Estimation of Cointegrated Systems in Continuous Time

Daniel Gonzalez Olivares

A thesis submitted in fulfilment of the requirements
for the degree of

Doctor of Philosophy

Department of Economics

University of Essex

December 4, 2017

"There is nothing either good or bad,

but thinking makes it so."

William Shakespeare

Acknowledgements

Studying this Ph.D has been one of the greatest experiences I have ever had in my life. Thought all these years, I have walked and enjoyed my carrier as a graduate student and this thesis is the result of all those years of work and dedication, of course, this journey has not been an easy one and, naturally, many people have played and important role.

Institutionally, I would to thank my two sponsors: The University of Guadalajara, and the National Council of Science and Technology in Mexico (CONACyT) whose support through their scholarship programs allowed me to complete my studies. Personally and professionally, I would like to thank all my teachers in Mexico whose encouragement supported my to pursue the Ph.D.

My family in general, my parents Benjamin Gonzalez and Maria del Rosario Olivares whose support has been invaluable and were always there to keep motivating me.

My supervisor, Professor Marcus J. Chambers has been a key element in this experience. I am deeply grateful for all his support, his patience and the uncountable hours he spend with me.

My wife, Martha Juliana, who has been always together with me in this journey and whose love and support made it infinitely easier.

Finally, our God, the person that allows us to achieve and succeed any challenge we face in our path.

Abstract

In this thesis we derive exact discrete time representation models that correspond to cointegrated systems in continuous time. At the same time, for the parameters of those models, estimation procedures are outlined. The representations are applicable for data observed as both stock or flow variables and with the use of some simulated data, the performance of the estimation procedure is assessed. More importantly, with the aim of analysing the costs, if there are any, of ignoring aggregation in the specification, the results of our estimation procedure are also compared with the ones we would have obtained by applying instead Johansen's estimation methodology. In the first part (Chapter 2), we detail the analysis for a first-order stochastic differential equation system, as a result, baseline finding are outlined. In the second part (Chapter 3) the analysis is generalized and not only includes higher order specifications in the system but also incorporates deterministic components on it. Finally, in the last part (Chapter 4) of this thesis, three applications of that estimation procedure are presented.

In the results, when the system is entirely comprised by stock variables and the specification follows a first order system, both Johansen's methodology and ours perform very well, with virtually identical estimates and, for the simulated data, improvements as the sample size increases. However, when the variables of interest are flows or the specification follows a higher order system, given that our exact discrete time representation includes moving average components in the error term, Johansen's estimates show a persistent bias in estimation, consequently, they reflected the cost of ignoring aggregation in the specification.

Contents

Chapter 1

Introduction

As it is well known, some decades ago, continuous time modelling, was not as popular in empirical economic studies as estimation in discrete time. In order to estimate the parameters of a model written in a continuous fashion, undoubtedly, it is needed to relate those parameters to the observed data. However, as the data are available at discrete intervals of time rather than on a continuous basis, such an estimation becomes challenging.

In attempting to solving such a complication Bergstrom [1966] following Phillips [1959], utilized an approximate discrete time model to estimate the parameters of the system. At the end, it was found that the accuracy in estimation of the approximation increased as the time interval between observations decreased. As a result, additional work was needed. In subsequent publications, many important improvements to the challenge were added. Phillips [1972], for example, considered a simple three-equation trade cycle model with five parameters and utilized the minimum distance (MD) procedure to provide consistent estimates of the structural parameters of the system. In his document, by the use of Monte Carlo simulations, he studied the small sample distributions of the estimates. Also, he compared those estimates with the ones obtained by applying instead the three stage least squares procedure to a discrete time approximation of the system. At the end, he found that the MD method gave superior estimates. Therefore, the author showed that there were considerable gains in efficiency from taking account of the exact restrictions on the distribution of the discrete data implied by the continuous time model. However, although the author said that the results may remain valid for more involved systems, they were only tested in that simple model.

Similarly, P.M. Robinson (see Robinson [1976a], Robinson [1976b]), by using the spectral representation of the continuous time model in terms of the Fourier transforms of the data, derived, for estimation, an approximate discrete time model (ADTM). In his document, the author considered a very general open continuous time dynamic model (a system of higher-order stochastic differential equations) and by applying the nonlinear least square method to that ADTM, he showed that the estimates of the structural parameters of the system were efficient, nevertheless, that conclusion remained valid only for the case when the exogenous variables of the system were generated by a stationary random process and satisfied certain aliasing conditions. As a result, the procedure could not be easily utilized in applied work.

It was not until Bergstrom's seminal paper in 1983 (Bergstrom, 1983) that continuous time modelling truly took off. As it is known, in his seminal paper, the author derived the first higher order exact discrete time model that held exactly for the data generated by the continuous time system regardless of the frequency with which that data was observed (the invariance property). Therefore, he gave econometricians the possibility to avoid the temporal aggregation problems that models naively specified in terms of the observation interval have.

Additionally to that, in the same document, the author specified the solution to the system considering that both stock and flow variables were included into the specification. In the solution, while stock variables were considered as observations taken at specific points of time, flows were measured as the accumulation of the underlying rate over a time interval. As a result, the serial correlation in the disturbances induced by the use of flow variables in the specification of the system was finally controlled, correctly analysed and accurately incorporated into the estimation of the parameters of the model.

In more recent years, at the same time, following that passion, important contributions and extensions of the original model have been added to the existent literature (see for example, Bergstrom, 1986, 1990, 1996, Wymer, 1993, Gandolfo, 1993 and Phillips, 1991) and the advantages of such procedures over its discrete time counterpart have been broadly underlined (see for example, Bergstrom, 1996 and Bergstrom and Nowman, 2007). Applications of such models, similarly, started with contributions in finance with the modelling of interest rates using linear and non-linear models as well as in macroeconomics with macro-systems of differential equations based on extensive economic theory that tried to derive the steady state solution of the economy (see, for example, Nowman [1997], Yu and Phillips [2001] and Bergstrom and Nowman [2007]).

For the non-stationary cointegrated case, the focus of this document, for example, Phillips [1991] introduced a triangular cointegrated system representation and derived an exact discrete time model in the form of a first-order triangular error correction model (ECM) format that could be regarded as the continuous time counterpart of the discrete time vector ECM popularized by Johansen [1991]. Chambers [1999], at the same time, derived the formula for the exact discrete time model corresponding to a continuous time higher-order system and presented the basis for its estimation and inference and Chambers and McCrorie [2007], using Phillips' triangular representation, introduced an error correction model using a frequency domain technique that approximates the likelihood function and outlined the asymptotic properties of the resulting estimators.

Among others, these papers have outlined the non stationary cointegrated case for models in continuous time, they have also underlined the characteristics of the disturbances and moreover, they have provided general conditions through which the estimates of the structural parameter of the system can be obtained. However, even after these important contributions, empirical applications of those systems have not appeared in the literature yet. Moreover, the few documents that outline the theory do not completely control dynamics. That is to say, when the triangular cointegrated representation is formulated, it is considered that the long run equilibrium relationships between the variables are embodied only in one part of the system, therefore, general dynamic adjustments are immediately ruled out (see, particularly, Phillips [1991] and Chambers [1999] equations (2) and (3)).

Therefore, in this thesis, we exploit that opportunity, try to expand in that line and focusing on the non-stationary cointegrated case, we develop and apply an estimation procedure for cointegrated systems in continuous time that incorporates full dynamics into the system. Also, with the aim of analysing the effects of temporal aggregation over the model specification, we assess the performance in estimation of such models by comparing and contrasting its estimated parameters with the ones we would have obtained by applying, instead, Johansen's general VECM to the same data.

Particularly, this thesis explores the topic in the following manner: Chapter 2, *Estimation of First Order Cointegrated Systems*, derives an exact discrete time error correction model, very much like in the Bergstrom tradition, that corresponds to a cointegrated continuous time system that is entirely comprised of stock or flow variables. For each specification, the chapter also outlines an estimation procedure that leads to the Gaussian estimates of our model's parameters. In an application, through the use of some simulated data, at the same time, it assesses the performance of such an estimation procedure and, with the aim of analysing the costs, if there are any, of ignoring aggregation in the model's specification, it compares and contrasts our continuous time estimates with the ones we would have obtained by applying instead Johansen's discrete time methodology to the same simulated data.

In the results, for stocks, on the one hand, it is found that both Johansen's methodology and ours perform very well, with reasonably small bias in estimation and improvements as the sample size increases. For flows, on the other hand, as the exact discrete time model includes a moving average component in the error term, when it comes to the dynamics of the system, Johansen's estimates show a persistent bias in estimation with almost no improvement as the sample size increases. Consequently, we can say that this persistent bias reflects the cost of ignoring aggregation in the specification.

Chapter 3, *Estimation of Higher Order Cointegrated Systems*, as an extension of the analysis provided in Chapter 2, presents an estimation procedure for cointegrated systems in continuous time that not only allows for higher order specifications in the system but also incorporates deterministic components on it. As before, the system in this chapter is also allowed to be entirely comprised of stock or flow variables.

The application follows a similar methodology of that specified in Chapter 2 and in the results, in almost all that cases, our continuous time estimation procedure shows superiority in estimation against Johansen's with smaller bias in the estimates and improvements as the sample size increases, as a result, it can be concluded that estimation bias in cointegrated systems does not only depend on whether the variables in the model suffer some sort of temporal aggregation, but also, on whether the system requires a higher order specification.

Chapter 4, *Empirical Applications*, finally, presents three multivariate applications of that estimation methodology. The analysis is carried out by comparing the estimates of the model's parameters considering two different time specification; Johansen's general VECM for discrete time and our exact discrete time VECM for continuous time. The applications evaluate, for the United States, the market efficiency hypothesis on the foreign exchange rate, the term structure of interest rates and the main implication of the rational-expectation permanent income hypothesis. In the results, it is shown that estimation bias in cointegrated systems does not only depend on whether the variables in the model suffer some sort of temporal aggregation, but also, on whether the system requires a higher order specification.

Chapter 2

Estimation of First Order Cointegrated Systems

This chapter derives an exact discrete time error correction model, very much like in the Bergstrom tradition, that corresponds to a cointegrated continuous time system that is entirely comprised of stock or flow variables. At the same time, for each specification, an estimation procedure that leads to the estimates of our model's parameters is outlined. In an application, through the use of some simulated data, the performance of such an estimation procedure is assessed. More importantly, with the aim of analysing the costs, if there are any, of ignoring aggregation in the model's specification, we compare and contrast our results with the ones we would have obtained by applying instead Johansen's discrete time methodology to the same simulated data.

In the results, for stocks, on the one hand, we find that both Johansen's methodology and ours perform very well, with reasonably small bias in estimation and improvements as the simple size increases. For flows, on the other hand, as our exact discrete time model includes a moving average component in the error term, when it comes to the dynamics of the system, Johansen's estimates show a persistent bias in estimation with almost no improvement as the sample size increases, consequently, we can say that this persistent bias reflects the cost of ignoring aggregation in the specification. Our methodology, instead, shows superiority in estimation with smaller bias and improvements as the sample size increases.

2.1 Introduction

Our economies and major financial markets are continuously operating through the year. They also involve millions of agents making decisions continuously and despite the fact of being observed and recorded in variables at particular points of time, these economic activities undoubtedly vary in a continuous fashion. It is intuitively obvious, then, that if we try to model and predict the behaviour of an economy with an econometric model, we should use a continuous time model rather than a discrete one.

In practice, however, due to the lack of accuracy in estimation inherited by the use of discrete time approximations of the systems[1] and also the inclusion of more sophisticated mathematical techniques in the method, continuous time modelling, some decades ago, was not as popular in empirical economic studies as estimation in discrete time. Nevertheless, even with those difficulties, Albert Rex Bergstrom, a distinguished New Zealand econometrician, continued working in the field and in Bergstrom [1983], he derived, together with outlining an estimation procedure, the first higher order exact discrete time model that held exactly for the data generated by the continuous time system regardless of the frequency with which that data was observed. As a result, he did not only bring econometricians the possibility to avoid the temporal aggregation problems that models naively specified in terms of the observation interval have, but also, built the basis over which continuous time modelling would take off.

In more recent years, thanks to the development of more sophisticated mathematical techniques and the improvement of computing power, that contribution has been widely explored and established in econometrics[2]. New extensions started with the inclusion of mixtures of stock and flow variables in the systems (Bergstrom, 1986) as well as the possibility for non-stationarity in the variables. Phillips [1991], for instance, introduced a triangular cointegrated system representation and, using a frequency domain regression technique for the estimation of the cointegrating vectors, derived an exact discrete time model in the form of a first-order triangular error correction model (ECM) format that could be regarded as the continuous time counterpart of the discrete time vector

[1]See Bergstrom [1976] for details.
[2]See Bergstrom [1990, 1996], Wymer [1993] and Gandolfo [1993] for extensive discussions.

ECM popularized by Johansen [1991]. In a generalization of Bergstrom's work and continuing with the non-stationary case, Chambers [1999] derived the formula for the exact discrete time model corresponding to a continuous time higher-order system and presented the basis for its estimation and inference and Chambers and McCrorie [2007], using Phillips' triangular representation, introduced an error correction model using a frequency domain technique that approximates the likelihood function and outlined the asymptotic properties of the resulting estimators.

Applications of such models, at the same time, started with many important contributions in finance with the modelling of interest rates using linear and non-linear models as well as in macroeconomics with macro-systems of differential equations based on extensive economic theory that tried to derive the steady state solution of the economy[3], however, even after this explosive development, empirical applications of cointegrated systems in continuous time have not appeared in the literature yet.

Among others, these papers have outlined the non stationary cointegrated case for models in continuous time, they have also underlined the characteristics of the disturbances and moreover, they have provided general conditions through which the estimates of the structural parameter of the system can be obtained. However, even after these important contributions, empirical applications of those systems have not appeared in the literature yet. Moreover, the few documents that outline the theory do not completely control dynamics. That is to say, when the triangular cointegrated representation is formulated, it is considered that the long run equilibrium relationships between the variables are embodied only in one part of the system, therefore, general dynamic adjustments are immediately ruled out (see, particularly, Phillips [1991] and Chambers [1999] equations (2) and (3)).

Therefore, the aim of this chapter is not only to follow that line and develop and apply an estimation procedure for cointegrated systems in continuous time that incorporates full dynamics into the system, but also, to measure to what extend using a continuous time specification yields more accurate estimates of the unknown parameters of the model, i.e., we will measure the gains, if there are any,

[3]See, for example, Nowman [1997], Yu and Phillips [2001] and Nowman [2006] as well as Donaghy [1993] and Bergstrom and Nowman [2007].

of considering aggregation in the model's specification.

There are three main contributions in this chapter. The first is the derivation of an exact discrete time error correction model that corresponds to a cointegrated continuous time system whose observations are allowed to be either stock variables, observable at points in time, or flow variables, observable as the integral of the underlying rate of flow over the observation interval. The second is the complete characterization of the properties of the discrete time disturbance vector as well as the derivation of an estimation procedure that provides the Gaussian estimates of the unknown parameters in our model and the third contribution, through the use of some simulated data in an application, is the possibility to compare and contrast those estimates with the ones we would have obtained by applying Johansen's methodology to the same data.

In the results, when the system is entirely comprised by stock variables, both our estimates and Johansen's perform very well with reasonably small bias in estimation and improvements as the sample size increases. However, in the flow variables case, given that our exact discrete time model includes a moving average component in the error term, when it comes to the dynamics of the system, Johansen's estimates show a persistent bias in estimation with no improvement as the sample size increases. Consequently, we can say that this persistent bias reflects the cost of ignoring aggregation in the specification.

The chapter is organized as follows. Section 2.2 defines the continuous time system and specifies the exact discrete time representation for stock and flow variables, deriving a VECM representation for non-stationary systems very much in the Bergstrom tradition. Section 2.3 concentrates on the derivation of the covariance properties of the discrete time disturbance vector for the two representations and outlines the estimation procedure. Section 2.4 summarizes the simulation results and compares both the estimates of our exact discrete time representation and those obtained by applying Johansen's methodology and section 2.5 concludes. Supplementary results are given in Appendix A and all proofs are in Appendix B.

2.2 The model

To examine the issues raised above, we shall consider a continuous time random n-vector $y(t)$ that is partitioned into two subvectors $y_1(t)$, $y_2(t)$ of dimensions $n_1 \times 1$ and $n_2 \times 1$, respectively, where $n_1 + n_2 = n$. We shall assume that the elements of $y(t)$ are $I(1)$ processes but that there exist n_1 stationary linear cointegrating relationships of the form $y_1(t) - B_1 y_2(t)$, where B_1 is a $n_1 \times n_2$ matrix of cointegrating parameters. In order to achieve identification in the system, we normalize these relationships on the elements of $y_1(t)$. It is also assumed that $y(t)$ satisfies the following first-order stochastic differential equation system

$$dy(t) = AB'y(t)dt + \zeta(dt), \qquad t > 0, \tag{2.1}$$

where $y(t)$ is a $n \times 1$ vector of continuous time random variables, $B = (I_{n_1}, -B_1)'$ and $A = (A_1', A_2')'$ are $n \times n_1$ reduced rank matrices with n_1 linearly independent vectors (A_1 is a $n_1 \times n_1$ matrix and A_2 is a $n_2 \times n_1$ matrix) and $\zeta(dt)$ is a vector of random measures defined on all subsets of the line $0 < t < \infty$ having finite Lebesgue measure such that

- $\mathrm{E}[\zeta(dt)] = 0$,
- $\mathrm{E}[\zeta(dt)\zeta(dt)'] = \Sigma dt$ and
- $\mathrm{E}[\zeta(\Delta_1)\zeta(\Delta_2)'] = 0$ for disjoint intervals Δ_1 and Δ_2.

In terms of the vectors $y_1(t)$ and $y_2(t)$, the system (2.1) implies that

$$\begin{aligned} dy_1(t) =& A_1 [\, y_1(t) - B_1 y_2(t) \,]\, dt + \zeta_1(dt), \quad t > 0, \\ dy_2(t) =& A_2 [\, y_1(t) - B_1 y_2(t) \,] dt + \zeta_2(dt), \quad t > 0, \end{aligned} \tag{2.2}$$

where the vector $\zeta(dt)$ has been also partitioned conformably with y_1 and y_2. The first equation relates the changes in y_1 to the disequilibrium error $B'y(t) = y_1(t) - B_1 y_2(t)$ while the second equation relates the changes for y_2. Notice that the reactions of y_1 and y_2 to the disequilibrium errors are captured by the adjustment coefficient matrices A_1 and A_2. Then, the system can be considered as a generalization of Phillips' triangular representation.

Finally, in order to explore the effects of temporal aggregation, we consider that the vector $y(t)$ can be entirely comprised by stock or flow variables. If we define stock variables in continuous time as $y^s(t)$, then its observed values at specific points in time are given by $y_t^s = y^s(t)$, also, if $y^f(t)$ is defined as flow variables in continuous time, then its observed rate of flow is given by

$$y_t^f = \int_{t-1}^{t} y^f(r)dr,$$

where in each case $t = 1, 2, \cdots, T$ and T denotes sample size.

After the set up and before the derivation of an exact discrete time representation of (2.1) that can be used for estimation, let's define, first, the unique mean square solution of the system, which initialized at $t = 0$ is given by Bergstrom [1984] and can be written as

$$y(t) = e^{tAB'}y(0) + \int_0^t e^{(t-s)AB'}\zeta(ds), \qquad t > 0, \tag{2.3}$$

where

$$e^{tAB'} = \sum_{j=0}^{\infty} \frac{(tAB')^j}{j!}.$$

As pointed out before by Bergstrom [1997], Chambers [1999] as well as McCrorie [2000], given that our system in (2.1) specifies a cointegrated relationship between the variables, it can be shown that the $n \times n$ exponential matrix $e^{tAB'}$ can be rewritten in a more simplified form that incorporates the reduced rank specification of AB' and as a result, simplifies the unique mean square solution to the system; next Lemma summarise such result.

Lemma 2.2.1 (Exponential Representation)**.**

Assuming that the $n_1 \times n_1$ matrix $M = B'A$ is non singular, the exponential matrix $e^{tAB'}$ can be re-written as

$$e^{tAB'} = I_n + AM^{-1}(e^{tM} - I_{n_1})B', \tag{2.4}$$

where all the vectors and matrices are specified as before.

Proof. See Appendix B. ∎

Then, using (2.4) into (2.3), our unique mean square solution to (2.1), which incorporates the reduced rank specification of AB' and is used in the derivation of the exact discrete time representation of our system, becomes

$$y(t) = (I_n + G(e^{tM} - I_{n_1})B')y(0) + \int_0^t (I_n + G(e^{(t-s)M} - I_{n_1})B')\zeta(ds), \qquad t > 0, \tag{2.5}$$

where $G = AM^{-1}$.

2.2.1 The discrete time representation

For each type of data, the exact discrete time model representation of our system is presented below in Lemma 2.2.2. For the derivation, considering (2.5), it is important to mention that standard manipulations of the type utilized in the proof of Theorem 2(c) of Bergstrom [1984] were applied to (2.1) (see Appendix B for details).

Lemma 2.2.2 (Exact Discrete Time Representations).

Let $y(t)$ satisfy the continuous time cointegrated system defined by (2.1), *then, the observed vector y_t^s evolves according to the discrete time VECM*

$$\Delta y_t^s = GJB'y_{t-1}^s + \eta_t, \qquad t = 1, \dots, T, \tag{2.6}$$

where $J = e^M - I_{n_1}$ and the disturbance vector η_t is defined as follows

$$\eta_t = \int_{t-1}^t \Big(I_n + G(e^{(t-s)M} - I_{n_1})B'\Big)\zeta(ds).$$

Also, the observed vector y_t^f evolves according to the discrete time VECM

$$\Delta y_t^f = GJB'y_{t-1}^f + v_t, \qquad t = 2, \dots, T, \tag{2.7}$$

where J is as before and

$$v_t = \int_{t-1}^t \int_{r-1}^r \Big(I_n + G(e^{(r-s)M} - I_{n_1})B'\Big)\zeta(ds)dr.$$

For the relationship between $y(0)$ and the observed vector y_1^f, finally, we have

$$y_1^f - y(0) = GEB'y(0) + v_1, \tag{2.8}$$

where $E = \int_0^1 (e^{rM} - I_{n_1})dr$ and the disturbance vector v_1 is defined as follows

$$v_1 = \int_0^1 \int_0^r \Big(I_n + G(e^{(r-s)M} - I_{n_1})B' \Big) \zeta(ds)dr.$$

Proof. See Appendix B. ■

From the lemma, it is easy to see that regardless of the observations being stocks or flows, the discrete time VECM representation is common. However, the specification of the observable vector y_t as well as the specification of the unobservable disturbance vector differs in each scheme.

At the same time, looking at equation (2.6) and/or (2.7), we can notice that our exact discrete time model does retain the original structure of the system as in (2.2), however, it now relates the change in y_t (Δy_t) to the lagged disequilibrium error $B'y_{t-1} = y_{1,t-1} - B_1 y_{2,t-1}$ and more importantly, it captures the reactions of y_{1t} and y_{2t} to the disequilibrium errors with more complicated functions involving all the parameters in the continuous time model. To see this, finding explicit representations of G, J and M

$$\begin{aligned} M &= B'A = (I_{n_1} - B_1) \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} = A_1 - B_1 A_2, \\ G &= AM^{-1} = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} (A_1 - B_1 A_2)^{-1}, \end{aligned} \tag{2.9}$$

then

$$\Delta y_t = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} (A_1 - B_1 A_2)^{-1} \left(e^{A_1 - B_1 A_2} - I_{n_1} \right) \begin{pmatrix} I_{n_1} & -B_1 \end{pmatrix} y_{t-1} + \epsilon_t,$$

where ϵ_t is either η_t or v_t from (2.6) or (2.7). Then, partitioning these matrices and vectors conformably with y_{1t} and y_{2t}, the corresponding structures are given by

$$\begin{aligned}
\Delta y_{1t} =& \underbrace{A_1(A_1 - B_1A_2)^{-1}\left(e^{A_1-B_1A_2} - I_{n_1}\right)}_{\text{Reaction parameter}}\Big[\ y_{1,t-1} - B_1y_{2,t-1}\ \Big] + \epsilon_{1t}, \\
\Delta y_{2t} =& \underbrace{A_2(A_1 - B_1A_2)^{-1}\left(e^{A_1-B_1A_2} - I_{n_1}\right)}_{\text{Reaction parameter}}\Big[\ y_{1,t-1} - B_1y_{2,t-1}\ \Big] + \epsilon_{2t}.
\end{aligned} \tag{2.10}$$

On top of that, we can see that the lemma also states that if we want to talk about appropriate methods of estimating the unknown parameters of the model, it is necessary to derive, firstly, the precise properties satisfied by the disturbance vectors η_t and v_t. The next section characterizes these properties, outlines the derivation of the Gaussian likelihood function and at the same time, specifies the estimation procedure for each type of data (stocks or flows).

2.3 Estimation and the properties of the discrete time disturbances

2.3.1 Stock representation

For stocks, equation (2.6) in the previous section, the problem of obtaining exact maximum likelihood estimates of the structural parameters of the continuous time model from a given sample, includes, as an important step, the derivation of the autocovariance matrix of the disturbance vector. As shown in Bergstrom [1984][4], this matrix can be calculated as follows

$$\begin{aligned}
\mathrm{E}[\eta_t\eta_t'] =& \mathrm{E}\left\{\left[\int_{t-1}^{t}\Big(I_n + G(e^{(t-s)M} - I_{n_1})B'\Big)\zeta(ds)\right] \times \right.\\
&\left.\left[\int_{t-1}^{t}\Big(I_n + G(e^{(t-s)M} - I_{n_1})B'\Big)\zeta(ds)\right]'\right\},\\
=&\int_{t-1}^{t}\Big(I_n + G(e^{(t-s)M} - I_{n_1})B'\Big)\Sigma\Big(I_n + G(e^{(t-s)M} - I_{n_1})B'\Big)' ds,\\
=&\int_{0}^{1}\Big(I_n + G(e^{(s)M} - I_{n_1})B'\Big)\Sigma\Big(I_n + G(e^{(s)M} - I_{n_1})B'\Big)' ds,\\
=&W,
\end{aligned} \tag{2.11}$$

where Σ is defined in the properties of the random vector and the last line follows from a simple change of variable in the integration.

Also,

$$\begin{aligned}
\mathrm{E}[\eta_u\eta_t'] =& \mathrm{E}\left\{\left[\int_{u-1}^{u}\Big(I_n + G(e^{(u-s)M} - I_{n_1})B'\Big)\zeta(ds)\right] \times \right.\\
&\left.\left[\int_{t-1}^{t}\Big(I_n + G(e^{(t-s)M} - I_{n_1})B'\Big)\zeta(ds)\right]'\right\},\\
=&0, \qquad u \neq t.
\end{aligned} \tag{2.12}$$

Then, assuming that η_t is normally distributed, the logarithm of the likelihood function is

$$L(\theta, \Sigma) = -\frac{nT}{2}\ln(2\pi) - \frac{nT}{2}\ln|W| - \frac{1}{2}\sum_{t=1}^{T}\eta_t' W^{-1}\eta_t, \tag{2.13}$$

[4]See Doob [1953] for an extensive discussion of this result.

where θ denotes the vector of unknown parameters to be estimated comprised in A and B.

Finally, the Gaussian estimates in $\hat{\theta}$ are then the elements of the vector θ that maximises L.

2.3.2 Flow representation

For flows, equation (2.7) in the previous section, maximum likelihood estimates, as well as for stocks, involve the characterization of the autocovariance properties of the discrete time disturbance vector v_t, however, as v_t involves a double integral of the vector of random measures $\zeta(dt)$, its autocovariance derivation requires some additional simplifications.

First, using Bergstrom [1997], McCrorie [2000] and Chambers [1999], the double integral following (2.7) and (2.8) can be divided into two single integrals using the following interchange of the orders of integration[5]

$$\begin{aligned} v_1 &= \int_0^1 \int_0^r f(r-s)\zeta(ds)dr, \\ &= \int_0^1 \left(\int_s^1 f(r-s)dr \right) \zeta(ds), \\ v_t &= \int_{t-1}^t \int_{r-1}^r f(r-s)\zeta(ds)dr, \qquad t = 2, \cdots, T, \\ &= \int_{t-1}^t \left(\int_s^t f(r-s)dr \right) \zeta(ds) + \int_{t-2}^{t-1} \left(\int_{t-1}^{s+1} f(r-s)dr \right) \zeta(ds), \end{aligned} \tag{2.14}$$

were $f(r-s) = I_n + G(e^{(r-s)M} - I_{n_1})B'$.

And second, solving the integrals in brackets (see Appendix B for details), the final expression of v_t is

$$\begin{aligned} v_1 &= \int_0^1 \left\{(1-s)(I_n - GB') + GM^{-1}\left[e^{(1-s)M} - I_{n_1}\right]B'\right\}\zeta(ds), \\ v_t &= \int_{t-1}^t \left\{(t-s)(I_n - GB') + GM^{-1}\left[e^{(t-s)M} - I_{n_1}\right]B'\right\}\zeta(ds), \\ &+ \int_{t-2}^{t-1} \left\{(s-t+2)(I_n - GB') + GM^{-1}\left[e^{M} - e^{(t-s-1)M}\right]B'\right\}\zeta(ds). \end{aligned} \tag{2.15}$$

Then, once the discrete time disturbance vector v_t is expressed in terms

[5]See Rozanov [1967] Theorem 2.4, p 12 for details.

of single integrals, we can easily apply a generalization of (2.11) to (2.15) and obtain the desired autocovariance matrix of the disturbance vector. Next Lemma presents the exact representation.

Lemma 2.3.1 (Autocovariance representation of v_t).

Following the assumptions of $\zeta(ds)$, the representation of v_t in equation (2.15) and defining the $nT \times 1$ vector $v = [v_1', \ldots, v_T']'$, the autocovariance representation of v_t is

$$\Omega = E[vv'], \tag{2.16}$$

where

$$\Omega = \begin{pmatrix} \Omega_{00} & \Omega_{01}' & 0 & 0 & \cdots & 0 \\ \Omega_{01}' & \Omega_0 & \Omega_1' & 0 & \cdots & 0 \\ 0 & \Omega_1 & \Omega_0 & \Omega_1' & \cdots & 0 \\ 0 & 0 & \Omega_1 & \Omega_0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \Omega_1 & \Omega_0 \end{pmatrix},$$

and

$$\begin{aligned} E[v_1 v_1'] =& \Omega_{00} = \int_0^1 \Xi_1(1-s)\,\Sigma\Xi_1'(1-s)\,ds, \\ E[v_1 v_2'] =& \Omega_{01} = \int_0^1 \Xi_2(1-s)\,\Sigma\Xi_1'(1-s)\,ds, \\ E[v_t v_t'] =& \Omega_0 = \int_0^1 \Xi_1(s)\,\Sigma\Xi_1'(s)\,ds + \int_0^1 \Xi_2(s)\,\Sigma\Xi_2'(s)\,ds, \\ E\left[v_t v_{t-1}'\right] =& \Omega_1 = \int_0^1 \Xi_2(s)\,\Sigma\Xi_1'(s)\,ds, \end{aligned} \tag{2.17}$$

with

$$\begin{aligned} \Xi_1(s) &= (s)(I_n - GB') + GM^{-1}\left[e^{sM} - I_{n_1}\right]B', \\ \Xi_2(s) &= (1-s)(I_n - GB') + GM^{-1}\left[e^{M} - e^{sM}\right]B'. \end{aligned} \tag{2.18}$$

Proof. See Appendix B. ■

As expected, the autocovariances of v_t depend directly on the autocovariance properties of $\zeta(ds)$, therefore all calculations are reduced to solve Ω_{00}, Ω_{01}, Ω_0 and

Ω_1. Also, from (2.17), it is easy to notice that the discrete time disturbance vector v_t follows a moving average process of order one.

Finally, with Ω as in Lemma 2.3.1, the logarithm of the Gaussian likelihood function can be written as

$$L(\theta, \Sigma) = -\frac{nT}{2}\ln(2\pi) - \frac{1}{2}\ln|\Omega| - \frac{1}{2}v'\Omega^{-1}v, \tag{2.19}$$

where θ and Σ are specified as before.

As with stocks, the computation of the likelihood function, for flows, requires the determinant and inverse of the respective covariance matrix, however, as we can see from (2.16), this matrix is now a $nT \times nT$ sparse matrix whose elements are very complicated expressions involving Ω_{00}, Ω_{01}, Ω_0 and Ω_1, as a result, for computational purposes, (2.19) may be not very convenient in our context. For that reason, rather than attempting to optimize directly L, we will follow Bergstrom [1985] and provide an alternative recursive algorithm that avoids these calculations, exploits the sparse nature of Ω and more importantly, yields the estimates of the parameter in our model; the algorithm proceeds as follows

Let P be the real lower triangular matrix, with positive elements along the diagonal, such that

$$PP' = \Omega, \tag{2.20}$$

$$P = \begin{pmatrix} P_{11} & 0 & 0 & \cdots & 0 \\ P_{21} & P_{22} & 0 & \cdots & 0 \\ 0 & P_{32} & P_{33} & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & P_{T-1,T-2} & P_{T-1,T-1} & 0 \\ 0 & 0 & 0 & P_{T,T-1} & P_{TT} \end{pmatrix},$$

where the matrices $P_{11}, P_{21}, P_{22}, P_{t,t-1}, P_{tt}$ $(t = 3, \ldots, T)$ can be computed recursively using

$$P_{11}P'_{11} = \Omega_{00}, \tag{2.21}$$

$$P_{21} = \Omega_{01}(P'_{11})^{-1}, \tag{2.22}$$

$$P_{22}P'_{22} = \Omega_0 - P_{21}P'_{21}, \tag{2.23}$$

$$P_{32} = \Omega_1 (P'_{22})^{-1}, \tag{2.24}$$

$$P_{33}P'_{33} = \Omega_0 - P_{32}P'_{32}, \tag{2.25}$$

for $t = 4, 5, \cdots, T$,

$$P_{t,t-1} = \Omega_1 (P'_{t-1,t-1})^{-1}, \tag{2.26}$$

$$P_{tt}P'_{tt} = \Omega_0 - P_{t,t-1}P'_{t,t-1}. \tag{2.27}$$

Let the $nT \times 1$ vector $\varepsilon = [\varepsilon'_1, \ldots, \varepsilon'_T]'$ be defined as

$$P\varepsilon = v, \tag{2.28}$$

so that

- $\mathrm{E}[\varepsilon] = 0, \qquad \mathrm{E}[\varepsilon\varepsilon'] = I_{nT\times nT},$
- $\mathrm{E}[\varepsilon_t] = 0, \qquad \mathrm{E}[\varepsilon_t\varepsilon'_t] = I_{n\times n}\ , \qquad (t = 1, \ldots, T),$
- $\mathrm{E}[\varepsilon_t\varepsilon'_s] = 0\ , \qquad (s \neq t; s, t = 1, \ldots, T).$

Then, minus twice the logarithm of the likelihood function L (ignoring the constant) is given by

$$\mathcal{L} = \sum_{i=1}^{nT} (\varepsilon_i^2 + 2\log p_{ii}), \tag{2.29}$$

where p_{ii} is the ith diagonal element of P and the nT elements of ε are computed in T vectors of size n using recursively the following procedure

$$\begin{aligned} \varepsilon_1 &= \Big(\varepsilon_{11}, \ldots, \varepsilon_{1n}\Big)' = P_{11}^{-1} v_1, \\ \varepsilon_t &= \Big(\varepsilon_{t1}, \ldots, \varepsilon_{tn}\Big)' = P_{tt}^{-1}(v_t - P_{t,t-1}\varepsilon_{t-1}), \qquad t = 2, 3, \cdots, T. \end{aligned} \tag{2.30}$$

Indeed, by using equation (2.29), the computation of our Gaussian estimates becomes simpler; it not only avoids the calculations of the inverse and determinant of Ω, but also, as we can see from (2.20), uses a Cholesky factorization of Ω that automatically takes into account its sparse nature, therefore, is computationally more efficient than the standard Cholesky factorization. Also, all computations are further simplified by the fact that the sequence of $n \times 2n$ matrices $\overline{P_t} = \big(P_{tt}, P_{t,t-1}\big)$ $t = 2, 3, \ldots, T$ converges very rapidly to a constant limit matrix which is the solution to the nonlinear second-order difference equations system given in (2.22) and (2.23). See Bergstrom [1990] chapter 7 for details.

Therefore, the Gaussian estimates $\hat{\theta}$ can be calculated following the next list of steps.

1. Compute A, B and Σ using the specified forms of the functions defining their elements.

2. Given the specific representations of the matrices G, B and Σ compute (see Appendix A for details) e^M together with Ω_{00}, Ω_{01}, Ω_0 and Ω_1 from their specifications as given in Lemma 2.3.1.

3. Given the numerical representations of Ω_{00}, Ω_{01}, Ω_0 and Ω_1, compute the Cholesky factorization of the matrix Ω (the matrix P) following recursively the steps on the set of equations from (2.21) to (2.27) as follows

   (a) Calculate the Cholesky factorization of Ω_{00} as in (2.21).

   (b) Calculate P_{21} as in (2.22) and the Cholesky factorization of $(\Omega_0 - P_{21}P'_{21})$ as in (2.23).

   (c) Calculate P_{32} as in (2.24) and the Cholesky factorization of $(\Omega_0 - P_{32}P'_{32})$ as in (2.25).

   (d) Setting a stop value sufficiently small, repeat x times step (c) for P_{xx} and $P_{x,x-1}$ until the differences between their values are equal or less than the stop value.

   (e) Generate the Cholesky factorization of Ω from (2.20) by using the different matrices from steps (a), (b) ,(c) and (d) and complete the $(x + 4, x + 5, \ldots, T)$ remaining matrices in P as a copy of the limit matrix obtained in step (d).

4. For the minimization of $\mathcal{L}$, with the data and allowing the model's parameters to vary, obtain a new P as in step 3 and ε recursively as in (2.30).

5. Set this new P and ε into (2.29) and calculate $\mathcal{L}$.

6. Repeat steps 4 and 5 until a minimum is achieved and take those $\hat{\theta}$ and $\hat{\Sigma}$ as the elements that minimize $\mathcal{L}$.

2.4 Simulation and comparison

Up until now, we have developed an exact discrete time model that corresponds to a cointegrated continuous time system. Also, with the characterization of the covariance properties of the discrete time disturbance vectors, we have stated, for estimation, a computationally efficient likelihood function for each of our model's specifications. More importantly, we have outlined a set of steps that leads to the Gaussian estimates of our model's parameters. Then, in this section, we evaluate its performance and accuracy against one of the most commonly used methodologies among cointegration literature; Johansen's approach to cointegration modelling[6].

The interest of such assessment, as pointed out before, rests mainly in the two different specifications of the systems under consideration. While the primary representations, given by VECM, are almost identical in both the exact discrete time representation of our cointegrated continuous time system and Johansen's VECM procedure for cointegration, the latter is naively specified in terms of the observation interval and the former is temporally aggregated. Therefore, given that Johansen's methodology ignores temporal aggregation, it is of interest to measure how accurate its estimated parameters are in our continuous time system.

Of course, given the inherent nature of the systems, such comparison cannot be directly measured, nevertheless, as we shall see below, we can accomplish the task by deriving an exact link function between the two specifications that translates the values of the estimated parameters in our cointegrated continuous time system to those we would have obtained by using, instead, Johansen's methodology.

First of all, for this application, let's define the model under consideration as a simplification of (2.1) in which there is only one cointegrating relationship, $(1, -b_1)$, contained in the matrix B' and only two speed of adjustment parameters, a_1 and a_2, contained in the matrix A, at the same time, we set $n = 2$ so that $n_1 = n_2 = 1$ and $y(0) = 0$, then, using Lemma 2.2.2, our exact discrete time

[6]See Johansen [1988, 1991] for details.

VECMs for estimation are given by

$$\begin{pmatrix} \Delta y_{1t} \\ \Delta y_{2t} \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} (a_1 - b_1 a_2)^{-1}(e^{a_1 - b_1 a_2} - 1)(y_{1,t-1} - b_1 y_{2,t-1}) + \eta_t, \quad t = 1, \ldots, T,$$

for stocks, and,

$$\begin{aligned} y_1 &= v_1, \\ \begin{pmatrix} \Delta y_{1t} \\ \Delta y_{2t} \end{pmatrix} &= \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} (a_1 - b_1 a_2)^{-1}(e^{a_1 - b_1 a_2} - 1)(y_{1,t-1} - b_1 y_{2,t-1}) + v_t, \\ t &= 2, \ldots, T. \end{aligned} \tag{2.31}$$

for flows.

Also, for the comparison, let's define the VECM representation in discrete time as

$$\Delta y_t = \gamma \lambda' y_{t-1} + w_t, \tag{2.32}$$

where γ contains the two speed of adjustment parameters (γ_1 and γ_2) of the system, λ' the cointegrating relationship $(1, -\lambda_1)$ and w_t is assumed to be *iid*.

Therefore, the exact link function, the function that relates the parameters in our cointegrated continuous time system with those in its discrete time counterpart, equating term by term (2.31) and (2.32), is given by

$$\begin{aligned} \begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix} &= \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} (a_1 - b_1 a_2)^{-1} \left(e^{a_1 - b_1 a_2} - 1\right), \\ \gamma_1 &= a_1 \left[(a_1 - b_1 a_2)^{-1} \left(e^{a_1 - b_1 a_2} - 1\right)\right], \\ \gamma_2 &= a_2 \left[(a_1 - b_1 a_2)^{-1} \left(e^{a_1 - b_1 a_2} - 1\right)\right], \end{aligned} \tag{2.33}$$

$$\lambda_1 = b_1. \tag{2.34}$$

As a result, if we want to measure how accurate Johansen's estimates are, in terms of closeness with the ones we get by applying our methodology, all we need to compute are the *implied*[7] values through (2.33) and (2.34). Of course,

[7]Subsequently, these estimates will be referred as the discrete time counterpart of those we would have obtained in (2.13) and/or (2.29).

in the derivation of this link function, identification in the system is achieved by considering the required normalization of the cointegrating relationship following (2.1).

As we can see from (2.33) and (2.34), more than just indicating the relationship between the two specifications, this link function provides many interesting insights of them. For instance, if we look at (2.34), we see that the cointegrating parameter appears exactly matched in the two specifications. As a result, the long run parameter of our cointegrated continuous time model can be estimated directly from its corresponding discrete time VECM representation (2.32), i.e., at least for this parameter, there is no aliasing or identification problem (see Phillips [1991] for details). Also, if we look at (2.33), aggregation becomes evident in our specification and given that Johansen's specification ignores it, equation (2.33) plays a crucial role in our analysis.

For this application, we generate our own observations through a simulation technique that specifies (2.31) as the data generating process (DGP) with three different parametric designs. For simplicity, in all three designs, we normalize the cointegrating parameter to be 1 so that the cointegrating relationship is given by $y_{1t} - y_{2t}$ and as a result, in this experiment, we stress mainly the implications of dynamics over the performance of estimation in the system. The exact representations are as follows

$$\begin{aligned} &\text{Design 1:} \quad \theta^0 = [a_1^0, a_2^0, b_1^0] = [1, 2, 1], \\ &\text{Design 2:} \quad \theta^0 = [a_1^0, a_2^0, b_1^0] = [-2, -1, 1], \\ &\text{Design 3:} \quad \theta^0 = [a_1^0, a_2^0, b_1^0] = [-0.4, 0.6, 1]. \end{aligned} \tag{2.35}$$

As we can see from (2.35), each parametric design explores a particular effect of dynamic adjustments over the system; while design 1 allows for a positive feedback between y_1 and y_2, design 2 changes it to be negative and design 3 makes it reciprocal. As a result, given the symmetry of the specification in our DGP, with these 3 experimental designs, we cover all possible combinations and analyse robustness in our estimation procedure.

Here, it is important to point out that the chosen values in these designs were obtained as representative elements of the feasible set of values in the parameter space that, at the same time, fulfilled all the requirements[8] for the system to be as stated in the assumptions of (2.1), therefore, the results in this application are generalizable.

In all three experimental designs, at the same time, the covariance matrix is considered as follows

$$\Sigma = \begin{bmatrix} \sigma_1 & \sigma_3 \\ \sigma_3 & \sigma_2 \end{bmatrix} = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}, \tag{2.36}$$

and results, then, are reported for $\rho = -0.5$ and $\rho = 0.5$ so that there is positive and negative correlation in the system. Note that in estimation, Σ is ensured to be positive definite by computing, instead, estimates of the the lower triangular matrix R, such that, $\Sigma = RR'$; these matrices are related as follows:

$$r_1 = 1, \quad r_2 = \rho, \quad r_3 = \sqrt{1-\rho^2}.$$

Then, the estimates of the structural parameters of our system $\hat{\theta}$ and $\hat{\Sigma}$ are obtained through the application of the methodologies as described before in subsections 2.3.1 and 2.3.2.

For the comparison, the *implied* true parametric designs for (2.32), which are calculated using (2.35) on (2.33) and (2.34), are given by:

$$\begin{aligned} &\text{Design 1:} \quad \theta_j^0 = [\gamma_1^0, \gamma_2^0, \lambda_1^0] = [0.632, 1.26, 1.0], \\ &\text{Design 2:} \quad \theta_j^0 = [\gamma_1^0, \gamma_2^0, \lambda_1^0] = [-1.26, -0.632, 1.0], \\ &\text{Design 3:} \quad \theta_j^0 = [\gamma_1^0, \gamma_2^0, \lambda_1^0] = [-0.253, 0.379, 1.0]. \end{aligned} \tag{2.37}$$

Then, the *implied* estimated values $\hat{\theta}_j$ are obtained by applying Johansen's methodology to the same simulated data as if it was generated by the VECM representation of order 1 given in (2.32).

Finally, performance of the method is analysed by measuring accuracy in estimation, which for our purposes, is defined as closeness between the estimated parameters $\hat{\theta}$ as well as $\hat{\theta}_j$ and their true values in (2.35) and (2.37). The procedure

[8]These requirements are as follows: (1) The eigenvalues of M' have non positive real parts, (2) the VAR that follows the system is stable and (3) the orthogonal complement matrices of the VAR are non-singular.

is described below, and as we have two different types of data, the first subsection focuses on the case when the variables of interest are stocks and the second when they are flows.

2.4.1 VECM simulations with stocks

For stocks, given by (2.6) and (2.31), the data generating process, which is used to generate ten thousand simulations of 50, 100 and 200 sample sizes, follows a VECM representation of order 1 that can be written as

$$\begin{pmatrix} \Delta y_{1t} \\ \Delta y_{2t} \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} (a_1 - b_1 a_2)^{-1}(e^{a_1 - b_1 a_2} - 1)(y_{1,t-1} - b_1 y_{2,t-1}) + \eta_t, \tag{2.38}$$

where η_t is assumed to be $N(0, W)$ with W as in (2.11).

The true parameter values are specified as in (2.35) and estimation of (2.38) follows the methodology described in subsection 2.3.1. For the discrete time cointegrated counterpart, additionally, the model is assumed to follow the VECM(1) specification given in (2.32) with true parameter values as in (2.37), as a result, for estimation, Johansen's methodology is applied to the same simulated data. Comparison, then, is established by measuring closeness between the estimated parameters of the two different methods with the true values in (2.35) and (2.37).

Results have been grouped and appear in tables 2.1, 2.2 and 2.3. Table 2.1 shows the results for design 1, table 2.2 for design 2 and 2.3 for design 3. Each table is divided into two sub tables for positive and negative correlation. Each sub table, in turn, is showing the true value, the bias and the standard error of each of the parameters for the 3 sample sizes in this exercise. Also, these sub tables are divided according to the estimation methodology so that the upper part of each sub table refers particularly to the results obtained through the application of our methodology whereas, the lower part refers to the result when Johansen's methodology is applied.

As we can see from tables 2.1, 2.2 and 2.3, in all the cases and for all the parameters, both Johansen's methodology and ours perform very well, with reasonably small bias in estimation and improvements as the sample size increases. However, Johansen's methodology shows a small superiority in estimation, with slightly lower bias in the parameters and less dispersion. If we look particularly at the cointegrating parameter in the system, for all the three designs, a similar but better pattern emerged with very small bias in estimation and improvements as the sample size increases, which is expected due to the consistency of this parameter (see equation 2.34). Also, in all three different parametric designs, for both our methodology and Johansen's and almost in all the cases, when the correlation parameter changes from positive ($\rho = 0.5$) to negative ($\rho = -0.5$), we see important reductions in the bias and the standard error of the estimates. Then, we can say that our methodology is robust against changes in the parametric specification of the system.

Indeed, these results are not surprising and they are as expected due to the fact that the models under consideration are exactly specified; both are expressed as VECMs of order 1, the disturbance structures are correct both being $i.i.d$, and even though ours is temporally aggregated, it is absorbed by the exact link function. As a result, the model in (2.32) together with (2.33) and (2.34), can be referred as the discrete time counterpart of the exact discrete time representation of our continuous time model in (2.1).

Table 2.1 Design 1 estimates for stock variables
(positive and negative correlation)

Positive correlation							
		Sample size			Sample size		
Continuous time estimates		50	100	200	50	100	200
			Bias			Standard Error	
Parameter	True value						
a_1	1	0.02793	0.00818	0.00585	0.38077	0.24538	0.17494
a_2	2	0.18196	0.09215	0.06470	0.56784	0.34062	0.23272
b_1	1	-0.00022	-0.00014	-0.00002	0.01844	0.00839	0.00406
σ_1	1	-0.03065	-0.01278	-0.00293	0.14638	0.09860	0.07124
ρ	0.5	-0.08600	-0.02819	-0.00819	0.27279	0.16786	0.11522
σ_2	1	-0.13083	-0.03177	0.01264	0.09560	0.03666	0.01828
Johansen's VECM(1) estimates							
γ_1	0.632	-0.04891	-0.01864	-0.00910	0.26869	0.18408	0.12623
γ_2	1.264	0.00808	0.00956	0.00716	0.21034	0.14322	0.09903
λ_1	1	-0.00048	-0.00010	0.00001	0.02672	0.01188	0.00565

Negative correlation							
		Sample size			Sample size		
Continuous time estimates		50	100	200	50	100	200
			Bias			Standard Error	
Parameter	True value						
a_1	1	-0.00101	0.00073	0.00104	0.20521	0.14131	0.09887
a_2	2	0.04057	0.01490	0.00097	0.29733	0.20920	0.14642
b_1	1	0.00003	0.00001	0.00003	0.01246	0.00569	0.00276
σ_1	1	-0.02933	-0.01271	-0.00384	0.16895	0.11953	0.08460
ρ	-0.5	-0.02038	0.01709	0.01382	0.33288	0.26346	0.19586
σ_2	1	-0.29236	-0.19993	-0.1409	0.23056	0.13924	0.07100
Johansen's VECM(1) estimates							
γ_1	0.632	-0.05100	-0.02256	-0.01078	0.19493	0.13695	0.09449
γ_2	1.264	-0.00303	-0.00070	0.00018	0.11194	0.07667	0.05321
λ_1	1	-0.00014	-0.00001	0.00002	0.01774	0.00796	0.00383

Table 2.2 Design 2 estimates for stock variables
(positive and negative correlation)

Positive correlation							
		Sample size			Sample size		
Continuous time estimates		50	100	200	50	100	200
		Bias			Standard Error		
Parameter	True value						
a_1	-2	-0.13135	-0.02401	0.02329	0.53704	0.29661	0.18538
a_2	-1	0.02491	0.02450	0.02440	0.37222	0.23130	0.15778
b_1	1	-0.00022	-0.00010	-0.00009	0.01853	0.00834	0.00408
σ_1	1	-0.03437	-0.01665	-0.00588	0.14314	0.09445	0.06618
ρ	0.5	-0.06975	-0.02176	0.00253	0.25794	0.15814	0.10214
σ_2	1	-0.11745	-0.03784	-0.00508	0.08511	0.03301	0.01432
Johansen's VECM(1) estimates							
γ_1	-1.264	-0.00859	-0.00151	-0.00279	0.21535	0.14256	0.09963
γ_2	-0.632	0.04393	0.02479	0.01042	0.27269	0.18154	0.12715
λ_1	1	-0.00020	-0.00007	0.00002	0.02603	0.01185	0.00568

Negative correlation							
		Sample size			Sample size		
Continuous time estimates		50	100	200	50	100	200
		Bias			Standard Error		
Parameter	True value						
a_1	-2	-0.03990	-0.01885	-0.00983	0.27842	0.18621	0.12893
a_2	-1	0.00570	0.00663	0.00239	0.20891	0.13908	0.09655
b_1	1	-0.00010	-0.00005	-0.00002	0.01262	0.00563	0.00276
σ_1	1	-0.03071	-0.01630	-0.00721	0.16558	0.11722	0.08121
ρ	-0.5	-0.02310	-0.02149	-0.01125	0.32103	0.24428	0.18066
σ_2	1	-0.21880	-0.11457	-0.05758	0.19560	0.10791	0.05322
Johansen's VECM(1) estimates							
γ_1	-1.264	0.00252	0.00252	-0.00002	0.11390	0.07606	0.05311
γ_2	-0.632	0.04925	0.02659	0.01184	0.19793	0.13400	0.09480
λ_1	1	-0.00015	-0.00003	0.00002	0.01748	0.00801	0.00383

Table 2.3 Design 3 estimates for stock variables (positive and negative correlation)

Positive correlation							
		Sample size			Sample size		
Continuous time estimates		50	100	200	50	100	200
		Bias			Standard Error		
Parameter	True value						
a_1	-0.4	-0.17154	-0.04965	-0.03150	0.84871	0.27097	0.17151
a_2	0.6	0.18117	0.04915	0.01092	1.04096	0.37710	0.17149
b_1	1	-0.00276	-0.00037	0.00029	0.06970	0.02943	0.01471
σ_1	1	0.00436	-0.00385	-0.00147	0.15559	0.08653	0.05844
ρ	0.5	-0.06202	-0.02142	-0.01140	0.24066	0.12323	0.07934
σ_2	1	-0.05931	-0.02604	-0.01788	0.08352	0.02432	0.01031
Johansen's VECM(1) estimates							
γ_1	-0.25	-0.04377	-0.01763	-0.01063	0.22907	0.14628	0.09823
γ_2	0.38	0.02306	0.01690	0.00818	0.21502	0.13990	0.09542
λ_1	1	-0.00495	-0.00072	0.00003	0.13685	0.04336	0.02055
Negative correlation							
		Sample size			Sample size		
Continuous time estimates		50	100	200	50	100	200
		Bias			Standard Error		
Parameter	True value						
a_1	-0.4	-0.12986	-0.04561	-0.01936	0.50061	0.18279	0.10804
a_2	0.6	0.15264	0.05325	0.02402	0.52942	0.19113	0.11320
b_1	1	-0.01912	-0.00440	-0.00039	0.19997	0.08622	0.03927
σ_1	1	0.02925	0.00908	0.00297	0.20371	0.11445	0.07429
ρ	-0.5	-0.06503	-0.02272	-0.01006	0.26653	0.15176	0.09883
σ_2	1	0.02567	0.00032	-0.00138	0.08249	0.02809	0.01231
Johansen's VECM(1) estimates							
γ_1	-0.25	-0.03908	-0.01995	-0.01090	0.12158	0.07538	0.04965
γ_2	0.38	0.02879	0.01497	0.00768	0.10121	0.06337	0.04327
λ_1	1	-0.03981	-0.00812	-0.00090	0.73388	0.12263	0.05499

2.4.2 VECM simulations with flows

For flows, the data generating process, which is used to generate ten thousand simulations of 50, 100 and 200 observations, follows a VECM representation with a moving average component in the error term of order 1 given by (2.7), (2.28) and (2.31) that can be written as

$$\begin{aligned} y_1 &= P_{11}\varepsilon_1, \\ \begin{pmatrix} \Delta y_{1t} \\ \Delta y_{2t} \end{pmatrix} &= \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} (a_1 - b_1 a_2)^{-1}(e^{a_1 - b_1 a_2} - 1)(y_{1,t-1} - b_1 y_{2,t-1}) + \\ &+ P_{tt}\varepsilon_t + P_{t,t-1}\varepsilon_{t-1}, \\ t &= 2, \ldots, T, \end{aligned} \tag{2.39}$$

where ε_t is coming from (2.28) and v in this same expression is assumed to be $N(0, \Omega)$ with Ω as in Lemma 2.3.1.

The true parametric designs, as with stocks, are given in (2.35) as well as (2.37) and for estimation, we apply our methodology as described in subsection 2.3.2 to the generated data as well as Johansen's to it as if it was generated by the VECM(1) specification given in (2.32).

Additionally, as there is a moving average component included in (2.39) that is being ignored by Johansen's VECM(1) specification, we take a step further in the analysis and in looking for a more accurate discrete time estimation of (2.39), we also apply Johansen's methodology to our generated data as if it was specified through a VECM representation of order 2 given by

$$\Delta y_t = \gamma \lambda' y_{t-1} + \Gamma \Delta y_{t-1} + s_t, \tag{2.40}$$

where s_t is assumed to be *iid*, γ, as before, is the matrix that includes the two speed of adjustment parameters (γ_1 and γ_2), λ' contains the unique cointegrating relationship $(1, -\gamma_1)$ and Γ is the matrix of coefficients that relates Δy_t with its lagged value and can be written as

$$\Gamma = \begin{pmatrix} \Gamma_1 & \Gamma_2 \\ \Gamma_3 & \Gamma_4 \end{pmatrix}$$

Finally, for comparison, we consider the relevant parameter estimates of this system ($\hat{\gamma}_1, \hat{\gamma}_2$ and $\hat{\lambda}_1$) and measure how close they are with the respective true parameter values in (2.37).

Results have been grouped and appear in tables 2.4 to 2.9 so that each design is explained in two tables with the first focusing in the results when the correlation is positive and the second when it is negative. As with stocks, each table is showing the true value, the bias and the standard error of each of the parameters for the 3 sample sizes in the exercise, however, in this new scenario, each table is divided in three sections; the upper part displays the results obtained through the application of our methodology, the middle part displays them when instead Johansen's methodology is applied and a VECM(1) specification is considered and finally, the lower part shows them when Johansen's is applied and a VECM(2) specification is considered. It is important to notice, for the lower part of these tables, that the *implied true* values for the matrix of coefficients Γ are not known, as a result, we cannot report the bias, instead, we are reporting only the mean value and the standard error of the estimates of these particular parameters.

As we can see from the tables, in almost all the cases, our methodology shows superiority in estimation against Johansen's with smaller bias in the estimates, however, they are concentrated about a mean with greater dispersion. Additionally, as expected, the estimates of the long run equilibrium parameter of the model in all three parametric designs show the smallest bias and standard deviation. For the change in the correlation from positive to negative, in almost all the cases and all three experimental designs, as with stocks, we see an important reduction in the standard error of the estimates.

Considering Johansen's VECM(1) specification (the middle part of the tables) and focusing only on the dynamics of the system (γ_1 and γ_2), the tables show a persistent bias in estimation with almost no improvement as the sample size increases, consequently, even though these estimates are reasonably close to the true parameter value, they clearly reflect the cost of ignoring aggregation in the specification.

For Johansen's VECM(2) specification (the lower part of the tables), additionally, we see a mixture of effects; on the one hand, when the correlation parameter is positive and the experimental designs are 1 and 2 (tables 2.4 and

2.6), the estimates of such specification are better than those we get by applying Johansen's VECM(1) specification, however, in some of the cases they are showing an increasing bias as the sample size increases. On the other hand, when the correlation parameter is negative, for the same tables, results are exactly the opposite and moreover, for design 3 (tables 2.8 and 2.9), regardless of the value of the correlation parameter, Johansen's VECM(2) estimates are always better even though they are showing an increasing bias. As a result, we cannot claim that the VECM(2) specification is capturing better the moving average component in our continuous time model. However, it is important to notice that the true link function between our system and the discrete time VECM(2) specification is not exactly given by (2.33) and (2.34), then, our reported values in these table may not be precisely measured, hence, the previous claim has to be taken with care.

Finally, as with stocks, given these results, we can say that our methodology is robust against changes in the parametric specifications of the system.

Table 2.4 Design 1 estimates for flow variables (positive correlation)

Positive correlation							
		Sample size			Sample size		
Continuous time estimates		50	100	200	50	100	200
		Bias			Standard Error		
Parameter	True value						
a_1	1	-0.07894	-0.00389	0.00383	3.09537	0.21924	0.15004
a_2	2	0.08740	0.00389	-0.00182	1.59327	0.22278	0.16290
b_1	1	0.00008	0.00005	-0.00006	0.01582	0.00737	0.00356
σ_1	1	0.00478	-0.00432	0.00210	1.21466	0.09008	0.06361
ρ	0.5	-0.04918	0.01079	0.00587	0.66906	0.12899	0.10139
σ_2	1	0.02407	0.02173	0.02065	0.46770	0.02557	0.01503
Johansen's VECM(1) estimates							
γ_1	0.633	0.35797	0.35615	0.35182	0.23715	0.16250	0.11179
γ_2	1.264	0.20694	0.19501	0.19229	0.19988	0.13786	0.09529
λ_1	1	0.00088	0.00047	0.00018	0.01587	0.00736	0.00353
Johansen's VECM(2) estimates							
γ_1	0.633	0.02929	0.02499	0.02217	0.46279	0.29801	0.20389
γ_2	1.264	-0.11707	-0.14728	-0.16168	0.33035	0.21656	0.14898
λ_1	1	-0.00022	-0.00004	-0.00006	0.01821	0.00776	0.00361
Parameter		Mean of the estimated parameter					
Γ_1		0.43392	0.44544	0.45090	0.39311	0.26412	0.17982
Γ_2		-0.32141	-0.32184	-0.32000	0.29069	0.20016	0.13405
Γ_3		0.34440	0.36586	0.37739	0.29323	0.19825	0.13486
Γ_4		-0.18554	-0.19311	-0.19632	0.22758	0.15544	0.10358

Table 2.5 Design 1 estimates for flow variables (negative correlation)

Negative correlation							
		Sample size			Sample size		
Continuous time estimates		50	100	200	50	100	200
		Bias			Standard Error		
Parameter	True value						
a_1	1	-0.03964	-0.01711	-0.01105	0.21341	0.11510	0.07037
a_2	2	0.03360	-0.01500	-0.00524	0.20825	0.12314	0.09757
b_1	1	0.00015	0.00011	0.00005	0.00688	0.00323	0.00160
σ_1	1	0.01062	-0.01019	-0.01061	0.15413	0.07978	0.03982
ρ	-0.5	-0.03760	-0.01427	-0.00438	0.16877	0.07769	0.03575
σ_2	1	0.03580	0.03537	0.03414	0.03910	0.00927	0.00225
Johansen's VECM(1) estimates							
γ_1	0.633	0.35100	0.35606	0.36005	0.14464	0.10016	0.06803
γ_2	1.264	0.19460	0.19035	0.18949	0.07215	0.04987	0.03423
λ_1	1	-0.00066	-0.00030	-0.00016	0.00689	0.00319	0.00152
Johansen's VECM(2) estimates							
γ_1	0.633	0.97165	0.96148	0.95883	0.63772	0.42227	0.29473
γ_2	1.264	0.37377	0.34880	0.33810	0.29051	0.19751	0.13876
λ_1	1	0.00028	0.00013	0.00004	0.00739	0.00328	0.00154
Parameter		Mean of the estimated parameter					
Γ_1		-0.32657	-0.31503	-0.31159	0.53499	0.35721	0.25038
Γ_2		-0.15450	-0.15223	-0.14839	0.23435	0.15874	0.10911
Γ_3		-0.05739	-0.03930	-0.03063	0.24100	0.16464	0.11604
Γ_4		-0.10174	-0.10460	-0.10575	0.10949	0.07369	0.05031

Table 2.6 Design 2 estimates for flow variables (positive correlation)

Positive correlation							
		Sample size			Sample size		
Continuous time estimates		50	100	200	50	100	200
		Bias			Standard Error		
Parameter	True value						
a_1	-2	-0.05117	0.03825	0.03760	0.39590	0.20067	0.12588
a_2	-1	0.04802	0.03007	0.02276	0.41006	0.20915	0.15303
b_1	1	-0.00011	-0.00009	-0.00005	0.01589	0.00711	0.00351
σ_1	1	-0.01263	-0.01136	-0.01142	0.14871	0.07928	0.05128
ρ	0.5	-0.03921	0.03121	0.02255	0.22600	0.11949	0.08264
σ_2	1	-0.05680	-0.02179	-0.01370	0.07247	0.02364	0.01224
Johansen's VECM(1) estimates							
γ_1	-1.26	-0.21074	-0.20055	-0.19622	0.20460	0.13811	0.09585
γ_2	-0.63	-0.36416	-0.36512	-0.36449	0.23959	0.16286	0.11329
λ_1	1	-0.00083	-0.00037	-0.00018	0.01593	0.00714	0.00350
Johansen's VECM(2) estimates							
γ_1	-1.26	0.10622	0.13911	0.15784	0.33226	0.22023	0.14778
γ_2	-0.63	-0.03466	-0.03353	-0.02495	0.45871	0.30220	0.20411
λ_1	1	0.00006	0.00011	0.00005	0.01843	0.00755	0.00359
Parameter		Mean of the estimated parameter					
Γ_1		-0.18468	-0.19241	-0.19618	0.22772	0.15182	0.10425
Γ_2		0.34075	0.36406	0.37740	0.29134	0.19724	0.13466
Γ_3		-0.32410	-0.32047	-0.32035	0.29043	0.19673	0.13659
Γ_4		0.43802	0.44315	0.45102	0.38907	0.26274	0.18182

Table 2.7 Design 2 estimates for flow variables (negative correlation)

Negative correlation							
		Sample size			Sample size		
Continuous time estimates		50	100	200	50	100	200
		Bias			Standard Error		
Parameter	True value						
a_1	-2	-0.02608	0.01861	0.01533	0.20047	0.12529	0.08739
a_2	-1	0.02654	-0.02477	-0.02379	0.19672	0.11316	0.07265
b_1	1	-0.00009	-0.00002	-0.00005	0.00689	0.00313	0.00160
σ_1	1	-0.03113	-0.02132	-0.02035	0.14494	0.09080	0.05993
ρ	-0.5	-0.03060	-0.00486	0.00167	0.17585	0.08646	0.04269
σ_2	1	-0.07362	-0.06578	-0.06411	0.04216	0.01105	0.00292
Johansen's VECM(1) estimates							
γ_1	-1.26	-0.19686	-0.19523	-0.19381	0.07394	0.04976	0.03444
γ_2	-0.63	-0.37097	-0.36124	-0.36082	0.14330	0.09945	0.06955
λ_1	1	0.00066	0.00035	0.00016	0.00690	0.00309	0.00152
Johansen's VECM(2) estimates							
γ_1	-1.26	-0.37980	-0.35617	-0.34230	0.28752	0.20112	0.13801
γ_2	-0.63	-0.97440	-0.97292	-0.96231	0.63017	0.42936	0.29417
λ_1	1	-0.00026	-0.00009	-0.00004	0.00741	0.00321	0.00154
Parameter		Mean of the estimated parameter					
Γ_1		-0.10059	-0.10403	-0.10544	0.10710	0.07307	0.05051
Γ_2		-0.06165	-0.04164	-0.03102	0.23697	0.16757	0.11547
Γ_3		-0.15811	-0.15007	-0.14856	0.22802	0.15822	0.11109
Γ_4		-0.32635	-0.32078	-0.31216	0.52569	0.36225	0.25043

Table 2.8 Design 3 estimates for flow variables (positive correlation)

Positive correlation							
		Sample size			Sample size		
Continuous time estimates		50	100	200	50	100	200
		Bias			Standard Error		
Parameter	True value						
a_1	-0.4	-0.10648	-0.05195	-0.04251	0.50057	0.24179	0.15348
a_2	0.6	0.09732	0.01664	-0.01085	0.95310	0.24091	0.15681
b_1	1	-0.00048	0.00028	-0.00005	0.06430	0.02855	0.01354
σ_1	1	0.00537	0.00033	0.00032	0.17827	0.08957	0.05990
ρ	0.5	-0.05466	-0.02138	-0.01187	0.42222	0.12133	0.07892
σ_2	1	-0.04768	-0.03189	-0.02970	0.21730	0.02387	0.01058
Johansen's VECM(1) estimates							
γ_1	-0.25	0.08622	0.09455	0.09960	0.22392	0.14753	0.10032
γ_2	0.38	-0.04686	-0.05985	-0.06413	0.21862	0.14524	0.09920
λ_1	1	0.00093	0.00132	0.00048	0.07010	0.02898	0.01355
Johansen's VECM(2) estimates							
γ_1	-0.25	0.01667	0.02382	0.03285	0.24905	0.15595	0.10205
γ_2	0.38	-0.03964	-0.05649	-0.06352	0.23580	0.15047	0.09966
λ_1	1	-0.00500	-0.00006	-0.00020	0.35901	0.03050	0.01375
Parameter		Mean of the estimated parameter					
Γ_1		0.22229	0.22051	0.21948	0.24203	0.16458	0.11077
Γ_2		0.00966	0.01810	0.02409	0.23695	0.16323	0.10979
Γ_3		0.06270	0.07298	0.07914	0.24030	0.16398	0.11027
Γ_4		0.17355	0.16872	0.16747	0.22980	0.15807	0.10709

Table 2.9 Design 3 estimates for flow variables (negative correlation)

Negative correlation							
		Sample size			Sample size		
Continuous time estimates		50	100	200	50	100	200
		Bias			Standard Error		
Parameter	True value						
a_1	-0.4	-0.08184	-0.03460	-0.01510	0.43756	0.15319	0.09780
a_2	0.6	0.09326	0.03624	0.01508	0.52194	0.15102	0.09765
b_1	1	-0.01682	-0.00232	-0.00088	0.20442	0.07892	0.03767
σ_1	1	0.02600	0.00870	0.00268	0.29671	0.11341	0.07491
ρ	-0.5	-0.05568	-0.02069	-0.00883	0.39576	0.15084	0.10082
σ_2	1	0.02458	0.00102	-0.00230	0.16792	0.02808	0.01279
Johansen's VECM(1) estimates							
γ_1	-0.25	0.08851	0.09533	0.09969	0.09845	0.06480	0.04330
γ_2	0.38	-0.04567	-0.05853	-0.06451	0.09018	0.05933	0.04025
λ_1	1	-0.00893	0.00560	0.00324	0.52871	0.07940	0.03752
Johansen's VECM(2) estimates							
γ_1	-0.25	0.01182	0.02668	0.03503	0.12654	0.07922	0.05163
γ_2	0.38	-0.03985	-0.05665	-0.06535	0.10481	0.06763	0.04580
λ_1	1	-0.02478	-0.00429	-0.00141	0.36611	0.08811	0.03826
Parameter		Mean of the estimated parameter					
Γ_1		0.21527	0.21198	0.21014	0.16734	0.11175	0.07705
Γ_2		0.00506	0.01167	0.01682	0.14844	0.10326	0.07015
Γ_3		0.06605	0.07770	0.08481	0.14917	0.10244	0.07004
Γ_4		0.17435	0.17284	0.17117	0.13104	0.08902	0.06170

2.5 Concluding remarks

In this chapter, a comparison between the estimates of a cointegrated continuous time model and those obtained in its discrete time cointegrated counterpart has been outlined. For that end, an exact discrete time representation that corresponded to a cointegrated continuous time model was derived. The model allowed observations to be entirely comprised by stock or flow variables. When the variables of interest were stocks, it was shown that the exact discrete time specification followed a VECM(1) and a VECM(1) with a moving average component in the error term when the variables were flows. For each specification, an estimation procedure that involved the derivation of the autocovariance properties of the discrete time disturbance vector was stated. In the application, a model that contained only one cointegrating relationship and two speed of adjustment parameters was stated and used to generate some simulated data. For the simulation, three different parametric designs were considered and comparison, was carried out over the estimates we obtained by applying Johansen's methodology and ours to the same simulated data. Of course, due to the inherent nature of the methodologies, for comparison, we employed an exact link function that specified the *implied* estimates we would have obtained by using Johansen's VAR specification into our data.

In the results, as expected, when the system was entirely comprised by stock variables, both Johansen's methodology and ours performed very well, with reasonably small bias in estimation and improvements as the simple size increases. However, when the variables of interest were flows, given that our exact discrete time representation included a moving average component in the error term, Johansen's VECM(1) estimates showed a persistent bias in estimation with almost no improvement as the sample size increased. Consequently, they reflect the cost of ignoring aggregation in the specification. Our methodology, instead, showed superiority in estimation with smaller bias and improvements as the sample size increases, however, they were concentrated about a mean with greater dispersion.

Looking for a more accurate discrete time representation of our model when the system was entirely comprised by flow variables, we also applied Johansen's methodology to our simulated data as if it was generated by a VECM(2) with only one cointegrating relationship and two speed of adjustment parameters. At the end, comparing these estimates with the ones we got in the VECM(1) representation and using the same link function, results showed mixed effects; on one the hand, when the correlation parameter is positive and the experimental designs are 1 and 2 (tables 2.4 and 2.6), the estimates of such specification are better than those we get in Johansen's VECM(1) specification. However, when the correlation parameter is negative, for the same tables, results are exactly the opposite. As a result, it was not possible to claim that this VECM(2) specification was capturing better the moving average component in our continuous time model.

For all the parametric designs and regardless of the variables being stocks of flows, we saw that our estimation results were broadly consistent with relatively small bias and clear improvements as the sample size increased, as a result, we could say that our estimation methodology was robust against changes in the true parametric specification of the system.

Finally, given these outcomes, we can say that even though Johansen's VECM(1) estimates are showing reasonably small bias in estimation in all the cases, when dynamics play an important role in the specification, they are substantially coming from a misspecified model and are contaminated by temporal aggregation bias.

Appendix A

Supplementary Results

For the computations of the covariance matrix W, we first notice that many elements of (2.11) can be simplified by grouping the constant matrices into single terms and solving their constant integrals as follows

$$
\begin{aligned}
W = \int_0^1 & \Big(I_n + G(e^{(s)M} - I_{n_1})B'\Big)\Sigma\Big(I_n + G(e^{(s)M} - I_{n_1})B'\Big)' ds, \\
= \quad & \Sigma(I_n - BG') + GB'\Sigma(BG' - I_n) \\
+ \quad & GM^{-1}(e^M - I_{n_1})B\Sigma(I_n - BG') \\
+ \quad & (I_n - BG')\Sigma B(M')^{-1}(e^{M'} - I_{n_1})G' \\
+ \quad & \Psi,
\end{aligned}
$$

where $\Psi = \int_0^1 Ge^{sM}B'\Sigma Be^{sM'}G'ds$.

As a result, we see that the computations of W are mainly reduced to calculate Ψ and the exponential matrix e^M, which, following Jewitt and McCrorie [2005] and Van Loan [1978], can be further simplified by computing instead the following exponential matrix

$$
\varpi = exp\begin{pmatrix} -M & B'\Sigma B \\ 0 & M' \end{pmatrix} = \begin{pmatrix} \varpi_{11} & \varpi_{12} \\ 0 & \varpi_{22} \end{pmatrix}
$$

where, $\Psi = G\big(\varpi'_{22}\varpi_{12}\big)G'$ and $e^M = \varpi'_{22}$. Hence, our problem is reduced to calculate either e^M or e^{ϖ}, and a number of methods exist for this purpose.

In our particular case, two different procedures are considered; first, we use the truncation of the infinite series representation of the exponential matrix e^M to a some sufficiently large integer N such that the elements of the difference $(e^M)_N - (e^M)_{N-1}$ are small enough to be neglected, therefore

$$e^M \approx e^M_N = I_{n_1} + \sum_{j=1}^{N} \frac{M^j}{j!},$$

second, we use a numerical calculation that is based on the approximation

$$e^A = (e^{2^{-s}A})^{2^s} \approx (r_m(2^{-s}A))^{2^s},$$

where $r_m(x)$ is the $[m/s]$ Pade approximation to e^x and the integers m and s are to be chosen in a prescribed way that aims to achieve full machine accuracy at minimal cost.[9]

At the end, the differences in the computations between the two procedures were small enough to be neglected[10], as a result, for computation efficiency, the Pade approximation approach was applied in the analysis.

In here, it is important to notice that alternative methods for this calculation do exist (see, for example, Moler and Van Loan [1978] and Ward [1977]), however, due to its efficiency (which makes it the most popular method for computing the matrix exponential), the Pade approximation method was considered. At the same time, the infinite series truncation method was considered not only for being the most straightforward, but also, as pointed out by Jewitt and McCrorie [2005], for being sufficiently precise in the calculations when a system like ours is being analysed, therefore, it serves for cross-validation and also as a threshold.

For Ω_{00}, Ω_{01}, Ω_0 and Ω_1 in (2.17), finally, a similar procedure was followed.

[9] See Al-Mohy and Higham [2009] for details.
[10] To be precise, the order of these differences was e^{-16}.

Appendix B

Proofs

Proof of Lemma 2.2.1.

Considering that

$$e^{tAB'} = I_n + \sum_{j=1}^{\infty} t^j \frac{(AB')^j}{j!} = I_n + \frac{AB'}{1!}t + \frac{(AB')(AB')}{2!}t^2 + \cdots .$$

We can write

$$\begin{aligned} (AB')^j &= \underbrace{(AB') \times (AB') \times \cdots \times (AB')}_{j \text{ times}}, \\ &= A \times \underbrace{(B'A) \times (B'A) \times \cdots \times (B'A)}_{j-1 \text{ times}} \times B'. \end{aligned}$$

Let $B'A = M$, then

$$(AB')^j = A \times M^{j-1} \times B'.$$

Hence

$$\begin{aligned} e^{tAB'} &= I_n + A\sum_{j=1}^{\infty} \frac{t^j}{j!} M^{j-1}B' = I_n + AM^{-1}\sum_{j=1}^{\infty} \frac{(tM)^j}{j!}B', \\ &= I_n + AM^{-1}(e^{tM} - I_{n_1})B'. \end{aligned}$$

which is used into (2.3) and the rest follows as in the Lemma. ■

Proof of Lemma 2.2.2.

First of all, following Bergstrom [1984], we can argue that (2.3) is the unique mean square solution to (2.1) due to the fact that the integral $\int_0^t e^{(t-s)AB'}\zeta(ds)$ exist.

It exist because for any matrix A, the series defining e^A is convergent, the elements of the matrix $e^{tAB'}\Sigma e^{tBA'}$ are convergent power series in t and they are integrable over any interval on the real line, as a result

$$\int_0^t e^{(t-s)AB'}\zeta e^{(t-s)BA'}ds < \infty,$$

were Σ is defined in the properties of the random vector $\zeta(dt)$. (See Bergstrom [1984] for details and a more comprehensive treatment of the proof).

Now, rewriting (2.5) as

$$\begin{aligned} y(t) &= \Big(I_n + G(e^{tM} - I_{n_1})B'\Big)y(0) + \int_0^{t-1}\Big(I_n + G(e^{(t-s)M} - I_{n_1})B'\Big)\zeta(ds) + \\ & \int_{t-1}^{t}\Big(I_n + G(e^{(t-s)M} - I_{n_1})B'\Big)\zeta(ds)), \\ y(t) &= \Big(I_n + G(e^{tM} - I_{n_1})B'\Big)y(t-1) + \int_{t-1}^{t}\Big(I_n + G(e^{(t-s)M} - I_{n_1})B'\Big)\zeta(ds), \end{aligned}$$

and then

$$y(t) - y(t-1) = GJB'y(t-1) + \int_{t-1}^{t}\Big(I_n + G(e^{(t-s)M} - I_{n_1})B'\Big)\zeta(ds),$$

$$\Delta y(t) = GJB'y(t-1) + \eta_t, \tag{B.1}$$

where all the matrices and η_t are defined in the Lemma. Finally, the exact discrete model when the sample is comprised entirely of stocks variables is given by

$$\Delta y_t^s = GJB'y_{t-1}^s + \eta_t,$$

Also, for the exact discrete model when the sample is comprised entirely of flows variables, if we integrate (B.1) over $(t-1,t)$

$$\Delta y_t^f = GJB'y_{t-1}^f + \int_{t-1}^{t}\int_{r-1}^{r}\Big(I_n + G(e^{(r-s)M} - I_{n_1})B'\Big)\zeta(ds)dr, \qquad \text{(B.2)}$$

$$\Delta y_t^f = GJB'y_{t-1}^f + v_t.$$

Finally, for the relationship between $y(0)$ and the observed vector y_1^f, integrating (2.5) over the interval $(0,1)$

$$\int_0^1 y(r)dr = \int_0^1 (I_n + G(e^{rM} - I_{n_1})B')y(0)dr + \int_0^1\int_0^r (I_n + G(e^{(r-s)M} - I_{n_1})B')\zeta(ds)dr,$$

$$y_1^f - y(0) = GEB'y(0) + v_1,$$

■

Proof of (2.15).

The integral of the function $f(r-s)$ in (2.15) can be obtained as follows:

$$\int_s^t f(r-s)dr = \int_s^t (I_n + G(e^{(r-s)M} - I_{n_1})B')dr,$$

since I_n, G and B are all constants and M^{-1} exist, the resulting expression is

$$\int_s^t f(r-s)dr = (t-s)(I_n - GB') + GM^{-1}[e^{(t-s)M} - I_{n_1}]B',$$

and the reaming integral

$$\int_{t-1}^{s+1} f(r-s)dr = (s-t+2)(I_n - GB') + GM^{-1}[e^{M} - e^{(t-s-1)M}]B'.$$

Combining these results gives the expression in (2.15). ■

Proof of Lemma 2.3.1.

Following the assumptions of $\zeta(ds)$ and the representation of v_t in equation

(2.15)

$$
\begin{aligned}
\mathrm{E}\left(v_1 v_1'\right) &= \Omega_{00} \\
&= \mathrm{E}\left[\left(\int_1^0 \Xi_1(1-s)\,\zeta(ds)\right)\left(\int_0^1 \Xi_1(1-s)\,\zeta(ds)\right)'\right], \\
&= \int_0^1 \Xi_1(1-s)\,\Sigma\Xi_1'(1-s)\,ds, \\
\mathrm{E}\left(v_2 v_1'\right) &= \Omega_{10} \\
&= \mathrm{E}\left[\left(\int_1^2 \Xi_1(2-s)\,\zeta(ds) + \int_0^1 \Xi_2(1-s)\,\zeta(ds)\right)\left(\int_0^1 \Xi_1(1-s)\,\zeta(ds)\right)'\right], \\
&= \int_0^1 \Xi_2(1-s)\,\Sigma\Xi_1'(1-s)\,ds, \\
\mathrm{E}\left(v_t v_t'\right) &= \Omega_0 \\
&= \mathrm{E}\left[\left(\int_{t-1}^t \Xi_1(t-s)\,\zeta(ds) + \int_{t-2}^{t-1} \Xi_2(t-s-1)\,\zeta(ds)\right)\right. \\
&\left.\left(\int_{t-1}^t \Xi_1(t-s)\,\zeta(ds) + \int_{t-2}^{t-1} \Xi_2(t-s-1)\,\zeta(ds)\right)'\right], \\
&= \int_{t-1}^t \Xi_1(t-s)\,\Sigma\Xi_1'(t-s)\,ds + \int_{t-2}^{t-1} \Xi_2(t-s-1)\,\Sigma\Xi_2'(t-s-1)\,ds, \\
&= \int_0^1 \Xi_1(s)\,\Sigma\Xi_1'(s)\,ds + \int_0^1 \Xi_2(s)\,\Sigma\Xi_2'(s)\,ds, \\
\mathrm{E}\left(v_t v_{t-1}'\right) &= \Omega_1 \\
&= \mathrm{E}\left[\left(\int_{t-1}^t \Xi_1(t-s)\,\zeta(ds) + \int_{t-2}^{t-1} \Xi_2(t-s-1)\,\zeta(ds)\right)\right. \\
&\left.\left(\int_{t-2}^{t-1} \Xi_1(t-s-1)\,\zeta(ds) + \int_{t-3}^{t-2} \Xi_2(t-s-2)\,\zeta(ds)\right)'\right], \\
&= \int_{t-2}^{t-1} \Xi_2(t-s-1)\,\Sigma\Xi_1'(t-s-1)\,ds, \\
&= \int_0^1 \Xi_2(s)\,\Sigma\Xi_1'(s)\,ds,
\end{aligned}
$$

where Ξ_1 and Ξ_2 are given in the Lemma and the last line of Ω_0 and Ω_1 follows from a simple change of variable in the integration. ■

Chapter 3

Estimation of Higher Order Cointegrated Systems

In this Chapter, as an extension of the analysis we provided in Chapter 2, we develop an estimation procedure for cointegrated systems in continuous time that not only allows for higher order specifications in the system but also incorporates deterministic components on it. At the same time, in order to provide as much generality as possible, we allow the system to be entirely comprised of stock or flow variables. For the analysis, we closely follow Bergstrom's tradition and, for each type of data, we derive an exact discrete time model and characterize entirely the properties of the discrete time disturbance vector. Also, with the use of an alternative exponential matrix factorization, we outline the autocovariance representations of the discrete time disturbances and obtain the Gaussian likelihood function.

As an application, we assess the performance of our estimation procedure over some simulated data and with the aim of measuring the costs, if there are any, of ignoring aggregation in the specification, we compare our results with the ones we would have obtained by imposing instead a discrete time specification (Johansen's specification) into the system.

In the results, in all cases, our estimation procedure shows superiority in estimation against Johansen's with smaller bias in the estimates and improvements as the sample size increases, however, they are concentrated about a mean with greater dispersion.

3.1 Introduction

In practice, econometricians have to work with time series that usually are not restricted to a linear specification nor to a zero mean or a mean and a trend either. The Gross Domestic Product (GDP) of United States, for example, is always conceived as a series with positive trend and a mean different than zero. Unemployment as well as inflation, at the same time, are perceived as series that can be modelled as autoregressive moving average processes (ARMA) of order higher than one.

In discrete time, as it is well known, all of those processes can be easily analysed and estimated with the usual econometric techniques (for example, Johansen's estimation procedure for cointegrated systems), however, as pointed out by Chambers and McCrorie [2007], if the model for estimation is naively specified in terms of the observation interval, it can be misspecified and its estimates can be contaminated by temporal aggregation bias. As a result, for estimation, it is needed to develop a model that has the property of holding exactly the process under consideration regardless of the frequency with which the data are observed. Such model is referred to as an exact discrete time model and is obtained by imposing, instead, a continuous time specification into the system.

The purpose of this chapter, therefore, is to generalize the analysis we provided in Chapter 2 and, following Bergstrom's tradition, derive an exact discrete time model, together with its estimation procedure, for higher order systems in continuous time that estimates the parameters of processes such as the ones mentioned above. In this chapter, particularly, we focus our attention on the non stationary cointegrated variables case and consider a higher order stochastic differential equation system that incorporates deterministic components (a constant and a linear trend) and is entirely comprised of stock or flow variables.

In here, it is important to mention that for the estimation of cointegrated systems in continuous time, alternative approaches exist. Harvey and Stock [1985, 1988], for example, proposed Kalman filter methods and Phillips [1991] proposed frequency domain regression techniques. In their analysis, considering a higher order system, Harvey and Stock handled irregularly spaced observations and accommodated moving average disturbances, however, they did not provide an exact discrete time model for the system under consideration. Phillips, alternatively, did accommodate moving average disturbances in the system but focused only on a simple first order system. Therefore, given our generalization, these methods should be viewed as alternatives to those outlined here and considered as references for future work.

Similarly than in Chapter 2, we assess the performance of our estimation procedure over some simulated data and with the aim of measuring the costs, if there are any, of ignoring aggregation in the specification, we compare our results with the ones we would have obtained by imposing instead a discrete time specification (Johansen's specification) into the system.

In the results, as expected, in all cases, our methodology shows superiority in estimation against Johansen's with smaller bias in the estimates, however, they are concentrated about a mean with greater dispersion. Additionally, the estimates of the long run equilibrium parameter of the model in all three parametric designs show the smallest bias and standard deviation.

Since we are considering two types of variables, we present our analysis by duplicate. Thus, this chapter is organized as follows: Section 3.2 frames the coitegrated continuous time system under consideration. Section 3.3 specifies the exact discrete time representation for stock and flow variables. Section 3.4 concentrates on the derivation of the covariance properties of the discrete time disturbance vector for the two representations and outlines the estimation procedure. Section 3.5 summarizes the simulation results and compares both the estimates of our exact discrete time representation and those obtained by applying Johansen's methodology and section 3.6 concludes. Supplementary results are given in Appendix C and all proofs in Appendix D.

3.2 The Model

The focus of this chapter is the continuous time random n-vector $y(t)$ that satisfies the stochastic differential equation system

$$d[D^{k-1}y(t)] = \Big[A_{k-1}D^{k-1}y(t) + \cdots + A_1 Dy(t) + A_0 y(t) \\ +a + bt\Big]dt + \zeta(dt), \qquad t > 0, \tag{3.1}$$

were k is a positive integer larger than or equal to 2, $A_0, \ldots, A_{k-1}$ are $n \times n$ matrices of unknown coefficients and a and b are $n \times 1$ vectors of unknown constants. For our purpose, we assume that $y(t)$ is $I(1)$ and is partitioned into two subvectors $y_1(t)$, $y_2(t)$ of dimensions $n_1 \times 1$ and $n_2 \times 1$, respectively $(n_1 + n_2 = n)$ with n_1 stationary linear cointegrating relationships of the form $y_1(t) - B_1 y_2(t)$, where B_1 is a $n_1 \times n_2$ matrix of cointegrating parameters. As a result, A_0 is singular and can be written as $A_0 = \text{Ç}B'$ where $B = (I_{n_1}, -B_1)'$ and $\text{Ç} = (\text{Ç}_1', \text{Ç}_2')'$ are reduced rank matrices of dimensions $n \times n_1$ that contain n_1 linearly independent vectors.

We also assume that the vector of random measures $\zeta(dt)$ is defined on all subsets of the line $0 < t < \infty$ having finite Lebesgue measure such that

- $\mathrm{E}[\zeta(dt)] = 0$,
- $\mathrm{E}[\zeta(dt)\zeta(dt)'] = \Sigma dt$ and
- $\mathrm{E}[\zeta(\Delta_1)\zeta(\Delta_2)'] = 0$ for disjoint intervals Δ_1 and Δ_2,

Considering this set up and in order to pursue our goal, we shall next find a solution to (3.1) and use its properties to derive an econometrically implementable model (also known as the exact discrete time model) that relates the unknown parameters of our system to the discrete time observations. For that, we first follow Chambers [1999] and rewrite (3.1) in a state space form as

$$dx(t) = [Ax(t) + a^* + b^* t]dt + \zeta^*(dt) \tag{3.2}$$

were the $nk \times nk$ matrix A and the $nk \times 1$ vectors $x(t)$, a^*, b^* and $\zeta^*(dt)$ are given by

$$x(t) \equiv \begin{bmatrix} y(t) \\ Dy(t) \\ \vdots \\ D^{k-1}y(t) \end{bmatrix}, \quad A = \begin{bmatrix} 0 & I_n & 0 & \cdots & 0 \\ 0 & 0 & I_n & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & I_n \\ A_0 & A_1 & A_2 & \cdots & A_{k-1} \end{bmatrix}, \quad a^* = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ a \end{bmatrix},$$

$$b^* = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ b \end{bmatrix}, \quad \zeta^*(dt) = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ \zeta(dt) \end{bmatrix}.$$

Then, given by Bergstrom [1984], the unique mean square solution to (3.2), initialized at $t = 0$, is presented as

$$x(t) = e^{tA}x(0) + \int_0^t e^{(t-s)A}(a^* + b^*s)ds + \int_0^t e^{(t-s)A}\zeta^*(ds), \qquad t > 0, \qquad (3.3)$$

where

$$x(0) \equiv [y(0)', Dy(0)', \ldots, D^{k-1}y(0)']' \quad \text{and} \quad e^{tA} = \sum_{j=0}^{\infty} \frac{(tA)^j}{j!}.$$

In here, it is important to notice that this solution can be considered as a generalization of that given by Bergstrom. It not only allows stochastic trends to interact with the system as a whole, but also specifies the particular treatment of $I(1)$ and cointegrated variables in the system by dealing explicitly with the singularity of A.[11] As a result, the use of such solution to the derivation of the exact discrete time model will require additional and slightly more complex mathematical derivations than the ones proposed by the author. Next section outlines such derivations more precisely.

In order to explore the effects of temporal aggregation in our specification and attempting to provide as much generality as possible in the applicability of our results, the vector $y(t)$ is allowed to be entirely comprised by stock or

[11] To see this, from the set up of the model, we know that A_0 is singular, as a result, $|A_0| = 0$ and hence A is singular.

flow variables. If we define stock and flow variables in continuous time as $y^s(t)$ and $y^f(t)$, respectively, then, the observed values at specific points in time, for stock variables, are $y_t^s = y^s(t)$ and the observed rate of flows, for flow variables, $y_t^f = \int_{t-1}^t y^f(r)dr$, where, in each case $t = 1, 2, \ldots, T$ and T denotes sample size.

For the derivation of the exact discrete time model, at the same time, our strategy will consists in solving out the unobservable components of both $x_t^s = x^s(t)$ and $x_t^f = \int_{t-1}^t x^f(r)dr$ from (3.3) by compacting that system into a single equation depending only on observable components as well as their lagged values. As a result, for easier exposure, we define, in advance, selection matrices that specify such separation and also, rewrite (3.3) in a vector autoregressive form (VAR) that automatically incorporates the lagged values into the system.

For the selection matrices, then, let's define S_1 and S_2 as the matrices that divide x_t^s (x_t^f) into two subvectors y_t^s (y_t^f) and w_t^s (w_t^f) that contain, respectively, its observable and unobservable components. These matrices and vectors are given by

$$\begin{aligned}
S_1 =& [I_n \quad 0_{n\times n^r}], & S_2 =& [0_{n^r\times n} \quad I_{n^r}],\\
y_t^s = S_1 x_t^s =& \begin{bmatrix} y^s(t) \end{bmatrix}, & y_t^f = S_1 x_t^f =& \begin{bmatrix} \int_{t-1}^t y^f(r)dr \end{bmatrix},\\
w_t^s = S_2 x_t^s =& \begin{bmatrix} Dy^s(t) \\ \vdots \\ D^{k-1}y^s(t) \end{bmatrix}, & w_t^f = S_2 x_t^f =& \begin{bmatrix} \int_{t-1}^t Dy^f(r)dr \\ \vdots \\ \int_{t-1}^t D^{k-1}y^f(r)dr \end{bmatrix},\\
x_t^s = x^s(t) =& \begin{bmatrix} y^s(t) \\ Dy^s(t) \\ \vdots \\ D^{k-1}y^s(t) \end{bmatrix}, & x_t^f = \int_{t-1}^t x^f(r)dr =& \begin{bmatrix} \int_{t-1}^t y^f(r)dr \\ \int_{t-1}^t Dy^f(r)dr \\ \vdots \\ \int_{t-1}^t D^{k-1}y^f(r)dr \end{bmatrix},
\end{aligned} \tag{3.4}$$

where $n^r = n(k-1)$, $0_{a\times b}$ is a null matrix of dimensions $a \times b$ and y_t^s (y_t^f) and w_t^s (w_t^f) are, respectively, the observable and unobervable components of x_t when the variables of interest are stocks (x_t^s) or flows (x_t^f).

For the VAR form, also, let's rewrite (3.3) as

$$\begin{aligned} x(t) = e^{tA}x(0) + \int_0^{t-1} e^{(t-s)A}(a^* + b^*s)ds + \int_{t-1}^{t} e^{(t-s)A}(a^* + b^*s)ds, \\ + \int_0^{t-1} e^{(t-s)A}\zeta^*(ds) + \int_{t-1}^{t} e^{(t-s)A}\zeta^*(ds), \end{aligned} \tag{3.5}$$

and let's lag (3.3) one period and get

$$x(t-1) = e^{(t-1)A}x(0) + \int_0^{t-1} e^{(t-1-s)A}(a^* + b^*s)ds + \int_0^{t-1} e^{(t-1-s)A}\zeta^*(ds). \tag{3.6}$$

Therefore, the VAR, by substituting (3.6) into (3.5) and rearranging terms, is given by

$$\begin{aligned} x(t) = e^{A}x(t-1) + \int_{t-1}^{t} e^{(t-s)A}(a^* + b^*s)ds + \int_{t-1}^{t} e^{(t-s)A}\zeta^*(ds), \\ x(t) = e^{A}x(t-1) + m_t + \varepsilon_t, \end{aligned} \tag{3.7}$$

were $m_t = \int_{t-1}^{t} e^{(t-s)A}(a^* + b^*s)ds$ and $\varepsilon_t = \int_{t-1}^{t} e^{(t-s)A}\zeta^*(ds)$.

3.3 The Discrete Time Representation

Once the VAR representation of the solutions to the system has been obtained and the observable and unobservable components of both x_t^s and x_t^f have been accurately divided, the derivation of the exact discrete time model can be finally outlined. Such derivation is described below and for simplicity, as two types of data are considered, two subsections are utilized.

3.3.1 Stock Variables

For stock variables, using the fact that $x_t^s = x^s(t)$, $y_t^s = S_1 x_t^s$ as well as $w_t^s = S_2 x_t^s$, the observable part of the system, by premultiplying (3.7) by S_1 and noting that $S_1'S_1 + S_2'S_2 = I$, is given by

$$y_t^s = S_1 e^{A}(S_1'S_1 + S_2'S_2)\big(x_{t-1}^s\big) + S_1 m_t + S_1 \varepsilon_t$$

which can be written as

$$y_t^s = C_{11}y_{t-1}^s + C_{12}w_{t-1}^s + m_{1t}^s + \varepsilon_{1t}^s \qquad (3.8)$$

where $C_{11} = S_1 e^A S_1'$, $C_{12} = S_1 e^A S_2'$, $m_{1t}^s = S_1 m_t$ and $\varepsilon_{1t}^s = S_1 \varepsilon_t$.

The unobservable part, by premultiplying now (3.7) by S_2 and following the same procedure, is given by

$$w_t^s = C_{21}y_{t-1}^s + C_{22}w_{t-1}^s + m_{2t}^s + \varepsilon_{2t}^s \qquad (3.9)$$

where $C_{21} = S_2 e^A S_1'$, $C_{22} = S_2 e^A S_2'$, $m_{2t}^s = S_2 m_t$ and $\varepsilon_{2t}^s = S_2 \varepsilon_t$.

Then, the desired exact discrete time model, as pointed out before, is the equation that solves the system of relationships summarized in (3.8) and (3.9) by eliminating out the unobservable components w_{t-1} of the system so that we finish with a single expression depending only on y_t^s and its lagged values. The precise form is given below in Lemma 3.3.1 and our strategy in the derivation follows closely the steps outlined in Chambers [1999] with some important differences related mainly to the inclusion of stochastic trends in the system as well as the particular characteristics of the exponential matrix e^{tA}.

It is important to notice, at the same time, that for such derivation, regardless of the observations being stocks or flows, a particular set of assumptions has to be considered; first we need to ensure invertibility in the system which is achieved by assuming that the coefficient matrix C_{22} of the unobservable elements of the system (w_{t-1}), and its lagged respective matrices, are non singular (Assumption 1 and 3 below) and secondly, we need to ensure that there are no linear dependencies between y_t and w_t which is achieved by assuming full row rank in the coefficient matrix, C_{12}, that is relating these two sub vectors (Assumption 2 below).

Assumption 1: The $n(k-1) \times n(k-1)$ matrix C_{22} is nonsingular.

Assumption 2: The $n \times n(k-1)$ matrix C_{12} is has full row rank n.

Assumption 3: The $n(k-1) \times n(k-1)^2$ matrix $[C_{22}^{-1}, \ldots, C_{22}^{-(k-1)}]$ has full row rank $n(k-1)$.

All together, these assumptions are usually considered when deriving exact discrete time representations of the observed variables in a system, they are closely related to the concepts of reconstructibility and detectability employed in optimal control theory for linear systems and overall, they set the minimal requirements for mathematically deriving our discrete time representation (see Chambers [1999] for an extensive discussion).

Lemma 3.3.1 (Exact Discrete Time Representation for Stock Variables).

Let $x(t)$ satisfy the continuous time cointegrated system defined in (3.2)*, then, under assumptions (1) - (3), the exact discrete time model under our observed vector y_t^s evolves according to the discrete time vector error correction model representation*

$$\begin{aligned}\Delta y_t^s &= \Pi_k(\theta) y_{t-1}^s + \Gamma_1(\theta)\Delta y_{t-1}^s + \cdots + \Gamma_{k-1}(\theta)\Delta y_{t-(k-1)}^s + g_t^s(\theta) + \eta_t^s, \\ t &= k+1, \cdots, T,\end{aligned} \tag{3.10}$$

where

$$\Pi_k = F_1 + \cdots + F_k - I, \qquad \Gamma_h = -\sum_{j=h+1}^{k} F_j, \quad h = 1, \cdots, k-1,$$

$$F_1 = C_{11} + C_{12} M N_1, \qquad F_j = C_{12} M N_j, \quad j = 2, \cdots, k$$

$$g_t^s = m_{1t}^s + C_{12} M \bar{m}_t^s, \qquad \eta_t^s = \varepsilon_{1t}^s + C_{12} M \bar{\varepsilon}_t^s,$$

$$M = \hat{M}^{-1}[-I_{n(k-1)} \quad M^*], \qquad \hat{M} = \begin{bmatrix} C_{12}C_{22}^{-1} \\ C_{12}C_{22}^{-2} \\ \vdots \\ C_{12}C_{22}^{-(k-1)} \end{bmatrix},$$

$$\bar{m}_t^s = \left[(m_{1,t-1}^s)', \cdots, (m_{1,t-(k-1)}^s)', (m_{2,t-1}^s)', \cdots, (m_{2,t-(k-1)}^s)' \right]',$$

$$\bar{\varepsilon}_t^s = \left[(\varepsilon_{1,t-1}^s)', \cdots, (\varepsilon_{1,t-(k-1)}^s)', (\varepsilon_{2,t-1}^s)', \cdots, (\varepsilon_{2,t-(k-1)}^s)' \right]'.$$

$$M^* = \begin{bmatrix} C_{12}C_{22}^{-1} & 0 & \cdots & 0 & 0 \\ C_{12}C_{22}^{-2} & C_{12}C_{22}^{-1} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ C_{12}C_{22}^{-(k-2)} & C_{12}C_{22}^{-(k-3)} & \cdots & C_{12}C_{22}^{-1} & 0 \\ C_{12}C_{22}^{-(k-1)} & C_{12}C_{22}^{-(k-2)} & \cdots & C_{12}C_{22}^{-2} & C_{12}C_{22}^{-1} \end{bmatrix},$$

$$N = \begin{bmatrix} -I_n & C_{11} & 0 & \dots & 0 & 0 \\ 0 & -I_n & C_{11} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots - I_n & C_{11} \\ 0 & C_{21} & 0 & \dots & 0 & 0 \\ 0 & 0 & C_{21} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & C_{21} \end{bmatrix} \equiv \begin{bmatrix} N_1 & N_2 & \cdots & N_k \end{bmatrix},$$

Proof. See Appendix D. ■

3.3.2 Flow Variables

Following a similar procedure than before, for flows, the system of equations under consideration is given by

$$\begin{aligned} y_t^f &= C_{11} y_{t-1}^f + C_{12} w_{t-1}^f + u_{1t}^f + v_{1t}^f, \\ w_t^f &= C_{21} y_{t-1}^f + C_{22} w_{t-1}^f + u_{2t}^f + v_{2t}^f, \end{aligned} \tag{3.11}$$

where C_{11}, C_{12}, C_{21} and C_{22} are as in (3.8) and (3.9) and equation (3.7) was integrated from $t-1$ to t so that

$$y_t^f = S_1 x_t^f = S_1 \int_{t-1}^{t} x^f(r)dr, \qquad w_t^f = S_2 x_t^f = \int_{t-1}^{t} x^f(r)dr,$$

$$u_{1t}^f = S_1 u_t^f = S_1 \int_{t-1}^{t} \int_{r-1}^{r} e^{(r-s)A}(a^* + b^* s)dsdr, \quad v_{1t}^f = S_1 v_t^f = S_1 \int_{t-1}^{t} \int_{r-1}^{r} e^{(r-s)A}\zeta^*(ds)dr,$$

$$u_{2t}^f = S_2 u_t^f = S_2 \int_{t-1}^{t} \int_{r-1}^{r} e^{(r-s)A}(a^* + b^* s)dsdr, \quad v_{2t}^f = S_2 v_t^f = S_2 \int_{t-1}^{t} \int_{r-1}^{r} e^{(r-s)A}\zeta^*(ds)dr,$$

Then, the exact discrete time model, as with stocks, is the equation that solves (3.11) and its precise form is given below in Lemma 3.3.2 (Note that as this derivation follows almost immediately from the previous Lemma, we also follow closely the steps outlined by Chambers [1999] as well as Assumptions 1-3 above).

Lemma 3.3.2 (Exact Discrete Time Representation for Flow Variables)**.**

Let $x(t)$ satisfy the continuous time cointegrated system defined in (3.2)*, then, under assumptions (1) - (3), the exact discrete time model under our* *observed vector y_t^f evolves according to the discrete time vector error correction model representation*

$$\begin{aligned}\Delta y_t^f &= \Pi_k(\theta) y_{t-1}^f + \Gamma_1(\theta)\Delta y_{t-1}^f + \cdots + \Gamma_{k-1}(\theta)\Delta y_{t-(k-1)}^f + g_t^f(\theta) + \eta_t^f, \\ t &= k+1, \cdots, T,\end{aligned} \tag{3.12}$$

where $\Pi_k, \Gamma_h, F_1, F_j, M, N$ *and* $\hat{M}$ *are specified as in Lemma 3.3.1 and*

$$g_t^f = u_{1t}^f + C_{12} M \bar{u}_t^f, \qquad \eta_t^f = v_{1t}^f + C_{12} M \bar{v}_t^f,$$

$$\bar{u}_t^f = \left[\ (u_{1,t-1}^f)', \cdots, (u_{1,t-(k-1)}^f)', (u_{2,t-1}^f)', \cdots, (u_{2,t-(k-1)}^f)'\ \right]',$$

$$\bar{v}_t^f = \left[\ (v_{1,t-1}^f)', \cdots, (v_{1,t-(k-1)}^f)', (v_{2,t-1}^f)', \cdots, (v_{2,t-(k-1)}^f)'\ \right]'.$$

Proof. See Appendix D. ■

As we can see from these lemmas, our exact discrete time models, regardless of the observations being stocks or flows, follow a common VECM representation, however, the specific forms of the disturbance vector as well as the stochastic trend component, differ in each scheme, as a result, in estimation, two different specifications are required.

At the same time, it is important to notice that these lemmas specify a solution to the system that holds only for $t = k+1, \cdots, T$ and if want to talk about appropriate methods of estimation, before deriving the specific properties of the discrete time disturbances, we need to derive an appropriate set of supplementary equations that relates $y_1, \cdots, y_k$ to the initial state vector $x(0)$. Again, as two different types of data are considered, two different specification are required and as before the proofs are heavily relying on the results of Chambers [1999].

For stocks, then, given that the initial observed value of the variables is directly specified in the system ($x_1^s = x^s(1)$), the set of supplementary equations is fixed to $x(0)$ and its representation, therefore, comes directly from (3.3). Next Lemma shows this result.

Lemma 3.3.3 (Supplementary Model for Stock Variables).

Under assumptions (1) - (3), the exact discrete time model under our observed vector y_t^s, that holds for $t = 1, \cdots, k$, evolves according to the discrete *time vector error correction model representation*

$$\begin{aligned} y_1^s &= G_1 x^s(0) + q_1^s + \eta_1^s, \\ \Delta y_t^s &= \Lambda_t(\theta) y_{t-1}^s + \Upsilon_1(\theta)\Delta y_{t-1}^s + \Upsilon_2(\theta)\Delta y_{t-2}^s + \\ &\quad \cdots + \Upsilon_{t-2}(\theta)\Delta y_2^s + G_t x^s(0) + q_t^s(\theta) + \eta_t^s, \\ t &= 2, \cdots, k, \end{aligned} \tag{3.13}$$

where

$$\begin{aligned} \Lambda_t &= J_1 + \cdots + J_{t-1} - I, & \Upsilon_h &= -\sum_{j=h+1}^{t-1} J_j, \quad \begin{matrix} h = 1, \cdots, t-2, \\ t = 3, \cdots, k, \end{matrix} \\ \eta_1^s &= S_1 \int_0^1 e^{(1-s)A} \zeta^*(ds), & q_1^s &= S_1 \int_0^1 e^{(1-s)A}(a^* + b^* s) ds, \\ \eta_2^s &= \varepsilon_{12}^s + C_{12} S_2 \int_0^1 e^{(1-s)A} \zeta^*(ds), & q_2^s &= m_{12}^s + C_{12} S_2 \int_0^1 e^{(1-s)A}(a^* + b^* s) ds, \\ J_1 &= C_{11}, & J_t &= C_{12} \sum_{j=1}^{t-1} C_{22}^{j-1} C_{21}, \quad t = 2, \cdots, k-1, \\ G_1 &= S_1 e^A, & G_t &= C_{12} C_{22}^{t-2} S_2 e^A \quad t = 2, \cdots, k, \end{aligned}$$

$$\begin{aligned} q_t^s &= m_{1t}^s + C_{12} \sum_{j=0}^{t-3} C_{22}^j m_{2,t-1-j}^s + C_{12} C_{22}^{t-2} S_2 \int_0^1 e^{(1-s)A}(a^* + b^* s) ds, \quad t = 3, \cdots, k, \\ \eta_t^s &= \varepsilon_{1t}^s + C_{12} \sum_{j=0}^{t-3} C_{22}^j \varepsilon_{2,t-1-j}^s + C_{12} C_{22}^{t-2} S_2 \int_0^1 e^{(1-s)A} \zeta^*(ds), \quad t = 3, \cdots, k, \end{aligned}$$

Proof. See Appendix D. ∎

For flows, on the contrary, the initial observed value of the variables is driven by $\int_0^1 x^f(r)dr$, as a result, the representation of the supplementary equations comes now from the integration of (3.3) from 0 to 1. Next Lemma shows this result.

Lemma 3.3.4 (Supplementary Model for Flow Variables)**.**

Under assumptions (1) - (3), the exact discrete time model under our observed vector y_t^f, that holds for $t = 1, \cdots, k$, evolves according to the discrete time vector error correction model representation

$$\begin{aligned} y_1^f &= Q_1 x^f(0) + q_1^f + \eta_1^f, \\ \Delta y_t^f &= \Lambda_t(\theta) y_{t-1}^f + \Upsilon_{h,1}(\theta)\Delta y_{t-1}^f + \Upsilon_{h,2}(\theta)\Delta y_{t-2}^f + \\ &\cdots + \Upsilon_{h,t-2}(\theta)\Delta y_2^f + Q_t x^f(0) + q_t^f(\theta) + \eta_t^f, \\ t &= 2, \cdots, k, \end{aligned} \tag{3.14}$$

where $\Lambda_t, \Upsilon_{h,t}, J_1$ and J_t are as in Lemma 3.3.3 and

$$\eta_1^f = S_1 \int_0^1 \int_0^r e^{(r-s)A} \zeta^*(ds) dr, \qquad q_1^f = S_1 \int_0^1 \int_0^r e^{(r-s)A}(a^* + b^* s) ds dr,$$

$$\eta_2^f = v_{12}^f + C_{12} S_2 \int_0^1 \int_0^r e^{(r-s)A} \zeta^*(ds) dr, \quad q_2^f = u_{12}^f + C_{12} S_2 \int_0^1 \int_0^r e^{(r-s)A}(a^* + b^* s) ds dr,$$

$$Q_1 = S_1 \int_0^1 e^{rA} dr, \qquad Q_t = C_{12} C_{22}^{t-2} S_2 \int_0^1 e^{rA} dr \quad t = 2, \cdots, k,$$

$$q_t^f = u_{1t}^f + C_{12} \sum_{j=0}^{t-3} C_{22}^j u_{2,t-1-j}^f + C_{12} C_{22}^{t-2} S_2 \int_0^1 \int_0^r e^{(r-s)A}(a^* + b^* s) ds dr, \qquad t = 3, \cdots, k,$$

$$\eta_t^f = v_{1t}^f + C_{12} \sum_{j=0}^{t-3} C_{22}^j v_{2,t-1-j}^f + C_{12} C_{22}^{t-2} S_2 \int_0^1 \int_0^r e^{(r-s)A} \zeta^*(ds) dr, \qquad t = 3, \cdots, k,$$

Proof. See Appendix D. ■

Once the required set of supplementary equations have been derived and incorporated into the exact discrete time models, the next step in the derivation of the Gaussian likelihood function is the complete characterization of the properties of the discrete time disturbance vector which, in our particular case, includes a specific treatment of the exponential matrix e^A as given in (C.2). The next section frames precisely these properties, outlines the derivation of the Gaussian Likelihood function and at the same time, specifies the estimation procedure.

3.4 The Properties of the Discrete Time Disturbances and the Estimation Procedure

Considering that we are working with two different types of data and for easier and more fluent presentation, we also divide this section into two subsections; the first focuses on stocks and the second on flows.

3.4.1 Discrete Time Disturbances for Stock Variables

For stock variables, as the general form of the discrete time disturbance vector η_t^s, given in Lemmas 3.3.1 and 3.3.3, is a function of the vectors ε_t^s and involves single integrals of the vector of random measures $\zeta^*(dt)$, its moving average representation as well as its autocovariance properties can be easily derived and depend only on those of $\zeta^*(dt)$.

The precise form of the autocovariances and their derivations are given below in Lemma 3.4.1 and as before, these results rely on those of Chambers [1999] but incorporate, at the same time, important differences in the computations caused mainly by the use of our alternative representation of e^{At} as given in (C.2).

Lemma 3.4.1 (Moving average representation of η_t^s).

Following the assumptions of $\zeta(dt)$, the moving average representation of the discrete time disturbance vectors $\eta_1^s, \eta_2^s, \cdots, \eta_T^s$ are given by

$$\begin{aligned} \eta_t^s &= \sum_{i=0}^{t-1} P_i^s \varepsilon_{t-i}, \qquad t = 1, \cdots, k, \\ \eta_t^s &= \sum_{i=0}^{k-1} R_i^s \varepsilon_{t-i}, \qquad t = k+1, \cdots, T, \end{aligned} \tag{3.15}$$

where

$$P_0^s = S_1, \quad P_i^s = C_{12}C_{22}^{i-1}S_2, \qquad i = 1, \cdots, t-1,$$
$$R_0^s = S_1, \quad R_i^s = C_{12}(M_{1i}S_1 + M_{i+1}S_2), \qquad i = 1, \cdots, k-1,$$
$$M = [M_1 \quad M_2 \quad \cdots \quad M_k],$$
$$M_1 = [M_{11} \quad M_{12} \quad \cdots \quad M_{1,k-1}],$$
$$\varepsilon_t = \int_{t-1}^{t} e^{(t-s)A}\zeta^*(ds) = \int_{t-1}^{t} \Big[I_{nk} + UH^{-1}(e^{H(t-s)} - I_{nk-n_2})V'\Big]\zeta^*(ds),$$
$$A = UV' \quad \textit{and} \quad H = V'U \qquad \textit{(See Appendix D for details).}$$

Also, if we define the $nT \times 1$ *vector* $\eta^s = [(\eta_1^s)', (\eta_2^s)', \cdots, (\eta_T^s)']'$, *then, its autocovariance representation is given by*

$$\Omega^s = E[\eta^s(\eta^s)'], \tag{3.16}$$

where

$$E[\eta_t^s(\eta_{t-j}^s)'] = \Omega_{t,t-j} = \sum_{i=j}^{t-1} P_i^s\Omega_\varepsilon(P_{i-j}^s)', \quad t = 1, \cdots, k, \qquad j = 0, \cdots, t-1,$$
$$E[\eta_t^s(\eta_{t-j}^s)'] = \Omega_{t,t-j} = \sum_{i=j}^{k-1} R_i^s\Omega_\varepsilon(P_{i-j}^s)', \quad t = k+1, \cdots, 2k-1, \quad j = t-k, \cdots, k-1,$$
$$E[\eta_t^s(\eta_{t-j}^s)'] = \Omega_j = \sum_{i=j}^{k-1} R_i^s\Omega_\varepsilon(R_{i-j}^s), \qquad t = k+1+j, \cdots, T, \quad j = 0, \cdots, k-1,$$
$$E[\eta_t^s(\eta_{t-j}^s)'] = 0, \qquad t = k+2, \cdots, T, \qquad j > k-1,$$

and

$$\Omega_\varepsilon = E[\varepsilon_t\varepsilon_t'] = \int_0^1 \big(e^{As}\big)\Sigma^*\big(e^{As}\big)'ds,$$
$$= \int_0^1 \big[I_{nk} + UH^{-1}(e^{Hs} - I_{nk-n_2})V'\big]\Sigma^*\big[I_{nk} + UH^{-1}(e^{Hs} - I_{nk-n_2})V'\big]'ds.$$

with

$E[\zeta^*(dt)] = 0,$

$E[\zeta^*(dt)\zeta^*(dt)'] = \Sigma^* dt,$ *and*

$E[\zeta^*(\Delta_1)\zeta^*(\Delta_2)'] = 0$ *for disjoint intervals* Δ_1 *and* Δ_2

also

$$\Sigma^* = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \Sigma \end{bmatrix}.$$

Proof. See Appendix D. ∎

As expected, Lemma 3.4.1 shows that the discrete time disturbances η_t^s follow a moving average process of order $k-1$ and, more importantly, it also shows that the autocovariances of η_t^s depend on the covariances Ω_ε.

The logarithm of the Gaussian likelihood function, then, with Ω^s as in the Lemma and assuming that η_t^s is normally distributed, can be written as

$$L(\theta, \Sigma) = -\frac{nT}{2}\ln(2\pi) - \frac{1}{2}\ln|\Omega^s| - \frac{1}{2}(\eta^s)'(\Omega^s)^{-1}\eta^s \tag{3.17}$$

where θ as well as Σ denote the unknown parameters of the system and our estimated parameters $\hat{\theta}$ and $\hat{\Sigma}$ are the values that maximize L.

As mentioned in Chapter 2, however, due to the sparse nature and size of Ω^s, obtaining those estimates by optimizing directly L is not very convenient in our context, as a result, we similarly follow Bergstrom [1985] and obtain, instead, the Gaussian estimates throughout an alternative recursive algorithm that not only avoids the computations of both the determinant and inverse of Ω^s, but also, exploits its sparse nature and reduces dramatically the number the computations and their complexity; the algorithm proceeds as follows

Let P be a real lower triangular matrix, with positive elements along the diagonal, such that

$$PP' = \Omega^s \tag{3.18}$$

$$P =$$

$$
\left(\begin{array}{ccccc|ccccc|c|cccc}
P_{11} & 0 & \cdots & \cdots & 0 & 0 & \cdots & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\
P_{21} & P_{22} & \cdots & \ddots & \vdots & 0 & 0 & \ddots & \ddots & 0 & \vdots & \vdots & \ddots & \vdots & 0 \\
\vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
P_{k-1,1} & \cdots & \cdots & P_{k-1,k-1} & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & 0 & \vdots & \vdots & \vdots & 0 \\
P_{k1} & P_{k2} & \cdots & \cdots & P_{kk} & 0 & \cdots & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\
\hline
0 & P_{k+1,2} & \cdots & \cdots & P_{k+1,k} & P_{k+1,k+1} & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\
\vdots & 0 & \ddots & \ddots & \vdots & P_{k+2,k+1} & P_{k+2,k+2} & 0 \cdots & & \cdots & 0 & 0 & \ddots & \ddots & 0 \\
\vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\
\vdots & \vdots & \vdots & \ddots & P_{2k-1,k} & \vdots & \ddots & \ddots & \ddots & 0 & \ddots & \ddots & \ddots & 0 & \vdots \\
0 & \cdots & \cdots & \cdots & 0 & P_{2k,k+1} & \ddots & \ddots & \cdots & P_{2k,2k} & 0 & \ddots & \ddots & 0 & 0 \\
\hline
0 & 0 & \cdots & 0 & 0 & 0 & P_{2k+1,k+2} & \ddots & \ddots & P_{2k+1,2k} & \ddots & 0 & \ddots & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & 0 & 0 & 0 & \ddots & \ddots & \ddots & \ddots & & 0 & \vdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \cdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\
\hline
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots & \cdots & 0 & P_{T-k+1,2k} & \cdots & P_{T-k+1,T-k+1} & \ddots & 0 & 0 \\
0 & \vdots & \vdots & \vdots & 0 & \vdots & \cdots & \cdots & 0 & \ddots & \ddots & \ddots & \ddots & P_{T-1,T-1} & 0 \\
0 & 0 & \cdots & 0 & 0 & 0 & \cdots & \cdots & 0 & 0 & 0 & P_{T,T-k+1} & \cdots & P_{T,T-1} & P_{TT}
\end{array}\right)
$$

where the matrices P_{ij} $(i = 1, 2\cdots, 2k-1, j = 1, 2, \cdots, k)$ are calculated, respectively, from the multiplication of the ith row and jth column of the matrices P and P' as follows

$$
\begin{aligned}
P_{11}P'_{11} &= \Omega_{11}, & P_{k+1,2} &= \Omega_{k+1,2}(P'_{22})^{-1}, \\
P_{21} &= \Omega_{21}(P'_{11})^{-1}, & P_{k+1,3} &= (\Omega_{k+1,3} - P_{k+1,2}P'_{32})(P'_{33})^{-1}, \\
&\vdots & &\vdots \\
P_{k1} &= \Omega_{k1}(P'_{11})^{-1}, & P_{k+1,k} &= (\Omega_{k+1,k} - P_{k+1,2}P'_{k2} - \cdots \\
& & &\quad - P_{k+1,k-1}P'_{k-1,k})(P'_{kk})^{-1}, \\
&\vdots & &\vdots \\
P_{kk}P'_{kk} &= (\Omega_{k,k} - P_{k,1}P'_{k,1} - \cdots, & P_{2k-1,k} &= \Omega_{2k-1,k}(P'_{kk})^{-1} \\
&\quad - P_{k,k-1}P'_{k,k-1}) & &
\end{aligned}
\tag{3.19}
$$

and the rest of the matrices P_{ij} $(i, j = k+1, k+2, \cdots, T)$ are computed recursively from

$$
\begin{aligned}
\Omega_0 =& \Omega_{i,i} = P_{i,i-k+1}P'_{i,i-k+1} + P_{i,i-k+2}P'_{i,i-k+2} + \cdots + P_{i,i}P'_{i,i}, \\
& \qquad (i = k+1, k+2, \cdots, T), \\
\Omega_1 =& \Omega_{i+1,i} = P_{i+1,i-k+2}P'_{i+1,i-k+2} + P_{i+1,i-k+3}P'_{i+1,i-k+3} + \cdots + P_{i+1,i}P'_{i,i}, \\
& \qquad (i = k+1, k+2, \cdots, T-1), \\
\vdots \qquad & \qquad \vdots \qquad\qquad \vdots \qquad\qquad \vdots \\
& \qquad \Omega_{k-1} = \Omega_{i+k-1,i} = P_{i+k-1,i}P'_{i,i}, \\
& \qquad (i = k+1, k+2, \cdots, T-k+1),
\end{aligned}
\tag{3.20}
$$

Also, let the $nT \times 1$ vector $\varepsilon = [\varepsilon_1', \cdots, \varepsilon_T']$ be defined as

$$P\varepsilon = \eta^s, \tag{3.21}$$

so that

- $\mathrm{E}[\varepsilon] = 0, \qquad \mathrm{E}[\varepsilon\varepsilon'] = I_{nT\times nT},$
- $\mathrm{E}[\varepsilon_t] = 0, \qquad \mathrm{E}[\varepsilon_t\varepsilon_t'] = I_{n\times n}\ , \qquad (t = 1, \ldots, T),$
- $\mathrm{E}[\varepsilon_t\varepsilon_s'] = 0\ , \qquad (s \neq t; s, t = 1, \ldots, T).$

Then, minus twice the logarithm of the likelihood function L (ignoring the constant) is given by

$$\mathcal{L}^s = \sum_{i=1}^{nT}(\varepsilon_i^2 + 2\log p_{ii}), \tag{3.22}$$

where p_{ii} is the ith diagonal element of P and the nT elements ε_i are computed in T vectors of size n using recursively the following procedure

$$\begin{aligned}
\varepsilon_1 &= \Big(\varepsilon_{11}, \ldots, \varepsilon_{1n}\Big)' = P_{11}^{-1}\eta_1^s, \\
\varepsilon_2 &= \Big(\varepsilon_{21}, \ldots, \varepsilon_{2n}\Big)' = P_{22}^{-1}(\eta_2^s - P_{21}\varepsilon_1), \\
&\vdots \qquad\qquad \vdots \qquad\qquad \vdots \\
\varepsilon_k &= \Big(\varepsilon_{k1}, \ldots, \varepsilon_{kn}\Big)' = P_{k,k}^{-1}(\eta_k^s - P_{k,k-1}\varepsilon_{k-1} - \cdots - P_{k,1}\varepsilon_1), \\
\varepsilon_{k+1} &= \Big(\varepsilon_{k+1,1}, \ldots, \varepsilon_{k+1,n}\Big)' = P_{k+1,k+1}^{-1}(\eta_{k+1}^s - P_{k+1,k}\varepsilon_k - \cdots - P_{k+1,2}\varepsilon_2), \\
&\vdots \qquad\qquad \vdots \qquad\qquad \vdots \\
\varepsilon_T &= \Big(\varepsilon_{T,1}, \ldots, \varepsilon_{T,n}\Big)' = P_{T,T}^{-1}(\eta_T^s - P_{T,T-1}\varepsilon_{T-1} - \cdots - P_{T,1}\varepsilon_{T-k+1}),
\end{aligned} \tag{3.23}$$

Therefore, the Gaussian estimates ($\hat{\theta}$ and $\hat{\Sigma}$) are obtained by optimizing (3.22) which, as mentioned in Chapter 2, due to fact of taking into account the sparse nature of Ω^s and also the convergence of the sequence of matrices $\bar{P}_t = (P_{tt}, P_{t,t-1}, \cdots, P_{t,t-k+1})$ $(t = k+1, k+2, \cdots, T)$, is computationally more efficient.

The optimization procedure follows closely the set of steps outlined in Chapter 2 and in our context they can be summarized as follows

(i) Compute A and Σ using the specified forms of the functions defining their elements.

(ii) Given the specific representations of the matrices U, H, V and Σ compute (see Appendix C for details) e^H together with Ω_{tj} $(t = 1, \cdots, 2k-1, j = 0, \cdots, k-1)$ and Ω_j $(j = 0, \cdots, k-1)$, from their specifications as given in Lemma 3.4.1.

(iii) Given these numerical representations, compute the Cholesky factorization of the matrix Ω^s (the matrix P) following recursively the steps on the set of equations following (3.18).

(iv) For the minimization of $\mathcal{L}^s$, with the data and allowing the model's parameters to vary, obtain a ε recursively as in (3.23).

(v) Set P and this new ε into (3.22) and calculate $\mathcal{L}^s$.

(vi) Repeat steps (iv) and (v) until a minimum is achieved and take those $\hat{\theta}$ and $\hat{\Sigma}$ as the elements that minimize $\mathcal{L}^s$.

3.4.2 Discrete Time Disturbances for Flow Variables

For flow variables, contrary than with stocks, as the general form of the discrete time disturbance vector η_t^f, given in Lemmas 3.3.2 and 3.3.4, is a function of v_t^f and involves now double integrals of ζ_t^*, its autocovariance properties not only depend on those of ζ_t^*, but also on those of v_t^f, as a result, before the actual derivations, additional simplifications, which involve reductions of the double integrals, are needed. Of course, similarly than with stocks, for the computations, we also take into account our particular representation of e^{At}.

The expressions of v_t^f, given by the equations following Lemmas 3.3.2 and 3.3.4, can be written as

$$\begin{aligned} v_1^f &= \int_0^1 \int_0^r e^{(r-s)A} \zeta^*(ds) dr, \\ v_t^f &= \int_{t-1}^t \int_{r-1}^r e^{(r-s)A} \zeta^*(ds) dr. \end{aligned} \tag{3.24}$$

Then, using Bergstrom [1997], McCrorie [2000] and Chambers [1999], the double integral in (3.24) can be divided into two single integrals using the following interchange of the orders of integration.

$$\begin{aligned} v_1^f &= \int_0^1 \Big[\int_s^1 e^{(r-s)A} dr\Big] \zeta^*(ds), \\ v_t^f &= \int_{t-1}^t \Big[\int_s^t e^{(r-s)A} dr\Big] \zeta^*(ds) + \int_{t-2}^{t-1} \Big[\int_{t-1}^{s+1} e^{(r-s)A} dr\Big] \zeta^*(ds). \end{aligned} \tag{3.25}$$

Finally, using our particular representation of e^{At} as given in (C.2), it is possible to analytically compute (see Appendix D for details) the integrals in square brackets of (3.25) and reduced them to single integrals as follows

$$\begin{aligned} v_1^f &= \int_0^1 \Big\{(1-s)(I_{nk} - UH^{-1}V') + UH^{-2}\big[e^{(1-s)H} - I_{nk-n_2}\big]V'\Big\}\zeta^*(ds), \\ &= \int_0^1 \phi(1-s)\zeta^*(ds), \\ v_t^f &= \int_{t-1}^t \Big\{(t-s)(I_{nk} - UH^{-1}V') + UH^{-2}\left[e^{(t-s)H} - I_{nk-n_2}\right]V'\Big\}\zeta^*(ds) \\ &+ \int_{t-2}^{t-1} \Big\{(s-t+2)(I_{nk} - UH^{-1}V') + UH^{-2}\left[e^{H} - e^{(t-s-1)H}\right]V'\Big\}\zeta^*(ds), \\ &= \int_{t-1}^t \phi(t-s)\zeta^*(ds) + \int_{t-2}^{t-1} \Big[\phi(1) - \phi(t-s-1)\Big]\zeta^*(ds), \end{aligned} \tag{3.26}$$

where $\phi(s) = (s)(I_{nk} - UH^{-1}V') + UH^{-2}\big[e^{sH} - I_{nk-n_2}\big]V'$.

Thus, as v_t^f involves now only single integrals, the autocovariances of η_t^f can be derived by using a generalization of the procedure outlined before in the stock variables case and depend only on those of $\zeta^*(dt)$. The precise form and their derivation are given below in Lemma 3.4.1.

Lemma 3.4.2 (Moving average representation of η_t^f)**.**

Following the assumptions of $\zeta(dt)$, the moving average representation of the discrete time disturbance vectors $\eta_1^f, \eta_2^f, \cdots, \eta_T^f$ are given by

$$\begin{aligned} \eta_t^f &= \sum_{i=0}^{t-1} P_i^f \xi_{t-i}, \qquad t = 1, \cdots, k, \\ \eta_t^f &= \sum_{i=0}^{k} R_i^f \xi_{t-i}, \qquad t = k+1, \cdots, T, \end{aligned} \tag{3.27}$$

where

$$
\begin{aligned}
P_0^f &= \Big[S_1 \;\; 0\Big], \qquad P_1^f = \Big[C_{12}S_2 \;\; S_1\Big], \\
P_i^f &= \Big[C_{12}C_{22}^{i-1}S_2 \;\; C_{12}C_{22}^{i-2}S_2\Big], \qquad\qquad i = 2, \cdots, t-1, \\
R_0^f &= \Big[S_1 \;\; 0\Big], \qquad R_1^f = \Big[C_{12}(M_{11}S_1 + M_2S_2) \;\; S_1\Big], \\
R_i^f &= \Big[C_{12}(M_{1i}S_1 + M_{i+1}S_2) \;\; C_{12}(M_{1,i-1}S_1 + M_iS_2)\Big], \qquad i = 2, \cdots, k-1, \\
R_k^f &= \Big[0 \;\; C_{12}(M_{1,k-1}S_1 + M_kS_2)\Big], \\
M &= [M_1 \;\; M_2 \;\; \cdots \;\; M_k], \\
M_1 &= [M_{11} \;\; M_{12} \;\; \cdots \;\; M_{1,k-1}], \\
\xi_t &= \Big[(v_{a,t}^f)' \;\; (v_{b,t}^f)'\Big]', \qquad t = 1, \cdots, T, \\
v_{a,t}^f &= \int_{t-1}^{t} \Xi_1(t-s)\zeta^*(ds), \qquad t = 1, \cdots, T, \qquad \Xi_1(s) = \phi(s), \\
v_{b,t}^f &= \int_{t-1}^{t} \Xi_2(t-s)\zeta^*(ds), \qquad t = 1, \cdots, T, \qquad \Xi_2(s) = \phi(1) - \phi(s).
\end{aligned}
$$

Also, if we define the $nT \times 1$ *vector* $\eta^f = [(\eta_1^f)', (\eta_2^f)', \cdots, (\eta_T^f)']'$*, then, its autocovariance representation is given by*

$$\Omega^f = E[\eta^f(\eta^f)'], \tag{3.28}$$

where

$$
\begin{aligned}
E[\eta_t^f(\eta_{t-j}^f)'] &= \Omega_{t,t-j} = \sum_{i=j}^{t-1} P_i^f\Omega_\xi(P_{i-j}^f)', \quad t = 1, \cdots, k, \qquad j = 0, \cdots, t-1, \\
E[\eta_t^f(\eta_{t-j}^f)'] &= \Omega_{t,t-j} = \sum_{i=j}^{k} R_i^f\Omega_\xi(P_{i-j}^f)', \quad t = k+1, \cdots, 2k, \qquad j = t-k, \cdots, k, \\
E[\eta_t^f(\eta_{t-j}^f)'] &= \Omega_j = \sum_{i=j}^{k} R_i^f\Omega_\xi(R_{i-j}^f)', \qquad t = k+1+j, \cdots, T, \quad j = 0, \cdots, k, \\
E[\eta_t^f(\eta_{t-j}^f)'] &= 0, \qquad t = k+2, \cdots, T, \qquad j > k,
\end{aligned}
$$

and

$$\Omega_\xi = E[\xi_t\xi_t'] = \begin{bmatrix} \int_0^1 \Xi_1(s)\Sigma^*\Xi_1(s)'ds & \int_0^1 \Xi_1(s)\Sigma^*\Xi_2(s)'ds \\ \int_0^1 \Xi_2(s)\Sigma^*\Xi_1(s)'ds & \int_0^1 \Xi_2(s)\Sigma^*\Xi_2(s)'ds \end{bmatrix}.$$

Proof. See Appendix D. ■

Differently than with stocks, Lemma 3.4.2 shows that the discrete time disturbances η_t^f follow now a moving average process of order k and, more importantly, it also shows that the autocovariances of η_t^f depend on the covariance matrix Ω_ξ of ξ_t.

For estimation, with Ω^f as in the Lemma and assuming that η_t^f is normally distributed, the logarithm of the Gaussian likelihood function can be written as

$$L(\theta, \Sigma) = -\frac{nT}{2}\ln(2\pi) - \frac{1}{2}\ln(\Omega^f) - \frac{1}{2}(\eta^f)'(\Omega^f)^{-1}\eta^f \tag{3.29}$$

where θ as well as Σ are specified as before.

In here, as mentioned before, getting our Gaussian estimates through the direct optimization of L is not convenient and similarly, an alternative procedure is required. This procedure is a mirror image of the one outlined above with the difference of considering instead a moving average representation of order k in the discrete time disturbances, as a result, the computationally efficient Cholesky factorization, matrix P, of Ω^f is given by

$$P = \left(\begin{array}{ccccc:ccccc:c:cccc}
P_{11} & 0 & \cdots & \cdots & 0 & 0 & \cdots & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\
P_{21} & P_{22} & \cdots & \ddots & \vdots & 0 & 0 & \ddots & \ddots & 0 & \vdots & \vdots & \ddots & \vdots & 0 \\
\vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
P_{k-1,1} & \cdots & \cdots & P_{k-1,k-1} & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & 0 & \vdots & \vdots & \vdots & 0 \\
P_{k1} & P_{k2} & \cdots & \cdots & P_{kk} & 0 & \cdots & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\
\hdashline
P_{k+1,1} & P_{k+1,2} & \cdots & \cdots & P_{k+1,k} & P_{k+1,k+1} & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\
0 & P_{k+2,2} & \ddots & \ddots & \vdots & P_{k+2,k+1} & P_{k+2,k+2} & 0 & \cdots & \cdots & 0 & 0 & \ddots & \ddots & 0 \\
\vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\
\vdots & \vdots & \vdots & \ddots & P_{2k-1,k} & \vdots & \ddots & \ddots & \ddots & 0 & \ddots & \ddots & \ddots & 0 & \vdots \\
0 & \cdots & \cdots & 0 & P_{2k,k} & P_{2k,k+1} & \ddots & \ddots & \cdots & P_{2k,2k} & 0 & \ddots & \ddots & 0 & 0 \\
\hdashline
0 & 0 & \cdots & 0 & 0 & P_{2k+1,k+1} & P_{2k+1,k+2} & \ddots & \ddots & P_{2k+1,2k} & \ddots & 0 & \ddots & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & 0 & 0 & P_{2k+2,k+2} & \ddots & \ddots & \ddots & \ddots & \ddots & 0 & \vdots & 0 \\
\vdots & \vdots & \ddots & \vdots & 0 & 0 & 0 & \ddots & \ddots & \ddots & \ddots & & 0 & \vdots & 0 \\
\hdashline
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \cdots & \ddots & & P_{T-k,2k} & \ddots & P_{T-k,T-k} & \ddots & \ddots & \vdots \\
0 & \vdots & \vdots & \vdots & 0 & \vdots & \cdots & \cdots & 0 & \ddots & \ddots & \ddots & \ddots & P_{T-1,T-1} & 0 \\
0 & 0 & \cdots & 0 & 0 & 0 & \cdots & \cdots & 0 & 0 & 0 & P_{T,T-k} & \cdots & P_{T,T-1} & P_{TT}
\end{array}\right)$$

where the matrices P_{ij} $(i, j = 1, 2, \cdots, T)$ are calculated recursively using similar systems as outlined in (3.19) and (3.20).

At the end, the Gaussian estimates ($\hat{\theta}$ and $\hat{\Sigma}$) are obtained by optimizing

$$\mathcal{L}^f = \sum_{i=1}^{nT}(\varepsilon_i^2 + 2\log p_{ii}), \tag{3.30}$$

where p_{ii} as well as the nT elements of ε_i are computed similarly than those in equation (3.22).

This likelihood function, for the reasons outlined above, is also computationally more efficient and its optimization procedure, as expected, is similarly summarized in the set of steps described in the optimization of (3.22).

3.5 Simulation Evidence

In this section, as an application of all the results outlined before, through the use of some simulations, we evaluate the performance of our estimation procedure for cointegrated systems in continuous time. On top of that, with the aim of measuring the costs, if there are any, of ignoring aggregation in the specification, we compare our estimates with the ones we would have obtained if Johansen's estimation procedure had been applied instead.

For this exercise, we define the system under consideration as a simplification of (3.1) in which there is only one cointegrating relationship, $(1, -b_1)$, contained in the matrix B and only two speed of adjustment parameters, g_1 and g_2, contained in the matrix Ģ. Also, we set $n = 2$ so that $n_1 = n_2 = 1$ and we fix $k = 2$, $y(0) = 0$, $a = b = 0$, the elementes of the matrix A_1 as $A_{1,12} = A_{1,21} = 0$, $A_{1,11} = x$ and $A_{1,22} = w$, therefore, using lemmas 3.3.1 and 3.3.2, the exact discrete time VECMs for estimation are given by

$$\begin{aligned} \Delta y_t^s &= \Pi y_{t-1}^s + \Gamma_1 \Delta y_{t-1}^s + \eta_t^s, \qquad t = 3, \cdots, T, \\ \Delta y_t^f &= \Pi y_{t-1}^f + \Gamma_1 \Delta y_{t-1}^f + \eta_t^f, \qquad t = 3, \cdots, T, \end{aligned} \tag{3.31}$$

where $\Pi = F_1 + F_2 - I = KB'$ and $\Gamma_1 = -F_2$, with

$$\begin{aligned} F_1 &= S_1 e^A S_1' + (S_1 e^A S_2')(S_2 e^A S_2')(S_1 e^A S_2')^{-1}, & S_1 &= \begin{pmatrix} I_2 & 0_2 \end{pmatrix}, \\ F_2 &= -(S_1 e^A S_2')(S_2 e^A S_2')(S_1 e^A S_2')^{-1}(S_1 e^A S_1') + (S_1 e^A S_2')(S_2 e^A S_1'), & S_2 &= \begin{pmatrix} 0_2 & I_2 \end{pmatrix}, \\ e^A &= \left(I_4 + U H^{-1}(e^H - I_3) V' \right), \\ K &= (k_1, k_2)', & B' &= (1, -b_1), \end{aligned}$$

$$A = UV' = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \text{g}_1 & -\text{g}_1 b_1 & x & 0 \\ \text{g}_2 & -\text{g}_2 b_1 & 0 & w \end{pmatrix} \qquad H = V'U = \begin{pmatrix} 0 & 1 & -b_1 \\ \text{g}_1 & x & 0 \\ \text{g}_2 & 0 & w \end{pmatrix}$$

$$V' = \begin{pmatrix} 1 & -b_1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad U = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ g_1 & x & 0 \\ g_2 & 0 & w \end{pmatrix}$$

As $y(0) = 0$, the respective supplementary equations, at the same time, are completely driven by the discrete time disturbance vectors η_t^s and η_t^f $(t = 1, 2)$, respectively, hence, using lemmas 3.3.3 and 3.3.4 they are summarized as follows

$$\begin{aligned} y_1^s &= \eta_1^s, \\ \Delta y_2^s &= \Lambda_2(\theta) y_1^s + \eta_2^s, \\ y_1^f &= \eta_1^f, \\ \Delta y_2^f &= \Lambda_2(\theta) y_1^f + \eta_2^f, \end{aligned} \tag{3.32}$$

where $\Lambda_2 = J_1 - I$ and $J_1 = C_{11}$.

For the discrete time specification, we also define the system under consideration as the following VECM

$$\Delta y_t = \gamma \lambda' y_{t-1} + \Gamma_1^d \Delta y_{t-1} + \eta_t \tag{3.33}$$

where γ contains the two speed of adjustment parameters (γ_1 and γ_2) of the system, λ' the cointegrating relationship $(1, -\lambda_1)$ and η_t is assumed to be *iid*.

At the end, for the comparison, if we want to measure how accurate the discrete time estimates are, in terms of our continuous time specification, all we need to compute are the *implied* estimated parameters, which, equating (3.31) and (3.33), are given by

$$\begin{aligned} \gamma_1 &= k_1, & \gamma_2 &= k_2, \\ \lambda_1 &= b_1, & \Gamma_1^d &= \Gamma_1. \end{aligned} \tag{3.34}$$

In our simulated data, we specify (3.31) and (3.32) as the data generating process (DGP) and consider two different parametric designs. For simplicity and due to its superconsistency (see equation 2.34), in all designs, we normalize the cointegrating parameter to be 1 so that the cointegrating relationship is given by $y_{1t} - y_{2t}$ and as a result, in this application, we stress mainly the implications of dynamics over the performance of estimation in the system. The exact representations are as follows

$$\begin{aligned} \text{Design 1:} \quad \theta^0 &= [g_1, g_2, b_1, \Gamma_{11}, \Gamma_{12}, \Gamma_{21}, \Gamma_{22}] \\ &= [1, 2, 1, -3, 0, 0, -3], \\ \text{Design 2:} \quad \theta^0 &= [g_1, g_2, b_1, \Gamma_{11}, \Gamma_{12}, \Gamma_{21}, \Gamma_{22}] \\ &= [-1, -2, 1, -2, 0, 0, -8], \end{aligned} \tag{3.35}$$

$$\Gamma_1 = \begin{pmatrix} \Gamma_{1,11} & \Gamma_{1,12} \\ \Gamma_{1,21} & \Gamma_{1,22} \end{pmatrix}$$

Note that in order to provide as much generality as possible in the description of the results, the chosen values in the parametric designs were obtained as representative elements of the feasible set of values in the parameter space, then, the systems that are generated by them are stable and feasible, as a result, both Johansen's estimation procedures and ours can be applied. Also, for efficiency in optimization, we are assuming that $\Gamma_{11} = c\Gamma_{21}$ (c constant) so that only one parameter of this matrix is needed.

At the same time, the covariance matrix Σ is taken as follows

$$\Sigma = \begin{bmatrix} \sigma_1 & \sigma_3 \\ \sigma_3 & \sigma_2 \end{bmatrix} = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix},$$

and results are reported for $\rho = -0.5$ and $\rho = 0.5$ so that there is positive and negative correlation in the system. Note that in estimation, Σ is ensured to be positive definite by computing, instead, estimates of the the lower triangular matrix R, such that, $\Sigma = RR'$; these matrices are related as follows

$$r_1 = 1, \quad r_2 = \rho, \quad r_3 = \sqrt{1 - \rho^2}.$$

Considering all those specifications, the estimates of our cointegrated continuous time system ($\hat{\theta}$ and $\hat{\Sigma}$), then, are obtained through the application of the methodologies described in subsections 3.4.1 and 3.4.2 to the simulated data.

For the comparison, the *implied* true parametric designs, which are calculated using (3.35) on (3.34), are given by

$$\begin{aligned}
\text{Design 1:} \quad \theta_j^0 &= [\gamma_1, \gamma_2, \lambda_1, \Gamma_{11}^d, \Gamma_{12}^d, \Gamma_{21}^d, \Gamma_{22}^d] \\
&= [0.294, 0.588, 1.0, 0.05, 0, 0, 0.05], \\
\text{Design 2:} \quad \theta_j^0 &= [\gamma_1, \gamma_2, \lambda_1, \Gamma_{11}^d, \Gamma_{12}^d, \Gamma_{21}^d, \Gamma_{22}^d] \\
&= [-0.436, -0.251, 1.0, 0.141, -0.103, -0.068, -0.002],
\end{aligned} \tag{3.36}$$

$$\Gamma_1^d = \begin{pmatrix} \Gamma_{1,11}^d & \Gamma_{1,12}^d \\ \Gamma_{1,21}^d & \Gamma_{1,22}^d \end{pmatrix}$$

and the *implied* estimated parameter vector $\hat{\theta}_j$ is obtained by applying Johansen's methodology to the same simulated data as if it was generated by the VECM representation of order 1 given in (3.33).

Finally, performance of the method is analysed by measuring accuracy in estimation, which for our purposes, is defined as closeness between the estimated parameters $\hat{\theta}$ as well as the $\hat{\theta}_j$ and the true values in (3.35) and (3.36). The procedure is described below, and as we have two different types of data, the first subsection focuses on the case when the variables of interest are stocks and the second when they are flows.

3.5.1 VECM Simulations With Stocks

For stocks, the data generating process, which is used to generate ten thousand simulations of 50, 100 and 200 sample sizes, follows a VECM representation that can be written as

$$\begin{aligned}
y_1^s &= \eta_1^s, \\
\Delta y_2^s &= \Lambda_2(\theta) y_1^s + \eta_2^s, \\
\begin{pmatrix} \Delta y_{1t}^s \\ \Delta y_{2t}^s \end{pmatrix} &= \begin{pmatrix} k_1 \\ k_2 \end{pmatrix} (y_{1,t-1}^s - b_1 y_{2,t-1}^s) + \Gamma_1 \Delta y_{t-1}^s + \eta_t^s,
\end{aligned} \tag{3.37}$$

where η_t^s follows a moving average representation or order 1 as given in Lemma 3.4.1 and η^s, in that same Lemma, is assumed to be $N(0, \Omega^s)$.

The true parametric designs, are given as before and for estimation, we apply our methodology as described in subsection 3.4.1 to the generated data as well as Johansen's to it as if it was generated by the VECM(1) specification given in (3.33). At the same time, as there is a moving average component in our specification that is being ignored by Johansen's, we also apply Johansen's methodology to our generated data as if it was specified through a VECM representation of order 2 given by

$$\Delta y_t = \gamma\lambda' y_{t-1} + \Gamma_1^d \Delta y_{t-1} + \Gamma_2^d \Delta y_{t-2} + s_t \tag{3.38}$$

where s_t is assumer to be iid, γ, as before, is the matrix that includes the two speed of adjustment parameters (γ_1 and γ_2), λ' contains the unique cointegrating relationship $(1, -\lambda_1)$, Γ_1^d is the matrix of coefficients that relates Δy_t with its lagged value and Γ_2^d does it with the second order lagged value.

For comparison, finally, we consider the relevant parameter estimates of this system and measure how close they are with the respective true parameter values in (3.36).

Results have been grouped and appear in tables 3.1, 3.2, 3.3 and 3.4 so that each design is explained in two tables with the first focusing in the results when the correlation is positive and the second when it is negative. Each table is showing the true value, the bias and the standard error of each of the parameters for the 3 sample sizes in the exercise. Also, for better understanding, each table is divided in three sections; the upper part displays the results obtained through the application of our methodology to the simulated data, the middle part displays them when instead Johansen's methodology is applied to the same simulated data considering a VECM(1) specification and finally, the lower part presents the results when Johansen's is applied to the same simulated data but now considering a VECM(2) specification. It is important to notice, for the lower part of these tables, that the implied true values for the matrix of coefficients Γ_2^d are not known, as a result, we cannot report the bias, instead, we are reporting only the mean value and the standard error of the estimates of these particular parameters.

As we can see from the tables, in almost all the cases and for the two parametric designs, our methodology shows superiority in estimation against

Johansen's with smaller bias in the estimates and improvements as the sample size increases, however, they are concentrated about a mean with greater dispersion. Additionally, as expected, the estimates of the long run equilibrium parameter of the model in all the parametric designs show the smallest bias and standard deviation. For the change in the correlation from positive to negative and paying attention to the estimates we obtained by applying Johansen's VECM(1) methodology and ours, we see an important reduction in their standard errors.

Considering Johansen's VECM(1) specification (the middle part of the tables) and focusing only on the dynamics of the system (γ_1 and γ_2), we see a persistent bias in estimation with almost no improvement as the sample size increases, consequently, they clearly reflect the cost of ignoring aggregation in the specification.

For Johansen's VECM(2) specification (the lower part of the tables), additionally, we see that the inclusion of an additional lag into the specification not only does not improves the estimates, but in some cases, it makes them worse. As a result, we cannot claim that the inclusion of an additional lag captures better the moving average component in our continuous time model, however, it is important to notice that the true reported values in the tables may not be precisely measured, hence, the previous claim has to be taken with care.

Table 3.1 Design 1 estimates for stock variables (positive correlation)

		Sample size			Sample size		
Continuous time estimates		50	100	200	50	100	200
		Bias			Standard Error		
Parameter	True value						
g_1	1	0.00990	0.00913	0.00667	0.55593	0.35968	0.24651
g_2	2	-0.03531	-0.02164	-0.01428	0.58699	0.36453	0.24513
b_1	1	0.00134	0.00045	0.00013	0.05568	0.02072	0.00953
$\Gamma_{1,11}$	-3	0.05091	0.03589	0.02802	0.49238	0.32557	0.22336
ρ	0.5	-0.00697	-0.00269	-0.00046	0.14711	0.09981	0.06710
σ_2	1	-0.04719	-0.02360	-0.01516	0.04644	0.02113	0.00976
Johansen's VECM(1) estimates							
γ_1	0.294	-0.09554	-0.07525	-0.07249	0.25376	0.16725	0.11475
γ_2	0.588	0.29791	0.28615	0.27731	0.19432	0.12748	0.08728
λ_1	1	-0.01106	-0.00502	-0.00228	0.05469	0.02107	0.00975
$\Gamma^d_{1,12}$	0.05	-0.34742	-0.36342	-0.36600	0.23105	0.15459	0.10743
$\Gamma^d_{1,12}$	0	0.02294	0.03182	0.03249	0.20247	0.13430	0.09515
$\Gamma^d_{1,21}$	0.05	-0.33385	-0.32114	-0.31404	0.19305	0.13131	0.09038
$\Gamma^d_{1,22}$	0	-0.04375	-0.05873	-0.06758	0.17043	0.11217	0.07769
Johansen's VECM(2) estimates							
γ_1	0.294	0.02468	0.03437	0.03110	0.33084	0.21305	0.14441
γ_2	0.588	0.42133	0.40015	0.38485	0.25757	0.16640	0.11391
λ_1	1	-0.00566	-0.00262	-0.00113	0.09056	0.02133	0.00968
$\Gamma^d_{1,11}$	0.05	-0.47263	-0.48005	-0.47686	0.32398	0.21243	0.14548
$\Gamma^d_{1,12}$	0	0.10858	0.11195	0.10854	0.26861	0.17394	0.12202
$\Gamma^d_{1,21}$	0.05	-0.46222	-0.44059	-0.42651	0.25881	0.17227	0.11863
$\Gamma^d_{1,22}$	0	0.03158	0.01031	-0.00335	0.22352	0.14592	0.10059
Parameter		Mean of the estimated parameter					
$\Gamma^d_{2,11}$		-0.14247	-0.13329	-0.12726	0.24839	0.16552	0.11254
$\Gamma^d_{2,12}$		0.01772	0.02014	0.02085	0.20847	0.13785	0.09418
$\Gamma^d_{2,21}$		-0.14282	-0.12954	-0.12007	0.20637	0.13733	0.09490
$\Gamma^d_{2,22}$		-0.02278	-0.02757	-0.03014	0.17583	0.11422	0.07839

Table 3.2 Design 1 estimates for stock variables (negative correlation)

		Sample size			Sample size		
Continuous time estimates		50	100	200	50	100	200
		Bias			Standard Error		
Parameter	True value						
g_1	1	-0.01583	-0.01207	-0.00781	0.39518	0.27218	0.19660
g_2	2	-0.05923	-0.04106	-0.02956	0.41201	0.26678	0.18195
b_1	1	0.00142	0.00042	0.00016	0.02105	0.00904	0.00416
$\Gamma_{1,11}$	-3	0.08201	0.05448	0.03917	0.56018	0.37631	0.26120
ρ	-0.5	-0.01769	-0.00704	0.00126	0.15234	0.10473	0.07411
σ_2	1	-0.05435	-0.03242	-0.02517	0.05620	0.02645	0.01311
Johansen's VECM(1) estimates							
γ_1	0.294	-0.24455	-0.22836	-0.22221	0.22493	0.15273	0.10566
γ_2	0.588	0.31927	0.31252	0.30991	0.10956	0.07183	0.05039
λ_1	1	-0.00501	-0.00228	-0.00109	0.02212	0.00966	0.00438
$\Gamma^d_{1,11}$	0.05	-0.11417	-0.13214	-0.13729	0.29016	0.19668	0.13663
$\Gamma^d_{1,12}$	0	0.08863	0.09585	0.09947	0.14749	0.10017	0.07117
$\Gamma^d_{1,21}$	0.05	-0.38350	-0.37345	-0.37025	0.14498	0.09688	0.06852
$\Gamma^d_{1,22}$	0	-0.07698	-0.08368	-0.08655	0.07556	0.05094	0.03464
Johansen's VECM(2) estimates							
γ_1	0.294	-0.30662	-0.29581	-0.29829	0.36073	0.23803	0.16195
γ_2	0.588	0.47925	0.46642	0.45921	0.17355	0.11441	0.07908
λ_1	1	-0.00269	-0.00125	-0.00064	0.02288	0.00954	0.00428
$\Gamma^d_{1,11}$	0.05	-0.05387	-0.06636	-0.06300	0.41839	0.27811	0.18873
$\Gamma^d_{1,12}$	0	0.05692	0.06222	0.06071	0.22978	0.15416	0.10892
$\Gamma^d_{1,21}$	0.05	-0.54799	-0.53271	-0.52484	0.20275	0.13526	0.09400
$\Gamma^d_{1,22}$	0	0.02159	0.01134	0.00536	0.12151	0.08195	0.05568
Parameter		Mean of the estimated parameter					
$\Gamma^d_{2,11}$		0.06502	0.06959	0.07838	0.31212	0.21138	0.14460
$\Gamma^d_{2,12}$		0.06253	0.06427	0.06749	0.15195	0.10146	0.07002
$\Gamma^d_{2,21}$		-0.18261	-0.17487	-0.16844	0.15581	0.10535	0.07260
$\Gamma^d_{2,22}$		-0.04874	-0.04487	-0.04399	0.07919	0.05242	0.03636

Table 3.3 Design 2 estimates for stock variables (positive correlation)

		Sample size			Sample size		
Continuous time estimates		50	100	200	50	100	200
		Bias			Standard Error		
Parameter	True value						
g_1	-1	-0.03205	-0.01116	-0.00513	0.23602	0.14776	0.09864
g_2	-2	0.04407	0.03276	0.02806	0.45600	0.28684	0.18938
b_1	1	0.00274	0.00006	-0.00001	0.06866	0.02758	0.01266
$\Gamma_{1,11}$	-2	0.07511	0.04392	0.03032	0.32241	0.21507	0.14652
ρ	0.5	-0.00578	-0.00325	-0.00205	0.14827	0.10205	0.06947
σ_2	1	-0.08505	-0.04356	-0.02674	0.04382	0.02084	0.00987
Johansen's VECM(1) estimates							
γ_1	-0.437	-0.32491	-0.29901	-0.29062	0.16718	0.10921	0.07645
γ_2	-0.251	-0.07503	-0.07219	-0.07134	0.06100	0.03862	0.02650
λ_1	1	0.02093	0.00986	0.00441	0.07309	0.02970	0.01351
$\Gamma^d_{1,11}$	0.141	-0.08824	-0.10770	-0.11464	0.16829	0.11233	0.07955
$\Gamma^d_{1,12}$	-0.103	-0.43379	-0.42063	-0.41685	0.44903	0.30089	0.21372
$\Gamma^d_{1,21}$	-0.068	0.09285	0.09270	0.09222	0.06667	0.04495	0.03112
$\Gamma^d_{1,22}$	-0.002	-0.42835	-0.42861	-0.42894	0.15574	0.10702	0.07485
Johansen's VECM(2) estimates							
γ_1	-0.437	-0.43902	-0.40865	-0.39636	0.22926	0.14996	0.10191
γ_2	-0.251	-0.14058	-0.13334	-0.13032	0.08966	0.05699	0.03930
λ_1	1	0.00926	0.00462	0.00207	0.07574	0.02910	0.01302
$\Gamma^d_{1,11}$	0.141	0.01160	-0.01404	-0.02498	0.22118	0.14628	0.10092
$\Gamma^d_{1,12}$	-0.103	-0.69037	-0.66320	-0.65087	0.54082	0.35788	0.24753
$\Gamma^d_{1,21}$	-0.068	0.14984	0.14489	0.14229	0.08697	0.05617	0.03905
$\Gamma^d_{1,22}$	-0.002	-0.56827	-0.56069	-0.55713	0.20479	0.13651	0.09543
Parameter		Mean of the estimated parameter					
$\Gamma^d_{1,11}$		0.02926	0.02117	0.01635	0.17140	0.11371	0.07799
$\Gamma^d_{2,12}$		-0.37970	-0.35394	-0.33524	0.48865	0.32767	0.22270
$\Gamma^d_{2,21}$		0.01968	0.01362	0.01142	0.06841	0.04507	0.03105
$\Gamma^d_{2,22}$		-0.21319	-0.19535	-0.18593	0.18158	0.12206	0.08538

Table 3.4 Design 2 estimates for stock variables (negative correlation)

		Sample size			Sample size		
Continuous time estimates		50	100	200	50	100	200
		Bias			Standard Error		
Parameter	True value						
g_1	-1	0.00584	0.00570	0.00531	0.17369	0.11219	0.07584
g_2	-2	0.03748	0.02895	0.02142	0.40393	0.26895	0.19021
b_1	1	-0.00138	-0.00059	-0.00015	0.02943	0.01208	0.00566
$\Gamma_{1,11}$	-2	0.04703	0.03161	0.02141	0.27756	0.18476	0.12857
ρ	-0.5	-0.00191	-0.00164	0.00161	0.14287	0.09846	0.06869
σ_2	1	-0.07890	-0.04155	-0.02434	0.03843	0.01822	0.00890
Johansen's VECM(1) estimates							
γ_1	-0.437	-0.33063	-0.31778	-0.31414	0.16054	0.10668	0.07523
γ_2	-0.251	0.01812	0.01295	0.01122	0.10084	0.06707	0.04678
λ_1	1	0.00417	0.00214	0.00095	0.03086	0.01259	0.00585
$\Gamma^d_{1,11}$	0.141	-0.16853	-0.17980	-0.18385	0.10333	0.07021	0.04915
$\Gamma^d_{1,12}$	-0.103	-0.48063	-0.47301	-0.47141	0.44391	0.29705	0.20960
$\Gamma^d_{1,21}$	-0.068	0.06972	0.07622	0.07820	0.06375	0.04301	0.03010
$\Gamma^d_{1,22}$	-0.002	-0.26308	-0.26625	-0.26705	0.26219	0.17582	0.12407
Johansen's VECM(2) estimates							
γ_1	-0.437	-0.50976	-0.49318	-0.48542	0.26788	0.17778	0.12069
γ_2	-0.251	-0.00923	-0.00663	-0.00568	0.16854	0.11166	0.07668
λ_1	1	0.00204	0.00118	0.00052	0.03168	0.01253	0.00576
$\Gamma^d_{1,11}$	0.141	-0.05227	-0.06709	-0.07430	0.18061	0.12080	0.08291
$\Gamma^d_{1,12}$	-0.103	-0.75989	-0.74527	-0.73745	0.55758	0.36866	0.25347
$\Gamma^d_{1,21}$	-0.068	0.08895	0.09003	0.09014	0.10857	0.07233	0.05005
$\Gamma^d_{1,22}$	-0.002	-0.30286	-0.29519	-0.29222	0.34276	0.22659	0.15708
Parameter		Mean of the estimated parameter					
$\Gamma^d_{2,11}$		-0.02264	-0.02340	-0.02339	0.10588	0.07122	0.04920
$\Gamma^d_{2,12}$		-0.41608	-0.40323	-0.39098	0.48917	0.32821	0.22188
$\Gamma^d_{2,21}$		0.00730	0.00648	0.00679	0.06394	0.04391	0.03043
$\Gamma^d_{2,22}$		-0.06160	-0.04400	-0.03788	0.29537	0.19975	0.13698

3.5.2 VECM Simulations With Flows

For flows, similarly than with stocks, the data generating process, which is used to generate ten thousand simulations of 50, 100 and 200 sample sizes, follows a VECM representation that can be written as

$$\begin{aligned} y_1^f &= \eta_1^f, \\ \Delta y_2^f &= \mathit{\Lambda}_2(\theta) y_1^f + \eta_2^f, \\ \begin{pmatrix} \Delta y_{1t}^f \\ \Delta y_{2t}^f \end{pmatrix} &= \begin{pmatrix} k_1 \\ k_2 \end{pmatrix} (y_{1,t-1}^f - b_1 y_{2,t-1}^f) + \Gamma_1 \Delta y_{t-1}^f + \eta_t^f, \end{aligned} \tag{3.39}$$

where η_t^f follows a moving average representation or order 2 as given in Lemma 3.4.2 and η^f, in that same Lemma, is assumed to be $N(0, \Omega^f)$.

The true parametric designs are given as before and for estimation, we apply our methodology as described in subsection 3.4.2 to the generated data as well as Johansen's to it as if it was generated by the VECM(1) specification given in (3.33). At the same time, for the reasons outlined before, we also apply Johansen's methodology to our generated data as if it was specified through a VECM representation of order 3 given by

$$\Delta y_t = \gamma \lambda' y_{t-1} + \Gamma_1^d \Delta y_{t-1} + \Gamma_2^d \Delta y_{t-2} + \Gamma_3^d \Delta y_{t-3} + s_t \tag{3.40}$$

where all matrices and components are similarly described as above.

For comparison, finally, we consider the relevant parameter estimates of this system and measure how close they are with the respective true parameter values in (3.36).

Similarly than in the stock variable case, results have been grouped and appear in tables 3.5, 3.6, 3.7 and 3.8 so that all the designs and all the correlations are explained. As before, each table shows the true value, the bias and the standard error. For better understanding of the results, each table is analogously divided in three sections with a lower part that reports only the mean value and the standard error of the estimates of Γ_2^d and Γ_3^d.

Looking at the tables, in general, a similar picture of the results outlined in the previous section emerges; for all the parameters and in all the designs, our methodology shows superiority in estimation against Johansen's, with smaller bias and improvements as the sample size increases. Similarly, Johansen's VECM(1) estimates show a persistent bias with no improvements as the sample size increases.

Also, looking at Johansen's VECM(3) parameters, we see that the inclusion of two additional lags into the specification does not improves the estimates.

Table 3.5 Design 1 estimates for flow variables (positive correlation)

		Sample size			Sample size		
Continuous time estimates		50	100	200	50	100	200
		Bias			Standard Error		
Parameter	True value						
g_1	1	-0.00815	-0.00803	-0.00081	0.46657	0.29346	0.20024
g_2	2	0.16410	0.06564	0.00881	0.48858	0.30117	0.20368
b_1	1	0.00407	0.00075	0.00003	0.05716	0.02370	0.01159
$\Gamma_{1,11}$	-3	-0.12608	-0.04951	-0.00219	0.46667	0.30418	0.20983
ρ	0.5	-0.01146	-0.00532	-0.00166	0.13321	0.09244	0.06379
σ_2	1	-0.02970	-0.01765	-0.01582	0.03405	0.01604	0.00765
Johansen's VECM(1) estimates							
γ_1	0.294	-0.19534	-0.19328	-0.19598	0.12378	0.07460	0.04845
γ_2	0.588	-0.24790	-0.27190	-0.28561	0.10619	0.06545	0.04313
λ_1	1	-0.01137	-0.00512	-0.00210	0.09245	0.02796	0.01223
$\Gamma^d_{1,11}$	0.05	0.72401	0.74631	0.75772	0.18951	0.12286	0.08280
$\Gamma^d_{1,12}$	0	-0.23331	-0.23769	-0.23970	0.16057	0.10525	0.07277
$\Gamma^d_{1,21}$	0.05	0.19793	0.22581	0.24056	0.19922	0.13214	0.09098
$\Gamma^d_{1,22}$	0	0.28409	0.27781	0.27506	0.15040	0.09881	0.06776
Johansen's VECM(3) estimates							
γ_1	0.294	-0.12260	-0.13103	-0.13887	0.17485	0.09843	0.06204
γ_2	0.588	-0.20662	-0.24466	-0.26436	0.15582	0.08688	0.05576
λ_1	1	-0.00853	-0.00115	-0.00042	0.07688	0.04785	0.01258
$\Gamma^d_{1,11}$	0.05	0.78813	0.83513	0.85745	0.25104	0.15785	0.10468
$\Gamma^d_{1,12}$	0	-0.13756	-0.14970	-0.15600	0.20791	0.13314	0.09090
$\Gamma^d_{1,21}$	0.05	0.17424	0.21623	0.23645	0.23097	0.14325	0.09696
$\Gamma^d_{1,22}$	0	0.47049	0.47004	0.46836	0.18825	0.12119	0.08285
Parameter		Mean of the estimated parameter					
$\Gamma^d_{2,11}$		-0.41316	-0.42336	-0.42889	0.24184	0.15945	0.11009
$\Gamma^d_{2,12}$		0.07776	0.07737	0.07831	0.21974	0.14269	0.09995
$\Gamma^d_{2,21}$		-0.16739	-0.16232	-0.15845	0.22367	0.14979	0.10366
$\Gamma^d_{2,22}$		-0.16975	-0.19008	-0.19914	0.19211	0.12726	0.08798
$\Gamma^d_{3,11}$		0.11220	0.12408	0.13323	0.24791	0.16302	0.11171
$\Gamma^d_{3,12}$		-0.03621	-0.03579	-0.03880	0.18371	0.11884	0.08114
$\Gamma^d_{3,21}$		0.00236	0.02981	0.04448	0.23027	0.14996	0.10182
$\Gamma^d_{3,22}$		0.06264	0.05681	0.05143	0.17232	0.10913	0.07416

Table 3.6 Design 1 estimates for flow variables (negative correlation)

		Sample size			Sample size		
Continuous time estimates		50	100	200	50	100	200
		Bias			Standard Error		
Parameter	True value						
g_1	1	-0.01974	-0.01767	-0.01595	0.322717	0.211658	0.149154
g_2	2	0.120239	0.040049	-0.00435	0.350868	0.225176	0.157702
b_1	1	0.000432	0.000136	2.85E-06	0.024264	0.010464	0.004884
$\Gamma_{1,11}$	-3	-0.09527	-0.02392	0.022929	0.466546	0.305166	0.212509
ρ	-0.5	-0.0096	0.001111	0.010321	0.134515	0.094515	0.073658
σ_2	1	-0.02262	-0.01734	-0.0167	0.035504	0.016857	0.009399
Johansen's VECM(1) estimates							
γ_1	0.294	-0.19564	-0.19303	-0.19552	0.165902	0.104053	0.068729
γ_2	0.588	-0.2509	-0.27616	-0.28673	0.120085	0.076759	0.051074
λ_1	1	-0.00425	-0.00156	-0.00062	0.037351	0.011827	0.005069
$\Gamma^d_{1,11}$	0.05	0.727211	0.744228	0.756423	0.280306	0.17729	0.116838
$\Gamma^d_{1,12}$	0	-0.24005	-0.23997	-0.2402	0.090226	0.059502	0.041292
$\Gamma^d_{1,21}$	0.05	0.195626	0.226388	0.239487	0.220618	0.144509	0.097024
$\Gamma^d_{1,22}$	0	0.271093	0.272527	0.273356	0.066212	0.043149	0.029412
Johansen's VECM(3) estimates							
γ_1	0.294	-0.11677	-0.12616	-0.13405	0.243996	0.140089	0.089845
γ_2	0.588	-0.21224	-0.25295	-0.26934	0.177095	0.103692	0.066322
λ_1	1	0.003236	-0.00024	-0.00013	0.042674	0.019973	0.005182
$\Gamma^d_{1,11}$	0.05	0.77188	0.817266	0.840393	0.359667	0.215289	0.139773
$\Gamma^d_{1,12}$	0	-0.15067	-0.16033	-0.16622	0.226636	0.145045	0.099403
$\Gamma^d_{1,21}$	0.05	0.180908	0.226529	0.245348	0.262437	0.159497	0.105271
$\Gamma^d_{1,22}$	0	0.472515	0.474915	0.476194	0.167358	0.107834	0.074083
Parameter		Mean of the estimated parameter					
$\Gamma^d_{2,11}$		-0.39624	-0.40569	-0.40798	0.320506	0.212405	0.147212
$\Gamma^d_{2,12}$		0.086175	0.0857	0.086925	0.240912	0.15501	0.107252
$\Gamma^d_{2,21}$		-0.18255	-0.17355	-0.17002	0.247671	0.16166	0.112372
$\Gamma^d_{2,22}$		-0.17772	-0.19855	-0.2083	0.169023	0.112392	0.07681
$\Gamma^d_{3,11}$		0.101788	0.114524	0.12079	0.343592	0.222291	0.153595
$\Gamma^d_{3,12}$		-0.04369	-0.04074	-0.04094	0.114576	0.074117	0.051292
$\Gamma^d_{3,21}$		0.013289	0.040803	0.052853	0.257374	0.166765	0.113462
$\Gamma^d_{3,22}$		0.051173	0.051489	0.051315	0.08668	0.055806	0.038389

Table 3.7 Design 2 estimates for flow variables (positive correlation)

		Sample size			Sample size		
Continuous time estimates		50	100	200	50	100	200
		Bias			Standard Error		
Parameter	True value						
g_1	-1	-0.10616	-0.03733	-0.00173	0.28608	0.17456	0.11633
g_2	-2	-0.10765	-0.01890	0.00667	0.40225	0.25446	0.17687
b_1	1	0.00060	-0.00011	0.00005	0.07384	0.03189	0.01527
$\Gamma_{1,11}$	-2	-0.08448	-0.01864	0.01324	0.30043	0.20079	0.14114
ρ	0.5	-0.02145	-0.01149	-0.00690	0.13102	0.09086	0.06533
σ_2	1	-0.02117	-0.02057	-0.02022	0.03641	0.01750	0.00882
Johansen's VECM(1) estimates							
Parameter	True value						
γ_1	-0.437	0.24452	0.26518	0.27490	0.13159	0.08169	0.05464
γ_2	-0.251	0.08682	0.09487	0.09906	0.04125	0.02558	0.01710
λ_1	1	0.00037	0.00008	-0.00046	0.07393	0.06662	0.01573
$\Gamma^d_{1,11}$	0.141	0.23816	0.23075	0.22743	0.15501	0.10186	0.07038
$\Gamma^d_{1,12}$	-0.103	0.60925	0.64488	0.66244	0.53093	0.34466	0.23552
$\Gamma^d_{1,21}$	-0.068	-0.08106	-0.08659	-0.08983	0.05782	0.03688	0.02502
$\Gamma^d_{1,22}$	-0.002	0.44105	0.47737	0.49563	0.16629	0.10821	0.07378
Johansen's VECM(3) estimates							
Parameter	True value						
γ_1	-0.437	0.18701	0.21442	0.22776	0.17949	0.10182	0.06517
γ_2	-0.251	0.05872	0.07596	0.08428	0.06692	0.03838	0.02524
λ_1	1	0.00204	0.00098	0.00022	0.09299	0.03824	0.01636
$\Gamma^d_{1,11}$	0.141	0.55330	0.55920	0.55990	0.22971	0.14258	0.09633
$\Gamma^d_{1,12}$	-0.103	0.27203	0.31165	0.33108	0.56116	0.34895	0.23421
$\Gamma^d_{1,21}$	-0.068	-0.06848	-0.08562	-0.09433	0.09156	0.05682	0.03894
$\Gamma^d_{1,22}$	-0.002	0.48415	0.54339	0.57021	0.21575	0.13595	0.09191
Parameter		Mean of the estimated parameter					
$\Gamma^d_{2,11}$		-0.32726	-0.34923	-0.36038	0.24471	0.16276	0.11173
$\Gamma^d_{2,12}$		-0.20423	-0.19095	-0.18450	0.54595	0.36199	0.24710
$\Gamma^d_{2,21}$		0.07709	0.07284	0.07068	0.09761	0.06398	0.04531
$\Gamma^d_{2,22}$		-0.22054	-0.20764	-0.20155	0.20323	0.13762	0.09612
$\Gamma^d_{3,11}$		0.10559	0.10188	0.09893	0.18374	0.11887	0.08102
$\Gamma^d_{3,12}$		0.07197	0.09198	0.10784	0.52417	0.33593	0.22636
$\Gamma^d_{3,21}$		-0.01142	-0.01482	-0.01657	0.07719	0.04921	0.03348
$\Gamma^d_{3,22}$		0.01054	0.02693	0.03678	0.19663	0.12897	0.08931

Table 3.8 Design 2 estimates for flow variables (negative correlation)

		Sample size			Sample size		
Continuous time estimates		50	100	200	50	100	200
		Bias			Standard Error		
Parameter	True value						
g_1	-1	-0.08000	-0.03129	-0.00790	0.22456	0.14499	0.10000
g_2	-2	-0.02919	0.04446	0.00769	0.36065	0.24342	0.17615
b_1	1	-0.00063	-0.00030	-0.00012	0.03371	0.01423	0.00661
$\Gamma_{1,11}$	-2	-0.03802	0.01758	0.04784	0.29204	0.19863	0.14390
ρ	-0.5	0.00616	0.00644	0.00144	0.13160	0.09104	0.06701
σ_2	1	-0.04779	-0.04538	-0.03398	0.03631	0.01738	0.00917
Johansen's VECM(1) estimates							
γ_1	-0.437	0.14133	0.16213	0.17437	0.17523	0.11262	0.07645
γ_2	-0.251	0.10834	0.11483	0.11766	0.07152	0.04594	0.03105
λ_1	1	0.00713	0.00302	0.00129	0.04069	0.01507	0.00683
$\Gamma^d_{1,11}$	0.141	0.32888	0.32443	0.32178	0.07918	0.05182	0.03588
$\Gamma^d_{1,12}$	-0.103	0.07459	0.10810	0.13255	0.59987	0.39069	0.26723
$\Gamma^d_{1,21}$	-0.068	-0.08473	-0.08694	-0.08836	0.03984	0.02621	0.01816
$\Gamma^d_{1,22}$	-0.002	0.49320	0.52712	0.54240	0.24387	0.15759	0.10649
Johansen's VECM(3) estimates							
γ_1	-0.437	0.17146	0.19651	0.20999	0.24731	0.14373	0.09233
γ_2	-0.251	0.06372	0.08020	0.08698	0.11626	0.06830	0.04443
λ_1	1	0.00230	0.00117	0.00050	0.05659	0.01640	0.00700
$\Gamma^d_{1,11}$	0.141	0.56700	0.57603	0.57793	0.22157	0.14059	0.09427
$\Gamma^d_{1,12}$	-0.103	0.14397	0.17715	0.19457	0.60090	0.36943	0.24622
$\Gamma^d_{1,21}$	-0.068	-0.06776	-0.08024	-0.08618	0.10760	0.06858	0.04570
$\Gamma^d_{1,22}$	-0.002	0.48912	0.54599	0.56976	0.28491	0.17664	0.11820
Parameter		Mean of the estimated parameter					
	$\Gamma^d_{2,11}$	-0.35158	-0.37033	-0.38100	0.20719	0.13497	0.09177
	$\Gamma^d_{2,12}$	-0.06005	-0.04671	-0.03765	0.56702	0.37257	0.25515
	$\Gamma^d_{2,21}$	0.06101	0.05798	0.05780	0.10099	0.06685	0.04566
	$\Gamma^d_{2,22}$	-0.20449	-0.19765	-0.19374	0.26421	0.17635	0.12156
	$\Gamma^d_{3,11}$	0.13016	0.12772	0.12673	0.11769	0.07800	0.05279
	$\Gamma^d_{3,12}$	-0.03523	-0.01984	-0.00676	0.58203	0.37388	0.25164
	$\Gamma^d_{3,21}$	-0.01104	-0.01171	-0.01177	0.05972	0.03943	0.02644
	$\Gamma^d_{3,22}$	0.00855	0.02831	0.03462	0.26734	0.17599	0.12136

3.6 Concluding remarks

In this chapter, as an extension of the analysis we provided before, we have developed an estimation procedure for cointegrated systems in continuous time that not only allows for higher order specifications in the system but also incorporates deterministic components on it.

With this new specification the structure of the chapter is maintained; we allow the system to be entirely comprised of stock or flow variables, derive, for each type of data, an exact discrete time model and, with the use of an alternative exponential matrix factorization, outline the autocovariance representations of the discrete time disturbances.

At the end, with a simple version of the original cointegrated system and using some simulated data as well as two different parametric designs, we evaluate the performance of our estimation procedure by measuring closeness between our estimated parameters and the true parametric designs. Also, with the aim of measuring the costs, if there are any, of ignoring aggregation in the specification, we compare our results with the ones we would have obtained by imposing instead a discrete time specification (Johansen's specification) into the system.

In the results, we strengthened the observations we outlined in the previous Chapter and showed that regardless of the variables being stocks of flows, when dynamics play an important role in the specification, our estimation procedure is always superior to Johansen's with more accurate parameters and improvements as the sample size increases. In other words, when dynamics play an important role in the specification, Johansen's estimates suffer from temporal aggregation bias.

Also, when Johansen's methodology was considered, we saw that the inclusion of additional lags into the specification did not improve the estimates and in some cases, it even made them worse.

Appendix C

Supplementary Results

Reduced rank factorization and exponential representation

Given the reduced rank assumptions of our system following (3.1), it can be shown (see Appendix D for details) that the matrix A can rewritten in terms of two reduced rank matrices U and V such that

$$A = UV' \tag{C.1}$$

where V' is a $(nk - n_2) \times nk$ matrix whose rows are the non zero rows of the reduced row echelon form of A and U is a $nk \times (nk - n_2)$ matrix whose elements are known functions of the given parameters in the system.

Then, it can also be shown (see Appendix D for details) that our exponential matrix e^{tA} can be rewritten as

$$\begin{aligned} e^{tA} = e^{tUV'} = \sum_{j=0}^{\infty} \frac{(tUV')^j}{j!} = I_{nk} + \frac{UV'}{1!}t + \frac{(UV')(UV')}{2!}t^2 + \cdots , \\ = I_{nk} + UH^{-1}(e^{Ht} - I_{nk-n_2})V'. \end{aligned} \tag{C.2}$$

where $H = V'U$ (See Appendix D for details).

Covariance matrix computation.

For the computations of the covariance matrices Ω^s and Ω^f, we follow the procedure as outlined in Chapter 1 and reduce the computations to calculate either the exponential matrix e^H or the integral $\Psi = \int_0^1 UH^{-1}e^{sH}V'\Sigma^* V e^{sH'}(H^{-1})'U'ds$, which was shown to be further reduced to compute only the following exponential matrix

$$\varpi = exp\begin{pmatrix} -H & V'\Sigma V \\ 0 & H' \end{pmatrix} = \begin{pmatrix} \varpi_{11} & \varpi_{12} \\ 0 & \varpi_{22} \end{pmatrix}$$

where, $\Psi = UH^{-1}\big(\varpi'_{22}\varpi_{12}\big)(H^{-1})'U'$ and $e^{H} = \varpi'_{22}$.

For validation, two different procedures were considered; the truncation of the infinite series representation of the exponential matrix and the Pade approximation method.

At the end, as mentioned in Chapter 1, the differences were small enough to be neglected and, as a result, for computation efficiency, the Pade approximation approach was applied in the analysis.

Appendix D

Proofs

Proof of (C.1).

Considering the assumptions of our system following (3.1)

$$A = \begin{bmatrix} 0 & I_n & 0 & \cdots & 0 \\ 0 & 0 & I_n & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & I_n \\ A_0 & A_1 & A_2 & \cdots & A_{k-1} \end{bmatrix}$$

$$= \begin{bmatrix} 0_{n_1\times n_1} & 0_{n_1\times n_2} & I_{n_1\times n_1} & 0_{n_1\times n_2} & \cdots & 0_{n_1\times n_1} & 0_{n_1\times n_2} \\ 0_{n_2\times n_1} & 0_{n_2\times n_2} & 0_{n_2\times n_1} & I_{n_2\times n_2} & \cdots & 0_{n_2\times n_1} & 0_{n_2\times n_2} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0_{n_1\times n_1} & 0_{n_1\times n_2} & 0_{n_1\times n_1} & 0_{n_1\times n_2} & \cdots & I_{n_1\times n_1} & 0_{n_1\times n_2} \\ 0_{n_2\times n_1} & 0_{n_2\times n_2} & 0_{n_2\times n_1} & 0_{n_2\times n_2} & \cdots & 0_{n_2\times n_1} & I_{n_2\times n_2} \\ Ç_1 & -Ç_1 B_1 & A_{1,11} & A_{1,12} & \cdots & A_{k-1,11} & A_{k-1,12} \\ Ç_2 & -Ç_2 B_1 & A_{1,21} & A_{1,22} & \cdots & A_{k-1,21} & A_{k-1,22} \end{bmatrix}$$

where the matrix A_0 has been rewritten in terms of $ÇB'$ and the matrices $A_1, \ldots, A_{k-1}$ have been partitioned according to n_1 and n_2 so that $A_{i,11}$ corresponds to the first n_1 rows and n_1 columns of the matrix A_i $(i = 1, \cdots, k-1)$, $A_{i,12}$ corresponds also to the first n_1 rows and the last n_2 columns of the matrix A_i and $A_{i,21}$ and $A_{i,22}$, respectively, correspond to the last n_2 rows of the same matrix A_i.

As the matrix A_0 is reduced rank, it is easy to see that A is also reduced rank and can be rewritten in terms of two reduced rank matrices U and V such that $A = UV'$. The $(nk - n_2) \times nk$ matrix V' will always consists of the non zero rows of the reduced row echelon form of A and is given by

$$V' = \begin{bmatrix} I_{n_1\times n_1} & -B_1 & 0_{n_1\times n_1} & 0_{n_1\times n_2} & \cdots & 0_{n_1\times n_1} & 0_{n_1\times n_2} \\ 0_{n_1\times n_1} & 0_{n_1\times n_2} & I_{n_1\times n_1} & 0_{n_1\times n_2} & \cdots & 0_{n_1\times n_1} & 0_{n_1\times n_2} \\ 0_{n_2\times n_1} & 0_{n_2\times n_2} & 0_{n_2\times n_1} & I_{n_2\times n_2} & \cdots & 0_{n_2\times n_1} & 0_{n_2\times n_2} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0_{n_1\times n_1} & 0_{n_1\times n_2} & 0_{n_1\times n_1} & 0_{n_1\times n_2} & \cdots & I_{n_1\times n_1} & 0_{n_1\times n_2} \\ 0_{n_2\times n_1} & 0_{n_2\times n_2} & 0_{n_2\times n_1} & 0_{n_2\times n_2} & \cdots & 0_{n_2\times n_1} & I_{n_2\times n_2} \end{bmatrix}. \tag{D.1}$$

The $nk \times (nk - n_2)$ matrix U, then, is obtained by removing from A all non pivot columns of V' and can be written as

$$U = \begin{bmatrix} 0_{n_1\times n_1} & I_{n_1\times n_1} & 0_{n_1\times n_2} & \cdots & 0_{n_1\times n_1} & 0_{n_1\times n_2} \\ 0_{n_2\times n_1} & 0_{n_2\times n_1} & I_{n_2\times n_2} & \cdots & 0_{n_2\times n_1} & 0_{n_2\times n_2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0_{n_1\times n_1} & 0_{n_1\times n_1} & 0_{n_1\times n_2} & \cdots & I_{n_1\times n_1} & 0_{n_1\times n_2} \\ 0_{n_2\times n_1} & 0_{n_2\times n_1} & 0_{n_2\times n_2} & \cdots & 0_{n_2\times n_1} & I_{n_2\times n_2} \\ Ç_1 & A_{1,11} & A_{1,12} & \cdots & A_{k-1,11} & A_{k-1,12} \\ Ç_2 & A_{1,21} & A_{1,22} & \cdots & A_{k-1,21} & A_{k-1,22} \end{bmatrix}. \tag{D.2}$$

Note that for existence of these matrices, not additional assumptions are required, therefore, Ç, particularly, can be singular or not as long as it follows the definitions of our system as in (3.1). Also, note that V', given the assumptions of the reduced row echelon form, is always unique.

■

Proof of (C.2).

Considering that

$$e^{tUV'} = I_{nk} + \sum_{j=1}^{\infty} t^j \frac{(UV')^j}{j!} = I_{nk} + \frac{UV'}{1!}t + \frac{(UV')(UV')}{2!}t^2 + \cdots .$$

We can write

$$\begin{aligned}(UV')^j &= \underbrace{(UV') \times (UV') \times \cdots \times (UV')}_{j \text{ times}},\\ &= U \times \underbrace{(V'U) \times (V'U) \times \cdots \times (V'U)}_{j-1 \text{ times}} \times V'.\end{aligned}$$

Let $V'U = H$, then

$$(UV')^j = U \times H^{j-1} \times V'.$$

Hence

$$\begin{aligned}e^{tUV'} &= I_{nk} + U\sum_{j=1}^{\infty} \frac{t^j}{j!} H^{j-1} V' = I_{nk} + UH^{-1} \sum_{j=1}^{\infty} \frac{(tH)^j}{j!} V',\\ &= I_{nk} + UH^{-1}(e^{tH} - I_{nk-n_2})V'.\end{aligned}$$

Note that, without loss of generality and considering the expression of U and V' in (D.2) and (D.1), the $(nk - n_2) \times (nk - n_2)$ matrix H can be written as

$$H = V'U = \begin{bmatrix} 0_{n_1\times n_1} & I_{n_1\times n_1} & -B_1 & \cdots & 0_{n_1\times n_1} & 0_{n_1\times n_2}\\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots\\ 0_{n_1\times n_1} & 0_{n_1\times n_1} & 0_{n_1\times n_2} & \cdots & I_{n_1\times n_1} & 0_{n_1\times n_2}\\ 0_{n_2\times n_1} & 0_{n_2\times n_1} & 0_{n_2\times n_2} & \cdots & 0_{n_2\times n_1} & I_{n_2\times n_2}\\ Ç_1 & A_{1,11} & A_{1,12} & \cdots & A_{k-1,11} & A_{k-1,12}\\ Ç_2 & A_{1,21} & A_{1,22} & \cdots & A_{k-1,21} & A_{k-1,22}\end{bmatrix}. \tag{D.3}$$

■

Proof of Lemma 3.3.1.

Given by Chambers [1999], the solution to the system of equations expressed in (3.8) and (3.9), can be written in a VAR format as

$$\begin{aligned} y_t^s &= C_{11}y_{t-1}^s + C_{12}\left[MN\bar{y}_t^s + M\bar{m}_t^s + M\bar{\varepsilon}_t^s\right] + m_{1t}^s + \varepsilon_{1t}^s, \\ &= C_{11}y_{t-1}^s + C_{12}MN_1y_{t-1}^s + C_{12}MN_2y_{t-2}^s + \cdots \\ &\quad + C_{12}MN_ky_{t-k}^s + C_{12}M\bar{m}_t^s + m_{1t}^s + C_{12}M\bar{\varepsilon}_t^s + \varepsilon_{1t}^s, \\ &= F_1y_{t-1}^s + F_2y_{t-2}^s + \cdots + F_ky_{t-k}^s + g_t^s(\theta) + \eta_t^s, \end{aligned}$$

where all the matrices are defined in the Lemma.

Finally, the VECM form is obtained by rewriting the last line of the previous expression as follows

$$\begin{aligned} y_t^s - y_{t-1}^s = F_1y_{t-1}^s - y_{t-1}^s + F_2y_{t-1}^s - F_2y_{t-1}^s + F_2y_{t-2}^s + \\ F_3y_{t-1}^s - F_3y_{t-1}^s + F_3y_{t-2}^s - F_3y_{t-2}^s + F_3y_{t-3}^s + \\ \cdots + F_ky_{t-k}^s + g_t^s(\theta) + \eta_t^s, \end{aligned}$$

$$\Delta y_t^s = \Pi_k(\theta)y_{t-1}^s + \Gamma_1(\theta)\Delta y_{t-1}^s + \cdots + \Gamma_{k-1}(\theta)\Delta y_{t-(k-1)}^s + g_t^s(\theta) + \eta_t^s. \quad \text{(D.4)}$$

■

Proof of Lemma 3.3.2.

This proof is exactly a mirror image of the proof in Lemma 3.3.1 with the interchange of stock variables for flow variables following the definitions in (3.11).

■

Proof of Lemma 3.3.3.

Considering the definition of the observed vector for stock variables, the solution to the system in (3.3) can be rewritten as

$$x_1^s = e^Ax(0) + \int_0^1 e^{(1-s)A}(a^* + b^*s)ds + \int_0^1 e^{(1-s)A}\zeta^*(ds), \quad \text{(D.5)}$$

which premultiplied by S_1 gives the expression for y_1^s.

At the same time, following Chambers [1999], the VAR expression for y_t^s $(t = 2, \ldots, k)$ is given by

$$y_t^s = J_1y_{t-1}^s + \cdots + J_{t-1}y_1^s + C_{12}C_{22}^{t-2}S_2e^Ax^s(0) + q_t^s + \eta_t^s, \quad t \geq 2. \quad \text{(D.6)}$$

Finally, the VECM representation in the lemma is achieved by applying to (D.6) a similar procedure to that in the derivation of equation (D.4).

■

Proof of Lemma 3.3.4.

This proof is exactly a mirror image of the proof in Lemma 3.3.3 with the interchange of stock variables for flow variables following the definitions in (3.11) and the fact that $x_1^f = \int_0^1 x^f(r)dr$.

■

Proof of Lemma 3.4.1.

Following their definitions as mentioned in Lemma 3.3.1, the expression for $\eta_{k+1}^s, \cdots, \eta_T^s$ can be rewritten as

$$\begin{aligned}\eta_t^s =& S_1\varepsilon_t + C_{12}\sum_{i=1}^{k-1} M_{1i}S_1\varepsilon_{t-i} + C_{12}\sum_{i=1}^{k-1} M_{i+1}S_2\varepsilon_{t-i},\\ =& \sum_{i=0}^{k-1} R_i^s\varepsilon_{t-i}, \qquad t = k+1, \cdots, T,\end{aligned} \tag{D.7}$$

where the matrices M_i, M_{ij} and R_i^s are defined in the Lemma.

The expression for $\eta_1^s, \cdots, \eta_k^s$, at the same time, with their definitions as in Lemma 3.3.3 are given by

$$\begin{aligned}\eta_t^s =& S_1\varepsilon_t + C_{12}\sum_{j=0}^{t-3} C_{22}^j S_2\varepsilon_{t-1-j} + C_{12}C_{22}^{t-2}S_2\varepsilon_1,\\ =& \sum_{i=0}^{t-1} P_i^s\varepsilon_{t-i}, \qquad t = 1, \cdots, k,\end{aligned} \tag{D.8}$$

where the matrices P_i^s are defined in the Lemma.

For the autocovariances of η_t^s, finally, we consider the autocovariances of ε_t

so that Ω_ε is obtained as

$$\begin{aligned}
E[\varepsilon_t\varepsilon_t'] =& E\Big[\int_{t-1}^{t} e^{A(t-s)}\zeta^*(ds)\Big]\Big[\int_{t-1}^{t} e^{A(t-s)}\zeta^*(ds)\Big]',\\
=& \int_{t-1}^{t}\Big(e^{A(t-s)}\Big)\Sigma^*\Big(e^{A(t-s)}\Big)'ds,\\
=& \int_{0}^{1}\Big(e^{As}\Big)\Sigma^*\Big(e^{As}\Big)'ds,\\
=& \int_{0}^{1}\big[I_{nk}+UH^{-1}(e^{Hs}-I_{nk-n_2})V'\big]\Sigma^*\big[I_{nk}+UH^{-1}(e^{Hs}-I_{nk-n_2})V'\big]'ds,
\end{aligned} \tag{D.9}$$

where $E[\zeta^*(ds)\zeta^*(ds)'] = \Sigma^* ds$.

■

Proof of (3.26).

Considering the alternative representation of our exponential matrix e^{tA} as in (C.2), the integrals inside the square brackets of (3.25) are calculated as follows

$$\begin{aligned}
\int_a^b e^{(r-s)A}dr =& \int_a^b \Big\{I_{nk}+UH^{-1}(e^{H(r-s)}-I_{nk-n_2})V'\Big\}dr,\\
&= (b-a)(I_{nk}-UH^{-1}V') + UH^{-2}\big[e^{(b-s)H}-e^{(a-s)H}\big]V',
\end{aligned} \tag{D.10}$$

where the constants a and b are changed as needed.

■

Proof of Lemma 3.4.2.

This proof is exactly a mirror image of the proof of Lemma 3.4.1 with the interchange of η_t^s for η_t^f and the fact that $v_t^f = v_{a,t}^f + v_{b,t-1}^f$, where $v_{a,t}^f$ and $v_{b,t-1}^f$ are defined in the Lemma and are calculated from equation (3.26).

■

Chapter 4

Empirical Applications

This chapter presents three multivariate applications of the estimation methodology for cointegrated systems in continuous time developed in the previous chapters. The analysis is carried out by comparing the estimates of the model's parameters considering two different time specification; Johansen's general VECM for discrete time and our exact discrete time VECM for continuous time. The applications evaluate, for the United States, the market efficiency hypothesis on the foreign exchange rate, the term structure of interest rates and the main implication of the rational-expectation permanent income hypothesis. In the results, it is shown that estimation bias in cointegrated systems does not only depend on whether the variables in the model suffer some sort of temporal aggregation, but also, on whether the system requires a higher order specification.

4.1 Introduction

With the aim of analysing the effects of temporal aggregation over the estimates of a model's parameters, focusing on the non stationary cointegrated case, this chapter presents three multivariate applications of the estimation methodology for systems in continuous time developed in the previous chapters.

In the analysis, for each application, the estimated parameters of the model are compared through the use of two different time specifications; Johansen's general VECM for discrete time and our exact discrete time VECM for continuous time. Given that the representation of the estimates differs dramatically with the time specification, a one-to-one comparison cannot be directly considered, as a result, in here, such comparison is carried out by utilizing the estimates of one specification and the translated values (the *implied values*) of the estimates of the other.

For the United States, the applications evaluate the market efficiency hypothesis on the foreign exchange rate, the term structure of interest rates and the main implication of the rational expectation-permanent income hypothesis. Since it is not obvious that these relationships exist, we also provide statistical justifications for the analysis. In all the cases, standard theoretical models are utilized to present the cointegrating relationships and, if required, likelihood ratio (LR) tests are applied to identify the specification of the model that fits the data the best. In each application, the sampling frequency of the variables as well as the aggregation method (if any) is considered, therefore, when applying our continuous time methodology, the most suited specification (flows or stock variables) is used.

In the results, the first application considers a first order system and a stock variables specification. As a result, both our continuous time methodology as well as Johansen's produce virtually identical estimates and base line conclusion are drawn. The second application presents a first order system but considers a flow variables specification instead. As a result, when applying Johansen's methodology, the estimates of the adjustment parameters show the cost of ignoring aggregation in the specification and leads to inappropriate conclusions. Finally, the third application considers also a flow variables specification but presents a second order system. As a result, it generalizes the analysis and further support for the use of our continuous time methodology is found.

It is important to notice, at the same time, that this document is not the first in contrasting models in continuous time with their discrete time counterpart; McCrorie and Chambers [2006], for example, in an application that analyses the money-income causality, state that formulating the model in continuous time offers a basis for correcting the effects of temporal aggregation in observed discrete data through a discrete time analogue. In a more general framework, Chambers and Thornton [2011], use a continuous-time autoregressive moving-average model (CARMA) to analyse sun-spot data and short-term interest rate. In the document, the authors develop an exact discrete time representation of the system under consideration and find out that the presence of a moving average component of order 1 (MA(1)) in the continuous time system has a dramatic impact on eradicating unaccounted-for serial correlation that is present in the discrete time model. In the non-stationary case, additionally, they only look at the situation where one of the roots of the characteristic equation of the system is identically equal to zero, but do not consider any cointegrated variables case.

This chapter, therefore, aims to extend the range of continuous time models that can be estimated using an exact discrete time representation by incorporating the non stationary cointegrated variables case into the existent literature.

The chapter is structured as follows: section 4.2 introduces the theoretical framework for the analysis as well as the comparison strategy. Then, section 4.3 presents the applications of our estimation methodology and section 4.4 concludes.

4.2 The Modelling Framework

In this section we specify the framework under which estimation and comparison is carried out. For that end, both the discrete and the continuous time systems are presented. For the continuous time analysis, we follow Bergstrom [1984] and rely on the results obtained in the previous chapters of this thesis, as a result, we consider a stochastic differential equation system and estimate an exact discrete time vector error correction model. For the discrete time analysis, at the same time, we consider Johansen's general VAR specification and also estimate a VECM.

4.2.1 Continuous time

The system under consideration is the continuous time random n-vector $y(t)$ that satisfies the stochastic differential equation system

$$\begin{aligned} d[D^{k-1}y(t)] = \Big[&A_{k-1}D^{k-1}y(t) + \cdots + A_1 Dy(t) + A_0 y(t) \\ &+a + bt\Big]dt + \zeta(dt), \qquad t > 0, \end{aligned} \tag{4.1}$$

where k is a positive integer larger or equal to 2 (for $k = 1$ the simplest version of the system is considered, see Chapter 2), $A_0, \ldots, A_{k-1}$ are $n \times n$ matrices of unknown coefficients and a and b are $n \times 1$ vectors of unknown constants. For our purpose, it is assumed that $y(t)$ is integrated of order one ($I(1)$) and that it is partitioned into two subvectors $y_1(t)$ and $y_2(t)$ of dimensions $n_1 \times 1$ and $n_2 \times 1$ respectively ($n_1 + n_2 = n$). It is also assumed that $y(t)$ contains n_1 stationary linear cointegrating relationships of the form $y_1(t) - B_1 y_2(t)$, where B_1 is a $n_1 \times n_2$ matrix of cointegrating parameters. $\zeta(dt)$, at the same time, is assumed to be a vector of random measures that is defined on all subsets of the line $0 < t < \infty$, has finite Lebesgue measure and satisfies

- $\mathrm{E}[\zeta(dt)] = 0$,
- $\mathrm{E}[\zeta(dt)\zeta(dt)'] = \Sigma dt$ and
- $\mathrm{E}[\zeta(\Delta_1)\zeta(\Delta_2)'] = 0$ for disjoint intervals Δ_1 and Δ_2.

In order to provide as much generality as possible in the applicability of our results, the vector $y(t)$ is allowed to be entirely comprised by stock or flow variables. If we define stock and flow variables in continuous time as $y^s(t)$ and $y^f(t)$, respectively, then, the observed values at specific points in time, for stock variables, are $y_t^s = y^s(t)$ and the observed rate of flows, for flow variables, $y_t^f = \int_{t-1}^t y^f(r)dr$, where, in each case $t = 1, 2, \ldots, T$ and T denotes sample size.

As shown in the previous chapters, the econometrically implementable models (the models that relate the unknown parameters of our system to the discrete time observations and are known as exact discrete time models) are written in a vector error correction form and, for stocks, are given by

$$\begin{aligned}\Delta y_t^s &= \Pi_k(\theta)y_{t-1}^s + \Gamma_1(\theta)\Delta y_{t-1}^s + \cdots + \Gamma_{k-1}(\theta)\Delta y_{t-(k-1)}^s + g_t^s(\theta) + \eta_t^s,\\ t &= k+1, \cdots, T,\end{aligned} \tag{4.2}$$

and

$$\begin{aligned}y_1^s &= G_1 x^s(0) + q_1^s + \eta_1^s,\\ \Delta y_t^s &= \mathit{\Lambda}_t(\theta)y_{t-1}^s + \Upsilon_1(\theta)\Delta y_{t-1}^s + \Upsilon_2(\theta)\Delta y_{t-2}^s +\\ &\cdots + \Upsilon_{t-2}(\theta)\Delta y_2^s + G_t x^s(0) + q_t^s(\theta) + \eta_t^s,\\ t &= 2, \cdots, k,\end{aligned} \tag{4.3}$$

where all the matrices and proofs are presented in lemmas 3.3.1 and 3.3.3 of Chapter 3 and θ is the vector that contains all the unknown parameter of the system (for $k = 1$, the simplest version of the system is considered, see Chapter 2 for details).

For flows, at the same time, the models are given by

$$\begin{aligned}\Delta y_t^f &= \Pi_k(\theta)y_{t-1}^f + \Gamma_1(\theta)\Delta y_{t-1}^f + \cdots + \Gamma_{k-1}(\theta)\Delta y_{t-(k-1)}^f + g_t^f(\theta) + \eta_t^f,\\ t &= k+1, \cdots, T,\end{aligned} \tag{4.4}$$

and

$$\begin{aligned}y_1^f &= Q_1 x^f(0) + q_1^f + \eta_1^f,\\ \Delta y_t^f &= \mathit{\Lambda}_t(\theta)y_{t-1}^f + \Upsilon_{h,1}(\theta)\Delta y_{t-1}^f + \Upsilon_{h,2}(\theta)\Delta y_{t-2}^f +\\ &\cdots + \Upsilon_{h,t-2}(\theta)\Delta y_2^f + Q_t x^f(0) + q_t^f(\theta) + \eta_t^f,\\ t &= 2, \cdots, k,\end{aligned} \tag{4.5}$$

where, similarly than before, all the matrices and proofs are developed in lemmas 3.3.2 and 3.3.4 of Chapter 4.

For the discrete time disturbances of these models, at the same time, we also follow the results of the previous chapters of this dissertation and, for the analysis, we utilize their moving average forms as well as their covariance properties as in Lemma 3.4.1 for η_t^s and Lemma 3.4.2 for η_t^f.

For estimation, finally, we use the methodologies as described in subsections 3.4.1 and 3.4.2 of Chapter 3 and optimize the alternative form of the likelihood function (equations (3.22) and (3.30)) to get the Gaussian estimates of our model's parameters

4.2.2 Discrete time

For the discrete time counterpart, we follow Johansen [1988, 1991, 1995] and consider a VAR with k lags

$$y_t = v + \delta t + \omega_1 y_{t-1} + \omega_2 y_{t-2} + \cdots + \omega_k y_{t-k} + \epsilon_t, \tag{4.6}$$

where y_t is a $n \times 1$ vector of variables, v and δ are $n \times 1$ vectors of parameters, $\omega_1 \cdots \omega_k$ are $n \times n$ matrices of parameters, and ϵ_t is a $n \times 1$ vector of disturbances. ϵ_t has mean 0, has covariance matrix Σ, and is *i.i.d.* normal over time.

The discrete time VECM form of (4.6) can be written as

$$\Delta y_t = v + \Phi y_{t-1} + \sum_{i=1}^{k-1} \Psi_i \Delta y_{t-i} + \delta t + \epsilon_t, \tag{4.7}$$

where $\Phi = \sum_{j=1}^{j=k} \omega_j - I_n$ and $\Psi_i = -\sum_{j=i+1}^{j=k} \omega_i$.

Similar to our continuous time specification, it is assumed that y_t is $I(1)$. Φ, therefore, has a reduced rank $n_1 < n$ $(n_1 + n_2 = n)$ and can be expressed as $\Phi = \alpha\beta'$, where α and β are $n \times n_1$ matrices of parameters with $rank(\alpha) = rank(\beta) = n_1$. The rows of $\beta' = [I, -\beta_1]$[12] form a basis for the n_1 cointegrating vectors and the elements of α distribute the impact of those cointegrating vectors

[12]In accordance with the notation used in (4.1) and given the specification of the systems, it is important to notice that $\beta_1 = B_1$.

to the evolution of Δy_t.

For an even more generalized form of the system, it is possible to exploit the properties of the matrix α and rewrite (4.7) as

$$\begin{aligned} \Delta y_t &= \alpha(\beta' y_{t-1} + \mu_0 + \rho t) + \sum_{i=1}^{k-1} \Psi_i \Delta y_{t-i} + \gamma + \tau t + \epsilon_t, \\ v &= \alpha\mu_0 + \gamma, \\ \delta t &= \alpha\rho t + \tau t, \end{aligned} \tag{4.8}$$

where μ_0 and ρ are $n_1 \times 1$ vectors of parameters and τ as well as γ are $n \times 1$ vectors of parameters with γ being orthogonal to $\alpha\mu_0$ ($\gamma'\alpha\mu_0 = 0$) and τ orthogonal to $\alpha\rho$ ($\tau'\alpha\rho = 0$).

Therefore, by applying restriction into the parameters, five different econometrically implementable models can be considered

Model I **unrestricted trend**

If no restrictions are placed on the trend parameters, our VECM implies that there are quadratic trends in the levels of the variables and that the cointegrating equations are stationary around time trends (trend stationary).

Model II **restricted trend**, $\tau = 0$

If $\tau = 0$ we assume that the trends in the levels of the data are linear but not quadratic. This specification allows the cointegrating equations to be trend stationary.

Model III **unrestricted constant**, $\tau = 0$ and $\rho = 0$

If $\tau = 0$ and $\rho = 0$ we exclude the possibility that the levels of the data have quadratic trends, and we restrict the cointegrating equations to be stationary around constant means. Because γ is not restricted to zero, this specification still puts a linear time trend in the levels of the data.

Model IV **restricted constant**, $\tau = 0, \rho = 0$ and $\gamma = 0$

If $\tau = 0, \rho = 0$ and $\gamma = 0$, we assume there are no linear time trends in the levels of the data. This specification allows the cointegrating equations to be stationary around a constant mean, but it allows no other trends or constant terms.

Model V **no trend**, $\tau = 0, \rho = 0, \gamma = 0$ and $\mu_0 = 0$

This specification assumes that there are no nonzero means or trends. It also assumes that the cointegrating equations are stationary with means of zero and that the differences and the levels of the data have means of zero.

Finally, for estimation, Likelihood Ratio (LR) tests are applied to the models so that the one that fits the data the best is considered. At the end, Johansen's maximum likelihood methodology is applied and the estimates of the model's parameters are obtained.

4.2.3 Comparison strategy

As mentioned in the previous chapters, comparison between the estimates of the two methodologies is not an easy task. If we look at equations (4.2), (4.4) and (4.8), we see that the models under consideration are more or less similar; they share a common VECM form and the cointegrating parameters can always be factored out, however, the estimates of the adjustment coefficients in our exact discrete time specification involve much more complicated expressions of the original parameters in the system than those of Johansen's and also, the representation of the discrete time disturbance processes changes dramatically with the specification from a MA process for the former to an $i.i.d$ process for the latter. As a result, a direct comparison of the estimates cannot be utilized and the implied values are required.

Considering that, for the comparison, firstly, we use discrete time tests to identify the specification of the model that fits the data the best; secondly, assuming that such model is correctly specified, we use both our continuous time methodology and Johansen's to estimate the model's parameters; and, finally, by equating term by term the elements of the econometrically implementable systems, we derive the implied values and compare them with one another.

Of course, if we suspect that the data under consideration presents some sort of temporal aggregation, Johansen's estimates will be biased and the implied values, as a result, are not going to be similar to our continuous time estimates.

Therefore, the cost of ignoring aggregation in the specification will be evident.

4.3 Empirical Applications

In this section we present the main results of this chapter. The first application considers stock variables and specifies the system in its simplest form, then, given the conclusions of Chapter 1, baseline findings are drawn. The second application, at the same time, considers flow variables and also specifies the system in its simplest form, as a result, direct implications of the cost of ignoring aggregation in the specification are expected. The third application, finally, deals with a higher order specification in the system and considers flow variables as well, therefore, given the conclusions of Chapter 2, further support for the use of our continuous time framework is anticipated.

4.3.1 Market efficiency and Cointegration

For this application, we extend the seminal work of Fama [1970], and follow Kühl [2007] to analyse the Efficiency Market Hypothesis (EMH) on the foreign exchange rate.

Our analysis focuses on a three-country model and we argue that the foreign exchange market is efficient if no cross-sectional arbitrage opportunities exist. In other words, if transaction costs are neglected, such a market is efficient if a specific amount of money in currency 1 retains its value, even if it is converted across the two other currencies.[13] Using the variables in logarithms, then, market efficiency means

$$s_t^{32} = s_t^{12} - s_t^{13}, \tag{4.9}$$

where s_t^{ij} are the exchange rates expressed in the same currency and i is the domestic currency in terms of the foreign currency j.

Equation (4.9) describes the so-called *no arbitrage* condition without transactions costs and states that a foreign exchange market is efficient provided that cointegration cannot be rejected if the cointegrating vector is $\beta' = (1, -1)$. Therefore, we can see that the EMH requires both cointegration and proportional cross-rate adjustments.

In this application, as mentioned earlier, in order to assess such a hypothesis,

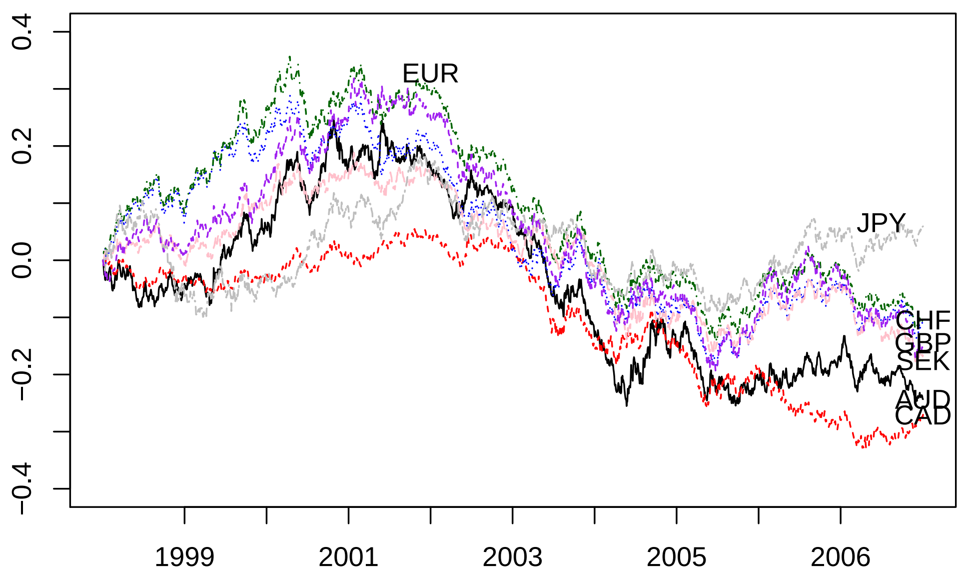
[13]See Frenkel and Levich [1975] and Levich [1985] for details.

Johansen's estimation procedure as well as likelihood ratio tests are applied to the data. At the same time, with the aim of analysing the effects, if there are any, of ignoring aggregation in the specification, our continuous time methodology is also applied to the same data. Next part describes precisely the analysis.

Empirical Results

The period under consideration runs from 4 January 1999 to 29 December 2006 and covers the daily exchange rates of the US-Dollar expressed in foreign currencies. We use the Australian Dollar (AUD), the Canadian Dollar (CAD), the Swiss Franc (CHF), the British Pound Sterling (GBP), the EURO (EUR), the Japanese Yen (JPY) and the Swedish Krona (SEK). [14] The exchange rates are taken from the database of the Federal Reserve Bank of St. Louis and are noon buying rates in New York City. Figure 4.1 depicts them.

Figure 4.1 Logarithm of Daily Exchange Rates (US-Dollar in Foreign Currencies)
standardised on 4 January 1999



Source: Federal Reserve Bank of St. Louis (2016)

As we can see from the figure, it seems to be a more or less parallel movement between some of the exchange rates. CHF, GBP and EUR, for instance, seem to co-move over time, therefore, they may support the EMH.

[14]The abbreviations refer to the USD expressed in units of foreign currency.

As it is well known, two or more time series are said to be cointegrated if each of them is $I(1)$ (nonstationary with a unit root) and there exist a linear combination that makes them stationary. As a result, before presenting the cointegrating analysis, we need to test the unit root properties of the series. To this end, two tests are applied to all exchange rates in natural logarithms: the Phillips-Perron (PP) test and the DF-GLS test by Elliott et al. [1996]. Results are described in Table 4.1.

As can be seen, both the PP and the DF-GLS test are not able to reject the null hypothesis of a unit root in levels in any of the exchange rates, as a result, they can be considered as nonstationary or $I(1)$ processes. Considering that, cointegration analysis can be applied.

Table 4.1 Unit Root Tests for the USD exchange rates

	PP test	DF-GLS test[a]	
	Statistic[b]	Lags[c]	Statistic[d]
EUR/USD	-0.571	4	-0.726
GBP/USD	-0.515	4	-0.598
JPY/USD	-2.137	1	-1.94
CHF/USD	-0.803	4	-0.907
AUD/USD	-0.56	4	-0.338
CAD/USD	-0.366	4	0.691
SEK/USD	-0.586	1	-0.740

[a] Stationarity around a mean is assumed.
[b] Critical values are 5 % -2.8, 1% -3.4.
[c] Number of lag is chosen by the modified AIC (MAIC).
[d] Critical values are 5 % -1.95, 1% -2.58.

For the cointegration test, Johansen's trace statistic is used. For our three-country model, we present the analysis for the USD-GBP-all-other-exchange-rates interactions. For the specification of the model, given the actual movement of the exchange rates over time (see Figure 4.1), we assume that a quadratic time trend in levels is negligible, as a result, only three possible specifications of the VECM (equation (4.8)) are considered. Table 4.2 presents the results.

Table 4.2 Cointegration Test using Johansen's approach

USD	Lags[a]	$H_o : r \leq 0, 1$	Model III Unrestricted Constant	Model IV Restricted Constant	Model V No Trend
GBP-EUR	1	0	24.83 **	25.38 **	2.815
		1	1.299	1.307	1.161
GBP-JPY	1	0	5.37	6.137	4.695
		1	0.086	0.837	0.089
GBP-CHF	1	0	19.04 *	19.597 *	10.6031
		1	3.03	3.0349	3.012
GBP-AUD	1	0	16.023 *	16.842 *	7.7246
		1	0.115	0.912	0.100
GBP-CAD	1	0	8.369	10.316	3.5108
		1	0.054	1.998	0.0006
GBP-SEK	1	0	16.481 *	17.021 *	7.2993
		1	0.442	0.705	0.5248

[a] Determination of lag length based on Schwarz bayesian information criterion (SBIC) and Lagrange Multiplier tests for autocorrelation in the residuals.

$*$ Rejection of the null hypothesis at 5 % .

$**$ Rejection of the null hypothesis at 1 % .

From Table 4.2, we can see that the null hypotheses of no cointegrating relationship can be rejected for model III as well as model IV for the exchange rate pairs GBP-EUR at 5 % percent level and GBP-CHF, GBP-AUD and GBP-SEK at 1 % percent level. For model V, on the contrary, for all exchange rate combinations, at least at a 5 % significance level, the null hypothesis cannot be rejected.

Considering that, there are two possible specifications of the model and, following Johansen and Juselius [1990] and Johansen [1994], a likelihood ratio test is utilized to choose the specification that fits the data the best. Results are reported in Table 4.3.

In the second column of Table 4.3, we test the absence of a linear trend in levels in the specification of the model and as we can see, in all currency pairs, the null hypothesis cannot be rejected. In the third column, at the same time, we test for the absence of a constant term in the cointegrating relationship and, for all pairs of currencies, we can see that a non-zero mean cannot be neglected.

Therefore, model IV represents the data better.

Table 4.3 Likelihood Ratio test on model specification of the VECM

	Absence of a linear trend in levels	Absence of a constant
	Model IV in Model III	Model V in Model IV
GBP -EUR	0.01	22.42 **
GBP-CHF	0.01	8.97 **
GBP-AUD	0.8	8.31 **
GBP-SEK	0.26	9.54 **

Critical values 3.84 for 5 % and 6.63 for 1%.

* Rejection of the null hypothesis at the 5 %.

** Rejection of the null hypothesis at the 1 %.

Table 4.4, lastly, presents the results of the evaluation of the market efficiency hypothesis as well as full estimates of the VECM. In the table, the first column presents the estimates of the constant term within the cointegrating relationship, the second presents them for the speed of adjustment parameters, the thrird for the cointegrating parameter and the forth, at the end, shows the likelihood ratio test for the restriction of the cointegrating vector.

As can be seen, all constant terms are statistically different from zero. All cointegrating parameters are also different from zero and more importantly, they present the correct sign for the no arbitrage condition. For the market efficiency hypothesis, finally, column IV validates it only for the pairs of currency GBP-CHF.

Table 4.4 Estimation of the VECM
Model IV: Restricted Constant

	μ_0	α	β	LR test (1,-1) [a]
GBP -EUR	0.059 **	-0.022 **	1	13.5 **
		-0.012 *	-0.712 **	(0.001)
GBP-CHF	0.032 **	-0.012 **	1	3.297
		-0.007	-0.794 **	(0.069)
GBP-AUD	-0.033 **	-0.011 **	1	9.312 **
		0.006	-0.674 **	(0.002)
GBP-SEK	0.029 **	-0.013 **	1	6.812 **
		-0.001	-0.759 **	(0.009)

[a] Test statistic for the hypothesis of a restricted cointegrated vector. P values in brackets.

* Rejection of the null hypothesis at 5 %.

** Rejection of the null hypothesis at 1 %.

For the continuous time framework, considering the results of Table 4.2 and Table 4.3, we define $y_1(t) = s^{12}(t)$ and $y_2(t) = s^{13}(t)$ as the exchange rates expressed in the same currency, where $s^{ij}(t)$ $(i = 1,\ j = 2, 3)$ is the domestic currency i in terms of the foreign currency j (note that t is being treated as a continuous time parameter), therefore, the continuous time system for this application can be written as

$$dy(t) = [c + A_0 y(t)]dt + \zeta(dt), \qquad t > 0, \tag{4.10}$$

where $y(t) = [y_1(t), y_2(t)]'$, $c = [c_1, c_2]'$ is a 2×1 vector of intercepts, $A_0 = ab'$, $a = [a_1, a_2]'$ is a 2×1 vector of adjustment parameters, $b = [1, -b_1]'$ is a 2×1 vector of cointegrating parameters and $\zeta(dt)$ is the vector of random measures that follows its definitions as in (4.1).

The exact discrete time system for stock variables, as a result, following (4.2) as well as the results from chapter 2, is defined as

$$\Delta y_t^s = c\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} + \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \frac{e^{(a_1 - b_1 a_2)} - 1}{a_1 - b_1 a_2}(1, -b_1)y_{t-1}^s + \eta_t^s, \qquad t = 1, \ldots, T, \tag{4.11}$$

where c as well as the disturbance vector η_t^s are

$$c = \int_0^1 \left(I_2 + \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \frac{e^{r(a_1 - b_1 a_2)} - 1}{a_1 - b_1 a_2} (1, -b_1) \right) dr,$$

$$\eta_t^s = \int_{t-1}^t \left(I_2 + \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \frac{e^{(t-s)(a_1 - b_1 a_2)} - 1}{a_1 - b_1 a_2} (1, -b_1) \right) \zeta(ds).$$

Considering equation (4.11), then, the estimates of our model's parameters are the elements of the vector θ that maximizes (see Chapter 2 for details) the following function

$$L(\theta, \Sigma) = -\frac{nT}{2} \ln(2\pi) - \frac{nT}{2} \ln|W| - \frac{1}{2} \sum_{t=1}^T \eta_t' W^{-1} \eta_t, \tag{4.12}$$

where $n = 2$ and $E[\eta_t \eta_t'] = W$.

For the comparison, at the same time, with the estimates already defined, we calculate the implied values by equating term by term (4.11) and model IV of equation (4.8) and get

$$\begin{gathered} \mu_0 \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = c \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}, \\ \alpha_1 = a_1 \frac{e^{(a_1 - b_1 a_2)} - 1}{a_1 - b_1 a_2}, \\ \alpha_2 = a_2 \frac{e^{(a_1 - b_1 a_2)} - 1}{a_1 - b_1 a_2}, \\ \beta_1 = b_1. \end{gathered} \tag{4.13}$$

Table 4.5 contains the results and for easier exposure, it is divided into three parts. The upper part presents the implied values[15], which, as mentioned above, are the continuous time equivalent to the discrete time values in our exchange rate models as summarized in Table 4.4. The middle part, additionally, shows the estimates we get by applying our continuous time methodology to those models and the lower part, more importantly, reports the likelihood ratio test statistics (and their associated p-value) and evaluates the efficiency market hypotheses in our application.

[15]It is important to stress that these values are for comparison purposes only and are obtained by solving the system of equations (4.13).

To gain further insights into these results, the table also contains estimates of the elements of the Cholesky matrix, M, corresponding to Σ so that

$$M = \begin{bmatrix} m_{11} & 0 \\ m_{21} & m_{22} \end{bmatrix}, \qquad \Sigma = \ MM', \qquad \Sigma = \begin{bmatrix} \sigma_1 & \sigma_2 \\ \sigma_2 & \sigma_3 \end{bmatrix},$$

$$\sigma_1 = m_{11}^2, \quad \sigma_2 = m_{11}m_{21}, \quad \sigma_3 = m_{21}^2 + m_{22}^2.$$

As we can see from the table, our estimates are virtually identical to the implied discrete time ones. All signs remain stable and the standard errors are reasonable small relative to their values. Also, all cointegrating parameters present the correct sign for the no arbitrage condition. Moreover, if we look at the lower part of the table, the efficiency market hypothesis in this continuous time specification is also validated only for the pair of currencies GBP-CHF.

Table 4.5 Implied Discrete Time Values and Continuous Time Model Estimates

		GBP-EUR		GBP-CHF		GBP-AUD		GBP-SEK	
Implied values	c_1	-0.0013		-0.0003		0.0003		-0.0003	
	c_2	-0.0006		-0.0002		-0.0001		-0.00002	
	a_1	-0.02		-0.01		-0.01		-0.01	
	a_2	-0.01		-0.007		0.006		-0.001	
		1		1		1		1	
	b_1	-0.712		-0.794		-0.674		-0.759	
CTM VECM Estimates	c_1	-0.0013	(0.00029)	-0.00038	(0.00014)	0.00030	(0.00017)	-0.00037	(0.00015)
	c_2	-0.0007	(0.00036)	-0.00022	(0.00019)	-0.00012	(0.00023)	-0.000017	(0.00020)
	a_1	-0.0217	(0.00502)	-0.01241	(0.00372)	-0.01149	(0.00402)	-0.01443	(0.00434)
	a_2	-0.0117	(0.00610)	-0.00759	(0.00508)	0.00586	(0.00526)	-0.00133	(0.00544)
		1	-	1	-	1	-	1	-
	b_1	-0.713	(0.03742)	-0.79097	(0.07931)	-0.66580	(0.05073)	-0.75642	(0.05236)
	m_{11}	0.0051	(0.00006)	0.00510	(0.00006)	0.00513	(0.00006)	0.00512	(0.00006)
	m_{21}	0.0043	(0.00011)	0.00461	(0.00012)	0.00294	(0.00014)	0.00410	(0.00012)
	m_{22}	0.0043	(0.00005)	0.00476	(0.00006)	0.00611	(0.00008)	0.00512	(0.00006)
LR test		16.41 **		3.26		6.82 **		6.15 *	
p-value		0.0001		0.070		0.001		0.013	

* Rejection of the null hypothesis at 5 %.

** Rejection of the null hypothesis at 1 %.

Numbers in parentheses denote standard errors.

As mentioned earlier, these results are not surprising and as expected, when the variables of interest are stocks and the systems under consider follow a VECM of order one, both our continuous time methodology as well as Johansen's perform similarly and yield virtually identical estimates of the model's parameters. Therefore, this particular application serves as the basis for our analysis and provides a benchmark reference.

4.3.2 The Term Structure of Interest Rates

Theoretical studies[16] of the modern term structure of interest rates suggest that there is an equilibrium relationship between interest rates at different maturities.

The segmented-market hypothesis, for example, states that investors have a preference for debt securities of a given term but that they are willing to substitute away from their preferred terms if they expect to be compensated for doing so through earning a risk or term premium. However, it does not specify whether the risk premium will be positive or negative and therefore, does not provide a closed form of the long-run relationship.

The expectation hypothesis, on the contrary, does not require risk premium to motivate an investor to mismatch his debt holding and planning horizon. Instead, it assumes that investors are rational and risk neutral, so that payment of the premium would not occur. Additionally, it assumes that transactions costs are zero, which means that the cost for investing in or buying an n-periods bond and holding it until maturity is the same as that of buying a series of one-period bonds. It follows, therefore, that long term rate can easily be represented as the present value of the expected future short rate.

Considering that, in this application, we utilize the expectation hypothesis to assess the long-run relationship between long and short term interest rates. For our purposes and given the present value model, we argue that such a relationship can be written as

$$\mathrm{L}_t - \beta_1 \mathrm{s}_t = \varepsilon_t, \tag{4.14}$$

where L_t represents the long term interest rate, s_t represents the short term interest rate, β_1 is the cointegrating parameter and ε_t is a random disturbance term.

If the expectation hypothesis holds, at the same time, the system is not only going to be structurally stable and show cointegration between the variables, but more importantly, it will embrace the proportional restriction $\beta' = (1, -1)$. In here, similarly than before, in order to assess such a hypothesis, a LR test is applied to the data.

[16] See for example Vasicek [1977] and Hall et al. [1992].

In applied work, especially in the one that involves macro time series, it is almost inevitable to face some sort of temporal aggregation (period averages, for example) in the data. If that is the case, as mentioned earlier, utilizing a model that is specified in terms of the observation interval can lead to estimates being contaminated by temporal aggregation bias and, as a consequence, it can arise owing to researchers making inappropriate economic interpretations of those parameter estimates. As a result, in order to avoid such complications, as pointed out by Bergstrom [1984], it is needed to formulate the econometric model for estimation in a continuous time fashion.

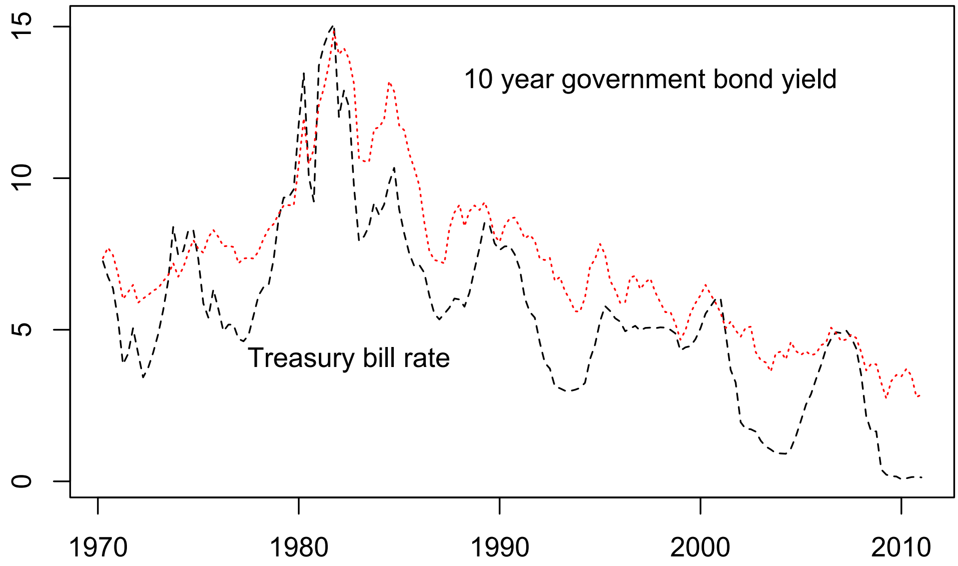
In this application, in order to illustrate such a scenario, rather than using interest rates as stock variables whose realisations are obtained daily on a fixed schedule, we aggregate them to quarterly average observations so that a flow variable is mimicked. Considering that, the long-run relationship (equation (4.14)) between long and short term interest rates as well as the expectations hypothesis are assessed utilizing both our continuous time methodology as well as Johansen's estimation procedure. At the end, comparison is presented and the costs, if there are any, of ignoring aggregation in the specification are measured. As in the previous application, before the actual comparison, discrete time analysis of the series are carried out. Next part describes precisely the procedure.

Empirical Results

The time series we use are the long-term government bond yield (10 years) and the treasury bill rate (government securities 4 weeks coupon equivalent) from 1970:Q1 to 2010:Q4. They are taken from the Federal Reserve Bank of St. Louis and the Department of the Treasury for the United States and, as mentioned earlier, they are aggregated to quarterly average from daily observations.[17] Figure 4.2 displays the series.

[17]The daily treasury bill rate is the daily secondary market quotation on the most recently auctioned treasury bill for the 4 weeks maturity tranche. The quotation is obtained at a closing hour each business day by the Federal Reserve Bank of New York. The daily long-term government bond is the unweighted average of bid yields obtained at a closing hour each business day on all outstanding fixed-coupon bonds neither due nor callable in less than 10 years.

Figure 4.2 Long-term and Short-term interest rates
Quarterly average from daily observations, U.S.



Source: Federal Reserve Bank of St. Louis (2016).

The statistical properties of the series as well as the specification tests are presented in Table 4.6. As can be seen, both the PP and the DF-GLS test are not able to reject the null hypothesis of a unit root in levels in either of the interest rates, as a result, they can be considered as nonstationary or $I(1)$ processes.

Table 4.6 Statistical tests and Specification analysis for the Interest Rates

Unit Root Tests

Levels	PP test	DF-GLS test[a]	
	Statistic[b]	Lags[c]	Statistic[d]
Short Term	-1.663	7	-0.960
Long Term	-1.099	7	-0.555

[a] Stationarity around a mean is assumed.

[b] Critical values are 5 % -2.86, 1% -3.43.

[c] Number of lag is chosen by the modified AIC (MAIC).

[d] Critical values are 5 % -1.95, 1% -2.58.

Cointegration Test using Johansen's approach

Lags[a]	$H_o : r \leq 0, 1$	Model III	Model IV	Model V
		Unrestricted Constant	Restricted Constant	No Trend
1	0	17.1154 *	17.6613	13.8356 *
	1	1.1412	1.6679	1.1

[a] Determination of lag length based on Schwarz bayesian information criterion (SBIC) and Lagrange Multiplier tests for autocorrelation in the residuals.

$*$ Rejection of the null hypothesis at 5 % .

$**$ Rejection of the null hypothesis at 1 % .

Likelihood Ratio test on model specification of the VECM

Absence of a constant and a linear trend in levels Model V in Model III
3.78

Critical values 3.84 for 5 % and 6.63 for 1%.

For the cointegration test, we use Johansen's trace statistic and assume that a quadratic time trend in levels is negligible, therefore, only three possible specifications of the VECM are considered. From the table, it can be seen that a first order specification is confirmed and the null hypothesis of no cointegrating relationship can be rejected for model III as well as model V at 5 % significance level. Using a likelihood ratio test to choose the specification of the model that fits the data the best, it can also be seen that the absence of constant term in the cointegrating relationship as well as a linear trend in levels in the system cannot be rejected. As a result, model V represents the data better.

With the confirmation of a cointegrating relationship in the model, we turn the analysis to fully estimate the VECM and assess the effects of short term rates on long term rates. Table 4.7 presents the results.

As can be seen, the cointegrating parameter is statistically different from zero and displays the correct sign, as a result, it can be said that the central bank can influence long-term rates by operating at the short end of the market. Looking at the LR test statistics, at the same time, we confirm that the expectations hypothesis is not consistent with the United States. For the speed of adjustment parameters, finally, it can also be seen that only one of them is statistically different from zero, however, the signs alternate, as a result, if one of the variables deviates from the long run relationship, an adjustment will always occur and the convergence rate is actually specified by the coefficients themselves.

Table 4.7 Estimation of the VECM
Model V: No Trend

α	β	LR test (1,-1)[a]
-0.042 *	1	6.16 **
0.036	-1.241 **	(0.013)

[a] Test statistic for the hypothesis of a restricted cointegrated vector. P values in brackets.

* Rejection of the null hypothesis at 5 %.

** Rejection of the null hypothesis at 1 %.

For the continuous time framework, considering the results of Table 4.6, we define $y_1(t) = L(t)$ and $y_2(t) = s(t)$ as the long and short term interest rates, respectively (note that t is being treated as a continuous time parameter). Therefore, the continuous time system for this application can be written as

$$dy(t) = A_0 y(t)dt + \zeta(dt), \qquad t > 0, \tag{4.15}$$

where $y(t) = [y_1(t), y_2(t)]'$, $A_0 = ab'$, $a = [a_1, a_2]'$ is a 2×1 vector of adjustment parameters, $b = [1, -b_1]'$ is a 2×1 vector of cointegrating parameters and $\zeta(dt)$ is the vector of random measures that follows its definitions as in (4.1).

The exact discrete time model for flow variables, as a result, following (4.4) as well as the results of chapter 2, is defined as

$$\Delta y_t^f = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \frac{e^{(a_1 - b_1 a_2)} - 1}{a_1 - b_1 a_2}(1, -b_1) y_{t-1}^f + \eta_t^f, \qquad t = 2, \ldots, T, \tag{4.16}$$

$$\eta_t^f = \int_{t-1}^{t} \int_{r-1}^{r} \left(I_2 + \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \frac{e^{(r-s)(a_1 - b_1 a_2)} - 1}{a_1 - b_1 a_2}(1, -b_1) \right) \zeta(ds) dr.$$

For the observed vector y_1^f, at the same time, we have

$$y_1^f = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \frac{1}{a_1 - b_1 a_2} \left(\int_0^1 (e^{r(a_1 - b_1 a_2)} - 1) dr \right) (1, -b_1) y(0) + \eta_1^f, \tag{4.17}$$

where $y(0)$ is the boundary condition of the data and

$$\eta_1^f = \int_0^1 \int_0^r \left(I_2 + \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \frac{e^{(r-s)(a_1 - b_1 a_2)} - 1}{a_1 - b_1 a_2}(1, -b_1) \right) \zeta(ds) dr.$$

Then, the estimates of our model's parameters are the elements of θ that optimize

$$\mathcal{L}(\theta, \Sigma) = \sum_{i=1}^{nT} (\varepsilon_i^2 + 2 \log p_{ii}), \tag{4.18}$$

where $n = 2$, θ denotes de vector of unknown parameters to be estimated, p_{ii} is the ith diagonal element of P, P is a real lower triangular matrix, with positive elements along the diagonal. $PP' = \Omega^f$, $E[\eta^f (\eta^f)'] = \Omega^f$, $\eta^f = [(\eta_1^f)', (\eta_2^f)', \cdots, (\eta_T^f)']'$, η_t^f follows a moving average process of order one and the nT elements of ε are computed in T vectors of size n using recursively the following procedure

$$\begin{aligned} \varepsilon_1 &= \left(\varepsilon_{11}, \varepsilon_{12} \right)' = P_{11}^{-1} \eta_1^f, \\ \varepsilon_t &= \left(\varepsilon_{t1}, \varepsilon_{t2} \right)' = P_{tt}^{-1} (\eta_t^f - P_{t,t-1} \varepsilon_{t-1}), \qquad t = 2, 3, \cdots, T. \end{aligned} \tag{4.19}$$

For the comparison, finally, we follow (4.13) and similarly calculate the implied values. Results are presented in Table 4.8 and the exposure follows that of Table 4.5.

As can be seen, the continuous time cointegrating parameter is almost identical to its discrete time counterpart. Looking at the LR test, at the same, our continuous time specification also rejects the expectations hypothesis for the United States. For the adjustment parameters, finally, although the signs remain, we see that bias is evident.

Considering that, some conclusions can be drawn. First, no matter what the time specification is, the long run relationship between the variables will always be identically measured by the cointegrating parameter and second, if it is suspected that the data contains some sort of temporal aggregation and Johansen's methodology is used, the estimates of the adjustment parameters will suffer from temporal aggregation bias and will lead to inaccurate conclusions. Therefore, in these particular cases, in order to make appropriate economic decisions, our continuous time methodology has to be used.

Table 4.8 Implied Discrete Time Values and Continuous Time Model Estimates

Implied values	a_1	-0.0429	
	a_2	0.0367	
	b_1	-1.241	
CTM VECM Estimates	a_1	-0.030	(0.0304)
	a_2	0.081	(0.0456)
		1	
	b_1	-1.242	(0.0836)
	m_{11}	0.645	(0.0357)
	m_{21}	0.637	(0.0732)
	m_{22}	0.784	(0.0464)
LR test		6.82 **	
p-value		0.009	

* Rejection of the null hypothesis at 5 %.
** Rejection of the null hypothesis at 1 %.
Numbers in parentheses denote standard errors.

4.3.3 Permanent Income Hypothesis

For this application, we extend the seminal propositions of Friedman [1957], Hall [1978], and Flavin [1981], and follow Campbell [1987] to evaluate, for the United States, the main implication of the rational expectations-permanent income hypothesis (PIH), that is to say, in this application, we evaluate the long run relationship between consumption and total disposable income.

In the analysis, as in Campbell [1987], we consider that such a relationship is given by the optimal path of consumption of the infinitely lived representative consumer of the economy, as a result, we write it as

$$C_t - \beta_1 Y_t^p = \varepsilon_t, \tag{4.20}$$

where $\beta_1 \leq 1$ is the propensity to consume, C_t is consumption, ε_t is the random disturbance term at time t and Y_t^p, at the same time, is known as the permanent income and is defined as

$$Y_t^p \equiv r \left[W_t + \left(\frac{1}{1+r} \right) \sum_{i=0}^{\infty} \left(\frac{1}{1+r} \right)^i \mathrm{E}_t Y_{t+i} \right],$$

where E_t is the expectations operator, W_t is non-human wealth, r is the constant rate of return of this non-human wealth and Y_t is the income at time t.

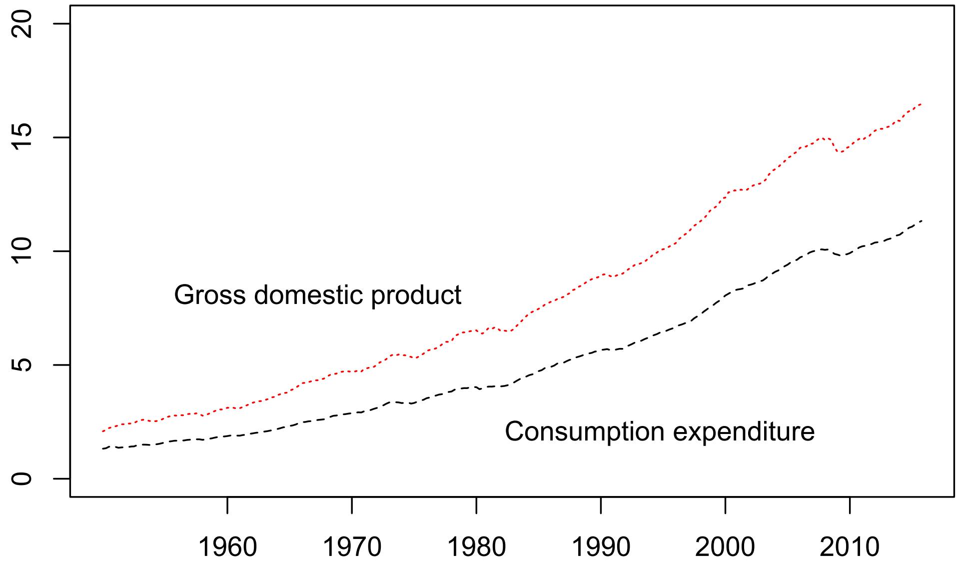
Looking at equation (4.20), if cointegration exists, we see that consumption will never deviate more than a certain fraction (given by the propensity to consume) of the expected present value of the future income. Moreover, if the rational expectations hypothesis holds, the cointegrating vector must be $\beta' = (1, -1)$, therefore, white noise deviations aside, consumption must proportionally adjust to permanent income. In here, as before, in order to test such hypothesis, Johansen's cointegration procedure as well as LR tests are applied to the data.

As in the previous applications, with the aim of analysing the effects, if there are any, of ignoring aggregation in the specification, we also apply our continuous time methodology to the data and assess the long run relationship between consumption and total disposable income. Of course, given that consumption as well as income are better conceived as flow variables, in our model, the flow variables specification is considered. Next part describes precisely the methodology.

Empirical Results

Our data comes from the Federal Reserve Bank of St. Louis and consists of Real Gross Domestic Product and Real Household Consumption Expenditure at 2009 prices for the United States. The observations are quarterly measures of the variables and the period runs from 1950:Q1 to 2015:Q4. Figure 4.3 depicts the series.

Figure 4.3 Gross Domestic Product and Consumption Expenditure at 2009 prices
Trillions US dollars



Source: Federal Reserve Bank of St. Louis (2016).

As we can see, the co-movement between the series seems to exist, therefore, the long run relationship may occur. For the cointegration analysis, Table 4.9 presents the statistical properties of the series as well as the tests for cointegration. As can be seen, the unit root tests are not able to reject the null hypothesis of a unit root in levels for any of the series, therefore, they are non stationary processes.

Table 4.9 Statistical tests and Specification analysis
Permanent Income Hypothesis

Unit Root Tests

Levels	PP test	DF-GLS test[a]	
	Statistic[b]	Lags[c]	Statistic[d]
GDP	2.656	2	-0.464
Consumption	4.199	3	-0.376

[a] Stationarity around a linear time trend is assumed.

[b] Critical values are 5 % -2.86, 1% -3.43.

[c] Number of lag is chosen by the modified AIC (MAIC).

[d] Critical values are 5 % -2.86, 1% -3.43.

Cointegration Test using Johansen's approach

Lags[a]	$H_o : r \leq 0, 1$	Model III	Model IV	Model V
		Unrestricted Constant	Restricted Constant	No Trend
2	0	35.515***	75.928**	67.528**
	1	0.316	5.581	0.4403

[a] Determination of lag length based on Schwarz bayesian information criterion (SBIC) and Lagrange Multiplier tests for autocorrelation in the residuals.

$*$ Rejection of the null hypothesis at 5 % .

$**$ Rejection of the null hypothesis at 1 % .

Likelihood Ratio test on model specification of the VECM

Absence of a constant and a linear trend in levels Model V in Model III
8.52**

Critical values 3.84 for 5 % and 6.63 for 1%.

* Rejection of the null hypothesis at the 5%.

** Rejection of the null hypothesis at the 1 %.

Turning over to cointegration, it can also be seen that a second order VECM specification is confirmed and the null hypothesis of no cointegrating relationship can be rejected for model IV as well as model V at 1 % significance level. From the LR test, it can be seen that a constant term in the cointegrating relationship as well as a linear trend in levels cannot be neglected. As a result, model III fits the data better.

Full estimates of the VECM as well as the implications of the PIH for the United States are presented in Table 4.10. As can be seen, the cointegrating parameter is statistically different from zero and displays the correct sign. For the speed of adjustment coefficients and the intercepts, on the contrary, we see that not all of them are statistically significant and on top of that, some are also showing the wrong sign. Looking at the LR test statistic, at the same time, we see that the standard rational expectations hypothesis is not consistent with the United States.

Table 4.10 Estimation of the VECM
Model III: Unrestricted Constant

μ_0	α	β	γ	Ψ_1	LR test (1,-1)[a]
				0.2591 **	
0.1838	-0.0399 **	1	-0.0006	0.0921 *	7.082 *
	-0.001	-0.7442 **	0.0163*	0.9221 **	(0.007)
				0.0433	

[a] Test statistic for the hypothesis of a restricted cointegrated vector. P vales im brackets.

* Rejection of the null hypothesis at 5 %.

** Rejection of the null hypothesis at 1 %.

For the continuous time specification, considering the results of Table 4.9, we define, respectively, $y_1(t) = C(t)$ and $y_2(t) = Y^p(t)$ as consumption and total income, therefore, the continuous time system for this application can be written as

$$d[Dy(t)] = [A_1 Dy(t) + A_0 y(t) + c]dt + \zeta(dt), \quad t > 0, \tag{4.21}$$

where $y(t) = [y_1(t), y_2(t)]'$, A_1 is a $2{\times}2$ matrix of parameters, $A_0 = ab'$, $a = [a_1, a_2]'$ is a 2×1 vector of adjustment parameters, $b = [1, -b_1]'$ is a 2×1 vector of cointegrating parameters, $c = [c_1, c_2]'$ is the vector of intercepts and $\zeta(dt)$ is the vector of random measures that follows its definitions as in (4.1). In here, it is important to notice that a second order specification of the system is considered because Johansen's test confirmed a second order VECM.

The exact discrete time system for flow variables, as a result, following (4.4), is defined as

$$\Delta y_t^f = \Pi(\theta) y_{t-1}^f + \Gamma_1(\theta)\Delta y_{t-1}^f + g_t^f(\theta) + \eta_t^f, \qquad t = 3, \cdots, T, \tag{4.22}$$

where θ is the vector of unknown parameters of the system comprised in A_1, A_0 and c. Also, η_t^f is the discrete time disturbance vector that follows a moving average process of order 2 (see Chapter 3 Lemma 3.3.2 for details), $\Pi(\theta) = F_1 + F_2 - I = k(\theta)b'$ and $\Gamma_1(\theta) = -F_2$, with

$$\begin{aligned}
F_1 &= S_1 e^A S_1' + (S_1 e^A S_2')(S_2 e^A S_2')(S_1 e^A S_2')^{-1},\\
F_2 &= -(S_1 e^A S_2')(S_2 e^A S_2')(S_1 e^A S_2')^{-1}(S_1 e^A S_1') + (S_1 e^A S_2')(S_2 e^A S_1'),\\
S_1 &= \begin{pmatrix} I_2 & 0_2 \end{pmatrix}, & S_2 &= \begin{pmatrix} 0_2 & I_2 \end{pmatrix},\\
e^A &= \Big(I_4 + UH^{-1}(e^H - I_3)V'\Big), & C_{ij} &= S_i e^A S_j' \;\; i,j = 1,2,\\
k(\theta) &= (k_1, k_2)', & b' &= (1, -b_1),
\end{aligned}$$

$$g_t^f(\theta) = S_1\Big(\int_0^1 e^{Ar} ds\Big)c^* - \Big(C_{12}C_{22}C_{12}^{-1}S_1 - C_{12}S_2\Big)S_1\Big(\int_0^1 e^{Ar} ds\Big)c^*$$

$$A = UV' = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ a_1 & -a_1 b_1 & x & z \\ a_2 & -a_2 b_1 & r & w \end{pmatrix}, \qquad H = V'U = \begin{pmatrix} 0 & 1 & -b_1 \\ a_1 & x & z \\ a_2 & r & w \end{pmatrix},$$

$$V' = \begin{pmatrix} 1 & -b_1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad U = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ a_1 & x & z \\ a_2 & r & w \end{pmatrix}, \qquad c^* = \begin{pmatrix} 0 \\ 0 \\ c_1 \\ c_2 \end{pmatrix}.$$

The complementary equations for $t = 1, 2$, at the same time, are given by

$$\begin{aligned}
y_1^f &= Q_1 x^f(0) + g_1^f(\theta) + \eta_1^f,\\
\Delta y_2^f &= \Lambda_1(\theta) y_1^f + Q_2 x^f(0) + g_2^f(\theta) + \eta_2^f,
\end{aligned} \tag{4.23}$$

where $x^f(0)$ is the vector of boundary conditions and

$$\Lambda_1 = (S_1 e^A S_1') - I_2 \qquad Q_1 = S_1 \int_0^1 e^{rA} dr, \qquad Q_2 = (S_1 e^A S_2') S_2 \int_0^1 e^{rA} dr,$$

$$g_1^f(\theta) = S_1 \Big(\int_0^1 \int_0^r e^{(r-s)A} ds dr \Big) c^*,$$

$$g_2^f(\theta) = S_1 \Big(\int_0^1 e^{Ar} ds \Big) c^* + C_{12} S_2 \Big(\int_0^1 \int_0^r e^{(r-s)A} ds dr \Big) c^*.$$

Then, the estimates of our model's parameters are the elements of θ that optimize

$$\mathcal{L}^f(\theta, \Sigma) = \sum_{i=1}^{nT} (\varepsilon_i^2 + 2 \log p_{ii}), \tag{4.24}$$

where $n = 2$, p_{ii} is the ith diagonal element of P, P is a real lower triangular matrix, with positive elements along the diagonal, such that $PP' = \Omega^f$, $E[(\eta^f)(\eta^f)'] = \Omega^f$, $\eta^f = [(\eta_1^f)', (\eta_2^f)', \cdots, (\eta_T^f)']$ and the nT elements ε are computed in T vectors of size n using recursively the following procedure

$$\begin{aligned} \varepsilon_1 &= \Big(\varepsilon_{11}, \varepsilon_{12}\Big)' = P_{11}^{-1} \eta_1^f, \\ \varepsilon_2 &= \Big(\varepsilon_{21}, \varepsilon_{22}\Big)' = P_{22}^{-1} (\eta_2^f - P_{2,1} \varepsilon_1), \\ \varepsilon_t &= \Big(\varepsilon_{t1}, \varepsilon_{t2}\Big)' = P_{tt}^{-1} (\eta_t^f - P_{t,t-1} \varepsilon_{t-1} - P_{t,t-2} \varepsilon_{t-2}), \qquad t = 3, \cdots, T. \end{aligned} \tag{4.25}$$

For the comparison, finally, we follow (4.13) and similarly calculate the implied values of our system's parameters by equating term by term (4.22), (4.23) and model III of equation (4.8), therefore

$$\begin{aligned} \alpha\mu_0 + \gamma &= g(\theta), & \beta_1 &= b_1, \\ \alpha_1 &= k_1(\theta), & \Psi_1 &= \Gamma_1(\theta), \\ \alpha_2 &= k_2(\theta). \end{aligned} \tag{4.26}$$

Results are presented in Table 4.11 and, as before, exposure follows that of Table 4.5. As can be seen, both the adjustment coefficients and the cointegrating parameter maintain the sign, however, the magnitude of the former changes with the specification. In here, it is important to notice that the parameter that differs the most from its implied value (a_2) is also the one that represents the estimate that is not statistically significant in the discrete time analysis (see α_2 in Table 4.10).

At the same time, if we look at the estimates of the intercepts (c_1 and c_2), a similar picture emerges; the magnitude changes with the specification and the parameter that switches the sign is also the one that represents the estimate that is not statistically significant in the discrete time analysis (see γ_1 in Table 4.10).

Finally, looking at the LR test, we see that our continuous time specification also rejects the rational expectations hypothesis for the United States.

As a result, it can be pointed out that estimation bias in cointegrated systems does not only depend on whether the variables in the model suffer some sort of temporal aggregation, but also, on whether the system requires a higher order specification. Therefore, as mentioned earlier, in any of those cases and in order to make appropriate conclusions of the model, our continuous time methodology is recommended.

Table 4.11 Implied Discrete Time Values and Continuous Time Model Estimates

	c_1	-0.0678				
	c_2	0.0712		Γ_1	-3.9222	
Implied values	a_1	0.1412		Γ_2	6.8484	
	a_2	0.3173		Γ_3	9.1451	
				Γ_4	-5.0536	
	b_1	-0.74				
	c_1	0.0457	(0.0209)			
	c_2	0.3805	(0.0578)	Γ_1	3.5187	(0.0167)
	a_1	0.1559	(0.0187)	Γ_2	9.4991	(0.0212)
	a_2	0.1306	(0.0113)	Γ_3	-5.4225	(0.0099)
CTM VECM Estimates		1		Γ_4	-4.6069	(0.0223)
	b_1	-0.759	(0.0082)			
	m_{11}	0.7513	(0.0173)			
	m_{12}	0.0304	(0.0133)			
	m_{22}	0.1778	(0.0081)			
LR test		12.005 **				
p-value		0.0005				

* Rejection of the null hypothesis at 5 %.

** Rejection of the null hypothesis at 1 %.

Numbers in parentheses denote standard errors.

4.4 Concluding remarks

With the aim of analysing the effects of temporal aggregation over the estimates of a model's parameters, focusing on the non stationary cointegrated case, this document has presented three multivariate applications of the estimation methodology for systems in continuous time developed in the previous chapters.

For the analysis, Johansen's general VECM as well as our exact discrete time VECM have been used. For the comparison, given that the representation of the estimates differs dramatically with the specification, the estimated parameters of the two specifications have been contrasted with one another through the use of the implied values. In there, discrete time test were used to identify the specification of the model that fits the data the best and such model was assumed to be correct.

The first application (**market efficiency and cointegration**) considered a first order system and a stock variables specification. In the results, as expected, given the simple specification and the fact that there was no temporal aggregation in the data, both our continuous time methodology as well as Johansen's produced virtually identical estimates.

The second application (**the term structure of the interest rate**), at the same time, presented a first order system but considered a flow variables specification instead, as a result, when applying Johansen's methodology, the estimates of the adjustment parameters showed the cost of ignoring aggregation in the specification and led to inappropriate conclusions.

The third application (**the permanent income hypothesis**), finally, considered also a flow variables specification but presented a second order system, as a result, it generalized the analysis. In the results, further support for the use of our continuous time methodology was found. In there, not only the adjustment parameters and the intercepts showed bias, but more importantly those who differed the most from their implied values were also the ones that represented the estimates that were not statistically significant in the discrete time analysis (see Table 4.11 together with Table 4.10).

Considering that, it has been concluded that estimation bias in cointegrated systems does not only depend on whether the variables in the model suffer some sort of temporal aggregation, but also, on whether the system requires a higher order specification. Therefore, as it was shown, in any of those cases and in order to make appropriate conclusions of the model, our continuous time methodology is recommended.

For further research, considering the results reported here, there are a number of directions that emerge. Perhaps the most obvious extension is to consider an exact discrete time representation for mixed sample. For that area, some progress has already being done, Chambers [2009], for example, has presented mixed sample first order cointegrated systems' analysis and although his specification varies from the one analysed here, he has settled the basis. Therefore, a generalization of that result is a natural extension and will be explored in future work.

References

Al-Mohy, A. H. and N. J. Higham (2009). A new scaling and squaring algorithm for the matrix exponential. *SIAM Journal on Matrix Analysis and Applications 31*, 970–989.

Bergstrom, A. and K. Nowman (2007). *A Continuous Time Econometric Model of the United Kingdom with Stochastic Trends*. Cambridge University Press.

Bergstrom, A. R. (1966). Nonrecursive models as discrete approximations to systems of stochastic differential equations. *Econometrica 1*, 173–182.

Bergstrom, A. R. (1976). *Statistical Inference in Continuous Time Economic Models*. North-Holland, Amsterdam.

Bergstrom, A. R. (1983). Gaussian estimation of structural parameters in higher order continuous time dynamic models. *Econometrica 51*, 117–152.

Bergstrom, A. R. (1984). Continuous time stochastic models and issues of aggregation over time. In Z. Griliches and M. D. Intriligator (Eds.), *Handbook of Econometrics*, Volume 2, pp. 1145–1212. North-Holland, Amsterdam.

Bergstrom, A. R. (1985). The estimation of parameters in nonstationary higher order continuous-time dynamic models. *Econometric Theory 1*, 369–385.

Bergstrom, A. R. (1986). The estimation of open higher-order continuous time dynamic models with mixed stock and flow data. *Econometric Theory 2*, 350–373.

Bergstrom, A. R. (1990). *Continuous Time Econometric Modelling*. Oxford University Press.

Bergstrom, A. R. (1996). Survey of continuous-time econometrics. In *Dynamic Disequilibrium Modeling: Theory and Applications: Proceedings of the Ninth International Symposium in Economic Theory and Econometrics*, Volume 9, pp. 1. Cambridge University Press.

Bergstrom, A. R. (1997). Gaussian estimation of mixed-order continuous-time dynamic models with unobservable stochastic trends from mixed stock and flow data. *Econometric Theory 13*, 467–505.

Campbell, J. Y. (1987). Does saving anticipate declining labor income? an alternative test of the permanent income hypothesis. *Econometrica 1*, 1249–1273.

Chambers, M. J. (1999). Discrete time representation of stationary and non-stationary continuous time systems. *Journal of Economic Dynamics and Control 23*, 619–639.

Chambers, M. J. (2009). Discrete time representations of cointegrated continuous time models with mixed sample data. *Econometric Theory 25*, 1030–1049.

Chambers, M. J. and R. J. McCrorie (2007). Frequency domain estimation of temporally aggregated gaussian cointegrated systems. *Journal of Econometrics 136*, 1–29.

Chambers, M. J. and M. A. Thornton (2011). Discrete time representation of continuous time arma processes. *Econometric Theory 28*, 219–2038.

Donaghy, K. P. (1993). A continuous-time model of the united states economy. In G. Gandolfo (Ed.), *Continuous-Time Econometrics*, pp. 151–193. Chapman and Hall, London.

Doob, J. L. (1953). *Stochastic Processes*. John Wiley, New York.

Elliott, G., T. J. Rothenberg, and J. H. Stock (1996). Efficient tests for an autoregressive unit root. *Econometrica 64*, 813–836.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance 25*, 383–417.

Flavin, M. A. (1981). The adjustment of consumption to changing expectations about future income. *Journal of Political Economy 89*, 974–1009.

Frenkel, J. A. and R. M. Levich (1975). Covered interest arbitrage: Unexploited profits? *Journal of Political Economy 83*, 325–338.

Friedman, M. (1957). *A Theory of the Consumption Function.* Princeton, NJ: Princeton University Press.

Gandolfo, G. (1993). *Continuous-Time Econometrics: Theory and Applications.* Chapman and Hall, London.

Hall, A. D., H. M. Anderson, and C. W. Granger (1992). A cointegration analysis of treasury bill yields. *The Review of Economics and Statistics 1*, 116–126.

Hall, R. E. (1978). Stochastic implications of the life cycle-permanent income hypothesis: Theory and evidence. *Journal of Political Economy 86*, 971–87.

Harvey, A. C. and J. H. Stock (1985). The estimation of higher-order continuous time autoregressive models. *Econometric Theory 1*, 97–117.

Harvey, A. C. and J. H. Stock (1988). Continuous time autoregressive models with common stochastic trends. *Journal of Economic Dynamics and Control 12*, 365–384.

Jewitt, G. and J. R. McCrorie (2005). Computing estimates of continuous time macroeconometric models on the basis of discrete data. *Computational Statistics and Data Analysis 49*, 397 – 416.

Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control 12*, 231–254.

Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models. *Econometrica 59*, 1551–1580.

Johansen, S. (1994). The role of the constant and linear terms in cointegration analysis of nonstationary variables. *Econometric Reviews 13*, 205–229.

Johansen, S. (1995). *Likelihood-based Inference in Cointegrated Vector Autoregressive Models.* Oxford University Press on Demand.

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

Kühl, M. (2007). Cointegration in the foreign exchange market and market efficiency since the introduction of the euro: Evidence based on bivariate cointegration analyses. Technical report, CEGE Discussion Paper.

Levich, R. M. (1985). Empirical studies of exchange rates: Price behavior, rate determination and market efficiency. *Handbook of International Economics 2*, 979–1040.

McCrorie, J. R. (2000). Deriving the exact discrete analog of a continuous time system. *Econometric Theory 16*, 998–1015.

McCrorie, J. R. and M. J. Chambers (2006). Granger causality and the sampling of economic processes. *Journal of Econometrics 132*, 311–336.

Moler, C. and C. Van Loan (1978). Nineteen dubious ways to compute the exponential of a matrix. *SIAM Review 20*, 801–836.

Nowman, K. B. (1997). Gaussian estimation of single-factor continuous time models of the term structure of interest rates. *The Journal of Finance 52*, 1695–1706.

Nowman, K. B. (2006). Continuous time interest rate models in japanese fixed income markets. In J. A. Batten et al. (Eds.), *Japanese Fixed Income Markets: Money, Bond and Interest Rate Derivatives*, pp. 321–346. Emerald Group Publishing Limited.

Phillips, A. W. (1959). The estimation of parameters in systems of stochastic differential equations. *Biometrika 46*, 67–76.

Phillips, P. C. B. (1972). The structural estimation of a stochastic differential equation system. *Econometrica 40*, 1021–1041.

Phillips, P. C. B. (1991). Error correction and long-run equilibrium in continuous time. *Econometrica 59*, 967–980.

Robinson, P. M. (1976a). The estimation of linear differential equations with constant coefficients. *Econometrica 44*, 751–764.

Robinson, P. M. (1976b). Fourier estimation of continuous time models. In A. R. Bergstrom (Ed.), *Statistical Inference in Continuous Time Economic Models*, pp. 215–226. North Holland, Amsterdam.

Rozanov, U. A. (1967). *Stationary Random Processes.* Holden-Day, New York.

Van Loan, C. (1978). Computing integrals involving the matrix exponential. *IEEE Transactions on Automatic Control 23*, 395–404.

Vasicek, O. (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics 5*, 177–188.

Ward, R. C. (1977). Numerical computation of the matrix exponential with accuracy estimate. *SIAM Journal on Numerical Analysis 14*, 600–610.

Wymer, C. R. (1993). Continuous-time models in macroeconomics: specification and estimation. *Continuous Time Econometrics: Theory and Applications 1*, 35–79.

Yu, J. and P. C. Phillips (2001). A gaussian approach for continuous time models of the short-term interest rate. *The Econometrics Journal 4*, 210–224.