EUROPEAN UNIVERSITY INSTITUTE
Department of Economics

Essays on Applications of I(1) and I(2) Cointegrated VAR Models on Issues in International Price Parities

Michael Pedersen

Thesis submitted for assessment with a view to obtaining the degree of Doctor of the European University Institute

Florence
June 2003

EUROPEAN UNIVERSITY INSTITUTE
Department of Economics

Essays on Applications of I(1) and I(2) Cointegrated VAR Models on Issues in International Price Parities

Michael Pedersen

The Thesis Committee consists of:

Søren Johansen, Supervisor, External EUI and Copenhagen
Anindya Banerjee, EUI
Marius Ooms, Free University Amsterdam, Dept. of Econometrics
Andrew Harvey, University of Cambridge, Faculty of Economics and Politics

Contents

Acknowledgments

After getting three years of leave of absence (which was later extended with two months) from my position in Danmarks Nationalbank, I arrived in Florence in the end of August 1999 ready to start my new life as a Ph.D. student. During that time I spent two weeks on summer schools at Pompeu Fabra in Barcelona and the European University Institute, two months on a Summer Internship in Bank of England, and four months on an exchange program at University of California at Berkeley. The thesis was completed after I returned to Danmarks Nationalbank.

The main influence on this thesis has naturally come from my supervisor Søren Johansen. I am very grateful to Søren for all the fruitful discussions we have had and for his patience in explaining me cointegration, especially the complicated math. Even the times when I needed to get the explanation two or three times before understanding it, Søren has always been very encouraging. I would also like to thank my second supervisor, Anindya Banerjee, for his guidance and encouragement during the process of writing the papers. Also thanks to the external members of the jury, Andrew Harvey and Marius Ooms, for very good comments which have improved the present thesis and will benefit my future research to a great extent.

During my stay at Berkeley, I had the opportunity to discuss my work with my appointed advisor Roger Crane, as well as Maurice Obstfeld, Andrew Rose, and Michael Jansson. I am grateful for all the comments I got. Thank you also Rebecca Driver and James Nixon from the Bank of England for good discussions during my stay there.

Furthermore, I thank Katarina Juselius, Andreas Beyer, Ronald MacDonald, and Lucca Fanelli for discussing and commenting on my work. Also thanks to Hans Christian Kongsted and Pieter Omtzig for sharing their software to test hypotheses in the $I(2)$ model.

My work has benefitted from comments from participants at various workshops at the European University Institute and at Berkeley. Furthermore, I have made presentations at seminars in Danmarks Nationalbank and Bank of England and acknowledge the comments made by the participants. Also comments from the participants at the conference 'Bridging Economics and Econometrics. Empirical Applications and Econometric Methods' held in Florence are appreciated.

Finally - but no least - I would like to thank my mother, Lissie, and my brothers Claus and Kenneth (and of course Jannie too), for their support during the writing of the thesis. And also thank you to all the new friends, including flat mates and travel companions, I have made during the three years: Fred, Katalina, Katherine, Magnus,

Oscar, Reina, and Sigrid, just to mention a few. You have helped making these three years one of the best investment I have ever made in myself.

CHAPTER 1

Introduction and overview

For more than two decades econometricians have applied vector autoregressive (VAR) models to analyze data in systems of variables. At the same time the theory has developed rapidly and the concept of cointegration was introduced by Granger (1983) in order to capture stationary relations between non-stationary time series. Much evidence suggest that many macro economic time series are not stationary but may be integrated of order one or two. The cointegrating vectors have been interpreted as long-run equilibrium relationships.

In international economics issues which have been of interest for researchers include tests of parities such as the Uncovered Interest Parity (UIP) and the Purchasing Power Parity (PPP). Cointegration methods have turned out to be useful in these studies of non-stationary variables. Also studies of market integration - interpreted as sharing of stochastic trends - have benefitted from the development of cointegration procedures.

Several methods have been developed to detect if two or more series cointegrate and to estimate the coefficients of cointegration. These include ordinary least squares (Engle and Granger, 1987), nonlinear least squares (Stock, 1987), principal components (Stock and Watson, 1988), canonical correlations (Bossaert, 1988), maximum likelihood in a fully specified error correction model (Johansen, 1988), instrumental variables (Hansen and Phillips, 1990), spectral regression (Phillips, 1988), three-step estimator (Engel and Yoo, 1989), and single-equation error correction model with leads and lags (Phillips and Loretan (1991), Saikkonen (1991), and Stock and Watson (1993)). The essays in the present thesis apply Johansen's method, which has been shown by Gonzalo (1994) to have the better properties of the first five methods mentioned.

More recently it has been discovered that for particular variables, for example prices, the difference is integrated of order one so that the levels are integrated of order two. This has led to the development of VAR models for $I(2)$ variables, which have now been applied in several empirical studies.

Chapter 2 aims at linking two directions studying integration of stock markets. One direction of literature investigates market integration - or the law of one price - under the assumption that stocks are priced according to the Capital Asset Pricing Model (CAPM). Another direction looks for common stochastic trends among stock prices. In chapter 2 it is shown that if stocks are priced according to the two factor model developed by Black (1972), then tests of restrictions in the cointegrated VAR can provide evidence for or against integrated markets. An empirical illustration is provided. This shows evidence that the stock markets of Denmark and the US are integrated. Chapter 2 is also printed as EUI working paper no. 2002/17 with the title 'Finding evidence of stock market integration applying a CAPM or testing for common stochastic trends. Is there a connection?'.

The empirical example in chapter 2 is a very simple one with only two variables, the real stock prices. Hence, an underlying assumption is that the PPP holds between the two countries. Evidence from many empirical studies, however, suggests that the PPP only holds rarely. This raises the very simple question "why?". There are several answers to this. One might be that many studies are based on the assumption that prices are $I(1)$. Since much recent research suggests that inflation rates are non-stationary, wrong methods have been applied. Another possible answer is that researchers may have been too optimistic in the sense that they test among countries too far away from each other. A third answer could be that prices are still adjusting to each other and have not yet reached a sustainable level.

In chapter 3 we investigate if these are indeed some of the reasons for the PPP not to hold. The PPP is tested within four areas in the US allowing for prices to be integrated of order two. In the $I(2)$ model, the stationary relations are multicointegrating, which means that they include levels as well as differences. These can be explained with a Linear Quadratic Adjustment Cost (LAQC) model, where policy makers aim at market integration and minimization of the inflation rate. The optimal solution implies a relation between price levels and inflation rates, which is called 'PPP with adjustment' in this thesis. The name refers to the fact that prices are still adjusting to each other and have not yet reached a sustainable PPP level. In the empirical example, evidence is provided that among the four areas considered, the PPP with adjustment holds between the three areas which are geographically closer. The main part of chapter 3 is also printed as EUI working paper no. 2002/18 with the title 'Does the Purchasing Power Parity hold within the US?'.

Testing the PPP in an $I(2)$ model is quite simple when the exchange rate is the same in the areas under consideration. When including an exchange rate in the system, the analysis becomes more complicated. Chapter 4 generalizes the analysis in chapter 3 by including an exchange rate.

An empirical discussion of whether or not inflation rates are stationary is provided in chapter 4. Using US price data for a period of almost 90 years in a univariate framework it turns out that stationarity depends very much on the sample investigated. This also suggests that risk of structural breaks in a long sample is very apparent, which leads to the problem of constant parameters in estimations. Another issue raised in this chapter is what outcome we can expect when analyzing systems which include prices and exchange rates. A comprehensive scenario analysis is provided to deal with this. An empirical example indicates that the PPP with adjustment holds between the US and the UK in the post Bretton Woods period. The main part of chapter 4 is based on a paper with the title 'Testing the PPP in a cointegrated VAR when inflation rates are non-stationary. With an application to the UK and the US'.

The analyses in chapter 3 and 4 are based on transformations of the systems from the $I(2)$ to the $I(1)$ space. Papers have treated this question in the stochastic sense but what is the role of the deterministic terms? The purpose of chapter 5 is to analyze this question.

Chapter 5 contains an analysis which is more technical than those of the previous three chapters. Explicit expressions are given for the deterministics in the $I(1)$ and the $I(2)$ models in stationary as well as non-stationary directions. Then the transformation of an $I(2)$ system into the $I(1)$ space is discussed and explicit relations between the stationary and non-stationary directions are found. With these in hand, the relationships between the deterministic terms in the two spaces are found. It turns out that the deterministics in general are not the same in the transformed model and the original one. It is suggested that the tests for deterministics should be done in the original $I(2)$ model. A strategy for testing is proposed. This chapter is based on a paper with the title 'On the transformation from $I(2)$ to $I(1)$ in the cointegrated VAR with focus on the role of the deterministic terms'.

In chapter 6 the conclusions of analyses are summarized and possible directions for future research are given.

CHAPTER 2

Stock market integration and the cointegrated VAR

Abstract: In this paper it is demonstrated that if assets are priced according to Black's (1972) CAPM, then tests on the cointegrated VAR can reveal evidence for or against integration of financial markets. If the market portfolios cointegrate one-to-one and share the same deterministic long-run trend, then the markets obey the law of one price. Furthermore, it is demonstrated how the driving force of the prices can be found. Evidence from an empirical example suggests that the Danish and American stock markets are integrated because US stock prices drive those of Denmark.

1. Introduction

The following analysis aims at combining two directions in the existing literature of testing for stock market integration. It will be demonstrated how tests for integration assuming a capital asset pricing model (CAPM) and tests for common stochastic trends among stock price indices can be linked by restrictions on the cointegrated vector autoregressive (VAR) model. The source of the common trends is rarely discussed in the existing literature. The present chapter takes up that discussion and shows how tests in the cointegrated VAR can provide information about it.

Relations between national financial markets have been of interest to researchers for more than a decade. Much effort has been made to find empirical evidence for or against links among stock markets in different countries, especially since the crash in stock markets in October 1987. Several studies of theoretical macroeconomic models have established many potential economic benefits from the integration of financial markets. The general consensus is that, because of better opportunities for risk sharing, the integration of financial markets can stimulate growth (see, for example, Pagano (1993) and Obstfeld (1989, 1998)). Indeed, much effort has been made at the political level in order to put legal conditions in place which, in turn, facilitate the integration of national financial markets. The European Union is an obvious example here (See Licht, 1997). This has led to the important question of whether or not financial markets have in fact become more integrated.

Various authors have suggested ways of testing for market integration. One direction of the literature argues (assumes) that markets are integrated if similar assets - i.e. assets with the same risk-adjusted payoff profile - are priced identically. A CAPM can be used to determine whether such assets have the same (theoretical) price. Another direction of research tests integration in terms of common stochastic trends (or cointegration) among international markets, which are measured by indices representing the whole market. The more markets are cointegrated - i.e. the fewer common stochastic trends the markets share - the stronger the evidence of integration.

In the present paper it is assumed that assets are priced according to a version of the CAPM developed by Black (1972). Furthermore, it is assumed that the data generating process (DGP) of the market portfolios can be described by a VAR model. With these assumptions in hand it is shown how evidence can be found for or against market integration simply by testing restrictions on the VAR model.

The rest of the chapter is organized as follows: Section 2 discusses the concept of the integration of financial markets on the basis of proposed definitions in the financial literature. In section 3, the statistical model is set up and discussed followed by a discussion of the theoretical asset pricing model in section 4. The restriction on the statistical model, based on the CAPM, is treated in section 5, and section 6 provides a discussion of the source of the stochastic trends in stock prices. It is also demonstrated how a simple test on the cointegrated relations can reveal information of where the common trends come from. In section 7, an empirical example is given, and the analysis is summarized in section 8.

2. Integration of financial markets

In the financial literature a formal definition of market integration does not seem to exist. Nevertheless, many proposals have been put forward: see, for example, Jorion and Schwartz (1986); Wheatley (1988); Gultekin et al. (1989); Bekaert and Harvey (1995); Chen and Knez (1995); and Hardouvelis et al. (1999). In general there seems to be some consensus that two financial markets are considered integrated if assets with the same risk-adjusted return cash-flows are priced similarly. Some authors also refer to this as the law of one price. It follows from this that integration is related to convergence of risk aversion across markets in the sense that the difference between investors' degree of aversion on different markets narrows.

The question of integration is not only relevant in an international context. When considering national markets, tests of integration among, for example, IT-stocks and industrial stocks are also relevant.[1] In the empirical analysis which follows, the focus is on integration between stock markets in different countries, but a similar analysis could easily be conducted using data for different industries. In fact, it might be the case that if national markets are perfectly integrated, then investors might prefer to diversify their portfolio between industries rather than countries.[2]

In practice, many stocks are not traded at more than one (or occasionally a few) stock exchange(s), which complicates the testing integration. So, how can we even talk about stock markets being integrated and, furthermore, is it possible to test this at all? This leads to the question of how stock prices are determined. If, for example, stock prices in general are determined only by domestic fundamental factors, then an examination of convergence in the development of appropriate fundamentals could serve as a test for integration. In practice, however, there seems to be more factors involved than fundamental variables in terms of influencing stock prices. The October 1987 crash in the American stock market, which spread to many other countries despite the fact that the development in the fundamentals were very different, is an example of this. Furthermore, the sharp increase in US stock prices, which started in 1995, seems to have spread to other countries as well, as can be seen in Figure 2.1, which shows the development in real stock prices in the US, the UK and Germany.

[1] Of course, comparing sectors in a national market is somewhat easier since fluctuations in exchange rates do not exist.

[2] This point was also noted by Hardouvelis et al. (1999).

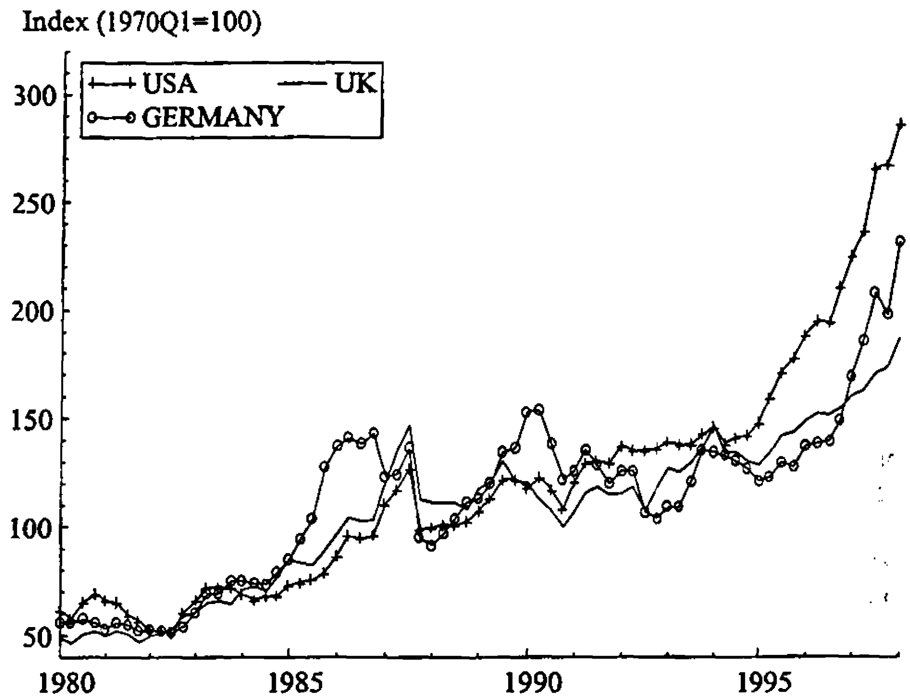


Figure 2.1. Real stock price indices. Quarterly data.

Testing integration empirically raises the difficult question of how to measure risk-adjusted return cash-flows. Several studies test integration by applying theoretical pricing models such as CAPM and APT (Arbitrage Pricing Theory). Both models price assets in general equilibrium. In CAPM the rates of return on all risky assets are functions of their covariances with the market portfolio (the portfolio consisting of all assets in the market). In APT models the returns of risky assets are linear combinations of various factors that affect asset returns. Hence, APT is more general than CAPM and it can be shown that CAPM is a special case of APT.[3] In the analysis below, a version of the CAPM is used to illustrate the meaning of risk-adjusted prices, and to describe the link between cointegration among prices and market integration.

[3]The CAPM and the APT model are described in standard text books on financial theory such as Copeland and Weston (1988).

3. The statistical model

We consider the unrestricted VAR(k) model (k is the number of lags), which is written in error correction form:[4]

$$\Delta X_t = \Pi X_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} + \mu_0 + \mu_1 t + \varepsilon_t, \varepsilon_t \sim iid(0, \Omega),$$

where X_t represents the data vectors of dimension p, for example consisting of stock price indices from different countries. The matrices Π, Γ_i, μ_0 and μ_1 include coefficients to be estimated. To simplify notation possible dummies are disregarded. The first thing to investigate is whether X_t is $I(1)$ and if any of the time series share the same stochastic trend(s). These hypotheses can be formulated as a reduced rank condition on Π:

$$\begin{aligned} H_1 :&\quad \Pi = \alpha\beta' \text{ has reduced rank } r < p, \\ H_2 :&\quad \alpha'_\perp \Gamma \beta_\perp \text{ has full rank } p - r, \end{aligned}$$

where $\Gamma = I - \sum_{i=1}^{k-1} \Gamma_i$. The $_\perp$ notation indicates an orthogonal complement such that $\alpha'_\perp \alpha = 0$ and $\beta'_\perp \beta = 0$. The hypothesis H_2 is about the $I(1)$ space having full rank such that there are no stochastic $I(2)$ trends present. This hypothesis ensures that we have $p-r$ independent common stochastic trends besides r cointegration relations. If H_1 and H_2 are accepted the prices share at least one common trend.

A test for the number of common trends is developed by Johansen (1988, 1991). He shows how to apply Anderson's (1951) technique of reduced rank regression to form a likelihood ratio test, where the maximized likelihood functions are found by solving an eigenvalue problem. More precisely, by making the reduced rank regression we get p eigenvalues: $1 > \widehat{\lambda}_1 > ... > \widehat{\lambda}_p > 0$. The likelihood ratio (Trace) test for r cointegrating vectors (and hence $p - r$ common trends) is given by:

$$-2 \ln Q(r \mid p) = -T \sum_{i=r+1}^{p} \ln(1 - \widehat{\lambda}_i), \tag{3.1}$$

where T is the number of observations. Johansen and Juselius (1990) derive the asymptotic distribution of the test and present critical values.

In the case of cointegration the moving average representation is, according to Granger's representation Theorem, given by:

$$X_t = C \sum_{i=1}^{t} (\varepsilon_i + \mu_0 + \mu_1 i) + C_1(L)(\varepsilon_t + \mu_0 + \mu_1 t) + A, \tag{3.2}$$

[4]The approach used here was developed by Johansen (1988, 1991, 1996).

where $C = \beta_\perp(\alpha'_\perp \Gamma \beta_\perp)^{-1}\alpha'_\perp$, A depends on the initial conditions and $C_1(L)$ satisfies the conditions given in Johansen's (1996) Theorem 4.2. Since $\alpha'_\perp \sum_{i=1}^{t} \varepsilon_i$ is the only non-stationary part of the process, this is defined as the common stochastic trends.[5] The matrix $\alpha'_\perp$ are the coefficients for the common trends and $\beta_\perp$ are the loadings from the common trends into the variables. The latter indicates to what extent the variables are affected by the trends.

The cumulation of a deterministic trend is a quadratic trend. Since this is rarely (probably never) seen in economic time series this cumulation should be avoided. This can be achieved by restricting the trend to the cointegrating space. Formally this is done by decomposing μ_i $(i = 1, 2)$ such that $\mu_i = \alpha\rho_i + \alpha_\perp\gamma_i$. The restriction $\gamma_1 = 0$ is then imposed. The error correction model can then be rewritten:

$$\Delta X_t = \alpha \begin{pmatrix} \beta \\ \rho'_1 \end{pmatrix}' \begin{pmatrix} X_{t-1} \\ t \end{pmatrix} + ... \tag{3.3}$$

It can be shown that the linear trend in the process is given by[6]

$$\tau_1 = C\alpha_\perp\gamma_0 + (C\Gamma - I_p)\beta(\beta'\beta)^{-1}\rho'_1. \tag{3.4}$$

Considering a process on the form $X_t = \tau_0 + \tau_1 t + \textit{stoch. terms}$ implies that $E(\Delta X_t) = \tau_1$, which will be used later. Note that in the long-run relations, $\beta' X_t$, the coefficient for the deterministic trend is given by $\beta'\tau_1 = -\rho'_1$. Hence, a test on the long-run trend is simply a test about ρ_1.

From (3.2) it follows that a test of the same impact from the common trends should be performed on the C matrix. Tests should reveal whether the rows of C are identical. Since $(\alpha'_\perp\Gamma\beta_\perp)^{-1}$ is only a normalization, the test for identical rows in C can be performed as a test for identical loadings from the stochastic trends, i.e. a test for that $\beta_\perp$ is proportional to $(1, 1, ..., 1)'$. If the prices share only one common trend, $\beta_\perp$ will be a vector of p components. Alternatively we can perform the test on the β matrix. The test will be for one-to-one cointegration between the variables. For example, in the case with $p = 2$ the test can be formulated as the hypothesis:

$$\beta_\perp = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \implies \beta = \begin{bmatrix} 1 \\ -1 \end{bmatrix}. \tag{3.5}$$

It is well known that in general β is not unique. In the case $p = 3$, the test of similar loadings from the common trends can be formulated

[5]See Johansen (1996) Definition 3.7.
[6]See Johansen (1996) exercise 6.1.

differently on β as illustrated below:

$$(3.6)\quad \beta_{\perp}=\begin{bmatrix}1\\1\\1\end{bmatrix}\quad\Longrightarrow\quad\beta=\begin{bmatrix}1&0\\-1&1\\0&-1\end{bmatrix}\text{ or }\beta=\begin{bmatrix}1&1\\-1&0\\0&-1\end{bmatrix}.$$

Testing hypotheses formulated as (3.5) and (3.6) is straightforward using the procedure described in Johansen and Juselius (1994). Furthermore, testing hypotheses on β is standard procedure in software packages such as PCFIML (Doornik and Hendry, 1997) and CATS in RATS (Hansen and Juselius, 1995).

4. Black's CAPM

This chapter operates on the assumption that integration of markets is a long-run feature. It seems reasonable to believe that two markets remain integrated regardless of whether or not the integration is punctuated by short periods of divergence. Furthermore, when applying an equilibrium model to price assets, short-term divergence from the equilibrium prices is likely to occur in the data. Prices in the long run - or in equilibrium - should, however, be the same for similar assets traded on different markets if these are perfectly integrated.

The original version of the CAPM was - among others - developed by Sharpe (1963, 1964). In this model it is assumed that investors have the opportunity to invest in a risk-free asset, which gives a risk-free return. Black (1972) shows that the results of the standard model also apply if no such risk-free asset exists. Another asset (or portfolio) which is unrelated (zero-correlated) with the market in general can take the place of the risk-free asset.

For practical purposes there might be reasons not to consider investments in risk-free assets. One reason is that no unique definition of a risk-free asset exists. Empirical testing of the CAPM often uses a money market rate as a proxy for the risk-free rate of return. Whether this is really a risk-free return or not can be subject to discussion.

This issue will not be touched upon further in this analysis since it is doubtful whether investors really take into account the risk-free rate of return when trying to determine the price of a stock. Assuming that stocks are priced according to Black's CAPM has the advantage that one does not have to make any decision about the role of the risk-free asset. Hence, when looking for evidence of whether two different markets would price identical assets similarly, we need only consider the stochastic properties of the stock prices themselves and not take into account the developments in, say, money market rates.

The assumptions of the model are the following: (i) investors are risk-averse agents, who maximize end-of-period expected utility of wealth; (ii) they are price-takers with homogenous expectations of jointly normal distributed asset returns; (iii) assets are available in a fixed quantity and are all tradable and divisible; (iv) markets are frictionless and all information is free and available to all investors; (v) there are no market imperfections nor restrictions on short selling. As mentioned before, in the standard CAPM it is also assumed that a risk-free asset exists. In Black's version this is not the case. Instead there exists a portfolio with a return which is independent (zero-correlated) of that of the market portfolio. Examining a model with no risk-free asset allows us to investigate the integration of markets, in which all assets bear a risk. Black's version is more general than the standard one and has Sharpe's model as a special case.

The underlying assumption of the model is that an investor only demands additional return as compensation for any risk that is correlated with the market as a whole. This is referred to as systematic risk. For other risks, called unsystematic, the investor requires no compensation. The market as a whole is referred to by the market portfolio m. This is defined as the portfolio, which consists of all assets of the market held in proportion to their value weights. Hence, the proportion of asset i in the market portfolio is:

$$w_i = \frac{V_i}{\sum_{j=1}^{n} V_j},$$

where V_i is the market value of asset i and n is the total number of assets. The idea is to measure the price of an arbitrary risky asset i as a price adjusted for the systematic risk.

Given the assumptions mentioned above, all investors will hold an efficient portfolio, i.e. a portfolio which minimizes the risk given the required rate of return or equivalently maximizes the rate of return given the risk the investor is willing to take. Hence, the set of all efficient portfolios is the solution to the problem

$$\begin{array}{lll} & & \min_{W} W'\Sigma W (= \sigma^2) \\ \text{s.t.} & (i) & r^e = W'R^e \\ & (ii) & W'\mathbf{1} = 1 \\ & (iii) & r^e \geq r^e_{mvp}, \end{array} \tag{4.1}$$

where W is a $n \times 1$ vector consisting of the portfolio weights for each of the n assets, Σ is the covariance matrix of returns, r^e is the required expected return (a number), R^e is a $n \times 1$ vector containing the expected return of all assets, $\mathbf{1}$ is a $n \times 1$ vector of ones and r^e_{mvp} is the return of the minimum-variance-portfolio, i.e. the portfolio which solves the problem (4.1) disregarding the constraints (i) and (iii). In appendix A it is shown that the set of solutions to (4.1) disregarding (iii) graphically represents a hyperbola in the $\sigma - r^e$ space. This set is called the frontier. The constraint (iii) ensures that investors always choose a portfolio on the upper part - i.e. the efficient part - of the frontier.

Since all investors hold an efficient portfolio, the market portfolio is also efficient as this is just the sum of all portfolios. In fact, it can be shown that the set of efficient portfolios is convex and, hence that any linear combination of efficient portfolios is also efficient.[7]

As the existence of a portfolio uncorrelated with the market portfolio is essential for Black's CAPM, it will be shown formally that this does exist. The zero-correlation portfolio, which will be called z, fulfills $corr(r^e_m, r^e_z) = 0$, where r^e_m is the expected return of the market portfolio and r^e_z is the expected return of the z-portfolio. As demonstrated in appendix A, the necessary and sufficient conditions for the solution of (4.1) implies

$$W = g + hr^e \tag{4.2}$$

and (iii), where g and h are $n \times 1$ vectors defined in Appendix A. Using (4.2) the covariance between two arbitrary frontier portfolios, p and q, is given by[8]

$$cov(r^e_p, r^e_q) = W'_p \Sigma W_q = \frac{C}{D}\left((r^e_p - r^e_{mvp})(r^e_q - r^e_{mvp}) + \frac{D}{C^2}\right), \tag{4.3}$$

where C and D are real numbers defined in Appendix A. Setting (4.3) equal to zero defines the unique zero-correlation portfolio corresponding to an arbitrary frontier portfolio:

$$r^e_z = r^e_{mvp} - \frac{D/C^2}{r^e_p - r^e_{mvp}}. \tag{4.4}$$

From (4.4) it follows that all frontier portfolios, except the minimum-variance portfolio, have a corresponding zero-correlation portfolio.[9] Particularly, the market portfolio has a corresponding z portfolio, assuming that $W_m \neq W_{mvp}$.

---

[7]See Huang and Litzenberger (1988) Chapter 3.

[8]See Appendix B.

[9]The zero-correlation portfolio for the market portfolio can also be found by solving problem (4.1), disregarding constraint (iii) and changing (i) to $W'_p \Sigma W_m = 0$.

The idea of the CAPM is to form a portfolio consisting of the market portfolio m in proportion $1-a$ and the rest in a risky asset i. This portfolio, p, has the following mean and variance of the return:

$$r_p^e = ar_i^e + (1-a)r_m^e, \tag{4.5}$$

$$\sigma_p^2 = a^2\sigma_i^2 + (1-a)^2\sigma_m^2 + 2a(1-a)\sigma_{im}, \tag{4.6}$$

where σ_{im} is the covariance of the returns of asset i and the market portfolio. Note that p can be thought of as an artificial portfolio. This follows from the fact that in equilibrium all portfolios are efficient and hence, as argued above, m is also efficient. As m consists of all assets, including i, a must be the excess demand of asset i. Now the idea is to investigate what happens when approaching equilibrium, i.e. when a approaches zero.

The ratio between the partial derivatives of the mean in (4.5) and the standard deviation in (4.6) for $a = 0$ gives us the slope of the frontier evaluated in market equilibrium:

$$\frac{\partial r_p^e/\partial a}{\partial \sigma_p/\partial a} = \frac{r_i^e - r_m^e}{(\sigma_{im} - \sigma_m^2)/\sigma_m} \text{ for } a = 0. \tag{4.7}$$

We apply the same trick and form an (artificial) portfolio consisting of the portfolios m and z. The expected return and variance of this are:

$$r_q^e = ar_z^e + (1-a)r_m^e,$$

$$\sigma_q^2 = a^2\sigma_z^2 + (1-a)^2\sigma_m^2.$$

Evaluating the slope of the risk-return trade-off in market equilibrium yields:

$$\frac{\partial r_q^e/\partial a}{\partial \sigma_q/\partial a} = \frac{r_m^e - r_z^e}{\sigma_m} \text{ for } a = 0. \tag{4.8}$$

In equilibrium (4.7) must hold between the market portfolio m and the risky asset i. At the same time (4.8) must hold between m and the zero-correlation portfolio z. Equalizing (4.7) and (4.8) and reorganizing gives us the expected return of i, expressed as a linear combination between the expected return of z and the expected return of m:

$$r_i^e = \left(1 - \frac{\sigma_{im}}{\sigma_m^2}\right) r_z^e + \frac{\sigma_{im}}{\sigma_m^2} r_m^e. \tag{4.9}$$

Black's model is also called a two-factor model, since we can determine the expected return of any risky asset by two factors. The ratio $\beta_i = \sigma_{im}/\sigma_m^2$ is the quantity of risk in the model. Formula (4.9) is the expected risk-adjusted return of asset i. To find the risk-adjusted price, we apply the definition of the return:

$$p_{i0} = \frac{E(p_{ie})}{1 + (1 - \beta_i)r_z^e + \beta_i r_m^e}, \tag{4.10}$$

where p_{i0} is the start-of-period price and $E(p_{ie})$ is the expected end-of-period price.

Evaluating (4.10) for two different markets can tell us something about the degree of integration between the markets. The idea is that we want to price two identical assets in two different markets. By identical assets are meant that the expected end-of-period prices are the same in both markets and the covariances with the market portfolios are equal. If the start-of-period prices are equal then the markets are perfectly integrated.

5. Integration and the restrictions on the VAR model

For simplicity, integration between only two markets is considered here. This can, however, easily be extended to consider three or more markets. From (4.10) it appears that equally risky assets will have the same price in market 1 and market 2 if the denominators are the same. Since the assets are assumed to be similar ($\sigma_{1,im} = \sigma_{2,im}$ and $E(p_{1,ie}) = E(p_{2,ie})$), this will occur if (i) the expected returns of the market portfolios are the same, $r_{1,m}^e = r_{2,m}^e$; (ii) the variances of the returns of the market portfolios are equal, $\sigma_{1,m}^2 = \sigma_{2,m}^2$; and (iii) the expected returns of the zero-correlation portfolios are equal, $r_{1,z}^e = r_{2,z}^e$.

Consider the prices of two market portfolios and assume that the development in these can be described by a VAR model. Furthermore, it is assumed that the prices are integrated of order one, $I(1)$, and cointegrate, i.e. they share the same stochastic trend. Hence, the system has $p = 2$ and the restriction $r = rank(\Pi) = 1$ is imposed. This can be tested applying the Trace test (3.1). The system of prices can then be written:

$$\begin{bmatrix} p_{1,t} \\ p_{2,t} \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \sum_{i=1}^{t} u_i + \begin{bmatrix} d_1 + k_1 \\ d_2 + k_2 \end{bmatrix} t + stat,$$

where $p_{i,t}$, $i = 1, 2$ is the logarithm of the series of prices for the market portfolios and *stat* represents the stationary components of the process, which will be unimportant in the following. In general, $u_i = f(\varepsilon_{1i}, \varepsilon_{2i})$, i.e. some function of the residuals of the process such that $E(u_t) = 0$ and $var(u_t) = \sigma_u^2$. The coefficients d_1 and d_2 refer to that part of the deterministic trend which is not automatically eliminated in the cointegrating relation (the long-run trend), whereas k_1 and k_2 refer to the part which will disappear. The long-run trend is the important one here as we are considering pricing in a steady state. Hence, when relating this model to the cointegrated VAR, $(d_1, d_2)' = -\beta(\beta'\beta)^{-1}\rho_1$ such that $\rho_1 = -\beta'(d_1, d_2)'$ and $(k_1, k_2)' = \tau_1 + (d_1, d_2)'$. The coefficients c_i and $(d_i + k_i)$, $i = 1, 2$ are to be estimated. The returns of the two market portfolios are given by:

$$\begin{aligned} r_{1,t} &= \Delta p_{1,t} = c_1 u_t + d_1 + k_1 + ... \sim I(0), \\ r_{2,t} &= \Delta p_{2,t} = c_2 u_t + d_2 + k_2 + ... \sim I(0), \end{aligned}$$

which are stationary. Since the expected value of a stationary process is constant over time, the expected returns are:

$$\begin{aligned} r^e_{1,t} &= E(r_{1,t}) = d_1 + k_1, \\ r^e_{2,t} &= E(r_{2,t}) = d_2 + k_2. \end{aligned}$$

In the long run k_1 and k_2 are eliminated, i.e. $\beta'(k_1, k_2) = 0$, such that the expected returns of the two market portfolios are equal if the coefficients for the long-run trends are the same: $d_1 = d_2$. This implies that there will be no trend in the cointegrating relation if the cointegrating vector is $\beta = (1, -1)'$ (see below). Hence, in this case imposing the restriction of equal expected returns implies $\rho_1 = 0$ in (3.3).

For calculation of the variance of the returns it should be noted that, in the long run, the non-stationary part of the process dominates the stationary part with respect to the stochastic variation. The long-run variances of the returns are given by:

$$\begin{aligned} var_{LR}(r^e_{1,t}) &= c_1^2 var(u_t) = c_1^2 \sigma_u^2, \\ var_{LR}(r^e_{2,t}) &= c_2^2 var(u_t) = c_2^2 \sigma_u^2. \end{aligned}$$

Thus, the variation of the returns of the market portfolios are equal if $c_1 = \pm c_2$ and in particular if $c_1 = c_2$, which is the most natural case and the one investigated here. In other words, the impact from the common stochastic trend are the same on both variables, that is the prices cointegrate one-to-one. The restriction imposed on the cointegrating VAR is $sp(\beta) = sp(1, -1)$.

The return of the zero-correlation portfolios will by definition fulfill that $corr_{LR}(r^e_{i,t}, r^e_{iz,t}) = 0, i = 1, 2$. If $c_1 = c_2$ the expected return of the zero-correlation portfolios will be equal. To see this, notice that for $c_1 = c_2$, $r^e_{1,t}$ and $r^e_{2,t}$ are perfectly correlated in the long run: $corr_{LR}(r^e_{1,t}, r^e_{2,t}) = 1$. Hence $corr_{LR}(r^e_{1,t}, r^e_{2z,t}) = corr_{LR}(r^e_{2,t}, r^e_{1z,t}) = 0$. And as the zero-correlation portfolio is unique it must hold that $E(r^e_{1z,t}) = E(r^e_{2z,t})$ for $c_1 = c_2$.

The above discussion can be summed up as follows: If the time series of prices for two different stock markets have the same deterministic long-run trend and they cointegrate one-to-one, then two similar assets will be priced equally on both markets. In other words, the markets are integrated in the sense that two assets with the same risk profile will be priced identically. The hypothesis of integrated market can be formulated as $sp(\beta, \rho_1) = sp(1, -1, 0)$.

6. Where does the common trend come from?

This section focuses on the source of the common stochastic trend in the situation where markets are integrated. Here, finding the driving force behind stock prices is regarded as an empirical issue. By the driving force is meant the following: a variable, x_2 say, is driven by another variable, x_1 say, if the non-stationary part of x_2 is the cumulation of the errors of x_1 only.

In this situation, where we are interested in market integration, we might want to look for factors driving the larger market - which could be another (bigger) market - to be the driving force. For example, if one finds evidence of integration between two relatively small stock markets such as those of Denmark and Sweden, then the common stochastic trend might come from a larger market like Germany, for example. If this turns out to be the case, the next step could be to investigate what drives the German market. For example, one could include German GDP as a variable. If this is the (only) driving force of the prices, we should still find evidence of only one common trend.

In what follows we investigate whether one of the two markets considered drives the system. Formulated another way, we want to discover whether the common stochastic trend is coming from one of the variables already included in the system. This can be formulated as a test on the coefficients for the common stochastic trends, i.e. on $\alpha_\perp$. The moving average representation (3.2) for the case $(p, r) = (2, 1)$ can be written

$$\begin{bmatrix} x_{1,t} \\ x_{2,t} \end{bmatrix} = \begin{bmatrix} \widetilde{\beta}_\perp^1 \\ \widetilde{\beta}_\perp^2 \end{bmatrix} \begin{bmatrix} \alpha_\perp^1 & \alpha_\perp^2 \end{bmatrix} \begin{bmatrix} \sum_{i=1}^t \varepsilon_{1,i} \\ \sum_{i=1}^t \varepsilon_{2,i} \end{bmatrix} + \ldots,$$

where $\widetilde{\beta}_\perp = \beta'_\perp(\alpha'_\perp \Gamma \beta_\perp)^{-1}$. If the driving force of the system is x_1, say, the restriction on the VAR is $sp(\alpha_\perp) = sp(1, 0)$ or formulated on α, $sp(\alpha) = sp(0, 1)$. As with tests on the β vectors, tests on the α vectors are likelihood ratio and is a standard test in the software packages mentioned above. For a description of the procedure see Johansen (1996) Chapter 8.

7. An empirical example

In this example we investigate whether the Danish and the US stock markets are integrated, examining data from the post Bretton Wood period. A priori one might expect a big market like the American one to have considerable influence on a small one like that of Denmark. It is unlikely, however, that there is any reciprocal effect.

The source for the data is IMF's *International Financial Statistic (IFS)* and for a further description the reader is referred to the *IFS* manual. The stock price indices (ser. 62) do not cover the entire markets but are used as proxies for the market portfolios. To obtain real prices the stock prices are deflated with the consumer prices indices (ser. 64). The Danish data are converted to US dollars using the average exchange rate (ser. AF). The series are in logarithms and cover quarterly data from 1976Q1 to 1998Q4. The data are illustrated in levels and differences in Figure 2.2. A first look at these could suggest that the series are $I(1)$.

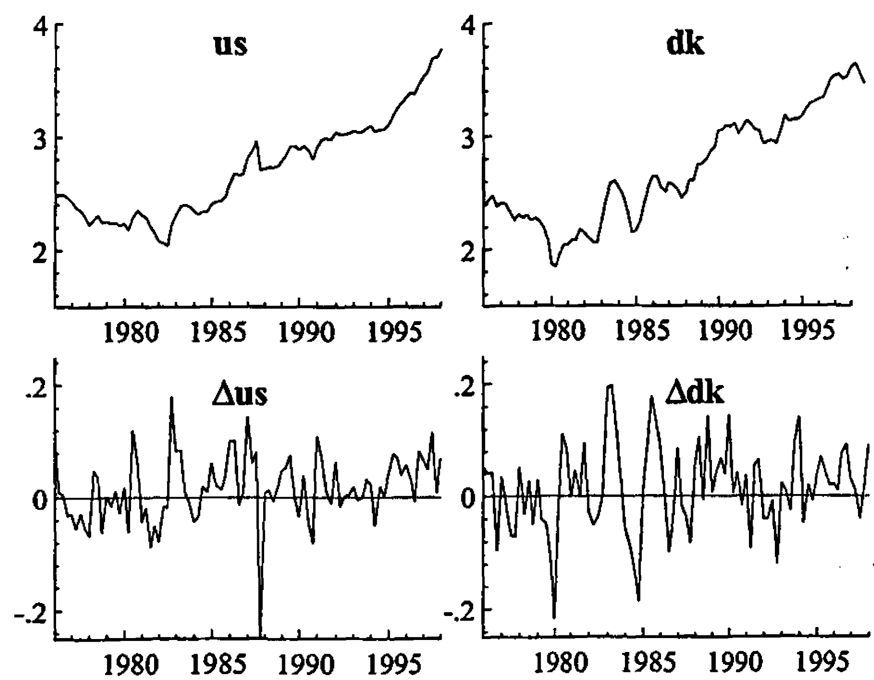


Figure 2.2. Data in levels and differences.

A VAR with 2 lags[10] is applied and misspecification tests indicate no autocorrelation and no ARCH in the residuals. There is only small evidence of normality, which is mainly due to the US data. Since normality is not crucial for the result obtained below, the model is considered well specified.

The critical values for the test of $rank(\Pi)$ depends on the specification of the deterministic terms. The first model considered includes a trend restricted to the cointegration space given in (3.3) This model will be named $H^*(r)$.[11] With a 10% significance level we can accept the hypothesis of one cointegrating vector. Accepting $rank(\Pi) = r = 1$ means restricting a root of 0.91 in the companion form matrix to one. It is tested whether the trend should be included in the cointegrated space, i.e. whether the long-run expected returns of the market portfolios are equal.[12] The hypothesis is formulated as:

$$\begin{array}{ccccc} & us & dk & trend & \\ \mathcal{H}_1: (& * & * & 0 &) \in sp(\beta). \end{array}$$

The test is accepted with test statistic $\chi^2(1) = 3.96$ ($p-value =$ 0.05). Hence the model is altered to include an unrestricted constant and is named $H_1(r)$. The outcome of the Trace test is a bit lower than the 90% critical quantile but with the two largest eigenvalues in the companion form matrix being 1.02 and 0.65 we accept the hypothesis of one cointegration vector. The hypothesis of restricting the constant to the β-space is accepted with $p-value = 0.22$ using the test statistic given in Theorem 6.3 in Johansen (1996) and the model is altered accordingly to $H_1^*(r)$. Also in this model evidence is in favor of $r = 1$ and the hypothesis of zero-coefficient for the constant is accepted with test statistic $\chi^2(1) = 0.33$ ($p - value = 0.56$). The acceptance of this hypothesis is due to the fact that the two indices have almost the same initial values. Hence, from the analysis of the deterministic terms it is concluded that the appropriate model in which to test hypotheses on β and a is without any deterministic terms. This model is named $H_2(r)$.

The Trace tests in $H_2(r)$ are given in Table 2.1 and are in favor of $r = 1$ but the hypothesis of two stationary relations is a borderline case. Imposing the restriction $r = 1$ seems reasonable as the second largest root in the companion form matrix is 0.65.

---

[10]The number of lags was determined by considering successive F-test from 5 lags and down as well as looking at the information criteria of Schwarz and Hannan-Quinn. All indications were in favor of 2 lags.

[11]The name of this model, as well as the following considered, are the same as in Johansen (1996).

[12]That is assuming we find evidence of one-to-one impact from the stochastic trend.

Table 2.1. Trace tests

H_0 for rank	Eigenvalues of Π	Trace	Asymp. 95% quant.
r=0	0.12	**15.44**	12.21
r=1	0.04	4.13	4.14

The first hypothesis to be tested in the model $H_2(1)$ is whether the common trend has the same impact on both markets. The hypothesis is formulated as

$$\mathcal{H}_2: \; sp\begin{pmatrix} us & dk \\ 1 & -1 \end{pmatrix} = sp(\beta).$$

With test statistic $\chi^2(1) = 1.48$ ($p - value = 0.22$) the hypothesis is accepted suggesting that the stock markets of the US and Denmark are indeed integrated. To examine the source of the stochastic trend, the following hypothesis is tested:

$$\mathcal{H}_3: \; \begin{pmatrix} us & dk \\ 0 & * \end{pmatrix} \in sp(\alpha).$$

The test is clearly accepted with test statistic $\chi^2(1) = 1.30$ ($p - value = 0.25$), indicating that the development in the Danish stock market is determined by the American stock market. The joint hypothesis $\{\mathcal{H}_2, \mathcal{H}_3\}$ is accepted with $\chi^2(2) = 1.69$ ($p - value = 0.43$) and the restricted error correction model with t-values is given by:

$$\begin{pmatrix} \Delta us \\ \Delta dk \end{pmatrix}_t = \begin{pmatrix} 0 \\ \underset{(3.034)}{0.155} \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix}' \begin{pmatrix} us \\ dk \end{pmatrix}_{t-1} + \ldots.$$

The conclusion of the empirical analysis is that the US and Danish stock markets do seem to be integrated. That is, assets with identical risk-adjusted payoff profile have the same price on both markets. The integration seems to be caused by the fact that Danish stock prices follow those of America.

To see to what extent the role of the deterministic terms affected the conclusion, hypotheses on β and α were also tested in all of the three other models: $H^*(r)$, $H_1(r)$, and $H_1^*(r)$. In all cases the hypotheses of integrated markets and no adjustment of the US market were accepted, supporting the conclusion in the present analysis.

8. Summary of results

The main question analyzed in this chapter can be formulated in the following manner. If we assume that a risky asset is priced according to the CAPM, is it then possible to find evidence for or against market integration, by looking for common stochastic trends in the asset prices? This chapter has demonstrated that the answer is "yes".

In the financial literature there is no commonly accepted definition of when two markets are integrated. However, there seems to be some consensus about the following: If the law of one price holds between two markets then they can be considered integrated. The present analysis shows that if the market portfolios cointegrate one-to-one and share the same long-run trend, then the LOP will hold.

To avoid taking a position on the issue of equalization of risk-free returns across markets, Black's version of the CAPM is used. It is assumed that a portfolio with no correlation with the market portfolio exists. Using this model the price of an arbitrary risky asset is found. This tells us what conditions need to be fulfilled in order for two markets to be integrated. It is then demonstrated how these conditions can be interpreted in terms of restrictions imposed on the VAR model.

An issue which seems to have been somewhat neglected in the literature so far is the question of where the stochastic trends in prices come from. It is demonstrated how tests on the cointegrated VAR model can reveal information about this.

An empirical example is provided in order to illustrate the issues discussed. Evidence from this suggests that the American and the Danish stock markets are integrated. The reason for the integration seems to be that the Danish prices follow the American.

The CAPM, which was used to illustrate the concept of risk adjusting, is a one-period model. It could be interesting to see if similar results apply when assets are priced with a dynamic model. Further empirical investigation should be undertaken to study the degree to which national stock markets are integrated. This issue might be of particular interest in the context of countries in the European Union. Furthermore, the actual timing of integration could be investigated making recursive analyses of the common stochastic trends. These issues will be left for future research.

9. Appendix A

We consider the problem (4.1) disregarding (*iii*). The problem can be solved using several methods. Here it is solved applying the Lagrange method. The Lagrangian is

$$\pounds(W, \lambda_1, \lambda_2) = W'\Sigma W - \lambda_1(W'R^e - r^e) - \lambda_2(W'\mathbf{1}-1),$$

where $\lambda_1, \lambda_2 > 0$. The necessary and sufficient first order conditions are:

$$\frac{\partial \pounds}{\partial W} = 2\Sigma W - \lambda_1 R^e - \lambda_2 \mathbf{1} = \mathbf{0}, \tag{9.1}$$

$$\frac{\partial \pounds}{\partial \lambda_1} = W'R^e - r^e = 0, \tag{9.2}$$

$$\frac{\partial \pounds}{\partial \lambda_2} = W'\mathbf{1} - 1 = 0. \tag{9.3}$$

Using (9.1) and (9.2) and (9.1) and (9.3) we find

$$r^e = \frac{1}{2}\lambda_1(R^{e\prime}\Sigma^{-1}R^e) + \frac{1}{2}\lambda_2(R^{e\prime}\Sigma^{-1}\mathbf{1}), \tag{9.4}$$

$$1 = \frac{1}{2}\lambda_1(\mathbf{1}'\Sigma^{-1}R^e) + \frac{1}{2}\lambda_2(\mathbf{1}'\Sigma^{-1}\mathbf{1}). \tag{9.5}$$

To simplify notation we define[13]

$$A = \mathbf{1}'\Sigma^{-1}R^e, \qquad B = R^{e\prime}\Sigma^{-1}R^e,$$
$$C = \mathbf{1}'\Sigma^{-1}\mathbf{1}, \qquad D = BC - A^2.$$

The formulas (9.4) and (9.5) form a system of two equations and two unknown variables. This can be solved with respect to the Lagrange multipliers:

$$\lambda_1 = 2\frac{Cr^e - A}{D}, \tag{9.6}$$

$$\lambda_2 = 2\frac{B - Ar^e}{D}. \tag{9.7}$$

Inserting (9.6) and (9.7) in (9.1) yields the optimal portfolio:

(9.8)

$$W = \frac{1}{D}\left[B(\Sigma^{-1}\mathbf{1}) - A(\Sigma^{-1}R^e)\right] + \frac{1}{D}\left[C(\Sigma^{-1}R^e) - A(\Sigma^{-1}\mathbf{1})\right] r^e,$$

which is formula (4.2) with $g = \frac{1}{D}\left[B(\Sigma^{-1}\mathbf{1}) - A(\Sigma^{-1}R^e)\right]$ and $h = \frac{1}{D}\left[C(\Sigma^{-1}R^e) - A(\Sigma^{-1}\mathbf{1})\right]$.

Let us take a look at the graphical representation of the so-called frontier, which is defined as the set of portfolios W which have the smallest possible variance $W'\Sigma W = \sigma^2$ given the expected return $W'R^e = r^e$. Using (9.8) we can find

$$\sigma^2 = \frac{C}{D}(r^e - \frac{A}{C})^2 + \frac{1}{C} \Longleftrightarrow \frac{(\sigma - 0)^2}{\left(\sqrt{1/C}\right)^2} - \frac{(r^e - A/C)^2}{\left(\sqrt{D}/C\right)^2} = 1,$$

which is a hyperbola in the $\sigma - r^e$ space with center in $(\sigma, r^e) = (0, \frac{A}{C})$ and asymptotes $r^e = \frac{A}{C} \pm \sqrt{\frac{D}{C}}\sigma$.

[13] The short-hand notation of Huang and Litzenberger (1988) is applied.

The minimum variance portfolio (mvp) is defined as the portfolio with the lowest possible variance, i.e. the portfolio with $(\sigma, r^e) = (\sqrt{\frac{1}{C}}, \frac{A}{C})$.

10. Appendix B

In this appendix it is demonstrated that the covariance between the rates of return of any two frontier portfolios, p and q, can be expressed as in (4.3). The covariance is defined as $cov(r_p^e, r_q^e) = W_p' \Sigma W_q$, with W_p and W_q given in (4.2). Using (4.2) we find

$$(g + hr_p^e)'\Sigma(g + hr_q^e) = g'\Sigma g + h'\Sigma g r_p^e + g'\Sigma h r_q^e + h'\Sigma h r_p^e r_q^e.$$

We apply the definitions of g and h given in appendix A and consider each of the terms separately:

$$\begin{aligned}
g'\Sigma g &= \left[\tfrac{B}{D}(\Sigma^{-1}\mathbf{1}) - \tfrac{A}{D}(\Sigma^{-1}R^e)\right]' \left[\tfrac{B}{D}\mathbf{1} - \tfrac{A}{D}R^e\right] \\
&= \tfrac{B^2C}{D^2} - \tfrac{A^2B}{D^2} - \tfrac{A^2B}{D^2} + \tfrac{A^2B}{D^2} \\
&= \tfrac{B^2C - A^2B}{D^2} = \tfrac{B}{D},
\end{aligned}$$

$$\begin{aligned}
h'\Sigma g r_p^e &= \left[\tfrac{C}{D}(\Sigma^{-1}R^e) - \tfrac{A}{D}(\Sigma^{-1}\mathbf{1})\right]' \left[\tfrac{B}{D}\mathbf{1} - \tfrac{A}{D}R^e\right] r_p^e \\
&= \left(\tfrac{ABC}{D^2} - \tfrac{ABC}{D^2} - \tfrac{ABC}{D^2} + \tfrac{A^3}{D^2}\right) r_p^e \\
&= \left(\tfrac{A^3 - ABC}{D^2}\right) r_p^e = -\tfrac{A}{D} r_p^e,
\end{aligned}$$

$$\begin{aligned}
g'\Sigma h r_q^e &= \left[\tfrac{B}{D}(\Sigma^{-1}\mathbf{1}) - \tfrac{A}{D}(\Sigma^{-1}R^e)\right]' \left[\tfrac{C}{D}R^e - \tfrac{A}{D}\mathbf{1}\right] r_q^e \\
&= \left(\tfrac{ABC}{D^2} - \tfrac{ABC}{D^2} - \tfrac{ABC}{D^2} + \tfrac{A^3}{D^2}\right) r_q^e \\
&= \left(\tfrac{A^3 - ABC}{D^2}\right) r_q^e = -\tfrac{A}{D} r_q^e,
\end{aligned}$$

$$\begin{aligned}
h'\Sigma h r_p^e r_q^e &= \left[\tfrac{C}{D}(\Sigma^{-1}R^e) - \tfrac{A}{D}(\Sigma^{-1}\mathbf{1})\right]' \left[\tfrac{C}{D}R^e - \tfrac{A}{D}\mathbf{1}\right] r_p^e r_q^e \\
&= \left(\tfrac{C^2B}{D^2} - \tfrac{A^2C}{D^2} - \tfrac{A^2C}{D^2} + \tfrac{A^2C}{D^2}\right) r_p^e r_q^e \\
&= \left(\tfrac{C^2B - A^2C}{D^2}\right) r_p^e r_q^e = \tfrac{C}{D} r_p^e r_q^e.
\end{aligned}$$

The terms are collected and with some algebra we find

$$\begin{aligned}
cov(r_p^e, r_q^e) &= \tfrac{B}{D} - \tfrac{A}{D} r_p^e - \tfrac{A}{D} r_q^e + \tfrac{C}{D} r_p^e r_q^e \\
&= \tfrac{C}{D}\left(r_p^e - \tfrac{A}{C}\right)\left(r_q^e - \tfrac{A}{C}\right) + \tfrac{1}{C},
\end{aligned}$$

which is the same as (4.3) since $r_{mvp}^e = \frac{A}{C}$ and $\frac{C}{D}\frac{D}{C^2} = \frac{1}{C}$.

CHAPTER 3

Does the PPP hold within the US?

Abstract: This chapter investigates whether the PPP holds between four metropolitan areas in the US applying a cointegrated VAR model. Three definitions form the basis of the empirical analysis. One relates the concept of underlying inflation to the sharing of common stochastic I(2) trends. The second is simply an econometric formulation of the PPP, whereas the third allows for adjustment of price levels. Evidence is found in favor of the same underlying inflation in the four areas. Adjustment of price levels, however, has only taken place between the three geographically closer areas.

1. Introduction

The present chapter investigates the extent to which the Purchasing Power Parity (PPP) holds within the US. Evidence is found that prices are integrated of order 2 making it possible to take advantage of the rich structure of the cointegrated Vector Autoregressive (VAR) model for I(2) variables. This chapter contributes to the existing literature by testing the PPP within an area with no trade barriers and where trade is associated with no risk of fluctuations in exchange rates. In this sense, it can also be considered as a benchmark case for what can be expected within the euro area with respect to the convergence of price levels.

Cassel (1922) was one of the first economists to pay attention to the question of whether or not goods are priced identically in different physical locations. He notes that "When two currencies have undergone inflation, the normal rate of exchange will be equal to the old rate multiplied by the quotient of the degree of inflation in the one country and in the other. There will naturally always be found deviations from this new normal rate, and during the transition period these deviations may be expected to be fairly wide.". Within the last two decades, in particular, extensive research has focused on how to measure and test the law of one price (LOP) between two identical goods traded at different physical locations. In the aggregated form this is known as the PPP. The development of strong econometric tools has intensified the research still further. Excellent reviews of literature and methods for empirical testing are given by Breuer (1994) and Froot and Rogoff (1995).

Much recent empirical testing of PPP has applied Johansen's (1991) econometric method to detect cointegration between non-stationary time series in a multivariate framework. Traditional studies focus on two or more countries with different currencies. These can be countries with either floating or fixed exchange rates. The present analysis, on the other hand, applies the Johansen technique to test whether the PPP holds within a currency union, i.e. an area with perfectly fixed exchange rates.

The data used here consists of monthly observations for the consumer price indices (CPI) for a period of 45 years from four metropolitan areas in the US. Based on evidence from earlier studies Froot and Rogoff (1995) note, among other things, the following: "... it is easier to reject the no-cointegration null across pairs of currencies that are fixed than across pairs that are floating."; "...tests based on CPI price levels tend to reject less frequently than tests based on WPIs."; and "... cointegration tests seem to yield much more reliable results when estimated over long sample periods...". Hence, a priori one might expect to find strong evidence in favor of PPP in the empirical analysis carried out in the present context. As will be revealed, however, this is not the case.

To my knowledge this is the first paper to study PPP within a single country applying a multivariate cointegration framework allowing prices to be integrated of order 2. Several papers have, however, investigated whether the LOP holds within a single country. Engle and Rogers (1996) investigate price differences between cities in America as well as in Canada. They find that the distance between cities matters in terms of relative prices, but geographical borders matter more. Parsley and Wei (1996) study 51 prices (traded and non-traded goods) from 48 US cities. They find evidence that the convergence rate to the PPP is higher than that which is typically found in cross-country studies. This rate, however, is slower the farther apart the cities are. Bayoumi and MacDonald (1998) apply a panel data framework and reach similar conclusions. They argue that relative price movements within countries are caused by real factors. Such results might throw doubt on the likelihood of finding strong evidence in favor of PPP in the present context. In a recent study Engle and Rogers (2001) surprisingly discover that variability in prices in the US are higher for traded than non-traded goods.

The rest of the chapter is organized as follows: The next section contains a brief discussion of theoretical and empirical considerations with respect to the PPP theory in the context applied in this chapter. A new definition of PPP which takes into account that price levels might adjust to each other is introduced and explained by a small economic model. In section 3, the econometric methods used in the empirical analysis in section 4 are discussed, while section 5 summarizes and concludes the analysis.

2. Theoretical and empirical aspects of PPP

In theory the price of a good should be the same even though it is traded in different physical locations. If this were not the case consumers would simply purchase the good where it is cheaper, forcing the producers to equalize prices. In reality, however, this is not the case for a number of reasons, such as cost of transportation, taxes etc. Rigidities in prices (such that they do not adjust as fast as exchange rates change) might also be a reason for the PPP not to hold in the short run. Yet prices should adjust in the long run to avoid arbitrage. Hence, PPP is considered a feature most likely to hold in the long run.

Let us first consider a simple general form of the parity in absolute form. Letting P^i be the price of a particular good - or more precisely a basket of goods - in country i and E^{ij} the exchange rate between country i and j (i.e. the price in country i's currency for one unit of country j's currency), the PPP between country i and j can be written as

$$P^i = P^j E^{ij}.$$

In the empirical analysis below it is investigated whether the PPP holds between four areas in the US. Since these use the same currency, we can simply set $E^{ij} = 1$. Letting small letters denote logarithm, the absolute PPP in the present context reads:

$$p^i = p^j. \tag{2.1}$$

According to (2.1) the price levels should be the same in all areas within the US. The relative form of the PPP states that the inflation rate should be the same in these areas:

$$\Delta p^i = \Delta p^j. \tag{2.2}$$

The relations (2.1) and (2.2) are based on theoretical relations and will only hold in specific cases. In general, there will exist a process h_t such that the following relation will hold:

$$p_t^i = p_t^j + h_t + const,$$

where $h_0 = 0$ and $const = p_0^i - p_0^j$. The relative version of PPP will hold if $h_t = 0$ for all t, and the absolute version holds if furthermore $const = 0$, i.e. if $p_0^i = p_0^j$. The term h_t can be interpreted as the deviation from the PPP.

When testing theoretical formulations such as (2.1) and (2.2) empirically, a more flexible formulation might be needed.[1] This refers to the stochastic properties of the time series. We will allow for the prices to be integrated of order 1 or 2. Hence, it follows that $h_t \sim I(d)$, for $d = 0, 1, 2$. If $d = 0$ the PPP holds. Since indices are considered the condition $const = 0$ is not straightforward to interpret. It will be argued below that if price levels are adjusting towards each other, the process h_t might be interpreted in terms of the inflation rate. Hence, we get a relation on the form

$$p^i = p^j + \kappa \Delta p^i, \tag{2.3}$$

where κ is a coefficient which could be positive, negative or possible zero in which case (2.1) and (2.3) coincide.

The econometric implications of (2.2) and (2.3) depend on the properties of the time series for the prices. Some empirical evidence suggests that inflation is non-stationary, i.e. that price levels are integrated of order 2. If this is the case and the impact from the $I(2)$ trend is the same, the properties of the series can be given interpretations related to the PPP. These are given in the three following (econometric) definitions:[2]

DEFINITION 1. *Let us consider two time series for prices and let* $p_t^i, p_t^j \sim I(2)$. *If* p_t^i *and* p_t^j *cointegrate such that* $p_t^i - p_t^j \sim I(1)$, *that is* $\Delta p_t^i - \Delta p_t^j \sim I(0)$, *then we will say that the underlying inflation is the same in areas* i *and* j.

DEFINITION 2. *Assume that either* $p_t^i, p_t^j \sim I(2)$ *or* $p_t^i, p_t^j \sim I(1)$ *and that* $p_t^i - p_t^j \sim I(0)$, *then we will say that the PPP holds.*

DEFINITION 3. *Let* $p_t^i, p_t^j \sim I(2)$. *If* $p_t^i - p_t^j + \kappa \Delta p_t^i \sim I(0)$ *for some* $\kappa \neq 0$, *then we will say that the PPP with adjustment holds.*

Whereas the economic intuition of Definition 2 is quite apparent, further comments on Definitions 1 and 3 might be in place. In Definition 1 it is important to notice that the stochastic variation in the price levels is dominated by the $I(2)$ness. Hence, if the levels share the same

[1] See also Haavelmo (1944) and Juselius (1995) for discussions of this point.

[2] The definitions are stated for the purpose of the present analysis such that the exchange rate is eliminated. They are, however, easy to extend to the case of areas with different currencies, see chapter 4.

stochastic $I(2)$ trend and the impact from this is the same, the development in the levels are similar from a stochastic point of view. This implies that the inflation rates are affected by the same impact from a common $I(1)$ trend and it will be referred to this as the underlying inflation is the same.[3] There is some relation between Definition 1 and the relative form of the PPP expressed in (2.2). In fact, Definition 1 can be thought of as the equivalence of the relative PPP when inflation is non-stationary. That the definition does not capture the economic concept of relative PPP is due to the fact that the price differential is $I(1)$ so that the two levels in principle could diverge. Note that for two price series to obey the relative PPP (in it's most strict form) it is required that $\Delta p_t^i - \Delta p_t^j \sim I(0)$ with mean zero.

To explain Definition 3 in greater detail, we consider two economies which seek to integrate their markets.[4] A measure for the degree of integration could be the difference of price levels: if markets for goods are perfectly integrated, prices should be the same. Assume for simplicity, that the price level in i is lower than in j. Hence, the problem is to minimize $p_t^i - p_t^j$. This, however, has a cost, namely inflation. The problem can be formulated in terms of a linear-quadratic adjustment cost model:

$$\min_{\{p_s^i\}} E_t \sum_{s=t}^{\infty} \beta^{s-t} \left[\delta \left(p_s^i - p_s^j\right)^2 + (\Delta p_s^i)^2 \right],$$

where $\delta > 0$ is the relative weight between the benefit of integration and the cost of inflation, $0 < \beta < 1$ is a discount factor making today more important than tomorrow.[5] In this type of model the variable p_t^j is referred to as the tracking variable and it will be assumed that $p_t^j = x_t'\theta + e_t$, where x_t is a vector of forcing variables and $e_t \sim iid(0, \sigma_e^2)$. The information set at time t is $\mathcal{F}_t = \{e_t, p_{t-j}^i, x_{t-j+1}\}_{j=1}^{\infty}$. However, e_t is assumed not to be known to the investigating econometrician, whose information set $\mathcal{G}_t \subset \mathcal{F}_t$.

The Euler equation gives the first-order condition for optimum:

$$\Delta p_t^i = \beta E_t \Delta p_{t+1}^i - \delta(p_t^i - p_t^j). \tag{2.4}$$

[3]The underlying inflation rate should not be mixed up with the core inflation, which is the inflation rate when the most volatile components are taking out of the price index.

[4]The following discussion is motivated by Gregory et al. (1993) and Gregory (1994).

[5]In this particular model β can be expected to be very close to one.

Formula (2.4) constitutes a second-order difference equation and the characteristic polynomial for this is

$$\beta z^2 - (1+\beta+\delta)z + 1 = 0.$$

This has two solutions: One larger than one and one smaller. Here we are only interested in the stable root, which will be named $\lambda(< 1)$. The solution of (2.4) is then given by

$$p_t^i = \lambda p_{t-1}^i + (1-\lambda)(1-\beta\lambda)E_t \sum_{s=t}^{\infty} (\beta\lambda)^{s-t} p_s^j \tag{2.5}$$

Two cases will now be considered. First, the one where the process for the tracking variable is integrated of order one and then where it is integrated of order two, i.e. $x_t \sim I(1)$ and $x_t \sim I(2)$. Case 2 is the interesting one in the present context. We consider for simplicity the case where x_t is a univariate process.

Case 1. $x_t \sim I(1)$. Let the stochastic process for x_t be given by

$$\Delta x_t = \varepsilon_t,\ \varepsilon_t \sim iid(0, \sigma_\varepsilon^2).$$

Then $E_t(x_{t+h}) = x_t$ for $h = 0, 1, ..., \infty$. It follows that $E_t \sum_{s=t}^{\infty} (\beta\lambda)^{s-t} p_s^j = \theta/(1-\beta\lambda)x_t + e_t$.[6] Inserting in (2.5) gives us the so-called partial adjustment model

$$p_t^i = \lambda p_{t-1}^i + (1-\lambda)\theta x_t + (1-\beta\lambda)(1-\lambda)e_t. \tag{2.6}$$

Rewriting (2.6) gives us the error correction form (ECM):

$$\Delta p_t^i = (\lambda - 1)(p_{t-1}^i - \theta x_{t-1}) + (1-\lambda)\theta\Delta x_t + (1-\beta\lambda)(1-\lambda)e_t. \tag{2.7}$$

In order to see the cointegrating relations, (2.7) is reparametrized:[7]

$$p_t^i = \theta x_t - \frac{\lambda}{1-\lambda}\Delta p_t^i + (1-\beta\lambda)e_t. \tag{2.8}$$

Since the two last terms in (2.8) are stationary it follows that p_t^i and x_t cointegrate with the vector $(1, -\theta)$. Hence also p_t^i and p_t^j cointegrate with this vector.

Case 2. $x_t \sim I(2)$. We now consider the case were x_t is integrated of order two:

$$\Delta^2 x_t = \varepsilon_t \Leftrightarrow x_t = 2x_{t-1} - x_{t-2} + \varepsilon_t, \varepsilon_t \sim iid(0, \sigma_\varepsilon^2).$$

[6]To obtain this the following relation has been applied

$$a + ak + ak^2 + ... + ak^n + ... = a/(1-k)\ if \mid k \mid < 1.$$

[7]In the present case it boils down to multiplying the equation with $1/(1-\lambda)$ and rearranging.

We have that $E_t(x_{t+h}) = (h+1)x_t - hx_{t-1} = x_t + h\Delta x_t$, for $h = 0, 1, ..., \infty$. We use this to find[8] $E_t \sum_{s=t}^{\infty} (\beta\lambda)^{s-t} p_s^j = \theta/(1-\beta\lambda) * x_t + \theta\beta\lambda/(1-\beta\lambda)^2 * \Delta x_t + e_t$, which is inserted in (2.5):

$$(2.9)\quad p_t^i = \lambda p_{t-1}^i + \theta(1-\lambda)x_t + \frac{\theta\beta\lambda(1-\lambda)}{1-\beta\lambda}\Delta x_t + (1-\lambda)(1-\beta\lambda)e_t.$$

Reorganizing (2.9) yields the ECM:

(2.10)

$$\Delta p_t^i = (\lambda-1)(p_{t-1}^i - \theta x_{t-1}) + \frac{\theta(1-\lambda)}{1-\beta\lambda}\Delta x_t + (1-\beta\lambda)(1-\lambda)e_t.$$

Note that in (2.10) $p_{t-1}^i, x_{t-1} \sim I(2)$, $\Delta p_t^i, \Delta x_t \sim I(1)$, and $e_t \sim I(0)$. Again we reparametrize:

$$(2.11)\qquad p_t^i = \theta x_t - \frac{\lambda}{1-\lambda}\Delta p_t^i + \frac{\beta\lambda}{1-\beta\lambda}\Delta x_t + (1-\beta\lambda)(1-\lambda)e_t.$$

From formula (2.11) it appears - since e_t is stationary - that cointegration from $I(2)$ to $I(1)$ exists such that $p_t^i - \theta x_t \sim I(1)$. Furthermore, $p_t^i - \theta x_t + \lambda/(1-\lambda)\Delta p_t^i - \beta\lambda/(1-\beta\lambda)\Delta x_t \sim I(0)$, which is a relation as in Definition 3 if $\theta = 1$. Since p_t^i and x_t share the same $I(2)$ trend, Δp_t^i and Δx_t share the same $I(1)$ trend and thus only one of the differences is needed to obtain stationarity.

Note that in definition 2 and 3 it will trivially hold that $\Delta p_t^i - \Delta p_t^j \sim I(0)$. Hence if the PPP holds (possibly with adjustment) then the underlying inflation rate in the areas must be the same. In other words, if prices are integrated of order 2 a necessary condition for the PPP (maybe with adjustment) to hold is that the areas have the same underlying inflation.

3. The statistical model

3.1. The I(1) case. For the statistical analysis, consider an unrestricted VAR(k) model which in error correction form is written as:[9]

$$(3.1)\qquad \Delta X_t = \Pi X_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} + \Phi D_t + \varepsilon_t,$$

where ε_t are identically independent distributed (*iid*) errors with mean zero and covariance matrix Ω. The data vector X_t consists of the four price indices such that $X_t = \{chi, ny, phil, la\}_t$.[10] The matrices Π, Φ

[8]See Appendix A.

[9]For a more formal and complete treatment of the procedures described below see, for example, the textbook by Johansen (1996).

[10]See section 4.1 for a detailed description of data.

and Γ_i are parameters to be estimated while D_t includes the deterministic terms of the model.

In the unrestricted version of (3.1), X_t will in general be $I(0)$. The $I(1)$ and $I(2)$ models are nested in this one. The hypothesis that X_t is integrated of order one can be formulated as the double requirement that $\Pi = \alpha\beta'$ has reduced rank $r < p(= 4)$ and $\alpha'_\perp \Gamma \beta_\perp$ ($\Gamma = I - \sum_{i=1}^{k-1} \Gamma_i$) has full rank $p - r$.[11] The second part of the requirement ensures that the $I(1)$ space has full rank and hence that X_t is not $I(2)$.

A test for the number of common trends in the I(1) model was developed by Johansen (1988, 1991). The principle of the so-called Trace-test is to maximize the likelihood functions under the null-hypothesis and the alternative hypothesis by applying the technique of reduced rank regression of Anderson (1951). The likelihood functions are maximized by solving eigenvalues problems and a likelihood ratio test for the hypothesis of r cointegrating vectors against the alternative of p can then be formulated as:

$$-2 \ln Q(r \mid p) = -T \sum_{i=r+1}^{p} \ln(1 - \widehat{\lambda}_i),$$

where $\widehat{\lambda}_i$ are the estimated eigenvalues and T is the total number of observations in the sample. The asymptotic distribution and critical values for the test statistic are derived by Johansen and Juselius (1990).

If there exists $p \times r$ matrices of full rank such that $\Pi = \alpha\beta'$ and $\alpha'_\perp \Gamma \beta_\perp$ has full rank, then the moving average representation for X_t is given by:[12]

$$X_t = C \sum_{i=1}^{t} (\varepsilon_i + \Phi D_i) + C(L)(\varepsilon_t + \Phi D_t) + A,$$

where $C = \beta_\perp (\alpha'_\perp \Gamma \beta_\perp)^{-1} \alpha'_\perp$, $C(L)$ is a convergent power series such that the effect of a shock at time t will die out and thus will have no long-run effect, and A depends on initial conditions so that $\beta' A = 0$. Since $\alpha'_\perp \sum_{i=1}^{t} \varepsilon_i$ is the only non-stationary part of the process these are referred to as the common stochastic trends. Notice that the C matrix has reduced rank $p - r$. The matrix $\alpha'_\perp$ represent the coefficients for the common trends and $\beta_\perp$ represent the loadings from the common trends into the process. The term $(\alpha'_\perp \Gamma \beta_\perp)^{-1}$ is simply a normalization.

[11] The notation $\perp$ indicates an orthogonal complement such that $\alpha'_\perp \alpha = 0$ and $\beta'_\perp \beta = 0$.

[12] According to Granger's representation Theorem. See Engle and Granger (1987).

The moving average representation gives a nice intuitive understanding of test of hypotheses formulated on β and α in the case of $r = p-1$ cointegration vectors. In this case both $\alpha_\perp$ and $\beta_\perp$ are $p \times 1$ vectors. When $r < p-1$ it is a more complicated matter since $\alpha_\perp$ and $\beta_\perp$ have more than one column and hence it is the span of these columns which should be interpreted.

An interesting hypothesis on α could be whether it has any zero-rows. For $r = p-1$ a zero-row for variable i in α implies zeros in all rows except for i in $\alpha_\perp$. This can be formulated as weak exogeneity of ΔX_{it} for β, see Engle et al. (1983).

The matrix $\beta_\perp$ represents the loadings from the common trends into the variables, i.e. the impact of the common trends. The dimension of $\beta_\perp$ is $p \times (p-r)$ and for the case $r = p-1$ it is simply a $p \times 1$ vector. In many econometric analyses an interesting hypothesis on β is for some kind of homogeneity between all variables, i.e. if the coefficients for every pair of variables sum to one. This hypothesis can also be formulated in terms of whether the variables are equally affected by the common trend. In other words, under this hypothesis the coefficients in $\beta_\perp$ are equal.

The techniques for testing hypotheses on α and β were developed by Johansen and Juselius (1990, 1992, 1994). A general hypothesis of no adjustment to long-run relations can be formulated as a linear restriction on the columns of α as $\alpha = H\psi$, where H is a $p \times m$ design matrix imposing $p-m$ restrictions and ψ includes $m \times r$ parameters to be estimated.

In the case of more than one cointegration vector, tests on β are often carried out in two steps. The first involves testing hypotheses on individual vectors. These can be formulated as $\beta = \{H\phi, \psi\}$, where H is a $p \times m$ design matrix imposing $p-m$ restrictions, ϕ is a $m \times 1$ matrix of free parameters and ψ is a $p \times (r-1)$ matrix of unrestricted coefficients. The second step is a joint test of hypotheses accepted in step one.

3.2. The I(2) case. For the error correction model in the I(2) case (3.1) is reparametrized in accelerations, velocity and levels:

$$\Delta^2 X_t = \Pi X_{t-1} - \Gamma \Delta X_{t-1} + \sum_{i=1}^{k-2} \Psi_i \Delta^2 X_{t-i} + \Phi D_t + \varepsilon_t.$$

The parameters to be estimated are Π, Γ, Ψ_i, and Φ. The data matrix X_t is $I(2)$ if $\Pi = \alpha\beta'$ has reduced rank $r < p$, $\alpha'_\perp \Gamma \beta_\perp = \xi\eta'$ has reduced rank $s_1 < p-r$ and the $I(2)$ space has full rank, i.e. $\alpha'_{\perp 2}\theta\beta_{\perp 2}$ has full rank.[13] In this case there are r (maybe multi or polynomial) cointegration vectors, s_1 common $I(1)$ trends and $s_2 = p - r - s_1$ common $I(2)$ trends. The process can be rotated to separate direct cointegrating, polynomial cointegrating and non-cointegrating directions. With the notation here $\beta_{\perp 2}$ is orthogonal to $(\beta, \beta_{\perp 1})$, $sp(\beta, \beta_{\perp 1}, \beta_{\perp 2}) = \mathrm{R}^p$ and $\beta = (\beta_0, \beta_1)$. The same notation holds with respect to α.

If $r > s_2$, $r - s_2$ vectors cointegrate directly from $I(2)$ to $I(0)$ in the β_0 direction. With the terminology of Engle and Granger (1987) these are $CI(2,2)$. There are s_2 vectors which multicointegrate - i.e. they require differences of data to obtain stationarity - from $I(2)$ to $I(0)$ in the β_1 direction. The associated stationary process for the multicointegrating relations are given by $\beta_1' X_t + \kappa' \Delta X_t \sim I(0)$, and this relates to Definition 3. Furthermore, s_1 directions cointegrate from $I(2)$ to $I(1)$ - i.e. are $CI(2,1)$ - in the $\beta_{\perp 1}$ direction.

If $X_t \sim I(2)$, then $\Delta X_t \sim I(1)$ and the number of unit roots in the characteristic polynomial is $s_1 + 2s_2$. Notice that the difference of a stochastic $I(2)$ trend is $I(1)$, whereas s_1 is the number of "independent" $I(1)$ trends, i.e. those which are not associated with any $I(2)$ trends.

The moving average representation of an $I(2)$ process is given by:[14]

$$\begin{aligned} X_t = \; & C_2 \textstyle\sum_{s=1}^{t} \sum_{i=1}^{s} (\varepsilon_i + \Phi D_i) + C_1 \sum_{i=1}^{t} (\varepsilon_i + \Phi D_i) \\ & + C_2(L)(\varepsilon_t + \Phi D_t) + A + Bt, \end{aligned}$$

where $C_2 = \beta_{\perp 2}(\alpha_{\perp 2}' \theta \beta_{\perp 2})^{-1} \alpha_{\perp 2}'$, $\beta' C_1 = -(\alpha'\alpha)^{-1} \Gamma C_2$ and $\beta_{\perp 1}' C_1 = -\overline{\alpha}_{\perp 1}'(I - \theta C_2)$. The terms A and B are functions of the initial conditions.[15] The interpretation of C_2 is equivalent to the C matrix in the $I(1)$ case. The matrix $\alpha_{\perp 2}$ are the coefficients to the common $I(2)$ trends and $\beta_{\perp 2}$ are the loadings. The C_1 matrix is a more complicated matter and does not have the same 'nice' interpretation.

Whereas tests for adjustment to long-run relations in the $I(1)$ model is simply a test on the α matrix, it is a more complicated matter in case of $I(2)$ness. Paruolo and Rahbek (1999) show that not only will this be a test for restrictions on α but also on $\alpha_{\perp 1}$ and $\beta_{\perp 1}$. The theory for testing hypothesis on $\beta_{\perp 1}$ and $\beta_{\perp 2}$ has been developed too (see Johansen (1997) and Paruolo (1998)).

---

[13] The short-hand notation of Johansen (1996) is used: $\theta = \Gamma \overline{\beta} \overline{\alpha}' \Gamma + \sum_{i=1}^{k-1} i \Gamma_i$, where in general - here and in the following - $\overline{\alpha} = \alpha(\alpha'\alpha)^{-1}$.

[14] See Johansen (1996) Theorem 4.6.

[15] The matrices A and B satisfy the conditions $(\beta, \beta_{\perp 1})' B = 0$ and $\beta' A - \overline{\alpha} \Gamma \overline{\beta}_{\perp 2} \beta_{\perp 2}' B = 0$.

3.3. Scenario analysis. When discussing price data there is some dispute between economists and econometricians. From a theoretical point of view, it is not possible for the inflation rate to be non-stationary in the long run. On the other hand much empirical research suggests that the inflation rate does indeed seem to be integrated of order one, i.e. the price acceleration is stationary when analyzing a sample which consists of course of less than infinitely many observations.[16]

Let us first consider the case where the prices are integrated of order 2. We consider a system with four variables: *chi*, *ny*, *phil*, and *la*. One implication of Definition 1 is that the four indices share the same $I(2)$ trend. In the case of one common $I(2)$ trend and, say, one (independent) common $I(1)$ trend, the system can be written as follows:

$$\begin{bmatrix} chi_t \\ ny_t \\ phil_t \\ la_t \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{bmatrix} \sum_{s=1}^{t} \sum_{i=1}^{s} u_i + \begin{bmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \\ d_{31} & d_{32} \\ d_{41} & d_{42} \end{bmatrix} \begin{bmatrix} \sum_{i=1}^{t} u_i \\ \sum_{i=1}^{t} v_i \end{bmatrix} + stat,$$

where in general u_i and v_i are functions of ε_{chi}, ε_{ny}, ε_{phil} and ε_{la}. The term *stat* refers to the stationary part of the process. Note that in this case we have $(r, s_1, s_2) = (2, 1, 1)$ which implies three unit roots in the characteristic polynomial but only two zero roots in the matrix Π. According to Definition 1, it is required that $c_1 = c_2 = c_3 = c_4$ if the underlying inflation is the same in the four areas. In this case the system can be transformed into the $I(1)$ space with no loss of information, i.e. the likelihood functions of the two systems will be approximately the same. A transformed system could look as follows:

$$\begin{bmatrix} chi_t - ny_t \\ chi_t - phil_t \\ chi_t - la_t \\ \Delta chi_t \end{bmatrix} = \begin{bmatrix} d_{11} - d_{21} & d_{12} - d_{22} \\ d_{11} - d_{31} & d_{12} - d_{32} \\ d_{11} - d_{41} & d_{12} - d_{42} \\ c_1 & 0 \end{bmatrix} \begin{bmatrix} \sum_{i=1}^{t} u_i \\ \sum_{i=1}^{t} v_i \end{bmatrix} + stat.$$

In the transformed model interesting hypotheses can be formulated based on whether or not cointegration vectors consist of a price differential and the inflation rate according to Definition 3. As an example we look for cointegration between $chi_t - ny_t$ and Δchi_t. The combination $(1, 0, 0, \kappa)X_t$ will be stationary if $\kappa c_1 = -(d_{11} - d_{21})$ and $d_{12} = d_{22}$. The implication is that not only is the underlying inflation the same in

[16]See for example Juselius (1999).

the areas of Chicago and New York (Definition 1), but also the PPP with adjustment (Definition 3) holds. This is the case because, even though the price levels do not cointegrate directly to $I(0)$, they do multicointegrate. Hence, the price levels seems to have adjusted, perhaps towards a sustainable PPP level.

In the case where $X_t \sim I(1)$ the underlying inflation is determined by the $I(1)$ part of the process. Hence, areas have the same underlying inflation if they share one common $I(1)$ trend with the same impact. But in this case this follows trivially from the fact that the PPP (Definition 2) holds.

It should be clear that if the question of whether there is an $I(2)$ trend in data is in doubt (maybe there is an almost $I(2)$ trend), then it is better to treat the model as $I(2)$ in this case of examining multiple PPP. To illustrate this, let us - falsely - assume that $X_t \sim I(2)$. The next step after detecting the nominal trend is to find the impact on each of the variables and figure out if it is possible to make a transformation of the model into the $I(1)$ space. Even though the variables in fact were all $I(1)$, this transformation would still be valid if the PPP holds. In this case we would end up with a stationary system. Hence, if the transformation cannot be made it is indeed evidence that the PPP does not hold between the areas.

4. Empirical analysis

4.1. Data and misspecification tests. The LOP is a theory stating that two identical goods should have the same price regardless of where they are traded. Arbitrage in good markets will equalize prices, in the case of areas with different currencies this could happen via the exchange rate. This of course makes no sense if the good is defined as an Arrow-Debreu good, which is - among other things - described by the time and location of the purchase.

In its aggregate form the LOP is referred to as the PPP. According to this all identical goods will be priced identically at different locations. A more flexible formulation of this - which might be useful when testing the PPP empirically - could state that some kind of geometric weighted average of prices should be the same at different locations. This formulation allows for different demands in different areas to have an influence. For example, it makes little sense to compare prices on winter clothes in Greenland and Florida, as the demands are very different. Furthermore, the more flexible formulation might also to a greater extent mirror consumer behavior. Often consumers do not bother to go shopping in two supermarkets, even though two different articles, say bread and wine, are cheaper in different markets. An economic argument for this behavior could be a kind of shoe-leather effect: 'it is not worth doing the extra walking to save a few pennies'.

How should the prices be aggregated then? This question does not have an obvious answer and is still an issue for discussion. One answer is to aggregate with the weights of the consumption. This has the advantage of comparing the cost of the consumption and in this sense a kind of 'cost-of-living', which takes account of the fact that geographical, cultural and other circumstances might require consumption of particular goods such as the example of winter clothes in Greenland. On the other hand, this could result in comparing prices on "apples" and "bananas", which is not appropriate when testing the PPP.

Hence, a discussion of the data used in empirical testing of LOP and PPP is certainly important. In the present analysis indices for the development in consumer prices are considered. More precisely, four indices from major metropolitan areas in the US are used on a monthly basis covering the period from 1953 to 1997, which gives a total of 45 years or 540 observations. The starting date of the period was chosen to avoid effects from World War II and to avoid using data from the beginning of the period which may be too imprecise (see below). The end-date is due to limitation in monthly observation in one of the series. The data was extracted from the US Bureau of Labor Statistic's (BLS) homepage[17] and is collected by local branches, as can be seen in Table 3.1.

Table 3.1. Description of data

Variable name	Metropolitan area	Region	Branch of BLS
chi	Chicago-Gary-Lake County	MW	Chicago
ny	New York-Northern N.J.-Long Island	NE	New York
phil	Philadelphia-Wilmington-Trenton	NE	Philadelphia
la	Los Angeles-Anaheim-Riverside	W	San Francisco

The indices - named CPI-U by the BLS - cover all urban consumers, which represents around 87 percent of the population. When using official indices covering long periods, problems may arise with respect to the measurement. This is due to the fact that the base year of the indices from time to time is changed for ease of interpretation of the later observations. For the oldest data considered here, changes of base year has occurred four times. First, with 1947-49=100, then 1957-59=100, 1967=100, and latest 1982-84=100. The indices used in the present analysis have 1967 as base year. Since the indices are published

[17] http://stats.bls.gov/.

with only one decimal, the accurateness of the changes in data in the beginning of the period - where the index number is relatively low - are less precise than for the later period. For example, the index from the Chicago area starts in January 1953 at 79.8 and ends in December 1997 at 486.5. An index change of 0.1 in the early period implies a monthly inflation of 0.125 per cent whereas the same change in the late period requires the prices to change only by 0.021 per cent over the month.[18] Logarithm of the data are illustrated in levels and differences in Figure 3.1.[19]

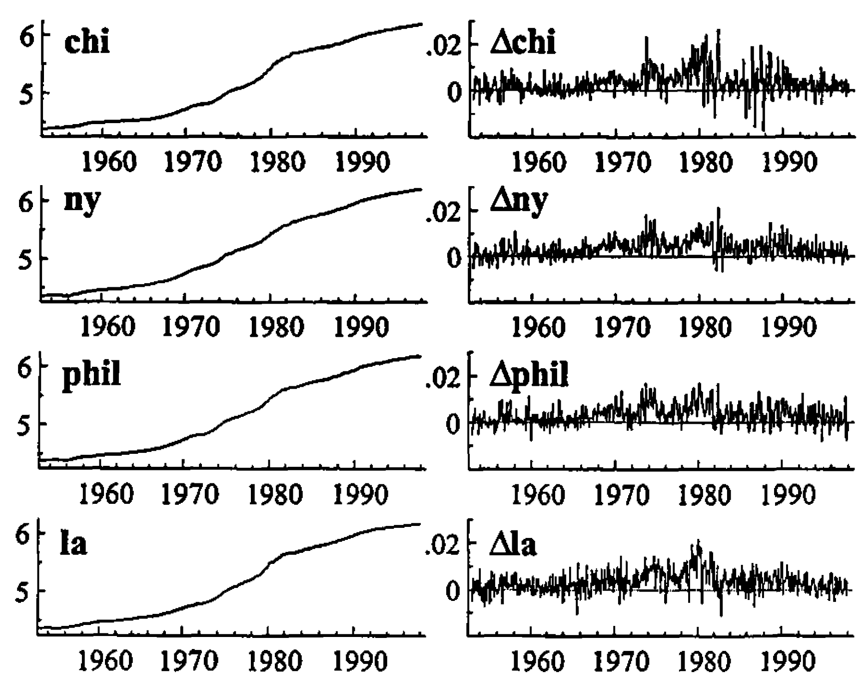


Figure 3.1. Logarithm of price indices in levels and differences.

So, what information can we get from an analysis of these indices? The indices contain information of prices for more than 2,000 articles. These are arranged in eight major groups: Food and Beverages; Housing; Apparel; Transportation; Medical Care; Recreation; Education

[18]The estimations were also done using data with base year 1982-84=100. The outcome was more or less the same as the one reported in the next subsection, suggesting that the choice of base year might not be of great importance.

[19]Figures 3.1 and 3.2 were made using GiveWin and PcFiml (see Doornik and Hendry (1996, 1997)).

and Communication; and Other Goods and Services.[20] Hence, they contain traded as well as non-traded goods. Furthermore, they include taxes which are directly related to the purchase such as sales and excise taxes. The indices, therefore, measure the price of a basket of consumption chosen by the average consumer. The baskets are, however, not the same in the areas according to Table 3.2, which means that a direct comparison of the indices should be made with some caution.

Table 3.2. Weights of components in the CPI-U, 1997

	chi	*ny*	*phil*	*la*
Food and beverages	15.283	16.493	17.199	16.213
Housing	40.227	43.336	39.446	43.141
Apparel	5.331	5.295	5.353	4.667
Transportation	17.959	14.127	16.450	17.021
Medical care	5.082	5.218	5.359	4.011
Recreation	5.648	5.234	6.339	5.731
Education and communication	5.623	6.074	5.318	5.311
Other goods and services	4.844	4.223	4.536	3.905

Source: BLS: 'Relative importance of components in the consumer price index, 1997'.
Note: The area 'ny' covers the same cities as in Table 3.1. For the other areas changes there have been made. In this Table 'chi' covers Chicago-Gray-Kenosha, 'phil' covers Philadelphia-Wilmington-Atlantic City, and 'la' covers Los Angeles-Riverside-Orange County.

The numbers in Table 3.2 are for a specific year, namely 1997. The BLS conducts consumer surveys every year and the weights are updated accordingly. As appears from the table the weights are a bit different between the areas. Since the baskets mirror the consumers' choices between articles, which are made for given prices, the development in the indices does describe some kind of purchasing behavior which can be analyzed comparing the indices. One way to think about it is the following: If a consumer moves from one area to another, her/his consumption (i.e. the personal basket) is likely to change. For example, if a consumer moves from Chicago to New York the percentage of the entire consumption which is spent on Housing is likely to increase. Thus in this setting, testing PPP using consumer price data also tells us something about the degree to which consumers are mobile in order to equalize prices.

[20] For a more detailed description of the indices the reader is refered to BLS' homepage. Especially a look at "Frequently Asked Questions" is worthwhile.

Based on prior testing, the cointegrated VAR considered includes a constant term restricted to the α-space. Hence the ECM reads

$$\Delta^2 X_t = \alpha(\beta', \rho_1)\begin{pmatrix} X_{t-1} \\ 1 \end{pmatrix} - \Gamma\Delta X_{t-1} + \sum_{i=1}^{k-2} \Psi_i \Delta^2 X_{t-i} + \varepsilon_t.$$

The econometric methods applied rely on the assumption that errors are *iid*. Recent research (Hansen and Rahbek, 1999), however, shows that the methods are robust to ARCH effects in the residuals. The main concern is whether there is evidence of autocorrelation in the errors. Investigation of the residuals indicates that a VAR model of order three ($k = 3$)[21] including four balanced impulse dummies and 12 centered seasonal dummies is sufficient to restore residuals with no autocorrelation. Lagrange-multiplier tests for non-autocorrelation of order one and four are accepted with test statistics $\chi^2(16) = 26.6$ ($p - value = 0.05$) and $\chi^2(16) = 16.6$ ($p - value = 0.41$).[22]

The four (balanced) impulse dummies are $d559$ (+1 in 1955:9; -1 in 1955:10; and 0 otherwise), $d658$ (+1 in 1965:8; -1 in 1965:9; and 0 otherwise), $d802$ (+1 in 1980:2; -1 in 1980:3; and 0 otherwise), and $d8710$ (+1 in 1987:10; -1 in 1987:11; and 0 otherwise).

4.2. Statistical analysis. The test for rank in the $I(2)$ model is a joint test for the number of cointegration vectors and the number of common stochastic trends. The low power of the test, however, has been demonstrated by Jørgensen (1998) and Johansen (2000). Hence, the final determination of the number of common $I(1)$ and $I(2)$ trends should not be based solely on the outcome of this test. A look at the number of unit roots in the companion matrix can help to provide insight into the properties of the process. These are reported in Table 3.3.

Table 3.3. Five largest roots of the companion matrix

Unrestricted	$r = 3$	$r = 2$	$r = 1$
1.00	1	1	1
0.99	1.00	1	1
0.97	0.96	1.00	1
0.93	0.92	0.92	0.99
0.63	0.63	0.63	0.63

[21] The choice of three lags was supported by Hannah-Quinn information criteria.

[22] All estimating and testing were performed using the program package CATS in RATS (see Hansen and Juselius, 1995). I(2) tests were performed using a routine developed by Clara Jørgensen.

In the unrestricted model the four largest roots are quite large (more than 0.9) whereas the fifth one is somewhat smaller. In fact, the largest root is unity. Restricting the model to three cointegration vectors raises the second root to one and the same is the case for $r = 2$. For $r = 1$ the fourth root is very close to one. In general, when restricting the number of unit roots in the process the following root raises towards unity which indicates $I(2)$ness as an $I(2)$ trend implies two unit roots. The fourth root seems to be quite stable, except for the case where it is the next root to the unit roots, which could suggest three unit roots and a root close to but less than one. Three unit roots are consistent with three $I(1)$ trends or one $I(2)$ trend and one $I(1)$ trend. From an economic point of view it might be reasonable to think that the indices might share a common nominal trend. The rank tests for the number of cointegration vectors and the number of common trends are given in Table 3.4.

Table 3.4. Rank test for the joint hypothesis $Q(s_1, r)$

$p-r$	r	$Q(s_1,r)$				$Q(r)$
4	0	880.0	599.9	355.8	142.3	115.2
		111.6	*90.3*	*72.7*	*59.5*	*49.9*
3	1		507.5	263.1	48.5	44.5
			67.0	*51.4*	*40.2*	*31.9*
2	2			198.9	**12.2**	**12.0**
				33.2	*23.6*	*17.8*
1	3				3.8	2.8
					11.1	*7.5*
$s_2 = p - r - s_1$		4	3	2	1	0

Note: Numbers in italics are 90% quantiles from Paruolo (1996) Table A1 and Johansen (1996) Table 15.2.

Applying the technique of Pantula (1989) of starting from the most restricted hypothesis, the first hypothesis to be accepted is the case with two cointegration vectors ($r = 2$); one $I(2)$ trend ($s_2 = p-r-s_1 = 1$) and one $I(1)$ trend ($s_1 = 1$). This is consistent with three unit roots, and will be the choice for the continuation of the analysis. Hence, the model is restricted to $(r, s_1, s_2) = (2, 1, 1)$.

To gain further insight into the $I(2)$ part of the process, the estimates of the loadings for the $I(2)$ trend: $\widetilde{\beta}_{\perp 2} = \beta_{\perp 2}(\alpha'_{\perp 2}\theta\beta_{\perp 2})^{-1}$ and the coefficient for the trend: $\alpha_{\perp 2}$ are considered. The estimates are given in Table 3.5.

Table 3.5. Loadings and coefficients for the $I(2)$ trend

Variable	$\tilde{\hat{\beta}}_{\perp 2}$	$\hat{\alpha}_{\perp 2}$
chi	1	-0.1187
ny	0.9948	0.2215
phil	0.9845	-0.1645
la	0.9806	0.0173

Note: The estimates are normalized on $\tilde{\hat{\beta}}_{\perp 2,1}$.

It is striking to note that the loadings into the variables are very similar. This is in line with economic intuition that the prices should share the same nominal trend and hence that the underlying inflation is the same in the areas (Definition 1). Indeed this indicates that it should be possible to transform the model to the $I(1)$ space by imposing restrictions of price homogeneity between the prices. The test for $sp(\beta_{\perp 2}) = sp(1,1,1,1)$ or alternatively that

$$sp(\beta, \beta_{\perp 1}) = sp \begin{pmatrix} 1 & 1 & 1 \\ -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

is $\chi^2(1)$ distributed, and with a test statistic very close to zero the hypothesis is strongly accepted.[23]

The unit root consistent with price homogeneity between all prices is imposed and we consider a system with the data vector $\widetilde{X}_t = [p_1, p_2, p_3, \Delta chi]_t$, where $p_1 = chi - ny$; $p_2 = chi - phil$; and $p_3 = chi - la$. This should give approximately the same likelihood function and we should lose no information about the price processes (see, for example, Juselius, 1996). To check if the errors in the transformed model are still *iid*, misspecification tests for autocorrelation of order one and four are performed. Both tests for non-autocorrelation are accepted with test statistics $\chi^2(16) = 22.9$ ($p - value = 0.12$) and $\chi^2(16) = 17.9$ ($p - value = 0.33$).

To perform a simple check on whether the $I(2)$ trend has indeed been removed, rank tests are performed on the transformed data set. The results are given in Table 3.6.

[23]The hypothesis was tested using Clara Jørgensen's $I(2)$ program with Hans Christian Kongsted's extensions. See Kongsted (1998).

Table 3.6. Rank test for the joint hypothesis $Q(s_1, r)$

$p-r$	r	$Q(s_1, r)$				$Q(r)$
4	0	1334.3	824.2	554.0	325.1	116.2
		111.6	*90.3*	*72.7*	*59.5*	*49.9*
3	1		739.9	458.7	237.9	**30.9**
			67.0	*51.4*	*40.2*	*31.9*
2	2			356.5	131.3	**8.3**
				33.2	*23.6*	*17.8*
1	3				126.5	**2.9**
					11.1	*7.5*
$s_2 = p - r - s_1$		4	3	2	1	0

Note: See Table 3.4.

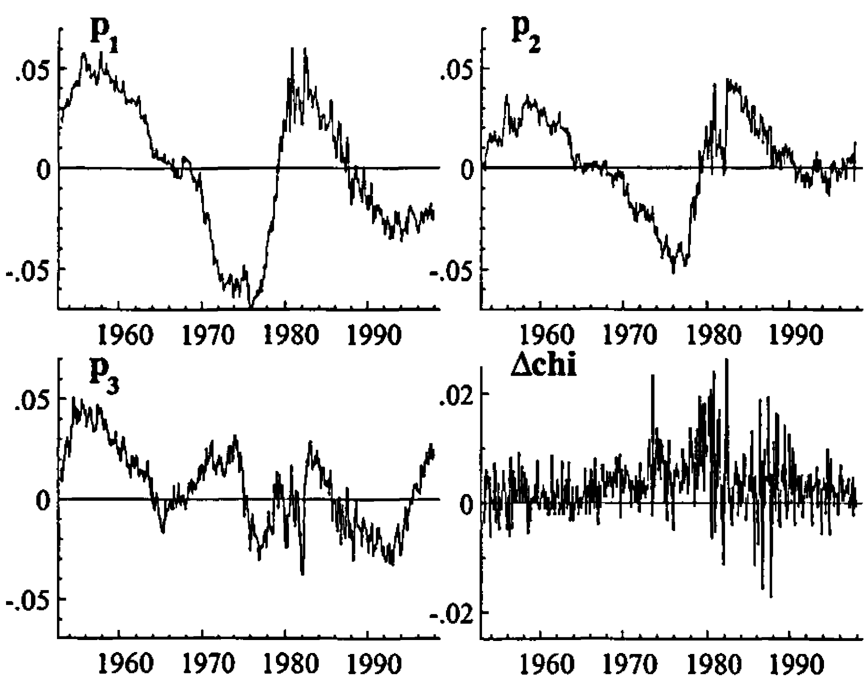


Figure 3.2. Transformed data.

Evidence from the tests reported in Table 3.6 suggests that indeed the $I(2)$ trend has been removed. The Trace test indicates $r = 1$ cointegration relation but with $r = 2$ as a borderline case. The four largest roots in the companion matrix in the unrestricted model are all real: 0.99; 0.98; 0.96; 0.68. The question is whether or not 0.96 is a unit root. Consistent with the results from the $I(2)$ analysis $r = 2$ is chosen for the further analysis. This choice seems to be appropriate, also given the evidence from the sensitivity analysis reported later. The data in the transformed system are illustrated in Figure 3.2. Notice that the transformation could be made using any other area as numeraire. In this case the inflation rate from the numeraire area would be included in the system. This was tried and did not alter the results to a great extent as reported in the sensitivity analysis in section 4.3.

To identify the cointegrating space, hypotheses concerning Definitions 2 and 3 as well as hypotheses regarding long-run exclusion, stationarity and adjustment to long-run relations are tested. The first four hypotheses tested concern exclusion from the cointegrating space. These are reported in Table 3.7.

Table 3.7. Tests for exclusion from stationary relations

	p_1	p_2	p_3	Δchi	$const.$			p-val.
$\mathcal{H}_1:$ (	0 0	* *	* *	* *	* *	) $\in sp(\beta)$	$\chi^2(2)$=15.56	[0.00]
$\mathcal{H}_2:$ (	* *	0 0	* *	* *	* *	) $\in sp(\beta)$	$\chi^2(2)$=17.88	[0.00]
$\mathcal{H}_3:$ (	* *	* *	0 0	* *	* *	) $\in sp(\beta)$	$\chi^2(2)$=5.85	[0.05]
$\mathcal{H}_4:$ (	* *	* *	* *	0 0	* *	) $\in sp(\beta)$	$\chi^2(2)$=79.32	[0.00]

It appears that p_3, i.e. the price differential between the Chicago and the Los Angeles areas, are not significant in the long-run relations. The hypothesis of exclusion of the constant was also tested and rejected with test statistic $\chi^2(2) = 57.02$ ($p - value = 0.00$).

Tests of hypotheses regarding the PPP (Definition 2) and (constant) stationarity of the inflation rate are given in Table 3.8. The hypotheses $\mathcal{H}_8 - \mathcal{H}_{10}$ are hypotheses about the PPP between areas not including Chicago. All of the hypotheses are rejected.

Table 3.8. Tests for the PPP and stationary inflation rate

	p_1	p_2	p_3	Δchi	$const.$			p-val.
$\mathcal{H}_5$: (	1	0	0	0	*	) $\in sp(\beta)$	$\chi^2(2) = 19.41$	[0.00]
$\mathcal{H}_6$: (	0	1	0	0	*	) $\in sp(\beta)$	$\chi^2(2) = 16.68$	[0.00]
$\mathcal{H}_7$: (	0	0	1	0	*	) $\in sp(\beta)$	$\chi^2(2) = 15.50$	[0.00]
$\mathcal{H}_8$: (	1	-1	0	0	*	) $\in sp(\beta)$	$\chi^2(2) = 16.71$	[0.00]
$\mathcal{H}_9$: (	1	0	-1	0	*	) $\in sp(\beta)$	$\chi^2(2) = 19.06$	[0.00]
$\mathcal{H}_{10}$: (	0	1	-1	0	*	) $\in sp(\beta)$	$\chi^2(2) = 18.06$	[0.00]
$\mathcal{H}_{11}$: (	0	0	0	1	*	) $\in sp(\beta)$	$\chi^2(2) = 6.24$	[0.04]

Hypotheses regarding the PPP with adjustment (Definition 3) are given in Table 3.9. These are tested with and without a constant in the relations. The presence of a constant could imply that the indices start at different levels. Note, however, that even though the indices start at the same level convergence can take place if the "true" price levels are different.

Table 3.9. Tests for the PPP with adjustment

	p_1	p_2	p_3	Δchi	$const.$			p-val.
$\mathcal{H}_{12}$: (	1	0	0	*	0	) $\in sp(\beta)$	$\chi^2(2) = 19.46$	[0.00]
$\mathcal{H}_{12}^c$: (	1	0	0	*	*	) $\in sp(\beta)$	$\chi^2(1) = 0.11$	[0.73]
$\mathcal{H}_{13}$: (	0	1	0	*	0	) $\in sp(\beta)$	$\chi^2(2) = 16.75$	[0.00]
$\mathcal{H}_{13}^c$: (	0	1	0	*	*	) $\in sp(\beta)$	$\chi^2(1) = 0.95$	[0.33]
$\mathcal{H}_{14}$: (	0	0	1	*	0	) $\in sp(\beta)$	$\chi^2(2) = 15.27$	[0.00]
$\mathcal{H}_{14}^c$: (	0	0	1	*	*	) $\in sp(\beta)$	$\chi^2(1) = 4.74$	[0.03]
$\mathcal{H}_{15}$: (	1	-1	0	*	0	) $\in sp(\beta)$	$\chi^2(2) = 14.47$	[0.00]
$\mathcal{H}_{15}^c$: (	1	-1	0	*	*	) $\in sp(\beta)$	$\chi^2(1) = 0.09$	[0.76]
$\mathcal{H}_{16}$: (	1	0	-1	*	0	) $\in sp(\beta)$	$\chi^2(2) = 17.85$	[0.00]
$\mathcal{H}_{16}^c$: (	1	0	-1	*	*	) $\in sp(\beta)$	$\chi^2(1) = 2.43$	[0.12]
$\mathcal{H}_{17}$: (	0	1	-1	*	0	) $\in sp(\beta)$	$\chi^2(2) = 18.02$	[0.00]
$\mathcal{H}_{17}^c$: (	0	1	-1	*	*	) $\in sp(\beta)$	$\chi^2(1) = 5.68$	[0.02]

Hypotheses of the PPP with adjustment are accepted between the areas $chi - ny$, $chi - phil$, $ny - phil$ and $ny - la$ when allowing for a constant in the stationary relation. The hypotheses are rejected for $chi - la$ and $phil - la$. All in all this indicates that when testing the PPP distance matters which, is in line with economic intuition and the findings in the studies of LOP mentioned in the introduction of this chapter. Furthermore, adjustment of price levels seems to have taken place between the areas of Chicago, New York and Philadelphia but not with the level in the Los Angeles area. There does, however, seem to be some adjustment between prices in New York and Los Angeles, but it should be kept in mind that p_3 is not significant in the stationary relations.

Finally, it is tested whether the variables adjust to the long-run relations. This seem to be the case for all of them as can be seen in Table 3.10.

Table 3.10. Tests for adjustment to long-run relations

	Δp_1	Δp_2	Δp_3	$\Delta^2 chi$			p-val.
$\mathcal{H}_{18}:$ (	0 0	* *	* *	* *	) $\in sp(\alpha)$	$\chi^2(2) = 11.86$	[0.00]
$\mathcal{H}_{19}:$ (	* *	0 0	* *	* *	) $\in sp(\alpha)$	$\chi^2(2) = 23.79$	[0.00]
$\mathcal{H}_{20}:$ (	* *	* *	0 0	* *	) $\in sp(\alpha)$	$\chi^2(2) = 10.49$	[0.00]
$\mathcal{H}_{21}:$ (	* *	* *	* *	0 0	) $\in sp(\alpha)$	$\chi^2(2) = 53.59$	[0.00]

The complete cointegration space is identified as the joint hypothesis $\{\mathcal{H}_{12}^c, \mathcal{H}_{13}^c\}$, which in turn coincides with $\mathcal{H}_3$ since it is just a matter of normalization. The identified space implies the following relations:[24]

$$\begin{aligned} chi &= ny - \underset{(4.390)}{41.212}\Delta chi + \underset{(0.019)}{0.144} \\ chi &= phil - \underset{(2.256)}{22.396}\Delta chi + \underset{(0.010)}{0.082}. \end{aligned} \tag{4.1}$$

4.3. Sensitivity analyses[25]. When making empirical analyses using econometric models, one always has to make some choices with respect to the model specification. Tests can help but it is rarely the case that it is absolutely clear how the model should be specified. This section contains a discussion of the sensitivity of the results regarding the numbers of cointegration relations and common $I(2)$ trends with respect to inclusion of the impulse dummies, the chosen lag length, the sample period and the transformation of the system with the Chicago area as numeraire. Also the sensitivity with respect to exclusion of p_3 in the transformed model - given $r = 2$ - is discussed.

First, however, a brief remark regarding the deterministic term. Prior to the analysis reported in this chapter, the same experiments

[24]Numbers in brackets are standard errors.

[25]A more thoroughly investigation of the stability of the cointegrating vectors is given in Appendix B.

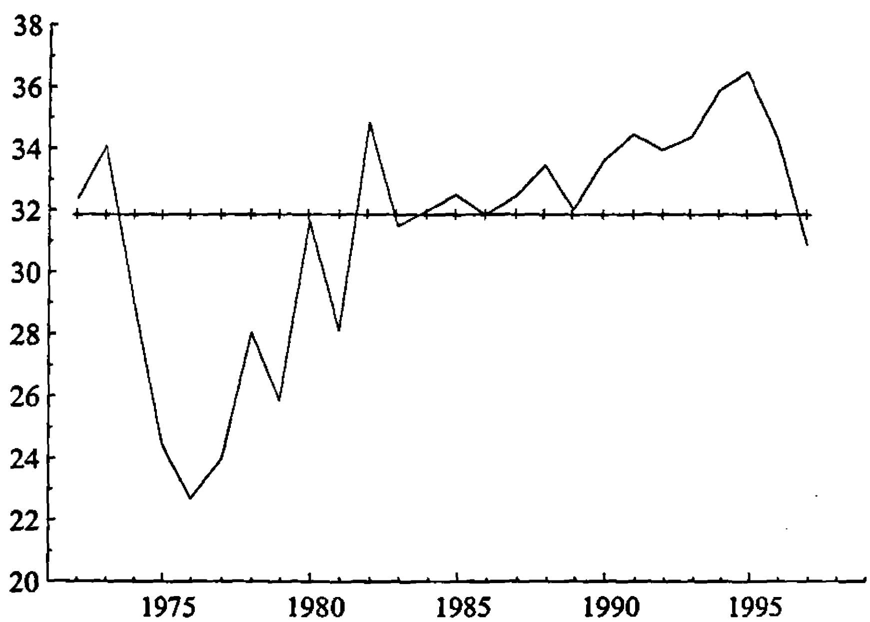
were conducted using a model allowing for a linear trend in the cointegrating relations and an unrestricted constant. Evidence was found in favor of one common $I(2)$ trend, one $I(1)$ trend and two cointegration relations. A transformation of the system similar to the one in the present analysis eliminated the $I(2)$ trend and the Trace statistics for the transformed system revealed clear evidence of two cointegrating relations. Furthermore, a χ^2 test strongly accepted the exclusion of p_3. The two identified stationary relation were similar to (4.1) and the trend was not significant in any of them. A test for restricting the constant to the β-space was strongly accepted with $p-value = 0.35$.

Dummies. It should be noted that without the inclusion of the four dummies in the system the Lagrange-Multiplier tests for non-autocorrelation of order one were rejected in the original as well as in the transformed system. Hence, it is doubtful whether the residuals are in fact *iid*. Having mentioned this, the inclusion of dummies does not change anything with respect to the conclusions. The first hypothesis accepted in the $I(2)$ case is for $(r, s_1, s_2) = (2, 1, 1)$. In the transformed system the Trace statistic for the hypothesis of more than one cointegration relation rises slightly but is still a borderline case. The hypothesis of exclusion of p_3 is accepted with $p - value = 0.05$.

Lags. The model was estimated with $k = 4, 5, 6$ lags and the relevant hypotheses were tested. Also with respect to the number of lags the results seem to be quite robust. In all cases the $I(2)$ analyses provided evidence of $(r, s_1, s_2) = (2, 1, 1)$. The Trace tests in the $I(1)$ model seem to decrease the more lags were included, and the relevant statistics were still borderline cases. For $k = 4$, the hypothesis for excluding p_3 from the stationary relations was rejected with $P\text{--}value =$ 0.03 but was accepted for $k = 5, 6$.

The choice of *chi* as numeraire. The model was estimated three times using each of the areas as transformation variable. In all cases rank tests indicated that the transformations removed the $I(2)$ness from the system. When using *ny* and *phil* as numeraire, the results were quite robust, i.e. the trace tests indicated that $r = 2$ was a borderline case and in both cases the variable including *la* could be tested out of the cointegration space. In the case where *la* was used as numeraire the test for $r = 2$ was clearly accepted. None of the variables were insignificant, which indicate that when making a transformation from the $I(2)$ to the $I(1)$ space similar to the one made in this analysis, one should carefully choose which variable to use as numeraire.

Sample period. The final sensitivity analysis performed here is with respect to the sample period. The analysis is made by fixing an initial period of 20 years and then adding one year of observations recursively. Hence, 25 point estimates were made. Furthermore, the model was estimated for the last 26 years of the sample only, i.e. from 1972 to 1997. Rank tests for the joint hypothesis of the number of cointegration relations and the number of $I(2)$ trends are performed successively following the same principle as for the $I(1)$ case of Hansen and Johansen (1999). In all cases but two - the periods ending in 1979 and 1980 - the rank tests indicated that one common $I(2)$ trend is present in the data. In the latter two cases the test statistics were borderline cases. The recursively performed Trace tests for the hypothesis of $r = 2$ in the transformed model are given in Figure 3.3.



Note: The "plus"-line indicates 90% critical value.

Figure 3.3. Recursive Trace tests for the hypothesis of $r = 2$.

In 14 cases the outcome of the Trace test supports the choice of two cointegration relations, and especially when the latter period is included in the sample the test seems supportive. This indicates that the choice of $r = 2$ in the transformed model is indeed appropriate, although it should be mentioned that the Trace statistic when estimating the model for the last 26 years is a borderline case with a value a bit lower than the critical. With respect to exclusion of p_3 the hypothesis is accepted in 17 out of the 25 cases in the recursive analysis and also when estimating the model only for the latter period.

All in all the results obtained in this chapter seem to be quite robust with respect to the specification of the model and the chosen sample period.

5. Conclusion

Does the PPP hold within the US? The evidence from the analysis in this paper indicates that when allowing for adjustment of price levels, it does hold between areas which are geographically close to each other but not between areas far apart.

The empirical analysis is made on the basis of three introduced definitions. Whereas one of them - although not formally defined - has been applied in more empirical studies the other two are new (to the best of my knowledge) in the literature. Definition 1 introduces the concept of sharing underlying inflation. This is the case if price indices share a common $I(2)$ trend with the same impact. Definition 3 allows for price levels to adjust and is related to the concept of multi-cointegration in the cointegrated VAR model for $I(2)$ variables. It can be interpreted in terms of optimizing agents (policy makers) seeking integration of markets.

The price indices analyzed here seem to share one common $I(2)$ trend with the same impact on all variables. Hence, it is concluded that the underlying inflation has been the same in the areas of Chicago, New York, Philadelphia and Los Angeles. Furthermore, evidence is found in favor of two cointegration vectors. These can be identified as relations between the areas of Chicago and New York and between Chicago and Philadelphia. This suggests that price levels in these areas have adjusted to each other whereas the price level in the Los Angeles area seems to be more independent from the others. Hence, in accordance with other studies of the LOP and PPP, the evidence is that distance matters. A comprehensive sensitivity analysis reveals that the results are quite robust with respect to the specification of the statistical model.

The present analysis can be considered a benchmark case for what can be expected to happen with respect to PPP within the euro area. It suggests that there is reason to believe that the underlying inflation might be the same in the long run. On the other hand, there is no reason to expect that the PPP will hold between all the countries even though price levels can be expected to adjust between countries close to each other. Whether structural differences between the US and the euro area will lead to other conclusions is an issue which will be left for future research.

6. Appendix A

CLAIM 1. *Let* $p_t^* = \theta x_t + e_t$ *with* $E_t(e_t) = e_t$, $E_t(e_{t+h}) = 0$ $(h = 1, ..., \infty)$, *and* $E_t(x_{t+h}) = x_t + h\Delta x_t$, *for* $h = 0, 1, ..., \infty$. *Then it holds that* $E_t \sum_{s=t}^{\infty} (\beta\lambda)^{s-t} p_s^* = \theta/(1-\beta\lambda) * x_t + e_t + \theta\beta\lambda/(1-\beta\lambda)^2 * \Delta x_t$

PROOF. Using the formula in footnote 6 we find that

$$E_t \sum_{s=t}^{\infty} (\beta\lambda)^{s-t} p_s^* = e_t + \frac{\theta}{1-\beta\lambda} x_t + \theta\beta\lambda\Delta x_t + 2\theta(\beta\lambda)^2 \Delta x_t + ...$$

looking apart from the first two terms and applying the formula in footnote 6 on the remaining terms we find

$$\left(\frac{\theta\beta\lambda}{1-\beta\lambda} + \frac{\theta(\beta\lambda)^2}{1-\beta\lambda} + \frac{\theta(\beta\lambda)^3}{1-\beta\lambda} + ...\right)\Delta x_t$$
$$= \frac{\theta}{1-\beta\lambda}\left(\frac{1}{1-\beta\lambda)} - 1\right)\Delta x_t = \frac{\theta\beta\lambda}{(1-\beta\lambda)^2}\Delta x_t.$$

∎

7. Appendix B

This appendix expands the sensitivity analysis in section 4.3. It is investigated if the cointegrating vectors found in (4.1) are stable.

To investigate whether β is stable over the whole sample, the estimated cointegrated space is fixed, $\widehat{\beta}_T$, and it is tested recursively if $sp(\beta) = sp(\widehat{\beta}_T)$ for the first t observations. The period covering up to 1977 is used to obtain the initial estimate. The results are given in Figure 3.4 where the test statistics are scaled by the 10% critical values. The line BZ represents the test statistics when the short-run parameters are re-estimated and BR are test statistics when short-run parameters are not re-estimated.

Indeed, the evidence from Figure 3.4 suggests that the cointegrated space given in (4.1) is not included in $sp(\beta)$ when estimating over the early subsamples. A natural question then is if it is possible to find vectors which are stable. Investigation of the cointegrating vectors for different subsamples indicates that the coefficient for Δcpi and the constant may be too large in (4.1). Several tests were made and it turned out that the vectors arising when estimating the model for the period 1953:2-1986:6 could be considered stable when applying the test procedure described above cf. Figure 3.5.

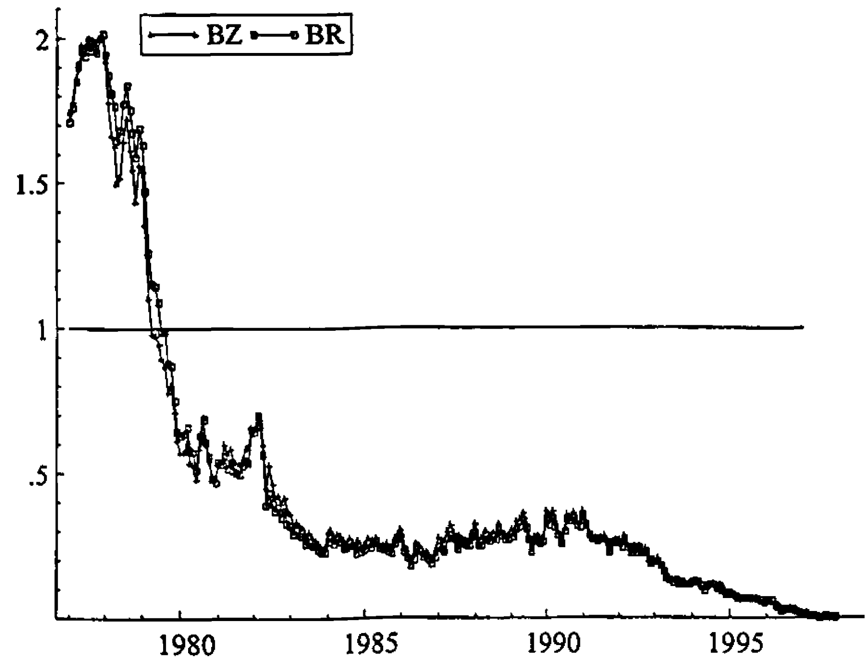


Figure 3.4. Tests for β constancy.

The relations implied are

$$\begin{aligned} chi &= ny - 26.990 \Delta chi + 0.094 \\ chi &= phil - 15.884 \Delta chi + 0.054, \end{aligned}$$

and the hypothesis that this space span the same space as β_T was accepted with test statistic $\chi^2(6) = 12.59$ ($p - value = 0.05$).

However, making estimations on different subsamples revealed that the observations from 1965 to 1968 have great influence on the parameter estimates. In Table 3.11 the estimates of the coefficients for Δchi given $r = 2$ are reported for different subsamples. All samples start in 1953:2. Especially something strange happens with the observations 1966:7 and 1966:8 where the coefficients even change the signs.

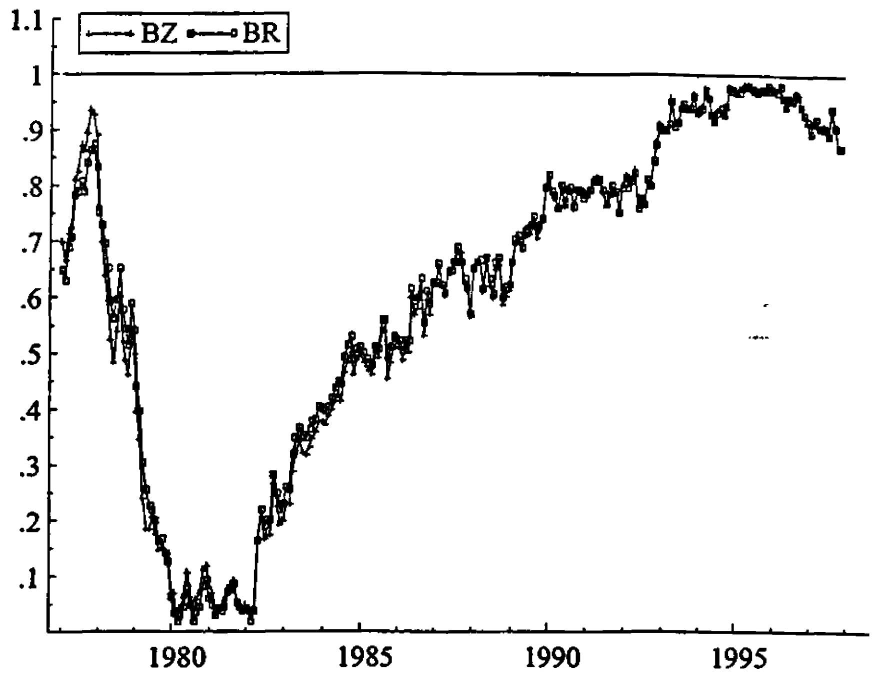


Figure 3.5. Tests for β constancy.

Table 3.11. Coefficients for Δchi

End date	Coeff. in CI1	Coeff. in CI2
63:12	−45.1 (7.8)	−20.0 (3.8)
64:12	−73.7 (11.9)	−45.6 (7.7)
65:12	−219.8 (35.2)	−131.1 (21.3)
66:7	−198.9 (28.2)	−116.8 (16.7)
66:8	303.3 (43.6)	176.6 (25.2)
66:12	228.6 (34.2)	130.2 (19.3)
67:12	78.1 (11.9)	42.7 (6.4)
68:12	48.9 (7.6)	28.1 (4.2)

Note: Numbers in brackets are standard deviations.

The period investigated in Table 3.11 was an economic turbulent one for the US economy with expansive fiscal policy, high budget deficits, high growth of money not to mentioned the war in Vietnam.

This seems to have affected the relative prices in the US to a high degree. Another turbulent period followed with the Smithsonian Agreement in 1971 which collapsed in 1973 and the first oil price shock. It is quite puzzling how one observation (1966:8) can have so great influence on the stationary relations. It might have to do with the normalization of the vectors but different normalizations were tried and it did not solve the problem. In the future it would be interesting to investigate this particular problem further.

It turned out that one can obtain a model with stable parameters if the period starts in 1976. The results of the empirical analysis for the period 1976-1997 are reported below.

Including an extra lag in the model for this period improved the test statistics for non-autocorrelation significantly. Hence, we consider a VAR(4) model with a constant restricted to the cointegrating space and 11 unrestricted seasonal dummies. The tests for non-autocorrelation of order one and four were accepted with p-values 0.32 and 0.38. The joint rank tests are given in Table 3.12, indicating $(r, s_1, s_2) = (2, 1, 1)$.

Table 3.12. Rank test for the joint hypothesis $Q(s_1, r)$

$p-r$	r	$Q(s_1,r)$				$Q(r)$
4	0	436.9	296.0	171.9	104.7	96.5
		111.6	*90.3*	*72.7*	*59.5*	*49.9*
3	1		263.7	135.6	44.0	43.0
			67.0	*51.4*	*40.2*	*31.9*
2	2			113.1	**20.3**	19.2
				33.2	*23.6*	*17.8*
1	3				28.8	**5.3**
					11.1	*7.5*
$s_2 = p - r - s_1$		4	3	2	1	0

Note: Numbers in italics are 90% quantiles from Paruolo (1996) Table A1 and Johansen (1996) Table 15.2.

The loadings from the common $I(2)$ trend normalized on the first entry are given by

$$\widehat{\widehat{\beta}}_{\perp 2} = (1; 0.96; 0.98; 1.00)'.$$

With the transformed data vector $\widetilde{X}_t = (p_1, p_2, p_3, \Delta chi)_t = (chi - ny, chi - phil, chi - la, \Delta chi)_t$ the joint rank tests reject presence of $I(2)$ness and support the hypothesis of $r = 2$. The hypothesis of exclusion of p_3 from the cointegration space is clearly accepted with p-value

0.43 and the identified space implies the relations

$$(7.1)\qquad \begin{aligned} chi &= ny + \underset{(3.34)}{25.83}\Delta chi - \underset{(0.02)}{0.10} \\ chi &= phil + \underset{(2.22)}{13.57}\Delta chi - \underset{(0.01)}{0.05}. \end{aligned}$$

Note that the signs for Δchi and the constant in (7.1) are opposite from those in (4.1). With this sample the cointegrating space seems to be stable cf. Figure 3.6.

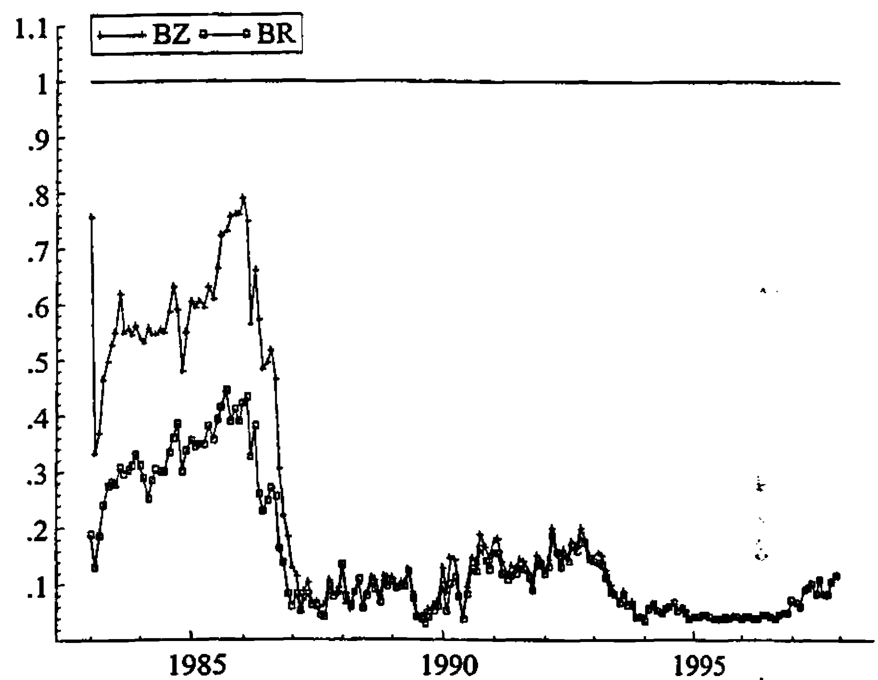


Figure 3.6. Tests for β constancy.

CHAPTER 4

PPP, cointegration, and non-stationary inflation

Abstract: It is demonstrated how to take advantage of the complex structure of the cointegrated VAR model for $I(2)$ variable in order to gain more information on relationships between prices. This is done in a comprehensive scenario analysis. An empirical example provides evidence that the underlying inflation is the same in UK and US but PPP does not hold. However, price levels have been adjusting.

1. Introduction

In the present analysis it is discussed how to test the Purchasing Power Parity (PPP) in the cointegrated vector autoregressive (VAR) model when inflation rates are non-stationary. Three definitions generalizing those made in chapter 3, are given in order to exploit the statistical properties of data in a way which can be interpreted in terms of the PPP. As the cointegrating properties between two or more variables are maintained when expanding the system, only systems with the variables in the PPP are considered in order to keep the analysis as simple as possible.

The literature of PPP has been developing for many decades. Economists have argued that, given standard assumptions, the theory states that prices, adjusted for the exchange rate, will be equal for similar goods, no matter where these are traded. Econometricians have applied statistical tools to find empirical evidence for or against the PPP, and in general the results have not been in favor. It has then been acknowledged that the PPP is a feature most likely to hold in the long run because, among other things, it takes time for firms and consumer to get information about changes in prices. But even as a long run feature it has been hard to find empirical evidence in favor of the PPP. This has led to one of the many puzzles in the literature of international economics, the so-called Purchasing Power Parity Puzzle (See Rogoff, 1996).

In many of the empirical studies testing the PPP it has been assumed that the inflation rates are stationary. That this may not be the case has been demonstrated in several recent empirical analyses including Juselius (1999) and the one in chapter 3 for the case of prices in the US. In order to apply the cointegrated VAR model for variables integrated of order one, some analyses include the real exchange rate as one variable instead of three variables: the two price levels and the nominal exchange rate. This is, for example, the case in the recent studies of Juselius and MacDonald (2000a, 2000b).

It might very well be the case, however, that important information about the time series is lost when including the real exchange rate as only one variable. Since statistical tools for treating variables integrated of order two have been developed, important information about price and exchange rate processes might be revealed applying these tools. On the other hand, well-tested software for testing hypotheses in the $I(2)$ model is only available to a limited extent, which makes the practical analysis of $I(2)$ systems more complicated. In the present chapter the impacts from the common $I(2)$ trends are investigated in order to find a suitable and economically interpretable transformation of the $I(2)$ model into the $I(1)$ space without too much loss of information. Well tested statistical software for making inference in $I(1)$ models has been available for several years.

The present essay contains eight sections. In the next one, a general relation between prices and the nominal exchange rate is stated and the standard PPP relations are derived from this. It is argued that if price levels adjust to each other, the adjustment process may be interpreted in terms of the inflation rate. Definitions are given in order to interpret the statistical properties of the time series in economic terms. Section 3 contains a discussion of whether inflation rates are stationary or not. US data for the consumer price index is used to give an empirical illustration. In section 4 the cointegrated VAR models for $I(1)$ and $I(2)$ variables are briefly described and section 5 contains a comprehensive scenario analysis describing possible outcomes when inflation rates are $I(1)$ or $I(2)$. In section 6 a possible scenario for testing the PPP among three countries is given and section 7 contains an application testing the PPP between the UK and the US. Section 8 summarizes the analysis.

2. Relationship between prices

Let P_t be the price of a particular basket of goods in a country at time t and let P_t^* be the exchange rate corrected price of the same basket in another country. In general, there will exist a process H_t such that the following relation will hold:

$$\frac{P_t}{P_0} = \frac{P_t^*}{P_0^*} H_t,$$

with $H_0 = 1$. Expressed in logarithms we get

$$p_t = p_t^* + h_t + const,$$

where small letters denote logarithms and $const = p_0 - p_0^*$. The relative version of the PPP holds if $h_t = 0$ for all t and the absolute version holds if furthermore $const = 0$ i.e. $p_0 = p_0^*$. The term h_t is the deviation from the PPP.

When testing the PPP empirically often indices are used and a more flexible formulation is needed. This refers to the stochastic properties of the time series. We will allow for the prices to be integrated of order 1 or 2. Hence, it follows that $h_t \sim I(d)$, for $d = 0, 1, 2$. If $d = 0$ the PPP holds. When price levels are expressed in indices, the condition $const = 0$ will in general not have any meaningful interpretation unless the underlying prices are exactly the same at the starting point.

If the price levels in two economies are adjusting toward each other, the process h_t might be interpreted in terms of the inflation rate when this is integrated of order one. To see this, consider two economies where policy makers want to integrate the markets. Let the measurement for integration be the difference in the prices such that the markets are perfectly integrated if the prices are the same. The problem to solve is to minimize the price differential as well as the inflation. This can be formulated by means of a linear-quadratic adjustment cost (LQAC) model:

$$\min_{\{p_s\}} E_t \sum_{s=t}^{\infty} \beta^{s-t} \left[\delta \left(p_s - p_s^* \right)^2 + \Delta p_s^2 \right],$$

where $\delta > 0$ is the relative weight between the benefit of integration and the cost of inflation, $0 < \beta < 1$ is a discount factor. Let the tracking variable $p_t^* = x_t'\theta + e_t$, where x_t is a vector of forcing variables and $e_t \sim iid(0, \sigma_e^2)$. In the present analysis we assume for simplicity that x_t is a univariate process. If $x_t \sim I(2)$ it can be shown that the solution to the problem implies that[1]

$$p_t = \theta x_t - \frac{\lambda}{1-\lambda} \Delta p_t + \frac{\beta\lambda}{1-\beta\lambda} \Delta x_t + (1-\beta\lambda)(1-\lambda) e_t, \tag{2.1}$$

where λ is the stable root in the characteristic polynomial related to the Euler-equation.

[1]See chapter 3 for details.

Since e_t is stationary it is required that $p_t \sim I(2)$ and cointegrates with x_t with the coefficient θ for the relation (2.1) to hold. In other words, $p_t - \theta x_t \sim I(1)$. Then Δp_t and Δx_t will share a common $I(1)$ trend and in fact, only one of them, say Δp_t, is needed in order to reach stationarity. In this case, however, the coefficient for Δp_t will not be the same as in (2.1).

Based on this discussion three "econometric" definitions are introduced. These are extensions of the ones given in chapter 3. When the exchange rate is introduced the interpretation becomes somewhat more complicated.

Let p_t^i and p_t^j be time series for price levels in country i and j, and let e_t^{ij} be the time series for the exchange rate between the countries.

DEFINITION 4. ***A:*** *If p_t^i, p_t^j, e_t^{ij} are integrated of order 2 and $p_t^i - p_t^j - e_t^{ij} \sim I(1)$, that is $\Delta p_t^i - \Delta p_t^j - \Delta e_t^{ij} \sim I(0)$, then we will say the exchange rate adjusted underlying inflation is the same in country i and j.* ***B:*** *If p_t^i, p_t^j are integrated of order 2, e_t^{ij} is at most integrated of order 1 and $p_t^i - p_t^j \sim I(1)$, that is $\Delta p_t^i - \Delta p_t^j \sim I(0)$, then we will say that the underlying inflation is the same in country i and j.*

DEFINITION 5. *If p_t^i, p_t^j, e_t^{ij} are at most integrated of order 2 and $p_t^i - p_t^j - e_t^{ij} \sim I(0)$, then we will say the PPP holds.*

DEFINITION 6. *If p_t^i is integrated of order 2, and either p_t^j or e_t^{ij} - or possibly both - are also integrated of order 2, and $p_t^i - p_t^j - e_t^{ij} - \kappa \Delta p_t^i \sim I(0)$ for some $\kappa \neq 0$, then we will say the PPP with adjustment holds.*[2]

The concept of underlying inflation in Definition 4 relates to the stochastic variation of the time series. In the case where the underlying inflation is the same in the two countries (Definition 4B), the price levels are affected with the same impact from one common $I(2)$ trend. Hence, the stochastic variation in the prices are the same, as the stochastic $I(2)$ part of the process is dominating. When the exchange rate adjusted underlying inflation is the same (Definition 4A), we have a situation where $p_t^i \sim I(2)$, $(p_t^j + e_t^{ij}) \sim I(2)$, and the impact from the common $I(2)$ trend is the same. In other words, p_t^i and $(p_t^j + e_t^{ij})$ cointegrate one-to-one from $I(2)$ to $I(1)$.

The situation where $p_t^i \sim I(1)$, $p_t^j, e_t^{ij} \sim I(2)$, and $p_t^j + e_t^{ij} \sim I(1)$ is not covered by Definition 4 but might occur in the empirical work.

[2] If $p_t^i, p_t^j, e_t \sim I(2)$, and they share two common $I(2)$ trends the PPP with adjustment holds if $p_t^i - p_t^j - e_t - \kappa_1 \Delta p_t^i - \kappa_2 \Delta p_t^j \sim I(0)$ for either $\kappa_1 \neq 0$ or $\kappa_2 \neq 0$. This situation is described in Case 6 in the scenario analysis in section 5.2.

The reason is that the stochastic variation in the (exchange rate adjusted) prices is the not same. Only in the case where p_t^i and $(p_t^j + e_t^{ij})$ cointegrate one-to-one from $I(1)$ to $I(0)$ does it make sense to say that Definition 4A holds. But in this case also the PPP, as stated in Definition 5, will hold. In fact, if Definition 5 holds, Definition 4A will hold trivially.

Also if Definition 6 holds, Definition 4A will hold trivially. The situation is that the real exchange rate is integrated of order 1, $p_t^i - p_t^j - e_t^{ij} \sim I(1)$, and it cointegrates with the inflation rate in country i. The name "PPP with adjustment" comes from the fact that prices are allowed to adjust towards each other maybe to a sustainable level. To simplify notation let $p_t^{j*} = p_t^j - e_t^{ij}$ be the exchange rate adjusted price level in country j. It follows that $p_t^i - p_t^{j*} \sim I(1)$, which implies that the price differential could be increasing or decreasing. In case it is increasing it could be because the level in j is lower than the one in i and is increasing towards the i-level. This is the adjustment term referred to in the definition. If the price differential cointegrates with the inflation rate, the adjustment term can be interpreted in terms of the inflation rate.

Note that if the underlying inflation in country i and j is the same (Definition 4B), that is, the stochastic variation in the price levels in the two countries are the same, then the term $\kappa\Delta p_t^i$ in Definition 6 can be replaced with $\kappa\Delta p_t^j$ keeping the cointegrating properties unchanged.

2.1. A comment on the deterministic term. In Definition 4-6 above it is not stated whether the combinations are trend-stationary, stationary around a constant or stationary with mean zero. Let us first consider Definition 4. If $\Delta p_t^i - \Delta p_t^j - \Delta e_t^{ij}$ is trend-stationary it means that a deterministic trend is present in at least one of the series. This implies a quadratic trend in the level which is most unlikely for prices and exchange rates. However, as a description of one of the series for a very limited period of time, it might be the case that it is described best as varying around a deterministic trend. For longer time series a deterministic trend in the changes seems very implausible. On the other hand, it could be the case that $\Delta p_t^i - \Delta p_t^j - \Delta e_t^{ij}$ is stationary around a constant. This would be the case if prices (corrected for the exchange rate) systematically grow more in one of the areas than in the other.

Now, let us turn to Definitions 5 and 6 and for simplicity keep the exchange rate constant. If $p_t^i - p_t^j$ is trend-stationary it implies that the series have different deterministic trends. If this is the case, the series will in fact diverge and it makes little sense to say that PPP holds. In Definition 6, however, it could be the case that - as a small sample property - the series have different deterministic trends and still converge because of the adjustment term $\kappa\Delta p_t^i$. Again, however, this is a description of the processes one should be very careful about interpreting and it should definitely not be used for forecasting.

3. Is the inflation rate non-stationary?

Several empirical studies find evidence that inflation rates are not stationary. For the case of the US these studies include Juselius (1999) and the one in chapter 3. Also in a recent study Stock (2001) notes that "... inflation has been highly persistent for the past three decades, and stably so.". In this section it is investigated whether the US inflation rate has been stationary or not since 1913. Furthermore, the statistical properties of the series are investigated with respect to the frequency of the data as well as stochastically structural breaks, i.e. parameter stability.

As is well-known, a time series x_t is said to be (covariance or weakly) stationary if the following requirements are fulfilled:

$$\begin{array}{ll} E(x_t) = \mu & \forall t \\ E(x_t - \mu)(x_{t-j} - \mu) = \gamma_j & \forall t \text{ and any } j. \end{array}$$

Or in words, the expected value at any time is constant and the covariance between x_t and x_{t-j} depends on the length between the dates and not on the dates themselves. In this section we refer to weak stationarity, whereas in the VAR model applied later, the cointegrating relations are strict stationary, i.e. for any $j_1, j_2, ..., j_n$ the joint distribution $(x_t, x_{t+j_1}, ..., x_{t+j_n})$ depends only on $j_1, j_2, ..., j_n$ and not on t. If the process is strict stationary and have finite second moment it is also covariance stationary, but the other way is not necessarily true.

A non-stationary time series, in the present context, is a series integrated of order one, which implies that the first order difference is stationary. Since the inflation rate itself is a difference, non-stationarity implies that the level is integrated of order two, i.e. a second order difference is needed to obtain stationarity.

A property of non-stationary time series is that the expected value and / or the variance vary over time. In fact, the variance goes asymptotically towards infinity. Because of this, macroeconomic authors have often claimed that inflation rates cannot be non-stationary. On the other hand, many econometric authors find statistical evidence that the inflation rate often is $I(1)$. One argument for this could be that the behavior in limited samples is like an $I(1)$ process. This in turn might imply that the inflation rate has different stochastic properties at different periods of time, which can be interpreted as different regimes. This should be taken into account when testing the hypothesis of a unit root, as the test requires parameters to be constant over time.

Since the concept of stationary is a statistical one referring to the stochastic properties of a time series, it should be tested empirically if the inflation rate is stationary applying statistical tools. To gain further insight in whether stationarity depends on the length of the sample, the period of time considered, or maybe the frequency of the data examined, an empirical investigation is provided. The longest price time series with monthly observations I could find, is the US consumer prices index from the Bureau of Labor Statistics. The series is named CPI-U and data is not adjusted for seasonality. The period considered here covers 1913 to 2000, which gives a total of 1056 monthly observations. In order to obtain quarterly and annual observations, indices are calculated as the average of the monthly data. This gives 352 quarterly and 88 annual observations. The inflation rates, given the three frequencies, are illustrated in Figure 4.1.

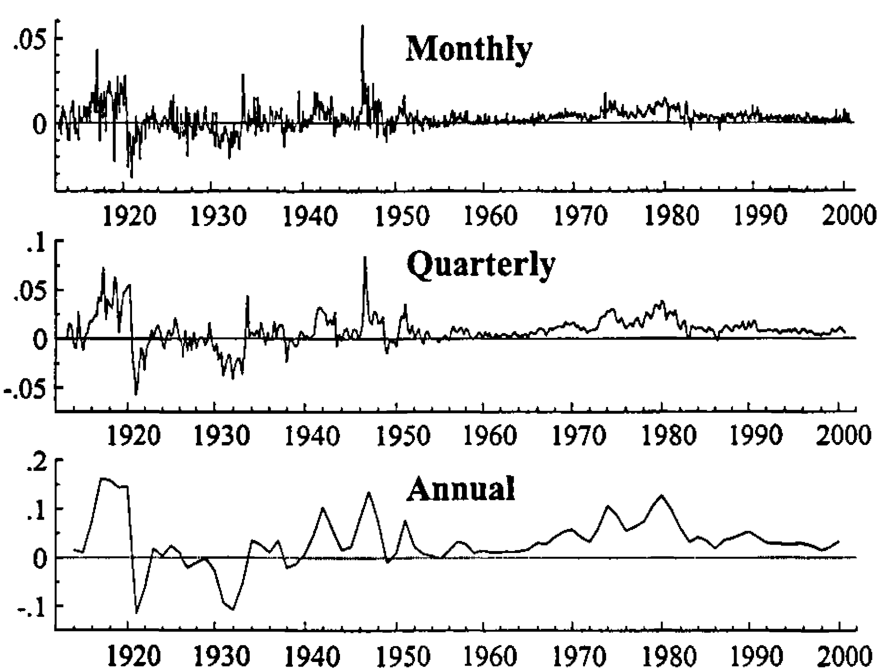


Figure 4.1. US inflation rates.

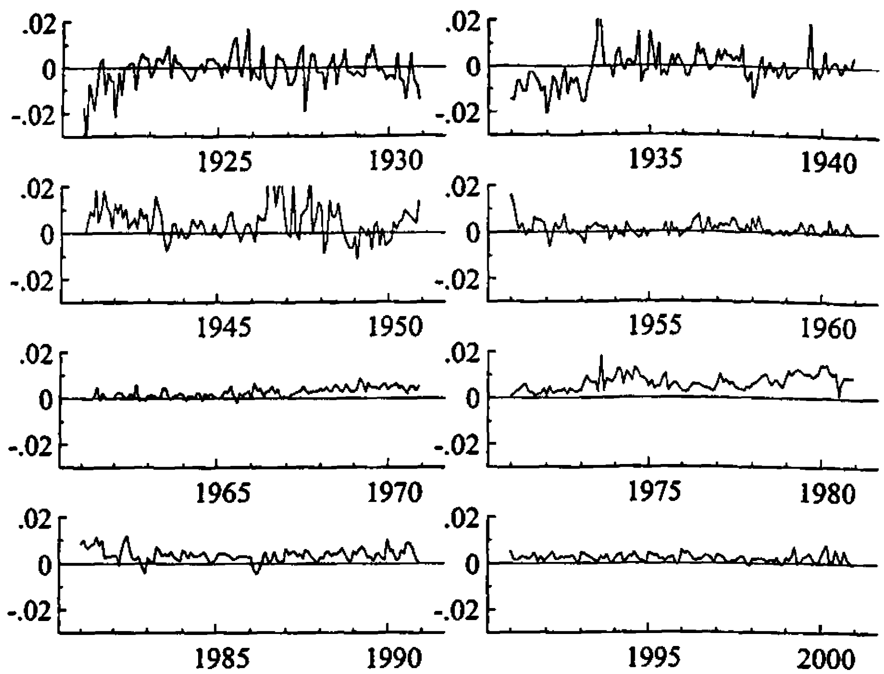


Figure 4.2. Inflation rates. Monthly observations.

In order to examine the stochastic small sample properties and the effects from the choice of sample period, the monthly and quarterly data are divided into periods of respectively 10 and 20 years. The inflation rates for these subsamples are shown in Figure 4.2 and 4.3.

The hypotheses of non-stationary inflation rates, or if the levels are $I(2)$, are tested with an augmented Dickey-Fuller (ADF) test in differences:

$$\Delta^2 x_t = (\rho - 1)\Delta x_{t-1} + \sum_{j=1}^{k-2} \varphi_j \Delta^2 x_{t-j} + \phi s_t + u_t,$$

where ρ, φ_j and ϕ are parameters to be estimated and s_t are seasonal dummies used in the analyses of monthly and quarterly observations. The choice of lag length k is based on Schwarz information criteria.

The null hypothesis that x_t is $I(2)$ is formulated as $\rho = 1$, implying a unit root in differences, and the alternative is $\rho < 1$. If the parameters are constant, the asymptotic distribution is the standard Dickey-Fuller and critical values for the relevant sample sizes are given in Table 4.1.

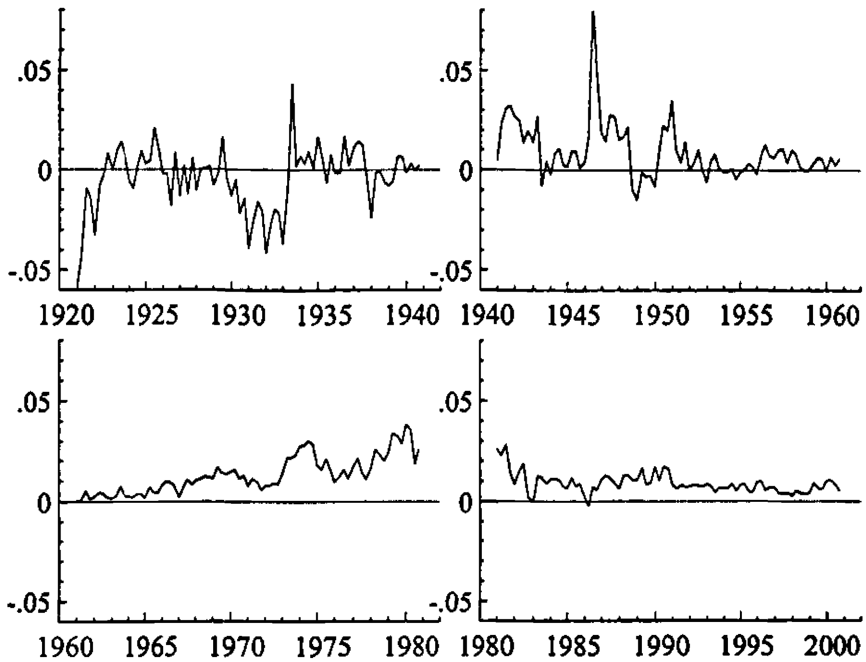


Figure 4.3. Inflation rates. Quarterly observations.

Table 4.1. Critical values in the DF distribution[3]

Observations	Critical values		
	1%	5%	10%
$T = 50$	−2.62	−1.95	−1.61
$T = 100$	−2.60	−1.95	−1.61
$T \rightarrow \infty$	−2.58	−1.95	−1.61

Source: Hamilton (1994).

Tables 4.2 to 4.4 report the tests for the hypotheses of $I(2)$ness for respectively monthly, quarterly and annual observations.

[3]Note that the simulations are made under the assumption that $\varphi_i = 0$ ($i = 1, ..., k-2$). Hence, it might not be quite accurate for the illustrations in this section but will be used as an approximation.

Table 4.2. Augmented Dickey-Fuller tests for monthly observations

Period	Estimate ($\widehat{\rho}$)	DF test statistic	T
Full sample	0.81	-6.98	1051
1921-1930	0.37	-4.44	115
1931-1940	0.64	-3.28	115
1941-1950	0.80	-2.11	115
1951-1960	0.57	-3.60	115
1961-1970	0.96	*-0.70*	115
1971-1980	0.98	*-0.52*	115
1981-1990	0.89	-2.09	115
1991-2000	0.94	*-1.09*	115

Note: Test statistics in italics indicate that the null is accepted with a 5% significance level.

Table 4.3. Augmented Dickey-Fuller tests for quarterly observations

Period	Estimate ($\widehat{\rho}$)	DF test statistic	T
Full sample	0.78	-6.62	350
1921-1940	0.54	-5.12	78
1941-1960	0.72	-3.56	78
1961-1980	0.98	*-0.50*	78
1981-2000	0.87	-2.58	78

Note: See Table 4.2.

Table 4.4. Augmented Dickey-Fuller tests for annual observations

Period	Estimate ($\widehat{\rho}$)	DF test statistic	T
Full sample	0.70	-3.89	85
1913-1956	0.55	-3.40	41
1957-2000	0.94	*-1.28*	41

Note: See Table 4.2.

When estimating over the full sample, regardless of the frequency, the null is rejected. In other words, the hypothesis of stationary inflation rate is accepted. It is doubtful, however, if we can apply the limited DF distribution, as the assumption of constant parameters is most likely violated. In many of the sub samples the tests indicate that the inflation rate is not stationary. Particularly in the latter sub samples, i.e. the subsamples after 1960, and in particular with monthly observations the hypothesis of unit root is accepted. In fact, there is strong evidence that parameters are not constant over the sample. In other words, the critical values reported in Table 4.1 are not valid, and the interpretation of the test statistics for the full sample cannot be made on this basis. This indicates different regimes of inflation rates in the period considered. A simple look at the estimates of ρ could indicate that a possible split point is around 1960. Further investigation of this point is beyond the scope of the present analysis and will not be carried out here.

The conclusion from this analysis seems to be, that when analyzing inflation rates with a limited data sample, it could very well be the case that the inflation rate is non-stationary. Furthermore, when considering a very long sample it is very likely that parameters are not constant and the test statistics cannot be compared with the standard distributions.

3.1. Robustness of the results. To examine the robustness of the results obtained above it is investigated if similar conclusions can be obtained applying different methods. First, we look into the description of the DGP. In the section above it was assumed that x_t could be described by an AR process. In the next section we examine if in fact also MA terms should be included to describe the process and if this will change the conclusion of non-constant parameters. Then we look further into the question of non-stationary inflation rates applying a KPSS test. The hypothesis is in this case turned around and the null is that the inflation rate is stationary as opposed to the ADF test, where the null hypothesis is non-stationarity.

3.1.1. *Fitting an ARIMA model.* In this section we try to fit an $ARIMA(p,d,q)$ to explain the variation in the US consumer price index for the period 1913-2000. We consider monthly observations and to get a first impression of the properties of the time series, we take a look at the autocorrelations and the partial autocorrelations. These are given for $n = 1, 2, ..., 24$ in Figure 4.4 where the dotted lines indicate 95% confidence intervals.[4] The autocorrelations decrease very slowly whereas the partial autocorrelations are insignificant for $n \geq 2$ indicating that AR is present in the time series, and that there is at least one unit root.

[4]The Box-Jenkins procedure in RATS by ESTIMA was used for estimations in this section.

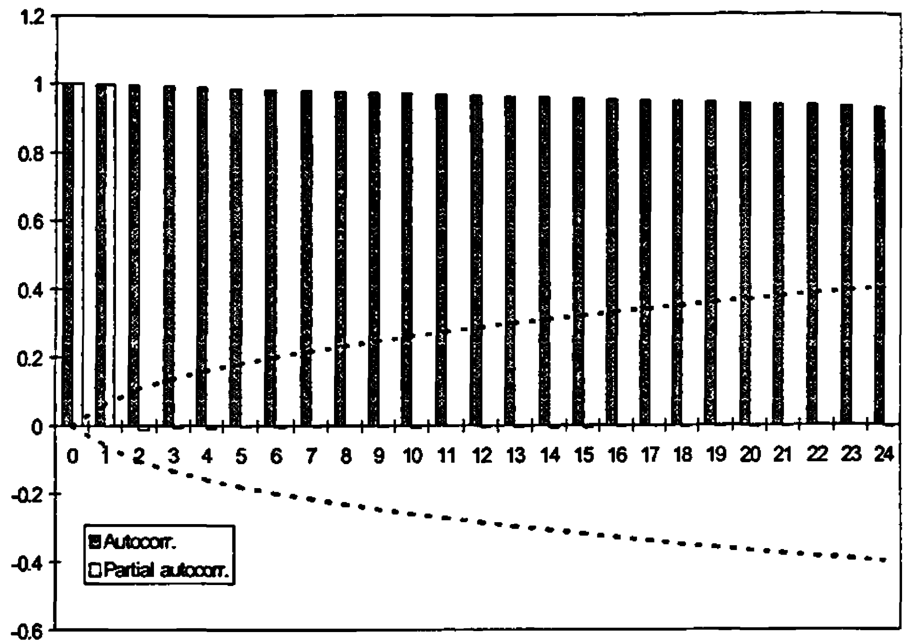


Figure 4.4. Autocorrelations and partial autocorrelations

An $ARIMA(p,1,q)$ model is fitted starting with $p = q = 2$ and testing down by means of likelihood ratio tests. To examine stability of the coefficients, $ARIMA(p,1,q)$ models are fitted for each decade in the sample. The results are given in Table 4.5 where the p-values are for the reduction of the model in the row given the alternative of an $ARIMA(2,1,2)$ model. In two cases convergence could not be obtained. This may have to do with lack of invertibility of the MA terms of the model. In these cases the alternative model has $q < 2$.

The evidence from Table 4.5 seems to suggest that indeed the parameters are not constant over the full sample. Furthermore, in one case (1941-1950) it is not clear if the DGP should be described by an $ARIMA(1,1,0)$ or an $ARIMA(0,1,1)$. In four decades there does not seem to be MA components present in the process and in the last decade the process seems to be $ARIMA(0,1,1)$.

Table 4.5. Estimations of $ARIMA(p, 1, q)$ models

	const	$AR(1)$	$AR(2)$	$MA(1)$	p-value
Full sample	0.0027 (0.0009)	1.2023 (0.0574)	−0.2413 (0.0475)	−0.7970 (0.0445)	0.28
1921-1930	−0.0011 (0.0008)	0.2703 (0.0826)	-	-	0.43
1931-1940	−0.0007 (0.0011)	0.4825 (0.0799)	-	-	0.16*
1941-1950	0.0049 (0.0012)	0.4159 (0.0848)	-	-	0.26
	0.0048 (0.0010)	-	-	0.3837 (0.0859)	0.07
1951-1960	0.0012 (0.0003)	0.2231 (0.0844)	-	-	0.20
1961-1970	0.0026 (0.0004)	0.3107 (0.0896)	0.3000 (0.0892)	-	0.06**
1971-1980	0.0072 (0.0013)	0.9227 (0.0454)	-	−0.6101 (0.0979)	0.79
1981-1990	0.0035 (0.0005)	0.5158 (0.0775)	-	-	0.80
1991-2000	0.0022 (0.0002)	-	-	0.2780 (0.0897)	0.08

Notes: Numbers in brackets are standard deviations. * The alternative model is ARIMA(1,1,1). ** The alternative model is ARIMA(3,1,0).

3.1.2. *Testing for stationarity.* In this section it is discussed whether the conclusions in section 3 depend on the testing procedure applied, i.e. the ADF test. The same question is here analyzed when the hypothesis is the opposite. Here the null is that the inflation rate is stationary. The test applied is developed by Kwiatkowski et al. (1992), commonly referred to as the KPSS test.

Assume that the time series x_t can be decomposed into the sum of a deterministic trend, a random walk, and a stationary error:

$$x_t = \xi t + r_t + \varepsilon_t, \varepsilon_t \sim iidN(0, \sigma_\varepsilon^2). \tag{3.1}$$

The term r_t serves the role of the random walk: $r_t = r_{t-1} + u_t$, with r_0 fixed and $u_t \sim iid(0, \sigma_u^2)$. In general (3.1) is an $I(1)$ process unless $\sigma_u^2 = 0$ in which case it is trend-stationary. Hence, the null of stationarity is formulated as the hypothesis $\sigma_u^2 = 0$.

Taking the difference of (3.1) yields:

$$\Delta x_t = \xi + w_t,$$

where $w_t = u_t + \Delta\varepsilon_t$ is the error for Δx_t. If u_t and ε_t are independent of each other, w_t can be expressed as an MA(1) process: $w_t = v_t + \theta v_{t-1}$ and (3.1) is seen to be equivalent to the ARIMA model

$$x_t = \xi + \beta x_{t-1} + w_t,$$

with $\beta = 1$. The difference between the DF test and the KPSS test is that the first tests the hypothesis $\beta = 1$ with $\theta = 0$, whereas the KPSS test considers the hypothesis $\theta = -1$ with $\beta = 1$.

Since it is unlikely that prices have a quadratic trend in levels, we consider the test where $\xi = 0$, in which case the KPSS test coincides with the locally best invariant (LBI) test of Nyblom and Makelainen (1983). KPSS refer to the test statistic for the null in this model as η_τ. The same tests as in section 3 are carried out. The values of the lag truncation parameter, which is used for estimation of the long-run variance, is allowed to vary from 0 to 8. The results for monthly observations are reported in Table 4.6, for quarterly observations in Table 4.7, and for annual data in Table 4.8. The 5% critical value is 0.146 and the 10% critical value is 0.119. A '*' indicates acceptance of the null at the 5% level.

Table 4.6. KPSS tests statistics. Monthly observations.

	Lag truncation parameter (L)								
Period	0	1	2	3	4	5	6	7	8
Full sample	0.883	0.568	0.439	0.364	0.312	0.275	0.248	0.226	0.209
1921-1930	0.420	0.333	0.298	0.281	0.263	0.244	0.229	0.216	0.206
1931-1940	0.548	0.381	0.311	0.272	0.248	0.233	0.224	0.213	0.201
1941-1950	0.233	0.167	0.141*	0.122*	0.107*	0.098*	0.093*	0.089*	0.085*
1951-1960	0.223	0.182	0.160	0.149	0.144*	0.144*	0.144*	0.142*	0.137*
1961-1970	0.147	0.143*	0.140*	0.143*	0.146*	0.143*	0.140*	0.142*	0.142*
1971-1980	0.421	0.289	0.221	0.183	0.159	0.143*	0.131*	0.121*	0.113*
1981-1990	0.496	0.328	0.273	0.249	0.235	0.226	0.223	0.219	0.214
1991-2000	0.068*	0.059*	0.065*	0.073*	0.083*	0.087*	0.083*	0.077*	0.075*

Table 4.7. KPSS test statistics. Quarterly observations.

	Lag truncation parameter (L)								
Period	0	1	2	3	4	5	6	7	8
Full sample	0.481	0.285	0.213	0.174	0.150	0.133*	0.121*	0.112*	0.105*
1921-1940	0.252	0.174	0.139	0.117*	0.102*	0.092*	0.087*	0.083*	0.080*
1941-1960	0.110*	0.073*	0.061*	0.055*	0.051*	0.049*	0.048*	0.048*	0.050*
1961-1980	0.124*	0.076*	0.059*	0.049*	0.043*	0.040*	0.039*	0.039*	0.040*
1981-2000	0.135*	0.097*	0.089*	0.080*	0.073*	0.069*	0.068*	0.067*	0.066*

Table 4.8. KPSS test statistics. Annual observations.

	Lag truncation parameter (L)								
Period	0	1	2	3	4	5	6	7	8
Full sample	0.189	0.116*	0.095*	0.086*	0.080*	0.076*	0.072*	0.069*	0.067*
1914-1956	0.263	0.166	0.140*	0.129*	0.125*	0.121*	0.116*	0.113*	0.111*
1957-2000	0.668	0.366	0.271	0.224	0.194	0.173	0.157	0.146*	0.139*

The results from the present analysis differ at certain points from the analysis in section 3. When using monthly observations the full sample test rejects that the inflation rate is stationary, which differ from the outcome of the test using ADF. This is also the case when testing in the subsamples covering the first two and the seventh decades in line with the outcome of the ADF tests. Also the outcome of the test for the third decade is in line with the ADF outcome whereas in the fifth and last decade the hypotheses are accepted where the ADF indicates non-stationary inflation rates. In the rest of the subsamples the outcomes depend on the value of L.

When using quarterly data there is more evidence supporting the null, but for the full sample and the first subsample the outcomes depend on the value of L. The outcome here is more or less in line with the ADF test which, however, rejected stationary inflation rates in the period 1961-1980.

When using annual data the outcome of the KPSS test depends on the value of L in all periods. However, there is some evidence that the inflation rate might be stationary in the early period and non-stationary in the late in line with the ADF test.

The overall conclusions in section 3 are not altered by the present analysis. Also the KPSS test for the full sample is very doubtful since it seems that parameters are not constant. Furthermore, also the present analysis suggests that inflation rates may indeed be non-stationary when considering certain periods.

4. The VAR models for I(1) and I(2) variables

In this section a brief description is given of the $I(1)$ and the $I(2)$ models. The intention is by no means to give a fully detailed description of estimation and testing procedures but simply to present the error correction and moving average representations in order to introduce some notation, which will be useful for the scenario analyses in section 5 and 6, and the empirical analysis in section 7. For a more detailed description - which also include estimation and testing in the $I(1)$ case - the reader is referred to the textbook by Johansen (1996)

4.1. The I(1) model. It is assumed that the data generating process (DGP) can be described by a VAR model. Let us consider the most general case of an unrestricted VAR(k) model. The error correction model (ECM) reads:

$$\Delta X_t = \Pi X_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} + \Phi D_t + \varepsilon_t, \tag{4.1}$$

where $\varepsilon_t \sim iid(0, \Omega)$. The data vector X_t is of dimension p and D_t are the deterministic terms of the model. The matrices Π, Γ_i ($i = 1, ..., k-1$) and Φ are of dimension $p \times p$ and include parameters to be estimated.

In the unrestricted form of (4.1) $X_t \sim I(0)$. The $I(1)$ and $I(2)$ models are nested in this one. If $rank(\Pi) = rank(\alpha\beta') = r < p$, cointegration exists and the number of cointegrating relations are r, or equivalently, there are $p - r$ common $I(1)$ trends. The vectors in β are the cointegration vectors and α includes the adjustment coefficients. When cointegration is present, the process can be formulated in the moving average (MA) representation:

$$X_t = C \sum_{i=1}^{t} (\varepsilon_i + \Phi D_i) + C(L)(\varepsilon_t + \Phi D_t) + A,$$

where $C = \beta_\perp(\alpha'_\perp \Gamma \beta_\perp)^{-1}\alpha'_\perp$, $\Gamma = I - \sum_{i=1}^{k-1} \Gamma_i$, $C(L)$ is a convergent power series and A depends on the initial conditions such that $\beta' A = 0$.[5] The term $\alpha'_\perp \sum_{i=1}^{t} \varepsilon_i$ is the only non-stationary part of the process and is referred to as the common stochastic $I(1)$ trends. The term $\beta_\perp$ includes the loadings from the stochastic trends into the process.

Besides the reduced rank requirement mentioned above it is required that $rank(\alpha'_\perp \Gamma \beta_\perp) = p - r$, i.e. the matrix has full rank. If this is not the case the process contains stochastic I(2) trends.

4.2. The I(2) model. The ECM for the $I(2)$ model is written as

$$\Delta^2 X_t = \Pi X_{t-1} - \Gamma \Delta X_{t-1} + \sum_{i=1}^{k-2} \Psi_i \Delta^2 X_{t-i} + \Phi D_t + \varepsilon_t,$$

where again $\varepsilon_t \sim idd(0, \Omega)$. The matrices Π, Γ, Ψ_i ($i = 1, ..., k-2$) and Φ include the coefficients to be estimated and D_t contains the deterministic terms. The requirement for cointegration in this model is

[5] The notation $\perp$ indicates an orthogonal complement such that $\alpha'_\perp \alpha = 0$ and $\beta'_\perp \beta = 0$.

$rank(\Pi) = rank(\alpha\beta') = r < p$ and $rank(\alpha'_\perp \Gamma \beta_\perp) = rank(\xi\eta') = s_1 < p-r$. In this case the MA representation reads

$$\begin{aligned} X_t = \; & C_2 \textstyle\sum_{s=1}^t \sum_{i=1}^s (\varepsilon_i + \Phi D_i) + C_1 \sum_{i=1}^t (\varepsilon_i + \Phi D_i) \\ & + C_2(L)(\varepsilon_t + \Phi D_t) + A + Bt, \end{aligned}$$

where $C_2 = \beta_{\perp 2}(\alpha'_{\perp 2}\theta\beta_{\perp 2})^{-1}\alpha'_{\perp 2}$, the terms $\theta, C_1, C_2(L), A$ and B are described in Johansen (1996) Theorem 4.6. The β and $\beta_\perp$ matrices are decomposed such that β, $\beta_{\perp 1}$ and $\beta_{\perp 2}$ are mutual orthogonal and $sp(\beta, \beta_{\perp 1}, \beta_{\perp 2}) = \mathbf{R}^p$. The same decomposition is made for α and $\alpha_\perp$ Furthermore, β is decomposed such that $\beta = (\beta_0, \beta_1)$. The common stochastic $I(2)$ trends are $\alpha'_{\perp 2} \sum_{s=1}^t \sum_{i=1}^s \varepsilon_i$ and $\alpha_{\perp 2}$ are the coefficients whereas $\widetilde{\beta}_{\perp 2} = \beta_{\perp 2}(\alpha'_{\perp 2}\theta\beta_{\perp 2})^{-1}$ are the (normalized) loadings.

In this model we have r cointegration vectors, s_1 common $I(1)$ trends and $s_2 = p-r-s_1$ common $I(2)$ trends. If $r > s_2$, $r-s_2$ relations cointegrate from the $I(2)$ space to the $I(0)$ space such that $\beta'_0 X_t \sim I(0)$. Furthermore s_2 directions multicointegrate to stationarity, i.e. $\beta'_1 X_t \sim I(1)$ but $\beta'_1 X_t + \kappa'\Delta X_t \sim I(0)$. If, on the other hand, $r \leq s_2$ then there will be no directly cointegrating relations and r directions multicointegrate from $I(2)$ to $I(0)$. The cointegrating relations are summarized in Table 4.9.

Table 4.9. Cointegrating relations

Direction	Dimension	Associated stationary process
$\beta'_0 X_t \sim I(0)$	$r - s_2$ (if $r > s_2$)	
$\beta'_1 X_t \sim I(1)$	$\min(s_2, r)$	$\beta'_1 X_t + \kappa'\Delta X_t \sim I(0)$
$\beta'_{\perp 1} X_t \sim I(1)$	s_1	$\beta'_{\perp 1}\Delta X_t \sim I(0)$
$\beta'_{\perp 2} X_t \sim I(2)$	s_2	$\beta_{\perp 2}\Delta^2 X_t \sim I(0)$

From Table 4.9 it appears that if $I(2)$ is present in a system of variables and cointegration exist, then it will always be the case that at least one of the cointegrating relations is multicointegrating, whereas there might not be any direct cointegration relations. But, strictly speaking, this would rule out the case that in a system with two variables they cointegrate to stationarity, which of course is very possible. In this case the estimate of κ in the multicointegrating relation will be insignificantly different from zero. Hence, in the multicointegrating relations it might be the case that $\kappa = 0$ and they will, in fact, be directly cointegrating.

5. Possible scenarios

In this section possible scenarios are described. Juselius (2002) chapter 12 also contains a scenario analysis for testing PPP in the cointegrated VAR. The value-added of the following analysis is a more extensive treatment of the multicointegrating relations in case of $I(2)$ness in order to explain Definition 4 and 6 in greater details. Furthermore, more possible scenarios are treated and emphasis is on transformation of the system from the $I(2)$ to the $I(1)$ space.

First, the cases with stationary inflation rates in both countries are considered. In these cases the exchange rate can be at most $I(1)$ for the PPP to hold. Then, the cases where inflation rates are not necessarily stationary are discussed. As will appear, in these situations it might be required that the exchange rate is $I(2)$ for the PPP to hold. The notation used in this section is p^1 and p^2 for (the logarithm of) the price levels in countries 1 and 2 and e is (the logarithm of) the exchange rate. Hence, the data matrix considered is $X_t' = [p^1, p^2, e]_t$. The scenarios considered are summarized in Figure 4.5.

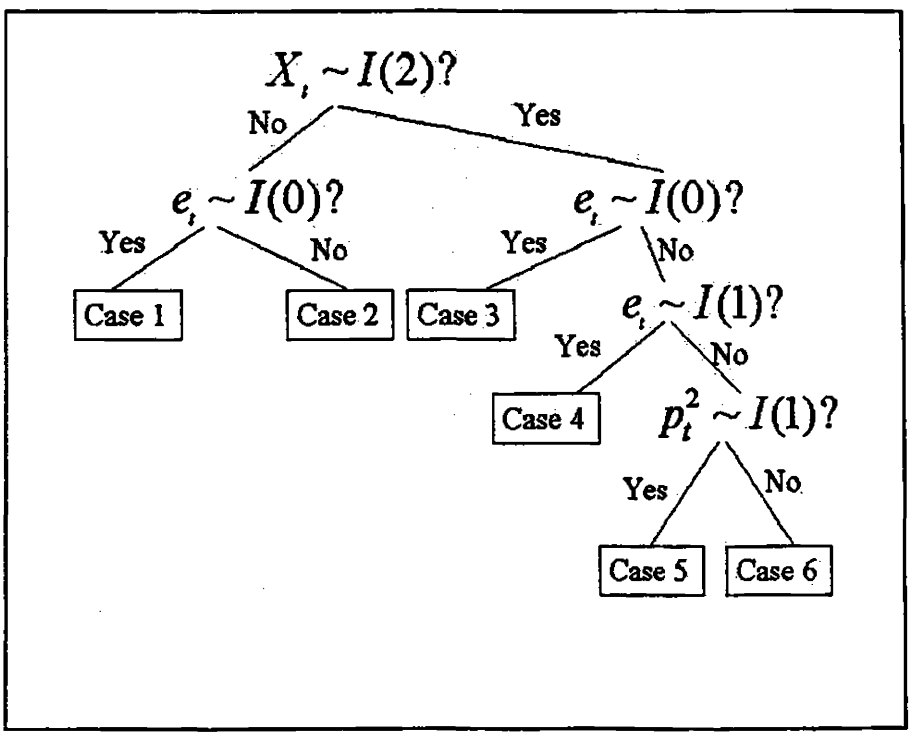


Figure 4.5. Possible scenarios.

5.1. Stationary inflation rates and exchange rate changes. The situation where price levels are stationary is ruled out, since this is unlikely to occur in reality. The exchange rate, on the other hand, will be allowed to be stationary, which could be the case if the two countries have the same currency, or if they have a fixed exchange rate policy. Note that for the cases described in this subsection, the relevant hypothesis to test is that of the PPP (Definition 5). Testing for same underlying inflation and the PPP with adjustment makes little sense as none of the variables are integrated of order 2.

For the PPP to hold it must be the case that there exists at least one cointegration relation or equivalently maximum two common stochastic trends. Thus, the general system considered can be written as

$$\begin{bmatrix} p^1 \\ p^2 \\ e \end{bmatrix}_t = \begin{bmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \\ d_{31} & d_{32} \end{bmatrix} \begin{bmatrix} \sum_{i=1}^{t} v_{1,i} \\ \sum_{i=1}^{t} v_{2,i} \end{bmatrix} + stat,$$

where $v_{k,t} = f_t^k(\varepsilon_{p^1}, \varepsilon_{p^2}, \varepsilon_e)$, $k = 1, 2$ for linear functions f^k, and *stat* refers to the stationary components of the process. Two cases are considered: Case 1 where the exchange rate is stationary and Case 2 where the exchange rate is $I(1)$.

Case 1: $p^1, p^2 \sim I(1)$ **and** $e \sim I(0)$. In this case no adjustment takes place through the exchange rate and it is therefore required that the prices move in a similar manner. As the exchange rate is stationary ($d_{31} = d_{32} = 0$), the system can only have one common trend for the PPP to hold ($d_{12} = d_{22} = 0$). Hence for the PPP (Definition 5) to hold it is required that $d_{11} = d_{21}$.

Case 2: $p^1, p^2, e \sim I(1)$. If the prices do not move together the difference should be captured in the development of the exchange rate. Hence, the exchange rate has to be integrated of order one to capture different movements in two non-cointegrated $I(1)$ series. In the case of only one cointegrating relation it is required that $d_{11} = d_{21} + d_{31}$ and $d_{12} = d_{22} + d_{32}$ for the PPP to hold.

If, on the other hand, only one common trend is driving the system ($d_{12} = d_{22} = d_{32} = 0$), we have two cointegrating vectors. An example of two cointegrating vectors that fulfill the requirement of the PPP are $\beta_1' = [1, -\frac{1}{2}, 0]$ and $\beta_2' = [0, -\frac{1}{2}, -1]$. More generally, it is required that $(1, -1, -1) \in sp(\beta)$ or that $d_{11} = d_{21} + d_{31}$.[6]

[6]The latter folows from the fact that we can normalize the two cointegration vectors such that $\beta_1' = [1, 0, -\frac{d_{11}}{d_{31}}]$ and $\beta_2' = [0, -1, -\frac{d_{21}}{d_{31}}]$. It then follows that $\frac{d_{21}}{d_{31}} + \frac{d_{11}}{d_{31}} = 1 \Leftrightarrow d_{11} = d_{21} + d_{31}$.

5.2. Non-stationary inflation rates and/or exchange rate changes. When inflation rates are non-stationary, price levels are integrated of order two, i.e. we need to take differences twice in order to reach stationarity. If the impact from one common $I(2)$ trend is not the same on the prices, the exchange rate also needs to be $I(2)$ for the PPP to hold.

When the system includes a variable which is $I(2)$ the analysis becomes somewhat more complicated. In fact, if one variable is $I(2)$ at least two must be, and cointegration is needed between them for the PPP to hold. Two general situations need to be considered. One where only two variables are $I(2)$ and one where all three variables are. In the first one, there must be at most one common $I(2)$ trend and one independent $I(1)$ trend for the PPP to hold. In the latter also the situation with two $I(2)$ trends and no independent $I(1)$ trend is a possibility. Case 3 to 5 below describe the first situation and the general system reads:

$$\begin{bmatrix} p^1 \\ p^2 \\ e \end{bmatrix}_t = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} \sum_{s=1}^{t} \sum_{i=1}^{s} u_i + \begin{bmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \\ d_{31} & d_{32} \end{bmatrix} \begin{bmatrix} \sum_{i=1}^{t} u_i \\ \sum_{i=1}^{t} v_i \end{bmatrix} + stat,$$

where $u_t = f^1(\varepsilon_{p_t^1}, \varepsilon_{p_t^2}, \varepsilon_{e_t}), v_t = f^2(\varepsilon_{p_t^1}, \varepsilon_{p_t^2}, \varepsilon_{e_t})$ and f^1 and f^2 are linear functions. Note that when one considers a three dimensional system with $I(2)$ness present it is only possible for the PPP with adjustment to hold since there will be (at least) one multicointegrating relation according to section 4.2. However, if $\kappa = 0$ we can apply Definition 5. In the case where the system is designed for testing PPP between three or more countries, direct cointegration is possible. An example of this situation is described in section 6.

Four cases are considered: Case 3, where both inflation rates are non-stationary and the exchange rate is stationary; Case 4, in which we consider a situation similar to Case 3 but with the exchange rate being $I(1)$. The fifth case is in terms similar to Case 4. However, in Case 5 one of the price series is assumed to be $I(2)$ whereas the other is $I(1)$. In this case it is needed for the exchange rate to be $I(2)$ for the PPP to be fulfilled. The last case considered is where all three variables are $I(2)$. Case 3 could be relevant for two countries with fixed exchange rates. The last three cases are more standard in the sense that exchange rate adjustment is needed to capture (stochastic) difference in the price developments.

Case 3: $p^1, p^2 \sim I(2)$ **and** $e \sim I(0)$. As the exchange rate is stationary it is neither affected by the $I(2)$ nor the $I(1)$ trend. Hence,

$c_3 = d_{31} = d_{32} = 0$. Or in other words, there is a cointegration vector consisting only of the exchange rate: $(0,0,1) \in sp(\beta)$ or with the notation in Table 4.9, $\beta_0' = (0,0,1)$. For the PPP with adjustment to hold it is therefore required that the system has two cointegrating vectors, $(r, s_1, s_2) = (2,0,1)$. Thus, neither of the prices can be affected by the independent $I(1)$ trend: $d_{12} = d_{22} = 0$. Note that in this situation it trivially follows that $\Delta e_t \sim I(0)$. The underlying inflation is the same in the countries if $p_t^1 - p_t^2 \sim I(1)$. Hence, Definition 4B holds if $c_1 = c_2$. To make the analysis more illustrative, we impose the restriction of the unit root associated with the $I(2)$ trend. The system is then mapped into the $I(1)$ space:

$$\begin{bmatrix} p^1 - p^2 \\ e \\ \Delta p^1 \end{bmatrix}_t = \begin{bmatrix} d_{11} - d_{21} \\ 0 \\ c_1 \end{bmatrix} \sum_{i=1}^{t} u_i + stat. \tag{5.1}$$

It appears at once that the PPP with adjustment holds if $d_{11} - d_{12} - \kappa c_1 = 0$ or $\kappa = (d_{11} - d_{12})/c_1$. If $\kappa = 0$ the PPP holds.

Case 4: $p^1, p^2 \sim I(2)$ **and** $e \sim I(1)$. In this case we have $c_3 = 0$. We have to consider two situations, namely those of one and two cointegration relations ($r = 1$ and $r = 2$). First, the case of $r = 1$ and hence $(s_1, s_2) = (1,1)$. As in the former case, Definition 4B holds if $c_1 = c_2$. Imposing the $I(2)$-unit root, the $I(1)$ system becomes:

$$\begin{bmatrix} p^1 - p^2 \\ e \\ \Delta p^1 \end{bmatrix}_t = \begin{bmatrix} d_{11} - d_{21} & d_{12} - d_{22} \\ d_{31} & d_{32} \\ c_1 & 0 \end{bmatrix} \begin{bmatrix} \sum_{i=1}^t u_i \\ \sum_{i=1}^t v_i \end{bmatrix} + stat.$$

For $\beta' = (1,1,\kappa)$, it must hold that $\kappa = (d_{11} - d_{21} - d_{31})/c_1$ and $d_{12} - d_{22} = d_{32}$ in which case Definition 6 holds. Furthermore, if $\kappa = 0$ Definition 5 holds.

Now, let us consider the case $(r, s_1, s_2) = (2,0,1)$. Note that in this situation we will have one directly cointegrating relation and one multicointegrating. Also note, that the exchange rate share the $I(1)$ trend, which is associated to the $I(2)$ trend. I.e. we have $d_{12} = d_{22} = d_{32} = 0$. Hence, the transformed system (when imposing a unit root on the prices) looks like the one in (5.1) except for the zero on the right hand side of the equality which is d_{31}. Formulated another way, we have $p_t^1 - p_t^2 = (d_{11} - d_{21}) \sum_{i=1}^t u_i + stat$, $e_t = d_{31} \sum_{i=1}^t u_i + stat$, and $\Delta p_t^1 = c_1 \sum_{i=1}^t u_i + stat$. Following the same line of arguments as in case 2, it follows that the PPP holds if $(1,1,0) \in sp(\beta)$, i.e. $d_{31} = d_{21} - d_{11}$. The PPP with adjustment holds for $(1,1,\kappa) \in sp(\beta)$, i.e. for $\kappa = (d_{21} - d_{11} - d_{31})/c_1$. In this case we have that $d_{31} \neq d_{21} - d_{11}$ and $c_1 \neq 0$ such that $\kappa \neq 0$.

Case 5: $p^1, e \sim I(2)$ and $p^2 \sim I(1)$. This case is in turn very similar to Case 4 and is just included for the sake of completeness. The desired transformation can be made only if the impact from the $I(2)$ trend is the same on p^1 and e but with opposite sign, i.e. $c_1 = -c_3$.[7] If this is the case the system can be mapped into the $I(1)$ space:

$$\begin{bmatrix} p^1+e \\ p^2 \\ \Delta p^1 \end{bmatrix}_t = \begin{bmatrix} d_{11}+d_{31} & d_{12}+d_{32} \\ d_{21} & d_{22} \\ c_1 & 0 \end{bmatrix} \begin{bmatrix} \sum_{i=1}^t u_i \\ \sum_{i=1}^t v_i \end{bmatrix} + stat.$$

Applying an analysis similar to the one above it follows that in the case $(r, s_1, s_2) = (1, 1, 1)$ the PPP with adjustment holds for $\kappa = (d_{11}+d_{31}-d_{21})/c_1$ and $d_{22} = d_{12}+d_{32}$, and the PPP holds if furthermore $\kappa = 0$.

For $(r, s_1, s_2) = (2, 0, 1)$ Definition 6 applies if $(1, -1, \kappa) \in sp(\beta)$ or expressed another way, if $\kappa = (d_{21} - d_{11} - d_{31})/c_1$. The PPP holds if $d_{21} = d_{11} + d_{31}$ in which case $\kappa = 0$.

Case 6: $p^1, p^2, e \sim I(2)$. When all three variables are integrated of order 2, it is possible that the PPP holds even though the system is driven by two common $I(2)$ trends. Let us first, however, consider the more simple case of only one $I(2)$ trend. For $c_1 = c_2 + c_3$ we have a situation were the exchange rate adjusted underlying inflation is the same in the countries (Definition 4A), i.e. $p^1 - p^2 - e \sim I(1)$ or $\Delta p^1 - \Delta p^2 - \Delta e \sim I(0)$. One unit root is imposed and the transformed system could look like the following:

$$\begin{bmatrix} p^1-p^2-e \\ \Delta p^1 \\ c_2p^1 - c_1p^2 \end{bmatrix}_t = \begin{bmatrix} d_{11}-d_{21}-d_{31} & d_{12}-d_{22}-d_{32} \\ c_1 & 0 \\ c_2d_{11} - c_1d_{21} & c_2d_{12} - c_1d_{22} \end{bmatrix} \begin{bmatrix} \sum_{i=1}^t u_i \\ \sum_{i=1}^t v_i \end{bmatrix}$$
$$+stat.$$

The last variable is put in the system in order not to loose any information there might be in the independent $I(1)$ trend. The situation illustrated in the formula above is with $(r, s_1, s_2) = (1, 1, 1)$ in the original system and $(r, s) = (1, 2)$ in the transformed system. The requirement for the PPP with adjustment to hold is that the $sp(\beta) = sp(1, \kappa, 0)$ which implies that $c_2d_{11} = c_1d_{21}$, $c_2d_{12} = c_1d_{22}$, $d_{12} = d_{22} + d_{32}$, and $\kappa = (d_{21} + d_{31} - d_{11})/c_1$. The PPP implies further that $\kappa = 0$.

If there is no independent $I(1)$ trend present, i.e. with $r = 2$ cointegrating vectors ($d_{12} = d_{22} + d_{32}$ and $c_2d_{12} = c_1d_{22}$) two situations can

[7] If we instead had considered the situation $p^2, e \sim I(2)$ and $p^1 \sim I(1)$, the impact from the $I(2)$ trend should be the same, i.e. $c_2 = c_3$.

arise: one where $c_2p^1 - c_1p^2$ is significant and one where it is insignificant. In the first case, one of the cointegrating vectors includes this variable and the interpretation might be difficult. For Definition 6 to apply it is then required that $(1, \kappa, 0)$ is the other vector. If, on the other hand, $c_2p^1 - c_1p^2$ is insignificant the PPP with adjustment holds if $(1, \kappa, 0) \in sp(\beta)$ and - as usual - the PPP holds if $\kappa = 0$.

Now let us turn to the situation with two common $I(2)$ trends and no independent $I(1)$ trend in the system, which then reads:

$$\begin{bmatrix} p^1 \\ p^2 \\ e \end{bmatrix}_t = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \\ c_{31} & c_{32} \end{bmatrix} \begin{bmatrix} \sum_{s=1}^{t} \sum_{i=1}^{s} u_{1i} \\ \sum_{s=1}^{t} \sum_{i=1}^{s} u_{2i} \end{bmatrix} + \begin{bmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \\ d_{31} & d_{32} \end{bmatrix} \begin{bmatrix} \sum_{i=1}^{t} u_{1i} \\ \sum_{i=1}^{t} u_{2i} \end{bmatrix} + stat.$$

In this case there can be only one cointegration vector. Hence, if the PPP (possibly with adjustment) holds it must be the case that the combination $p^1 - p^2 - e$ eliminates both $I(2)$ trends, i.e. $c_{11} = c_{21} + c_{31}$ and $c_{12} = c_{22} + c_{32}$. In this case Definition 4A applies. Imposing the two unit roots from the $I(2)$ trends the transformed system could look like the following

$$\begin{bmatrix} p^1 - p^2 - e \\ \Delta p^1 \\ \Delta p^2 \end{bmatrix}_t = \begin{bmatrix} d_{11} - d_{21} - d_{31} & d_{12} - d_{22} - d_{32} \\ c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \begin{bmatrix} \sum_{i=1}^{t} u_{1i} \\ \sum_{i=1}^{t} u_{2i} \end{bmatrix} + stat.$$

Note that to keep as much information as possible in the transformed system two differences are included. For example, it might be the case that $c_{12} = 0$ in which case Δp^1 would only contain information of the first $I(2)$ trend. Now the PPP with adjustment holds if $\kappa_1 c_{11} + \kappa_2 c_{21} = d_{31} + d_{21} - d_{11}$ and $\kappa_1 c_{12} + \kappa_2 c_{22} = d_{32} + d_{22} - d_{12}$ possibly with either κ_1 or κ_2 equal to zero.[8] The PPP holds if $\kappa_1 = \kappa_2 = 0$.

6. Testing multiple PPP: An example with three countries

A test of multiple PPP when inflation rates are non-stationary can provide evidence of how prices adjust between three countries. For example, it could be the case, that two of the countries have the same level of prices, whereas the third country is still adjusting towards the

[8] Solving these two equations with respect to κ_1 and κ_2 gives $\kappa_1 = (c_{22}(d_{31} + d_{21} - d_{11}) - c_{21}(d_{32} + d_{22} - d_{21}))/(c_{11}c_{22} - c_{12}c_{21})$ and $\kappa_2 = (c_{11}(d_{32} + d_{22} - d_{12}) - c_{12}(d_{31} + d_{21} - d_{11}))/(c_{11}c_{22} - c_{12}c_{21})$.

common level. In this case one might expect the PPP (Definition 5) to hold between country 1 and 2 and the PPP with adjustment (Definition 6) between these and area 3.

Testing the multiple PPP between areas where the currency is the same or fixed is quite simple compared to the case where the countries in question have different currencies. When the countries have different currencies, the system easily becomes very big. Here a simple three country model is considered. The data vector consists of five variables: $X_t' = [p^1, p^2, p^3, e^{12}, e^{13}]_t$, where p_t^i $(i = 1, 2, 3)$ is (the logarithm of) the price level in country i and e_t^{1j} $(j = 1, 2)$ is (the logarithm of) the exchange rate between country 1 and j. For simplicity, we consider a situation where price levels are $I(2)$ but exchange rates are $I(1)$. Furthermore, we assume that the prices share the same $I(2)$ trend and there are two independent $I(1)$ trends in the system:

$$\begin{bmatrix} p^1 \\ p^2 \\ p^3 \\ e^{12} \\ e^{13} \end{bmatrix}_t = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ 0 \\ 0 \end{bmatrix} \sum_{s=1}^{t} \sum_{i=1}^{s} u_i + \begin{bmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \\ d_{31} & d_{32} & d_{33} \\ d_{41} & d_{42} & d_{34} \\ d_{51} & d_{52} & d_{35} \end{bmatrix} \begin{bmatrix} \sum_{i=1}^{t} u_i \\ \sum_{i=1}^{t} v_{1i} \\ \sum_{i=1}^{t} v_{2i} \end{bmatrix} + stat,$$

where $u_t = f_t^1(\varepsilon_{p^1}, \varepsilon_{p^2}, \varepsilon_{p^3}, \varepsilon_{e^{12}}, \varepsilon_{e^{13}})$, and $v_{kt} = f_t^k(\varepsilon_{p^1}, \varepsilon_{p^2}, \varepsilon_{p^3}, \varepsilon_{e^{12}}, \varepsilon_{e^{13}})$, $k = 2, 3$, with f^1 and f^k being linear functions. We have a situation with two cointegration relations, one directly cointegrating and one multicointegrating. In this model country 1 is taken as the numeraire country without loss of generality.

If the impact from the $I(2)$ trend is the same on all prices ($c_1 = c_2 = c_3$) then the underlying inflation is the same in all of the countries (Definition 4B). Imposing this unit root, the system can be mapped into the $I(1)$ space:

$$\begin{bmatrix} p^1 - p^2 \\ p^1 - p^3 \\ \Delta p^1 \\ e^{12} \\ e^{13} \end{bmatrix}_t = \begin{bmatrix} d_{11} - d_{21} & d_{12} - d_{22} & d_{13} - d_{23} \\ d_{11} - d_{31} & d_{12} - d_{32} & d_{13} - d_{33} \\ c_1 & 0 & 0 \\ d_{41} & d_{42} & d_{43} \\ d_{51} & d_{52} & d_{53} \end{bmatrix} \begin{bmatrix} \sum_{i=1}^{t} u_i \\ \sum_{i=1}^{t} v_{1i} \\ \sum_{i=1}^{t} v_{2i} \end{bmatrix}$$
$$+stat.$$

The PPP (Definition 5) holds between 1 and 2 if $d_{41} = d_{11} - d_{21}$, $d_{42} = d_{12} - d_{22}$ and $d_{43} = d_{13} - d_{23}$. The PPP with adjustment holds between 1 and 3 if $\kappa = (d_{11} - d_{31} - d_{51})/c_1$, $d_{52} = d_{12} - d_{32}$, and $d_{53} = d_{13} - d_{33}$ for $\kappa \neq 0$. Hence, this illustrates an example of two countries maintaining the same relative price level and a third country in which prices adjust toward the common level.

7. Empirical example: The UK and the US

Other research studying the PPP relation between the UK and the US include Edison (1987), Kim (1990), and Crowder (1996a, 1996b). Applying annual data for the period 1890-1978, Edison (1987) sets up a theoretical monetary model and tests hypotheses with univariate autoregressive (AR) models. She does not find evidence in favor of the PPP but the results do indicate that forces in the economy drive the exchange rate towards the PPP equilibrium. She also split the sample in two subsamples, one covering the period of the gold standard and one covering the Bretton Woods period. Estimations in these periods reinforce her results.

Kim (1990) supplies an extensive study of PPP between the US and five industrial countries, including the UK. Using data for wholesale prices (WPI) as well as CPI covering the period from 1900-1987 (1914-1987 for the case of the CPI) he reports evidence that both exchange rates and prices are integrated of order one (he does not test for $I(2)$ness). In the case of the UK he finds, by means of cointegrating regressions, that the exchange rate expressed in dollars cointegrates with both WPI and CPI whereas the pound rate only cointegrates with the WPI. With univariate test procedures he furthermore finds that the real exchange rate follows a random walk when CPI is used but not when it is the WPI.

In a study with annual WPI data (1900-1991) and CPI data (1869-1991) Crowder (1996a) analyzes the PPP relations between Canada, the UK and the US. He finds evidence of cointegration - using the Johansen method - in the three dimensional system which includes the WPI from the UK and the US but not in the system which includes the CPI.

The PPP between 15 OECD countries including the UK and the US is the subject of another paper by Crowder (1996b). He applies monthly data from January 1973 to February 1992 using the CPI as price variables. For the cases of the US and the UK he reports evidence, by means of Augmented Dickey-Fuller (ADF) tests, that prices in both countries contain two unit roots. Applying Johansen's test for cointegration he finds no support for cointegration (from $I(2)$ to $I(1)$)

among the two price series. But estimating a relative price level in any case, he does find some evidence of cointegration between this and the nominal exchange rate.

The present analysis is more in line with the one by Crowder (1996b) as it is allowed for the variables to be integrated of order two. Also because monthly data is applied for a period starting in 1972.

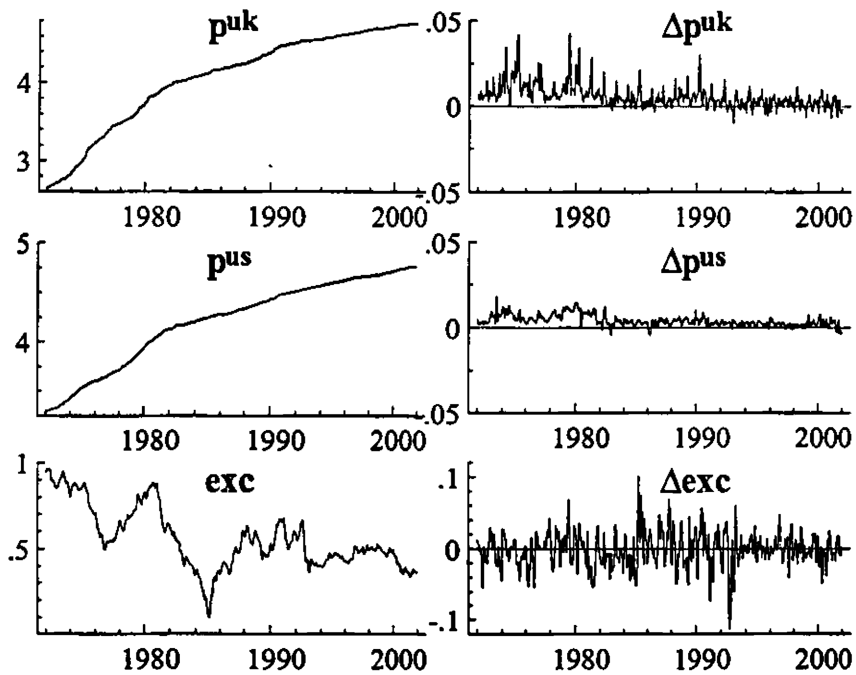


Figure 4.6. Data in level and differences.

7.1. Data and test for misspecification. Data is extracted from IMF's CD-rom containing International Financial Statistics (IFS). The British and American consumer prices (p^{uk} and p^{us}) have the series number 64 and the exchange rate (*exc*), which is the average rate of GBP expresses in USD, is series ah. According to the written version of the IFS, the American CPI series coincide with the one used in Section 3. Observations are monthly and cover the post-bretton wood, January 1972 to December 2001. Hence, a priori there are $T = 360$ observations. The CPI data is not adjusted for seasonality and logarithms is taken of the data prior to the analysis. Illustrations of data in levels and differences are given in Figure 4.6 from which it seems hard to determine whether the inflation rates are stationary. Also it is not clear if the exchange rate is $I(1)$. Thus the analysis starts from the assumption that the data might be integrated of order 2.

As is well-known, the critical values of the test statistic for the number of cointegrating vectors and common $I(1)$ trends depend on the deterministic terms. The analysis here starts with the most general model which does not allow for quadratic trends in levels. The restrictions imposed on the model are given in Rahbek et. al (1999). Tests for exclusion of the linear drift were clearly accepted regardless of the number of cointegrating vectors.[9] Also the test of the model with a constant restricted to the β space against the alternative of an unrestricted constant was strongly accepted. Complete exclusion of the constant was rejected. Hence, the model considered for the further analysis contains a constant restricted to the β space. Furthermore, 11 centered seasonal dummies are included as well as four unrestricted impulse dummies $d738$ (1 in August 1973; -1 in September 1973; and 0 otherwise); $d754$; $d798$; and $d9210$[10] were included. Hence, the cointegrated $I(2)$ model considered - with $k = 2$ lags - is in the ECM:

$$\Delta^2 X_t = \alpha(\beta', \rho_1') \begin{pmatrix} X_{t-1} \\ 1 \end{pmatrix} - \Gamma \Delta X_{t-1} + \Phi_0 S_t + \Phi_1 d_t + \varepsilon_t,$$

where $X_t' = [p^{uk}, p^{us}, exc]_t$ is the data matrix, ρ_1 is an $r \times 1$ vector of coefficients, S_t are the seasonal dummies, d_t includes the impulse dummies, and $\varepsilon_t \sim iid(0, \Omega)$. Misspecification test of the models indicated no autocorrelation in the residuals. The Lagrange Multiplier test for non-autocorrelation of order one and four were accepted the $p - values$ 0.09 and 0.32. Tests for Gaussian distributed errors were rejected mainly due to excess kurtosis in the price series, and also tests for ARCH could not be rejected. However, since normality and ARCH are not crucial for the results obtained below, the model is considered well-specified for the purpose here.

7.2. Statistical analysis. To get a first insight into how many unit roots there might be in the process, we take a look at the roots of the companion form matrix. These are given in Table 4.10 for different restriction on $r = rank(\Pi)$.

[9] Testing and estimation was done in CATS in RATS, see Hansen and Juselius (1995). The $I(2)$ testing and estimation was done with routine made by Clara Jørgensen which is available from Estima's homepage at www.estima.com.

[10] The three last mentioned dummies are constructed with the same pattern as $d738$.

Table 4.10. Roots of the companion form matrix

Unrestricted	$r=2$	$r=1$
0.99	1	1
0.99	0.99	1
0.96	0.97	0.99
0.60	0.60	0.60
0.39	0.38	0.35
0.31	0.32	0.35

It seems that three unit roots are present in the process. When imposing the restriction $r = 1$ the third root moves up a bit closer to one. Three unit roots could be a case with three independent $I(1)$ trends (and hence no cointegration) or one common $I(2)$ trend and one $I(1)$ trend. The joint rank tests for the number of r and s_1 are given in Table 4.11.

Table 4.11. Rank tests for the joint hypotheses $Q(r, s_1)$

r	$Q(r, s_1)$			$Q(r)$
0	378.59 (70.87)	229.84 (54.53)	96.24 (42.91)	84.64 (34.80)
1		166.36 (36.12)	**21.81** (26.00)	**18.11** (19.99)
2			67.88 (12.93)	**6.22** (9.13)
s_2	3	2	1	0

Note: Numbers in brackets are 95% quantiles from Paruolo (1996) Table A1 and Johansen (1996) Table 15.2.

In accordance with the number of unit roots in the process the evidence from Table 4.11 suggests one $I(2)$ and one $I(1)$ trend, $(r, s_1, s_2) = (1, 1, 1)$. Given this choice the loadings from the $I(2)$ process are estimated. Normalizing the vector on the first coefficient, the estimate is

$$\widehat{\beta}_{\perp 2} = \begin{pmatrix} 1 \\ 0.65 \\ 0.07 \end{pmatrix}.$$

The coefficient for the exchange rate is very small and there seems to be no $I(2)$ness in this variable. On the other hand, both prices seem to be affected by the $I(2)$ trend. Whether the impact is the same is more questionable. To test this, we assume for a minute that the impact is the same and consider the transformed system $\widetilde{X}_t' =$

$[p^{uk} - p^{us}, exc, \Delta p^{uk}]_t$ and perform rank tests on this.[11] These are given in Table 4.12.

Table 4.12. Rank tests for the joint hyphotheses $Q(r, s_1)$

r	$Q(r,s_1)$			$Q(r)$
0	1027.72 (70.87)	327.76 (54.53)	190.04 (42.91)	114.55 (34.80)
1		246.54 (36.12)	110.90 (26.00)	43.51 (19.99)
2			138.65 (12.93)	7.70 (9.13)
s_2	3	2	1	0

Note: See Table 4.11.

Indeed it seems like the $I(2)$ness has been removed from the system. However, in the transformed system the rank tests indicates $r = 2$ cointegrating relations and not one, suggesting that the transformation could have removed two unit roots. The roots of the companion form matrix in new system are given in Table 4.13.

Table 4.13. Roots of the companion form matrix

Unrestricted	$r = 2$	$r = 1$
0.99	1	1
0.95	0.99	1
0.67	0.67	0.73
0.47	0.45	0.53
0.20	0.17	0.28
0.01	0.03	0.13

In the transformed system two unit roots seems to be present and also it seems that the $I(2)$ness has been removed. Thus, despite the rank tests it seems more likely that the transformed system includes $r = 1$ cointegrating vector and therefore that it is a fair approximation of the original $I(2)$ system. In other words, evidence from this part of the analysis is that the UK and the US have had the same underlying inflation (Definition 4B) in the period considered. Furthermore, the scenario which might describe this situation is described as case 4 in section 5.2 (case 3 if the exchange rate unexpectedly turns out to be stationary). The estimated β vector is

[11] Tests indicate no evidence of autocorrelation in the transformed system. The tests for non-autocorrelation of order one and four were accepted with $p - values$ 0.17 and 0.39.

$$\begin{array}{lcccc} & p^{uk}-p^{us} & exc & \Delta p^{uk} & const. \\ \widehat{\beta}=(& 1 & -8.6 & 884.0 & 1.5 \quad) \end{array}$$

The next step of the analysis is to determine if the PPP (maybe with adjustment) holds between the countries. Relevant hypotheses and test statistics are given in Table 4.14. The hypotheses $\mathcal{H}_1$ and $\mathcal{H}_2$ concern stationarity of the exchange rate and the inflation rate. The latter is a kind of sensitivity test to see if it was a mistake to consider the inflation rate to be $I(1)$. The two next hypotheses $\mathcal{H}_3$ and $\mathcal{H}_4$ concern Definition 5 and 6.

Table 4.14. Tests on the cointegrating vector

	p^{uk}-p^{us}	exc	Δp^{uk}	c			p-$val.$
$\mathcal{H}_1$: $sp($	0	1	0	$*$	$)=sp(\beta)$	$\chi^2(2)=42.52$	[0.00]
$\mathcal{H}_2$: $sp($	0	0	1	$*$	$)=sp(\beta)$	$\chi^2(2)=15.56$	[0.00]
$\mathcal{H}_3$: $sp($	1	-1	0	$*$	$)=sp(\beta)$	$\chi^2(2)=36.17$	[0.00]
$\mathcal{H}_4$: $sp($	1	-1	$*$	$*$	$)=sp(\beta)$	$\chi^2(1)=3.10$	[0.08]

The hypotheses $\mathcal{H}_1 - \mathcal{H}_3$ are rejected implying that indeed it was the appropriate choice to treat the system as being integrated of order 2. Furthermore, there is no evidence that the PPP holds between the UK and the US. The hypothesis $\mathcal{H}_4$ is accepted suggesting that the PPP with adjustment (Definition 4) holds between the two countries. Tests of hypotheses regarding weak exogeneity (zero rows in the α vector) suggested that the price differential and the inflation rate might be weakly exogenous. The hypotheses that both variables are weakly exogenous, however, was strongly rejected. The joint hypothesis of $\mathcal{H}_4$ and $p^{uk} - p^{us}$ weakly exogenous were accepted with $p - value = 0.10$, implying that short-run adjustment to the stationary relations is through the variables exc and Δp^{uk}. The identified cointegrating vector (with standard errors in brackets) implies the relation

$$p^{uk} = p^{us} + exc - \underset{(14.84)}{154.92}\Delta p^{uk} - \underset{(0.079)}{0.034}$$

The evidence from the empirical analysis suggests that the underlying inflation have been the same in the UK and the US and thus disagreeing with the result of Crowder (1996b). The reason for the difference can probably be attributed to the tools applied but maybe also to the longer span of data used in the present analysis. The rejection of the hypotheses of the PPP is in line with most other studies on these two countries. Finally, it does seem that the price levels have been adjusting.

7.3. Parameter stability. As discussed in section 3, empirical work including price variables might be affected by parameter instability when a longer period of time is considered. Tests for constant parameters in the cointegrated VAR have been suggested by Hansen and Johansen (1999). Here, two test are made: one that tests whether the estimated β vector has been constant and one that tests whether the non-zero eigenvalue is constant. The tests are based on recursive estimates and the first third of the period is used for the initial estimate. Hence, the recursive estimations start in January 1982 and one observation is added each time.

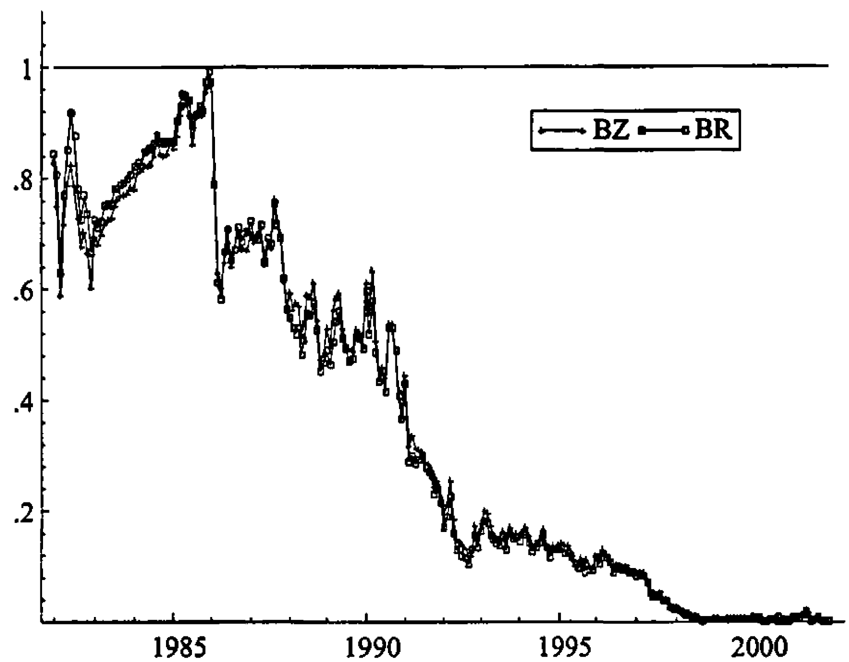


Figure 4.7. Test for β constancy.

The test for constant β vector consist of testing whether the full-sample estimate $\widehat{\beta}_T$ are contained in the all the sub-sample estimates $\widehat{\beta}_t$. The test statistic is χ^2 distributed with $(p-r)r$ degrees of freedom. The test statistic are calculated both with and without re-estimating the short-run parameters. The results are given in Figure 4.7, where the test statistics are scaled by the 5% critical value and BZ is the test statistic with re-estimation of the short-run parameters and BR is without. In none of the case does the test statistic exceed the 5%

critical value indicating the indeed the estimate of β is constant in the period.

The recursively estimated non-zero eigenvalue with 95% confidence bands is given in Figure 4.8. At no point does the point estimates of the eigenvalue exceed values higher than the minimum of the upper band or values lower than the maximum of the lower band. Hence, given the evidence from Figure 4.7 and 4.8, there is no reason to believe that the parameters should not be constant in the sample analyzed.

Figure 4.8. Recursively estimated eigenvalue.

8. Summary and conclusion

When inflation rates are non-stationary, testing of PPP relations in the cointegrated VAR model becomes more complicated. On the other hand, admitting that inflation rates are not necessary stationary gives the possibility to take advantage of the VAR model for $I(2)$ variables, which in turns can reveal more information about the price processes.

The PPP has been tested in many applications within the last couple of decades and, in general, the results have been quite negative. This might have to do with the fact, that prices are adjusting toward a sustainable level, in which case the statistical properties of the series are not necessarily in accordance with the traditional PPP requirements. When variables are $I(2)$ the multicointegrating relations can capture this phenomena and in the present chapter (and the previous one) this is defined as the PPP with adjustment. An economic interpretation is given with an optimization problem for policy makers aiming at market integration.

That inflation rates indeed might not be stationary was demonstrated in an analysis applying US data for a period of 88 years. Tests were conducted in a univariate framework with monthly, quarterly and annual observations. Strong evidence was found that parameters were not stable over the sample suggesting different regimes of inflation. Furthermore, for several of the subsamples, where parameters could be constant, it could not be rejected that the inflation rate was non-stationary.

A scenario analysis was made in order to explore different situation which can arise when testing the PPP in the cointegrating VAR model. The cases with stationary inflation rates was shown to be quite simple to handle compared to the situations where price levels were integrated of order 2. In these cases it was demonstrated how appropriate transformations could be made in order to perform estimation and testing in the $I(1)$ model. Focus was on exploration of the multicointegrating relations and interpretation with respect to underlying inflation rates and the PPP. Also an example of a scenario with three countries was provided in order to demonstrate how it could be the case, that two countries are at a sustainable PPP level whereas the third one is still adjusting.

In an empirical application evidence was found that the underlying inflation in the UK and the US was the same. Tests revealed, however, that the PPP did not hold but that prices have been adjusting to each other. Future research might focus on PPP tests allowing prices to be potentially described by $I(2)$ processes.

CHAPTER 5

$I(2)$-to-$I(1)$ transformation and deterministic terms

Abstract: This chapter is concerned with an explicit formulation for the deterministic terms in the $I(1)$ and $I(2)$ models. Explicit links are found between the models we find by transformation from $I(2)$ to $I(1)$. It is shown that the deterministics change when making the transformation. Furthermore, it is argued that testing for deterministics should be done in the original system and a procedure for simultaneously testing for presence of deterministic terms and the number of cointegrating relations is proposed.

1. Introduction

The present chapter discusses the role of the deterministic terms when transforming a cointegrated $I(2)$ vector autoregressive (VAR) model into the $I(1)$ space. Hence, this study is related to the literature discussing deterministic terms of the VAR models and transformations which decrease the order of integration from two to one.

The role of the deterministic terms in cointegrated VAR models has been discussed by other authors. Johansen (1994) discusses different restrictions for the deterministic terms in the $I(1)$ and $I(2)$ models. In the same paper, and in Johansen and Juselius (1990), it is discussed how to test for the presence of deterministic terms. Paruolo (1996) discusses how to find the number of $I(2)$, $I(1)$ and $I(0)$ trends under different assumptions of the constant term, and Rahbek et al. (1999) discuss the case of trend stationarity. Rahbek (1997) finds explicit expressions for the deterministic terms in the $I(1)$ and $I(2)$ models.

Transformations of cointegrated $I(2)$ systems into the $I(1)$ space have been applied by many authors. This has turned out to be useful, for example, when examining systems of variables including nominal variables such as prices, money, and GDP. For examples of studies making such transformations, see Juselius (1995) and Kongsted (1998).

Kongsted (2002) describes a general procedure of making "nominal-to-real" transformations and how to test it. The wording "nominal-to-real" refers to the fact, that the systems examined often include nominal variables which are $I(2)$ - for example money and prices -

and a transformation which might eliminate the $I(2)$ trend could be made in order to consider an $I(1)$ system including real money and inflation rate. In the present chapter the terminology "$I(2)$ to $I(1)$" transformations will be used, since in earlier studies, see chapter 3 and 4, it has proven useful to consider a transformation of two series of prices where the price differential is $I(1)$. Kongsted deals with the stochastic part of the VAR models when making transformation. This chapter differs from his paper by focusing on the coefficients for the deterministic terms.

In the next section the deterministic terms of the $I(1)$ model is discussed, and Section 3 deals with the deterministic terms in the $I(2)$ model. The coefficients for the linear drift and the constant are derived in all directions. These are summarized in tables, which are useful when discussing transformation of an $I(2)$ system into the $I(1)$ space. This is done in Section 4 where the relationship between the cointegrating vectors in the $I(2)$ and $I(1)$ cases are derived and simple examples are given. The role of the deterministic terms in the transformation is also discussed in this section. It turns out that there is a clear relation between the models for the deterministic terms in the $I(1)$ and the $I(2)$ cases, which leads to a proposal for a strategy of testing outlined in Section 5. Section 6 concludes the analysis.

In general the following notation used: $\alpha_\perp$ denotes an orthogonal complement such that $\alpha'\alpha_\perp = 0$, and $\overline{\alpha} = \alpha(\alpha'\alpha)^{-1}$ such that $\alpha'\overline{\alpha} = I_r$.

2. The deterministic terms in the I(1) model

We consider the cointegrated VAR(2) model in error correction form (ECM):[1]

$$\Delta X_t = \alpha\beta' X_{t-1} + \Gamma_1 \Delta X_{t-1} + \mu_t + \varepsilon_t,$$

with the data vector X_t of dimension p, and where $\alpha\beta'$ and Γ_1 includes parameters to be estimated. The matrices α and β have full rank r, and Γ_1 is a $p \times p$ matrix. It is assumed that $rank(\alpha'_\perp \Gamma \beta_\perp) = p - r$, where $\Gamma = I_p - \Gamma_1$, such that X_t is not $I(2)$. The process has errors ε_t which are *iid* with mean zero and constant variance-covariance matrix Ω, and the deterministic terms (a constant and a trend) are collected in $\mu_t : \mu_t = \mu_0 + \mu_1 t$. The parameters μ_0 and μ_1 are decomposed in the α and $\alpha_\perp$ directions:

$$\mu_i = \alpha\rho_i + \alpha_\perp\gamma_i, i = 0, 1, \tag{2.1}$$

[1]For simplicity we consider models with only $k = 2$ lags.

where ρ_i are r-vectors and γ_i are s-vectors. It follows that $\rho_i = \overline{\alpha}'\mu_i$ and $\gamma_i = \overline{\alpha}'_\perp\mu_i$. Three model for the deterministic terms are considered:[2]

$$H^*(r) : \mu_t = \mu_0 + \alpha\rho_1 t, \tag{2.2}$$

$$H_1(r) : \mu_t = \mu_0, \tag{2.3}$$

$$H_1^*(r) : \mu_t = \alpha\rho_0. \tag{2.4}$$

Hence, the following restriction are imposed on the deterministic terms: $H^*(r) : \alpha'_\perp\mu_1 = 0$; $H_1(r) : \mu_1 = 0$; and $H_1^*(r) : \mu_1 = 0$ and $\alpha'_\perp\mu_0 = 0$. The model (2.2) includes a trend restricted to the cointegrating space, (2.3) includes a unrestricted constant, and (2.4) contains a restricted constant. With these restrictions on the deterministic terms, the ECM representations can be reformulated to respectively

$$H^*(r) : \Delta X_t = \alpha(\beta', \rho_1)\begin{pmatrix} X_{t-1} \\ t \end{pmatrix} + \Gamma_1\Delta X_{t-1} + \mu_0 + \varepsilon_t,$$

$$H_1(r) : \Delta X_t = \alpha\beta' X_{t-1} + \Gamma_1\Delta X_{t-1} + \mu_0 + \varepsilon_t,$$

$$H_1^*(r) : \Delta X_t = \alpha(\beta', \rho_0)\begin{pmatrix} X_{t-1} \\ 1 \end{pmatrix} + \Gamma_1\Delta X_{t-1} + \varepsilon_t.$$

We see that the model $H_1(r)$ coincides with $H^*(r)$ for $\rho_1 = 0$. The moving average (MA) representation for the most general model, $H^*(r)$, is given by

$$H^*(r) : X_t = C\sum_{i=1}^{t}\varepsilon_i + C(L)\varepsilon_t + \tau_0 + \tau_1 t + A, \tag{2.5}$$

where $C = \beta_\perp(\alpha'_\perp\Gamma\beta_\perp)^{-1}\alpha'_\perp$ and A depends on the initial conditions and satisfies $\beta' A = 0$. In this model, r directions are stationary, $\beta' X_t \sim I(0)$, and $s = p - r$ directions are integrated of order one, $\beta'_\perp X_t \sim I(1)$. Rahbek (1997) shows that the vectors τ_0 and τ_1 are given by[3]

$$\tau_0 = \overline{\beta}\overline{\alpha}'(\Gamma C - I_p)\mu_0 + \overline{\beta}\vartheta\rho_1, \tag{2.6}$$

$$\tau_1 = (C\Gamma - I_p)\overline{\beta}\rho_1 + C\mu_0, \tag{2.7}$$

where $\vartheta = \overline{\alpha}'\Gamma C\Gamma\overline{\beta} - I_r - \overline{\alpha}'\Gamma\overline{\beta}$. As demonstrated in Appendix A, τ_0 and $\tau_1 t$ refer to that part of the deterministic terms that does not depend on the initial conditions. It should be noted that $\beta'_\perp A \neq 0$.

[2]The models are given the same names as in Johansen (1996).

[3]See also Appendix A.

First we consider the model $\mathbf{H}^*(\mathbf{r})$. Inserting μ_0 given in (2.1), we find that the coefficients for the deterministic terms are given by

$$\tau_0 = \overline{\beta}\overline{\alpha}'\Gamma C\alpha_\perp\gamma_0 - \overline{\beta}\rho_0 + \overline{\beta}\vartheta\rho_1,$$

$$\tau_1 = (C\Gamma - I_p)\overline{\beta}\rho_1 + C\alpha_\perp\gamma_0.$$

The combinations $\beta' X_t$ are trend-stationary, with the coefficients for linear trend and constant given by

$$\beta'\tau_1 = -\rho_1,$$

$$\beta'\tau_0 = \overline{\alpha}'\Gamma C\alpha_\perp\gamma_0 - \rho_0 + \vartheta\rho_1.$$

In the $I(1)$ directions, $\beta'_\perp X_t$, there are no constants, $\beta'_\perp\tau_0 = 0$, and the coefficients for the linear trends are given by

$$\beta'_\perp\tau_1 = \beta'_\perp C(\alpha_\perp\gamma_0 + \Gamma\overline{\beta}\rho_1).$$

We now turn to the model $\mathbf{H}_1(\mathbf{r})$, i.e. we impose the additional restriction $\rho_1 = 0$. The coefficients in the MA representation (2.5) are then given by

$$\tau_0 = \overline{\beta}\overline{\alpha}'\Gamma C\alpha_\perp\gamma_0 - \overline{\beta}\rho_0$$

$$\tau_1 = C\alpha_\perp\gamma_0.$$

The stationary combinations are $I(0)$ around constants given by

$$\beta'\tau_0 = \overline{\alpha}'\Gamma C\alpha_\perp\gamma_0 - \rho_0,$$

and the $I(1)$ directions have no constant, $\beta'_\perp\tau_0 = 0$, but deterministic trends with the coefficients

$$\beta'_\perp\tau_1 = \beta'_\perp C\alpha_\perp\gamma_0.$$

In the last model, $\mathbf{H}_1^*(\mathbf{r})$, where the restrictions on the deterministic terms imply $\mu_1 = 0$ and $\gamma_0 = 0$, there is no trend in any directions, $\tau_1 = 0$. The stationary directions contain constants given by

$$\beta'\tau_0 = -\rho_0,$$

whereas there are no constant in the $I(1)$ directions, $\beta'_\perp\tau_0 = 0$.

The coefficients for the deterministic terms in the $I(1)$ model are summarized in Table 5.1.

Table 5.1. Deterministic terms in the $I(1)$ model

Model	Direction	Coeff. linear trend	Constant
$H^*(r)$	$\beta' X_t \sim I(0)$	$-\rho_1$	$\overline{\alpha}'\Gamma C\alpha_\perp\gamma_0 - \rho_0 + \vartheta\rho_1$
(2.2)	$\beta'_\perp X_t \sim I(1)$	$\beta'_\perp C(\alpha_\perp\gamma_0 + \Gamma\overline{\beta}\rho_1)$	0
$H_1(r)$	$\beta' X_t \sim I(0)$	0	$\overline{\alpha}'\Gamma C\alpha_\perp\gamma_0 - \rho_0$,
(2.3)	$\beta'_\perp X_t \sim I(1)$	$\beta'_\perp C\alpha_\perp\gamma_0$	0
$H_1^*(r)$	$\beta' X_t \sim I(0)$	0	$-\rho_0$
(2.4)	$\beta'_\perp X_t \sim I(1)$	0	0

3. The deterministic terms in the I(2) model

In this section three models are considered: $H^*(r, s_1)$ which allows for a linear trend in the stationary and difference-stationary directions; $H_1(r, s_1)$ which only allows for trend in the difference-stationary directions; and $H_1^*(r, s_1)$ which allows for a constant in the stationary relations only and thus no linear trends.

In the $I(2)$ case the ECM is formulated in accelerations, differences and levels. For the cointegrated VAR(2) it is

$$\Delta^2 X_t = \alpha\beta' X_{t-1} - \Gamma\Delta X_{t-1} + \mu_0 + \mu_1 t + \varepsilon_t, \tag{3.1}$$

with $rank(\alpha'_\perp \Gamma\beta_\perp) = rank(\xi\eta') = s_1 < p - r$. Now, define $\theta = \Gamma\overline{\beta}\overline{\alpha}'\Gamma + \Gamma$, and $\alpha_{\perp 1} = \overline{\alpha}_\perp\xi$, $\alpha_{\perp 2} = \alpha_\perp\xi_\perp$, $\beta_{\perp 1} = \overline{\beta}_\perp\eta$, and $\beta_{\perp 2} = \beta_\perp\eta_\perp$ such that $sp(\alpha, \alpha_{\perp 1}, \alpha_{\perp 2}) = sp(\beta, \beta_{\perp 1}, \beta_{\perp 2}) = \mathbf{R}^p$. The condition $rank(\alpha'_{\perp 2}\theta\beta_{\perp 2}) = s_2 = p - r - s_1$ secures that $X_t \nsim I(3)$.

As in the $I(1)$ case, the term μ_0 is decomposed in the α and $\alpha_\perp$ directions:

$$\mu_i = \alpha\rho_i + \alpha_\perp\gamma_i, i = 0, 1,$$

In order to avoid accumulation to a quadratic trend, not only the linear drift but also the constant needs to be restricted. This can be accomplished with the restrictions (see appendix B)

$$\mu_1 = \alpha\rho_1, \quad \text{and} \quad \alpha'_{\perp 2}\mu_0 + \alpha'_{\perp 2}\Gamma\overline{\beta}\rho_1 = 0. \tag{3.2}$$

The ECM representations can then be written:

$$H^*(r, s_1) : \Delta^2 X_t = \alpha(\beta', \rho_1)\begin{pmatrix} X_{t-1} \\ t \end{pmatrix} - \Gamma\Delta X_{t-1} + \mu_0 + \varepsilon_t, \tag{3.3}$$

$$H_1(r, s_1) : \Delta^2 X_t = \alpha\beta' X_{t-1} - \Gamma\Delta X_{t-1} + \mu_0 + \varepsilon_t, \tag{3.4}$$

$$H_1^*(r,s_1): \Delta^2 X_t = \alpha(\beta',\rho_0)\begin{pmatrix} X_{t-1} \\ 1 \end{pmatrix} - \Gamma\Delta X_{t-1} + \varepsilon_t, \tag{3.5}$$

where μ_0 in $H^*(r,s_1)$ fulfills the requirement (3.2) and in $H_1(r,s_1)$, $\alpha'_{\perp 2}\mu_0 = 0$. The MA representation for $H^*(r,s_1)$ is given by

$$X_t = C_2 \sum_{s=1}^{t}\sum_{i=1}^{s}\varepsilon_i + C_1\sum_{i=1}^{t}\varepsilon_i + C^*(L)\varepsilon_t + \tau_0 + \tau_1 t + A + Bt, \tag{3.6}$$

where $C_2 = \beta_{\perp 2}(\alpha'_{\perp 2}\theta\beta_{\perp 2})^{-1}\alpha'_{\perp 2}$, $\beta' C_1 = \overline{\alpha}'\Gamma C_2$, $\beta'_{\perp 1}C_1 = \overline{\alpha}'_{\perp 1}(I - \theta C_2)$, $C^*(z)$ is convergent for $| z |< 1+v$, for some $v > 0$, A and B depends on initial values and satisfy $\beta' A - \overline{\alpha}'\Gamma\overline{\beta}_{\perp 2}\beta'_{\perp 2}B = 0$, and $(\beta, \beta_{\perp 1})'B = 0$.

Let $\delta = \overline{\alpha}'\Gamma\overline{\beta}_{\perp 2}$. In the $I(2)$ model, $\beta' X_t - \delta\beta'_{\perp 2}\Delta X_t \sim I(0)$. If $r > s_2$, we decompose β in the $\delta_\perp$ and δ directions, such that $\beta_0 = \beta\delta_\perp$, $\beta_1 = \beta\delta$ and $\beta = \beta_0\overline{\delta}'_\perp + \beta_1\overline{\delta}'$. We see that $\alpha\beta' = \alpha(\overline{\delta}_\perp\beta'_0 + \overline{\delta}\beta'_1) = \alpha_0\beta'_0 + \alpha_1\beta'_1$, where $\alpha_0 = \alpha\overline{\delta}_\perp$ and $\alpha_1 = \alpha\overline{\delta}$. Hence, $\beta'_0 X_t - \delta'_\perp\delta\beta'_{\perp 2}\Delta X_t = \beta'_0 X_t \sim I(0)$ and $\beta'_1 X_t - \delta'\delta\beta'_{\perp 2}\Delta X_t \sim I(0)$, i.e. $\beta'_0 X_t$ are $CI(2,0)$ relations whereas $\beta'_1 X_t$ are $C(2,1)$ but multicointegrate. Note that in general, β_1 and α_1 have dimension $p \times s_2$. This is, however, only the case if $r \geq s_2$. If $r < s_2$ the dimension is $p \times r$. Hence, the correct number of 'real' multicointegrating relations are $\min(r, s_2)$. Besides the stationary relations, s_1 directions are $CI(2,1)$, $\beta'_{\perp 1}X_t \sim I(1)$, and s_2 directions are $I(2)$, $\beta'_{\perp 2}X_2 \sim I(2)$.

As in the last section, τ_0 and $\tau_1 t$ refer to that part of the deterministic terms, which does not depend on the initial conditions. It should be noted that $\beta'_{\perp 1}A$, $\beta'_{\perp 2}A$ and $\beta'_{\perp 2}B$ might be different from zero.

Let $\kappa^* = \overline{\alpha}'(I_p - \Gamma\overline{\beta}_{\perp 1}\overline{\xi}'\alpha'_\perp)\mu_0 - (I_r + \overline{\alpha}'\Gamma\overline{\beta} - \overline{\alpha}'\Gamma\overline{\beta}_{\perp 1}\overline{\xi}'\alpha'_\perp\Gamma\overline{\beta})\overline{\alpha}'\mu_1$. In Appendix B it is shown that the coefficients for the deterministic terms in (3.6) are given by

$$\tau_0 = -\overline{\beta}\kappa^*, \tag{3.7}$$

$$\tau_1 = -\overline{\beta}_{\perp 1}\overline{\xi}'\alpha'_\perp\mu_0 + (\overline{\beta}_{\perp 1}\overline{\xi}'\alpha'_\perp\Gamma\overline{\beta} - \overline{\beta})\overline{\alpha}'\mu_1. \tag{3.8}$$

In the $I(2)$ directions it is not possible to separate the linear drift from the initial conditions, so we have imposed the restriction $\beta'_{\perp 2}\tau_1 = 0$. Furthermore, it is not possible to separate the constant from the initial values in the $I(1)$ and the $I(2)$ directions. This implies the restrictions $\beta'_{\perp 1}\tau_0 = 0$ and $\beta'_{\perp 2}\tau_0 = 0$, which are fulfilled by assumption.

Let us first consider the model $\mathbf{H}^*(\mathbf{r}, \mathbf{s}_1)$. The $r - s_2$ directly cointegrating relations have linear trends and constants with the coefficients

$$\beta'_0\tau_1 = \beta'_0\overline{\beta}\overline{\alpha}'\mu_1 = -\delta'_\perp\rho_1,$$

$$\beta_0'\tau_0 = -\beta_0'\overline{\beta}\kappa^* = -\delta_\perp'\kappa^*.$$

For the s_2 multicointegrating relations the coefficients for the deterministic trends are $\beta_1'\tau_1 = -\delta'\rho_1$ and the constants are given by

$$\beta_1'\tau_0 - \delta'\delta\beta_{\perp 2}'\tau_1 = -\beta_1'\overline{\beta}\kappa^* - 0 = -\delta'\kappa^*.$$

In the s_1 $I(1)$ directions the deterministic trends have the coefficients

$$\begin{aligned}\beta_{\perp 1}'\tau_1 &= -\overline{\xi}'\alpha_\perp'\mu_0 + \overline{\xi}'\alpha_\perp'\Gamma\overline{\beta}\overline{\alpha}'\mu_1 \\ &= -\overline{\xi}'\alpha_\perp'\alpha_\perp\gamma_0 + \overline{\xi}'\alpha_\perp'\Gamma\overline{\beta}\rho_1.\end{aligned}$$

We now turn to the model $\mathbf{H}_1(\mathbf{r}, \mathbf{s}_1)$ in which the restriction $\rho_1 = 0$ implies

$$\mu_1 = 0 \quad \text{and} \quad \alpha_{\perp 2}'\mu_0 = 0 \tag{3.9}$$

The coefficients in the MA-representation (3.6) are given by

$$\tau_0 = -\overline{\beta}\kappa_1,$$

$$\tau_1 = -\overline{\beta}_{\perp 1}\overline{\xi}'\alpha_{\perp 1}'\mu_0,$$

where $\kappa_1 = \overline{\alpha}'(I_p - \Gamma\overline{\beta}_{\perp 1}\overline{\xi}'\alpha_{\perp 1}')\mu_0$ and we have used the fact that $\alpha_\perp'\mu_0 = \alpha_{\perp 1}'\mu_0 + \alpha_{\perp 2}'\mu_0$. Neither the directly nor the multicointegrating relations contain a deterministic trend: $\beta_0'\tau_1 = \beta_1'\tau_1 = 0$. The constants in the stationary relations are given by respectively

$$\beta_0'\tau_0 = -\beta_0'\overline{\beta}\kappa_1 = -\delta_\perp'\kappa_1,$$

$$\beta_1'\tau_0 - \delta'\delta\beta_{\perp 2}'\tau_1 = -\beta_1'\overline{\beta}\kappa_1 - 0 = -\delta'\kappa_1.$$

The $I(1)$ directions have linear trends with coefficients given by

$$\beta_{\perp 1}'\tau_1 = -\overline{\xi}'\alpha_{\perp 1}'\mu_0 = -\overline{\xi}'\xi'\overline{\alpha}_\perp'\alpha_\perp\gamma_0 = -\gamma_0.$$

For the model $\mathbf{H}_1^*(\mathbf{r}, \mathbf{s}_1)$ the following restrictions are imposed:

$$\mu_1 = 0, \quad \text{and} \quad (\alpha_{\perp 1}', \alpha_{\perp 2}')\mu_0 = 0. \tag{3.10}$$

Hence we have $\mu_t = \alpha\rho_0$. The coefficients in the MA-representation (3.6) are given by

$$\tau_0 = -\overline{\beta}\overline{\alpha}'\mu_0 = -\overline{\beta}\rho_0,$$

$$\tau_1 = 0.$$

Hence, no linear trend is present.

The $CI(2,0)$ and the multicointegrating relations contain constants given by respectively $\beta_0'\tau_0 = -\delta_\perp'\rho_0$ and $\beta_1'\tau_1 = -\delta'\rho_0$.

The findings above are summarized in Table 5.2.

Table 5.2. Deterministic terms in the $I(2)$ model

Model	Direction	Coeff. lin. trend	Const.
$H^*(r,s_1)$	$\beta_0'X_t \sim I(0)$	$-\delta_\perp'\rho_1$	$-\delta_\perp'\kappa^*$
(3.2)	$\beta_1'X_t - \delta'\delta\beta_{\perp2}'\Delta X_t \sim I(0)$	$-\delta'\rho_1$	$-\delta'\kappa^*$
	$\beta_{\perp1}'X_t \sim I(1)$	$\overline{\xi}'\alpha_\perp'(\Gamma\overline{\beta}\rho_1 - \alpha_\perp\gamma_0)$	0
$H_1(r,s_1)$	$\beta_0'X_t \sim I(0)$	0	$-\delta_\perp'\kappa_1$
(3.9)	$\beta_1'X_t - \delta'\delta\beta_{\perp2}'\Delta X_t \sim I(0)$	0	$-\delta'\kappa_1$
	$\beta_{\perp1}'X_t \sim I(1)$	$-\gamma_0$	0
$H_1^*(r,s_1)$	$\beta_0'X_t \sim I(0)$	0	$-\delta_\perp'\rho_0$
(3.10)	$\beta_1'X_t - \delta'\delta\beta_{\perp2}'\Delta X_t \sim I(0)$	0	$-\delta'\rho_0$
	$\beta_{\perp1}'X_t \sim I(1)$	0	0

4. "I(2)-to-I(1)" transformation

In this section we try to answer the question "can we find relations between the stationary and non-stationary relations in the original $I(2)$ system and transformed $I(1)$ system?". And furthermore, "do the deterministic terms change when transforming the model?".

We consider the cointegrated VAR model for variables integrated of order two. In the MA representation (3.6), $\beta_{\perp2}$ represent the loadings from the common $I(2)$ trends into the process. Hence, a matrix b satisfying the condition $sp(b) = sp((\beta_{\perp2})_\perp) = sp(\beta, \beta_{\perp1})$ can be used to transform X_t into the $I(1)$ space. However, $b'X_t$ only has dimension $p - s_2$, and in order not to loose the dynamics of the original system, $v'\Delta X_t$ is added to the new one, where v is of dimension $p \times s_2$ and satisfies the condition $\mid v'\beta_{\perp2} \mid \neq 0$. Hence, the transformed system is[4]

$$\widetilde{X}_t = \begin{pmatrix} b'X_t \\ v'\Delta X_t \end{pmatrix} \sim I(1). \tag{4.1}$$

An obvious choice is to set $v = \beta_{\perp2}$. This certainly fulfill the requirement. Another - perhaps more relevant for practical purposes - is to choose v such that $sp(v) = sp(I_{s2}, 0_{p-s_2})$, where the '1's in the columns correspond to a non-zero entry in the same column in $\beta_{\perp2}$.

The ECM for the transformed system is

$$\Delta\widetilde{X}_t = \widetilde{\Pi}\widetilde{X}_{t-1} + \widetilde{\Gamma}_1\Delta\widetilde{X}_{t-1} + \widetilde{\mu}_0 + \widetilde{\mu}_1 t + \widetilde{\varepsilon}_t, \tag{4.2}$$

[4]In what follows '$\sim$'s will denote matrices related to the transformed $I(1)$ system such that, for example, $\widetilde{\beta}'\widetilde{X}_t \sim I(0)$.

where the explicit expressions for $\widetilde{\Pi}$, $\widetilde{\Gamma}_1$, $\widetilde{\mu}_0$, $\widetilde{\mu}_1$ and $\widetilde{\varepsilon}_t$ are given in appendix C (see also Kongsted, 2002). Since there is an extra lag in the transformed model, an implicit restriction of zeros is put on $\widetilde{\Gamma}_1$.

To find a relationship between the cointegrating vectors in the $I(2)$ and the $I(1)$ systems, we first have to find the coefficients for $v'\Delta X_t$. As we are considering an $I(2)$ model, $s_2 > 0$. Thus, if cointegration exists, at least one of the relations will be multicointegrating. The coefficients for the difference part in the $I(1)$ model are given by $\omega = -(\overline{\beta}'_{\perp 2}v)^{-1}\delta'$. To see this consider the stationary relations $\beta' X_t - \delta\beta'_{\perp 2}\Delta X_t \sim I(0)$. We want to rewrite this to include the terms $b'X_t$ and $v'\Delta X_t$.

We can write

$$\beta_{\perp 2} = va_0 + ba_1, \tag{4.3}$$

where $sp(b) = sp(\beta, \beta_{\perp 1})$ and a_i $(i = 1, 2)$ are matrices of suitable dimensions. Premultiplying (4.3) with $\beta'_{\perp 2}$ we find

$$\begin{aligned} \beta'_{\perp 2}\beta_{\perp 2} &= \beta'_{\perp 2}va_0 \\ \Rightarrow \qquad a_0 &= (\beta'_{\perp 2}v)^{-1}\beta'_{\perp 2}\beta_{\perp 2} = (\overline{\beta}'_{\perp 2}v)^{-1}. \end{aligned}$$

Since $b = (\beta, \beta_{\perp 1})$, $\beta = \overline{b}b'\beta$ so that the stationary relations are given by $\beta'\overline{b}b'X_t - \delta(a_0'v' + a_1'b')\Delta X_t$. Since $b'\Delta X_t \sim I(0)$, we have

$$\beta'\overline{b}b'X_t - \delta(v'\overline{\beta}_{\perp 2})^{-1}v'\Delta X_t \sim I(0).$$

Hence, the cointegrating relations in the transformed system are given by

$$\widetilde{\beta} = \begin{pmatrix} \overline{b}'\beta \\ \omega \end{pmatrix}, \tag{4.4}$$

where $\omega = -(\overline{\beta}'_{\perp 2}v)^{-1}\delta'$. If $r > s_2$, we can multiply (4.4) with $(\delta, \delta_\perp)$ from the right, which gives us

$$\widetilde{\beta}(\delta, \delta_\perp) = \begin{pmatrix} \overline{b}'\beta_1 & \overline{b}'\beta_0 \\ -(\overline{\beta}'_{\perp 2}v)^{-1}\delta'\delta & 0 \end{pmatrix}. \tag{4.5}$$

Before turning to the non-stationary directions, let us consider a couple of examples.

Example 1: Consider a system of dimension $p = 3$ with $(r, s_1, s_2) = (1, 1, 1)$ and let $\beta'_{\perp 2} = (1, 1, 1)'$, i.e. the impact from the common $I(2)$ trend is the same on all variables. That means, that the impact from the common $I(1)$ trend on ΔX_t is the same on all variables too. The matrix δ' is of dimension 1×1 and we set $\delta' = d$. First, choose $v = \beta_{\perp 2}$. Then $v'\Delta X_t = \Delta X_{1t} + \Delta X_{2t} + \Delta X_{3t}$. Furthermore, $a_0 = 1$ such that $\omega = -d$.

Now choose $v = (1,0,0)'$ such that $v'\Delta X_t = \Delta X_{1t}$. We find that $a_0 = 3$ such that $\omega = -3d$.

Example 2: Let again $p = 3$ but now $(r, s_1, s_2) = (1, 0, 2)$. Let the impacts from the two common $I(2)$ trends be given by

$$\beta_{\perp 2} = \begin{pmatrix} 2 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix},$$

i.e. $X_{1t} - 2X_{2t} - X_{3t} \sim I(1)$ so that $\Delta X_{1t} - 2\Delta X_{2t} - \Delta X_{3t} \sim I(0)$ In this case $s_2 > r$ and δ' has dimension 2×1. We set $\delta' = (d_1, d_2)'$

First, choose $v = \beta_{\perp 2}$ such that the variables entering the transformed system are $2\Delta X_{1t} + \Delta X_{2t}$ and $\Delta X_{1t} + \Delta X_{3t}$. We find that $a_0 = I_{s_2}$ and $\omega = -(d_1, d_2)'$.

Now choose

$$v' = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

so that $X_{1t} - 2X_{2t} - X_{3t} - (5d_1 + 2d_2)\Delta X_{2t} - 2(d_1 + d_2)\Delta X_{3t}$ is stationary.

To see if we can find a relationship between the non-stationary directions, we first note that the non-stationary relations in the transformed system are given by $\widetilde{\beta}_\perp' X_t \sim I(1)$, where $\widetilde{\beta}_\perp$ has dimension $p \times s$ $(s = s_1 + s_2)$. Let the general $\widetilde{\beta}_\perp$ be given by

$$\widetilde{\beta}_\perp = \begin{pmatrix} A \\ B \end{pmatrix},$$

where A has dimension $(p - s_2) \times s$ and B has dimension $s_2 \times s$. Then A and B have to fulfill the requirements

$$\beta_0'\bar{b}A = 0_{(r-s_2)\times s} \tag{4.6}$$

$$\beta_1'\bar{b}A + \delta'\omega' B = 0_{s_2\times s} \tag{4.7}$$

The solution of (4.6) is $A = b'(\overline{\beta}\delta, \beta_{\perp 1})$. Then we see that (4.6) is fulfilled, $\beta_0'\bar{b}b'(\overline{\beta}\delta, \beta_{\perp 1}) = \delta_\perp'\beta'(\overline{\beta}\delta, \beta_{\perp 1}) = 0$. Inserting A in (4.7) yields

$$\begin{aligned} & \delta'\beta'\bar{b}b'(\overline{\beta}\delta, \beta_{\perp 1}) - \delta'\delta(v'\overline{\beta}_{\perp 2})^{-1}B = 0 \\ \iff\ & \delta'\delta(I_{s_2}, 0_{s_2\times s_1}) - \delta'\delta(v'\overline{\beta}_{\perp 2})^{-1}B = 0 \\ \iff\ & B = v'\overline{\beta}_{\perp 2}v(I_{s_2}, 0_{s_2\times s_1}). \end{aligned}$$

Hence, we find that the relationship between the non-stationary relations in the original and the transformed system is given by

$$\widetilde{\beta}_\perp = \begin{pmatrix} b'\overline{\beta}\delta & b'\beta_{\perp 1} \\ v'\overline{\beta}_{\perp 2} & 0_{s_2\times s_1} \end{pmatrix}. \tag{4.8}$$

We see that indeed $\widetilde{\beta}'\widetilde{\beta}_\perp = 0$ and we have found the link between the stationary and the non-stationary relations when making a transformation from the $I(2)$ space into the $I(1)$ space. Note that in (4.8) the $CI(2,1)$ relations given by $\beta'_{\perp 1}X_t$ in the original system are maintained in $\widetilde{\beta}'\widetilde{X}_t$ in the transformed model.

Before investigating the role of the deterministic terms, we consider a couple of practical examples, to illustrate the results given above.

In the following the notation given in (4.1) is used. Hence $X_t \sim I(2)$, $\widetilde{X}_t, \Delta X_t \sim I(1)$, and $\Delta\widetilde{X}_t, \Delta^2 X_t \sim I(0)$.

4.1. Stylized examples. The following - very simple - examples are inspired by the one in Kongsted (2002). Consider a vector consisting of three nominal time series: money stock, nominal income, and the price level, $X_t = \{m_t, y_t, p_t\}$, and all variables are assumed to be $I(2)$ with the same impact, $\beta_{\perp 2} = (1,1,1)'$. Two cases are considered: one with one multicointegrating relation ($r = 1$), and one with one directly cointegrating as well as one multicointegrating relation ($r = 2$).

Case 1: $(r, s_1, s_2) = (1,1,1)$. The chosen transformation is made with the matrices

$$b = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -1 & -1 \end{pmatrix}, \; v = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \tag{4.9}$$

such that $\widetilde{X}_t = \{m_t - p_t, y_t - p_t, \Delta p_t\} \sim I(1)$. Note that in the $I(1)$ system $(r, s) = (1,2)$. Let us assume that the multicointegrating relation is the inverse velocity with inflation, $m_t - y_t + d\Delta p_t \sim I(0)$. Hence, we have

$$\beta = \beta_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \; \beta_{\perp 1} = \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}.$$

Expressed in terms of the transformed system we find

$$\widetilde{\beta} = \begin{pmatrix} 1 \\ -1 \\ 3d \end{pmatrix}, \; \widetilde{\beta}_\perp = \begin{pmatrix} d/2 & 3 \\ -d/2 & 3 \\ 1/3 & 0 \end{pmatrix}.$$

Hence the matrix used for the transformation of the cointegrating relations is

$$\bar{b} = \begin{pmatrix} 2/3 & -1/3 \\ -1/3 & 2/3 \\ -1/3 & -1/3 \end{pmatrix}. \tag{4.10}$$

Case 2: $(r, s_1, s_2) = (2, 0, 1)$. Again the matrices b and v given in (4.9) are used to make the transformation. Let us assume that the multicointegrating relation is a relation between real money and the inflation rate, and that the real GDP is (perhaps trend-) stationary. I.e. $m_t - p_t + \kappa \Delta p_t \sim I(0)$ and $y_t - p_t \sim I(0)$. The cointegration relations in this system can be expressed by the matrices

$$\beta = (\beta_0, \beta_1) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \\ -1 & -1 \end{pmatrix}, \beta_{\perp 1} = 0.$$

We calculate

$$b'\overline{\beta} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \overline{\beta}'_{\perp 2} v = 1/3,$$

and with $\delta' = (d_1, d_2)'$, we can use (4.4), (4.8), and the transformation matrix $\overline{b}$ given in (4.10) to find the matrices in the transformed system

$$\widetilde{\beta} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \\ -3d_1 & -3d_2 \end{pmatrix}, \widetilde{\beta}_\perp = \begin{pmatrix} d_2 \\ d_1 \\ 1/3 \end{pmatrix}.$$

Note that in this example where $r > s_2$ we can decompose $\widetilde{\beta}$ in the directions of $\delta_\perp$ and δ as done in (4.5). Thus it follows that $d_1 = 0$ such that the first column in $\widetilde{\beta}$ contains the $C(2,0)$ relation from the original system.

4.2. The deterministic terms. In this section we investigate the role of the deterministic term in the transformation from the $I(2)$ space to the $I(1)$ space.

In the MA-representation of the $I(2)$ model given in (3.6), we see that ΔX_t contains a constant, τ_1, given in (3.8). However, ΔX_t has no drift such that implicit zero restrictions are put on the coefficients for the trends in the transformed model. Furthermore, the last s_2 columns in the matrix $\widetilde{\Gamma}_1$ in (4.2) are zeros, and the restriction (3.2), carried over from the original model, implies a seemingly unnecessary restriction on the parameters in the transformed system (see Appendix C). Because of this, the transformed model is not a 'star' model, and testing for rank and presence of deterministics should be done in the original model. It is, however, an approximate 'star' model, and will be treated as such in the following.

We consider the most general model, $H^*(r, s_1)$. As argued above and in Appendix A, the deterministics in the MA-representation are given by

$$\widetilde{\tau}_0 = \widetilde{\widetilde{\beta}\widetilde{\alpha}}'(\widetilde{\Gamma}\widetilde{C} - I_p)\widetilde{\mu}_0 + \widetilde{\widetilde{\beta}\vartheta\widetilde{\rho}}_1,$$

$$\widetilde{\tau}_1 = (\widetilde{C}\widetilde{\Gamma} - I_p)\widetilde{\widetilde{\beta}}\widetilde{\rho}_1 + \widetilde{C}\widetilde{\mu}_0.$$

In Appendix D these are expressed in terms of the parameters of the original $I(2)$ model. We find that the coefficients for the trends in the stationary directions are given by

$$\widetilde{\beta}'\widetilde{\tau}_1 = -\widetilde{\beta}'\widetilde{\widetilde{\beta}}\widetilde{\rho}_1 = \rho_1,$$

which are the same as in the original system. The constant terms, however are different. They are given by

$$\widetilde{\beta}'\widetilde{\tau}_0 = \widetilde{\beta}'\widetilde{\widetilde{\alpha}}'(\widetilde{\Gamma}\widetilde{C} - I_p)\widetilde{\mu}_0 + \widetilde{\beta}'\widetilde{\vartheta}\widetilde{\rho}_1,$$

where the $\widetilde{\beta}$ is given in (4.4) and the individual terms of $\widetilde{\widetilde{\beta}\widetilde{\alpha}}'\widetilde{\Gamma}\widetilde{C}\widetilde{\mu}_0$, $\widetilde{\widetilde{\beta}\widetilde{\alpha}}'\widetilde{\mu}_0$, and $\widetilde{\widetilde{\beta}\vartheta}\widetilde{\rho}_1$ can be found in appendix D. Also the coefficients for the trends in the non-stationary directions are different from those in the original model. They are given by

$$\widetilde{\beta}_\perp'\widetilde{\tau}_1 = \widetilde{\beta}_\perp'\widetilde{C}\widetilde{\Gamma}\widetilde{\widetilde{\beta}}\widetilde{\rho}_1 + \widetilde{\beta}_\perp'\widetilde{C}\widetilde{\mu}_0,$$

where $\widetilde{\beta}_\perp'$ is given in (4.8) and $\widetilde{C}\widetilde{\Gamma}\widetilde{\widetilde{\beta}}\widetilde{\rho}_1$ and $\widetilde{C}\widetilde{\mu}_0$ are given in appendix D.

Hence, it can be concluded that when transforming an $I(2)$ model into the $I(1)$ space, also the deterministics are transformed. In other words, the 'cost' of the transformation is that the coefficients for the deterministic terms change. This should be taken into account if the interpretation of the deterministics is based on the $I(2)$ system. Note, however, that the restriction $\alpha_\perp'\mu_1 = 0$ is invariant to the transformation since $\widetilde{\alpha}_\perp'\widetilde{\mu}_1 = \alpha_\perp'(b,v)'^{-1}(b,v)'\mu_1 = \alpha_\perp'\mu_1$. Also the restrictions in the models H_1 and H_1^* are invariant since $\mu_1 = 0 \Rightarrow \widetilde{\mu}_1 = (b,v)'\mu_1 = 0$ and $\widetilde{\alpha}_\perp'\widetilde{\mu}_0 = \alpha_\perp'(b,v)'^{-1}(b,v)'\mu_0 = \alpha_\perp'\mu_0$. This means that the transformation of an $H^*(r,s_1)$ model is (approximate) $H^*(r)$, the model $H_1(r,s_1)$ becomes $H_1(r)$, and $H_1^*(r,s_1)$ becomes $H_1^*(r)$.

Because of the restrictions on the transformed system, carried over from the $I(2)$ model, it is suggested that testing for the model of the deterministic terms should be done in the $I(2)$ model, i.e. before transforming the system into the $I(1)$ space. Since the critical values for the rank tests depend on the presence of the deterministic terms, determination of rank and deterministics should be done simultaneously. A simple strategy is proposed in the next section.

5. Testing hypotheses for the deterministic terms in the $I(2)$ model

Rahbek et al. (1999) show that the asymptotic distribution of the cointegration parameters is mixed Gaussian such that given (r, s_1) inference on these are asymptotically χ^2 distributed. In the models considered in this analysis, the deterministic terms are restricted to the α space such that tests for significant coefficients can be performed on the β matrix.

The following strategy for testing is proposed: **(1)** Assume the model is $H^*(r, s_1)$ and find r and s_1; **(2)** Given $(r, s_1) = (r^*, s_1^*)$ test the hypothesis that the model is $H_1(r^*, s_1^*)$ against the alternative of $H^*(r^*, s_1^*)$; **(3)** If the hypothesis in step (2) is accepted, determine $(r, s_1) = (\overline{r}, \overline{s_1})$ given the model is $H_1(r, s_1)$; **(4)** Test the hypothesis that the model is $H_1^*(\overline{r}, \overline{s_1})$ against the alternative of $H_1(\overline{r}, \overline{s_1})$; **(5)** If the hypothesis in step (4) is accepted, determine (r, s_1) given the model is $H_1^*(r, s_1)$.[5]

Test of hypotheses for the number of cointegrating relations and common stochastic $I(1)$ trends are outlined in Rahbek et al. (1999) for the case of $H^*(r, s_1)$ and in Paruolo (1996) for the models $H_1(r, s_1)$ and $H_1^*(r, s_1)$. Simulations of critical values are also supplied in these papers.

Procedures for testing different models for the deterministic terms in the $I(1)$ model are given by Johansen and Juselius (1990) and Johansen (1994). Since - as mentioned above - the distribution in the $I(2)$ model is mixed Gaussian too, the standard likelihood ratio tests can be applied. The hypothesis of model $H_1(r^*, s_1^*)$ against the alternative of $H^*(r^*, s_1^*)$ can be formulated as $\mathcal{H}_0 : \rho_1' = 0$ against $\mathcal{H}_a : \rho_1' \neq 0$. From the representation (3.3) we see that this is a test on the β matrix and the likelihood ratio test statistic is χ^2 distributed with r degrees of freedom. The hypothesis of model $H_1^*(\overline{r}, \overline{s_1})$ against the alternative of $H_1(\overline{r}, \overline{s_1})$ can be tested with the test statistic

$$LR = -T \sum_{i=r+1}^{p} \ln \left[\frac{1 - \widehat{\lambda}_i^*}{1 - \widehat{\lambda}_i} \right],$$

where $\widehat{\lambda}_i^*$ and $\widehat{\lambda}_i$ are the $p - r$ smallest eigenvalues in the eigenvalue problem arising from the reduced rank regression which can be used

[5]Note that the model $H^*(r, s_1)$ applies for all values of ρ_1 and ρ_2. Hence, one can expect to find $(r^*, s_1^*) = (\overline{r}, \overline{s}_1)$. I think, however, it is a good idea to follow the steps to confirm the results from the first rank test, especially if it turns out to be a borderline case.

to estimate the model.[6] The test statistic is χ^2 distributed with $p-r$ degrees of freedom.

6. Conclusion

The purpose of this analysis was to discuss the role of the deterministic terms when making transformations from the $I(2)$ to the $I(1)$ space. Explicit formulations for the drift and the constant were found for the $I(1)$ as well as the $I(2)$ model in stationary and non-stationary directions. These were used to discuss the role of the deterministics when making transformations.

Also the $I(2)$-to-$I(1)$ transformation was discussed and explicit links between stationary and non-stationary directions were found. It turned out, that there exist matrices by which one can go from one space to the other - and back again - supporting the view that if a valid transformation can be made, then there is no loss of information and inference on the stationary relations and the adjustment coefficients can be made in the $I(1)$ model. Also links between the deterministic models in the two spaces were found. It was shown that when making a transformation also the deterministic terms are transformed. It was argued, that testing for presence of deterministics should be done in the original $I(2)$ system, as implicit restrictions are put on the transformed model, such that this is not a 'star' model.

Based on these observations a strategy for simultaneous determination of the presence of deterministic terms and the number of cointegrating relations was proposed.

7. Appendix A

The following is based on Johansen (1992) and Rahbek (1997). We consider the cointegrated VAR(k) model for $I(1)$ variables in the ECM form:

$$\Delta X_t = \alpha\beta' X_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} + \mu_0 + \mu_1 t + \varepsilon_t, \tag{7.1}$$

and let $\Gamma = I_p - \sum_{i=1}^{k-1} \Gamma_i$. The characteristic polynomial for (7.1) is given by

$$A(z) = I_p(1-z) - \alpha\beta' z - \sum_{i=1}^{k-1} \Gamma_i (z^i - z^{i+1}). \tag{7.2}$$

The analysis is based on the following assumption:

[6] See Johansen (1996) for details.

Assumption: *Assume that the roots of $A(z)$ are either one or lie outside the unit circle.*

Hence, we consider only systems which do not have explosive roots. With (7.2) in hand, we find that

$$\dot{A}(z) = \frac{d}{dz}A(z) = -I_p - \alpha\beta' - \sum_{i=1}^{k-1}\Gamma_i i z^{i-1} + \sum_{i=1}^{k-1}\Gamma_i(i+1)z^i,$$

$$\ddot{A}(z) = \frac{d^2}{dz^2}A(z) = \sum_{i=1}^{k-1}\Gamma_i(i-i^2)z^{i-2} + \sum_{i=1}^{k-1}\Gamma_i(i^2+i)z^{i-1}.$$

The expansion of $A(z)$ around $z=1$ is

$$\begin{aligned} & -\alpha\beta'X_t - (-\alpha\beta' - \Gamma)\Delta X_t + \Phi\Delta^2 X_t + A^*(L)\Delta^3 X_t = e_t \\ \Leftrightarrow\quad & -\alpha\beta'X_t + (\alpha\beta' + \Gamma)\Delta X_t + \Phi\Delta^2 X_t + A^*(L)\Delta^3 X_t = e_t, \end{aligned} \tag{7.3}$$

where $\Phi = \sum_{i=1}^{k-1} i\Gamma_i$, $e_t = \mu_0 + \alpha\rho_1' t + \varepsilon_t$ and $A^*(z)$ is a suitable polynomial.

Premultiplying (7.3) by $\overline{\alpha}'$ and $\overline{\alpha}_\perp'$ respectively yields

$$-\beta'X_t + \beta'\Delta X_t + \overline{\alpha}'\Gamma\Delta X_t + \overline{\alpha}'\Phi\Delta^2 X_t + \overline{\alpha}'A^*(L)\Delta^3 X_t = \overline{\alpha}'e_t, \tag{7.4}$$

$$\overline{\alpha}_\perp'\Gamma\Delta X_t + \overline{\alpha}_\perp'\Phi\Delta^2 X_t + \overline{\alpha}_\perp'A^*(L)\Delta^3 X_t = \overline{\alpha}_\perp' e_t. \tag{7.5}$$

Now, define the process

$$\begin{pmatrix} Y_t \\ U_t \end{pmatrix} = \begin{pmatrix} \beta'X_t \\ \beta_\perp'\Delta X_t \end{pmatrix} \sim I(0).$$

We use the fact that $\Delta X_t = \overline{\beta}\beta'\Delta X_t + \overline{\beta}\beta_\perp'\Delta X_t$ and rewrite (7.4) and (7.5)

$$-Y_t + \Delta Y_t + \overline{\alpha}'\Gamma\overline{\beta}\Delta Y_t + \overline{\alpha}'\Gamma\overline{\beta}_\perp U_t + \overline{\alpha}'\Phi\overline{\beta}_\perp\Delta U_t + \ldots = \overline{\alpha}'e_t, \tag{7.6}$$

$$\overline{\alpha}_\perp'\Gamma\overline{\beta}\Delta Y_t + \overline{\alpha}_\perp'\Gamma\overline{\beta}_\perp U_t + \overline{\alpha}_\perp'\Phi\overline{\beta}_\perp\Delta U_t + \ldots = \overline{\alpha}_\perp' e_t, \tag{7.7}$$

where '...' refers to terms which are unimportant in the following. Written in matrix form we see that (7.6) and (7.7) is the expansion of the characteristic polynomial for (Y_t, U_t), named $B(z)$, around $z=1$:

$$\begin{aligned} & \begin{pmatrix} -I_r & \overline{\alpha}'\Gamma\overline{\beta}_\perp \\ 0 & \overline{\alpha}_\perp'\Gamma\overline{\beta}_\perp \end{pmatrix}\begin{pmatrix} Y_t \\ U_t \end{pmatrix} + \begin{pmatrix} I_r + \overline{\alpha}'\Gamma\overline{\beta} & \overline{\alpha}'\Phi\overline{\beta}_\perp \\ \overline{\alpha}_\perp'\Gamma\overline{\beta} & \overline{\alpha}_\perp'\Phi\overline{\beta}_\perp \end{pmatrix}\begin{pmatrix} \Delta Y_t \\ \Delta U_t \end{pmatrix} \\ & + B^*(L)\begin{pmatrix} \Delta^2 Y_t \\ \Delta^2 U_t \end{pmatrix} = (\overline{\alpha}, \overline{\alpha}_\perp)'e_t = B(L)\begin{pmatrix} Y_t \\ U_t \end{pmatrix}. \end{aligned}$$

Since $X_t \sim I(1)$, $\alpha_\perp' \Gamma \beta_\perp$ has full rank and $B(z)$ is invertible (see Johansen, 1992). Hence, $(Y_t, U_t)' = B(z)^{-1} e_t$. We find that

$$B(1)^{-1} = \begin{pmatrix} -I_r & \overline{\alpha}'\Gamma\overline{\beta}_\perp(\overline{\alpha}'\Gamma\overline{\beta}_\perp)^{-1} \\ 0 & (\overline{\alpha}_\perp'\Gamma\overline{\beta}_\perp)^{-1} \end{pmatrix}$$

$$\dot{B}(1)^{-1} = -B(1)^{-1}\dot{B}(1)B(1)^{-1} = -\begin{pmatrix} \dot{B}(1)_{11}^{-1} & \dot{B}(1)_{21}^{-1} \\ \dot{B}(1)_{12}^{-1} & \dot{B}(1)_{22}^{-1} \end{pmatrix},$$

where

$$\begin{aligned}
\dot{B}(1)_{11}^{-1} &= I_r + \overline{\alpha}'\Gamma\overline{\beta} - \overline{\alpha}'\Gamma\overline{\beta}_\perp(\overline{\alpha}_\perp'\Gamma\overline{\beta}_\perp)^{-1}\overline{\alpha}_\perp'\Gamma\overline{\beta} \\
\dot{B}(1)_{12}^{-1} &= -(\overline{\alpha}_\perp'\Gamma\overline{\beta}_\perp)\overline{\alpha}_\perp'\Gamma\overline{\beta} \\
\dot{B}(1)_{21}^{-1} &= -(\overline{\alpha}'\Gamma\overline{\beta}_\perp)(\overline{\alpha}_\perp'\Gamma\overline{\beta}_\perp)^{-1} - (\overline{\alpha}'\Gamma\overline{\beta})(\overline{\alpha}'\Gamma\overline{\beta}_\perp)(\overline{\alpha}_\perp'\Gamma\overline{\beta}_\perp)^{-1} \\
&\quad -(\overline{\alpha}'\Phi\overline{\beta}_\perp)(\overline{\alpha}_\perp'\Gamma\overline{\beta}_\perp)^{-1} + (\overline{\alpha}'\Gamma\overline{\beta}_\perp)(\overline{\alpha}_\perp'\Gamma\overline{\beta}_\perp)^{-1} \\
&\quad [(\overline{\alpha}_\perp'\Gamma\overline{\beta})(\overline{\alpha}'\Gamma\overline{\beta}_\perp)(\overline{\alpha}_\perp'\Gamma\overline{\beta}_\perp)^{-1} + (\overline{\alpha}_\perp'\Phi\overline{\beta}_\perp)(\overline{\alpha}_\perp'\Gamma\overline{\beta}_\perp)^{-1}] \\
\dot{B}(1)_{22}^{-1} &= (\overline{\alpha}_\perp'\Gamma\overline{\beta}_\perp)^{-1}(\overline{\alpha}_\perp'\Gamma\overline{\beta}\overline{\alpha}'\Gamma\overline{\beta}_\perp + \overline{\alpha}_\perp'\Phi\overline{\beta}_\perp)(\overline{\alpha}_\perp'\Gamma\overline{\beta}_\perp)^{-1}.
\end{aligned}$$

We can always write $X_t = \sum_{i=1}^t \Delta X_i + X_0$. Decomposing this in the $\overline{\beta}\beta'$ and $\overline{\beta}_\perp\beta_\perp'$ directions, we find

$$\begin{aligned}
X_t &= \overline{\beta}\beta' X_t + \overline{\beta}_\perp\beta_\perp' \textstyle\sum_{i=1}^t \Delta X_t + \overline{\beta}_\perp\beta_\perp' X_0 \\
&= \overline{\beta}Y_t + \textstyle\sum_{i=1}^t \overline{\beta}_\perp U_t + \overline{\beta}_\perp\beta_\perp' X_0.
\end{aligned} \tag{7.8}$$

We can now find the deterministic terms in the $\overline{\beta}Y_t$ directions:

$$\begin{aligned}
&(\overline{\beta}, 0)\left[B(1)^{-1}(\overline{\alpha}, \overline{\alpha}_\perp)'(\mu_0 + \alpha\rho_1' t) - \dot{B}(1)^{-1}(\overline{\alpha}, \overline{\alpha}_\perp)'\alpha\rho_1'\right] \\
&= \overline{\beta}(-\overline{\alpha}' + \overline{\alpha}'\Gamma\overline{\beta}_\perp(\overline{\alpha}'\Gamma\overline{\beta}_\perp)^{-1}\overline{\alpha}_\perp')(\mu_0 + \alpha\rho_1' t) \\
&\quad -\overline{\beta}(I_r + \overline{\alpha}'\Gamma\overline{\beta} - \overline{\alpha}'\Gamma\overline{\beta}_\perp(\overline{\alpha}_\perp'\Gamma\overline{\beta}_\perp)^{-1}\overline{\alpha}_\perp'\Gamma\overline{\beta})\overline{\alpha}'\alpha\rho_1' \\
&= \overline{\beta}\overline{\alpha}'(-I_p + \Gamma C)\mu_0 + \overline{\beta}\rho_1' t + \overline{\beta}(-I_r - \overline{\alpha}'\Gamma\overline{\beta} + \overline{\alpha}'\Gamma C\Gamma\overline{\beta})\rho_1',
\end{aligned} \tag{7.9}$$

where we have used the fact that $sp(\overline{\alpha}) = sp(\alpha)$ for all variables such that $sp(\overline{\beta}_\perp(\overline{\alpha}_\perp'\Gamma\overline{\beta}_\perp)^{-1}\overline{\alpha}_\perp') = sp(\beta_\perp(\alpha_\perp'\Gamma\beta_\perp)^{-1}\alpha_\perp') = sp(C)$. The same way we can find the deterministic terms in the $\overline{\beta}_\perp U_t$ directions:

(7.10)

$$\begin{aligned}
&(0, \overline{\beta}_\perp)\left[B(1)^{-1}(\overline{\alpha}, \overline{\alpha}_\perp)'(\mu_0 + \alpha\rho_1' t) - \dot{B}(1)^{-1}(\overline{\alpha}, \overline{\alpha}_\perp)'\alpha\rho_1'\right] \\
&= \overline{\beta}_\perp(\overline{\alpha}_\perp'\Gamma\overline{\beta}_\perp)^{-1}\overline{\alpha}_\perp'(\mu_0 + \alpha\rho_1' t)) + \overline{\beta}_\perp[(\overline{\alpha}_\perp'\Gamma\overline{\beta}_\perp)^{-1}\overline{\alpha}_\perp'\Gamma\overline{\beta}\overline{\alpha}' \\
&\quad +(\overline{\alpha}_\perp'\Gamma\overline{\beta}_\perp)^{-1}(\overline{\alpha}_\perp'\Gamma\overline{\beta}\overline{\alpha}'\Gamma\overline{\beta}_\perp + \overline{\alpha}_\perp'\Phi\overline{\beta}_\perp)(\overline{\alpha}_\perp'\Gamma\overline{\beta}_\perp)^{-1}\overline{\alpha}_\perp']\alpha\rho_1' \\
&= C\mu_0 - C\Gamma\overline{\beta}\rho_1',
\end{aligned}$$

Inserting (7.9) and (7.10) in (7.8) we find that the deterministic terms are given by (2.6) and (2.7). Also we see that $A = \overline{\beta}_\perp\beta_\perp' X_0$ in the MA representation (2.5) such that $\beta' A = 0$.

8. Appendix B

We consider the ECM form of the cointegrated $I(2)$ model (also given in (3.1))

$$\Delta^2 X_t = \alpha\beta' X_{t-1} - \Gamma\Delta X_t + \mu_0 + \mu_1 t + \varepsilon_t, \tag{8.1}$$

and set

$$X_t = \tau_0 + \tau_1 t + Y_t, \tag{8.2}$$

where

$$\Delta^2 Y_t = \alpha\beta' Y_{t-1} - \Gamma\Delta Y_{t-1} + \varepsilon_t. \tag{8.3}$$

We see from (8.2) that $Y_{t-1} = X_{t-1} - \tau_0 - \tau_1(t-1)$, $\Delta Y_{t-1} = \Delta X_{t-1} - \tau_1$, and $\Delta^2 Y_t = \Delta X_t$. Inserting this in (8.3) we find

$$\Delta^2 X_t = \alpha\beta' X_{t-1} - \alpha\beta'(\tau_0 - \tau_1 + \tau_1 t) - \Gamma\Delta X_{t-1} + \Gamma\tau_1 + \varepsilon_t,$$

and comparing with (8.1) that

$$\mu_0 = -\alpha\beta'(\tau_0 - \tau_1) + \Gamma\tau_1, \tag{8.4}$$

$$\mu_1 = -\alpha\beta'\tau_1. \tag{8.5}$$

To solve these two equations we need to assume $\alpha'_\perp\mu_1 = 0$. Hence we can write $\mu_1 = \alpha\rho_1$, where ρ_1 is a r-vector, which is the first restriction in (3.2). We find when premultiplying (8.5) with $\overline{\alpha}'$ and $\alpha'_\perp$ that

$$\overline{\alpha}'\mu_1 = -\beta'\tau_1. \tag{8.6}$$

Premultiplying (8.4) with $\overline{\alpha}'$ and $\alpha'_\perp$ yields

$$\overline{\alpha}'\mu_0 = -\beta'(\tau_0 - \tau_1) + \overline{\alpha}'\Gamma\tau_1 \tag{8.7}$$

$$\alpha'_\perp\mu_0 = \alpha'_\perp\Gamma\tau_1. \tag{8.8}$$

Decomposing τ_1 in the $\overline{\beta}\beta'$ and $\overline{\beta}_\perp\beta'_\perp$ directions and using (8.6) we get

$$\tau_1 = \overline{\beta}\beta'\tau_1 + \overline{\beta}_\perp\beta'_\perp\tau_1 = -\overline{\beta}\overline{\alpha}'\mu_1 + \overline{\beta}_\perp\beta'_\perp\tau_1, \tag{8.9}$$

Inserting (8.9) in (8.7) and (8.8) we find

$$\overline{\alpha}'\mu_0 = -\beta'\tau_0 - \overline{\alpha}'\mu_1 - \overline{\alpha}'\Gamma\overline{\beta}\overline{\alpha}'\mu_1 + \overline{\alpha}'\Gamma\overline{\beta}_\perp\beta'_\perp\tau_1 \tag{8.10}$$

$$\alpha'_\perp\mu_0 = -\alpha'_\perp\Gamma\overline{\beta}\overline{\alpha}'\mu_1 + \xi\eta'\overline{\beta}'_\perp\tau_1, \tag{8.11}$$

where we have used the reduced rank condition $\alpha'_\perp\Gamma\beta_\perp = \xi\eta'$ in (8.11).

Premultiplying (8.11) with $(\overline{\xi}, \xi_\perp)'$ and applying $\beta_{\perp 1} = \overline{\beta}_\perp\eta$ yields

$$\overline{\xi}'\alpha'_\perp\mu_0 = -\overline{\xi}'\alpha'_\perp\Gamma\overline{\beta}\overline{\alpha}'\mu_1 + \beta'_{\perp 1}\tau_1, \tag{8.12}$$

$$\xi_\perp'\alpha_\perp'\mu_0 = -\xi_\perp'\alpha_\perp'\Gamma\overline{\beta}\overline{\alpha}'\mu_1. \tag{8.13}$$

Note that (8.13) is the second restriction in (3.2). Rewriting and reorganizing (8.12) and (8.10) we find

$$\beta_{\perp 1}'\tau_1 = \overline{\xi}'\alpha_\perp'\mu_0 + \overline{\xi}'\alpha_\perp'\Gamma\overline{\beta}\overline{\alpha}'\mu_1, \tag{8.14}$$

$$\beta'\tau_0 - \overline{\alpha}'\Gamma\overline{\beta}_{\perp 2}\beta_{\perp 2}'\tau_1 = -\overline{\alpha}'\mu_0 - (I_r + \overline{\alpha}'\Gamma\overline{\beta})\overline{\alpha}'\mu_1 + \overline{\alpha}'\Gamma\overline{\beta}_{\perp 1}\beta_{\perp 1}'\tau_1, \tag{8.15}$$

where we have used the fact that $\overline{\beta}_\perp\beta_\perp' = \overline{\beta}_{\perp 1}\beta_{\perp 1}' + \overline{\beta}_{\perp 2}\beta_{\perp 2}'$. Now, in the $I(2)$ directions it is not possible to separate the linear drift from the initial conditions. Since we are only interested in the deterministic terms, which do not depend on initial values, we set $\overline{\beta}_{\perp 2}'\tau_1 = 0$. With this in hand, we separate τ_1 in the $\overline{\beta}\beta'$ and $\overline{\beta}_{\perp 1}\beta_{\perp 1}'$ directions

$$\tau_1 = \overline{\beta}\beta'\tau_1 + \overline{\beta}_{\perp 1}\beta_{\perp 1}'\tau_1,$$

and insert the expressions (8.6) and (8.14) to find

$$\tau_1 = \overline{\beta}_{\perp 1}\overline{\xi}'\alpha_\perp'\mu_0 + \overline{\beta}_{\perp 1}\overline{\xi}'\alpha_\perp'\Gamma\overline{\beta}\overline{\alpha}'\mu_1 - \overline{\beta}\overline{\alpha}'\mu_1.$$

Premultiplying (8.15) with $\overline{\beta}$ and inserting (8.14) and the restriction $\overline{\beta}_{\perp 2}\tau_1 = 0$ yields

$$\begin{aligned}\tau_0 \;=\; & -\overline{\beta}\overline{\alpha}'\mu_0 - \overline{\beta}(I_r + \overline{\alpha}'\Gamma\overline{\beta})\overline{\alpha}'\mu_1 + \overline{\beta}\overline{\alpha}'\Gamma\overline{\beta}_{\perp 1}(\overline{\xi}'\alpha_\perp'\mu_0 + \overline{\xi}'\alpha_\perp'\Gamma\overline{\beta}\overline{\alpha}'\mu_1) \\ & -\overline{\beta}\overline{\alpha}'(I_p - \Gamma\overline{\beta}_{\perp 1}\overline{\xi}'\alpha_\perp')\mu_0 - \overline{\beta}(I_r + \overline{\alpha}'\Gamma\overline{\beta} - \overline{\alpha}'\Gamma\overline{\beta}_{\perp 1}\overline{\xi}'\alpha_\perp'\Gamma\overline{\beta})\overline{\alpha}'\mu_1.\end{aligned}$$

9. Appendix C

We consider the cointegrated $I(2)$ ECM model given in (3.1)

$$\Delta^2 X_t = \alpha\beta' X_{t-1} - \Gamma\Delta X_{t-1} + \mu_0 + \mu_1 t + \varepsilon_t. \tag{9.1}$$

In the transformed system, given in (4.1), we set $\widetilde{X}_{1t} = b'X_t$ and $\widetilde{X}_{2t} = v'\Delta X_t$ such that $\Delta\widetilde{X}_{1t} = b'\Delta X_t$, $\Delta^2\widetilde{X}_{1t} = b'\Delta^2 X_t$, and $\Delta\widetilde{X}_{2t} = v'\Delta^2 X_t$. Then we can write

$$\begin{aligned}\Delta X_t &= v_\perp(b'v_\perp)^{-1}b'\Delta X_t + b_\perp(v'b_\perp)^{-1}v'\Delta X_t \\ &= v_\perp(b'v_\perp)^{-1}\Delta\widetilde{X}_{1t} + b_\perp(v'b_\perp)^{-1}\widetilde{X}_{2t},\end{aligned} \tag{9.2}$$

and the same way

$$\Delta^2 X_t = v_\perp(b'v_\perp)^{-1}\Delta^2\widetilde{X}_{1t} + b_\perp(v'b_\perp)^{-1}\Delta\widetilde{X}_{2t}. \tag{9.3}$$

Inserting (9.2) and (9.3) in (9.1), we find

$$\begin{aligned}
& v_\perp(b'v_\perp)^{-1}\Delta^2\widetilde{X}_{1t} + b_\perp(v'b_\perp)^{-1}\Delta\widetilde{X}_{2t} \\
= \; & \alpha(I_r,0)\widetilde{X}_{1t-1} - \Gamma(v_\perp(b'v_\perp)^{-1}\Delta\widetilde{X}_{1t-1} + b_\perp(v'b_\perp)^{-1}\widetilde{X}_{2t-1}) \\
& +\mu_0 + \mu_1 t + \varepsilon_t,
\end{aligned}$$

where we have used $\beta' X_{t-1} = \beta'\overline{b}b'X_{t-1} = (I_r,0)\widetilde{X}_{1t-1}$ since

$$\beta'\overline{b} = \beta'(\beta,\beta_{\perp 1})((\beta,\beta_{\perp 1})'(\beta,\beta_{\perp 1}))^{-1} = (I_r,0).$$

Applying $\Delta^2\widetilde{X}_{1t} = \Delta\widetilde{X}_{1t} - \Delta\widetilde{X}_{1t-1}$ and $[v_\perp(b'v_\perp)^{-1}, b_\perp(v'b_\perp)^{-1}]^{-1} = (b,v)'$, we find

$$\begin{aligned}
& [v_\perp(b'v_\perp)^{-1}, b_\perp(v'b_\perp)^{-1}]\Delta\widetilde{X}_t \\
= \; & [\alpha(I_r,0), -\Gamma b_\perp(v'b_\perp)^{-1}]\widetilde{X}_{t-1} \\
& -[(\Gamma - I_p)(v_\perp(b'v_\perp)^{-1}, 0]\Delta\widetilde{X}_{t-1} \\
& +\mu_0 + \mu_1 t + \varepsilon_t \\
\Leftrightarrow \; \Delta\widetilde{X}_t = \; & (b,v)'[\alpha(I_r,0), -\Gamma b_\perp(v'b_\perp)^{-1}]\widetilde{X}_{t-1} \\
& -(b,v)'[(\Gamma - I_p)(v_\perp(b'v_\perp)^{-1}, 0]\Delta\widetilde{X}_{t-1} \\
& +(b,v)'(\mu_0 + \mu_1 t + \varepsilon_t).
\end{aligned}$$

Hence, the parameters in the transformed model can be expressed as

$$\begin{aligned}
\widetilde{\Pi} &= (b,v)'[\alpha(I_r,0), -\Gamma b_\perp(v'b_\perp)^{-1}] \\
\widetilde{\Gamma}_1 &= -(b,v)'[(\Gamma - I_p)(v_\perp(b'v_\perp)^{-1}, 0] \\
\widetilde{\mu}_0 &= (b,v)'\mu_0 \\
\widetilde{\mu}_1 &= (b,v)'\mu_1 = (b,v)'\alpha\rho_1 \\
\widetilde{\varepsilon}_t &= (b,v)'\varepsilon_t.
\end{aligned}$$

Note that $\widetilde{\Gamma} = I_p - \widetilde{\Gamma}_1$ so that

$$\begin{aligned}
\widetilde{\Gamma} &= (b,v)'[v_\perp(b'v_\perp)^{-1}, b_\perp(v'b_\perp)^{-1}] + (b,v)'[(\Gamma - I_p)(v_\perp(b'v_\perp)^{-1}, 0] \\
&= (b,v)'[\Gamma v_\perp(b'v_\perp)^{-1}, b_\perp(v'b_\perp)^{-1}].
\end{aligned}$$

Furthermore, since $\widetilde{C} = \widetilde{\beta}_\perp(\widetilde{\alpha}'_\perp\widetilde{\Gamma}\widetilde{\beta}_\perp)^{-1}\widetilde{\alpha}'_\perp$, we find

$$\begin{aligned}
\widetilde{C} = \; & \begin{pmatrix} b'\overline{\beta}\delta & b'\beta_{\perp 1} \\ v'\overline{\beta}_{\perp 2} & 0 \end{pmatrix} \\
& \left[\alpha'_\perp[\Gamma v_\perp(b'v_\perp)^{-1}, b_\perp(v'b_\perp)^{-1}] \begin{pmatrix} b'\overline{\beta}\delta & b'\beta_{\perp 1} \\ v'\overline{\beta}_{\perp 2} & 0 \end{pmatrix}\right]^{-1} \alpha'_\perp(b,v)'^{-1} \\
= \; & \begin{pmatrix} b'\overline{\beta}\delta & b'\beta_{\perp 1} \\ v'\overline{\beta}_{\perp 2} & 0 \end{pmatrix} \\
& \left[\alpha'_\perp[\Gamma v_\perp(b'v_\perp)^{-1}b'\overline{\beta}\delta + \overline{\beta}_{\perp 2}, \Gamma v_\perp(b'v_\perp)^{-1}b'\beta_{\perp 1}]\right]^{-1} \alpha'_\perp(b,v)'^{-1}.
\end{aligned}$$

It can be verified that $\widetilde{\Pi}\widetilde{\beta}_\perp = 0$. Since $\widetilde{\Pi} = \widetilde{\alpha}\widetilde{\beta}'$ such that $\widetilde{\alpha} = \widetilde{\Pi}\widetilde{\beta}(\widetilde{\beta}'\widetilde{\beta})^{-1}$, we can also find

$$
\begin{aligned}
\widetilde{\alpha} &= (b,v)'[\alpha(I_r,0), -\Gamma b_\perp(v'b_\perp)^{-1}] \begin{pmatrix} \bar{b}'\beta \\ \omega \end{pmatrix} (\beta'\bar{b}\bar{b}\beta + \omega'\omega)^{-1} \\
&= (b,v)'[\alpha(I_r,0)\bar{b}'\beta - \alpha\bar{\alpha}'\Gamma\overline{\beta}_{\perp 2}(v'\overline{\beta}_{\perp 2})^{-1}\omega](I_r + \omega'\omega)^{-1} \\
&= (b,v)'[\alpha(I_r,0)\bar{b}'\beta + \alpha\delta(v'\overline{\beta}_{\perp 2})\omega](I_r + \omega'\omega)^{-1} \\
&= (b,v)'\alpha[I_r + \omega'\omega](I_r + \omega'\omega)^{-1} \\
&= (b,v)'\alpha.
\end{aligned}
$$

Thus it follows that

$$\widetilde{\alpha}_\perp = (b,v)'^{-1}\alpha_\perp.$$

To investigate the implications of the restriction (3.2), we express the parameters in the original model in terms of the parameters of the transformed model. It is straightforward to find

$$
\begin{aligned}
\mu_0 &= (b,v)'^{-1}\widetilde{\mu}_0 \\
\mu_1 &= (b,v)'^{-1}\widetilde{\mu}_1 \\
\alpha &= (b,v)'^{-1}\widetilde{\alpha} \\
\alpha_\perp &= (b,v)\widetilde{\alpha}_\perp.
\end{aligned}
$$

Since $\alpha_\perp = (\alpha_{\perp 1}, \alpha_{\perp 2})$ it follows that

$$\alpha_{\perp 2} = (b,v)\widetilde{\alpha}_\perp(0, I_{s_2})'.$$

From (4.4) we find

$$(b,0)\widetilde{\beta} = b\bar{b}'\beta = \beta.$$

Given the expression of $\widetilde{\Pi}$ and $\widetilde{\Gamma}$ we find

$$
\begin{aligned}
-(b,v)'^{-1}\widetilde{\Pi}(0,v)' &= \Gamma b_\perp(v'b_\perp)^{-1}v' \\
(b,v)'^{-1}\widetilde{\Gamma}(b,0)' &= \Gamma v_\perp(b'v_\perp)^{-1}b',
\end{aligned}
$$

such that

$$\Gamma = (b,v)'^{-1}[\widetilde{\Gamma}(b',0) - \widetilde{\Pi}(0,v)'].$$

The first part of (3.2), $\mu_1 = \alpha\rho_1$, then implies

$$
\begin{aligned}
& (b,v)'^{-1}\widetilde{\mu}_1 = (b,v)'^{-1}\widetilde{\alpha}\rho_1 \\
\Rightarrow\ & \widetilde{\mu}_1 = \widetilde{\alpha}\rho_1,
\end{aligned}
$$

which has the same consequence as in the original system. The second part, $\alpha'_{\perp 2}(\mu_0 + \Gamma\overline{\beta}\rho_1) = 0$, implies

$$\begin{aligned}
& (0, I_{s_2})\tilde{\alpha}'_{\perp}\left[\tilde{\mu}_0 + [\tilde{\Gamma}(b',0) - \tilde{\Pi}(0,v)'](b,0)\tilde{\beta}\left(\tilde{\beta}'\begin{pmatrix} b'b & 0 \\ 0 & 0 \end{pmatrix}\tilde{\beta}\right)^{-1}\rho_1\right] \\
= \; & (0, I_{s_2})\tilde{\alpha}'_{\perp}\left[\tilde{\mu}_0 + \tilde{\Gamma}\begin{pmatrix} b'b & 0 \\ 0 & 0 \end{pmatrix}\tilde{\beta}\left(\tilde{\beta}'\begin{pmatrix} b'b & 0 \\ 0 & 0 \end{pmatrix}\tilde{\beta}\right)^{-1}\rho_1\right] = 0,
\end{aligned}$$

which is a seemingly unnecessary restriction on the transformed parameters, which is carried over from the $I(2)$ model.

10. Appendix D

In this appendix we find the coefficients for the deterministics in the MA-representation for the transformed model expressed in terms of the parameters of the original $I(2)$ model. We separate $\tilde{\mu}_i$ $(i = 0, 1)$ in the $\tilde{\alpha}$ and $\tilde{\alpha}_{\perp}$ directions:

$$\begin{aligned}
& \tilde{\mu}_i = \tilde{\alpha}\tilde{\rho}_i + \tilde{\alpha}_{\perp}\tilde{\gamma}_i \\
\Rightarrow \; & (b,v)'\mu_i = (b,v)'\alpha\tilde{\rho}_i + (b,v)^{-1}\alpha_{\perp}\tilde{\gamma}_i.
\end{aligned}$$

Hence, we find

$$\begin{aligned}
\tilde{\rho}_i &= [\alpha'\Lambda\alpha]^{-1}\alpha'\Lambda\mu_i \\
\tilde{\gamma}_i &= [\alpha'_{\perp}(b,v)'^{-1}(b,v)^{-1}\alpha_{\perp}]^{-1}\alpha_{\perp}(b,v)'^{-1}(b,v)'\mu_i,
\end{aligned}$$

where $\Lambda = (b,v)(b,v)'$. The coefficients for the constants and the drifts are then given by

$$\tilde{\tau}_0 = \tilde{\overline{\beta}}\tilde{\overline{\alpha}}'(\tilde{\Gamma}\tilde{C} - I_p)\tilde{\mu}_0 + \tilde{\overline{\beta}}\tilde{\vartheta}\tilde{\rho}_1, \tag{10.1}$$

$$\tilde{\tau}_1 = (\tilde{C}\tilde{\Gamma} - I_p)\tilde{\overline{\beta}}\tilde{\rho}_1 + \tilde{C}\tilde{\mu}_0, \tag{10.2}$$

where $\tilde{\vartheta} = \tilde{\overline{\alpha}}'\tilde{\Gamma}\tilde{C}\tilde{\Gamma}\tilde{\overline{\beta}} - I_r - \tilde{\overline{\alpha}}'\tilde{\Gamma}\tilde{\overline{\beta}}$.

We consider each term in(10.1) separately:

$$\begin{aligned}
\tilde{\overline{\beta}}\tilde{\overline{\alpha}}'\tilde{\Gamma}\tilde{C}\tilde{\mu}_0 \; = \; & \begin{pmatrix} \overline{b}'\beta \\ \omega \end{pmatrix}(I_r + \omega'\omega)^{-1}[\alpha'\Lambda\alpha]^{-1}\alpha'(b,v) \\
& (b,v)'[\Gamma v_{\perp}(b'v_{\perp})^{-1}b'\overline{\beta}\delta + \overline{\beta}_{\perp 2}, b_{\perp}(v'b_{\perp})^{-1}b'\beta_{\perp 1}] \\
& [\alpha'_{\perp}[\Gamma v_{\perp}(b'v_{\perp})^{-1}b'\overline{\beta}\delta + \overline{\beta}_{\perp 2}, \Gamma v_{\perp}(b'v_{\perp})^{-1}b'\beta_{\perp 1}]]^{-1}\alpha'_{\perp}\mu_0, \\
-\tilde{\overline{\beta}}\tilde{\overline{\alpha}}'\tilde{\mu}_0 \; = \; & -\begin{pmatrix} \overline{b}'\beta \\ \omega \end{pmatrix}(I_r + \omega'\omega)^{-1}[\alpha'\Lambda\alpha]^{-1}\alpha'\Lambda\mu_0,
\end{aligned}$$

$$
\begin{aligned}
\widetilde{\overline{\beta\widetilde{\alpha}}}'\widetilde{\Gamma}\widetilde{C}\widetilde{\Gamma}\widetilde{\overline{\beta}}\widetilde{\rho}_1 \;=\; & \begin{pmatrix} \overline{b}'\beta \\ \omega \end{pmatrix} (I_r + \omega'\omega)^{-1}[\alpha'\Lambda\alpha]^{-1}\alpha'\Lambda \\
& [\Gamma v_\perp(b'v_\perp)^{-1}b'\overline{\beta}\delta + \overline{\beta}_{\perp 2}, \Gamma v_\perp(b'v_\perp)^{-1}b'\beta_{\perp 1}] \\
& \left[\alpha'_\perp[\Gamma v_\perp(b'v_\perp)^{-1}b'\overline{\beta}\delta + \overline{\beta}_{\perp 2}, \Gamma v_\perp(b'v_\perp)^{-1}b'\beta_{\perp 1}]\right]^{-1} \\
& \alpha'_\perp[\Gamma v_\perp(b'v_\perp)^{-1}\overline{b}'\beta, b_\perp(v'b_\perp)^{-1}\omega](I_r + \omega'\omega)^{-1} \\
& [\alpha'\Lambda\alpha]^{-1}\alpha'\Lambda\mu_1,
\end{aligned}
$$

$$
\begin{aligned}
\widetilde{\overline{\beta\widetilde{\alpha}}}'\widetilde{\Gamma}\widetilde{\overline{\beta}}\widetilde{\rho}_1 \;=\; & \begin{pmatrix} \overline{b}'\beta \\ \omega \end{pmatrix} (I_r + \omega'\omega)^{-1}[\alpha'\Lambda\alpha]^{-1}\alpha'(b, v) \\
& [\Gamma v_\perp(b'v_\perp)^{-1}\overline{b}'\beta, b_\perp(v'b_\perp)^{-1}\omega](I_r + \omega'\omega)^{-1} \\
& [\alpha'\Lambda\alpha]^{-1}\alpha'\Lambda\mu_1.
\end{aligned}
$$

We now make the same exercise for (10.2):

$$
\begin{aligned}
\widetilde{C}\widetilde{\Gamma}\widetilde{\overline{\beta}}\widetilde{\rho}_1 \;=\; & \begin{pmatrix} b'\overline{\beta}\delta & b'\beta_{\perp 1} \\ v'\overline{\beta}_{\perp 2} & 0 \end{pmatrix} \\
& \left[\alpha'_\perp[\Gamma v_\perp(b'v_\perp)^{-1}b'\overline{\beta}\delta + \overline{\beta}_{\perp 2}, \Gamma v_\perp(b'v_\perp)^{-1}b'\beta_{\perp 1}]\right]^{-1} \\
& \alpha'_\perp[\Gamma v_\perp(b'v_\perp)^{-1}\overline{b}'\beta, b_\perp(v'b_\perp)^{-1}\omega](I_r + \omega'\omega)^{-1} \\
& [\alpha'(b, v)(b, v)'\alpha]^{-1}\alpha'(b, v)(b, v)'\mu_1,
\end{aligned}
$$

$$
-\widetilde{\overline{\beta}}\widetilde{\rho}_1 \;=\; -\begin{pmatrix} \overline{b}'\beta \\ \omega \end{pmatrix} (I_r + \omega'\omega)^{-1}[\alpha'\Lambda\alpha]^{-1}\alpha'\Lambda\mu_1,
$$

$$
\begin{aligned}
\widetilde{C}\widetilde{\mu}_0 \;=\; & \begin{pmatrix} b'\overline{\beta}\delta & b'\beta_{\perp 1} \\ v'\overline{\beta}_{\perp 2} & 0 \end{pmatrix} [\alpha'_\perp[\Gamma v_\perp(b'v_\perp)^{-1}b'\overline{\beta}\delta + \overline{\beta}_{\perp 2}, \\
& \Gamma v_\perp(b'v_\perp)^{-1}b'\beta_{\perp 1}]]^{-1}\alpha'_\perp\mu_0.
\end{aligned}
$$

With the restriction of the $H^*(r, s_1)$ model, $\mu_1 = \alpha\rho_1$, we get

$$
\begin{aligned}
& [\alpha'\Lambda\alpha]^{-1}\alpha'\Lambda\mu_1 \\
= \; & [\alpha'\Lambda\alpha]^{-1}\alpha'\Lambda\alpha\rho_1 \\
= \; & \rho_1.
\end{aligned}
$$

CHAPTER 6

Conclusions

When making analyses in the field of economics, one tries to answer some questions, which might be completely new, or find support for or against existing results. Often the process of the analysis raises new questions which could be interesting to look into. This chapter summarizes some of the conclusions and suggestions for future research outlined in the previous four chapters.

Chapter 2 dealt with the issue of integration of stock markets. It was shown that if stocks are priced according to the CAPM, then it is possible to find evidence for or against market integration by looking for common stochastic trends in the time series for the market portfolios. Common trends, however, is not enough. Also we need to identify the cointegrating relations. An advantage of testing integration with a cointegrated VAR model is, that we might also be able to say something about whether the stochastic trend comes from some of the markets considered. An empirical example suggested integration of the US and the Danish stock markets, which was due to the fact that the Danish market followed the development in the American. An interesting question is if we can apply common trends analysis if stocks are not priced according to the CAPM but maybe by another theoretical model such as the APT. Finally, in the conclusion, it was suggested that timing of market integration could be investigated by recursive analysis. This issue has to do with the question of constant parameters, which was also discussed in chapter 4, and it is certainly important when discussing integration of markets.

Chapter 3 contained an analysis of the PPP within the borders of the US. Contrary to traditional analyses of the PPP, this one did not include an exchange rate. The prices were found to be integrated of order two, which raised the issue of how to interpret the multicointegrating relations in this context. This could be done with an LQAC model in which policy makers seek integration of markets and try to minimize inflation. It was argued that when in doubt whether prices are $I(1)$ or $I(2)$, then it is better to treat them as $I(2)$. Three definitions were introduced, of which two are new in the literature. They relate to underlying inflation - same stochastic variation - and PPP with adjustment. The latter takes into account that even though the PPP does not hold, it might be the case that prices are adjusting toward a sustainable PPP level. The empirical analysis provided evidence that the underlying inflation is the same in the four areas considered and that the PPP with adjustment holds between the three areas geographically closer to each other. The obvious question to ask after this analysis is, if similar conclusions can be made for the euro area, or other areas, where the exchange rate is fixed.

Chapter 4 generalized the analysis in chapter 3 by including the exchange rate. This makes the analysis somewhat more complicated. Three definitions of price parities generalized those of chapter 3 and, by means of an empirical illustration, it was argued that when considering a long span of data for prices, there is a high risk of non-constant parameters. There was some indication that the US inflation rate might have been stationary in the pre-1960 period but non-stationary in the recent period. A scenario analysis exposed possible cases when analyzing cointegrated VAR systems of variables including prices and exchange rates. An empirical analysis in this chapter suggested that the PPP with adjustment holds between the UK and the US, when analyzing data from the post Bretton woods period. The parameters seemed to be stable for this sample. Since inflation rates are often non-stationary future research in theoretical macroeconomics might take this into account. For example, the price mechanism could be described by an LQAC model.

Chapter 5 was concerned with the role of the deterministic terms when making transformations from the $I(2)$ space to $I(1)$. The deterministics were expressed by the parameters in the $I(1)$ as well as the $I(2)$ model. Then followed a discussion of the $I(2)$-to-$I(1)$ transformation and the parameters in the transformed model were expressed by the $I(2)$ parameters. A discussion of the deterministic terms in the transformed model was then given. It turned out that the drifts in the stationary directions were unchanged by the transformation whereas the constants, as well as the drifts in the non-stationary directions, changed. Because of implicit restrictions on the transformed model, carried over from the original model, it was suggested that tests for rank and deterministic terms should be carried out in the $I(2)$ model.

CHAPTER 7

Bibliography

Anderson, T.W. (1951) 'Estimating linear restrictions on regression coefficients for multivariate normal distributions', *Annals of Mathematical Statistics*, 22(3), 327-351.

Bayoumi, T. & MacDonald, R. (1998), 'Deviations of exchange rates from purchasing power parity: A story featuring two monetary unions', *Working Paper No. 98/69*, International Monetary Fond.

Bekaert, G. & Harvey, C.R. (1995), 'Time-varying world market integration', *Journal of Finance*, 50(2), 403-444.

Black, F. (1972), 'Capital market equilibrium with restricted borrowing', *Journal of Business*, 45, 444-455.

Bossaerts, P. (1988), 'Common nonstationary components of asset prices', *Journal of Economic Dynamics and Control*, 12, 347-364.

Breuer, J.B. (1994), 'An assessment of the evidence on purchasing power parity', ch. 7 in Williamson, J. (ed.), *Estimating Equilibrium Exchange Rates*, Institute for International Economics, Washington DC.

Cassel, G. (1922), *Money and Foreign Exchange after 1914*, Macmillan.

Chen, Z. & Knez, P.J. (1995), 'Measurement of market integration and arbitrage', *Review of Financial Studies*, 8(2), 287-325.

Copeland, T.E. & Weston, I.F. (1988), *Financial Theory and Corporate Policy*, 3^{rd} edition, Addison-Wesley Publishing Company.

Crowder, W.J. (1996a), 'A reexamination of long-run PPP: The case of Canada, the UK, and the US', *Review of International Economics*, 4(1), 64-78.

Crowder, W.J. (1996b), 'Purchasing power parity when prices are I(2)', *Review of International Economics*, 4(2), 234-246.

Doornik, J.A. & Hendry, D.F. (1996), *GiveWin. An Interface to Empirical Modelling*, International Thomson Business Press.

Doornik, J.A. & Hendry, D.F. (1997), *Modelling Dynamic Systems Using PcFiml 9.0 for Windows*, International Thomson Business Press.

Edison, H.J. (1987), 'Purchasing power parity in the long run: A test of the Dollar/pound exchange rate (1890-1978)', *Journal of Money, Credit, and Banking*, 19(3), 376-387.

Engel, R.F. & Yoo, B.S. (1989), 'Cointegrated economic time series: A survey with new results', *Discussion Paper 87-26R*, University of California, San Diego, CA.

Engle, C. & Rogers, J.H. (1996), 'How wide is the border?', *American Economic Review*, 86, 1112-1125.

Engle, C. & Rogers, J.H. (2001), 'Violating the law of one price: Should we make a federal case out of it?', *Journal of Money, Credit and Banking*, 33(1), 1-15.

Engle, R.F. & Granger, C.W.J. (1987), 'Co-integration and error correction: Representation, estimation and testing', *Econometrica*, 55(2), 251-276.

Engle, R.F., Hendry, D.F. & Richard J.-F. (1983), 'Exogeneity', *Econometrica*, 51, 277-304.

Froot, K.A. & Rogoff, K. (1995), 'Perspectives on PPP and long-run real exchange rates', ch. 32. in Grossman, G.M. & Rogoff, K., *Handbook of International Economics. Volume III*, Elsevier.

Gonzalo, J. (1994), 'Five alternative methods of estimating long-run equilibrium relationships', *Journal of Econometrics*, 60, 203-233.

Granger, C.W.J. (1983), 'Co-integrated variables and error-correction models', *Discussion paper 83-13*, University of California, San Diego, CA.

Gregory, A.W. (1994), 'Testing for cointegration in linear quadratic models', *Journal of Business and Economic Statistic*, 12(3), 347-360.

Gregory, A.W., Pagan, A.R. & Smith, G.W. (1993), 'Estimating linear quadratic models with integrated processes', Ch. 15 in Phillips, P.C.B. (ed.), *Models, Methods, and Applications of Econometrics: Essays in Honor of A.R. Bergstrom*, Blackwell.

Gultekin, M.N, Gultekin, N.B. & Penati, A. (1989), 'Capital controls and international capital market segmentation: The evidence from the Japanese and American stock markets', *Journal of Finance*, 44(4), 849-869.

Haavelmoo, T. (1944), 'The probability approach in econometrics', *Econometrica*, 12, suppl., 1-118.

Hamilton, J.D. (1994), *Time Series Analysis*, Princeton University Press.

Hansen, B. & Phillips, P.C.B. (1990), 'Estimation and inference in models of cointegration: A simulation study', *Advances in Econometrics*, 8, 225-248.

Hansen, H. & Johansen, S. (1999), 'Some test for parameter constancy in cointegrated VAR-models', *Econometrics Journal*, 2, 26-52.

Hansen, H. & Juselius, K. (1995), *CATS in RATS. Cointegration Analysis of Time Series*, Estima.

Hansen, E. & Rahbek, A. (1998), 'Stationarity and asymptotics of multivariate ARCH time series with an application to robustness of cointegration analysis', *Reprint 12*, Department of Theoretical Statistics, University of Copenhagen.

Hardouvelis, G., Malliaropulos, D. & Priestley, R. (1999), 'EMU and European stock market integration', *CEPR Discussion Paper No. 2124*.

Huang, C.-f. & Litzenberger, R.H. (1988), *Foundations for Financial Economics*, North-Holland.

Johansen, S. (1988), 'The statistical analysis of cointegration vectors', *Journal of Economic Dynamics and Control*, 12(2), 231-254.

Johansen, S. (1991), 'Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models', *Econometrica*, 59(6), 1551-1580.

Johansen, S. (1992), 'A representation of vector autoregressive processes integrated of order 2', *Econometric Theory*, 8(2), 188-202.

Johansen, S. (1994), 'The role of the constant and linear terms in cointegrated analysis of non-stationary variables', *Econometric Reviews*, 13(2), 205-229.

Johansen, S. (1996), *Likelihood Based inference in Cointegrated Vector Autoregressive Models*, 2nd edition, Oxford University Press.

Johansen, S. (1997), 'Likelihood analysis of the I(2) model', *Scandinavian Journal of Statistics*, 24, 433-462.

Johansen, S. (2000), 'A Bartlett correction factor for tests on the cointegrating relations', *Econometric Theory*, 16(5), 740-778.

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

Johansen, S. & Juselius, K. (1992), 'Testing structural hypotheses in a multivariate cointegration analysis of the PPP and the UIP for UK', *Journal of Econometrics*, 53, 211-244.

Johansen, S. & Juselius, K. (1994), 'Identification of the long-run and the short run structure. An application to the ISLM model', *Journal of Econometrics*, 63, 7-36.

Jorion, P. & Schwartz, E. (1986), 'Integration vs. segmentation in the Canadian stock market', *Journal of Finance*, 41(3), 603-614.

Juselius, K. (1995), 'Do the purchasing power parity and uncovered interest parity hold in the long run? An example of likelihood inference in a multivariate time-series model', *Journal of Econometrics*, 69, 211-240.

Juselius, K. (1996), 'A structured VAR in Denmark under changing monetary regime', Revised version of *Discussion Paper 96-02*, Institute of Economics, University of Copenhagen.

Juselius, K. (1999), 'Price convergence in the medium and the long run. An I(2) analysis of six price indices', ch. 13 in Engle, R.F. & White, H. (eds.), *Cointegration, Causality and Forecasting*, Oxford University Press.

Juselius, K. (2002), *The Cointegrated VAR Model: Econometric Methodology and Macroeconomic Applications*, Manuscript.

Juselius, K & MacDonald, R. (2000a), 'International parity relationships between Germany and the United Stated: A joint modelling approach', *Discussion Paper No. 00-10*, Institute of Economics, University of Copenhagen.

Juselius, K. & MacDonald, R. (2000b), 'Interest rate and price linkages between the USA and Japan: Evidence from the post- Bretton Woods period', *Japan and the World Economy*, forthcoming.

Jørgensen, C. (1998), 'A simulation study of the tests in the cointegrated VAR model', *Unpublished Report*, Institute of Economics, University of Copenhagen.

Kim, Y. (1990), 'Purchasing power parity in the long run: A cointegration approach', *Journal of Money, Credit and Banking*, 22(4), 491-503.

Kongsted, H.C. (1998), 'An I(2) cointegration analysis of small-country import price determination', *Discussion Paper 98-22*, Institute of Economics, University of Copenhagen.

Kongsted, H.C. (2002), 'Testing the nominal-to-real transformation', *Discussion Paper 02-06*, Institute of Economics, University of Copenhagen.

Kwiatkowski, D., Phillips, P.C.B, Schmidt, P. and Shin, Y. (1992), 'Testing the null hypothesis of stationarity against the alternative of a unit root', *Journal of Econometrics*, 54, 159-178.

Licht, A.N (1997), 'Stock market integration in Europe', *Paper No. 15 in Program on International Systems*, Havard Law School.

Nyblom, J. and Makelainen, T. (1983), 'Comparisons of the tests for presence of random walk coefficients in a simple linear model', *Journal of the American Statistical Association*, 78, 856-864.

Obstfeld, M. (1989), 'How integrated are world capital markets? Some new tests', Ch. 7 in Calvo, G., Findlay, R., Kouri, P. & de Macedo, J.B. (eds.), *Debt, Stabilization & Development*, The United Nations University.

Obstfeld, M. (1998), 'The global capital market: Benefactor or menace?', *NBER Working Paper No. 6559.*

Pagano, M. (1993), 'Financial markets and growth. An overview', *European Economic Review*, 37, 613-622.

Pantula, S.G. (1989), 'Testing for unit roots in time series data', *Econometric Theory*, 5, 256-271.

Parsley, D.C. & Wei, S.-J. (1996), 'Convergence to the law of one price without trade barriers or currency fluctuations', *The Quarterly Journal of Economics*, 111, 1211-1236.

Paruolo, P. (1996), 'On the determination of integration indices in I(2) systems', *Journal of Econometrics*, 72, 313-356.

Paruolo, P.(1998), 'Asymptotic efficiency of the two-step estimator in I(2) VAR systems', Forthcoming in *Journal of Econometrics.*

Paruolo, P. & Rahbek, A. (1999), 'Weak exogeneity in I(2) VAR systems', *Journal of Econometrics*, 93, 281-308.

Phillips, P.C.B. (1988), 'Spectral regression for cointegrated time series', in W. Barnett (ed.), *Non-parametric and Semiparametric Methods in Economics and Statistics*, Cambridge University Press, 413-436.

Phillips, P.C.B. & Loretan, M. (1991), 'Estimation long run economic equilibria', *Review of Economic Studies*, 58, 407-436.

Rahbek, A. (1997), 'Representation of cointegrated I(2) and I(1) processes with deterministic trends', *Manuscript.*

Rahbek, A., Kongsted, H.C. & Jørgensen, C. (1999), 'Trend stationarity in the I(2) cointegration model', *Journal of Econometrics*, 90, 265-289.

Rogoff, K. (1996) 'The purchasing power parity puzzle', *Journal of Economic Literature*, 34(2), 647-668.

Saikkonen, P. (1991), 'Asymptotically efficient estimation of cointegration regression', *Econometric Theory*, 7, 1-21.

Sharpe, W.F. (1963), 'A simplified model of portfolio analysis', *Management Science*, January, 277-293.

Sharpe, W.F. (1964), 'Capital asset prices: A theory of market equilibrium under conditions of risk', *Journal of Finance*, 19(3), 425-442.

Stock, J.H. (1987), 'Asymptotic properties of least squares estimators of cointegrating vectors', *Econometrica*, 55, 1035-1056.

Stock, J.H. (2001), 'Discussion of Cogley and Sargent, "Evolving post-world war II inflation dynamics"', forthcoming in *NBER Macroeconomics Annual 2001.*

Stock, J.H. & Watson, M.W. (1988), 'Testing for common trends', *Journal of the American Statistical Association*, 83, 1097-1107.

Stock, J.H. & Watson, M.W. (1993), 'A simple estimator of cointegrating vectors in higher order integrated systems', *Econometrica*, 61(4), 783-820.

Wheatley, S. (1988), 'Some tests of the consumption-based asset pricing model', *Journal of Monetary Economics*, 22, 193-215.

L1
31
.0
P6