Specifying and Estimating Partial Equilibrium Models for Use in Macro Models: A Road Map for the KIPPRA–Treasury Macro Model

Alemayehu Geda
Njuguna Ndung'u

Macroeconomic and Economic Modelling Division
Kenya Institute for Public Policy Research and Analysis

KIPPRA Discussion Paper No. 8

June 2001

This work is the outcome of the ISS-MMC-KIPPRA Collaborative Research Project

KIPPRA in brief

The Kenya Institute for Public Policy Research and Analysis (KIPPRA) is an autonomous institute whose primary mission is to conduct public policy research, leading to policy advice. KIPPRA's mission is to produce consistently high quality analysis of key issues of public policy and to contribute to the achievement of national long-term development objectives by positively influencing the decision-making process. These goals are met through effective dissemination of recommendations resulting from analysis and by training policy analysts in the public sector. KIPPRA therefore produces a body of well-researched and documented information on public policy, and in the process it assists in formulating long-term strategic perspectives. KIPPRA serves as a centralized source from which the government and the private sector may obtain information and advice on public policy issues.

Published 2001

Bishops Garden Towers, Bishops Road
PO Box 56445, Nairobi, Kenya
tel: +254 2 719933/4; fax: +254 2 719951
email addresses: admin@kippra.or.ke
for authors: alemayehu@excite.com and snjuguna@kippra.or.ke

ISBN 9966 949 16 X

The Discussion Paper Series disseminate results and reflections from ongoing research activities of the institute's programmes. The papers are not externally refereed and are disseminated to inform and invoke debate on policy issues. Opinions expressed in the papers are entirely those of the author or authors and do not necessarily reflect the views of the Institute.

KIPPRA acknowledges generous support by the European Union (EU), the African Capacity Building Foundation (ACBF), the United States Agency for International Development (USAID), the Department for International Development of the United Kingdom (DfID) and the Government of Kenya (GoK).

Abstract

Formulating research questions and carrying out empirical analysis is not an easy task. It is particularly difficult for practitioners in government ministries, who are bogged down with the day-to-day running of the ministry and have little time left for pursuing the frontier of research. Yet these economists need to use the latest available techniques in their trade so as to arrive at sensible policy analysis. Thus there is a clear gap between practical economists' work and that of academic economists. This paper is an attempt to bridge that gap. It provides the procedure for formulating a research question, shows how to locate the relevant background literature, how to specify the economic and econometric models relevant for the analysis, and how to write the final research report. It illustrates this procedure by using macroeconomic issues from Kenya. With the aim of reaching a wider audience, it is accompanied by an annex detailing the simplified, relevant basic mathematical and econometric principles employed in the main text. We believe it will serve as a guide for improving the KIPPRA Treasury Macro Model of Kenya and hope that it will serve as a valuable tool for those using the macro model and for training, both now and in the future. The paper may also serve as a guide for researchers in applied macroeconomics at the graduate level of their study.

Acknowledgement

We would like to thank Prof. Marc Muyts for constructive comments on an earlier version of this paper. However, any errors are ours.

Abbreviations

ADF	Augmented Dickey-Fuller
CI	Cointegration
ECM	Error Correction Model
IPAR	Institute of Policy Analysis and Research
KTMM	KIPPRA Treasury Macro Model
OLS	Ordinary Least Square
VAR	Vector Auto-Regressive

Contents

1 Introduction

Among the many tasks KIPPRA will have in the future is to make an in-depth analysis of each component of the KIPPRA Treasury Macro Model (KTMM) by re-examining the specifications, updating the data used, and experimenting with the robustness of the components of the KTMM using high-frequency data. This requires not only understanding the KTMM but also an approach to model specification and estimation. This paper provides a rough guide on how to carry out such an exercise. It will also serve as a training tool for both KIPPRA staff and others who will be trained using the KTMM and other macroeconomic analysis and research issues.

This paper also makes the point that the macro department of KIPPRA is engaged not only in working on models but also in building capacity both at KIPPRA and in line ministries and the private sector. The paper serves as an input in that process.

It is organized as follows. In section 2, we discuss how to motivate a particular study in question using background information. In section 3, we describe how the researcher can locate a research question in the existing literature on the topic. Section 4 discusses model specification issues. The section's emphasis is on how to move from research question and an examination of the relevant literature to model specification. Section 5 addresses the transition from model specification to estimation. Section 6 pinpoints issues faced in writing up the research report. To make the discussion tractable and easy to follow, an illustration of how to estimate an investment and inflation equations are used as examples throughout the paper. It should then be straightforward for the practitioner to apply this knowledge to other partial models of the KTMM, such as exchange rate, consumption and exports.

2 Motivating the partial model

Suppose the researcher wants to examine the investment equation specified and estimated in KTMM. When working on this topic you will use your knowledge of econometrics to investigate the macroeconomic determinants of private investment in Kenya, using data for the period, say, 1997–2000. So the model is already specified; the additions required are the type and frequency of the data and econometrics knowledge, especially of the time series.

Investment is central to the growth experience of any country. Here you may be required to investigate empirically the determinants of investment. From macroeconomic theory, you know that there are different theories about investment functions. Keynesian theory, for example, explains the dynamics of investment through the accelerator principle, which puts the emphasis on the demand side. In contrast, neoclassical theory emphasizes the supply side by looking at the user cost of capital. Another approach, associated with Tobin, looks at the discrepancy between the market value of productive assets vis-à-vis their replacement costs to explain new investments.

A further perspective on the dynamics of private investment can be obtained by investigating the interaction between public and private investment in an economy. A key question here is whether public investment has a crowding-out or a crowding-in effect on private investment: that is, whether public investment goes at the expense of private investment or, instead, whether it stimulates private investment. One of the currently celebrated empirical regularities is the complementarity between forms of public investment such as infrastructure and private investment. Finally, in a country like Kenya, determining to what extent political factors influence investment behaviour and investment decisions is a further uncertainty that causes complications.

Hence, we can tackle this topic in various ways. We shall return to this matter of formulating a more precise research question. At this point, suffice it to say that to write a research paper on this topic you will need to carry out a number of steps:

1) Specify a focused question within this general topic and relate it to relevant theories.
2) Study and summarize the relevant literature on your chosen question.
3) Carry out econometric estimation and hypothesis testing, and evaluate your results in the light of your research question.

So far, we have emphasized the role of investment in growth and theories of investment as possible motivation for your research. Another source of motivation is an examination of the pattern of the investment data in Kenya. For instance, Ronge and Kimuyu (1997) took up a good section of their paper to analyse investment using descriptive analysis and graphs. One aspect to enrich such an analysis is to break down investment into various components: fixed investment, inventory, infrastructure (for public investment), and so on. Such detailed descriptive and trend analysis is important to focus on major turning points and the source of such events. In figure 1, for instance, we observe a sharp rise in investment after 1992.

Just by looking at that figure we may not say much. If we turn to one of Ronge and Kimuyu's (1997) graphs, however, we note that this sharp rise is attributed to investment in the transport sector (in particular, investment on equipment). Similarly, Ndung'u and Ngugi (1999) motivated their study of the financial and exchange rate market and its impact on inflation in Kenya by using growth trends, major macro-economic indicators, and the evolution of exchange-rate policy (starting from some decades back and going to the recent era of liberalization) in Kenya. The major point is that you need to have a consistent story to tell—and background information is central to that.

3 Locating your study in the literature

To carry out your research you will need to broaden your theoretical reading through a literature survey. We have provided you with a list of such materials in the reading list at the back of this booklet. The set of materials comprises basic readings you should consult to enable you to formulate your econometric model on the investment function for Kenya. This is a basic set of readings; obviously, we encourage you to look for further materials as well.

For your convenience, we have classified these materials under three headings: 1) basic theoretical readings on investment behaviour, 2) readings on applying econometrics in estimating investment functions, and 3) empirical studies of investment behaviour in developing countries. This reading needs to be complemented by general reading on the Kenyan economy and on investment behaviour in Kenya. If you were estimating an export equation instead, you would have needed to undertake a similar set of readings about exports.

4 Model specification

4.1 Formulating testable empirical questions

Successful research requires formulating the question you want to investigate in a way that makes investigation possible. At the start of your work, you should spend time thinking about the research question you wish to study and how to make that question manageable. Then, after you have read basic literature such as that in the reading list and thought through the way authors have set out their research questions, you should try to formulate your questions. You may wish to repeat their research questions and relate them to Kenyan data. That is a legitimate scientific procedure. But in most cases, you will need to establish clear value added in the research questions you propose. This is not difficult. It may be that you want to update the data and include a recent period of analysis; it may be that there have been major policy shifts during the update period; or it may be that you have discovered weaknesses in other research works.

A further constraint on the type of research questions you can deal with in an econometric study relates to the data at your disposal. Clearly, the way you formulate your research question will depend on whether or not the data are available. For example, the database may not allow you to estimate sectoral investment functions (that is, separate functions for agriculture, mining, manufacturing, and so on), but you may distinguish between private and public investment. Similarly, you may find it difficult to estimate Tobin's (1969) *q*-model because data on replacement costs are not usually available. Hence, when thinking about a research question, take account of the data at your disposal and decide how to make the best use of them.

Before you read on as to how to formulate a more specific research question, we suggest that you take the time to have a good look at the evolution of the variable you seek to explain: for example, private investment. Since we are interested in the evolution of the volume of investment, it is best to use investment at constant prices. Some authors, however, also use investment as a share of GDP as their dependent variable. You might wish to consider both variants, although we suggest you use the former variant: investment at constant prices. Now plot this variable against time and take a good look at it. Better still, plot the logarithms of this variable against time. Why is this better? Recall from your econometrics study why a plot of the logarithm of an economic time series against time is superior for detecting trends and its implication for distribution.

If you have looked at the evolution of investment over time, you will have noted that the 1990s (the grey-coloured area) witnessed a marked upsurge in private investment (note, however, the degree of this upturn in the log version). That is, the growth rate picked up. This is not surprising in view of the huge reform and a major shift in major macro aggregates following the intensification of liberalization. One would expect that this might affect private investment behaviour. This gives you an interesting clue: do you expect the investment function to be stable over the period as a whole, or might there have been a structural break in the 1990s that political and reform factors caused? This is a question you may want to investigate econometrically once you have decided on a particular model.

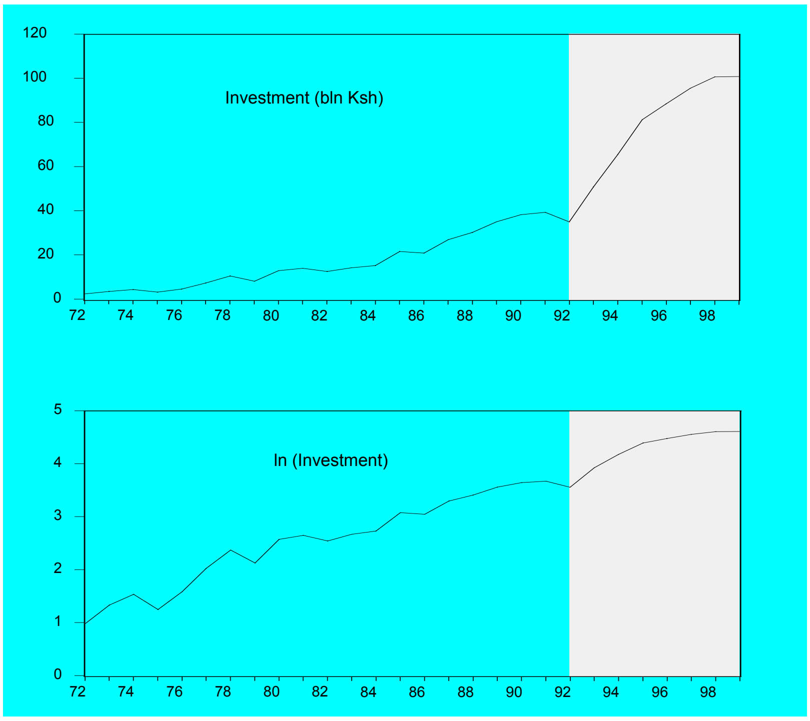

Figure 1. Pattern of investment in Kenya: 1972–1999.

Another question you may want to ask is why investment seems to have surged upward at this time. Is it the type of investment? Here you may want to separate short-term investment in financial assets from fixed investments. One could also separate fixed investment into 'productive' investment— machinery, and so on, from simply building. A further question would be to ask why investment is surging upward but the economy sinking into a recession in this period. Then the question of productivity becomes relevant.

4.2 Narrowing down your research question

Let's say your general question concerns the determinants of private investment in Kenya. Depending on the nature of your work, you may be expected to deal with the issue in all its forms, or you may not be expected to deal with this question exhaustively. If the latter, it is best to narrow down your task by setting yourself a more specific task within the confines of this general question. Here are two broad suggestions as to how you might do this.

1) One possible avenue is to compare and contrast demand-side with supply-side theories of investment. In this case, we suggest that you compare the accelerator and the user cost models in terms of how well they explain private investment behaviour in the Kenyan economy. You might even decide to end up with a mongrel (= mixed) specification that borrows elements from each in a single specification.

2) Another avenue is to examine whether public investment crowds out or crowds in private investment. That is, your specific research question will be to assess the impact of public investment on private investment.

Each of these avenues gives your research a more limited task. The former involves comparing how well different theories of investment explain the empirical evidence; the latter investigates a specific policy question. In each case, however, it is important to think carefully about the specification of your model. Which variables should you include? Should you use log-transformed variables or not? How should you deal with lagged variables? What are the expected signs of the coefficients in the model? Which hypotheses are interesting to test? Is the model stable across the period as a whole?

As to the last question, we suggest that you investigate specifically whether model stability pertained during the 1990s. To do this, you can use Chow's test or some other technique. Alternatively, you may wish to use dummy variables to single out the 1990s. The important point is not just to introduce dummies but also to explain their meaning. Before you introduce dummies, it is always advisable to estimate the model recursively, understand the profile of the coefficients estimated, and then locate the effect the dummies will have in stabilizing the coefficients.

4.3 Coping with model specifications

Specifying a model is a task that requires skill and experience. This is particularly true when we work with time-series data, where we often use lagged variables to denote that what happened in the past has repercussions for the present. For example, in dealing with investment functions, you may wish to try out a specific type of models that aims to capture the fact that adjustment is neither smooth nor immediate. This is the partial adjustment model explained in box 1.

Box 2 briefly discusses the basic specification of the naive and flexible accelerator models that you may need to use in your research.

Box 3 briefly reviews the user cost model, which, for example, underlies most of the World Bank type of estimations of investment functions in African economies.

To be able to deal with the problem of crowding in or out of private investment by public investment, it is necessary to bring public investment explicitly into the picture as an explanatory variable. By way of example, box 4 shows you how to do this in the context of a mixed specification using a partial adjustment model. You may wish to try out your own pet theory instead.

When dealing with the crowding-in–crowding-out hypothesis, it matters what assumptions you make about the lags of the explanatory variables. The important point is that your

Box 1. Partial adjustment models

Suppose you have dependent variable *Y*. The change in *Y* at time *t* is the difference between its current and its preceding period actual level, given as

$$\Delta Y = Y_t - Y_{t-1}$$

What determines the change in *Y* from period to period?

In a partial adjustment model we make two assumptions as to the nature of this change:

First, we assume that the desired level of *Y, Y**, depends on a number of variables: as follows:

$Y^* = f(X_1, X_2, ..., X_n);$

Second, we assume that in each period, *Y* only partially adjusts to its desired level, as follows:

$$\Delta Y = \lambda(Y_t^* - Y_{t-1})$$

where

$$0 \leq \lambda \leq 1$$

that is, ΔY is a certain (say λ) proportion of the difference between the desired and the preceding year's level of *Y*. The parameter λ is the partial adjustment coefficient. If $\lambda = 1$, it follows that adjustment is immediate. If $\lambda = 0$, however, no adjustment takes place. Within these boundaries, the larger the value of λ the faster will be the adjustment.

Substituting equation 2 into equation 3 yields,

$$\Delta Y = \lambda\left[f(X_1, X_2, ..., X_n) - Y_{t-1}\right]$$

and using equation 1 we obtain the following partial adjustment model:

$$\mathrm{Y_t} = \lambda\ f(\mathrm{X_1, X_2, ..., X_n}) + (1 - \lambda)\,\mathrm{Y_{t-1}}$$

Now, provided that equation 2 is linear in the **X** variables, equation 6 can be estimated with linear regression techniques after an error term is added to the equation. The specification of equation 2 obviously depends on your theoretical model (for example, accelerator, user costs).

assumptions about these lags should be reasonable. For example, on the one hand, private investors may immediately react to initiating construction of new infrastructure that may take time to complete, or they may respond with a certain lag after its completion. On the other hand, high levels of government investment may reduce private capital's access to finance.

Box 2. Naive and flexible accelerator models

The naive accelerator model

The 'naïve' accelerator model assumes that investment is determined by the change in the level of output. If we denote capital stock by *K* and output by *Q,*

$$K_t - K_{t-1} = \beta (Q_t - Q_{t-1}) \quad [1]$$

or equivalently:

$$\Delta K_t = I_t = \beta \ \Delta Q_t \quad [2]$$

where β is the acceleration coefficient that gives us the increase in the capital

rather naive model since it assumes that the level of present-day investment is influenced only by changes in output in the current period.

The flexible accelerator model

In modelling the accelerator mechanism, it is more appropriate to take account of lagged effects. This is the basic idea behind the 'flexible' accelerator model. Hence, this model assumes that investment is determined not only by the current change in output but also by its earlier changes. To model a flexible accelerator, you can introduce one or more lags of the right-hand variable, ΔQ, in equation 2. Alternatively, you can also try using the partial adjustment model.

One way to apply the partial adjustment model in this case is by assuming that the desired level of the capital stock depends on the current level of output, as follows:

$$K_t^* = \beta \ Q_t \quad [3]$$

or alternatively:

$$I^*_t = \Delta K_t^* = \beta \Delta Q_t \quad [4]$$

The partial adjustment mechanism (see box 1) is now given by

$$K_t = \lambda \beta Q_t + (1 - \lambda) K_{t-1} \quad [5]$$

Obviously, if you do not have data on the capital stock it is best to take first

$$I_t = \lambda \beta \Delta Q_t + (1 - \lambda) I_{t-1} \quad [6]$$

This resulting equation lends itself to estimation and hypothesis testing using

For these reasons, it is advisable to try out various lags (say, zero to five years) to check the effect of public investment on private investment. In Kenya, for instance, we found current and five-year lagged public investment have statistically significant (positive) impact on private investment.

Box 3. The user-cost model

Let Q be the real level of output, K the capital stock, L labour, and I investment. The investment function can then be derived, following

$$Q = AK^{\alpha} L^{\beta} \quad [1]$$

Assume diminishing returns to scale, that is, $\alpha + \beta < 1$.

The first-order condition for profit maximization is that the marginal product of capital, the partial derivative of Q with respect to K, equals the marginal real cost of capital given by c / p, where c is the rental cost of capital and p is the

$$\frac{\partial Q}{\partial K} = \frac{c}{p} \quad [2]$$

Since $\frac{\partial Q}{\partial K} = \alpha A L^{\beta} K^{\alpha-1} = \frac{\alpha Q}{K}$, equation 2 can be written as

$$\frac{\alpha Q}{K} = \frac{c}{p} \quad \Rightarrow \quad K = \alpha \frac{p\,Q}{c} \quad [3]$$

By taking first differences on both side of equation 3, we obtain

$$I_t = \alpha \Delta \left(\frac{p\,Q}{c} \right) \quad [4]$$

If we assume, however, that there is a delivery lag (that is, investment in the current period is the sum of investment orders made in preceding years),

$$I_t = \sum_{i=0}^{n} \lambda_i \Delta \left(\frac{pQ}{c} \right)_{t-i} \quad [5]$$

In econometric application, a useful proxy for c / p is the real lending rate.

Alternatively, you could use a partial adjustment model by re-expressing

this model, using the framework developed in box 1.

You may also be interested in checking or testing other new explanations or perceived wisdom on private investment pause. One such argument is that fixed investments are sunk costs once made, and they are also sector specific. The would-be investor has only one option—the timing of the investment. To the extent that there are uncertainties in future returns, this will bid the value of waiting and she will invest in information gathering to reduce the risk of her returns in future. This is a

Box 4. The crowding-in/out model

Let

$$I_p = \lambda \, (K_{pt}{}^{*} - K_{pt-1}) \quad [1]$$

and let

$$K_{pt}^{*} = \beta_1 Y_t + \beta_2 K_{gt-1} + \beta_3 r_t \quad [2]$$

substituting equation 2 into equation 1 and taking first difference yields

$$I_{pt} = \gamma_1 (Y_t - Y_{t-1}) + \gamma_2 I_{g\,t-1} + \gamma_3 (r_t - r_{t-1}) + \gamma_4 I_{p\ t-1} \quad [3]$$

Equation 3 is a model that allows you to check whether public investment has a crowding-in or crowding-out effect on private investment.

coordination problem and has been used to explain investment pause in sub-Saharan Africa countries. The model then has to introduce variability of key prices and macro-indicators that affect private investment in the model.

5 From model specification to estimation

It is one thing to have specified your theoretical model but quite another to estimate it. Estimation requires not only skill but also patience. Since most macro models rely on time-series data, we will focus on estimation based on time series—that is, on macroeconometrics. We will use Ndung'u and Ngugi's (1999) recent work on Kenya to motivate the exposition.

The first main point that you need to note is that econometrics is not a mechanical procedure. You need to withdraw yourself from the mechanics of it and see the whole issue from a wider perspective, and in particular, in light of the questions posed at the beginning of your inquiry. The second important point is that since you are confronting your theory with data at this stage, you have a lot of room to modify it in the light of those data—you should repeatedly move back and forth between data and theory. This means going back to your original theory to revise it further, coming back to your data with the revised theory . . . and so on. In this interactive process, you need to make sure that each of your steps in estimation is a justifiable move. You should not simply move mechanically from unit root test, to cointegration, to error-correction modelling, and so on. Each step should justify your conclusions at every stage.

To bring this argument home, let us see how Ndung'u and Ngugi (1999) estimated their mode of inflation in Kenya. They started by specifying a theoretical model where inflation is assumed to originate from both the demand and the supply side. The demand side is represented by the demand for money equation. The supply side is made up of the foreign sector. Before they actually specified the model, they set their work in the context of the literature of both Kenya and other developing countries. To move to estimation, they first specified the theoretical money demand equation as

$$M2_t - P_t = c_1 RY_t - c_2 TDR_t \qquad [1]$$

where *M*2 is the log of nominal money stock, *P* is the log of domestic price level, *RY* is the log of real income and *TDR* is the rate of interest.

Since their interest is in inflation, they have expressed equation 1 as a price equation in equation 2 by invoking an assumption that prices adjust to excess supply in the money market.

$$P_t = c_0 M2_t - c_1 RY_t + c_2 TDR_t \qquad [2]$$

The next step is a major leap from theory to estimation: they rewrote equation 2 in a general autoregressive-distributed lag form of order *m* [*AR*(*m*)] since the adjustment period is not known *a priori*, thus getting

$$\begin{aligned} P_t = & \sum_{i=1}^{m} \alpha_{1i} P_{t-i} + \sum_{i=0}^{m} \alpha_{2i} M2_{t-i} \\ & + \sum_{i=0}^{m} \alpha_{3i} RY_{t-i} + \sum_{i=0}^{m} \alpha_{4i} TDR_{t-i} + \varepsilon_{1t} \end{aligned} \qquad [3]$$

Notice the introduction of the ε_i term, which is a departure from exact theoretical specification to a stochastic one in equation 2.

Equation 3 would be readily estimated were it not for the spurious regression problem that characterizes estimating models with non-stationary series such as *P, M*2, *RY* and *TDR*. Spurious regression is regression with a non-stationary series that renders what are called nonsense results. This is because variables that are likely to be non-stationary will have a stochastic trend, and where variables share a trend it will be difficult to disentangle the true relationship from the trend influence. Thus you need stationary series or their linear combination, which is stationary. See below how to run a sensible regression. You could formally investigate whether a series is stationary or not (although invariably time-series macro variables are non-stationary) using various tests.[1] Ndung'u and Ngugi (1999) did that using the Augmented Dickey-Fuller (ADF) and the Phillips-Perron (PP) tests, which are among the popular such tests usually found in econometric software packages.

Differencing the series above would tackle the non-stationary problem. However, you would lose the long-run information in the data, which is central to your theoretical model. Thus you need to think of a mechanism by which you can tackle the problem of spurious regression and have the long-run information as well. This basically requires combining the short-run (differenced) equation with the long-run (level-based) equation in one model. This can be done provided you can find a vector that renders a linear combination of the level variables that is stationary. This is the cointegrating vector, and the method is termed 'cointegration analysis'. It is one of the recent

[1] In most cases, we need to perform these tests so that we can distinguish between macro variables that are trend stationary from those that are difference stationary. Care should be taken here in the tests.

developments in time-series econometrics. To carry out such an investigation, Ndung'u and Ngugi (1999) reset the parameters of equation 3 into equation 4, calling it the Error Correction Model.

(If you find the formulation of equation 4 below a bit complicated, see Thomas (1997) and Kennedy (1992) then move progressively to Enders (1996), Rao (1994), Engel and Granger (1991) and Greene (1993). See also the mathematical appendix attached for a simplified treatment of some of the basic issues. Software packages such as Eviews and PC-Give/ PC-Fimil are also very helpful. Here, we are concerned with an outline of the procedure and not the details of the derivation and the techniques involved as such.)

$$\Delta P_t = \sum_{i=1}^{m-1} \alpha_{6i} \Delta P_{t-i} + \sum_{i=0}^{m-1} \psi_{1i} \Delta M2_{t-i} + \sum_{i=0}^{m-1} \psi_{2i} \Delta RY_{t-i} + \sum_{i=0}^{m-1} \psi_{3i} \Delta TDR_{t-i} + \alpha_5 \left[M2 - \beta_1 P - \beta_2 RY + \beta_3 TDR \right]_{t-m} + \varepsilon_{1t} \quad [4]$$

where

$\alpha_{6i} = (\alpha_1 - 1)_i + \alpha_{1, i+1} + \ldots + \alpha_{1, m-1}$ for $i = 1, \ldots m - 1$

$\Psi_{j,i} = \alpha_{j,i} + \alpha_{j+1, i+1} + \ldots + \alpha_{4, m-1}$ for $j = 1 \ldots 3$ and $i = 1, \ldots m - 1$

$\alpha_5 = \alpha_{2,i} + \alpha_{2, i+1} + \ldots + \alpha_{4, m-1}$ for $i = 0, \ldots, m - 1$

The term in the square bracket [ ] shows the long-run model.

Equation 4 is the estimatable version of the theoretical model (equation 2) from the demand side. However, Ndung'u and Ngugi (1999) do not think that inflationary pressure comes from the demand side alone. They hypothesized that supply-side effects are also important and can be depicted by capturing the effect of imported input prices and the depreciation of the foreign exchange rate. Note here that they are effectively specifying another long-run (cointegrating vector) relationship. In fact, they found supporting evidence for this hypothesis from their empirical work, too—from their test for cointegration. This test does not reject the possibility of having two cointegrating vectors (although it does reject the possibility of having three). They thus specified this additional relationship as

$$P_t = c_3 NER_t + c_4 WP_t \quad [5]$$

where *NER* is the nominal exchange rate and *WP* is a foreign price index, all in log. This is what is called the exchange-rate pass-through equation, which can be written in autoregressive distributed lag mode of order *m* (similar to equation 3 above) as,

$$P_t = \sum_{i=1}^{m} \alpha_{7i} P_{t-i} + \sum_{i=0}^{m} \alpha_{8i} NER_{t-i} + \sum_{i=0}^{m} \alpha_{9i} WP_{t-i} + \varepsilon_{2t} \quad [6]$$

Similar to equation 4 above, the Error Correction Model (ECM) version of equation 6 can be given by

$$\begin{aligned}\Delta P_t = &\sum_{i=1}^{m-1} \alpha_{10i} \Delta P_{t-i} + \sum_{i=0}^{m-1} \psi_{4i} \Delta NER_{t-i} + \sum_{i=0}^{m-1} \psi_{5i} \Delta WP_{t-i} \\ &+ \alpha_{11} \left[NER - \beta_4 P + \beta_5 WP \right]_{t-m} + \varepsilon_{3t}\end{aligned} \quad [7]$$

where

$\alpha_{10i} = (\alpha_7 - 1)_i + \alpha_{7, I+1} + \ldots + \alpha_{7, m-1}$ for $I = 1, \ldots m - 1$
$\Psi_{j,i} = \alpha_{j,i} + \alpha_{j+1, I+1} + \ldots + \alpha_{9, m-1}$ for $j = 4, 5$ and $i = 1, \ldots m - 1$
$\alpha_{11} = \alpha_{8,i} + \alpha_{8, I+1} + \ldots + \alpha_{9, m-1}$ for $I = 0, \ldots, m - 1$

The term in the square bracket [ ] shows the second long-run model (error-correction term).

The final stage in your specification of the inflation model for Kenya, following Ndung'u and Ngugi's formulation, is to combine equation 4 (the demand-side effect) and equation 5 (the supply-side effect) in one model, shown in equation 8.

$$\begin{aligned}\Delta P_t = &\sum_{i=1}^{m-1} \theta_{1i} \Delta P_{t-i} + \sum_{i=0}^{m-1} \theta_{2i} \Delta M2_{t-i} + \sum_{i=0}^{m-1} \theta_{3i} \Delta RY_{t-i} \\ &+ \sum_{i=0}^{m-1} \theta_{4i} \Delta TDR_{t-i} + \sum_{i=0}^{m-1} \theta_{5i} \Delta NER_{t-i} + \sum_{i=0}^{m-1} \theta_{6i} \Delta WP_{t-i} \\ &+ \alpha_{21} \left[M2 - \beta_1 P + \beta_2 RY + \beta_3 TDR \right]_{t-m} \\ &+ \alpha_{22} \left[NER - \beta_4 P + \beta_5 WP \right]_{t-m} + \theta_0 + \psi_i D_{st} + \varepsilon_{0t}\end{aligned} \quad [8]$$

where θ is a constant and D_{st} is a vector of dummies.

This completes your journey of moving from theoretical model specification to estimation. Notice that in this journey other supply-side factors, like the labour market, are not important influences or that data for them are not available. Addressing these factors can be a useful starting point for updating the model in future and testing its conclusions further. Once you have used equation 8 above, you can proceed to estimate the long-run equations, handling the two error-correction forms separately. The two vectors, based on Ndung'u and Ngugi's estimates, are

$$M2_t - 1.72P_t - 1.42Ry_t + 0.32TDR_t \quad [9a]$$

$$NER_t - 1.195P_t + 0.95WP_t \quad [9b]$$

Having found these long-run (equilibrium) relationships—the cointegrating vectors—you can examine the plausibility of the theory outlined. For instance, does the homogeneity assumption in the money-demand equation hold? The result above does not seem to support that, since the coefficients of price in the first vector are different from unity (you can actually formally test that!). You may now also proceed to estimate the final error-correction model given by equation 8 by substituting the variables in the two square brackets (the error-correction terms) for the two residuals derived from the estimation of the two long-run equations shown in equations 9a and 9b. The next step is how to present your results to your readers.

6 Writing up your research output

Writing up the results of an econometric study requires a certain skill. It is an art you need to develop when you engage yourself in research. You need to pay attention to this task, as your efforts will remain buried and in vain if you fail to communicate. Extracts from a chapter in Griffith et al. (1993:

837–838) can help you organize your findings and write up your research.

At the start of your research, after you have read the available literature, it is useful to write a draft abstract that you can discuss with your colleagues. Box 5 gives you advice on what to include in your abstract and box 6 gives you advice on how to proceed (both from Griffith et al. 1993). These two boxes provide a summary of the steps and the thinking that should go into the process.

At the end of your paper, you will need to write an elaborate conclusion. Because your research agenda includes updating research work and informing policy, your conclusion will have to tie up these two aspects and deliver the message. As you understand the subject matter well, the challenge is how to

Box 5. Writing an abstract

Your abstract should include

1) a concise statement of the problem
2) comments on the information that is available, based on the contributions others may have made in this area
3) specification of the research design that includes
    a) the economic models relevant to providing information on the questions posed and the decision context in which the information will be used (economic model)
    b) specification of the corresponding statistical model that specified the sampling process by which you visualize that the underlying data should have been or may have been generated
    c) suggestion for collecting or obtaining sample observations consistent with the statistical models
    d) outlines of the estimation and the hypothesis-testing procedure and the corresponding sampling and power properties; it should be indicated whether or not point and interval estimates of the unknown parameters will be obtained along with point and interval forecasts
    e) the potential economic and statistical implication of the results

Source: Griffith et al. (1993:837)

present this knowledge in a language that policymakers can understand and will appreciate. After writing the conclusion, you will need to write an executive summary, which will appear at the beginning of the paper. It in turn can be repackaged to become a policy brief. These presentations provide a bridge—a useful link from research to advising on policy and to policy formulation.

Box 6. Writing up

Writing up the research results includes many of the items covered in the research abstract.

1) The place to start your research report is with the questions you have investigated, why they are important, who should be interested in the results, and how they relate to previous work done in this area. Identify some of the potential payoffs and state what is to be covered in each section.
2) Present the economic models you have used, define the economic variables, and state the assumptions that have been made and the hypotheses suggested.
3) Given the economic model, specify the corresponding statistical models, the variables to be included in each statistical model, the functional form(s), the stochastic characteristics of the equation errors, and any other relevant assumptions.
4) Identify the data you have used, state how they were generated or collected, discuss whether or not you believe that they are consistent with the sampling process specified in the statistical models, and make comments on their quality and the population(s) to which they apply.
5) Given the economic and statistical models, identify the estimation and inference procedures that you have used and be specific concerning the sampling properties of the estimation rules and the power of the tests. If alternative estimation rules or test statistics or both might have been appropriate, state why you made a particular choice.
6) Once you have carried through the econometric analysis, the next section should contain the empirical results. In the results section, you should present the estimated economic relations and the corresponding statistics and comment on their statistical significance (sampling precision, *t*-statistics, and so on). Summarize the statistical inferences as appropriate. *(continued on next page)*

(continued from previous page)

7) The next section should contain the economic implications or consequences of your results. These economic implications usually relate to a decision or a choice. For a firm, it might relate to pricing policy or scale. On an economic policy level, it might relate to Central Bank actions or taxing and spending alternatives.
8) Given the statistical and economic implications of your investigation, you should then comment on some of the shortcomings of your results. Have you used the correct economic and statistical models? Would an alternative statistical model have been more appropriate? What about possible measurement errors in your data? Have you used the most efficient estimation procedures and test statistics? Does the estimated sampling precision underlying your parameter estimates permit you to make economic decisions with the necessary level of confidence?
9) Finally, if you were going to carry on this research, what suggestions do you have for further study? What about the adequacy of the sample data? Should other economic and statistical models have been considered? Were the estimation procedures appropriate, or should other estimation rules have been used? What new questions has the research suggested or, knowing what you now know, what would be the nature of your research?

Source: Griffith et al. (1993:837–838)

Mathematical appendix

Points on matrix algebra for cointegration analysis

Time-series analysis, which is important in macroeconometric modelling, requires some basic concepts in matrix algebra. This annex is aimed at providing such input by simplifying the treatment of these concepts in major textbooks and picking the most relevant ones for cointegration (CI) analysis.

Linear dependence and rank of a matrix

The two concepts of linear dependence and rank of a matrix are widely used in CI analysis. Solution to a matrix requires that the matrix is non-singular. Non-singularity implies that the determinant of the matrix is different from zero. This in turn requires that each row of the matrix is *linearly independent.* A row or column is linearly independent if it cannot be formulated as a linear combination of another row or column. The number of linearly independent rows or columns[1] of a matrix is called the *rank of a matrix*. The most important question in the context of CI analysis is how do we know that a matrix has linearly independent rows?

One important method is to check whether its determinant is equal to or different from zero. If it is equal to zero (that is, $|\mathbf{A}| = 0$) definitely matrix **A** has linearly dependent rows. If on the other hand $|\mathbf{A}| \neq 0$, then matrix **A**'s rows are said to be linearly independent. Thus, a non-singular matrix (a matrix with a non-zero determinant) is said to have a *full rank*. Note also that for a matrix of dimension (n x m) the rank can have a maximum value of n or m, whichever is the minimum. It should be the minimum, because a determinant is defined only for a square matrix. Thus, $r(A) \leq Min\{m, n\}$. So we require that $m = n$ at all times in the type of analysis we consider.

An important concept in matrix algebra, which is widely used in cointegration analysis and related to the discussion above, is the

[1] Since using either a row or a column makes no difference, we will use the row in the exposition.

characteristic root (Eigen values) and the *characteristic vector (Eigen vector)*.[2] This is discussed in the next section.

The concept of Eigen values (and Eigen vectors)

Given an n x n matrix **D**, we may find a scalar r and an n x 1 vector **X**, **X** ≠ 0 such that the matrix equation below is satisfied, that is,

$$\underset{n\times n}{D}\,\underset{n\times 1}{X} = r\,\underset{n\times 1}{X} \qquad \text{[A1]}$$

If expression A1 is satisfied, then

- the scalar r is referred to as *the characteristic root (Eigen value)* of matrix **D**
- the vector **X** is referred as the *characteristic vector (Eigen vector)* of matrix **D**

Equation A1 can be rewritten as

$(\mathbf{D} - r\mathbf{I})\mathbf{X} = \mathbf{O}$ where $\mathbf{O} = n$ x 1 matrix [A2]

If we need a non-trivial solution, the coefficient matrix $(\mathbf{D} - r\mathbf{I})\mathbf{X}$, called the *characteristic matrix* of matrix **D**, needs to be singular (that is, its determinant vanishes):

$$|\mathbf{D} - r\mathbf{I}| = \begin{vmatrix} d_{11}-r & d_{12} & \dots & d_{1n} \\ d_{21} & d_{22}-r & \dots & d_{21} \\ . & . & . & . \\ d_{n1} & d_{n2} & \dots & d_{nn}-r \end{vmatrix} = 0 \qquad \text{[A3]}$$

The expression in equation A3 is what is referred as the *characteristic equation* of matrix **D**. This determinant, upon *Laplace expansion*,[3] gives an n^{th}-degree polynomial in the variable r. This will have n roots (r_1, r_2 … r_n), each of which is a *characteristic root (Eigen value)*.

---

[2] Characteristic roots (vectors), Eigen values (vectors), and latent roots (vectors) are synonymous. Time-series literature often uses the term Eigen value (vector). Every Eigen value will be related to an Eigen vector of the variables in the analysis.

[3]

Inserting such Eigen values into the equation system ($\mathbf{D} - r\mathbf{I})\mathbf{X} = \mathbf{O}$ will produce a vector that we term as *Eigen vector* (note that the determinant of the equation $|\mathbf{D} - r\mathbf{I}|$ will vanish). The $(\mathbf{D} - r\mathbf{I})\mathbf{X} = \mathbf{O}$ system, however, generates an infinite number of vectors corresponding to the Eigen values r_i because it is a homogenous system.[4] However, by normalization, a characteristic root can be selected from this set of infinite solution. Imposing a unit circle function on the solution usually does such normalization. The latter gives stability to the function. If the characteristic roots lie outside the unit circle, the solution for the polynomial in question will be explosive. Since this concept is important in CI analysis, we have illustrated its derivation using a numerical example (see Chiang 1984:327).

$$\text{If } \mathbf{D} = \begin{bmatrix} 2 & 2 \\ 2 & -1 \end{bmatrix} \qquad \text{[A4]}$$

$$\Rightarrow \quad |\mathbf{D} - r\mathbf{I}| \begin{vmatrix} 2-r & 2 \\ 2 & -1-r \end{vmatrix} = r^2 - r - 6 = 0 \qquad \text{[A5]}$$

with characteristics roots (Eigen values) $r_1 = 3$ and $r_2 = -2$

When these roots are used in equation $(\mathbf{D} - r\mathbf{I})\mathbf{X} = \mathbf{O}$ we get

$$\Rightarrow \quad \begin{bmatrix} 2-3 & 2 \\ 2 & -1-3 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} -1 & 2 \\ 2 & -4 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad \text{[A6]}$$

Note that the rows are linearly dependent. The system A6 results in

$$X_1 = 2X_2 \qquad \text{[A7]}$$

A unique solution can be found out of this infinite solution possibility by imposing the unit circle function $X_1^2 + X_2^2 = 1$

$\Rightarrow X_1^2 + X_2^2 = (2X_2)^2 + X_2^2 = 1 \quad \leftarrow$ inserting $2\mathbf{X}_2$ in place of $\mathbf{X}_1$, as given by equation A7.

$$\Rightarrow 5X_2^2 = 1 \text{ and, hence, } X_2 = 1/\sqrt{5} \text{ and } X_1 = 2/\sqrt{5}$$

---

[4] A matrix system is homogenous when it has zero (as opposed to a constant) on the right-hand side of the equality sign.

Thus, the first characteristic vector (Eigen vector) is given by

$$V_1 = \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix}$$

You can do the same using $r_2 = -2$.

The concept of rank and Eigen value or vector is central to the cointegration test using the Johansen approach. This is because testing the existence of a non-zero Eigen value is tantamount to testing to the rank of a matrix. The latter in turn indicates the number of linearly independent rows. This follows from the fact that the rank of a matrix is equal to the number of non-zero Eigen values.[5] Thus if we have

$$\mathbf{X}_t = \mathbf{A}_1 \mathbf{X}_{t-1} + \boldsymbol{\varepsilon}_t \qquad \text{[A8]}$$

we can have [A9] by subtracting X_{t-1} from both sides of equation [A8].

$$\begin{aligned} \Delta \mathbf{X}_t &= \mathbf{A}_1 \mathbf{X}_{t-1} - \mathbf{X}_{t-1} + \boldsymbol{\varepsilon}_t \\ &= (\mathbf{A}_1 - \mathbf{I})\mathbf{X}_{t-1} + \boldsymbol{\varepsilon}_t \qquad \text{[A9]} \\ &= \prod \mathbf{X}_{t-1} + \boldsymbol{\varepsilon}_t \end{aligned}$$

where $\mathbf{X}_t$ and $\boldsymbol{\varepsilon}_t$ are (n x 1) vectors, $\mathbf{A}_1$ and $\mathbf{I}$ (n x n) matrices of parameters and identity, respectively, and $\Pi = (\mathbf{A}_1 - \mathbf{I})$.

The rank of ($\mathbf{A}_1 - \mathbf{I}$) equals the number of cointegrating vectors. If the rank = 0, all the $\Delta\mathbf{X}_{it}$ are unit root processes. They are not cointegrated for we lack the linear combination of $\mathbf{X}_{it}$. The system A9 will also be a convergent system of differences equations if we rule out Eigen values that are greater than unity (or if we impose the unit

[5] If all rows of a matrix are linearly independent, the determinants of this matrix are not equal to zero. It follows that none of the Eigen values can equal zero when the determinants of the matrix are different from zero. If the rank equals zero, on the other hand, each element of the matrix must equal zero and all Eigen values will be equal to zero. In the intermediate case, r linearly independent rows ($0 <$ rank of the matrix $= r < n$) will contain r Eigen values (the other ($n - r$) Eigen values being zero). So the Johansen procedure requires testing for the most significant non-zero Eigen value, which has a corresponding linear combination of the variables in the system, the Eigen vector (see Enders 1995:412–418 for detail).

circle functional form for transformation). We hope you are now convinced that a test for non-zero Eigen values is tantamount to the test for the number of cointegration vectors. If yes, the next question is how are we going to conduct such test?

It follows from the previous discussion that the major issue in testing is to check the statistical significance of the Eigen values (that is, whether they are statistically different from zero or not). Here we will use our knowledge that the rank of a matrix equals the number of its Eigen values that differ from zero (see the discussion above). Suppose the *n* Eigen values of the matrix Π are given in the following order: $\lambda_1 > \lambda_2 > \ldots \lambda_n$. If the rank of Π is 0, then all these Eigen values (the λs) will equal zero. Applying our knowledge of logarithm that ln(1) = 0, the expressions (1 – λ) in matrix Π can be reduced to zero if we substitute the zero values of λ_is in the expression (1 – λ) and take its logarithm (that is, ln(1 – λ) = 0). If the rank of Π is unity and $0 < \lambda_1 < 1$, the first expression in $\ln(1 - \lambda_1)$ = negative (that is, log of a fraction is negative) all other $\lambda_i s = 0$. In applied work we can get the estimates of Π and the Eigen values. Then we can conduct a test for the number of Eigen values that are not insignificantly different from unity using what are called the λ_{trace} and λ_{max} tests given by

$$\lambda_{trace}(r) = -T \sum_{i=r+1}^{n} \ln\left(1 - \hat{\lambda}_i\right)$$

$$\lambda_{\max}(r, r+1) = -T \ln\left(1 - \hat{\lambda}_{r+1}\right)$$

where

$\hat{\lambda}$ are estimated Eigen values obtained from estimated Π matrix.

λ_{trace} tests the null hypothesis that the number of cointegrating vectors is ≤ = *r* against a general alternative[6]

λ_{max} tests the null that the number of cointegrating vector is *r* against the alternative *r* + 1

T is the sample size

[6] Notice that this statistics equals zero when $\lambda_i s = 0$. The further these Eigen values are from zero, the more negative the $\ln(1 - \lambda_i)$ will be; hence the higher the λ_{trace} statistics and hence the possibility of having cointegrating vectors.

The distribution of this test is non-standard and depends only on the degrees of freedom ($n - r$). It is tabulated by Osterward-Lenum (1992) and other recent economics textbooks. Most econometric software packages (such as Eviews or PC-Give/PC-Fimil) will give you the critical values developed by Johansen and Juselius (1990, 1992) using Monte Carlo studies or using the Cointegration Analysis for Time Series (CATS) which was developed for this. To test the hypothesis, compare the calculated statistics above with the tabulated ones.

Noting also the Granger representation theorem (Engel and Granger 1987) that states that a cointegrating system has an error-correction representation and vice versa, we can formulate a vector auto-regressive (VAR) (as given in the main text) into an error-correction form. We outline below how one can transform a VAR, say equation 3 in the main text, into an error-correction representation given in equation 4. This two-step procedure borrows from Johansen and Juselius (1990, 1992) and Juselius (1991).

Following Engel and Granger (1987), two or more time series are cointegrated of order *d, b;* they are denoted CI (d, b) if they are integrated of order d, but a linear combination of these variables exists that is integrated of order b, where $b < d$. There are two ways of testing the existence of cointegration between variables of interest and formulating an ECM.

The first approach (the Engel-Granger two-stage approach) begins by testing whether the variables of interest are stationary. If variables contemplated in the model follow an $I(1)$ process, then in the first stage, estimates of the long-run equilibrium equation (using OLS) are made. Then an ADF test on the residual of the long-run equation is conducted to determine if the variables in question are cointegrated—that is, whether the error term follows a stationary process. If the error term is stationary (taken as proof of cointegration) in the second stage, we could combine the error term with the first difference of the variables (short-run indicators) to estimate the final model.

This approach has a number of shortcomings. In particular, when we have more than two variables in an equation, the residual-based technique no longer has a unique vector. It should be understood that with *k* variables in the system, there must be at most *k* – 1 vectors (see Engel and Granger 1991, Banerjee et al. 1993, Enders 1996). Secondly, unless one has a clear theory, as for example a long-term money-demand equation that specifies the direction and magnitude of the coefficients, the interpretation of the outcome vector in this residual-based method is problematic. Finally, one has to make a strong assumption about a unique cointegration vector in the analysis.

These factors have led to the popularity of the Johansen approach, which handles multivariate systems in a better way (see Johansen 1988, 1991; Johansen and Juselius 1990). Following Johansen (1988, 1991) we may consider a VAR model given by equation A10.

$$X_t = \Pi_1 X_{t-1} + \ldots + \Pi_k X_{t-k} + \mu + \phi D_1 + \varepsilon_t \quad (t = 1,\ldots,T) \qquad \text{[A10]}$$

In general, an economic time series is a non-stationary process, and VAR systems like equation A10 can be expressed in the first-difference form. If we use $\Delta = 1 - L$, where L is the lag operator, we can rewrite equation A10 as

$$\Delta X_t = \Gamma_1 \Delta X_{t-1} + \ldots + \Gamma_{k-1} \Delta X_{t-k+1} + \Pi X_{t-k} + \mu + \Phi D_t + \varepsilon_t \quad \text{[A11]}$$

where

$$\Gamma_i = -(\mathrm{I} - \Pi_1 - \ldots - \Pi_i) \quad (i = 1,\ldots,k-1)$$
$$\text{and} \quad \Pi = -(\mathrm{I} - \Pi_1 - \ldots - \Pi_k)$$

Model equation A11 is a traditional first-difference VAR model except for the term ΠX_{t-1}. The Johansen procedure is based on an examination of matrix Π, which contains information about a long-run relationship. The analysis of the long-run relationship in the model is based on examining the rank of this matrix. If this matrix has a full rank, the vector process X_t is stationary. If the rank equals zero, the matrix is a null matrix and the equation A11 remains a traditional VAR. The variables are not cointegrated and so have no long-run solution in themselves. The third and most interesting possibility is when $0 < \text{rank}(\Pi) = r < p$, which implies there are $p \times r$ matrices α and β such that $\Pi = \alpha\beta'$. The cointegration vector β has the property that $\beta' X_t$ is stationary even though X_t itself is non-stationary. The Johansen procedure helps to determine and identify this (these) cointegrating vector (s).

The question is then how this can be done in a single equation. The algebra follows the same logic, but outside matrix algebra it becomes messy. We can thus demonstrate this using two variables, lagged one period each.

$$X_t = \beta_0 + \beta_1 X_{t-1} + \beta_2 Y_t + \beta_3 Y_{t-1} + \varepsilon_t \quad \text{[A12]}$$

Subtracting X_{t-1} from both sides, we get

$$\begin{aligned} X_t - X_{t-1} &= \beta_0 + \beta_1 X_{t-1} - X_{t-1} + \beta_2 Y_t + \beta_3 Y_{t-1} + \varepsilon_t \\ \Delta X_t &= \beta_0 + (\beta_1 - 1) X_{t-1} + \beta_2 Y_t + \beta_3 Y_{t-1} + \varepsilon_t \end{aligned} \quad \text{[A13]}$$

To have ΔX_{t-1} on the right-hand side, we need to subtract $(\beta_1 - 1)X_{t-2}$ on both sides:

$$\Delta X_t - (\beta_1 - 1) X_{t-2} = \beta_0 + (\beta_1 - 1)\Delta X_{t-1} + \beta_2 Y_t + \beta_3 Y_{t-1} + \varepsilon_t$$

$$\Delta X_t = \beta_0 + (\beta_1 - 1)\Delta X_{t-1} + (\beta_1 - 1) X_{t-2} + \beta_2 Y_t + \beta_3 Y_{t-1} + \varepsilon_t$$

[A14]

The same thing can be done for Y_t and then we get X_{t-k} combined with Y_{t-k} to form a vector as per the theory (such as the demand for money or the purchasing power parity in the text) to be tested. The most important thing is to assemble the variables in the Π matrix (as in equation A11) that provide the set of long-run information in a way that they are theoretically plausible, and the econometrics testing will follow to validate the relationship.

References

Alemayehu Geda. 1996. 'Dissertation guide for graduate study in financial economics.' University of London, School of Oriental and African Studies, Economics Department CIEE, London.

Banerjee, A., J. Dolado, J.W. Galbraith and F. Hendry. 1993. *Cointegration, Error-Correction and the Econometric Analysis of Non-stationary Data.* Oxford: Oxford University Press.

Chiang, Alpha C. 1984. *Fundamental Methods of Mathematical Economics.* London: McGraw-Hill Books.

Enders, Walter. 1996. *Applied Econometrics.* New York: John Wiley and Sons.

Engel, R.F., and C.W.J. Granger. 1987. 'Co-integration and error-correction: representation, estimation and testing.' *Econometrica* 55:251–276.

Engel, R.F., and C.W.J. Granger, eds. 1991. *Long-run Econometric Relationships: Readings in Co-integration.* Oxford: Oxford University Press.

Greene, William H. 1993. *Econometric Analysis.* New York: Macmillan.

Griffith, W., C. Hill, and G. Judge. 1993. *Learning and Practicing Econometrics.* New York: John Wiley and Sons.

Johansen, Soren. 1988. 'Statistical analysis of cointegration vectors.' *Journal of Economic Dynamics and Control* 12:231–254. (Also in Engel and Granger 1991, chap. 7.)

Johansen, Soren. 1991. 'Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models.' *Econometrica* 59(6):1551–1580.

Johansen, Soren, and K. Juselius. 1990. 'Maximum likelihood estimation and inference on cointegration: with applications to the demand for money.' *Oxford Bulletin of Economics and Statistics* 52(2): 169–210.

Johansen, Soren, and K. Juselius. 1992. *Identification of the Long-run and the Short-run Structure: An Application of the ISLM Model.* Discussion Paper. Copenhagen: Institute of Economics, University of Copenhagen.

Juselius, K. 1991. *Domestic and Foreign Effects on Prices in an Open Economy: The Case of Denmark.* Copenhagen: Institute of Economics, University of Copenhagen.

Kennedy, Peter. 1992. *A Guide to Econometrics*. Oxford: Blackwell.

Ndung'u, Njuguna, and Rose W. Ngugi. 1999. 'Adjustment and liberalization in Kenya: the financial and foreign exchange markets.' *Journal of International Development* 11:465–491.

Osterwald-Lenum, M. 1992. 'A note with quantiles of the asymptotic distribution of maximum likelihood cointegration rank test statistics.' *Oxford Bulletin of Economics and Statistics* 51:219–239.

Rao, B. Bhaskara, ed. 1994. *Cointegration for the Applied Economist*. London: Macmillan.

Ronge, Eric, and Peter Kimuyu. 1997. *Private Investment in Kenya: Trends, Composition and Determinants*. IPAR Discussion Paper No. 009/97. Nairobi: Institute of Policy Analysis and Research.

Sydsaeter, Knut, and Peter Hammond. 1994. *Mathematics for Economics*. Harare: Academic Books.

Thomas, R.L. 1997. *Modern Econometrics: An Introduction*. London: Addison-Wesley.

Tobin, J. 1969. 'A general equilibrium approach to monetary theory.' *Journal of Money, Credit and Banking,* 1(February):15–29.

Additional reading list for literature review on investment in Kenya

Basic theoretical materials

The first two listings are relevant for the accelerator model of investment.

Chenery, H. 1952. 'Over capacity and the acceleration principle.' *Econometrica* 20(1):1–28.

Keynes, J.M. 1936. 'The marginal efficiency of capital,' chapter 11, in *The General Theory of Employment, Interest and Money*. London: Macmillan. p. 135–146.

Jorgenson, D.W. 1963. 'Capital theory and investment behavior.' *American Economic Review,* supplement 53(2) May, p. 247–259. This is relevant for investment functions that are based on the user cost of capital or, broadly, the neoclassical theory. This is actually the theoretical base of the current KTMM version.

Surveys of empirical studies of investment functions

The following survey articles give a general picture of the empirical literature.

Chirinko, R.S. 1993. 'Business fixed investment spending: modeling strategies, empirical results, and policy implications.' *Journal of Economic Literature* 31(December):1875–1911.

Helliwell, J.F. 1976. 'Aggregate investment functions: a survey of issues.' In J.F Helliwell, ed., *Aggregate Investment,* Penguin Modern Economics Readings. Harmondsworth, England: Penguin Books. p. 13–49.

Jorgenson, D.W. 1971. 'Econometric studies of investment behaviour: a survey.' *Journal of Economic Literature* 9(4):1111–1147.

Precious, M. 1987. *Rational Expectations, Non-market Clearing, and Investment Theory.* Oxford: Clarendon Press.

Recent work related to developing countries

Elbadawi, Ibrahim, Benno J. Ndulu and Njuguna Ndung'u. 2000. 'Risk, uncertainties and debt overhang as determinants of private investment in sub-Saharan Africa.' Unpublished report.

FitzGerald, E.V.K., K. Jansen and R. Vos. 1994. 'External constraints on private investment decisions in developing countries.' In J.W. Gunning, Henk Kox, Woute Tims and Ynto de Wit, eds., *Trade, Aid and Development.* New York: St. Martin Press.

Greene, J., and D. Villanueva. 1991. 'Private investment in developing countries: an empirical analysis.' *IMF Staff Papers* 38:33–58.

Khan, Moshe S., and Manmohan S. Kumar. 1997. 'Public and private investment and growth process in developing countries.' *Oxford Bulletin of Economics and Statistics* 59(1).

Serven, L., and A. Solimano. 1992. 'Economic adjustment and investment performance in developing countries: the experience of the 1980s.' In V. Corbo, S. Fischer and S.B. Webb, eds., *Adjustment Lending Revisited: Policies to Restore Growth.* Washington, D.C.: World Bank.

Sundararajan, V., and S. Thakur. 1980. 'Public investment, crowding out, and growth: a dynamic model applied to India and Korea.' *IMF Staff Papers* 27:814–855.

Kenya

Oshikoya, Temitope W. 1992. 'Interest rate liberalization, savings, investment and growth in Kenya.' *Savings and Development* 26(3).

Soderbom, Mans. 1998. 'Investment behaviour.' In Arne Bigstern and Peter Kimuyu, eds., *Structure and Performance of Manufacturing in Kenya.* Nairobi: Departments of Economics, University of Nairobi and Gothenburg University.