and $h_g > 0$, as increasing government expenditure encourages investment in private capital by increasing the marginal product of private capital.

Although rising private capital and rising institutional capital are both predicted to boost output, they are predicted to do so in different ways. An increase in private capital exerts a positive level effect on output but reduces its growth rate, as private capital enters negatively into equation (9). In contrast, an increase in institutional capital exerts both a positive level effect and a positive growth rate effect on output, as institutional capital enters positively into equation (9).

Our model also predicts a potential role for interactions between infrastructure expenditure and institutional capital. Mechanically, institutional capital may exert a direct effect not only on the level and growth rate of output but also on the sign and magntiude of the infrastructure-output relationship, as

$$\frac{\partial y}{\partial g} = A \cdot \beta \cdot \left(\frac{g}{k}\right)^{1-\beta} \cdot \left(\frac{i}{k}\right)^{\gamma}.$$

Intuitively, we expect that the marginal productivity of infrastructure expenditure will be higher in the presence of good institutions that direct the expenditure to socially useful projects and create a climate in which the private sector is best able to make use of the infrastructure.

In view of this observation, functional forms of models (10) and (12) in which all variables are additively separable may constitute a misspecification of the true growth process. Such forms estimate a single marginal productivity of infrastructure expenditure for all levels of institutional capital, as opposed to a marginal productivity that varies with the level of institutional capital.

We thus estimate all of our regression models in additively separable form

$$y = \alpha_1 \cdot i + \alpha_2 \cdot g + F(k, X)$$

and with a multiplicative infrastructure-institution interaction term

$$y = \alpha_1 \cdot i + \alpha_2 \cdot g + \alpha_3 \cdot i \cdot g + F(k, X).$$

In the former case, the estimated marginal productivity of infrastructure expenditure is simply $\hat{\alpha}_2$, while in the latter case it is $\hat{\alpha}_2 + \hat{\alpha}_1 \cdot i$.

4 Data

The model above requires measures of output, capital stock, infrastructure and institutions. Per capita output (Y), private capital stock (K), physical infrastructure stock[4] (PI), social infrastructure (SI) stock and aggregate infrastructure stock (I) are sourced from the South African Reserve Bank, all measured in constant 2000 prices.[5] We include institutional measures of property rights (PROP), political fractionation (FRAC) and risk (RISK). The first two measures are obtained from Fedderke, *et al*, (2001) and updated to reflect the time period from 1997 to 2006. The third is obtained by splicing the indices developed by Fedderke, *et al*, (2001) and Fedderke and Pillay (2007), as these cover different time periods. (A detailed discussion on the updating and splicing processes is included in the appendix.)

In addition to these core variables, we include measures of human capital (H) and the efficacy of the financial system (F). The latter is simply the ratio of the $M3$ money supply, measured at constant 2000 prices and obtained from the Reserve Bank, to output. The former is the ratio of highly-skilled and

[4]The physical infrastructure stock measure includes stock held by both government and public sector corporations.

[5]Although the discussion above was phrased in terms of government expenditure on infrastructure services, we use a measure of government infrastructure stock in testing the model. This reflects the fact that the South African government has historically provided infrastructure services through investment in infrastructure stock, rather than purchasing infrastructure services from the private sector (Perkins, *et al*, 2005). As noted above, the predictions of the model are identical in both cases.

skilled to semi-skilled and unskilled workers sourced from Quantec Research. While all other data series were available from 1960 to 2006, the human capital measure is available only for the period 1970 to 2005.[6]

On the basis of the Ermini-Hendry test (Banerjee, *et al*, 1993) and Schwarz Bayesian criterion we log-transform the output, capital stock, infrastructure stock and human capital series. This facilitates comparison with prior South African studies in this field, almost all of which have used log-transformed data. Furthermore, as Banerjee, *et al* (1993), note, a cointegrating relationship between log-transformed variables implies a cointegrating relationship between those variables in levels, but not vice versa. Our empirical results thus have greater generality when derived from log-transformed variables. The formal testing procedures reject, however, the hypothesis that log-transformation is appropriate for our measures of financial depth, property rights, political fractionation and risk.

Tables (1) and (2) report the results of the Phillips-Perron test sequence for the core variables for the period 1960 to 2006, with critical values obtained from Dickey and Fuller (1981). The data series are clearly not $I(0)$, though K, I, PI, SI and RISK are characterised by trend and drift components. Y, F, H, PROP and RISK are clearly $I(1)$ series, but the test sequence cannot reject the hypothesis that K, I, PI, SI and FRAC are $I(2)$ series - an economically implausible result. However, appropriate controls for structural breaks lead us to reject the hypothesis that these series are nonstationary in first differences.[7] We thus conclude that all of our variables are $I(1)$. Repeating

[6]No single measure of human capital stock in South Africa is available for 1960 - 2006. Prior analyses have largely relied on the data series on human capital stock provided by Fedderke, *et al*, (2000) but this is only available up until 1997. Attempts to splice the two series were unsuccessful, most likely because they are measuring fundamentally different conceptions of human capital: the Quantec data measure the employment in different skills categories and the Fedderke, *et al*, data measure formal school-leaving qualifications.

[7]K, I, PI and SI in first differences appear to be characterised by structural breaks in the mid-1970s, corresponding to a period of rising political uncertainty and instability in South Africa. FRAC in first differences appears to be characterised by a structural

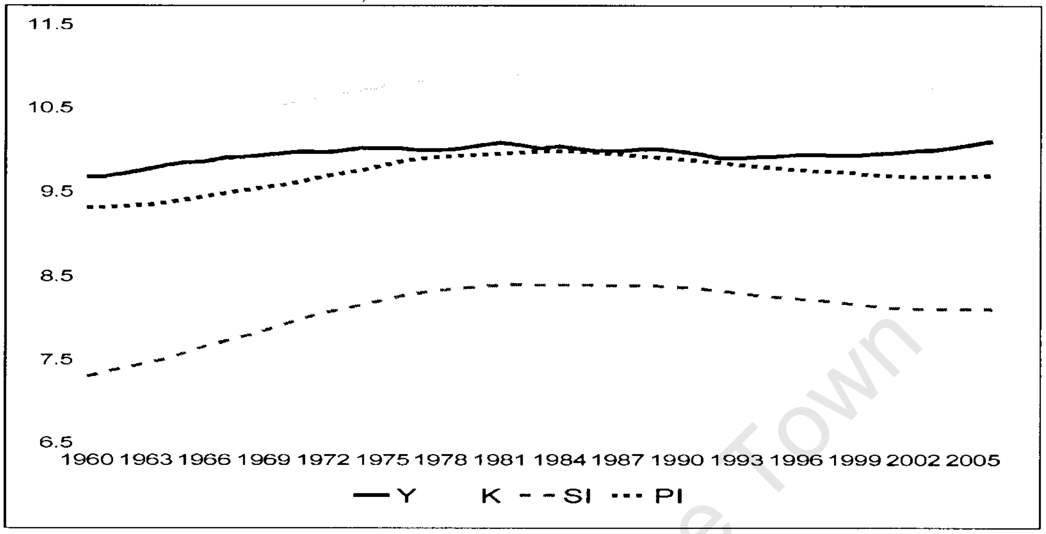
Series	Model	τ_τ	ϕ_3	τ_μ	ϕ_1	τ
Y	ADF(1)	-2.148	6.797*	-2.167	8.091*	1.290
K	ADF(1)	-2.948	10.629*	-2.860	10.481*	0.677
I	ADF(2)	-2.689	7.234*	-2.810	7.438*	0.292
PI	ADF(2)	-2.662	11.038*	-2.744	11.148*	0.398
SI	ADF(1)	-2.411	10.324*	-2.833	9.947*	-0.356
H	ADF(1)	-0.184	2.709	-1.873	6.529*	-1.097
F	ADF(0)	-0.192	5.970	0.041	0.650	0.524
PROP	ADF(0)	-1.591	2.788	-0.428	3.561	1.057
FRAC	ADF(1)	-1.248	3.372	-1.747	3.185	-0.532
RISK	ADF(0)	-2.925	10.889*	-2.732	9.173*	-1.817

Table 1: Test statistics of the Phillips-Perron test sequence for one unit root, 1960-2006. (Except F and H, 1970-2005.) * denotes significance at the 5% level.

Series	Model	τ_τ	ϕ_3	τ_μ	ϕ_1	τ
Y	ADF(0)	-3.923*	7.863*	-4.002*	7.844*	-2.728*
K	ADF(0)	-0.697	2.773	-1.310	1.545	-0.984
I	ADF(0)	-0.639	0.778	-1.302	0.615	-1.091
PI	ADF(0)	-0.975	1.850	-1.406	1.415	-1.110
SI	ADF(0)	-0.942	2.061	-1.492	1.817	-1.284
H	ADF(0)	-3.515*	7.393*	-3.220*	5.279*	-1.117
F	ADF(0)	-4.169*	8.698*	-3.514*	5.348*	-2.279*
PROP	ADF(0)	-6.413*	7.305*	-6.505*	7.303*	-2.414*
FRAC	ADF(0)	-2.379	4.244	-2.326	3.434	-1.846
RISK	ADF(2)	-6.057*	16.248*	-4.131*	16.254*	-4.930*

Table 2: Test statistics of the Phillips-Perron test sequence for two unit roots, 1960-2006. (Except F and H, 1970-2005.) * denotes significance at the 5% level.

Figure 1: Log-tranformed per capita output, capital stock and physical and social infrastructure stock, 1960-2006.



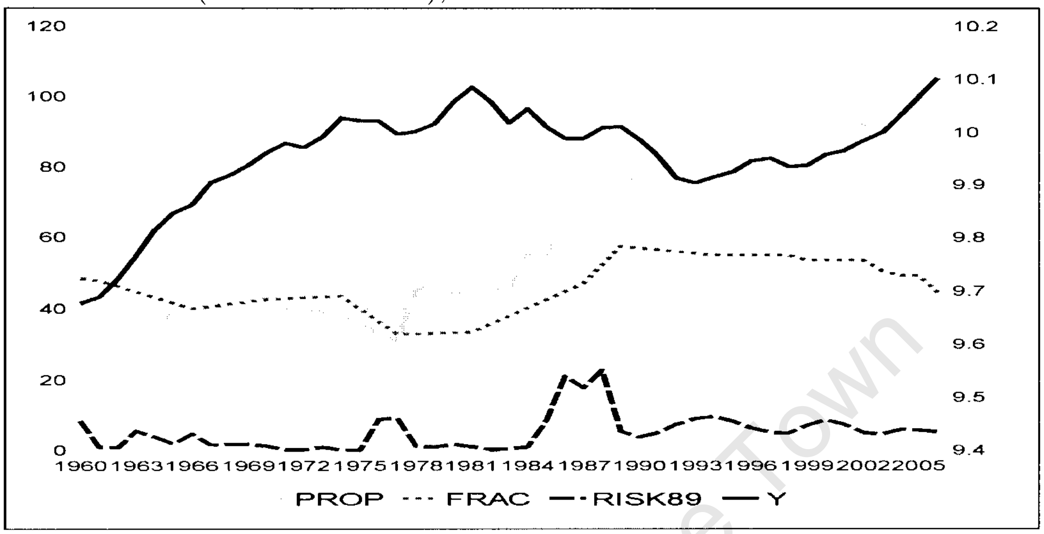
the above analysis for Y, I, K, PI, SI, PROP, FRAC and RISK for the period 1970 - 2005 produces consistent conclusions.

Figure (1) shows the relationship between log-transformed per capita output, capital stock and physical and social infrastructure stock for the period under investigation. The three series move together for much of the period, rising during the 1960s and 1970s, peaking between the late 1970s and early 1980s and then falling. However, they later begin to diverge, with output rising from the early 1990s, while capital stock and infrastructure stock continued to fall for almost another decade. This graphical comparison provides a suggestion that capital and infrastructure stock can explain much of the historical variation in South African output but that this relationship begins to break down in recent years.

Under the model developed above, this phenomenon might be explained by

break in the late 1980s. We employ both the innovational outlier and additive outlier tests proposed by Perron (1994), reflecting our uncertainty regarding the exact date of the structural break.

Figure 2: Log-tranformed per capita output (right vertical axis) and institutional indices (left vertical axis), 1960-2006.



changing institutional climate. Figure (2) shows the relationship between log-transformed per capita output and a range of institutional measures: property rights, political fractionation and risk. The steady rise in property rights during the past two decades provides a candidate explanation for the apparent break-down of the historical relationship between output on the one hand and capital and infrastructure stock on the other. Our measures of risk and political fractionation, on the other hand, show relatively little variation (in the context of their historical magnitudes) during this period, suggesting that they may play a smaller role in driving output.

5 Empirical results

5.1 Methodology

We employ two estimators in our empirical analysis: the univariate cointegrating autoregressive distributed lag (*ARDL*) estimator developed by Pesaran, Shin and Smith (1995, 1996) and the multivariate cointegrating vector error correction (*VECM*) estimator developed by Johansen (1988) and Johansen and Juselius (1992). This subsection briefly outlines the theory behind each estimator and then explains the process followed in implementing the estimation. Subsections (5.2) and (5.3) then report and discuss the estimation results obtained using the *ARDL* and Johansen estimators, respectively.

We begin with a univariate regression model of the form

$$Y_t = f(\sum_{j=1}^{J} Y_{t-j}, \sum_{j=0}^{\tilde{J}} X_{t-j})$$

where X is a vector of right-hand side variables. The *ARDL* estimator employs a linear structure for f and explicitly allows for the possibility that the variables of interest may be integrated of order zero or one.

The *ARDL* technique proceeds in two steps. In the first step, we test the hypothesis that a long-run forcing relationship exists from X to Y, which entails estimating

$$dY_t = \alpha + \beta \cdot T + \gamma \cdot Y_{t-1} + \delta \cdot X_{t-1} + \sum_{j=1}^{\tilde{J}} \mu_j \cdot dY_{t-j} + \sum_{j=1}^{J} \nu_j \cdot dX_{t-j} + \epsilon_t$$

where T is a time trend and the lag lengths J and $\tilde{J}$ are chosen to render the errors Gaussian. Under the null hypothesis of no long-run relationship, $\gamma = \delta = 0$. Only if we reject this null hypothesis is it appropriate to proceed to the second step of the analysis, in which we estimate the coefficients of this long-run relationship.

Here we estimate

$$Y_t = \alpha + \beta \cdot T + \sum_{j=1}^{J} \gamma_j \cdot Y_{t-j} + \sum_{j=0}^{\tilde{J}} \delta_j \cdot X_{t-j} \epsilon_t \tag{13}$$

where the lag lengths J and $\tilde{J}$ are chosen to render the errors Gaussian. The cointegrating relationships are then given by the equation

$$Y_t = \bar{\alpha} + \bar{\beta} \cdot t + \bar{\delta} \cdot X_t + \bar{\epsilon}_t$$

where the coefficients are computed using the formulae

$$\bar{\alpha} = \frac{\hat{\alpha}}{1 - \sum_{j=1}^{J} \hat{\gamma}_j}$$

$$\bar{\beta} = \frac{\hat{\beta}}{1 - \sum_{j=1}^{J} \hat{\gamma}_j}$$

$$\bar{\delta} = \frac{\sum_{j=0}^{J} \hat{\delta}_j}{1 - \sum_{j=1}^{J} \hat{\gamma}_j}$$

In the following subsection we report these long-run coefficients directly.

However, the *ARDL* estimator suffers from a number of limitations. Firstly, the estimation does not take into account the possibility of contemporaneous feedback from output to the right-hand side variables - it instead assumes that the right-hand side variables are all weakly exogenous. Secondly, it does not allow an explicit test of our model's prediction that the positive effect of infrastructure spending on output operates by raising the marginal productivity of capital. Finally, it presumes the existence of a single cointegrating relationship in the data, when in fact there may be several. The estimated cointegrating vector would then be a linear combination of the true cointegrating relationships, resulting in inefficient estimation.

The multivariate *VECM* estimator at least in part addresses all of these limitations. We begin with an n-dimensional vector-autoregressive model (*VAR*)

$$Z_t = \sum_{j=1}^{J} \cdot A_j \cdot Z_{t-j} + \delta + \epsilon_t$$

where the Z vector contains all n variables of interest, the δ vector contains all relevant deterministic components and ϵ is an error term rendered Gaussian by appropriate specification of the lag-length. This expression yields the system of difference equations

$$\Delta Z_t = \sum_{j=1}^{J} \cdot \Gamma_j \cdot \Delta Z_{t-j} + \Pi \cdot Z_{t-j+1} + e_t$$

where Γ_j is a n-dimensional vector of parameters capturing the short-run dynamics of the model and Π is an $n \times n$ matrix capturing the long-run relationship between the variables in Z. Under the null hypothesis that the model contains r cointegrating relationships

$$\Pi = \alpha \cdot \beta'$$

where α and β are $n \times r$ non-singular matrices. We can explicitly test for the value of r, though some of these tests have very low power in finite samples.

Where $r > 1$, we need to impose restrictions on the matrix representing the cointegrating space, β. Specifically, exact identification requires the imposition of r linearly independent restrictions on each of the r cointegrating vectors, each of which is a column in the β matrix. The long-run relationships between the elements of Z_t are then given by the elements of the restricted Π matrix and the short-run dynamics by the elements of the Γ matrix.

Although the *VECM* estimator addresses many of the shortcomings of the *ARDL* estimator, it is not without limitations of its own. Perhaps most importantly, estimating all of the elements of the α, β and Γ matrices is very expensive in terms of degrees of freedom and the approach is thus problematic in small samples. Furthermore, aspects of the estimator's finite sample performance are unknown.

In the current analysis, the identification requirements pose an additional challenge. Our model predicts that physical, infrastructural and institutional capital will all positively force output and should thus not be restricted in the β matrix. We therefore specify the cointegrating relationships

$$\Pi \cdot Z_{t-j+1} = \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \\ \alpha_{31} & \alpha_{32} \\ \alpha_{41} & \alpha_{42} \\ \alpha_{51} & \alpha_{52} \\ \alpha_{61} & \alpha_{62} \end{pmatrix} \cdot \beta' \cdot \begin{pmatrix} Y \\ K \\ F \\ H \\ INST \\ I \end{pmatrix}$$

where

$$\beta' = \begin{pmatrix} 1 & \beta_{12} & 0 & \beta_{14} & \beta_{15} & \beta_{16} \\ \beta_{21} & 1 & \beta_{23} & 0 & \beta_{25} & \beta_{26} \end{pmatrix}$$

and INST and I are measures of institutional and infrastructural capital, respectively. This model builds off the specification in subsection (3.4) and incorporates the findings of Kularatne (2002, 2006) that financial depth affects economic output primarily by encouraging capital accumulation while human capital accumulation affects economic output directly. While there is some evidence for certain choices of $INST$ and I that three cointegrating relationships may be present, the sample size is sufficiently small that estimating such a model would leave us with virtually no statistical power.

The inclusion of human capital stock, H, however, limits our sample period to the period between 1970 and 2005. Compared to the *ARDL* analysis, this simultaneously reduces the number of observations and increases the number of parameters requiring estimation. The *VECM* estimation is thus characterised by low power and the possibility of finite sample biases. In sum, both the *ARDL* and *VECM* estimators suffer from limitations in the context of our analysis, underlining the importance of employing both estimators and comparing their results.

Our empirical analysis proceeds as follows. Given the range of available institutional and infrastructural measures, we estimate each model using physical infrastructure stock, social infrastructure stock, aggregate infrastructure stock and both physical and social infrastructure stock, using measures of property rights, fractionation and risk (giving 12 specifications). We then re-estimate each model with infrastructure-institution interactions and, for the *ARDL* estimator only,[8] with quadratic infrastructure terms, allowing for the non-linear relationship between infrastructure and output implied by our model.

For each of these sets of variables, we:

[8]The Johansen estimator does not permit the inclusion of a quadratic relationship.

- test for the existence and direction of a long-run relationship
- if such a relationship exists, test for the number of cointegrating relationships; and
- estimate the long-run relationship(s) using the *ARDL* and/or *VECM* estimator(s), depending on the outcomes of the second test.

Included variables	Left-hand side variable							
	Y	K	I	PI	SI	PROP	FRAC	RISK
Y, K, I, PROP	4.67*	2.06	1.90			3.25		
Y, K, I, FRAC	7.84*	2.27	3.46				2.48	
Y, K, I, RISK	2.23	3.09	2.31					7.63*
Y, K, PI, PROP	4.67*	1.53		2.68		3.37		
Y, K, PI, FRAC	7.60*	1.89		4.64			0.27	
Y, K, PI, RISK	2.18	2.33		3.62				7.22*
Y, K, SI, PROP	5.33*	5.58*			0.42	2.47		
Y, K, SI, FRAC	6.95*	3.02			2.30		4.24	
Y, K, SI, RISK	2.25	2.04			3.25			6.06*

Table 3: Test results for the existence of a long-run relationship, 1960 - 2006. * denotes signifiance at the 5% level, using the critical values from Pesaran, Shin and Smith (1995).

Table (3) reports the results of first set of tests. We observe that the null hypothesis that no long-run forcing relationship exists to Y from the relevant subset of right-hand side variables is rejected for almost all specifications. We can proceed to the second stage of the testing procedure outlined above. These results also provide support for the idea that the appropriate left-hand side variable in our estimation of the coefficients of the long-run relationship is per capita output.[9] We report only a selection of results, which are broadly

[9]The tests also support the contention that risk may be the appropriate left-hand side variable. However, estimation of the long-run relationship with this specification yields economically nonsensical results, irrespective of whether we employ a univariate or multivariate estimator. We thus do not report these results.

representative of the entire set.

Table (4) reports the results of the second set of tests. (Again, we report only a selection of results, which are broadly representative.)[10] The results are clearly highly sensitive to the choice of which deterministic components to include in the VAR and the trace and maximal eignenvalue statistics routinely suggest different conclusions. Where such conflicts arise we place greater weight on the structure suggested by trace statistic (which has better finite sample performance) and by the specifications including a trend component, given the high probability of underspecification typically present in growth models.

Most specifications provide some support for the existence of either one or two cointegrating vectors, though a larger proportion of the specifications including interaction terms appear to contain at least three cointegrating vectors. However, these results should be interpreted with some degree of caution: the test structure is such that r cointegrating vectors are found only if the null hypothesis of $r+1$ cointegrating vectors is rejected. Hence, the low power arising from our small sample introduces a systematic bias in favour of finding large numbers of cointegrating vectors. This is particularly pertinent for the specifications containing more variables (such as those including interactions) where power is particularly low.

5.2 Univariate regression analysis

Examining the results obtained under univariate estimation clearly reveal a high degree of sensitivity to changes in both specification and lag length.

[10]The lag length in the underlying VAR used in these tests was chosen on the basis of the adjusted log-likelihood statistic and the Schwarz Bayesian criteria. These results are not reported here but they indicate that all models should be estimated with VAR $\in \{1, 2\}$. Where the test evidence is ambiguous we err on the side of parsimony, given the small sample.

Given the range of results obtained, we report only a selection of results. This sample is weighted toward the regression results that support the hypotheses developed in our theoretical model and so should be viewed as an upper bound on the extent to which the data cohere with the model's hypotheses.

The findings in table (5) suggest that capital stock is significantly positively associated with economic output, with an estimated elasticity between 1.2 and 1.9 (and an outlier of 2.5). This is broadly consistent with prior research in this area. However, the relationship between infrastructure stock and output is negative in most specifications. Aggregate infrastructure stock enters negatively and significantly with a elasticity of approximately -1.3, which is larger than prior empirical findings. Physical infrastructure stock is also negative and significant, with an even larger (in absolute value) elasticity. When both physical and social infrastructure stock are included, the latter has a positive output elasticity but this is not always significant. Social infrastructure stock alone enters with a negative elasticity but this is not significant and is typically closer to zero than the coefficient on physical infrastructure. When both measures are included, the negative coefficient on physical infrastructure is larger (in absolute magnitude) but social infrastructure now has a positive impact on output.

Although the the exact form of the output-infrastructure relationship is highly sensitive to specification changes, its negative sign is relatively robust. This may either reflect a genuine negative relationship between the two variables or may reflect the conflation of the (theoretically negative) direct relationship and the (theoretically positive) indirect relationship *via* the marginal productivity of capital. The level of infrastructure investment may thus have exceeded its optimum and the resultant overinvestment may be reducing both output and the growth rate of output. This is not, however, consistent with prior findings, which have generally found that infrastructure stock is well below its optimal level for most of South Africa's history.

Furthermore, the relationship between the institutional variables and economic output is highly sensitive to specification changes. None of the three institutional variables are significant determinants of output across all specifications and the most robust finding is a *positive* relationship between risk and output, which defies economic explanation. There is also more evidence in support of a negative relationship between fractionation and output than a positive relationship. This provides moderate support to the hypothesis that greater political fractionation induces delays and rent-seeking into the policy-making process, with detrimental effects on government efficacy. This coefficient should, however, be interpreted with a high degree of caution, given that South Africa's legislature has historically enjoyed relatively little power to influence the executive branch of government and so fractionation may be a poor indicator of the extent to which government faces checks and balances in its formulation and execution of infrastructure policy.

As a final note on these results, the absolute magnitude of the coefficients on the institutional variables is substantially smaller than the coefficients on physical capital and infrastructure capital, across all specifications. We cannot, however, use the magnitude of these coefficients to comment on the relative importance of different determinants of output, as the different series are measured on different scales and the institutional variables are in level form, whereas most other variables have been log-transformed.

All of the specifications including a quadratic term in infrastructure stock produce highly economically implausible coefficient estimates with elasticities routinely exceeding six in absolute value and so are not reported here. However, table (6) reports selected results obtained using infrastructure-institution interaction terms, which differ from those in table (5) in several respects.

Firstly, the coefficient on capital becomes more sensitive to specification changes: though it remains positive, its magnitude varies substantially across specifications and it is insigificant in several specifications. The infrastructural variables retain broadly the same signs and magnitudes as in the previous results but they are not significant in any specification. The institutional variables continue to be highly sensitive to specification changes and are seldom significant, though the positive association between output and risk remains relatively robust. Insignificance of these estimates should, however, be treated with caution, as the model's relatively low power may lead to inappropriate underrejection of the null hypothesis of insignificance.

The interaction terms themselves are almost invariably insignificant but their economic interpretation is intriguing. In the first column of table (6), for example, both fractionation and physical infrastructure stock are estimated to reduce economic output but their interaction term *raises* output. This suggests that infrastructure stock may be more effectively deployed to promote economic growth under a more fractionated polity. The magnitude of this interaction effect is not insubstantial: a rise in physical infrastructure stock is positively associated with economic output for values of fractionation greater than 6.8, which includes all observations in the sample. However, given the lack of precision of the estimates in this specification, such computations are highly speculative.

The estimates obtained under $ARDL$ estimation thus suggest four broad conclusions: physical capital is positively associated with output, infrastructure is negatively associated with output (though this finding is more true for physical than social infrastructure) with no plausible evidence of a quadratic relationship, institutional variables have relatively little explanatory power and interaction terms are potentially important components of our model. Above all these, however, is the conclusion that relatively few findings are robust across a range of specifications and that power limitations pose a substantial challenge to inference. Broadly speaking, the predictions of our model are not borne out by this modeling exercise: institutional factors do not appear to be important and the infrastructure-output relationship does not appear to be non-linear.

5.3 Multivariate regression analysis

For each of the specifications cited in subsection (5.1) we test for the number of cointegrating vectors using the trace and maximal eigenvalue statistics (with the addition of the financial depth and human capital measures we require for identification purposes). These estimates are broadly consistent with their equivalents in table (4) and so we do not report them separately.[11] On the basis of these tests we identify which specifications are amenable to the two-equation regression model proposed in subsection (3.4).

Given the detail involved in fully interpreting a VECM, we discuss in depth only a small selection of multivariate models. These are again broadly representative of the overall results, with a bias toward those indicating greater coherence of the data with the model and so should again be viewed as providing an informal upper bound of the coherence on the theory with the data.

Table (7) reports the estimates for the model using physical infrastructure and risk as measures of infrastructure and institutions, respectively. Three specifications are shown: (*i*) without an infrastructure-institution interaction term, (*ii*) with an interaction term and (*iii*) with an interaction term and the feedback channel from output to physical capital stock zero-restricted. The third specification illustrates the impact that overidentification has on many of the models: the overidentified model is characterised by more precise and more economically plausible coefficient estimates and, as we discuss below, is somewhat more stable than the just-identified model. The null hypothesis that the overidentification restriction is valid cannot be rejected at any standard level of significance.

The estimated elasticity of output with respect to capital is consistently pos-

[11] A notable exception to this trend is observed for the two interaction-augmented models we discuss below. An examination of table (4) may suggest that a two-equation specification is inappropriate here. However, the test procedure supports the hypothesis of two cointegrating relationships for both of these sets of variables for the period 1970 to 2005 with F and H included.

itive across all three specification and the magnitude between 0.95 and 1.55 is consistent with prior studies. However, the coefficient is not significant in any of the specifications, most likely reflecting the model's low power. There is also weak evidence suggesting that there is a positive feedback loop from output to physical capital stock. These results are consistent across a range of specifications, though the evidence in support of a positive feedback loop is stronger in most other specifications.

Human capital is positively associated with output in all three models, though its coefficient is not robust across different specifications and is seldom significant. More generally, both the sign and magnitude of this coefficient are *highly* sensitive to even minor specification changes and the balance of evidence is in fact in favour of a *negative* relationship, which at first glance appears entirely at odds with mainstream growth theory and empirics. However, this apparent discrepancy can be explained by a careful consideration of the economic content of the human capital measure. Our data measure the ratio of skilled to unskilled workers in the economy, which is not an indicator of the supply human capital available to the economy, but rather an indicator of the relative employment of human capital (i.e. a measure of equilibrium in the market for skilled labour, rather than the supply side of that market). Furthermore, if dips in output are associated with a disproportionate loss of employment for unskilled workers, we may observe an increase in the ratio of skilled to unskilled employment during economic slowdowns. This hypothesis receives indicative support from the fact that the covariance of the two series is in fact -0.371. Unfortunately, the lack of alternative measures of human capital stock in South Africa leaves no avenue for avoiding this constraint.

The consistently negative relationship between physical capital and financial depth (replicated in most other specifications) also warrants explanation. Our measure of financial depth (the ratio of M3 money supply to output) is vulnerable to a similar problem: a slowdown in output growth relative to money growth may well result in a disproportionate rise in observed financial depth, raising the possibility of a negative relationship between the two series. The covariance of 0.477 does not directly support this hypothesis but this bivariate measure of association does not take into account the influence of other variables in the fully specified multivariate model and we suspect that the construction of the financial depth indicator is indeed problematic. Here again, we do not have ready access to any measures of financial depth that measure the actual *capacity* of the financial system, rather than some equilibrium in the money market.

The estimated coefficients on physical infrastructure are representative of a pattern consistently found in almost all specifications: infrastructure positively affects capital stock and negatively affects output directly, with the indirect effect *via* capital accumulation dominating the direct effect. The indirect elasticities are 1.06, 1.15 and 0.86 in the first, second and third models, respectively, while the direct elasticities are -0.68, -1.05 and -0.66.[12] None of these differences are, however, statistically significant,[13] in view of the highly imprecise coefficient estimates in the first cointegrating vector in all three specifications.

This pattern of competing direct and indirect effects is replicated for relationship between risk and output. In contrast to the *ARDL* estimation results, which found a consistently positive relationship between risk and output, the VECM presents a considerably more nuanced picture. The typical pattern is that risk exerts a negative effect effect on physical capital stock (and hence a negative indirect effect on output) but a positive direct effect on output. The overall effect is invariably insignificant, which remains at odds with economic theory but is less controversial than the positive relationship found above.

The inclusion of a risk-infrastructure interaction term produces intriguing results.[14] The direct effect of risk on output is now negative for any value of

[12]We abstract away from the interaction terms at this stage of our discussion.

[13]Using a χ^2 test for equality of the estimates.

[14]We limit discussion in this paragraph to the overidentified model, given the extreme instability of the just-identified model.

PI less than 9.92, which is true of the period prior to 1978 and after 1988. Furthermore, the direct effect of physical infrastructure on output is now positive for the early 1970s and from the mid-1980s onward. The latter finding suggests that physical infrastructure in fact has positive direct and indirect effects on output, provided that the economy is experiencing a sufficiently low level of risk that the private sector is able to make appropriate use of the available infrastructure. This is a potentially highly salient result, as it runs contrary to the finding of a negative or insignificant effect of physical infrastructure on output in all of the prior literature employing multivariate specifications and implies that their results may have been driven by simulatenously high levels of risk and infrastructure stock between between the mid 1970s and mid 1980s.

However, this result should be interpreted with caution in view of some diagnostic checks on the stability of the estimated models. The bulk of the error correction terms in all three reported specifications are positive, implying that the model does not converge following a shock to either cointegrating vector and casting doubt on the stability of the model. In order to comment more confidently on the stability of the model, we must consider the impulse response functions, which map out the predicted response of the cointegrating vectors and the individual variables to shocks in either the cointegrating relationships or the variables themselves. Due to the number of graphs involved in this analysis, we simply summarize the broad findings, rather than presenting the graphical evidence directly.

In the specifications reported above, the cointegrating relationships themselves are stable to shocks of any form, though they routinely take up to two decades to converge back to their original values. The stability of the individual series varies: output, physical capital and infrastructure converge back to their original steady states, though output converges to a higher steady state following shocks to any of the institutional variables. This convergence again takes up to two decades. The institutional variables themselves are relatively unstable - they fail to converge after shocks to any of the other variables. These results are highly consistent across a range of specifications and this instability suggests that the empirical model is not an entirely accurate depiction of the underlying data generating process.

The second model we consider employs social infrastructure stock and property rights as measures of infrastructure and institutions, respectively, and is also estimated with and without interaction terms. We again observe a positive elasticity of output with respect to physical capital and a positive feedback loop. Financial capital is now positively associated with physical capital (which is not representative of the bulk of the specifications) and the relationship between output and human capital differs across the two specifications. The positive indirect and negative direct relationship between infrastructure and output is again observed, with elasticities of 0.6 and -0.37 in the first model and 1.85 and -0.64 in the second. Neither difference is, however, statistically significant, again reflecting the imprecision of the estimates. This contrasts with the only prior study of this relationship (Kularatne, 2006), which found both the direct and indirect effects to be positive.

The role of property rights in both cointegrating vectors is highly sensitive to specification changes and no clear pattern can be observed. With the inclusion of an interaction term, we see that property rights are positively associated with physical capital for all observed values of social infrastructure stock, though the effect is entirely insignificant, with an implied coefficient of 0.004 at the mean value of social infrastructure stock. The direct effect on output is overwhelmingly negative. These results are also, however, sensitive to specification changes.

The interaction terms tells a more interesting story for role of social infrastructure stock. Up until 1990, the implied indirect effect is positive and the implied direct effect is negative. For all values of property rights observed after 1990, however, the signs are reversed. This is surprising and does nor accord with our expectation regarding the interaction between property rights and physical capital - we would expect improved protection of property rights to facilitate private sector use of rising social infrastructure stock and thereby encourage investment in physical capital. This may therefore reflect a structural change in the model itself in the early 1990s, corresponding to the fundamental change in the relationship between the economy and government at the end of the apartheid period. However, as this result is not replicated in other specifications, we cannot read too much into it.

In interpreting these results we must bear in mind that these models suffer from the same stability problems noted above, casting doubt on the extent to which they cohere with the observed data and meaning that their output should be treated with a degree of caution.

6 Conclusion

Given the policy importance of the relationship between infrastructure and output in the current South African context, a clear academic understanding of this relationship is vital. While this relationship has been widely studied in recent years, with increasingly appropriate and sophisticated tools, little attention has been paid to its institutional dimension. This is perhaps surprising, in view of the clear *a priori* case for the importance of this institution linkage.

The contribution of this paper is, firstly, to provide a clear theoretical case for attending to the role of institutions when examining the infrastructure-output relationship and, secondly, to test for the existence and form of such a role. Our model provides a number of clear-cut predictions concerning the nature of this relationship: institutional capital positively affects both output and the relationship between infrastructure and output. These predictions, however, receive mixed support when tested against the data.

The data provide unambiguous support for the idea that the infrastructure-output relationship is highly sensitive to the inclusion of various measures of institutional capital: risk, fractionation and property rights. However, the relationship is also highly sensitive to the choice of measure and specification of the model - few results are robust across different specifications and our empirical models are routinely unstable. The identification of a stable and robust relationship, and the process of inference more broadly, is complicated by the short time series available and the limitations of some of the available measures.

The central conclusion of our study is that institutions appear to be an important factor in shaping the infrastructure-output relationship, but the exact form of this relationship is not clear. Ultimately, however, we may be attempting to extract more structure from the data than is currently possible.

7 Bibliography

1. Abedian, I. and Van Seventer, D. (1995). "Productivity and Multiplier Analysis of Infrastructure Investment in South Africa: An Econometric Investigation and Preliminary Policy Implications." Mimeo. Pretoria: Ministry of Finance.

2. Alves, P. and Edwards, L. (2005). "South Africa's Export Performance: Determinants of Export Supply." *South African Journal of Economics*, Vol. 74, No. 3, pp. 473 - 500.

3. Aschauer, D. (1989a). "Is Public Expenditure Productive?" *Journal of Monetary Economics*, Vol. 23, pp. 177 - 200.

4. Aschauer, D. (1989b). "Public Investment and Productivity Growth in the Group of Seven" *Economic Perspectives*, Vol. 13, pp. 17 - 25.

5. Aschauer, D. (1989c). "Does Public Capital Crowd Out Private Capital?" *Journal of Monetary Economics*, Vol. 24, pp. 171 - 188.

6. Banerjee, A., Dolado, J., Galbraith, J. and Hendry, D. (1993). *Co-Integration, Error-Correction and the Econometric Analysis of Non-Stationary Data.* Oxford University Press, Oxford.

7. Barro, R. (1990). "Government Spending in a Simple Model of Endogenous Growth." *Journal of Political Economy*, Vol. 98, No. 5, pp. 177 - 200.

8. Development Bank of Southern Africa. (1998). "Infrastructure: A Foundation for Development." *Development Report 1998.* Pretoria.

9. Dickey, D. and Fuller, W. (1981). "Likelihood Ratio Statistics for Autoregressive Time Series With a Unit Root." *Econometrica*, Vol. 47, pp. 1057 - 1072.

10. EDWARDS, L. and JOHNNY, B. (2006). "South Africa's Export Performance: Determinants of Export Supply." Presented at the Economic Research Southern Africa Infrastructure Workshop, May 2006.

11. ENDERS, W. (2004). *Applied Econometric Time Series.* John Wiley and Sons, Hoboken.

12. ESFAHANI, H. and RAMIREZ, M. (2003). "Institutions, Infrastructure and Economic Growth." *Journal of Economic Development*, Vol. 70, No. 2, pp. 443 - 477.

13. FEDDERKE, J. (2004). "Investment in Fixed Capital Stock: Testing for the Impact of Systemic and Sectoral Uncertainty." *Oxford Bulletin of Economics and Statistics*, Vol. 66, No. 2, pp. 165 - 187.

14. FEDDERKE, J. and BOGETIĆ, Ž. (2006). "Infrastructure and Growth in South Africa: Direct and Indirect Productivity Impacts of 19 Infrastructure Measures." Economic Research Southern Africa Working Paper No. 39.

15. FEDDERKE, J., DE KADT, R. and LUIZ, J. (2000). "Uneducating South Africa: The Failure to Address the 1910-1993 Legacy." *International Review of Education.* Vol. 46, No. 3/4, pp. 257 - 281.

16. FEDDERKE, J., DE KADT, R. and LUIZ, J. (2001). "Indicators of Political Liberty, Property Rights and Political Instability in South Africa: 1935 - 1997." *International Review of Law and Economics*, Vol. 21, pp. 103 - 134.

17. FEDDERKE, J. and GARLICK, R. (2007). "Infrastructure Development and Economic Growth in South Africa: A Review of the Accumulated Evidence." Unpublished manuscript, University of Cape Town.

18. FEDDERKE, J., PERKINS, P. and LUIZ, J.(2006). "Infrastructural Investment in Long-run Economic Growth: South Africa 1875 - 2001." *World Development*, Vol. 34, No. 6, pp. 1037 - 1059.

19. Fedderke, J. and Pillay, N. (2007). "A Theoretically Defensible Measure of Risk: Using Financial Data from a Middle Income Context." Economic Research Southern Africa Working Paper No. 64.

20. Fielding, D. (1997). "Aggregate Investment in South Africa: A Model with Implications for Political Reform." *Oxford Bulletin of Economics and Statistics*, Vol. 59, No. 3, pp. 349 - 369.

21. Gramlich, E. (1994). "Infrastructure Investment: A Review Essay." *Journal of Economic Literature*, Vol. XXXII, September, pp. 1176 - 1196.

22. Guasch, J. (2004). *Granting and Renegotiating Infrastructure Concessions: Doing It Right.* Washington, DC: World Bank Institute.

23. Johansen, S. (1988). "Statistical Analysis of Cointegrating Vectors." *Journal of Economic Dynamics and Control.* Vol. 12, pp. 231 - 254.

24. Johansen, S. and Juselius, K. (1992). "Testing Structural Hypotheses in a Multivariate Cointegrating Analysis of the PPP and the UIP for the UK." *Journal of Econometrics.* Vol. 53, pp. 211 - 224.

25. Kessides, I. (2004). *Reforming Infrastructure: Privatisation, Regulation and Competition.* Washington, DC: World Bank Institute and Oxford University Press.

26. Kularatne, C. (2002). "An Examination of the Impact of Financial Deepening on Long-run Economic Growth: An Application of a VECM Structure to a Middle-income Country Context." *South African Journal of Economics*, Vol. 70, No. 4, pp. 647 - 687.

27. Kularatne, C. (2006). "Social and Economic Infrastructure Impacts on Economic Growth in South Africa." Unpublished manuscript.

28. Lütkepohl, H. and Krätzig, M. (2004). *Applied Time Series Econometrics.* Cambridge University Press, New York.

29. MARIOTTI, M. (2002). "An Examination of the Impact of Economic Policy on Long-run Economic Growth: An Application of a VECM Structure to a Middle-income Country Context." *South African Journal of Economics*, Vol. 70, No. 4, pp. 688 - 724.

30. MUNNELL, A. (1990). "Why Has Productivity Growth Declined? Productivity and Public Investment." *New England Economic Review*, Jan/Feb, pp. 3 - 22.

31. PERKINS, P., FEDDERKE, J. and LUIZ, J. (2005). "An Analysis of Economic Infrastructure Investment in South Africa." *South African Journal of Economics*, Vol. 32, No. 2, pp. 211 - 228.

32. PERRON, P. (1994). "Trend, Unit Root and Structural Change in Macroeconomic Time Series." In RAO, B. (ed.) *Cointegration for the Applied Economists.* St. Martin's Press, New York.

33. PESARAN, H. and SHIN, Y. (1996). "An Autoregressive Distributed Lag Modeling Approach to Cointegration Analysis." DAE Working Paper No. 9514, Department of Applied Economics, University of Cambridge.

34. PESARAN, H., SHIN, Y. and SMITH, R. (1995). "Testing for the Existence of a Long-Run Relationship." Unpublished manuscript, University of Cambridge.

35. PHILLIPS, P. and PERRON, P. (1988). "Testing for a Unit Root in Time Series Regression." *Biometrika*, Vol. 75, pp. 333 - 346.

36. SCHWERT, G. (1989). "Tests for Unit Roots: a Monte Carlo Investigation." *Journal of Business and Economic Statistics*, Vol. 7, pp. 147 - 159.

37. SOUTH AFRICA. (2007a). *Accelerated and Shared Growth Initiative for South Africa: A Summary.* Available online: http://www.info.gov.za/asgisa/.

38. South Africa. (2007b). *National Budget Review.* Available online: http://www.treasury.gov.za/documents/budget /2007/review/default.htm.

8 Appendix

The construction of the institutional indices of risk, property rights and political fractionation is an important component of our analysis. All three measures have their genesis in the indices developed by Fedderke, De Kadt and Luiz (2001), hereinafter "FDKL," but these are unfortunately available only until the late 1990s and thus needed to be updated for the purposes of this paper.

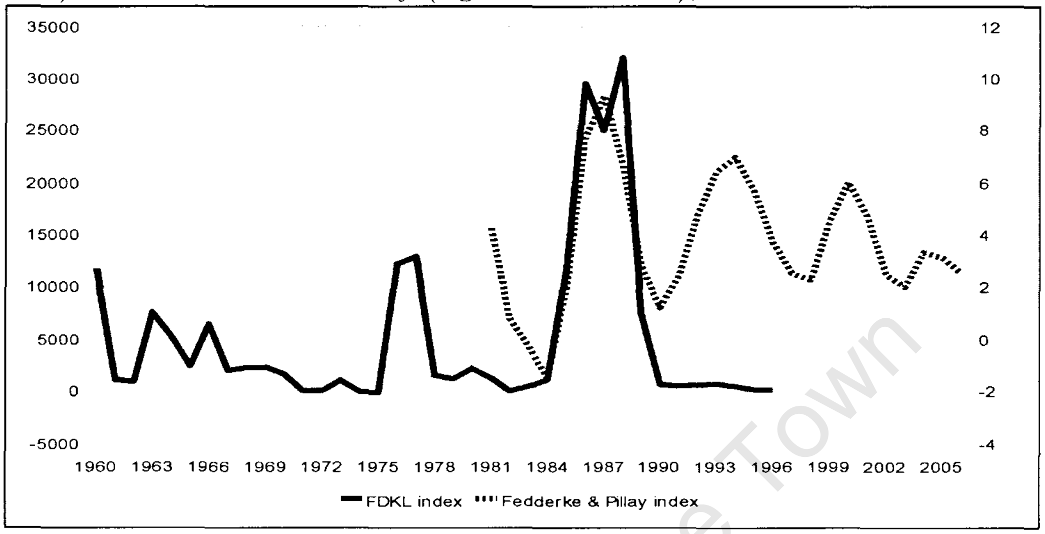
8.1 Risk

The FDKL index is explicitly an index of *political* risk and instability, constructed from data including prosecutions under apartheid-era "national security" legislation, bannings, police actions against "riots" and media censorships. Given South Africa's history during the period under investigation (1935 - 1997), it seems likely that the dimension of risk relevant to economic actors would be political in nature.

In the past decade, however, political risk of the form measured by the FDKL index has been of little relevance to South Africa, as almost all of the measures included in the index have dropped to zero. The more relevant dimension of risk during this period has arguably been economic in nature, reflecting phenomena such as currency fluctuations and commodity price movements. We therefore turn to the index developed by Fedderke and Pillay (2007), which uses deviations in bond yields from those predicted under the expectations hypothesis to capture investors' perceptions of overall risk levels in the South African economy.

Figure (3) depicts the two indices. As the discussion above suggests, the FDKL index appears to underestimate risk substantially from early 1990s onward. Furthermore, movements in the two indices are almost identical during the mid- to late-1980s, suggesting that political risk was the dominant

Figure 3: Risk indices developed by Fedderke, De Kadt and Luiz (left vertical axis) and Fedderke and Pillay (right vertical axis), 1960 - 2006.



feature of the economy at the time. This period of co-movement suggests an appropriate splicing approach is to normalise one index at some point during this period or for an average of several years during this period. We experiment with several splicing schemes of this form and find that the conclusions of the paper are robust to changes in the splicing scheme. The results reported above are those obtained normalising the Fedderke and Pillay index on the FDKL index in 1989.

8.2 Property Rights

The index of property rights used in FDKL measured the "right to possess, the right to use, the right to manage, the right to the capital, the right to security, the incident of transmissibility and liability to execution" of fixed (i.e. not intellectual) property, with an equal weighting attached to each component. The index was increasing in the degree of protection of property rights and bounded by 0 and 100. This was a *de jure* measure that was based purely on the content of legislation and did not reflect the *de facto* enforcement of these rights. Reflecting the different legislation applied to different racial groups for much of the relevant period, the index was constructed by assigning a property rights measure for each racial group and then weighting these by population shares.

The original index was the result of a three-step process in which (*i*) the relevant set of evaluation criteria were defined, *a priori*, (*ii*) two researchers independently constructed data series and (*iii*) the series were presented to a panel of South African social scientists and lawyers and modified on the basis of their input. The updating process, however, was significantly more constrained and thus followed a more modest process in which the third step was entirely omitted and the second step was conducted by only one researcher, rather than two operating independently.

We therefore conducted a desktop survey of all legislation passed by the National Assembly between 1996 and 2007 and identified those pieces of legislation affecting the components of property rights listed above. We excluded legislation that affected the property rights only of convicted criminals (eg. the Prevention of Organised Crime Act of 1998 and its subsequent amendments), and that constrained property owners' actions in ways not directly related to the control of their property (eg. the Broad-Based Black Economic Empowerment Act of 2003) and attached a low weighting to legislation affecting a very small proportion of South Africa's fixed capital stock (eg. the Communal Land Rights Act of 2004) .

8.3 Fractionation

The FDKL index of fractionation measures the degree of concentration in the legislative branch of government, using the formula

$$F = 1 - \sum_{j=1}^{J} \left(\frac{n_j}{N}\right) \left(\frac{n_j - 1}{N - 1}\right)$$

where n_j denotes the number of members of party j, N denotes the number of members of the legislature and J denotes the total number of parties in the legislature. This measure was updated for the period 1996 to 2006 (for the National Assembly only, not the National Council of Provinces, to ensure comparability with the original data series) using data from the Journal of Procedural Developments.

Variables	VAR length	No C, No T E'value	No C, No T Trace	Rst C, No T E'value	Rst C, No T Trace	Unrst C, No T E'value	Unrst C, No T Trace	Unrst C, Rst T E'value	Unrst C, Rst T Trace	Unrst C, Unrst T E'value	Unrst C, Unrst T Trace
Y, K, PI, FRAC	2	1	2	0	2	0	1	1	3	1	2
Y, K, PI, RISK	1	4	4	2	2	2	2	2	3	2	3
Y, K, SI, PROP	1	4	4	2	3	3	2	3	4	2	3
Y, K, SI, RISK	2	1	2	1	1	1	1	1	1	1	1
Y, K, PI, SI, FRAC	1	4	4	3	3	3	3	3	3	3	3
Y, K, I, PROP	2	3	4	0	3	0	1	1	1	0	1
Y, K, I, FRAC	2	1	2	1	2	1	2	1	3	1	2
Y, K, I, FRAC, I*FRAC	2	2	2	2	3	1	2	0	5	3	3
Y, K, PI, FRAC, PI*FRAC	2	2	2	2	3	2	2	1	4	1	3
Y, K, PI, RISK, PI*RISK	2	1	4	1	4	1	1	1	2	1	3
Y, K, SI, PROP, SI*PROP	1	5	5	3	4	3	3	4	4	3	3
Y, K, SI, RISK, SI*RISK	2	2	2	2	2	2	2	2	2	2	2

Table 4: Estimated number of cointegrating vectors, using the trace and maximal eigenvalue test statistics, 1960 - 2006. Columns correspond to different deterministic components (intercept, C, and trend, T) included in the VAR, with or without restrictions.

Dep. Var.	Y	Y	Y	Y	Y	Y	Y
Model	ARDL(2,1,0,1)	ARDL(2,1,0,0)	ARDL(1,1,0,1)	ARDL(2,0,0,0,1)	ARDL(1,0,0,1,2)	ARDL(2,1,0,0)	ARDL(2,1,0,1)
K	1.586*	1.222*	1.180*	2.547*	1.853*	1.879*	1.699*
	(0.373)	(0.619)	(0.585)	(0.480)	(0.331)	(0.725)	(0.430)
I						-1.366*	-1.283*
						(0.676)	(0.455)
PI	-1.281*			-2.640*	-1.759*		
	(0.434)			(0.507)	(0.308)		
SI		-0.622	-0.663	0.417*	0.234		
		(0.453)	(0.420)	(0.199)	(0.211)		
PROP		0.001		-0.005*		-0.002	
		0.001		(0.001)		(0.001)	
FRAC	-0.005*				-0.006*		-0.004
	(0.002)				(0.002)		(0.002)
RISK			0.006				
			(0.004)				
ECM	-0.321*	-0.218	-0.247*	-0.416*	-0.486*	-0.285*	-0.311*
	(0.111)	(0.116)	(0.113)	(0.138)	(0.142)	(0.118)	(0.112)

Table 5: Results obtained under *ARDL* estimation including a constant and trend. Reported coefficients are for the long-run relationships. Standard errors are in parentheses. * denotes significance at the 5% level. Sample period is 1960 to 2006.

Dep. Var.	Y	Y	Y	Y
Model	ARDL(1,1,0,1,0)	ARDL(2,1,0,2,2)	ARDL(1,1,0,0,1)	ARDL(1,1,0,1,0)
K	0.566	1.185*	0.479	2.0313*
	(1.116)	(0.601)	(0.271)	(0.429)
I				-2.326*
				(0.822)
PI	-1.229			
	(1.240)			
SI		-0.653	-0.252	
		(0.429)	(0.263)	
PROP			0.060*	
			(0.026)	
PROP*SIS			-0.007*	
			(0.003)	
FRAC	-0.185			-0.175
	(0.117)			(0.134)
FRAC*IS				0.017
				(0.013)
FRAC*PIS	0.018			
	(0.012)			
RISK		0.041		
		(0.155)		
RISK*SIS		0.042		
		(0.155)		
ECM	-0.409*	-0.486	-0.245*	-0.341*
	(0.128)	(0.135)	(0.115)	(0.127)

Table 6: Results obtained under *ARDL* estimation including a constant and trend. Reported coefficients are for the long-run relationships. Standard errors are in parentheses. * denotes significance at the 5% level. Sample period is 1960 to 2006.

	Model 1		Model 2		Model 3	
Variable	CV1	CV2	CV1	CV2	CV1	CV2
Y	1	-0.345†	1	0.139	1	0
	–	(0.305)	–	(0.352)	–	–
K	-1.544	1	-1.241	1	-0.964	1
	(1.932)	–	(8.998)	–	(4.477)	–
F	0	0.121†	0	0.282†	0	0.257†
	–	(0.118)	–	(0.204)	–	(0.185)
H	-0.235†	0	-2.691†	0	-1.195	0
	(0.586)	–	(11.295)	–	(2.099)	–
PI	0.684*	-0.689*	1.047	-0.934†	0.657	-0.893†
	(1.684)	(0.189)	(8.355)	(0.143)	(3.914)	(0.072)
RISK	-0.007*	0.002*	4.957	0.140	2.729*	-0.004†
	(0.004)	(0.002)	(18.152)	(0.396)	(3.482)	(0.288)
PI*RISK			-0.499†	-0.014	-0.275*	0.000†
			(1.827)	(0.040)	(0.345)	(0.029)
ECM(-1)	0.253	0.153*	0.034*	0.018*	0.060*	0.048*
	(0.150)	(0.045)	(0.013)	(0.004)	(0.024)	(0.008)
ECM(-1)	1.029	-0.000	0.011	0.285	-0.013	0.317*
	(0.854)	(0.254)	(0.056)	(0.017)	(0.067)	(0.021)

Table 7: Johansen estimation results with standard errors reported in brackets. * denotes significance at the 5% level. † denotes variables whose exclusion restriction could not be tested due to convergence problems. All models estimated for the period 1970 to 2005 with VAR = 1. Left-most model estimated with unrestricted intercepts and trends, others estimated with unrestricted intercepts and restricted trend.

	Model 1		Model 2	
Variable	CV1	CV2	CV1	CV2
Y	1	-0.392*	1	-0.791†
	–	(0.423)	–	(0.483)
K	-1.526†	1	-1.356†	1
	(0.663)	–	(0.622)	–
F	0	-0.553†	0	-0.322†
	–	(1.293)	–	(0.402)
H	-0.323†	0	0.615†	0
	(0.413)	–	(0.491)	–
SI	0.372	-0.391†	0.642	-1.366†
	(0.404)	(0.205)	(0.689)	(1.019)
PROP	-0.005*	0.002	0.652†	-0.144†
	(0.002)	(0.002)	(0.076)	(0.145)
SI*PROP			-0.008†	0.017†
			(0.009)	(0.017)
ECM1	0.237	0.203*	0.081	0.083*
	(0.044)	(0.044)	(0.119)	(0.032)
ECM2	0.236	0.069	-0.047	-0.127*
	(0.046)	(0.046)	(0.169)	(0.045)

Table 8: Johansen estimation results with standard errors reported in brackets. * denotes significance at the 5% level. † denotes variables whose exclusion restriction could not be tested due to convergence problems. All models estimated for the period 1970 to 2005 with VAR = 1. Left-hand model estimated with unrestricted intercepts and restricted trends, right-hand model estimated with unrestricted intercepts and trends.