Essays on non-linear aggregation in macroeconomics

Gianluca Moretti

Thesis submitted for Ph.D.

The University of York

Department of Economics and Related Studies

August 2006

Abstract

In this PhD thesis I investigate the implications of heterogeneity and aggregation in macroeconomic models. The importance of aggregation lies on the fact that when heterogeneity is allowed, we cannot expect macro models to have the same characteristics as the underlying micro models. In particular, a direct consequence of aggregation is that the dynamic properties of the micro model do not hold in general for the macro model. Despite this problem, modern macroeconomics tends to model aggregate data alone, through the construction of models where the individual consumer or firm is related to aggregate data under the guise of a "representative agent". In this thesis, I present a heterogeneous real business cycle model where I allow for cross sectional heterogeneity in the dynamics of the firm productivities. I show that heterogeneity allows the model to generate very persistent dynamics that can mimic impressively those of actual data. This is because, the dynamics of the model are now the result of the interactions between heterogeneous firms. Another problem that often arises with heterogeneity is that through aggregation, the dynamics that describe the co-movements between two variables can be more persistent and complex than the dynamics observed for the individual behaviour. Standard co-integration techniques are not able to deal with such persistent co-movements since they cannot distinguish between persistent deviations from the equilibrium and spurious relations. Therefore, many intuitive economic relations are often empirically rejected. To this purpose, I introduce in the thesis a methodology which can test robustly for co-integration between two variables, which deviate persistently from their long-run equilibrium. I test for a co-integration in the Uncovered Interest Parity and the Purchasing Power Parity with my approach and, unlike the standard approaches, it does not reject the hypothesis that they hold in the long run.

Contents

List of Tables

List of Figures

Acknowledgements

I am deeply grateful to my supervisor, Gabriel Talmain, for his support as well as invaluable help and patience throughout my Ph.D.

I wish to particularly thank Karim Abadir and the members of my thesis advisory group, Huw Dixon and Peter Spencer, for their technical support and helpful suggestions.

I would like to express my gratitude towards to Franco Peracchi and Franco Giusti for their encouragement and support during my first year in York and for their consistent trust when I was an undergraduate and a graduate student.

I gratefully acknowledge the ESRC (UK) grant RES-000-23-0176 and assume the responsibility for the conclusions of this thesis and any errors it may contain.

My best regards to people in Genesis 6 and other friends in York for providing such stimulating and fun environment during these years and making my time there very special.

I would like also to thank my colleagues at the Econometric Research Office of the Bank of Italy for their support and understanding while completing this thesis.

I am forever indebted to my parents for their continuous support during the last thirty years and in particular to my mother to whom I would like to dedicate this thesis.

Finally, a very special thank to mały Hania for being particularly supportive and patient, especially in the last few months of the revision of the thesis.

Declaration

I hereby declare that the dissertation, submitted in partial fulfilment of the requirements for the degree of Doctorate of Philosophy and entitled "Essays on non-linear aggregation in macroeconomics", represents my own work and has not been previously submitted to this or any other institution for any degree, diploma or other qualification.

Chapter 2 is currently being revised and extended with Gabriel Talmain. A previous version of this paper has been presented in a workshop at the University of York.

A previous version of chapter 4 has been presented at the Week of Doctoral Studies at the Charles University, Prague, June 2004.

Chapter 1

Introduction

Aggregation is an inevitable aspect of economics. On one side, most of the observed macroeconomic time series are the result of the aggregation among heterogeneous units, such as commodities, productivities, firms or households. On the other, modern macroeconomic theories link behavioral relations of individual agents to the dynamics of the macroeconomic variables.

Historically, given the close link between the concepts of aggregation and heterogeneity, we can trace its origin back to the Adam Smith's principle of the "invisible hand". According to this principle, the interaction of heterogeneous agents, each of them pursuing the selfish ends, leads to equilibrium which is socially satisfactory.

From the second half of this century, important authors such as Gorman (1953), Klein (1953), Theil (1954), Muellbauer (1976, 1981) and Stoker (1984, 1986) have studied aggregation in many areas of economics, including production, income, investment and employment. In particular, they analyzed the conditions under which macro models reflect and provide interpretable conditions on the behavior of the micro units. Their works are also known as the "deterministic approach" to aggregation in contraposition to the more recent "stochastic approach" that has become quite popular in the last two decades.

This modern theory of aggregation has developed quite rapidly and can be basically divided into two major approaches. The first involves modelling aggregate data alone, through the construction of models where individual consumer or firm is related to aggregate data under the guise of a "representative agent". These models have become the workhorses of modern real business cycle (RBC) theory. They explain aggregate behaviors as if they were a large scale magnification of the behavior of a representative rational agent. In the past years, these models have been extended to many areas of macroeconomics such as monetary analysis[1], international economics[2] and labor markets[3].

The other aggregation approach is based on adopting a framework that allows individual data and aggregate data to be modelled under one consistent format. In particular, an individual model is specified together with some assumptions that allow us to formulate an aggregate model which is consistent with the individual model. Such approach combines individual heterogeneity with a tractable and parsimonious model for the aggregate data. This idea was initially put forward by Robinson (1978) and Granger (1980) who were the first to study the asymptotic behavior of aggregate heterogeneous time series models. Further developments can be found in Forni and Lippi (1997), Chambers (1998), Lewbel (1994) and Zaffaroni (2004), just to mention some important contributions.

A characteristic of these models is that when heterogeneity is allowed, we cannot expect the macro model to be characterized by the same dynamic properties as the underlying micro model. For example, it is a well-known result[4] that averaging a large number of linear micro models would generate long memory in the aggregate process, even if its micro-components are strictly stationary.

[1]See Dotsey et.al. (1999), Altig, Christiano et. al. (2005) and Smets and Wouters (2003), just to mention some relevant work.

[2]See Backus and Kehoe (1992), Baxter (1995) and Ambler et.al (2004)

[3]See for instance, Merz (1995), Den Hann et. al. (2000).

[4]See Granger (1980) and more recently Zaffaroni (2004).

A direct consequence of this "aggregation effect" is that the basic properties of the micro model do not hold in general for the macro model. This is true for many econometric procedures, such as cointegration and Granger causality, that need therefore be redefined in such a context.

Although (linear) aggregation has been originally designed for time series models, in the last years there have been a few attempts to combine this approach with neoclassical representative agent models. In fact, since the reduced form of single agent models can be represented as a linear model with some autoregressive components, some authors have tried to merge the two approaches by imposing heterogeneity in the reduced form of these models. See Michelacci and Zaffaroni (2000) for an application to international economics, and Krussel and Smiths (1998) to labor market.

However, none of the proposed approaches have been able so far to develop a well grounded model of individual behaviors that trace at aggregate level the implications of such behaviors. In fact, they fail either to account for heterogeneity in modelling aggregate variables or to provide a framework that links individual heterogeneity and structural modelling of aggregate data.

On one side, representative agent models, despite their popularity, have not provided any conceptual foundation or realistic condition for ignoring compositional heterogeneity in the dynamics of aggregate data. They are also often characterized by properties that are at odds with empirical evidence[5] giving rise to behavior, so called "puzzles", that cannot be reconciled with stylized facts[6]. For instance, a well-known puzzle that curses representative agent models is the "weak transmission mechanism". In fact, according to the findings of Cogley and Nason (1995), these models have to rely on implausible exogenous sources of dynamics to replicate the persistence and the volatility observed in actual data.

[5]See Basu (1998), King and Rebello (1999), Forni and Lippi (1999) and Cogley and Nason (1995).

[6]Some very popular puzzles are the uncovered interest rate puzzle, the purchase parity puzzle and the "excess smoothness" of consumption puzzle.

On the other, although linear aggregation focuses on heterogeneity, it lacks a microfounded structure for the individual behaviors that allows us to trace the implications of such behaviors at aggregate level. In other words, it fails to provide a framework where individual heterogeneity is incorporated in a structural modelling of aggregate data. In fact, if on one side linear aggregation helps to reconcile the dynamics of micro data with those of aggregate data, see Forni and Lippi (1999), on the other it cannot provide by construction any explanation underlying these differences. For example, in a recent paper, Altissimo, Mojon and Zaffaroni (2005) show that linear aggregation can conciliate slow macroeconomic adjustment with fast adjustment at microlevel in the euro area inflation.

Despite the relevance of their findings, their framework cannot provide any explanation of the causes underlying the inflation persistence in the euro area. This is because linear aggregation lacks a structural model describing micro behaviors.

An important contribution in this direction has been recently given by Abadir and Talmain (2002). They show that the introduction of firm heterogeneity in a dynamic general equilibrium (DGE) model would generate a new form of aggregating micro behaviors. In particular, differently from the previous approaches that either assume no heterogeneity or impose heterogeneity "ad hoc" in the reduced form of the model, they exploit the aggregation formula that arises from the micro structure of the model. This creates a framework that allows us to study actual observed aggregate data patterns as they are related to the characteristic of a specific economy.

The purpose of this thesis is to make a "little step" toward investigating the implications of this structural aggregation approach. In particular, we show that by combining the structure of a general equilibrium model together with heterogeneity, it is possible to solve many of the problems that characterize the two single approaches: the representative agent approach and linear aggregation.

On one side, through heterogeneity we can reconcile the dynamic properties of a DGE model with those of actual data and give a new insight into the propagation mechanism of the models. We show that, in this framework, cycles and fluctuations emerge not as the result of the reaction of one individual to some substantial exogenous shocks (like in representative agent models), but as a natural result of the interactions between heterogeneous agents. Endogenous fluctuations arise very naturally in a context of many interacting agents.

On the other side, differently from linear aggregation we can link through the structure of the DGE model the properties of the microstructure of the model to the patterns of aggregate data. In fact, we are able to map and quantify the effect of structural parameters of the model on the time series patterns of the aggregate data.

We start by showing in chapter 2 how heterogeneity and structural aggregation can reconcile the dynamic properties of a DGE model with some stylized fact on actual data. In particular we focus on the dynamics of the U.S. GNP. A well-known problem that curses real business cycle models is the "weak propagation mechanism puzzle".

This puzzle was discovered in a recent paper by Cogley and Nason (1995). Testing through simulations a wide range of RBC models they showed that none of them is able to reproduce the persistence and the volatility that is observed in the U.S. GNP data. They concluded that these models are plagued with a weak propagation mechanism since they must rely on implausibly large exogenous sources of dynamics to replicate the dynamics of actual data. This point was also raised by Rotemberg and Woodford (1996) and Muellbauer (1997). Although more sophisticated versions of RBC models have been proposed to make its dynamics consistent with data, none of them has been able to give exhaustive results. A survey on these models can be found in King and Rebello (1999) and Rebello (2005).

In this chapter we show how heterogeneity and structural aggregation can give a new insight into the propagation mechanism of these models. For this purpose we present a heterogeneous RBC (HRBC) model where we allow for cross sectional heterogeneity in the dynamics of the model's firm productivities.

Specifically, we consider the monopolistic competition model introduced by Abadir and Talmain (2002) and model firm heterogeneity using micro data for 450 U.S. sectors. Assuming that the productivity of each firm evolves according to an autoregressive process, we estimate its parameters from the U.S. micro data. We then calibrate the other parameters of the model and compare the time series of the simulated data with the ones of the U.S. data using the same test proposed by Cogley and Nason (1995). In particular, we calculate the probability to observe the persistence and the volatility of the U.S. GNP under the null hypothesis that the data were generated by the HRBC. Surprisingly, we fail to reject the null hypothesis. This means that the HRBC model generates dynamics that can mimic those of the data. This result shows that heterogeneity in the firm productivity allows the model to build up an internal transmission mechanism that is actually missing in representative firm models. Another implication of this result is that whenever the conclusions of a single-agent model are empirically rejected, we cannot discern whether we are rejecting the implicit properties of the model or we are rejecting the assumption that the economy is represented by a single individual.

This chapter represents a novelty in the DGE literature for the following reasons. First, we propose an approach to estimate the microstructure of the economy and model parsimoniously firm heterogeneity. In this respect, even if heterogeneity of the economic agents is a very intuitive and plausible assumption, it has not been implemented so far in a structural model calibrated using micro data for a specific economy. Then, we test extensively the dynamics of the model and show that through heterogeneous firm the output dynamics of the model can reproduce[7] the persistence and volatility observed in the U.S. GDP. This is more ambitious achievement for a RBC model that can usually aims to match few correlations of the data. Finally, it has to be noticed

[7]Even if Abadir and Talmain showed that their model could theoretically match the properties of the U.S. GNP, they did not estimate the micro structure for these economies and test whether their intuition would hold empirically.

that our results are even more striking if we consider that our approach is demanding the model to replicate both the structure of the micro-data dynamics and the pattern of the macroeconomic data.

As already mentioned, a direct consequence of aggregation is that the properties of a micro model do not hold in general for the macro model. This is also true for the dynamics that describe the evolving of a certain (cointegrating) relation between two variable at the micro level. In particular, aggregation can make these dynamics change dramatically at macro level by generating long memory and non linearities. Evidence of long memory in the co-movements of many macroeconomic variables have already been found by Cheung (1993), Diebold et al. (1991) and Abadir et. al. (2006). This means that although there is an equilibrium relation between economic variables spanning the long-run, these variables can be away from such equilibrium for a very long length of time. As recently shown in a numbers of papers, see for instance Gonzalo and Lee (1998), standard cointegration techniques are not able to deal with such persistent co-movements since they cannot distinguish between long-memory equilibrium and spurious relations. In particular, both the Engle-Granger procedure and the Johansen's full-Information maximum likelihood display have very low power against the alternative of long memory co-movements.

To this purpose we present in chapter 3 a methodology to test for the presence of long memory co-movements when the two variables are non-stationary and there exists a linear combination that is characterized as a long memory process.

The contribution of this chapter is the following. First, we provide with a methodology that, differently from standard cointegration technique, is able to detect the presence of long-memory co-movements. In this respect, the test we propose displays both high size under the null (of spurious relation) and high power under the alternative (long memory relation). Then, it also reduces the small sample bias in the estimation of the long-run relation by more than 41% compared to standard ordinary least squares.

This methodology has the potential to help solve many of the puzzles that arise in macroeconomics. For instance, we apply our procedure to the Johansen and Juselius (1992) data base for the UK purchasing-power parity (PPP) and uncovered interest rate parity (UIP). It is a well-known result that cointegration is rejected for these variables even if it is economically plausible[8]. We show instead that with our approach the null of no cointegration is rejected at 95%, implying the existence of long memory equilibrium among these variables. This difference in the results can be interpreted in the following way. Testing with standard cointegration approaches imposes very strict conditions on the long-run relation among the variables: first, it assumes that the relation is linear, then deviations from the long-run equilibrium must be strictly stationary. In the light of the evidence produced in this thesis, these conditions are not very likely to be satisfied by actual data. Therefore, it should not be very surprising that we often reject cointegration even when a certain long-run relation emerges from economic theory. On the other side, the approach we propose is able to test for the existence of a long-run relation, whilst allowing for possible non-linearities and persistent deviations from its long-run that are not forced to be strictly stationary.

In chapter 4, we compare structural aggregation with standard linear aggregation. In particular, we confront the statistical properties of a process generated by Abadir and Talmain's framework with those of a linear aggregate process. The importance of such comparison arises from the different implications they have on dynamic properties of the aggregate process. Given a stochastic sequence $\{x_{i,t}\}_{t=1}^{\infty}$, whose logarithm evolves as linear process, a linear aggregate process is defined as

$$X_t = N^{-1} \sum_{i=1}^{N} \ln x_{i,t}$$

On the other side, the process that arises from Abadir and Talmain's formula is

$$Z_t = \ln \left(N^{-1} \sum_{i=1}^{N} x_{i,t} \right)$$

[8]See Harris (1992).

Abadir and Talmain's approach differs from the linear aggregation in two aspects. First, it is by construction coherent with the micro-structure of a monopolistic competition framework. Then, as can be readily seen, its aggregation formula is consistent with the way economic data is assembled in national accounts. In fact, most of the macroeconomic time series, available in the national statistical database, are obtained by adding up the levels rather than the logarithm of disaggregated data. Thus, the need to compare these two approaches arises from the fact that while the former is the one usually studied in the time series analysis, the latter is the one that arises from the "true" data generating process in national accountancy.

Although some authors[9] claim that the X_t can represent, as first order approximation, the dynamics of Z_t, we show that this is true only if certain conditions on the volatility of the microdata are satisfied. In fact, only under these conditions it is possible to show that the autocovariance functions of X_t represent the leading term of a small variance expansion of the autocovariance function of Z_t. We also assess the accuracy of this approximation through simulations. Specifically, we generate a panel of microdata $\{x_{i,t}\}_{t=1}^{\infty}$ for different distributions of the volatilities of the micro models. Then, we aggregate them according to both scheme in order to get the processes X_t and Z_t. Finally, we compare the autocovariance functions of the two aggregate processes. Our result show that when the distribution of the volatilities of the micro models is dense in a neighborhood of zero, then the two processes generate similar dynamics in terms of persistence and volatility. Conversely, if this condition does not hold, their dynamics can be dramatically different due to the effects of the non-linearities that characterize Z_t. In particular, linear aggregation tends to underestimate both the volatility and the persistence that would arise by a correct aggregation procedure. In this situation, the assumption of linearity of the aggregate process is not innocuous, not even as first order approximation.

[9] See for instance Forni and Lippi (1997)

Chapter 2

Output dynamics in a heterogeneous RBC model

2.1 Introduction

In their seminal paper on the time series properties of real business cycles (RBC) models, Cogley and Nason (1995) showed that standard representative-agent RBC models are plagued by a "weak propagation mechanism". Specifically, this kind of model lacks an internal propagation mechanism that would transform a low amplitude, weakly correlated input signal into a highly correlated output signal of significantly larger amplitude. In fact, by testing through simulations a wide range of RBC models they showed that none of them is capable to reproduce the persistence and volatility[1] that are observed in the actual time series for the U.S. GNP[2]. They conclude that the only way to close this gap between actual and simulated data dynamics is to allow

[1]Persistence is measured using the autocorrelations, while volatility, using the impulse-response function.

[2]They also considered models that rely on counterfactual external sources of dynamics such as temporal aggregation, serially correlated increments to the productivity, and high order AR representation for transitory shocks. Although these assumptions help to reduce the gap between the persistence in simulated and actual data it does not give any improvement in terms of volatility.

these models to rely on implausibly large exogenous sources of dynamics. More recently, this point has also been claimed by Rotemberg and Woodford (1996) and Muellbauer (1997).

In the past few years many different extensions of RBC models have been proposed to make its dynamics consistent with those of actual data (in particular with U.S. GNP), but none of them has been able to give exhaustive results. In fact, they managed at most to replicate a few correlations of the data rather than matching the entire pattern of the autocorrelation function (ACF). For this purpose, Hansen (1985), Burnside et al. (1993) and Burnside and Eichenbaum (1996) have incorporated factor hoarding, Rotemberg and Woodford (1996) considered oil shocks in order to capture supply-side shocks, and Baxter and King (1993) have incorporated habit formation. A detailed survey on these approaches can be found in King and Rebello (1999) and Rebello (2005). A common assumption that characterizes these RBC models is that the economy is composed of a representative firm[3] whose productivity dynamics are driven by a log-linear process, either as stationary fluctuations around a linear deterministic trend or as ARIMA process. As we will show below, this assumption plays a crucial role in determining the dynamics of these models.

Recently, an important contribution on this topic has been given by Abadir and Talmain (2002). They introduce firm heterogeneity in a general equilibrium model and show that it would lead to an aggregate output being a nonlinear aggregation of the firm total factor productivities. They derive the time series properties of this aggregate process and show that they are very different from those of the log-linear processes, that drive the dynamics of the individual firm. In fact, it is characterized by long memory, nonlinearities and mean reversion despite its persistence. Most importantly, they show that this process can successfully predict the ACF structure of the U.S. and U.K. GNP.

[3]or by a continuum of firms that face the same constant return technology.

The intuition behind their approach relies on a fundamental theorem in time series analysis: the aggregation of single linear dynamic micro-relationships generates a process that has very different properties from its single component. For instance, it is a well-known result[4] that the linear aggregation of a large number of ARMA processes gives rise to an aggregate process that displays long memory and requires fractional differencing to achieve a stationary ARMA representation. Earlier works on linear aggregation include Robinson (1978), Granger (1980) and Lewbel (1994). The novelty of Abadir and Talmain's (2002) approach is to consider structural CES aggregation, that arise from the microstructure of the model, rather than the arithmetic average of ARMA processes as is usually done in the literature. From a macroeconomic point of view this means considering heterogeneity in the structural form of the model rather than in the reduced form[5]. Further references on this topic are Forni and Lippi (1997), Sowell (1992) and Zaffaroni (2004) while some economic application can be found in Michelacci and Zaffaroni (2000), Pesaran (2003), and Altissimo, Mojon and Zaffaroni (2005).

In this paper we present a heterogeneous Real Business Cycle (HRBC) model where we introduce firm heterogeneity using micro-data for the U.S. manufacturing industries. We compare the time series properties of the model-generated data with those of the U.S. data following the approach in Cogley and Nason (1995). The main contribution is to show that by allowing for cross sectional firm heterogeneity, a simple RBC has a striking capability to replicate the dynamics of the U.S. GNP, both in terms of persistence and volatility. Specifically, we consider a monopolistic competition framework and model firm heterogeneity using micro data for 450 U.S. sectors. Assuming that the productivity of each sector evolves according to an auto regressive process of order one,

[4]See Granger (1980) and Lippi and Zaffaroni (1998).

[5]This difference might seem subtle but it has important implications for the time series properties of the model generated data. In fact, in the latter case it implies solving the form structural model for a representative agent and then allowing for heterogeneity in the reduced form (see Michelacci and Zaffaroni (2000)), while the former considers heterogeneity directly in the structural form.

we estimate the distributions of the autoregressive parameters and standard deviations for all sectors. We then parameterize these distributions and use the model to generate time series for the aggregate output. Finally, we calculate the probability to observe the persistence and volatility in the U.S. GNP under the null hypothesis that the data were generated by the HRBC model. Our simulations strongly reject the hypothesis that both the volatility and persistence of the simulated data are significantly different from those of U.S. data. This result is quite surprising and it has not been achieved so far with this kind of model. This means that heterogeneity in the dynamics of the firm productivities is fundamental to build up an endogenous transmission mechanism in the RBC model, which generates output dynamics that mimic those of the U.S. economy.

The approach presented in this paper can be seen as a novelty in the RBC literature for the following reason. Even if heterogeneity of the economic agents is a very intuitive and plausible assumption, it has not been implemented so far in a structural model calibrated using micro data for a specific economy. On one side, most macroeconomic models assume that the economy is composed of homogenous firms that face the same productivity dynamics. On the other, even if by allowing for heterogeneity Abadir and Talmain's model could theoretically match the ACF of U.S. and U.K. GNP, they did not estimate the microstructure for these economies and test whether their intuition would hold empirically.

In this respect, our results give a new insight into the "weak propagation mechanism puzzle". They clearly show that the inability of RBC models to replicate the observed dynamics is due to the non-neutral effects of the representative agent hypothesis that characterizes not only RBC models but most of the models employed in macroeconomics. In fact, this assumption crucially affects the dynamics of the model-simulated data and it does not allow for any interaction among agents within the same sectors. Consequently, it can produce misleading results and poor consistency of the model with the data.

The paper is organized as follows. In the next section, we describe the HRBC model. Section 3 analyzes the statistical properties of the aggregate process and compares them with those of log-linear processes. Section 4 presents the estimation procedure for the parameters of the distributions that account for firm heterogeneity. In section 5, we describe the approach to test for significant differences between the dynamic properties of actual and model generated data. Finally, in section 6 we present our simulation results.

2.2 The model

The model[6] is a dynamic general equilibrium framework with monopolistic competition and it is characterized by a representative agent and two firm sectors. The final sector is composed of competitive firms that produce an output Y_t using intermediate goods q_i as input. The intermediate sector is composed of heterogeneous firms that are monopolistic competitors and produce a differentiated output using capital and labor. It has to be mentioned that we tried to keep the model as standard as possible for two reasons. First, in order to isolate the effects of heterogeneity on the dynamic properties of the model, we did not want to incorporate any other endogenous mechanism that could have affected either its persistence or volatility, e.g. labor adjustment costs or gestation lags. These mechanisms would make the model more realistic but in the light of our results they would play a minor role, compared to heterogeneity, in explaining the persistence and volatility observed in the data. Then, in order to compare directly our results with those in Cogley and Nason (1995) we decided to use a model which is as close as possible to the one they used in their paper.

[6]Differently from Abadir and Talmain (2002) we remove the hypothesis of 100% capital depreciation. This modification is more realistic but it complicates the solution of the model since it can no longer be solved by recursive methods.

2.2.1 The household's problem

The economy is populated by an infinitely-lived individual that consumes a final good C_t and inelastically supplies labor l_t which has been normalized to one. His instantaneous utility function is given by

$$U_t = \frac{C_t^{1-\gamma} - 1}{1-\gamma} \tag{2.1}$$

with $\gamma > 0$.

The household's problem is to choose $\{C_{t+i}\}_{i=0}^{\infty}$ in order to maximize the present value of the expected future consumption

$$E_t \left\{ \sum_{i=0}^{\infty} \beta^i \left(\frac{C_{t+i}^{1-\gamma}}{1-\gamma} \right) \right\}$$

subject to his budget constraint

$$C_t = W_t + (R_t + (1-\delta)) K_t + \pi_t - K_{t+1}$$

which is composed by the real wage W_t, the rental cost of capital R_t, the profit π_t of the firms and the stock of capital K_t that depreciates every period at the rate δ.

The first order condition for the consumer's intertemporal problem is given by

$$C_t^{-\gamma} = \beta E_t \left[(R_{t+1} + (1-\delta)) C_{t+1}^{-\gamma} \right]$$

2.2.2 The final goods sector

The firm in the final good sector produces a homogeneous good, Y_t, using N intermediate goods q_i according to the CES aggregate production function

$$Y_t = \left[\frac{1}{N} \sum_{i=1}^{N} q_{i,t}^{\rho} \right]^{\frac{1}{\rho}} \tag{2.2}$$

where $\rho > 0$.

This kind of production function exhibits constant return to scale, diminishing marginal productivity and constant elasticity of substitution.

Furthermore, since this firm operates in a competitive environment it makes zero profit and it takes the price as given.

Therefore, the problem of the firm is to minimize its costs according to

$$\min_{q_i} \sum_{i=1}^{N} p_{i,t} q_{i,t}$$

subject to eq. 2.2.

The solution of this minimization problem is a demand for the intermediate good $q_{i,t}$ given by

$$q_{i,t} = (p_{i,t})^{-\frac{1}{1-\rho}} \left[\frac{N^{\frac{1-\rho}{\rho}} Y_t}{\left[\sum_{i=1}^{N} \left(\frac{1}{p_i}\right)^{\frac{\rho}{1-\rho}}\right]^{\frac{1}{\rho}}} \right] \tag{2.3}$$

Since we are assuming perfect competition, the aggregate price $\bar{p}_t$ is equal to the marginal (and average) cost and it can be derived using eq.2.3 as

$$\begin{aligned} \bar{p}_t &= \frac{\sum_{i=1}^{N} p_i q_i}{Y_t} = \sum_{i=1}^{N} p_i (p_{i,t})^{-\frac{1}{1-\rho}} \left[\frac{N^{\frac{1-\rho}{\rho}}}{\left[\sum_{i=1}^{N} \left(\frac{1}{p_i}\right)^{\frac{\rho}{1-\rho}}\right]^{\frac{1}{\rho}}} \right] \\ &= \sum_{i=1}^{N} (p_{i,t})^{-\frac{\rho}{1-\rho}} \left[\sum_{i=1}^{N} (p_{i,t})^{-\frac{\rho}{1-\rho}}\right]^{-\frac{1}{\rho}} N^{\frac{1-\rho}{\rho}} = \left[\sum_{i=1}^{N} (p_{i,t})^{-\frac{\rho}{1-\rho}}\right]^{1-\frac{1}{\rho}} N^{\frac{1-\rho}{\rho}} \\ &= \frac{N^{\frac{1-\rho}{\rho}}}{\left[\sum_{i=1}^{N} (p_{i,t})^{-\frac{\rho}{1-\rho}}\right]^{\frac{1-\rho}{\rho}}} \end{aligned} \tag{2.4}$$

Now, if we rearrange eq. 2.4 as

$$\left[\sum_{i=1}^{N} (p_{i,t})^{-\frac{\rho}{1-\rho}}\right] = \frac{N}{(\bar{p}_t)^{\frac{\rho}{1-\rho}}} \tag{2.5}$$

we can use eq. 2.5 in the firm demand q_i to express it as function of the sector and aggregate price, i.e.

$$
\begin{aligned}
q_{i,t} &= (p_{i,t})^{-\frac{1}{1-\rho}} \frac{N^{\frac{1-\rho}{\rho}}}{\left[\sum_{i=1}^{N}\left(\frac{1}{p_i}\right)^{\frac{\rho}{1-\rho}}\right]^{\frac{1}{\rho}}} Y_t \\
&= (p_{i,t})^{-\frac{1}{1-\rho}} \left[N^{-\rho}\bar{p}_t^{\frac{\rho}{1-\rho}}\right]^{\frac{1}{\rho}} Y_t = \\
&= \left(\frac{\bar{p}_t}{p_{i,t}}\right)^{\frac{1}{1-\rho}} \frac{Y_t}{N} \qquad (2.6)
\end{aligned}
$$

Eq. 2.6 expresses the first order condition of the firm demand for intermediate input q_i as function of the relative price of the input and aggregate demand Y_t.

2.2.3 The intermediate sector

The intermediate sector is composed of N monopolistically competitive firms owned by the household where each faces a demand curve given by 2.6. Every firm uses both labor l_t and capital K_t to produce output according to the constant return technology

$$q_{i,t} = \theta_{i,t} K_{i,t}^{\alpha} l_{i,t}^{1-\alpha} \qquad (2.7)$$

where $\theta_{i,t}$ is the technological process of the firm i which is assumed to be different for each firm, and as shown below, evolves as a geometric auto-regressive process of order one.

The firm i chooses $l_{i,t}$ and $K_{i,t}$ to minimize its total cost, given by

$$W_t l_{i,t} + R_t K_{i,t}$$

subject to

$$q_{i,t} = \theta_{i,t} K_{i,t}^{\alpha} l_{i,t}^{1-\alpha}$$

The first order conditions related to this problem give a capital/ labor ratio given by

$$\frac{K_{i,t}}{l_{i,t}} = \frac{\alpha}{(1-\alpha)} \frac{W_t}{R_t} \tag{2.8}$$

which tells us that the firm will choose the labor and capital according to their relative price. Since aggregate labour is standardized to one we have that the aggregate demand of capital is given by

$$\begin{aligned} K_t &= \sum_{i=1}^{N} K_{i,t} = \sum_{i=1}^{N} \frac{\alpha}{(1-\alpha)} \frac{W_t}{R_t} l_{i,t} \\ &= \frac{\alpha}{(1-\alpha)} \frac{W_t}{R_t} \sum_{i=1}^{N} l_{i,t} = \frac{\alpha}{(1-\alpha)} \frac{W_t}{R_t} \qquad (2.9) \\ &= \frac{K_{i,t}}{l_{i,t}} \qquad (2.10) \end{aligned}$$

where the last equality is obtained using the RHS of eq. 2.8 in eq. 2.9. Now, if we re-arrange the production function 2.7 for $l_{i,t}$ and use eq. 2.10 we get

$$\begin{aligned} l_{i,t}^{1-\alpha} &= \frac{q_{i,t}}{\theta_{i,t} K_{i,t}^{\alpha}} = \frac{q_{i,t}}{\theta_{i,t} (K_t l_{i,t})^{\alpha}} \\ l_{i,t} &= \frac{q_{i,t}}{\theta_{i,t} K_t^{\alpha}} \qquad (2.11) \end{aligned}$$

which expresses the demand for labor as function of the output $q_{i,t}$ and aggregate capital. In the same way, by substituting eq. 2.11 in eq. 2.10 we get the demand for capital as function of aggregate capital and firm output

$$K_{i,t} = \frac{K_t^{1-\alpha}}{\theta_{i,t}} q_{i,t}$$

It is now possible to derive the firm profit maximizing price. In fact, given firm demand for capital and labour, the production cost of a unit of output is given by

$$z_{i,t} = \frac{R_t K_t^{1-\alpha}}{\theta_{i,t}} + \frac{W_t}{\theta_{i,t} K_t^{\alpha}}$$

Since from 2.9 $R_t = \frac{\alpha}{(1-\alpha)}\frac{W_t}{K_t}$ then

$$\begin{aligned} z_{i,t} &= \frac{R_t K_t^{1-\alpha}}{\theta_{i,t}} + \frac{W_t}{\theta_{i,t} K_t^{\alpha}} \\ &= \frac{\alpha}{(1-\alpha)} W_t \frac{K_t^{-\alpha}}{\theta_{i,t}} + \frac{W_t}{\theta_{i,t} K_t^{\alpha}} \\ &= \left(\frac{1}{1-\alpha}\right) \frac{W_t}{\theta_{i,t} K_t^{\alpha}} \end{aligned} \tag{2.12}$$

We can now calculate the price that maximizes the firm profit by maximizing the profit

$$(p_{i,t} - z_{i,t})\, q_{i.t}$$

given the demand derived in eq. 2.6 and the unit cost in eq. 2.12. The solution to this problem is a price equal to

$$\begin{aligned} p_{i,t} &= \left(1 + \frac{1-\rho}{\rho}\right) z_{i,t} \\ &= \left(1 + \frac{1-\rho}{\rho}\right) \left(\frac{1}{1-\alpha}\right) \frac{W_t}{\theta_{i,t} K_t^{\alpha}} \end{aligned} \tag{2.13}$$

where the second term in $\left(1 + \frac{1-\rho}{\rho}\right)$ represents the mark-up of the firm added to the cost of producing an extra unit of output.

eq. 2.13 is very informative on how heterogeneity works in this model. Since, apart from the technological process $\theta_{i,t}$, all the terms in marginal cost $z_{i,t}$ are the same for all the firms, according to eq. 2.13 a firm with a technology above the average has the incentive to lower the price and therefore expand its market share. The inverse is true for firms that operate with obsolete technology, whose market share will consequently reduce. This effect would be completely excluded in models with a representative firm.

Finally, by substituting eq. 2.13 in the expression for the aggregate price defined in eq. 2.4, we can express the aggregate price as function of aggregate productivity,

$$\bar{p}_t = \frac{W_t}{\rho(1-\alpha) K_t^{\alpha} \theta_t^{\frac{1}{\nu}}} \tag{2.14}$$

where

$$\theta_t = \frac{1}{N}\sum_{i=1}^{N}\theta_{i,t}^{\nu} \tag{2.15}$$

is the aggregate productivity of the economy and $\nu = \frac{\rho}{1-\rho}$.

The aggregate productivity θ_t is the result of the structural aggregation of the single productivities and it is a nonlinear function of the firm productivities. The nonlinearities come from the parameter ν which is related to the mark-up of the economy. The properties of this process are described in more detail in the next section. However, eq. 2.15 shows that degree of imperfect competition of the economy, i.e. ν, affects the dynamics of the aggregate productivity θ_t. A complete description of these effects can be found in Abadir and Talmain (2002). They showed that, as the monopoly power increases, i.e. $\nu^{-1} \to \infty$, the memory of the aggregate process increases while, as the economy becomes more competitive, i.e. $\nu^{-1} \to 0$, the memory of θ_t decreases[7].

2.2.4 Aggregation and economy constraints

It is now possible to derive the aggregate production function showing the relation between total final goods output and total factors inputs. Recalling the demand for the intermediate good in eq. 2.6 and observing that by eq. 2.13 and eq. 2.14 the ratio $\frac{\bar{p}_t}{p_{i,t}}$ is equal to $\frac{\bar{p}_t}{p_{i,t}} = \frac{\theta_{i,t}}{\theta_t^{\frac{1}{\nu}}}$, we get

$$q_{i,t} = \left[\frac{\theta_{i,t}}{\theta_t^{\frac{1}{\nu}}}\right]^{\frac{1}{1-\rho}} \frac{Y}{N} \tag{2.16}$$

where θ_t is the aggregate productivity defined in the previous paragraph.

[7]In this respect, it is interesting to analyze the propagation effect of a firm productivity shock at aggregate level for different degree of competition ν. If the economy is characterized by small number of firms with high monopolistic power, the effects of a shock on one of these firm productivity will be considerable at aggregate level. On the other side, if the economy is very competitive, the importance of a single firm is irrelevant to the entire system and therefore, the effect productivity of a firm productivity shock is negligible at aggregate level. In other words, the higher the monopolistic power in the economy is, the stronger are the effects of single firm shock on the economy as a whole.

eq. 2.16 represents the demand for the good i as function of both aggregate and firm productivity.

Using this formulation for $q_{i,t}$ in the equation for the demand for labor in eq. 2.11 we get

$$
\begin{aligned}
l_{i,t} &= \frac{q_{i,t}}{\theta_{i,t} K_t^{\alpha}} \\
&= \left[\frac{\theta_{i,t}}{\theta_t^{\frac{1}{\nu}}}\right]^{\frac{1}{1-\rho}} \frac{Y}{\theta_{i,t} K_t^{\alpha} N}
\end{aligned}
\tag{2.17}
$$

and aggregating over i gives an aggregate production function defined as

$$
\begin{aligned}
1 &= \left[\frac{1}{\theta_t^{\frac{1}{\nu}}}\right]^{\frac{1}{1-\rho}} \frac{Y_t}{K_t^{\alpha}} \sum_{i=1}^{N} \frac{\theta_{i,t}^{\frac{\rho}{1-\rho}}}{N} \\
1 &= \left[\frac{1}{\theta_t^{\frac{1}{\nu}}}\right]^{\frac{1}{1-\rho}} \frac{Y_t}{K_t^{\alpha}} \theta_t = \left[\frac{1}{\theta_t^{\frac{1}{\nu}}}\right]^{\frac{1}{1-\rho}-1} \frac{Y_t}{K_t^{\alpha}} \\
&= \left[\theta_t^{-\frac{1}{\nu}}\right]^{\nu} \frac{Y_t}{K_t^{\alpha}}
\end{aligned}
$$

which implies

$$
Y_t = \theta_t K_t^{\alpha} \tag{2.18}
$$

The aggregate production function in eq. 2.18 gives an insight into the transmission mechanism generated by the model and can be used to draw a comparison with the representative firm model[8]. In the representative case, we have that all the productivities are the same, therefore eq. 2.18 becomes

$$
Y_t = \left(\tilde{\theta}_t^{\nu}\right)^{\frac{1}{\nu}} K_t^{\alpha} = \tilde{\theta}_t K_t^{\alpha} \tag{2.19}
$$

where $\tilde{\theta}$ is the aggregate technology for the representative case, i.e. $\frac{1}{N}\sum_{i=1}^{N} \tilde{\theta}_t^{\nu} = \tilde{\theta}_t^{\nu}$.

[8]This case is equivalent to assuming an intermediate sector that produces a single good.

An economic model is said to be characterized by a strong transmission mechanism if it can transform a weak input signal into a strong output signal.

Now, as eq. 2.18 and eq. 2.19 show, output dynamics are determined by movements in the productivity and movements in the capital that are induced by output dynamics. The effects of this last source of dynamics are negligible in a RBC model, see Cogley and Nason[9] (1993), therefore, in the representative firm model, productivity movements are completely transmitted to output without any structural modifications. This means that the transmission mechanism of the single firm model is very weak since an output innovation coincide with a firm productivity innovation.

The same is not true in the heterogenous firm model. First, as we show in detail below, an input signal on the firm productivity $\theta_{i,t}$ will generate an aggregate signal θ_t whose properties can be very different from those of its components. In other words, a weakly correlated innovation on firm productivity builds up persistence at aggregate level through aggregation; the transmission mechanism is generated by the aggregation of heterogeneous firms. This means that both the amplitude and the persistent of θ_t are determined by the structure of the firm heterogeneity. Therefore, we can have a variety of output dynamics according to different degrees of heterogeneity. Furthermore, by comparing eq. 2.18 with eq. 2.19, differently from the representative case, the parameter ν influences the properties of the productivity signal, in particular its amplitude.

Finally, to close the model, we need to define the aggregate economy resource constraint for the final goods

$$Y_t = C_t + I_t$$

where I_t is the investment and a law for the evolution of the capital

[9] They solve analytically the model and show that all output dynamics are determined by productivity dynamics and there is a very little contribution from the internal propagation mechanism of the model.

$$K_{t+1} = I_t + (1 - \delta) K_t$$

The equilibrium for this economy is a vector of quantities $(Y_t, C_t, I_{,t}, K_{t+1})$ such that all agents are maximizing subject to their respective constraints, supply equals demand in each market, and all the resource constraints are satisfied, given the values of the predetermined variable K_t and exogenous variable $\theta_{i,t}$.

In the next paragraph we describe the properties of the aggregate process θ_t.

2.3 The aggregate process and its statistical properties

In this paragraph we illustrate the dynamic properties of the aggregate productivity, compare them with those of log-linear processes and also draw the differences between our aggregation approach and standard linear aggregation. Let us recall the aggregate productivity as defined in eq. 2.16

$$\theta_t = \frac{\sum_{i=1}^{N} \theta_{i,t}^{\nu}}{N} \tag{2.20}$$

Following Abadir and Talmain, we assume that its single components $\theta_{i,t}$ evolve according to a geometric AR(1) process

$$\ln \theta_{i,t} = \mu + \alpha_i \ln \theta_{i,t-1} + \varepsilon_{i,t} \tag{2.21}$$

where $|\alpha_i| < 1$ for every i and $\varepsilon_{i,t} \sim IN(0, \omega_i^2)$. Both α_i and ω_i are *i.i.d.* random variables drawn from some underlying probability distribution function (pdf) $f(\alpha)$ and $f(\omega)$ that will be parameterized below. While the statistical properties (conditioned on α_i and ω_i) of the single productivities $\ln \theta_{i,t}$ are well-defined, the distribution of the $\alpha's$ and $\omega's$ will entirely determine the properties of θ_t as $N \to \infty$.

The statistical properties of θ_t were first discovered by Abadir and Talmain (2002) who showed that it is characterized by long memory, non-stationarity and mean reversion.

Before proceeding with the parameterization of the beta and generalized gamma densities, it is important to stress a point about the effects of a constant term in the single productivities. It is a well-known results in aggregation literature[10] that the statistical properties of the aggregate process θ_t can differ substantially from those of its single components $\theta_{i,t}$ and in particular, stationarity of the latter does not necessarily imply stationarity of the former. Here, although we are aggregating asymptotically stationary processes, if $\ln\theta_{i,t}$ has mean different from zero and the distribution of the micro parameters α_i is dense around the unit root, then the aggregate process θ_t will exhibit a drift because of the dominant effects of near unit roots on the aggregate process. The intuition behind this result can be roughly explained by noting that for α_i very close to one, the mean of $\ln\theta_{i,t}$ can be approximated, using L'Hopital's rule, by

$$\begin{aligned} E\left(\ln\theta_{i,t}\right) &= E\left(\mu\sum_{j=0}^{t-1}\alpha_i^j + \sum_{j=0}^{t-1}\alpha_i^j\varepsilon_{i,t-j}\right) \\ &= \mu\frac{1-\alpha_i^t}{1-\alpha_i} \\ &\simeq \mu t \end{aligned}$$

Thus, if we allow for a constant term and almost unit root in eq. 2.21, then $\ln\theta_t$ will be characterized by a mean whose leading term evolves according to a linear trend.

This argument is consistent with the findings of Diebold and Senhadji (1996), Rudebusch (1993) and Diebold and Rudebusch (1989) where, using a long span of annual data, they produce evidence in support of trend-stationarity and long memory rather than difference stationarity in the aggregate U.S. GNP.

In the light of the above considerations and since we are mainly interested in the cyclical properties of the aggregate process $\ln\theta_t$, in the simulations below we will also

[10] See Granger (1980), Lewbel (1994) and Lippi and Zaffaroni (1998).

consider a demeaned version of the process. In other words, using the transformation $\mu = (1 - \alpha_i) \ln \theta$ in eq. 2.21, we can consider the alternative formulation

$$\ln \hat{\theta}_{i,t} = \alpha_i \ln \hat{\theta}_{i,t-1} + \varepsilon_{i,t} \tag{2.22}$$

with $\ln \hat{\theta}_{i,t} = \ln \theta_{i,t} - \ln \theta$. This formulation has the effect of removing a drift from the aggregate process but it does not alter its memory.

The pdf $f(\alpha)$ and $f(\omega)$ have been chosen in the following way[11]. Since we exclude the presence of unit roots, we can assume $\alpha_i \in (0, 1)$ and parameterize $f(\alpha)$ as a Beta density with parameters g_α and h_α

$$f(\alpha) = \begin{cases} \frac{\alpha^{g_\alpha - 1}(1-\alpha)^{h_\alpha - 1}}{B(g_\alpha, h_\alpha)} & 0 \leq \alpha < 1 \\ 0 & \text{otherwise} \end{cases} \tag{2.23}$$

where $B(g_\alpha, h_\alpha)$is the Beta function. On the other side, since $\omega_i^2 \in \Re_+$ we can assume that the variate ω is distributed according to a Generalized Gamma density with parameters λ, g_ω and h_ω

$$f(\omega) = \begin{cases} \frac{\lambda h_\omega^{g_\omega} \omega^{\lambda g_\omega - 1}}{\Gamma(g_\omega)} \exp(-h_\omega \omega^\lambda) & \omega \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{2.24}$$

[11] Abadir and Talmain (2002) decompose the shock $\varepsilon_{i,t}$ into common and idiosyncratic components, i.e.

$$\varepsilon_{i,t} = \beta_i u_t + v_{i,t};\ u_t \sim IN(0, \sigma);\ v_{i,t} \sim IN(0, s_i)$$

and assume that both the loadings of the common shock β_i and the standard deviation s_i are distributed according to a generalized gamma density. We decided not to decompose the shock for two reasons. First, we are primarily interested in studying the effects of heterogeneity on the dynamics of aggregate process rather than in analyzing the decomposition of the U.S. TFP. Then, the assumption that the loadings β_i are distributed as a generalized gamma density is quite restrictive. In fact, since this density is defined only on $\Re_+$, it would not allow for any negative response of the productivities to a common shock.

where $\Gamma(g_\omega)$ is the Gamma function. Both distributions are plausible densities to represent heterogeneity as $N \longrightarrow \infty$. The beta density can take on many shapes in the interval $(0,1)$ as the parameters g_α and h_α vary. Particularly, the case $g_\alpha > 1$ and $h_\alpha \leq 1$ is very important since the pdf becomes strictly increasing and dense around 1. This implies that roots close to the unit value (but strictly smaller) are more likely to be drawn from such distribution, and consequently the persistence of the aggregate process will increase. The Generalized Gamma is a very rich distribution with support on $[0,\infty)$ that incorporates many known densities as special cases. As we will see below, the parameters of the distributions $f(\alpha)$ and $f(\omega)$ together with the parameter ν completely determine the dynamics of the aggregate process.

The statistical properties for the process θ_t in the set up just described can be found in the appendix, while a detailed analysis of the effects of the parameter's distribution on the persistence and volatility of the aggregate process is given in Abadir and Talmain (2002). It is only important to recall here that the leading term of the autocorrelation function of the logarithm of θ_t is given by

$$Cor(\ln\widehat{\theta}_t, \ln\widehat{\theta}_{t+k}) \simeq \left(1 + \left(4h_\omega^2\right)^{-\frac{2}{\lambda}} ak^2\right)^{-\frac{\nu^4}{4a}} \tag{2.25}$$

where a is an arbitrary constant. As it can be clearly seen, the parameters λ, h_ω and ν completely determine the rate of decay of the autocorrelation function of the process.

The functional form in eq. 2.25 gives rise to a wide range of ACF shapes depending on the degree of heterogeneity of the single components. According to the results of Abadir et al. (2006), this functional form characterizes the autocorrelation function of many economic aggregate variables such as GNP, inflation rates and exchange rates. Some examples of these autocorrelation functions are shown in the figures 2.1-2.3.

figure 2.1 to 2.3

Specifically, figure 2.1 plots the autocorrelation function of the logarithm of the U.S. GNP per capita that will be used in our analysis; figure 2.2, the autocorrelation function of the U.S. inflation rate[12]; figure 2.3, the autocorrelation function of the Federal Fund Rate.

Before proceeding with the estimation of the density $f(\alpha)$ and $f(\omega)$, we give some insight on the difference between our approach and linear aggregation (i.e. the arithmetic mean of the geometric AR(1) process in 2.21) that has also been proposed in similar RBC contexts[13]. As mentioned earlier, this distinction is quite important since linear aggregation implies heterogeneity in the reduced form model, while our approach, in the structural model. This difference can be easily seen by recalling the AR(1) process in eq. 2.21 and aggregating over i. Linear aggregation implies an aggregate process $Y_{N,t}$ given by

$$Y_{N,t} = N^{-1}\sum_{i=1}^{N} \ln \theta_{i,t} = N^{-1}\sum_{i=1}^{N} \frac{\varepsilon_{i,t}}{1-\alpha_i L} \tag{2.26}$$

(where L denotes the lag operator) while the logarithm of θ_t is equal to

$$\ln(\theta_t) = \ln\left(N^{-1}\sum_{i=1}^{N} \exp\frac{\varepsilon_{i,t}}{1-\alpha_i L}\right) \tag{2.27}$$

It can be readily seen that the dynamics of the two processes can be quite different and the statistical properties[14] of 2.26 can only represent, as first order approximation, those of $\ln\theta_t$, omitting eventual statistical non linearities that could arise from higher terms in a logarithm expansion.

2.4 Estimation of the heterogeneity parameters

[12] It is calculated as the logarithm of the first difference of the Consumer Price Index.

[13] See Michelacci and Zaffaroni (2000) for instance.

[14] Details on the statistical properties of linear aggregate process can be found in Granger (1980) or Zaffaroni (2004).

In this paragraph we describe the procedure to estimate the parameters of the cross sectional distributions $f(\alpha)$ and $f(\omega)$ defined above. For this purpose we use microdata for U.S. industry taken from the NBER manufacturing productivity database. This set contains annual informations on 459 manufacturing industries[15] for the period 1958-1996. A detailed description of the manufacturing establishments and the variables included in the database can be found in Bartelsman and Gray (1996). The original purpose of this database was to estimate the production functions of U.S. industries using industry level data, but recently it has been also used in a variety of research projects. We use it to estimate the parameters for the Beta and Generalized Gamma densities in the following way.

First, using the NBER data we construct for each sector the time series of the total factor productivity (TFP). The TFP gives a measure of the Solow residual for the U.S. industries, and consequently it might well represent the productivity of each firm (sector) in our model. Then, for each TPF series we fit an AR(1) process and estimate its autoregressive parameter and standard deviation. To save space, we do not report the diagnostic results, but it has to be mentioned that the AR(1) is an adequate representation for more than 95% of the sectors under consideration[16]. Finally, from the distributions of the autoregressive parameters and standard deviations we estimate respectively the parameters of the Beta (eq. 2.23) and the Generalized Gamma densities (eq. 2.24).

Specifically, the Solow residuals have been constructed following the approach in Olley and Pakes (1996) and Bartelsman and Dhrymes (1998). For each industry the Total Factor Productivity (TFP) has been calculated as

$$\ln TFP_{i,t} = \ln Q_{i,t} - \alpha_k \ln K_{i,t} - \alpha_L \ln L_{i,t} - \alpha_M \ln M_{i,t}$$

[15] A very few sectors have been removed from the sample because they had extremely large variance. Such outlier would affect the estimation of the parameters of the Gamma distribution.

[16] More specifically, 22 sectors are characterized by autocorrelated residuals while 13 sectors present omitted nonlinearities.

where $Q_{i,t}$ is the real gross output for the industry i in year t, and $K_{i,t}$, $L_{i,t}$, and $M_{i,t}$ are respectively capital, labor and intermediate inputs.

Then, for each series $\ln TFP_{i,t}$ we estimate the coefficient μ, the autoregressive parameter α_i and the standard deviation ω_i of $\varepsilon_{i,t}$, in the AR(1) representation[17]

$$lnTFP_{i,t} = \mu + \alpha_i \ln TFP_{i,t-1} + \varepsilon_{i,t} \tag{2.28}$$

In figure 2.4 we plot the histogram of the estimated $\hat{\alpha}_i$ and $\hat{\omega}_i$.

figure 2.4

Finally, using non-linear least squares we fit a beta and a generalized gamma densities to the empirical distribution of the $\hat{\alpha}'s$ and $\hat{\omega}'s$. In particular, for the estimation of the parameters of the generalized gamma distribution we follow the approach in Cohen & Whitten (1988). In the next table we report the estimated values for the parameters of the two densities.

$g_\alpha = 23.7133$	$h_\alpha = 1.1723$	$\lambda = 0.03358682$
$g_\omega = 4092.834$	$h_\omega = 4811.156$	$\nu = 10$

Table 2.1: Parameters for Generalized Gamma and Beta Density

The value for the elasticity[18] ν has been taken from Chari, Kehoe and McGrattan (2000). Visual inspection of the two plots in figure 2.5 confirms the goodness of fit between the empirical and fitted distributions and consequently the plausibility of the assumed density functions.

figure 2.5

[17]We recall that including a constant in eq.2.28 implies generating an aggregate process characterized by a drift. Therefore, in order to compare the dynamic of the simulated data with the detrended GDP we also remove the sample mean from each series $\ln TFP_{i,t}$ to transform it into a zero-mean process with $\mu = 0$.

[18]Since there is no unique agreement on the values of ν, we also run simulation with ν equal to 5 and 7 but we did not observe significant difference in the results.

It has to be mentioned that since quarterly data for TFP are not available, we use a frequency conversion to obtain estimates of the coefficient α_i and the variance ω_i^2 for the quarterly data[19]. Finally, using the estimated distributions we are able to generate the aggregate productivity according to the formula

$$\hat{\theta}_t = \frac{\sum_{i=1}^{\nu} \widehat{TFP}_{i,t}^{\nu}}{N}$$

It has to be mentioned that the estimated mean of the constant terms μ is equal to 0.006325. This value will be used in the simulations to generate time series with a drift for the artificial GNP.

2.5 Methodology

In this section, we present the approach to test for significant differences in the dynamics of actual and artificial data.

The aim of this chapter is to compare the persistence and volatility of the model-generated data with those of the U.S. data, and to test whether they could be created by the same data generating process. It has to be noticed that in our context this is a very demanding burden to put on the model as it is required to match the relationship between micro and macro data, in addition to the traditional demand to replicate the pattern of the macroeconomic data.

A drawback of the approach presented in this chapter is the high degree of non-linearities in the model. In fact, since the aggregate process does not have a recursive representation, the model has to be solved numerically[20]. For this reason, we used the Parametrized Expectation approach with a rate of convergence of 10^{-8}, see De Haan

[19]See chapter 16 of Wei (1990) for details on annual to quarterly frequency conversion.

[20]The only way to have a closed form solution with our aggregate process is to assume 100 % capital depreciation as in Abadir and Talmain. The advantage of this method is the immediate interpretation of the solution parameter, but unfortunately there are only a few situations when it can be applied.

and Marcet (1994) for further details on this solution method. It has to be mentioned that in this context log-linearizing is not possible. In fact, apart from altering the dynamic structure of the aggregate productivity[21], it would give rise to an extremely large system of equations to be solved; in the case considered in this chapter we would have more than 5000 linear equations.

Once the model has been solved, we use the reduced form to generate artificial data for the output and to compare its time series properties with the data for the U.S. real pre-capita GNP[22]. Specifically, for each artificial sample we calculate the autocovariances, the autocorrelations and the variogram[23] and collect them into empirical distributions. Finally, we use these distributions to calculate the probability to observe a certain statistic (precisely the generalized Q-statistics) for U.S. GNP under the null hypothesis that the data were generated by our RBC model.

Compared to Cogley and Nason (1995) we use different statistics to measure the persistence and volatility of the data. Regarding the volatility we choose the autocovariance function rather than the Impulse-Response (IR) function. We believe that, compared to the IR, the autocovariance function is a better measure of the amplitude of innovations in the aggregate productivity. In fact, the former would require the construction of a structural VAR whose results might be sensitive to its specification. Furthermore, even if the structural VAR were properly specified, the IR would not give more information than the autocovariance function on the volatility the process.

Regarding the persistence, we used the autocorrelation function together with another measure of persistence, the variogram, which is known to be more consistent than the ACF for very persistent data and short samples.

---

[21]The effect of ν would be completely lost by linearizing the aggregate productivity.

[22]Our definition of output is the quarterly real GNP per capita, taken from national economic accounts of the Bureau of Economic Accounts. The sample size is 228 observations.

[23]See below for the definition of variogram and the appendix for more details on its properties.

In fact, it is a widely known result[24] that the estimation of the autocorrelations for linear processes with root close to one can be very biased, even in samples of more than 200 observations, as shown in the appendix. Such bias can become dramatically large in the case of long memory in the data as shown in Newbold and Agiakloglou (1997)[25]. Since our aggregate process has long memory by construction, both the sample autocovariance and autocorrelations could be heavily biased even if we use samples with more than 200 observations. A solution to this problem is to employ a measure of time-dependence, known as the variogram, which is widely used in geostatistics. Given a stochastic process $\{x_t\}_{t=1}^{T}$, the variogram(VR) at lag k is defined as

$$VR(x_t, x_{t+k}) = \frac{var(x_{t+k} - x_t)}{2}$$

where $var(\cdot)$ indicates the variance of the increment $(x_{t+k} - x_t)$.

Therefore, instead of focusing on the moment of the process x_t, it considers the increments $(x_{t+k} - x_t)$ with $k = 0, 1, ...$ A more comprehensive description of its properties can be found in Beran (1998) and Haslett (1997).

The variogram presents many advantages compared to autocorrelation function. First, it is directly related to the autocovariance function, but differently from this one, it is unbiased for near unit root processes, and is also defined for certain non (variance) stationary processes including the random walk.

The testing procedure has been conducted in the following way. For each artificial time series generated by the model we calculate the autocovariances, the autocorrela-

[24]See Haslett (1997) for instance.

[25]Given the stochastic process $\{x_t\}_{t=1}^{T}$ we define the autocorrelation at lag k as

$$Cor(x_{t+k}, x_t) = \frac{E(x_{t+k}x_t) - E(x_{t+k})E(x_t)}{\sqrt{E(x_{t+k}^2) - E(x_{t+k})^2}\sqrt{E(x_t^2) - E(x_t)^2}}$$

This definition is slightly different from the usual definition $Cor(x_{t+k}, x_t) = \frac{Cov(x_{t+k}x_t)}{var(x_t)}$ since it allows for a mean correction. The two definitions are asymptotically equivalent in strictly stationary processes but the former can help to reduce the near unit root bias although it does not eliminate it completely.

tions and the variogram, and store them in empirical probability distributions. Then, we construct the generalized Q-statistics, as shown in the next table, and count the fraction of samples that yield significantly different results from those found in the U.S. data. The three generalized Q-statistics are defined as follow:

$Q_1 = (\hat{r} - r)' \hat{V}_1^{-1} (\hat{r} - r)$	$\hat{r} = \frac{\sum_{i=1}^{N} \hat{r}_i}{N}$	$\hat{V}_1 = N^{-1} \sum_{i=1}^{N} (\hat{r}_i - r)(\hat{r}_i - r)'$
$Q_2 = (\hat{c} - c)' \hat{V}_2^{-1} (\hat{c} - c)$	$\hat{c} = \frac{\sum_{i=1}^{N} c_i}{N}$	$\hat{V}_2 = N^{-1} \sum_{i=1}^{N} (c_i - c)(c_i - c)'$
$Q_3 = (\hat{v} - v)' \hat{V}_3^{-1} (\hat{v} - v)$	$\hat{v} = \frac{\sum_{i=1}^{N} \hat{v}_i}{N}$	$\hat{V}_3 = N^{-1} \sum_{i=1}^{N} (\hat{v}_i - v)(\hat{v}_i - v)'$

Table 2.2: Generalized Q-statistics for autocorrelations, autocovariances and variogram

where the vectors $\hat{r}$ and r are respectively the sample and model-generated autocorrelations, $\hat{c}$ and c are the sample and model-generated autocovariances, and $\hat{v}$ and v are the sample and model-generated variogram. Although the generalized Q-statistics are approximately chi-squared with degrees of freedom equal to the number of lags considered, we decided to use empirical critical values rather than asymptotic critical values. In fact, the large number of lags considered could seriously diminish the power of the test.

We consider two different cases. In the first, we generate data from the HRBC model assuming that the aggregate productivity is a zero mean process (i.e. no constant term in its AR components) and compare its dynamics with those of the linearly detrended[26] U.S. GNP. In the second case, we allow for a constant term in the single productivities[27] and compare the model dynamics with those of the logarithm of the U.S. GNP. This

[26]As mentioned above, a linear trend is consistent with the assumption of a constant term and almost unit roots in the single productivities.

[27]This time the aggregate productivity has been defined as

$$\begin{aligned} \theta_t &= \frac{\sum_{i=1}^{N} TFP_{i,t}}{N} \\ \ln TFP_{i,t} &= \mu + \alpha_i \ln TFP_{i,t-1} + \varepsilon_{i,t} \\ \varepsilon_{i,t} &\sim IN\left(0, \omega_i^2\right) \end{aligned}$$

where α_i and ω_i are drawn respectively from the estimated Beta and Generalized Gamma density.

means that in the first case we focus on the cyclical properties of actual and simulated data, while in the second on the dynamic properties of the data as a whole. For the simulations we set the number of firms N equal to 5000 and performed 30000 replications for each case. Finally, we set the lag order equal to 100 lags for the ACF of the detrended data and 150 lags for data in levels.

2.6 Simulation results

In this paragraph we present the simulation results for the exercise described in the previous paragraph. We start by reporting the estimated values for three Q-statistics together with their probability values. Then, we plot the empirical and model generated autocorrelations, autocovariances and variogram together with their 95% confidence bands. We show that in both cases we fail to reject the null hypothesis that the dynamics of the U.S. GNP could be generated by our HRBC model.

We start by showing in table 2.3 the values for the Q_1, Q_2, and Q_3 statistics with their probability values shown in parentheses. The second column refers to the case for detrended data while the third column to the case where we allow for a drift in the data.

	Q-statistics (model no drift)	Q-statistics (model with drift)
Q_1(autocorrelations)	95.8058	65.8545
	(0.3259)	(0.2318)
Q_2(autocovariances)	39.7783	131.7139
	(0.1572)	(0.2749)
Q_3(variogram)	29.6422	79.1103
	(0.2055)	(0.3539)

Table 2.3: Estimated generalized Q-statistics for autocorrelations, autocovariances and variogram

As it can be readily seen, none of the test considered is significant[28]. This implies that we can not reject the hypothesis that our simulated data can reproduce both the persistence and the volatility that is observed in the data. In particular, Q_2 and Q_3 are important since they imply that our artificial data can match not only the persistence but also the volatility of the actual data at all lags. It has to be mentioned that Cogley and Nason could not achieve these results even with models characterized by counterfactual external sources of dynamics. This shows that through heterogeneity and non-linear aggregation the model is able to generate a strong internal propagation mechanism which does not need to rely on implausibly large exogenous shocks to replicate the dynamics of actual data. The goodness of this fit is also confirmed by a visual inspection of the figures 2.6-2.11 where we plot the autocovariance, the autocorrelation function and the variogram for the artificial and U.S. data together with the 95% confidence bands. Figures 2.6-2.8 show the autocovariance, the autocorrelations and the variogram for the detrended data, while figures 2.9-2.11 for the data in levels.

figure 2.6 to 2.11

Although these figures contain the same information as the Q-statistics considered above, they allow us to see graphically how closely the simulated data match the persistence and the volatility of the actual data. In fact, it can be readily seen that the autocovariance function, the autocorrelation function and the variogram for the U.S. data lies inside the confidence band at all lags. This is a novel result for a RBC model that can usually match few correlations of the data.

Finally, it can be noticed that the variogram in figure 8 confirms the presence of long memory in the detrended U.S. GNP. In fact, it is characterized by a slow rate of convergence, that is typical of long memory process, as opposed to the very fast rate of stationary process or to the explosive rate of unit root process (see the appendix for more details).

[28] As already mentioned in a previous footnote, we got better results with $\nu = 7$ and similar results with $\nu = 5$.

We can conclude that, according to the results shown in this section, a possible solution to the "weak propagation mechanism" puzzle can be found in heterogeneity and structural aggregation. In fact, comparing our results with those of other representative RBC models, such a degree of goodness of fit for the persistence and volatility that characterize the actual data has not been achieved so far.

2.7 Conclusion

In this chapter we present a heterogeneous RBC model where firms face different productivity dynamics. We model such heterogeneity using the NBER manufacturing data base for 459 U.S. firms. By assuming that the total factor productivity can be represented by an AR(1) process, we estimate the distributions for the autoregressive parameters and the standard deviations for each sector. Then, we use these distributions to calibrate the microstructure of the model. We study the time series properties of the model-generated data and compare them with those of U.S. data. Our simulations strongly reject the hypothesis that the persistence and the volatility generated by HRBC model are significantly different from those of the U.S. GNP. This shows that, cross sectional heterogeneity in the firm dynamics allows the model to build up an endogenous strong propagation mechanism, which solves the "weak propagation mechanism puzzle" that characterizes standard RBC models.

In this chapter, we were mainly concerned with the effects of heterogenous productivity dynamics on the aggregate TFP and output dynamics. However, our approach could be extended to many areas of macroeconomics to build up more exhaustive and complete heterogeneous models. The extensions in these directions could be numerous. The ideal would be to construct a micro founded model with a complete heterogeneous microstructure and calibrate it using industry sector data. This could lead to, for instance, heterogeneous production functions with different degrees of capital utilizations and capital depreciations. On the consumer side, it would be interesting to consider the implications of consumers with different discount factors, risk aversion and preferences on the properties of a RBC model. This could lead to the aggregation of heterogeneous rational expectation models[29] that have the potential to help solve many of the modern economic puzzles such as inflation and unemployment persistence.

[29] An attempt in this direction has been done by Pesaran (2003).

2.8 List of figures

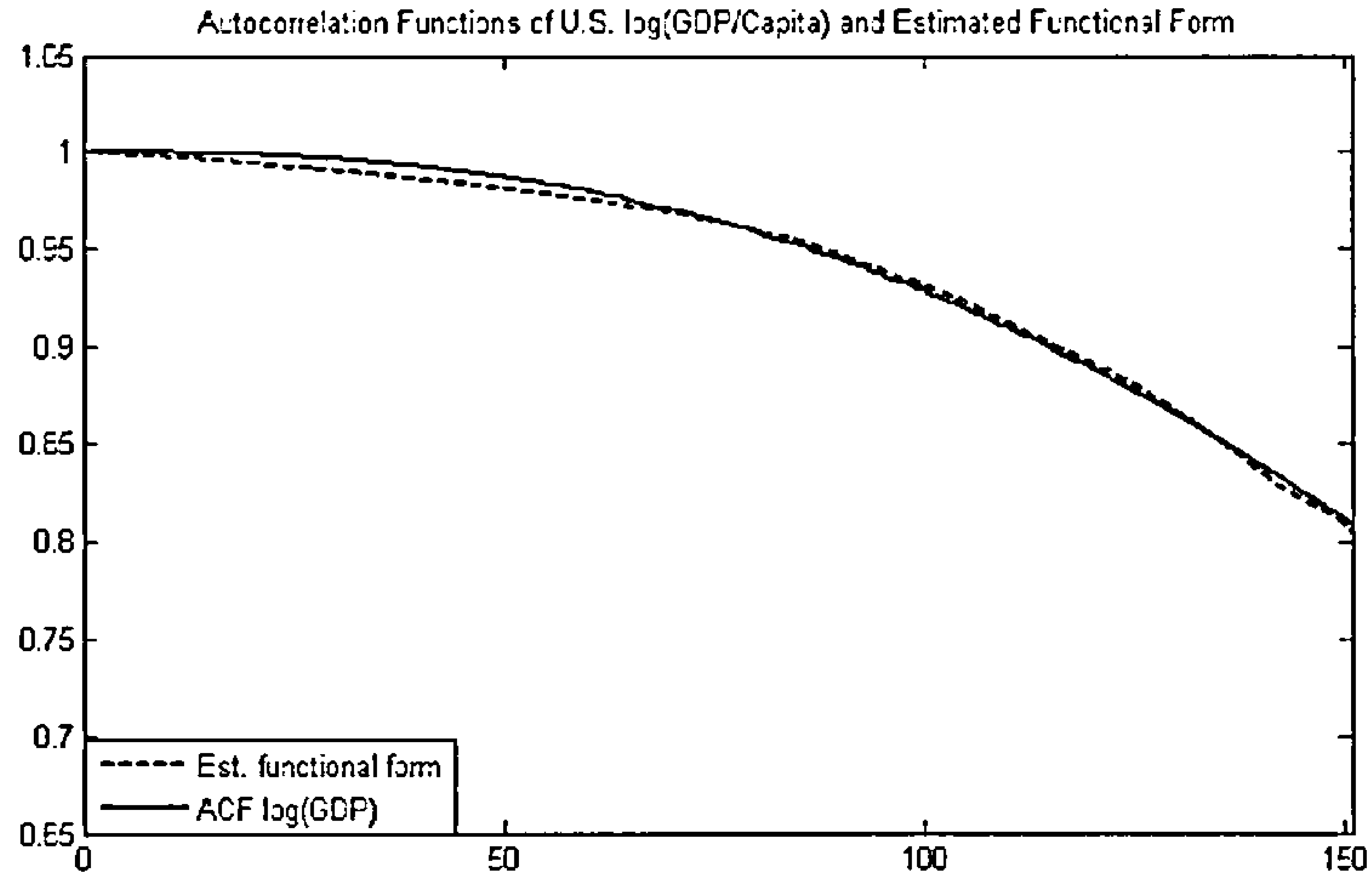

Figure 2.1: Autocorrelation function U.S. GNP and fitted functional form

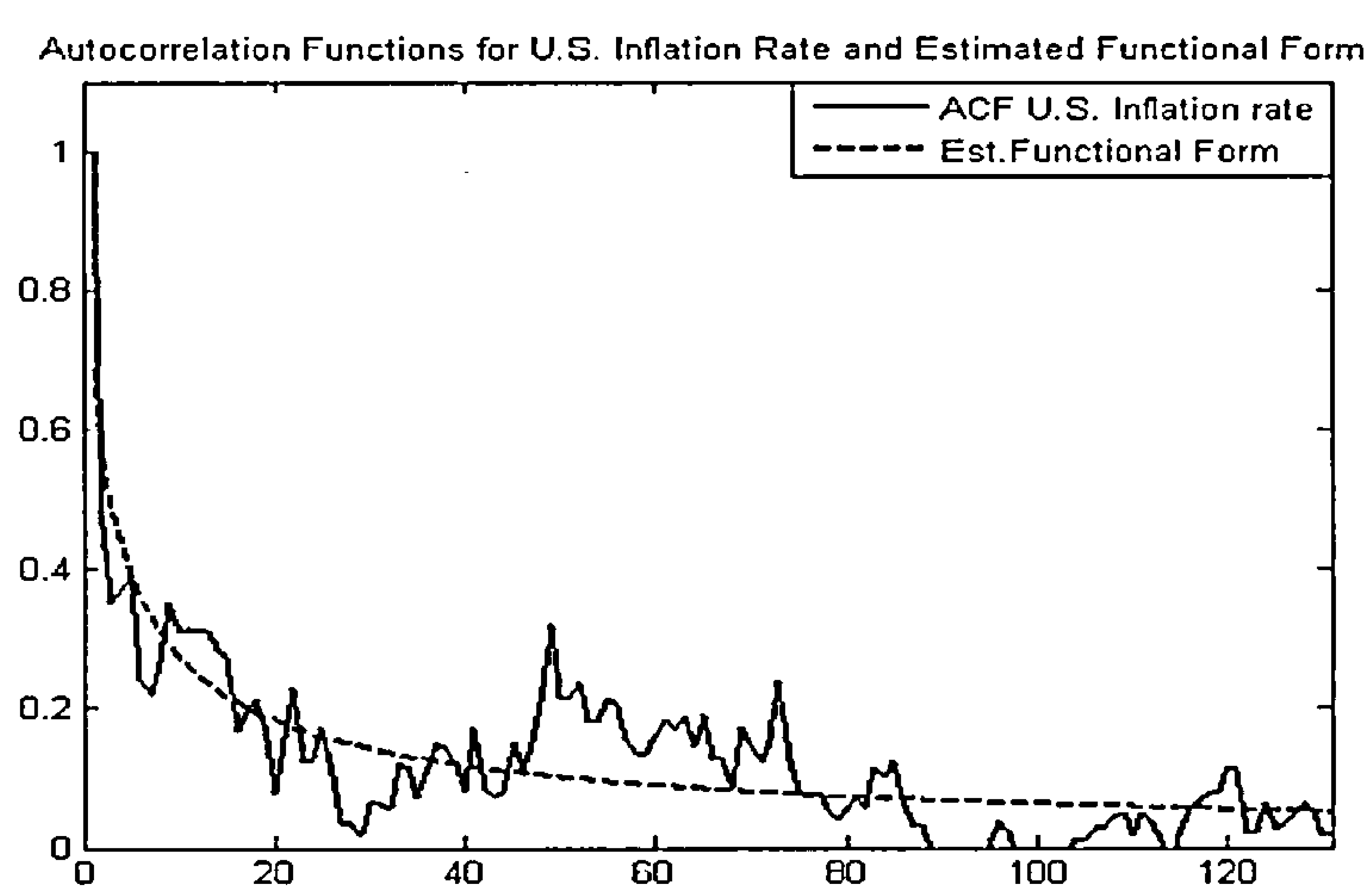

Figure 2.2: Autocorrelation function U.S. inflation rate and fitted functional form

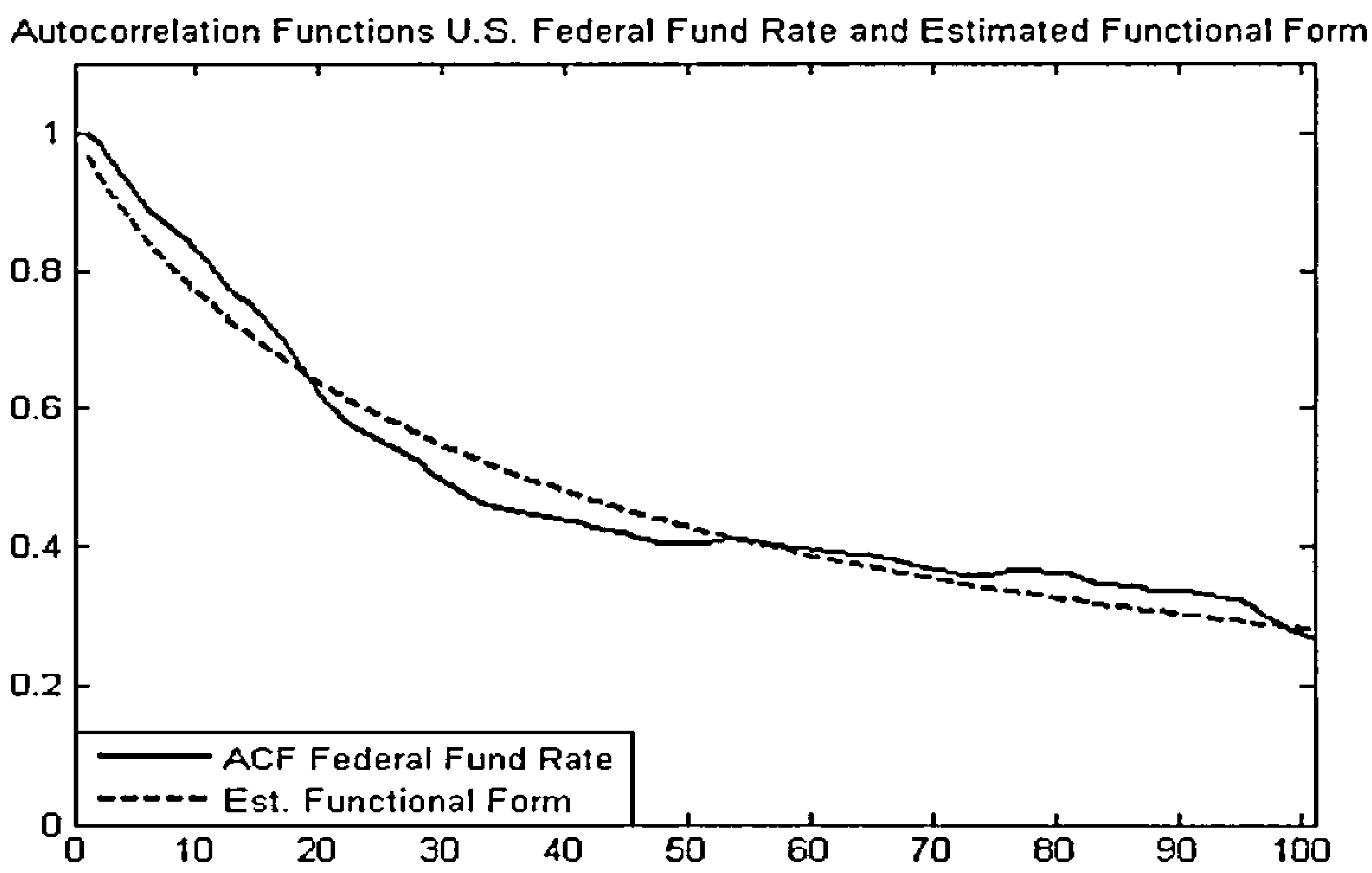

Figure 2.3: Autocorrelation function U.S. federal fund rate and Fitted Functional Form

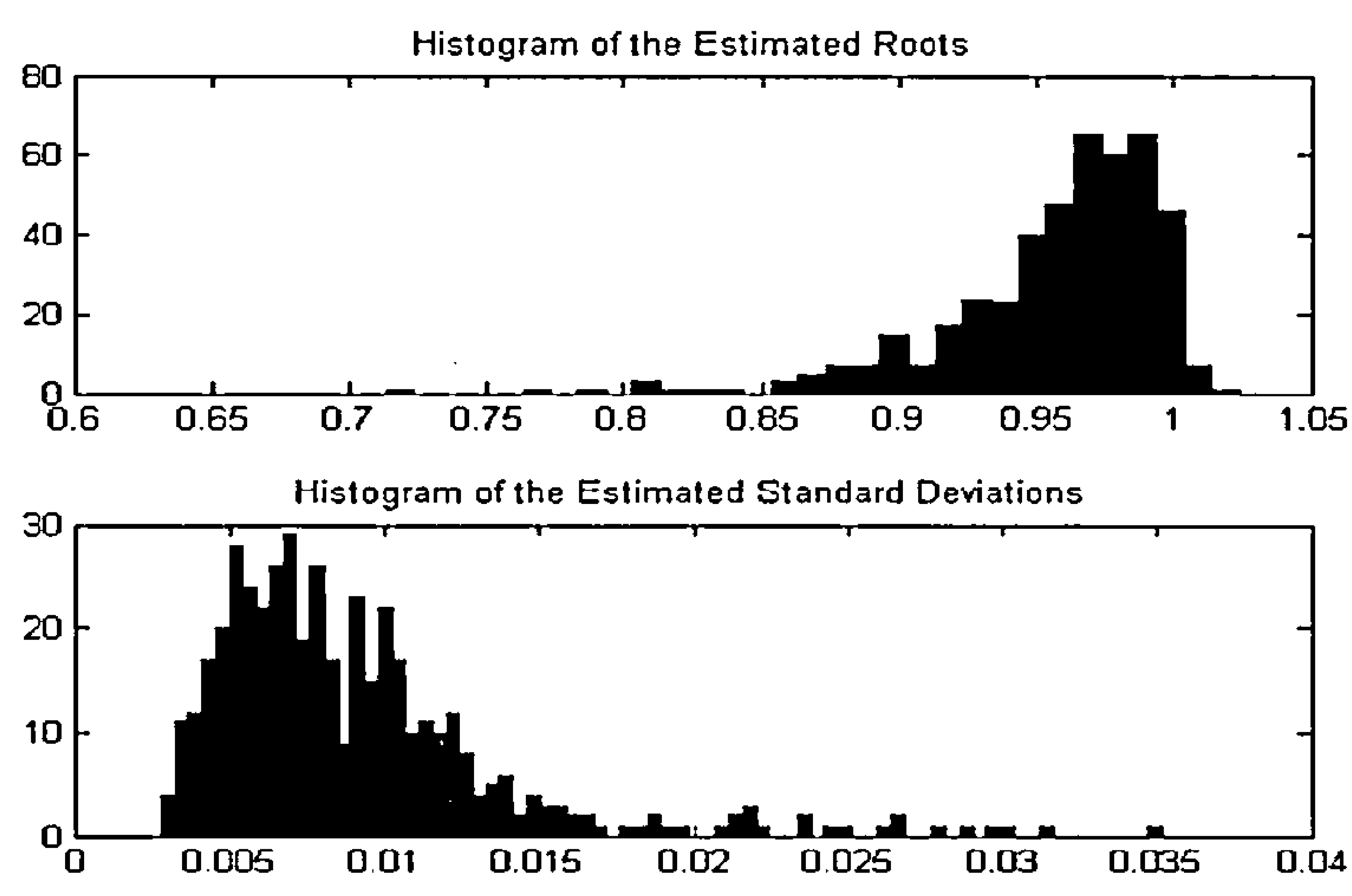

Figure 2.4: Histogram of the estimated autoregressive roots and standard deviations

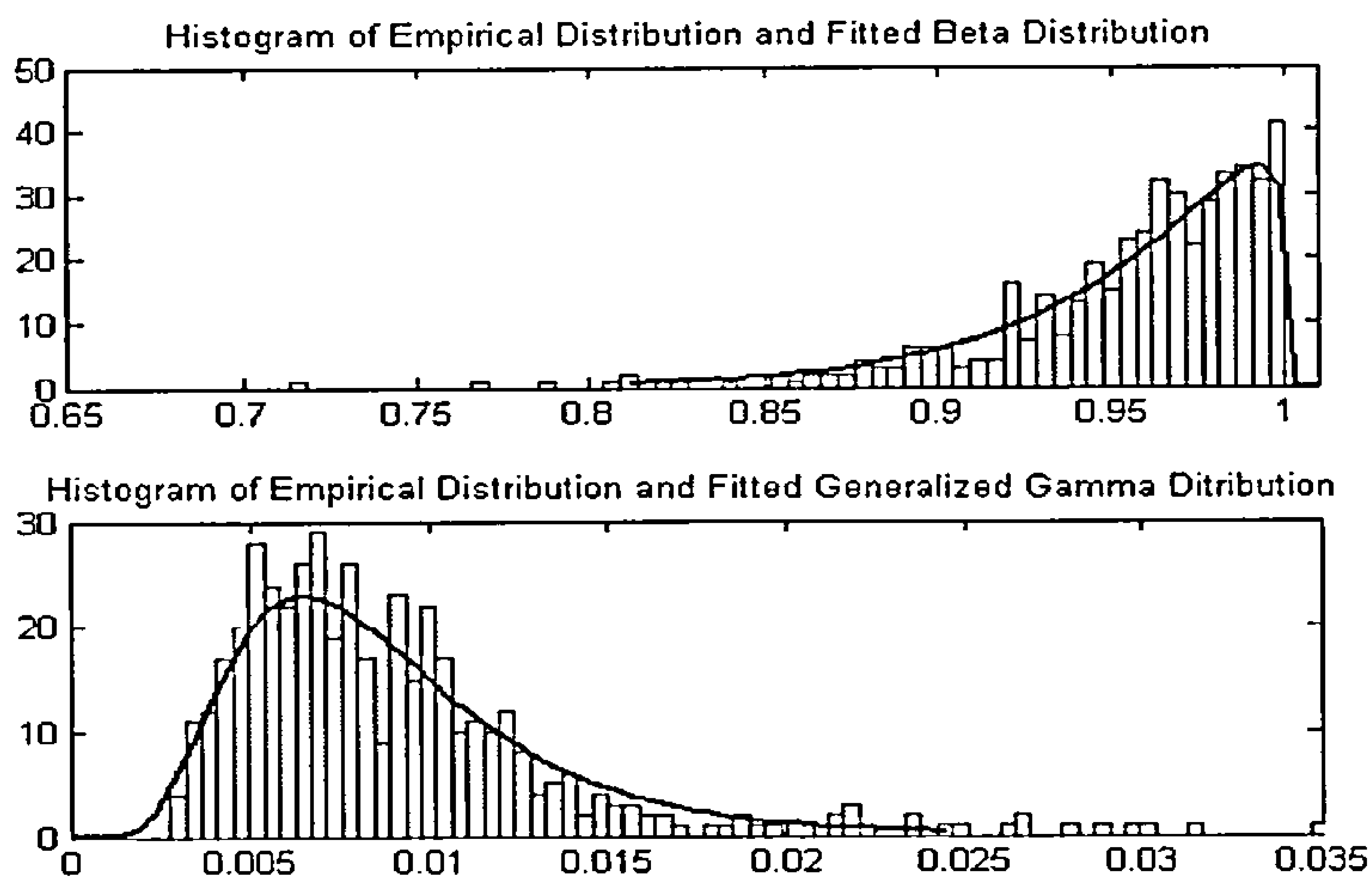

Figure 2.5: Histogram of estimated autoregressive roots and standard deviation and fitted Beta and Generalized Gamma density

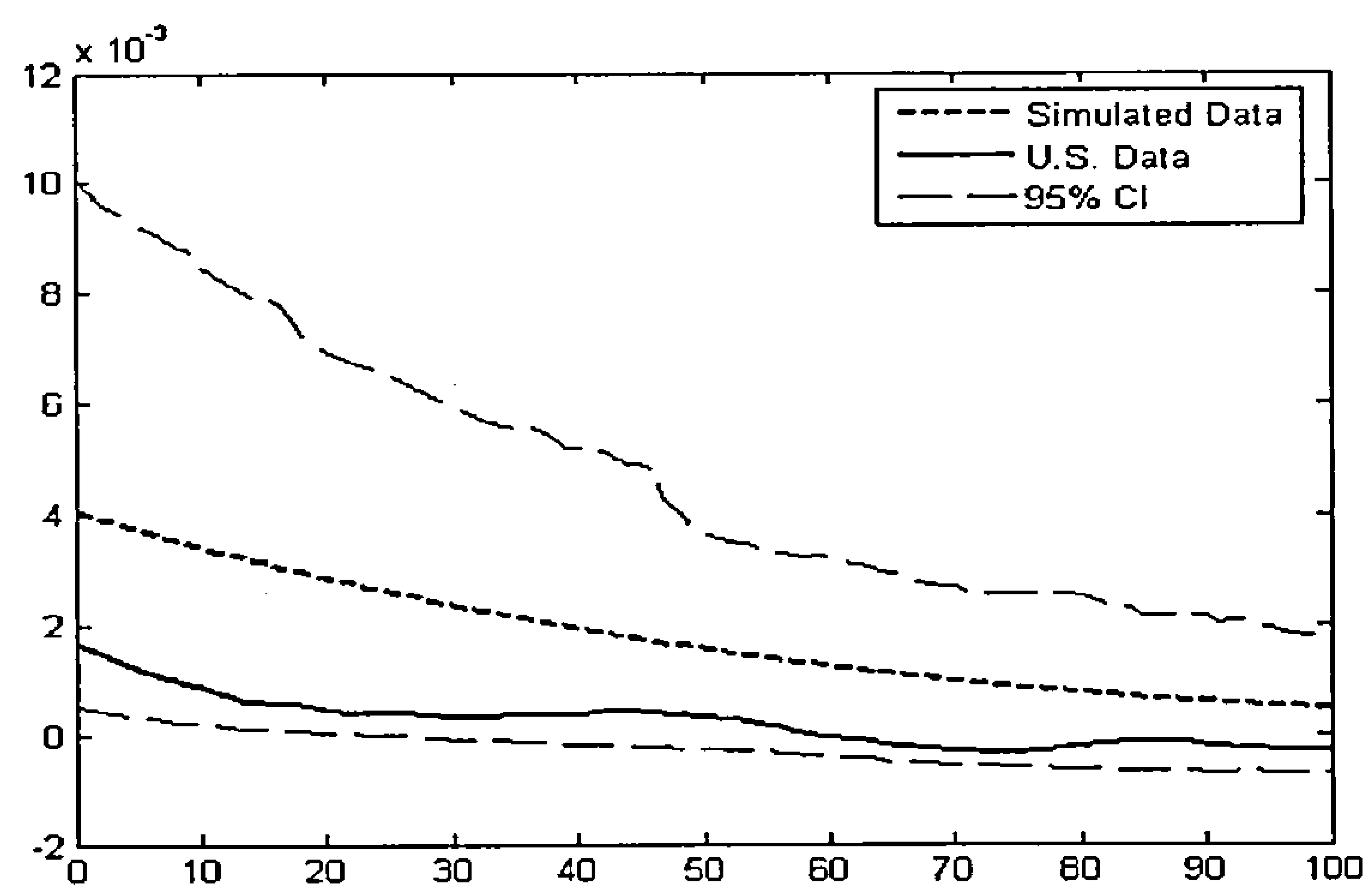

Figure 2.6: Autocovariance function of U.S. GDP and simulated output when no drift is present in the data

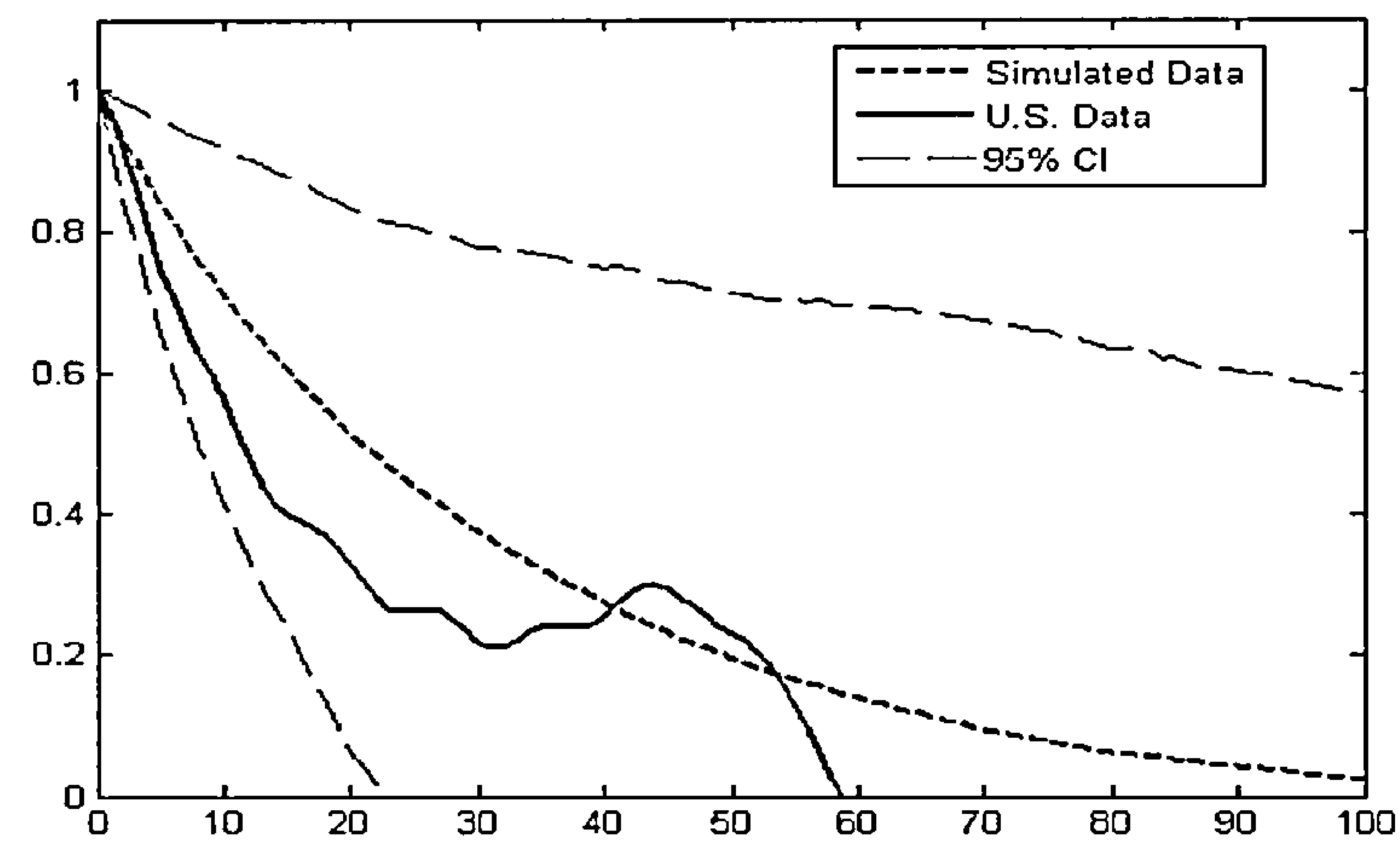

Figure 2.7: Autocorrelation function of U.S. GDP and simulated output when no drift is present in the data

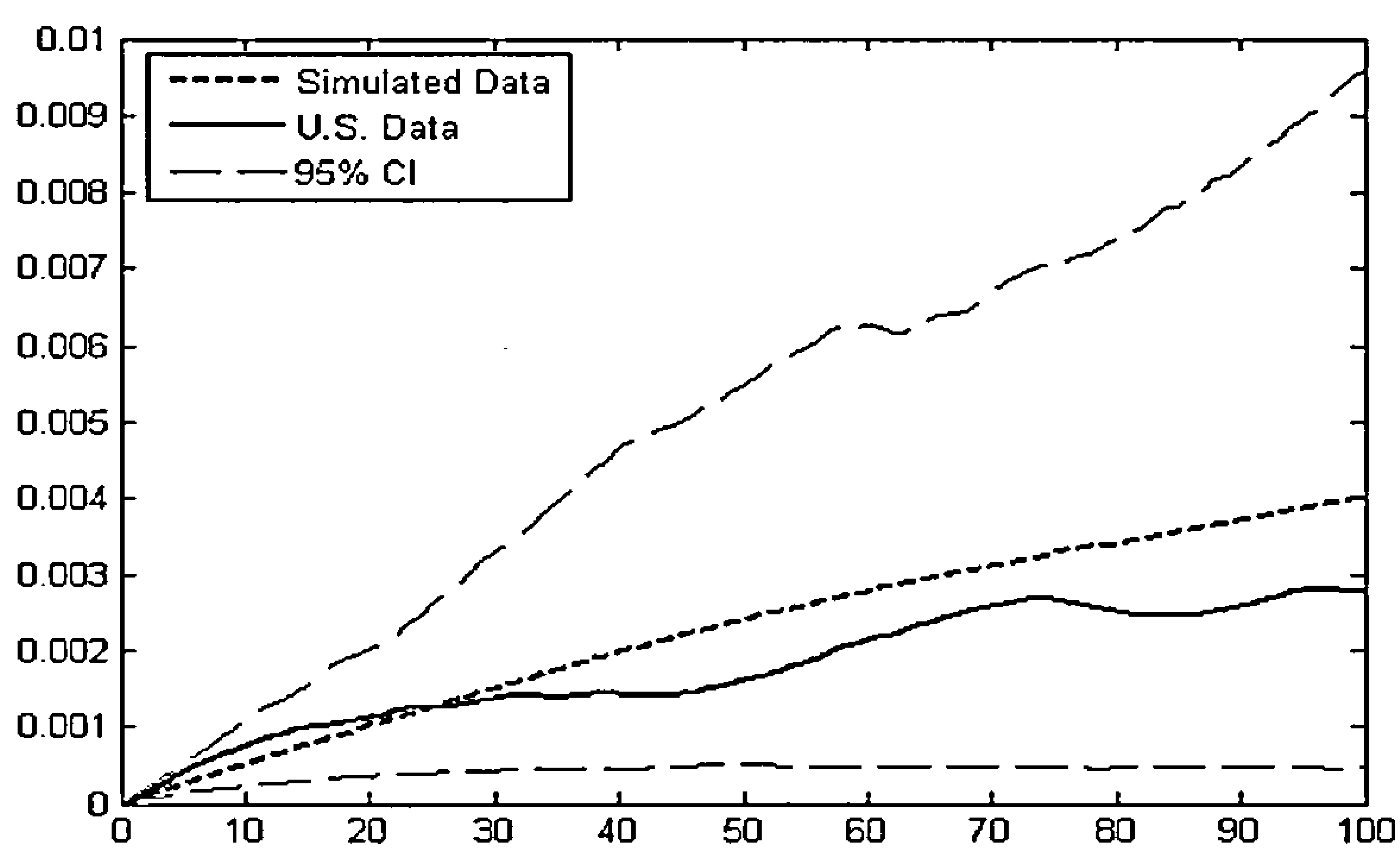

Figure 2.8: Variogram of U.S. GDP and simulated output when no drift is present in the data

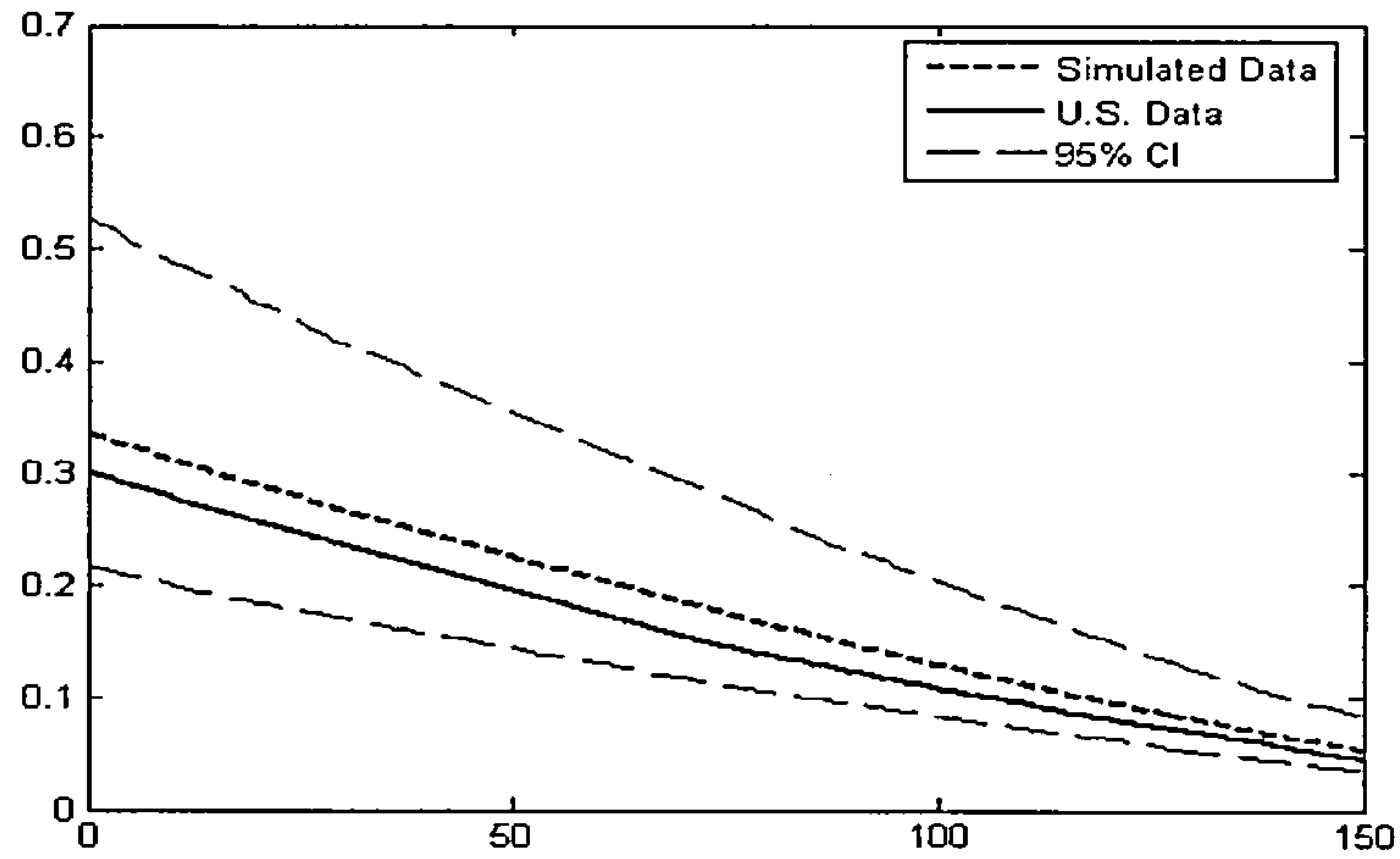

Figure 2.9: Autocovariance function of U.S. GDP and simulated output when drift in present in the data

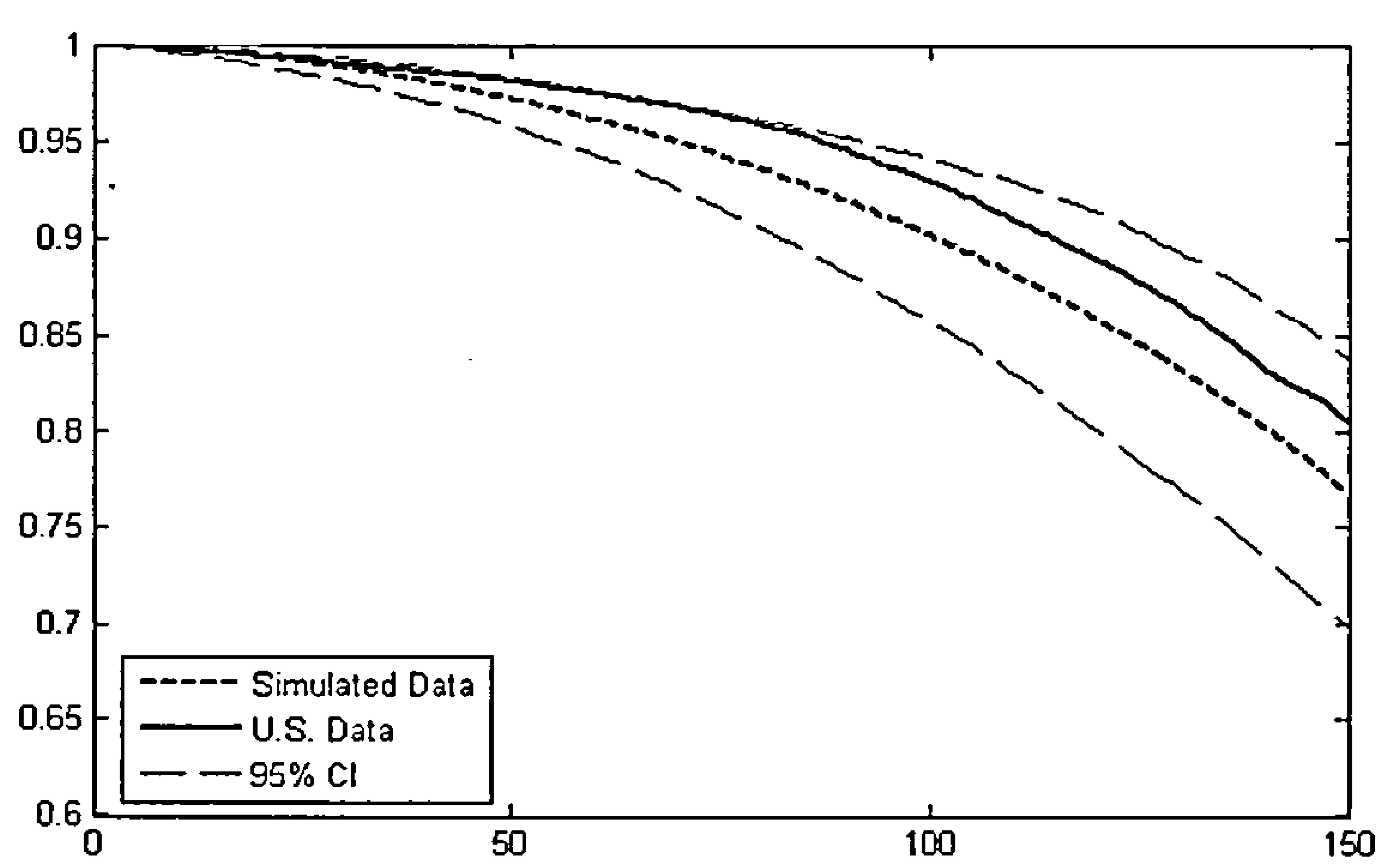

Figure 2.10: Autocorrelation function of U.S. GDP and simulated output when drift in present in the data

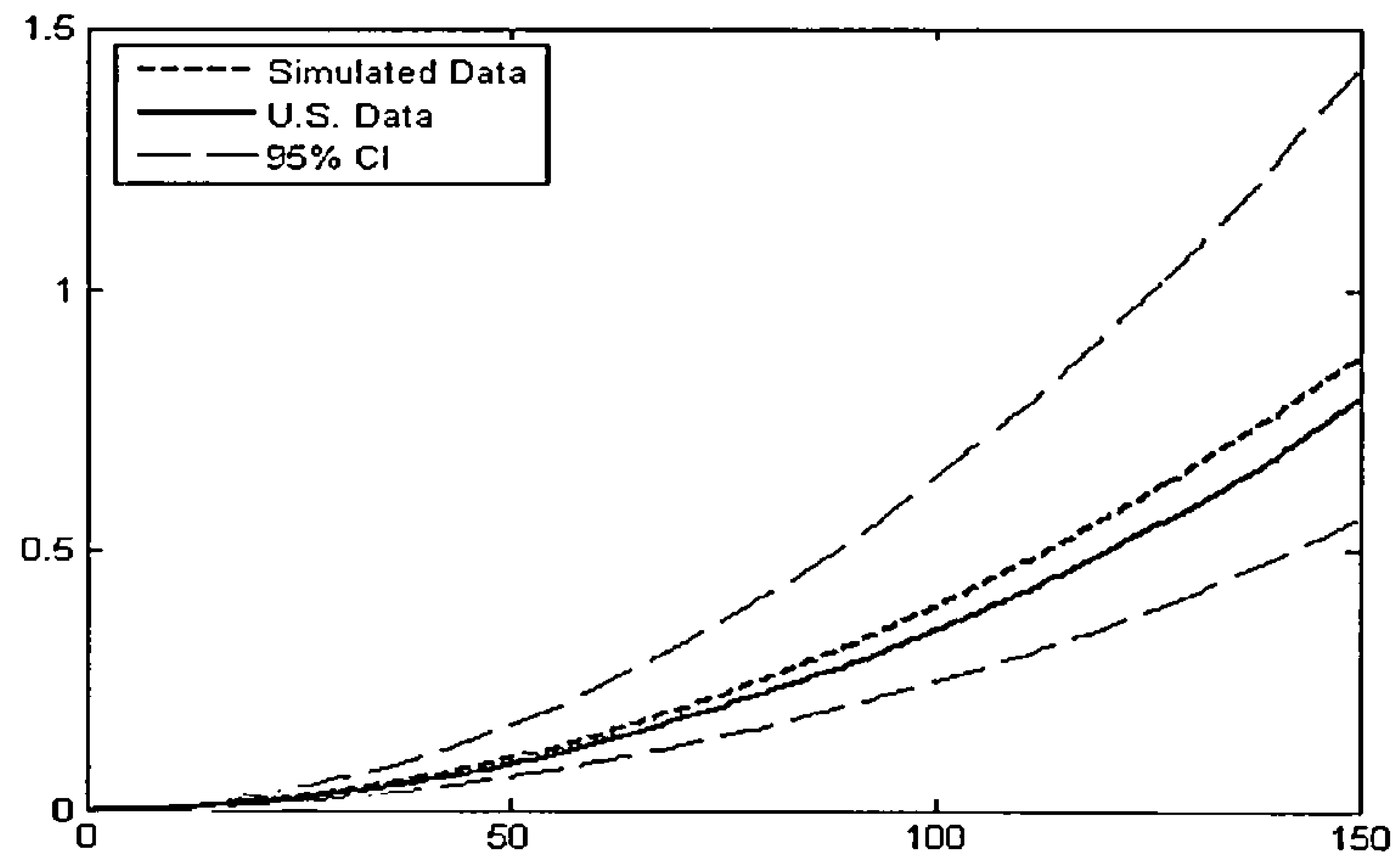

Figure 2.11: Variogram of U.S. GDP and simulated output when drift in present in the data

2.9 Appendix 1

In this appendix we derive the autocorrelation function of $\ln \theta_t$. During the derivation, we make a few uses of the following two lemmas.

Lemma 1 *For x a Generalized Gamma variate with density function*

$$f(x) = \frac{\lambda h_x^{g_x} x^{\lambda g_x - 1}}{\Gamma(g_x)} \exp(-h_x x^\lambda)$$

where $\lambda \in (2, \infty)$ and $g_x \in \Re_+$, then

$$E\left[x^\nu \exp\left(bx^2\right)\right] \simeq \frac{2^{\frac{3}{2}-g_x}\sqrt{\pi}}{\Gamma(g_x)\, b^{\frac{\nu}{2}}} \left(\frac{b^{\frac{\lambda}{2}}}{2h_x}\right)^{\frac{2\nu}{\lambda}+g_x-\frac{1}{2}} \exp\left(\frac{b^2}{(2h_x)^{\frac{4}{\lambda}}} - \frac{b^{\frac{\lambda}{2}}}{4h_x}\right)$$

for $\nu + \lambda g_x \in \Re_+$ and large $b \in \Re_+$.

Proof. See Abadir and Talmain(2002). ∎

Lemma 2 *Let $g(t) \in \Re_+$ be a strictly increasing function of t, then*

$$\int_0^1 \alpha^{g_\alpha - 1}(1-\alpha)^{h_\alpha - 1} \exp(g(t)\,\alpha)d\alpha \simeq \exp(g(t)) \frac{\Gamma(h_\alpha)}{[g(t)]^{h_\alpha}} \tag{2.29}$$

for large t and $h_\alpha > 0$.

Remark 1 *This lemma is of intrinsic mathematical interest since for $g_\alpha = a$, $h_\alpha = b - a$ and $g(t) = t$ it represents an asymptotic expansion for*

$$\int_0^1 \alpha^{b-1}(1-\alpha)^{a-b-1} \exp(t\alpha)d\alpha \equiv \frac{\Gamma(b-a)\,\Gamma(a)}{\Gamma(b)} M(b, a, t)$$

which is the integral representation of a confluent hypergeometric function where $M(b, a, t)$ is a Kummer's function; see pg.505 of Abramowitz and Stegun (1972) for more details.

Proof of Lemma 2. The above lemma can be easily proven by a direct application of Watson's lemma[30] after having replaced α by $(1-\alpha)$ in eq. 2.29. ■

We evaluate the autocorrelation function of $\ln \theta_t$ only for the case when its single components have zero mean (i.e. $\mu = 0$). In fact, this is the only case where the derivation of the ACF of $\ln \theta_t$ is different from the one in the appendix of Abadir and Talmain's paper. Let us recall the AR(1) processes defined above, $\ln \theta_{i,t} = \alpha_i \ln \theta_{i,t-1} + \varepsilon_{i,t}$. Since $\theta_{i,t}^{\nu} = \exp(\nu \sum_{t=1}^{T} \alpha^t \varepsilon_{i,t})$, it is easy to show that

$$\begin{aligned} E(\theta_t) &= \frac{1}{N}\sum_{i=1}^{N} E(\exp(\nu \sum_{t=1}^{T} \alpha^t \varepsilon_{i,t}) \\ &= \frac{1}{N}\sum_{i=1}^{N} (\exp(\frac{1-\alpha_i^{2t}}{2(1-\alpha_i^2)}\nu^2\omega_i^2)) \simeq E_\omega E_\alpha \left(\exp(\frac{1-\alpha_i^{2t}}{2(1-\alpha_i^2)})\nu^2\omega_i^2\right) \end{aligned}$$

where the approximation holds for large N. If we substitute for the Beta density function defined in eq. 2.23, and approximate the exponential term using l'Hopital's rule we get,

$$\begin{aligned} E(\theta_t) &= \frac{1}{B(g_a, h_\alpha)} E_\omega \left(\int_0^1 \alpha^{g_a-1}(1-\alpha)^{h_\alpha-1} \left(\exp(\frac{1-\alpha^{2t}}{2(1-\alpha^2)})\nu^2\omega_i^2\right) d\alpha \right) \\ &\simeq E_\omega \left(\frac{1}{B(g_a, h_\alpha)} \int_0^1 \alpha^{g_a-1}(1-\alpha)^{h_\alpha-1} \left(\exp(\frac{t\alpha^{2(t-1)}}{2})\nu^2\omega_i^2\right) d\alpha \right) \end{aligned}$$

Replacing α by $\alpha^{\frac{1}{2(t-1)}}$ and noticing that $1-\alpha^{\frac{1}{2(t-1)}} \simeq \frac{1-a}{2(t-1)}$ gives

$$\begin{aligned} E(\theta_t) &= E_\omega \left(\frac{1}{B(g_a, h_\alpha) 2(t-1)} \int_0^1 \alpha^{\frac{g_\alpha}{2(t-1)}-1}(1-\alpha^{\frac{1}{2(t-1)}})^{h_\alpha-1} \exp\left((\frac{t\alpha}{2})\nu^2\omega^2\right) d\alpha \right) \\ &\simeq E_\omega \left(\frac{1}{B(g_a, h_\alpha)(2(t-1))^{h_\alpha}} \int_0^1 \alpha^{b-1}(1-\alpha)^{h_\alpha-1} \exp\left((\frac{t\alpha}{2})\nu^2\omega^2\right) d\alpha \right) \end{aligned} \quad (2.30)$$

Applying Lemma 2 with $g(t) = \frac{\nu^2\omega^2}{2} t$ to the integral in eq. 2.30,

$$\begin{aligned} E(\theta_t) &\simeq E_\omega \left(\frac{\Gamma(h_\alpha)}{B(g_a, h_\alpha)(2(t-1))^{h_\alpha}} \left(\frac{1}{\frac{t\nu^2\omega^2}{2}}\right)^{h_\alpha} \exp(\frac{\nu^2 t^2}{2}\omega^2) \right) \\ &= \frac{\Gamma(h_\alpha)}{B(g_a, h_\alpha)(\nu^2 t(t-1))^{h_\alpha}} E_\omega \left(\omega^{-2h_\alpha} \exp(\frac{\nu^2 t^2}{2}\omega^2) \right) \end{aligned}$$

[30] See Keener (1988).

Finally, by lemma1 and the fact that $B(g_\alpha, h_\alpha) = \frac{\Gamma(g_\alpha)\Gamma(h_\alpha)}{\Gamma(g_\alpha+h_\alpha)}$,

$$\begin{aligned}
& E(\theta_t) \\
\simeq\ & \frac{2^{\frac{3}{2}-g_\omega}\sqrt{\pi}\Gamma(g_\alpha+h_\alpha)}{\Gamma(g_\alpha)\Gamma(g_\omega)(\nu^2 t(t-1))^{h_\alpha}} \frac{\left(\frac{\nu^2 t}{2}\right)^{\frac{\lambda}{2}\left(g_\omega-\frac{1}{2}\right)-h_\alpha}}{(2h_\omega)^{g_\omega-\frac{1}{2}-4\frac{h_\alpha}{\lambda}}} \exp\left(\left(\frac{\nu^2 t}{2}\right)^2\left(\frac{1}{2h_\omega}\right)^{\frac{4}{\lambda}} - \left(\frac{\nu^2 t}{2}\right)^{\frac{\lambda}{2}}\left(\frac{1}{4h_\omega}\right)\right)
\end{aligned}$$

Now, since $Cov(\theta_{t+k}, \theta_t) = E(\theta_{t+k}, \theta_t) - E(\theta_{t+k})E(\theta_t)$, we only need to evaluate $E(\theta_{t+k}, \theta_t)$.We can start by approximating the exponential term using l'Hopital rule,

$$\begin{aligned}
& E(\theta_{t+k}, \theta_t) \\
=\ & E_i E_s\left[\exp\left(\frac{1-\alpha_i^{2t}}{2(1-\alpha_i^2)}\nu^2\omega_i^2 + \frac{1-\alpha_s^{2t+k}}{2(1-\alpha_s^2)}\nu^2\omega_s^2 + \frac{1-\alpha_i^t\alpha_s^t}{(1-\alpha_i\alpha_s)}\nu^2\alpha_s^k\omega_s^2\right)\right] \\
\simeq\ & \frac{1}{B(h_\alpha, g_\alpha)^2} E_{\omega_i}[\int_0^1 \alpha_i^{g_\alpha-1}(1-\alpha_i)^{h_\alpha-1}\exp\left(\frac{t\alpha_i^{2(t-1)}}{2}\nu^2\omega_i^2\right) \\
& E_{\omega_i}[\int_0^1 \alpha_s^{g_\alpha-1}(1-\alpha_s)^{h_\alpha-1}\exp\left(\frac{(t+k)\alpha_s^{2(t+k-1)}}{2}\nu^2\omega_s^2 + \nu^2 t\alpha_i^{t-1}\alpha_s^{t+k-1}\omega_i\omega_s\right)]]d\alpha_i d\alpha_s
\end{aligned}$$

Replacing α_i by $\alpha_i^{\frac{1}{2(t-1)}}$ and α_s by $\alpha_s^{\frac{1}{2(t+k-1)}}$ and since $\alpha_i^{\frac{1}{2(t-1)}} \simeq \frac{1-\alpha_i}{2(t-1)}$,

$$\begin{aligned}
& E(\theta_{t+k}, \theta_t) \\
\simeq\ & \frac{1}{B(g_\alpha, h_\alpha)(4(t-1)(t+k-1))} E_{\omega_i}[\int_0^1 \alpha_i^{\frac{g_\alpha}{2(t-1)}-1}\left(1-\alpha_i^{\frac{1}{2(t-1)}}\right)^{h_\alpha-1}\exp\left(\frac{t\alpha_i}{2}\nu^2\omega_i^2\right) \\
& E_{\omega_s}\left(\int_0^1 \alpha_s^{\frac{g_\alpha}{2(t+k-1)}-1}\left(1-\alpha_s^{\frac{1}{2(t+k-1)}}\right)^{h_\alpha-1}\exp\left(\frac{(t+k)\alpha_s}{2}\nu^2\omega_s^2 + \nu^2 t\alpha_i^{\frac{1}{2}}\alpha_s^{\frac{1}{2}}\omega_i\omega_s\right)\right)]d\alpha_i d\alpha_s \\
\simeq\ & \frac{1}{B(g_\alpha, h_\alpha)(4(t-1)(t+k-1))^{h_\alpha}} E_{\omega_i}[\int_0^1 \alpha_i^{\frac{g_\alpha}{2(t-1)}-1}(1-\alpha_i)^{h_\alpha-1}\exp\left(\frac{t\alpha_i}{2}\nu^2\omega_i^2\right) \\
& E_{\omega_s}\left(\int_0^1 \alpha_s^{\frac{g_\alpha}{2(t+k-1)}-1}(1-\alpha_s)^{h_\alpha-1}\exp\left(\frac{(t+k)\alpha_s}{2}\nu^2\omega_s^2 + \nu^2 t\alpha_s^{\frac{1}{2}}\alpha_i^{\frac{1}{2}}\omega_i\omega_s\right)\right)]d\alpha_i d\alpha_s
\end{aligned}$$

Using the Lemma 2, the quantity inside the expectation $E_{w_s}(\cdot)$ can be approximated by

$$\begin{aligned}
& \int_0^1 \alpha_s^{\frac{g_\alpha}{2(t+k-1)}-1}(1-\alpha_s)^{h_\alpha-1}\exp\left(\frac{(t+k)\alpha_s}{2}\nu^2\omega_s^2 + \nu^2 t\alpha_i^{\frac{1}{2}}\alpha_s^{\frac{1}{2}}\omega_i\omega_s\right)]d\alpha_s \\
\simeq\ & \frac{\Gamma(h_\alpha)}{\left(\frac{t+k}{2}\nu^2\omega_s^2\right)^{h_\alpha}}\exp\left(\left(\frac{t+k}{2}\nu^2\omega_s^2\right) + \nu^2 t\alpha_i^{\frac{1}{2}}\omega_i\omega_s\right)
\end{aligned}$$

Hence,

$$
\begin{aligned}
& E(\theta_{t+k}, \theta_t) \\
\simeq\ & \frac{\Gamma(h_\alpha)}{B(g_\alpha, h_\alpha)(4(t-1)(t+k-1))^{h_\alpha}} E_{\omega_i}[\int_0^1 \alpha_i^{\frac{g_\alpha}{2(t-1)}-1}(1-\alpha_i)^{h_\alpha-1}\exp\left(\frac{t\alpha_i}{2}\nu^2\omega_i^2\right) \\
& E_{\omega_s}\left(\left(\frac{t+k}{2}\nu^2\omega_i^2\right)^{\frac{g_\alpha}{2(t-1)}-h_\alpha}\exp\left(\left(\frac{t+k}{2}\nu^2\omega_i^2\right)+\nu^2 t\alpha_i^{\frac{1}{2}}\omega_i\omega_s\right)\right)]d\alpha_i \\
\simeq\ & \frac{B(g_\alpha, h_\alpha)^{-1}\Gamma(h_\alpha)}{(4(t-1)(t+k-1))^{h_\alpha}} E_{\omega_i}E_{\omega_s}\left[\int_0^1 \alpha_i^{\frac{g_\alpha}{2(t-1)}-1}(1-\alpha_i)^{h_\alpha-1}\exp\left(\frac{t\alpha_i}{2}\nu^2\omega_i^2+\nu^2 t\alpha_i^{\frac{1}{2}}\omega_i\omega_s\right)\right. \\
& \left.\left(\frac{t+k}{2}\nu^2\omega_i^2\right)^{-h_\alpha}\exp\left(\left(\frac{t+k}{2}\nu^2\omega_i^2\right)\right)\right] d\alpha_i
\end{aligned}
$$

In the same way the integral with respect to α_i can be approximated by

$$
\begin{aligned}
& \int_0^1 \alpha_i^{\frac{g_\alpha}{2(t-1)}-1}(1-\alpha_i)^{h_\alpha-1}\exp\left(\frac{t\alpha_i}{2}\nu^2\omega_i^2+\nu^2 t\alpha_i^{\frac{1}{2}}\omega_i\omega_s\right) d\alpha_i \\
\simeq\ & \frac{\Gamma(h_\alpha)}{\left(\frac{t}{2}\nu^2\omega_i^2\right)^{h_\alpha}}\exp\left(\left(\frac{t}{2}\nu^2\omega_i^2\right)+\nu^2 t\omega_i\omega_s\right)
\end{aligned}
$$

and

$$
\begin{aligned}
& E(\theta_{t+k},\theta_t) \\
\simeq\ & \frac{\Gamma(h_\alpha)^2}{B(g_\alpha,h_\alpha)\left(4(t-1)(t+k-1)\right)^{h_\alpha}} E_{\omega_i}E_{\omega_s}\left[\left(\frac{t}{2}\nu^2\omega_i^2\right)^{-h_\alpha}\exp\left(\left(\frac{t}{2}\nu^2\omega_i^2\right)+\nu^2 t\omega_i\omega_s\right)\right. \\
& \left.\left(\frac{t+k}{2}\nu^2\omega_i^2\right)^{-h_\alpha}\exp\left(\left(\frac{t+k}{2}\nu^2\omega_i^2\right)\right)\right] \\
\simeq\ & \frac{\Gamma(h_\alpha)^2}{B(g_\alpha,h_\alpha)\left(4t(t-1)(t+k)(t+k-1)\right)^{h_\alpha}} E_{\omega_i}E_{\omega_s}\left[\left(\omega_i^2\right)^{-h_\alpha}\exp\left(\left(\frac{t}{2}\nu^2\omega_i^2\right)+\nu^2 t\omega_i\omega_s\right)\right. \\
& \left.\left(\omega_i^2\right)^{-h_\alpha}\exp\left(\left(\frac{t+k}{2}\nu^2\omega_i^2\right)\right)\right] \\
\simeq\ & \frac{\Gamma(h_\alpha)^2}{B(g_\alpha,h_\alpha)\left(4t(t-1)(t+k)(t+k-1)\right)^{h_\alpha}} E_{\omega_i}E_{\omega_s}\left[\left(\omega_i^2\right)^{-h_\alpha}\exp\left(\frac{t}{2}\nu^2\omega_i^2\right)\sum_{j=0}^{\infty}\frac{(\nu^2 t\omega_i\omega_s)^j}{j!}\right. \\
& \left.\left(\omega_i^2\right)^{-h_\alpha}\exp\left(\left(\frac{t+k}{2}\nu^2\omega_i^2\right)\right)\right] \\
=\ & \frac{\Gamma(h_\alpha)^2}{B(g_\alpha,h_\alpha)\left(t(t-1)(t+k)(t+k-1)\right)^{h_\alpha}} \sum_{j=0}^{\infty}\frac{(\nu^2 t)^j}{j!} E_{\omega_i}E_{\omega_s}\left[\omega_i^{-2h_\alpha+j}\exp\left(\frac{t}{2}\nu^2\omega_i^2\right)\right. \\
& \left.\omega_i^{-2h_\alpha+j}\exp\left(\left(\frac{t+k}{2}\nu^2\omega_i^2\right)\right)\right]
\end{aligned}
$$

where the infinite summation has been obtained using a binomial expansion.

We can evaluate the expectations with respect to ω_i and ω_s using the lemma 1,

$$
\begin{aligned}
& E(\theta_{t+k},\theta_t) \\
\simeq\ & \frac{\Gamma(h_\alpha)^2 4^{\frac{3}{2}-g_\omega}\pi}{B(g_\alpha,h_\alpha)\Gamma(g_\omega)^2(t(t-1)(t+k)(t+k-1))^{h_\alpha}} \qquad (2.31)\\
& \frac{\left[\left(\frac{t\nu^2}{2}\right)\left(\frac{(t+k)\nu^2}{2}\right)\right]^{\frac{\lambda}{2}\left(g_\omega-\frac{1}{2}\right)-h_\alpha}}{\left(\frac{1}{4h_\omega^2}\right)^{\frac{-4h_\alpha}{\lambda}+g_\omega-\frac{1}{2}}} \exp\left(\left(\frac{t\nu^2}{2}\right)^2\left(\frac{1}{2h_\omega}\right)^{\frac{4}{\lambda}}-\left(\frac{t\nu^2}{2}\right)^{\frac{\lambda}{2}}\left(\frac{1}{4h_\omega}\right)+\right.\\
& \left.+\left(\frac{(t+k)\nu^2}{2}\right)^2\left(\frac{1}{2h_\omega}\right)^{\frac{4}{\lambda}}-\left(\frac{(t+k)\nu^2}{2}\right)^{\frac{\lambda}{2}}\left(\frac{1}{4h_\omega}\right)\right)\sum_{j=0}^{\infty}\frac{(\nu^2 t)^j}{j!}\left(\frac{\left(\frac{(t+k)\nu^2}{2}\right)^{\frac{1}{2}}}{\left(\frac{1}{2h_\omega}\right)^{\frac{2}{\lambda}}}\frac{\left(\frac{t\nu^2}{2}\right)^{\frac{1}{2}}}{\left(\frac{1}{2h_\omega}\right)^{\frac{2}{\lambda}}}\right)^j \\
=\ & \frac{\Gamma(h_\alpha)^2 4^{\frac{3}{2}-g_\omega}\pi}{B(g_\alpha,h_\alpha)\Gamma(g_\omega)^2(t(t-1)(t+k)(t+k-1))^{h_\alpha}}\frac{\left[\left(\frac{t\nu^2}{2}\right)\left(\frac{(t+k)\nu^2}{2}\right)\right]^{\frac{\lambda}{2}\left(g_\omega-\frac{1}{2}\right)-h_\alpha}}{\left(\frac{1}{4h_\omega^2}\right)^{\frac{-4h_\alpha}{\lambda}+g_\omega-\frac{1}{2}}} \\
& \exp\left(\left(\frac{1}{2h_\omega}\right)^{\frac{4}{\lambda}}\left[\frac{\nu^4\left[t^2+(t+k)^2\right]}{4}\right]-\left(\frac{1}{4h_\omega}\right)\left(\frac{\nu^\lambda\left[t^{\frac{\lambda}{2}}+(t+k)^{\frac{\lambda}{2}}\right]}{2^{\frac{\lambda}{2}}}\right)\right) \qquad (2.32)\\
& \exp\left(\frac{\nu^4\sqrt{t^3(t+k)}}{2(4h_\omega^2)^{\frac{2}{\lambda}}}\right) \qquad (2.33)
\end{aligned}
$$

and recalling $E(Z_t)$ and $E(Z_{t+k})$ we get

$$
\begin{aligned}
Cov(\theta_{t+k},\theta_t) &= E(\theta_{t+k}\theta_t)-E(\theta_{t+k})E(\theta_t)\\
&\simeq \frac{\pi 4^{\frac{3}{2}-g_\omega}}{(\nu^4 t(t-1)(t+k)(t+k-1))^{h_\alpha}}\left(\frac{\Gamma(g_\alpha+h_\alpha)}{\Gamma(g_\alpha)\Gamma(g_\omega)}\right)^2\frac{\left(\frac{\nu^4 t(t+k)}{4}\right)^{\frac{\lambda}{2}\left(g_\omega-\frac{1}{2}\right)-h_\alpha}}{(4h_\omega^2)^{g_\omega-\frac{1}{2}-4\frac{h_\alpha}{\lambda}}}\\
&\quad \exp\left(\left(\frac{1}{2h_\omega}\right)^{\frac{4}{\lambda}}\left[\frac{\nu^4\left[t^2+(t+k)^2\right]}{4}\right]-\left(\frac{1}{4h_\omega}\right)\left(\frac{\nu^\lambda\left[t^{\frac{\lambda}{2}}+(t+k)^{\frac{\lambda}{2}}\right]}{2^{\frac{\lambda}{2}}}\right)\right)\\
&\quad \left(\exp\left(\frac{\nu^4\sqrt{t^3(t+h)}}{2(4h_\omega^2)^{\frac{2}{\lambda}}}\right)-1\right)
\end{aligned}
$$

Now the autocorrelation function can be obtained as $Cor(\theta_{t+k},\theta_t) = \frac{Cov(\theta_{t+k},\theta_t)}{\sqrt{var(\theta_{t+k})var(\theta_t)}}$,

$$Cor(\theta_{t+k},\theta_t) = \frac{\left(\exp\left(\frac{\nu^4\sqrt{t^3(t+h)}}{2(4h_\omega^2)^{\frac{2}{\lambda}}}\right)-1\right)}{\sqrt{\left(\exp\left(\frac{\nu^4 t^2}{2(4h_\omega^2)^{\frac{2}{\lambda}}}\right)-1\right)\left(\exp\left(\frac{\nu^4(t+k)^2}{2(4h_\omega^2)^{\frac{2}{\lambda}}}\right)-1\right)}}$$

It has to be noticed that the formula for the autocorrelation function is the same as the one in Abadir and Talmain.

This can be explained by the fact that in Abadir and Talmain's set up, neither μ nor the parameters of distribution of variance of the idiosyncratic shocks affect the autocorrelation function of θ_t. Finally, the leading term of the ACF of $\ln\theta_t$ in eq. 2.25 can be obtained by applying the small variance expansions in section 3.2 of Abadir and Talmain (2002).

2.10 Appendix 2

In this appendix we present the small sample properties of the variogram and the autocorrelation function for persistence processes. As already mentioned, a detailed description of the statistical properties of the variogram, which goes beyond the purpose of this paper, can be found in Haslett (1997) and Beran (1998). In this section, using simulations we show that for very persistent processes the sample variogram has a faster convergency rate than the sample autocorrelation function. As mentioned above the variogram is defined for a time series $\{x_t\}_{t=1}^T$ as

$$VR(x_t,x_{t+k}) = \frac{E(x_{t+k}-x_t)^2}{2}$$

It is easy to show that for strictly stationary processes it is directly related to the autocovariance function $c(k)$ and to the autocorrelation function $r(k)$ by

$$\begin{aligned} VR(x_t,x_{t+k}) &= c(0)\left(1-\frac{c(k)}{c(0)}\right) \\ &= c(0)(1-r(k)) \end{aligned}$$

When dealing with linear processes with root close to one, the estimation of the autocorrelation function can be very biased, even in reasonably large samples (more than 200 observations). In fact, as the root of the process approaches one the sample moments converge at a slower rate to the population moments. This bias becomes dramatically large for random walks or when long memory is present in the data (see Newbold and Agiakloglou (1997)).

This problem is shown in the figure below, where we plot the theoretical and sample autocorrelation functions for two AR(1) processes[31] with roots respectively equal to 0.95 and 1.

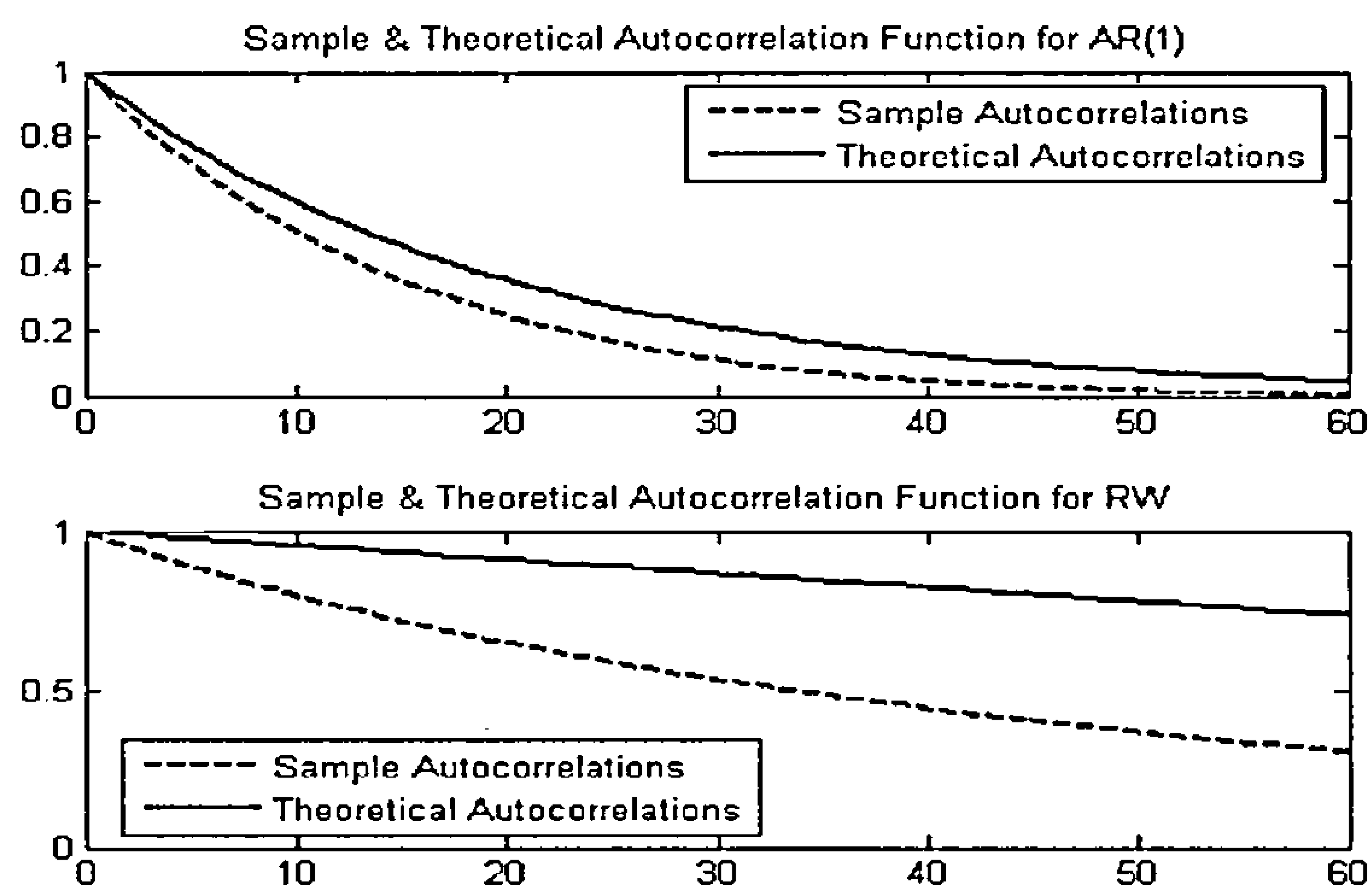

Figure 2.12: Sample and theoretical autocorrelation function of the AR(1) process and random walk

[31]The artificial series have been generated using the following data generating process:

$$x_t = \alpha x_{t-1} + \varepsilon_t \qquad \varepsilon_t \sim IN\,(0, 0.01)$$

with $x_0 = 0$ and $t = 1, ..., 228$.

The sample autocorrelations have been calculated averaging over 1000 replications.

As can be readily seen, as the root of the process approaches one, the bias in the ACF of the simulated data becomes extremely large. This result can have some implications when comparing the dynamics of the simulated data with actual data. In fact, it is common practice in macroeconomics to start from the autocorrelations structure of some economic variables, make some assumptions on their data generating process and then try to mimic their autocorrelations with the model simulated data. In the light of the above results, if the data are very persistent, the ACF of the simulated data will be biased even if the data generating process is correctly specified.

In the next plot we present the same result for the variogram. As it can be readily seen, the sample and theoretical variogram are indistinguishable at all lags and the reduction in the bias is remarkable.

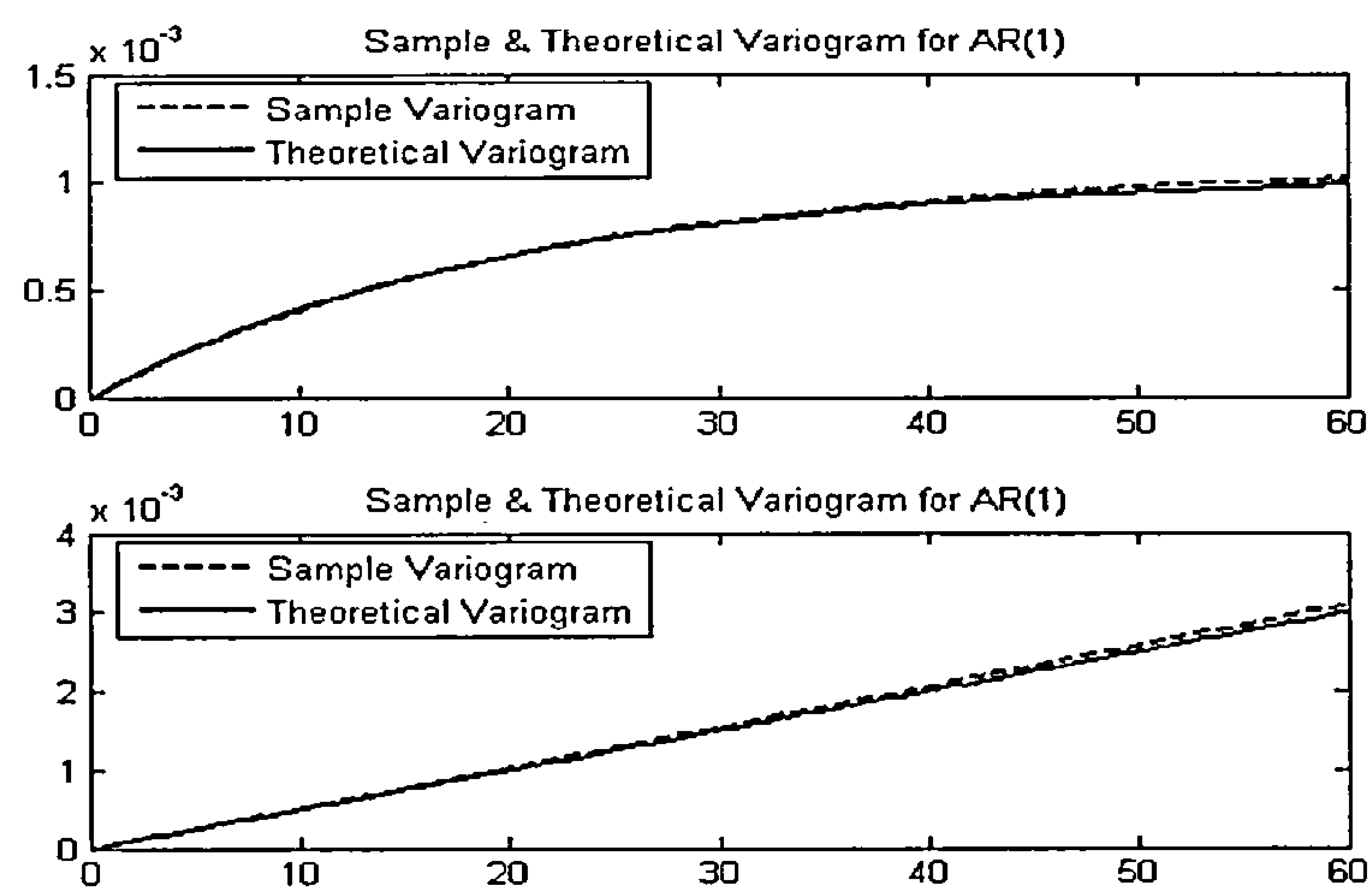

Figure 2.13: Theoretical and Sample Variogram for AR(1) process and random walk

Although the variogram for the aggregate process θ_t, has not been derived yet[32], we expect its sample variogram to be unbiased in the light of above results. In fact, even if the process has long memory and therefore is more persistent than the AR(1)

[32]Given the structure of the process it would be very difficult to derive it analytically.

process considered above, it still has less memory than a random walk. Since the sample variogram of the random walk is unbiased, as shown in figure 2.13, this implies that the variogram of the long memory process θ_t should also be unbiased.

Chapter 3

Long memory co-movements in macroeconomic time series: a modified E-G approach

3.1 Introduction

Since the seminal papers of Granger (1981) and Engle and Granger (1987) the concepts of integration and cointegration have developed in many areas of both econometrics and applied macroeconomics. By a well-known definition two time-series $x_{1,t}$ and $x_{2,t}$ are said to be cointegrated of order $CI(\delta, b)$ if they are individually integrated of order $I(\delta)$ and there exists a linear combination $\varepsilon_t = x_{1,t} - \beta x_{2,t}$ that is integrated of order $I(\delta - b)$.

In the past years, many approaches have been proposed to test for cointegration; see Watson (1995) for a survey on the these approaches. In particular, they have been designed for the case when $\delta = 1$ and $b = 1$. Under this assumption, $x_{1,t}$ and $x_{2,t}$ are $I(1)$ variables and they are cointegrated if there exists a linear combination ε_t that is $I(0)$. This case is very appealing for applied economist since it allows us to estimate long run steady states as linear combination of non stationary variables. Furthermore, movements from the steady state equilibrium can be represented using standard ARMA models.

Recently, this notion of cointegration has been criticized by a number of researchers who have claimed that the distinction between $I(0)$ and $I(1)$ is rather arbitrary. They have proposed instead to allow ε_t to be integrated of order $I(d)$ with $0 \leq d < 1$ (i.e. fractionally Integrated) or more generally to belong to the class of long memory processes.

Long memory cointegration implies that although there is an equilibrium relation between economic variables spanning the long-run, these variables can be away from such equilibrium for a very long period of time.

Standard cointegration techniques cannot be applied in this context since they cannot distinguish between long-memory co-movements and spurious relations. For instance, the two most popular cointegration approaches, the Engle-Granger (E-G) two-step procedure and the Johansen's full-Information maximum likelihood[1] (FIML) cannot deal with the hypothesis of long-memory cointegration. On one side, as Gonzalo and Lee (1998) and Anderson and Gredenhoff (1994) show, the FIML is not robust to dynamic misspecification and gives rise to pitfalls and spurious regressions if the dynamics of long-run regression residuals ε_t departs from the $I(1)$ assumption. On the other side, Diebold and Rudebusch (1991b), Lee and Schmidt (1996), Hassler and Wolters (1994) and Gonzalo and Lee (1998) show that the ADF test, which plays a crucial role in the E-G approach, has low power against the alternative of long memory despite its asymptotic consistency[2]. In particular, testing for individual unit root is not enough if the data generating process has long-memory or dynamic behaviors different from those of a unit root process.

[1] See Johansen (1988 and 1995).

[2] For the case where ε_t is a fractional ARMA process, few approaches have been proposed to take this type of alternative explicitly into consideration. A review of this subject can be found in Dolado, Gonzalo and Mayoral (2002). The approach proposed here does not force ε_t to belong strictly to the class of FARIMA process.

In this chapter we present a methodology to test for the presence of cointegration when two variables are $I(1)$ and there exists a linear combination that behaves as a long memory process. The contribution of this chapter is twofold. First, differently from standard cointegration technique, our approach is able to detect the presence of long-memory co-movements. Then, the test we propose is characterized by high size under the null (of spurious relation) and high power under the alternative (long memory relation). Furthermore, it also reduces the small sample bias in the estimation of the long-run relation by more than 41% compared to ordinary least square (OLS).

We apply our procedure to the Johansen and Juselius (1992) data base for the UK purchasing-power parity (PPP) and uncovered interest rate parity (UIP) and show that the null of no cointegration is rejected at 95% in contradiction with what was previously shown with standard single equation techniques, see Harris (1992).

Evidence of long memory in the co-movements of many macroeconomic variables was already found by Cheung and Lai (1993), Diebold et. al. (1991) and Abadir and Talmain (2005). On one side, Cheung and Lai suggest that deviations from the PPP equilibrium could follow a mean reverting long memory process. On the other, Abadir and Talmain show that most economic variables are characterized by a high degree of persistence and nonlinearities that, if not properly accounted for, can bias standard econometric techniques and give rise to economic puzzles and counterintuitive results. To this purpose, they propose a method to disentangle the relations between variables from the effects of their persistence. In this paper we go beyond their findings and propose a procedure to test for such long-memory co-movements and show that null of no cointegration is rejected at 95% for the PPP-UIP database.

Our approach combines the Engle and Granger cointegration approach with the quasi-maximum-likelihood (QML) procedure for long memory process introduced by Abadir and Talmain (2005).

The difference between our approach and standard cointegration technique can be understood by analyzing the assumption that characterized the two approaches.

Testing with standard cointegration techniques imposes very strict conditions on the long-run relation among the variables: first, it assumes that the relation is strictly linear, then the variables must adjust towards this equilibrium at a relatively fast rate ($I(0)$ hypothesis). Therefore, it should not be very surprising that such "cointegration" is rejected even when the long-run relation between variables seems economically plausible (as the PPP-UIP theorem). What is really important when testing for cointegration is not stationarity but mean reversion toward the long run equilibrium. Strict stationarity is a sufficient but not necessary condition to have mean reversion. In fact, there exist many processes that are not stationary but still mean reverting[3].

Conversely, the approach we propose is able to test for the existence of a long-run relation, whilst allowing for possible non-linearities and persistent deviations from its long-run that are not forced to be strictly stationary.

This chapter is organized as follows. In the next section, we recall the Engle and Granger approach and describe the moving block and the stationary bootstrap. In section 3, we test for cointegration for the PPP-UIP database using standard E-G approach. In section 4, we describe our modified testing procedure. Then, in section 5, we apply our procedure to the PPP-UIP to test for the presence of a long memory equilibrium relation. Finally, in section 6, we run some simulations to calculate the empirical size and power of our testing procedure and the reduction in the small sample estimation bias given by the QML estimator relatively to standard OLS estimator.

3.2 Testing for cointegration and Bootstrap ADF-

[3]For instance, any process which is $I(d)$ with $d \in (0, 1)$ would fall into this category. They are called fractionally integrated processes and details on their properties can be found in Beran (1998) and Hosking (1981).

test

In this section we recall the E-G cointegration's approach and describe how to calculate stationary bootstrap (SB) and moving block bootstrap (MBB) confidence intervals, under the null of no cointegration.

The E-G approach is a very intuitive two step procedure. In the first step we estimate by ordinary least square the relation

$$x_{1,t} = \beta x_{2,t} + \varepsilon_t \tag{3.1}$$

also called cointegrating vector, while in the second we test whether the regression residuals ε_t are strictly stationary. To this purpose, we run the regression

$$\Delta\varepsilon_t = \rho\varepsilon_{t-1} + \rho_1\Delta\varepsilon_{t-1} + ... + \rho_k\Delta\varepsilon_{t-k} + u_t \tag{3.2}$$

and construct the t-value statistic t_ρ for the estimated parameter $\hat{\rho}$, which is called ADF statistic of order k or $ADF(k)$. If we reject the hypothesis that $\hat{\rho} = 0$, then ε_t has an ARMA representation and the variables $x_{1,t}$ and $x_{2,t}$ are cointegrated with cointegrating vector $[1, \ -\beta]$. Otherwise, if we fail to reject the hypothesis $\hat{\rho} = 0$ then ε_t is non stationary and eq. 3.1 is a spurious relation.

It is a well-known result that the ADF statistics converges, under the null of no-cointegration, to a non-standard distribution. The critical values for this distribution have been derived by Said and Dickey (1984) using simulation. These critical values, as well as those of other cointegration approaches, are justified on asymptotic grounds. It is well-known that in a context of cointegration regression models, asymptotic critical values are not very reliable unless the sample size is very large[4]. This implies that testing for cointegration in small samples can give rise to substantial estimation bias as well as size distortion in the associated tests of significance. A solution to this problem has been proposed by Li and Maddala (1997) who suggest that bootstrap

[4]See Li and Maddala (1997) for more details.

methods can provide some corrections, in the E-G approach, to both estimation bias and size distortions. Using various type of bootstrap methods, in particular moving block bootstrap and stationary bootstrap, they show the superiority of bootstrap-based critical values over asymptotic critical values. On the basis of their results we decided to use bootstrapped rather than asymptotic critical values in all the ADF tests shown below.

We now describe briefly the moving block and stationary bootstrap which will be implemented in the next sections, and also discuss some general issues about their implementation in testing for the significance of the parameter ρ in eq. 3.2.

A complete exposition of the statistical properties of the bootstrap can be found in Hall (1992), Efron and Tibshirani (1992) and Shao and Tu (1995). The idea behind the bootstrap is the following. Let us consider a sample of i.i.d. variables $\{x_1, ..., x_n\}$ with underlying distribution F and a parameter of the population $\theta(F)$ on which we want to make inference. Let us define $\hat{\theta}(F)$ the estimated parameter from $\{x_1, ..., x_n\}$. Then, the bootstrap distribution of $\hat{\theta}(F)$ can be derived by re-sampling with replacement from $\{x_1, ..., x_n\}$ and by calculating from each re-sample the parameter $\tilde{\theta}(F)$. This generates a distribution $\tilde{F}$ of parameters $\tilde{\theta}(F)$ that provides, under general conditions, an approximation of true distribution F.

In time series analysis data are generally not i.i.d., therefore different approaches, such as moving block bootstrap (MMB) and stationary bootstrap (SB), have been proposed to capture the dependence structure of the data. The moving block bootstrap was introduced by Carlstein (1986) who considers non-overlapping blocks and further developed by Künsch (1989) who proposed instead overlapping blocks. In particular, given a time series sample $\{x_1, ..., x_n\}$, Künsch proposed to construct $n - l + 1$ blocks of data of length l, $B_j = \{x_j, x_{j+1}, ..., x_{j+l-1}\}$, $j = 1, ..., n - l + 1$ and to resample with replacement from those blocks. A different type of block bootstrap is the stationary bootstrap where the block length l is sampled from the geometric distribution $P(l = m) = (1-p)^{m-1} p$ with $m = 1, 2, ...$ and $p \in (0, 1)$, while the starting date j of the first observation of the block is chosen according to a uniform distribution on $[1, n]$. If $j + l - 1$ exceed the index n of the last observation x_n, then the block is constructed as $B_j = \{x_j, ..., x_n, x_{1,...,}x_{l-n+j-1}\}$.

This kind of bootstrap has been introduced by Politis and Romano (1994) after discovering that the time series generated by the MBB bootstrap may not be stationary even if the original series $\{x_1, ..., x_n\}$ is stationary.

We proceed now with the description step by step of the algorithm to calculate bootstrap critical values for the ADF t_ρ-statistics, under the null hypothesis of no cointegration:

1) Estimate the cointegrating vector $y_t = \alpha + \beta x_t + \varepsilon_t$ by OLS and get the regression residual $\hat{\varepsilon}_t = y_t - \hat{\alpha} + \hat{\beta} x_t$

2) Run the ADF(k) regression

$$\Delta\hat{\varepsilon}_t = \rho\hat{\varepsilon}_{t-1} + \rho_1\Delta\hat{\varepsilon}_{t-1} + ... + \rho_k\Delta\hat{\varepsilon}_{t-k} + u_t \tag{3.3}$$

and calculate the ADF-statistic for the estimated $\hat{\rho}$, defined as $t_\rho = \hat{\rho}/SE(\hat{\rho})$, where $SE(\hat{\rho})$ is the standard deviation of $\hat{\rho}$.

3) Estimate the ADF(k) regression under the null of no cointegration (i.e. imposing $\rho = 0$),

$$\Delta\hat{\varepsilon}_t = \rho_1^0\Delta\hat{\varepsilon}_{t-1} + ... + \rho_k^0\Delta\hat{\varepsilon}_{t-k} + u_t^0 \tag{3.4}$$

and calculate regression residuals $\hat{u}_t^0$.

4) We can use the residuals $\hat{u}_t^0$ to derive a bootstrap distribution for the t_ρ-statistics under the null hypothesis of no cointegration, in the following way. For the stationary bootstrap, choose a value of p and form N blocks $B_i = \{\hat{u}_t^0, \hat{u}_{t-1,}^0 ..., \hat{u}_{t-l+1}^0\}$, $i = 1, ..., N$ where for each B_i, l is sampled from a geometric distribution and t from an uniform distribution as described above. For the moving block bootstrap choose a block length l and construct $n - l + 1$ blocks $B_i = \{\hat{u}_j^0, \hat{u}_{j+1,}^0 ..., \hat{u}_{j+l-1}^0\}$, $j = 1, ..., n - l + 1$ and resample with replacement N times from these blocks.

5) For each resampled block[5] $B_i = \{\hat{u}_j^0, \hat{u}_{j+1}^0, ..., \hat{u}_{j+l-1}^0\}$, we can use the bootstrapped residuals $\{\hat{u}_j^0\}$ and equation 3.4 to construct a time-series of residuals $\{\hat{\varepsilon}_{i,t}^0\}_{t=1}^l$ under the null of no cointegration[6].

6) Calculate for each bootstrapped sample $\{\hat{\varepsilon}_{i,t}^0\}_{t=1}^l$, $i = 1, ..., N$ the ADF statistics $\tilde{t}_\rho^i$ and store it in order to generate an empirical distribution of $\tilde{t}_\rho^i$.

7) Finally, define $\tilde{t}_\rho^L$ and $\tilde{t}_\rho^H$ respectively as the 2.5% lower and upper quantile of the distribution of the $\tilde{t}_\rho^i$ and reject the null if $t_\rho > \tilde{t}_\rho^H$ or $t_\rho < \tilde{t}_\rho^L$.

A few points are worth mentioning about the procedure just described. It is unavoidable when using bootstrap technique in a regression context to face two fundamental choices: first, we need to choose what variable to bootstrap and then how to construct its simulated counterpart. Regarding the first issue, we can choose to bootstrap either the original data or the regression residuals; then whatever variable we have chosen, we must decide how to generate its simulated bootstrap sample.

This last point can be made more clear by recalling step 4. There, we choose to use, as bootstrap sample, the residuals $\hat{u}_t^0$ from the ADF regression estimated under the null of no cointegration. Another possible could have been to choose instead $\hat{u}_t$, the regression residuals from the original ADF regression in eq. 3.3, and use them in step 4-7 instead of $\hat{u}_t^0$.

Concerning the first choice, we decided to use the regression residuals rather than the actual data for the following reason. Although there are some authors[7] that suggest bootstrapping the regressors rather than the residuals, it is not considered to be the best choice in a cointegration context since it discards some useful information used in the residual based approach. For instance, in testing whether $y_t = \alpha + \beta x_t + \varepsilon_t$ cointegrating relation, if we choose to bootstrap the regressors y_t and x_t would not

[5]either moving block or stationary bootstrap re-sample.

[6]We need to assume that the first $\tilde{\varepsilon}_0^i, ..., \tilde{\varepsilon}_{-k+1}^i$ are equal to zero and $\tilde{\varepsilon}_1^i$is equal to the first observation $\tilde{u}_j^0$ of the Bootstrap block.

[7]See Li and Maddala (1997).

take into account the information that under the null hypothesis the residuals ε_t are $I(1)$, while under the alternative they are $I(0)$.

Regarding the second issue, the scheme we chose in step 4-6 is, according to the findings of Li and Maddala (1997), the most reliable under the null since it does not depend on the OLS estimated parameters $\hat{\rho}$ and $\hat{\rho}_1, ..., \hat{\rho}_k$. In fact, if the null is true but the OLS gives an estimate of ρ which is far away from zero, then the bootstrap distribution based for instance on $\hat{u}_t$ will give a poor approximation of the distribution of the true errors under the null. On the contrary, the bootstrap distribution of ρ based on the residuals $\hat{u}_t^0$ will be close to the true distribution if then null hypothesis is true. Conversely, if the null is not true then it will be very different from the true distribution. More details on these issues can be found in Li and Maddala (1997), van Giersbergen and Kiviet (2002) and Basawa et al.(1991).

3.3 The PPP-UIP cointegration analysis

In this section we show how standard methods can reject cointegration for a long run equilibrium that arise from economic theory.

We reconsider the PPP-UIP data discussed in Johansen and Juselius (1992), Harris (1995) and Rahbek and Mosconi (1999), and to test for cointegration using the E-G approach. A full description of this data, and the source, goes beyond the purpose of this chapter and it can be found in the given references. The database is composed by quarterly, seasonally adjusted, time series from 1972-1 to 1987-2 for the UK wholesale price index (p_t^{uk}), the UK trade weighted foreign wholesale price[8] index (p_t^{w}), the three-month UK treasury bill rate (i_t^{uk}), the three month Eurodollar interest rate (i_t^{ed}), and the UK effective exchange rate (e_t^{uk}). The purpose is to test whether the PPP and

[8]The wholesale price index is a more palusible variable than retail price index for an empirical analysis of the PPP. In fact, it is likely that for many goods, trading costs for the consumer are much larger than those faced by the wholesaler.

the UIP relations, that arise from economic theory, hold empirically. In fact, economic theory suggests that in the long run we should expect price differentials between two countries to be equal to the nominal exchange rate differential (PPP)[9], and the nominal interest rate differentials to be equal to the expected changes in the exchange rates (UIP).

In other terms, we would expect that

$$p_t^{uk} - p_t^{w} = e_t^{uk} \qquad \text{(PPP)} \qquad (3.5)$$

$$i_t^{uk} - i_t^{ed} = E\left(e_t^{uk}\right) - e_t^{uk} \qquad \text{(UIP)} \qquad (3.6)$$

Now, if the markets are efficient, we would expect that deviations from the PPP would influence expected changes in exchange rate, in other words

$$E\left(e_t^{uk}\right) = p_t^{uk} - p_t^{w}$$

Hence, the above equations 3.5 and 3.6 can be linked together as

$$i_t^{uk} - i_t^{ed} = p_t^{uk} - p_t^{w} - e_t^{uk} \qquad (3.7)$$

We can estimate this long-run relation[10] in eq. 3.7 by running the regression

$$p_t^{uk} = \alpha_0 + \alpha_1 p_t^{w} + \alpha_2 e_t^{uk} + \alpha_3 i_t^{uk} + \alpha_4 i_t^{ed} + \varepsilon_t \qquad (3.8)$$

The OLS estimates together with some diagnostics are shown in the following table

[9]This definition of the PPP relation is also known as the absolute version which is assumed to be valid only in the long run.

[10]In order to account for the possibility that the prices are $I(2)$, we also used the different specification suggested in Rahbek and Mosconi (1999)

$$p_t^{uk} - p_t^{w} = \alpha_0 + \alpha_1 \Delta p_t^{w} + \alpha_2 e_t^{uk} + \alpha_3 i_t^{uk} + \alpha_4 i_t^{ed} + \varepsilon_t$$

However, since we failed to reject the null of no cointegration we report the cointegration results only for the standard case.

$p_t^{uk} = \underset{(-2.78)}{-2.18} + \underset{(25.96)}{1.6}\, p_t^{w} + \underset{(1.132)}{0.13}\, e_t^{uk} + \underset{(0.77)}{0.33} i_t^{uk} - \underset{(-1.74)}{0.78}\, i_t^{ed} + \varepsilon_t$		
$R^2 = 0.985$	$\sigma^2 = 0.1464$	$D - W = 0.1464$
$L - B = \underset{(0.00)}{141}$	$RESET = \underset{(0.00)}{14.54}$	

Table 3.1: OLS estimate of the UIP-PPP regression

It can be readily seen that the residuals diagnostic reveals the presence of omitted nonlinearities and highly autocorrelated residuals. Furthermore, a very low *D-W* statistic together with a very high R^2 could be a sign that the above relation is a spurious regression.

Following the E-G approach, we use the estimated regression residuals $\hat{\varepsilon}_t$ to test whether the relation estimated in table 3.1 has any economic meaning or is simply a spurious relation

In the next table, we report the ADF[11] statistic for $\hat{\varepsilon}_t$ together with its asymptotic[12] critical values and both the stationary and moving block bootstrap 95% critical values, named respectively ACI, SBCI and MBBCI

t_ρ-*statistics*	95% ACI	95% SBCI	95% MBBCI
-0.965320	-2.937	-2.5558	-2.2027

Table 3.2: ADF statistics and asymptotic, stationary bootstrap and moving block bootstrap critical values

The null of no cointegration cannot be rejected at 95% and therefore we should conclude that the relation we found in table 3.1 is spurious.

This result is not surprising since it is quite well-known, see Harris (1995) for instance, that the hypothesis of cointegration is rejected for the PPP-UIP relation.

[11]We run the ADF test up to five lags in the ADF regression. Since we fail to reject the in all cases, we report the results on when one lag is considered.

[12]The asymptotic critical values are taken from Said and Dickey (1984).

Many economist[13] have tried to justify this counter-intuitive result and the issue is still controversial[14]. Some of the justification for such empirical failure are related to trade barriers, pricing to market, international trade costs[15] such as transport costs, product heterogeneity, indirect tax differences, etc.[16]. Furthermore, the presence of omitted nonlinearities in eq. 3.8 is empirically consistent with the presence of trade costs (see Micheal et al. (1997) and Taylor (2001)).

As we show below, the result presented in this section is an artifact caused by the long memory of the residuals and the incapability of standard methodology to account for such degree of persistence. In fact, we would find evidence of a long-run relation in the PPP-UIP database only if this relation were strictly linear and the variables converged toward their equilibrium values at a relatively fast rate. These hypotheses implies, in economic terms, perfect competition in the foreign (goods and exchange rates) markets which is rejected by empirical evidence. Hence, for these reasons, even if the existence of a PPP-UIP long run equilibrium seems economically plausible, we would expect the speed of convergence to be very slow due to market failures mentioned in the previous paragraph. The possibility that deviations from the UIP and PPP equilibrium could follow a mean reverting long memory process was already suggested by Cheung and Lai (1993), Abuaf and Jorian (1990), Imbs et al. (2005) and Abadir and Talmain (2005). Our results also confirm the presence of such long memory. In fact, by an inspection of fig.3.1, the autocorrelation function (ACF) of $\hat{\varepsilon}_t$ does not converge toward zero exponentially as implied by the I(0) assumption, but clearly it does not support evidence of a possible unit-root either. Its slow rate of decay could support

[13]See Obstfeld and Rogoff (2000) for a good survey on PPP failure.

[14]Abadir and Talmain (2005) have recently solved the UIP puzzle with the same quasi ML approach although in a different context since they did not test explicitly for cointegration.

[15]The effects of these costs would have been much bigger if we had used retail price rather than wholesale price.

[16]Other explanation that have been advocate to explain the PPP empirical failure are related to differences in the price index weight, in the productivity growths and in the proportion of tradeable to non tradeable.

the presence of long memory.

figure 3.1

3.4 A modified Engle-Granger approach to test for long-memory cointegration

In this section we present our approach to test for the existence of a long-memory relation between the PPP-UIP variables. As mentioned above, testing for cointegration in a E-G context is equivalent to test whether in the representation

$$(1-L)^d \Phi(L) \varepsilon_t = \Theta(L) e_t \qquad e_t \sim IN(0, \sigma_e^2) \tag{3.9}$$

the parameter d is equal to one (i.e. no cointegration) or equal to zero (i.e. cointegration) where $\Phi(L)$ and $\Theta(L)$ are polynomial on the lag operator L. This implies that in both cases we are assuming that deviations from the long run equilibrium evolves according to an ARIMA model.

Recently, some researchers have started to doubt the ability of ARIMA process to fit the dynamics of many economic variables, in favor of the more general class of long memory processes. Considerable evidence of long memory in macroeconomic times series has be found in the works of Sowell (1992), Rudebusch(1989), Diebold and Rudebusch(1991a), Baillie and Bollerslev (1994), Baillie et al.(1996), Haubrich and Lo (1993), Crato and Rothman(1994), Hassler and Wolters (1995) and Abadir et al.(2006).

In particular, Abadir and Talmain (2002, 2005), have recently shown that the dynamics of most macroeconomic variables follow a new type of mean-reverting long-memory process. This process is characterized by a very slow decay of the ACF whose leading term can be represented by the functional form

$$\rho(\tau) \simeq \frac{1 - a\left[1 - \cos(\omega\tau)\right]}{1 + b\tau^c} \tag{3.10}$$

where a, ω, b, c are parameters to be estimated. This functional form can closely fit the ACF of many macroeconomic time series, including all the variables studied in the seminal Nelson and Plosser (1982) paper; see Abadir et.al. (2006).

It turn out that can also describe very well the rate of decay of the ACF of $\hat{\varepsilon}_t$. In fact, estimating eq. 3.10 for the ACF of $\hat{\varepsilon}_t$ gives

$$\hat{\rho}_t \simeq \frac{1 - 1.047\left[1 - \cos\left(0.28762\tau\right)\right]}{1 + 0.3225\tau^{0.17045}} \tag{3.11}$$

which, as shown in figure 3.2, reveals a striking accuracy of this functional form in fitting the ACF of $\hat{\varepsilon}_t$.

figure 3.2

Abadir and Talmain (2005) show that it is possible to use this functional form to disentangle co-movements between variables from the effects of their own persistence. In fact, we can use it to correct the original variables to account for autocorrelations and eventual non-linearities that could otherwise lead to biased estimates and spurious results, as extensively discussed in the introduction of this paper. This can be easily done, by using eq. 3.11 to construct the estimated autocorrelation matrix $\hat{R}$ for $\hat{\varepsilon}_t$, i.e.,

$$\hat{R} \equiv \begin{pmatrix} 1 & \hat{\rho}_1 & \cdots & \hat{\rho}_{T-2} & \hat{\rho}_{T-1} \\ \hat{\rho}_1 & 1 & \ddots & \ddots & \hat{\rho}_{T-2} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \hat{\rho}_{T-2} & \ddots & \ddots & 1 & \hat{\rho}_1 \\ \hat{\rho}_{T-1} & \hat{\rho}_{T-2} & \cdots & \hat{\rho}_1 & 1 \end{pmatrix} \tag{3.12}$$

and implementing it either in a generalized least square approach or in a ML framework. In fact, since eq. 3.10 requires only four parameters it can make any GLS procedure easy to implement[17].

[17]In fact, the implementation of GLS would require the estimation of $(T-1)$ parameters, which is generally impossible with the exception of some special cases.

To the light of these consideration we have modified the E-G approach in the following way:

1) Estimate the long-run relation $y_t = \alpha + \beta x_t + \varepsilon_t$ by maximizing the likelihood

$$ML(\alpha,\beta) = (2\pi)^{-\frac{1}{2}} |R|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(y_t - \alpha - \beta x_t)' R^{-1}(y_t - \alpha - \beta x_t)\right)$$

with respect to α and β and the parameters $(a,\ b,\ c,\ \omega)$ of the functional form,

$$\rho_\varepsilon(\tau) \simeq \frac{1 - a[1 - \cos(\omega\tau)]}{1 + b\tau^c}$$

This can be done with any two step procedure[18] for a given starting values for $a,\ b,\ c$ and ω.

2) Calculate the regression residuals $\hat{\varepsilon}_t = y_t - \hat{\alpha} - \hat{\beta} x_t$ and estimate the ADF regression

$$\Delta\hat{\varepsilon}_t = \phi\hat{\varepsilon}_{t-1} + \phi_1\Delta\hat{\varepsilon}_{t-1} + ... + \phi_k\Delta\hat{\varepsilon}_{t-k} + u_t \tag{3.13}$$

maximizing the likelihood functions

$$ML(\rho,\rho_{1,...,}\rho_k) = (2\pi)^{-\frac{1}{2}} \left|\hat{\Omega}\right|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(\Delta\hat{\varepsilon}_t - \phi\hat{\varepsilon}_{t-1})' \hat{\Omega}^{-1}(\Delta\hat{\varepsilon}_t - \phi\hat{\varepsilon}_{t-1})\right)$$

where

$$\begin{aligned} \hat{\Omega} &= \{\hat{\omega}_{i,j} : \hat{\omega}_{i,j} = \hat{\rho}_u(|i-j|)\,;\ i = 1,...,T-1;\ \ j = 1,...,T-1\} \\ \hat{\rho}_u(\tau) &\simeq \frac{1 - a[1 - \cos(\omega\tau)]}{1 + b\tau^c} \end{aligned}$$

with respect to ϕ and the parameters of the functional form $\hat{\rho}_u(\tau)$ fitted to the ACF of u_t.

3) Evaluate the t-statistic t_ϕ for the parameter estimated $\hat{\phi}$ in eq. 3.13.

4) Calculate the bootstrap critical values $\tilde{t}_\rho^L$ and $\tilde{t}_\rho^H$ for t_ρ using the approach described in section 2.

5) Reject the null hypothesis of no cointegration if either $t_\rho < \tilde{t}_\rho^L$ or $t_\rho > \tilde{t}_\rho^H$.

[18] Although this can appear as a unconstrained optimization problem, it is not since $\hat{R}$ has to be positive definite.

In the last section of this chapter we investigate the improvement of this approach on the size and the power of the ADF test. In the next section we show how it can help to detect a long memory equilibrium between the PPP-UIP data.

3.5 The PPP-UIP cointegration test revised

In this section we apply the above procedure to the PPP-UIP database and show that we reject the hypothesis of no cointegration. We start with the estimation eq. 3.8 using the QML procedure, and then test for long-memory cointegration. The estimation results together with some residual diagnostics are shown in table 3.3

$p_t^{uk} = -\underset{(-2.45)}{1.48} + \underset{(23.56)}{1.47}\, p_t^w + \underset{(1.60)}{0.14} e_t^{uk} \underset{(-0.12)}{-0.02}\, i_t^{uk} - \underset{(-2.44)}{0.61}\, i_t^{ed} + \varepsilon_t$		
$R^2 = 0.9989$	$\sigma^2 = 0.0006$	$D-W = 2.011$
$L-B = \underset{(0.25)}{4.155}$	$RESET = \underset{(0.12)}{2.00}$	

Table 3.3: UIP-PPP regression with modified procedure

It can be clearly seen that all the problems that emerged with the standard E-G approach, specifically autocorrelated residuals and omitted non-linearities, have disappeared. Furthermore, a high R^2 together and a D-W very close to 2, eliminate any possibility of an eventual spurious regression.

We now estimate the ADF regression (using again the QML approach), and calculate the statistic t_ρ and its bootstrapped critical values $\tilde{t}_\rho^L$ and $\tilde{t}_\rho^H$.

We present in the next table the t_ρ-statistic together with the SB and MBB critical value[19] and its asymptotic critical values

[19] To get the bootstrapped critical value we choose a number of draws equal to 1000. For the moving block bootstrap we set a block length equal to 15, while for the stationary bootstrap a values of p (the parameter of the geometric distribution) equal to 0.04, which gives an average sample length

t_ρ-*statistics*	95% SBCI	95% MMBCI	95% ACI
-2.8012	-2.6028	-2.416	-2.937

Table 3.4: ADF statistics for modified regression with stationary bootstrap, moving block bootstrap and asymptotic critical values

We reject the null hypothesis of no cointegration. This result gives evidence that there exists a long memory cointegrating relation among the PPP and UIP variables. Although such relation is not strictly stationary, as required by standard cointegration, it is still mean reverting. This very slow adjustment, is not, however, completely unrealistic. As already mentioned, it is consistent with the assumption of market failures that characterize the foreign goods and exchange rate markets. Thus, as anticipated above, it is possible to detect a long-run stable relation between the variables if we allow for possible non-linearities and fluctuations from the long-run equilibrium, that are characterized by a strong persistence that cannot be captured by strictly stationary processes. In the next section, through simulation we give more support to this intuition.

3.6 Simulations

In this section we run some simulation to compare the standard E-G cointegration approach with the procedure described above. Specifically, we evaluate the empirical size and power[20] of the ADF test in the E-G original framework and in our modified procedure. Furthermore, we also calculate the difference in the small-sample estima-

of 25. We have also tried to simulate for different sample length but we didn't observe significant differences in the results.

[20]We remind that the size of a test is the probability of rejecting the null hypothesis when this is true while the power is the probability of rejecting the null when the alternative is true.

tion bias between the QML procedure and standard OLS. The simulation has been conducted as follows.

For the sake of comparison with similar works (Engle and Granger (1987), Cheung and Lai (1993) and Gil-Alana (2003) for instance) we use artificial data $x_{1,t}$ and $x_{2,t}$ generated by the bivariate system

$$\begin{aligned} x_{1,t} + x_{2,t} &= u_t \\ 2x_{1,t} + x_{2,t} &= v_t \end{aligned}$$

We consider two different data generating processes (DGP). In the first, we assume no cointegration and define $(1 - L)\, u_t = \xi_t$ and $(1 - L)\, v_t = \eta_t$ where both ξ_t and η_t are $IN\,(0, 1)$ variables. Given this data generating process, we evaluate the size of the ADF test for both the standard E-G approach and our modified procedure.

In the second, we assume that v_t is a long memory process with the same ACF structure as $\hat{\varepsilon}_t$, the equilibrium error from the PPP-UIP relation, and evaluate the power of the ADF test for the two approaches. Under this assumption, $x_{1,t}$ and $x_{2,t}$ are by construction non stationary variables but they are linked together by the cointegrating vector $[1, -0.5]$ which describes the long memory equilibrium between the two variables.

The artificial data for v_t is constructed in the following way. First, starting from the estimated functional form[21] in eq. 3.11, we construct the variance-covariance matrix $\hat{R}$ of $\hat{\varepsilon}_t$, as defined in 3.12. Then, using the Cholesky factorization we decompose $\hat{R}$ into $\hat{R} = \Gamma\Gamma'$ where Γ is lower triangular. Finally, given the sequence $\{\eta_t\}_{t=1}^{T}$ of $IN\,(0, 1)$

[21] Originally this exercise was conducted using the estimated functional form for the UIP regression in Abadir and Talmain (2005).After we decided to introduce the empirical application for the PPP-UIP database we believe it is more coherent to use the functional form estimated above. In any case, using the functional form in Abadir and Talmain did not give significantly different results from the one presented below.

variables we construct v_t as

$$v_t = \Gamma \eta_t$$

This transformation as the effect to generate in v_t the same autocorrelation structure as $\hat{\varepsilon}_t$.

In the first simulation we evaluate the empirical size and the power for the standard E-G cointegration approach. Following Li and Maddala (1997) we have set the number of replication to 500. For each replication, we apply the original E-G approach, as described in previous section, and evaluate the ADF t_ρ-statistic and its bootstrapped 2.5% lower and upper quantiles $\tilde{t}_\rho^L$ and $\tilde{t}_\rho^H$. When x_1 and x_2 are generated according to the first DGP described above, we calculate the percentage of time that the null hypothesis is rejected when it is true (size of the test). Conversely, when the artificial data is generated according to the second DGP, we calculate the percentage of time that the test rejects the false null hypothesis of no cointegration (power of the test). In the following tables we report the empirical size and power of the ADF test[22] for the standard E-G approach, calculated respectively using moving block and stationary bootstrap for a sample length respectively equal to 100 and 200.

[22]It has to be mentioned that when the null is true we have set k, the number of lags in the ADF-regression, equal to 1 (i.e. the true data generating process under the null). On the other side, when the alternative is true we have chosen a number of lags equal to three in order to remove all the autocorrelations from the residuals. This was done since using the same number of lags for both cases it would have reduced even more the power and the size of the ADF-test. It is important to note that the modified approach is not sensitive to the choice of k since the procedure is designed to account for any autocorrelations in the residuals.

No cointegration (Standard E-G)		
Sample Size	MBB	SB
$T = 100$	0.130	0.126
$T = 200$	0.056	0.070

Table 3.5: Sample size of the ADF test in E-G approach when null of no cointegration is true. Nominal size: 0.05

Long memory cointegration (Standard E-G)		
Sample Power	MBB	SB
$T = 100$	0.014	0.002
$T = 200$	0.148	0.072

Table 3.6: Sample power of the ADF test in E-G approach when null of no cointegration is true. Nominal size: 0.05

First, using bootstrapped critical values the ADF test has a size which is close to the nominal, especially for a sample length of 200 observations. This result confirms the findings in Li and Maddala (1997) that bootstrap critical values improve the size of the ADF test under the null hypothesis. Then, most importantly, when the alternative hypothesis of long-memory cointegration is true, the ADF test has very low power. In fact, it rejects the null hypothesis at most 15 % of times when the alternative is true. This means that the ADF test, in the standard E-G context, is not capable of distinguishing between long memory and unit root even in fairly large samples. This result has an important implication for the macroeconomist. It shows that rejection of a long run equilibrium by the ADF test is not conclusive evidence for excluding any relations between economic data. In fact, the E-G approaches would lead to the conclusion that no equilibrium relation exists between the variables, any time that this is not strictly $I(0)$.

We now repeat the same exercise for our modified approach and show that it can deal properly with this problem. In table 3.7 and 3.8 we report the empirical size and power of the ADF test in our modified framework.

No cointegration (Modified E-G)		
Sample Size	MBB	SB
$T = 100$	0.09	0.08
$T = 200$	0.08	0.03

Table 3.7: Sample size of the ADF test in modified approach when null of no cointegration is true. Nominal size: 0.05

Long memory cointegration (Modified E-G)		
Sample Power	MBB	SB
$T = 100$	0.864	0.89
$T = 200$	0.98	0.984

Table 3.8: Sample power of the ADF test in E-G approach when null of no cointegration is true. Nominal size: 0.05

The rejection frequencies under the null does not present any significant different with the standard case; in fact, we can reject at most 10% of times the null of no cointegration when it is true. On the other side, the power of the test has a striking improvement. In fact, the rejection frequency is very high and close to its nominal values even for short samples. Already with a sample length of 100 observations we are able to reject 90% of the times the null of no cointegration. Therefore, on one side our approach is as reliable as the standard approach when there is no relation among the variables; on the other, is capable to detect a long run equilibrium when the fluctuation from such equilibrium are not strictly stationary.

This shows that our approach provides a robust procedure for disentangling long memory co-movements in the variables from their persistence and eventual non linearities.

Therefore, allowing (and accounting) for possible non linearities and long memory it is possible to detect the true cointegrating relation and long memory fluctuations around the long run equilibrium.

In the last simulation we evaluate the small sample estimation bias for both the approaches and show that our modified procedure leads to a substantial improvement.

We use the OLS estimate β^{ols} and QML estimate β^{QML} for the cointegrating parameter (equal to -0.5 in the data generating process), from the previous exercise, and evaluate the mean of the small-sample bias for both approaches, defined as

$$\begin{aligned} B\left(\beta^{ols}\right) &= \frac{\sum_{i=1}^{N}\left(\beta_i^{ols}+0.5\right)}{N} \\ B\left(\beta^{QML}\right) &= \frac{\sum_{i=1}^{N}\left(\beta_i^{QML}+0.5\right)}{N} \end{aligned} \tag{3.14}$$

where N is the number of replication in the simulation.

In table 3.9 we report the small sample bias, for the estimation of the cointegrating vector respectively for the QML approach and OLS estimation.

Bias in cointegrating parameter estimation		
Sample size	$B\left(\beta^{QML}\right)$	$B\left(\beta^{ols}\right)$
$T = 100$	0.0098423	0.016869
$T = 200$	0.0064476	0.008687

Table 3.9: Sample power of the ADF test in E-G approach when null of no cointegration is true. Nominal size: 5

For a sample size of 100 observation, the small sample bias for the least square estimator β^{ols} is about 0.017 while the bias for the QML estimator β^{QML} is 0.0098 which is

41% smaller. When number of observation double the bias decreases for both estimator but the β^{ols} is still significantly more biased than β^{QML}. In both cases our procedure is able to reduce significantly the estimation bias compared to standard OLS that is usually implemented in E-G approach. This improvement can be justified considering that with the QML procedure we account for eventual omitted non-linearities and strong autocorrelations in the regression residuals. To the light of the results presented in this section, we can conclude that the inability of the standard E-G approach to test for long memory co-movements relies on the strict conditions, linearity and strict stationarity, that it imposes on the long-run relation among the variables. Once the approach is modified to account for long memory fluctuation around their equilibrium relation, it becomes a robust tool to test for cointegration.

Summarizing, the main results of this section are the followings. Consistently with findings of Li and Maddala (1997), the ADF test based on bootstrap critical values is characterized by high size under the null hypothesis. On the other side, when the alternative of long memory co-movements is true, the ADF test displays very low power in the standard E-G approach. By accounting for the long memory in the data it is possible to increase significantly the empirical power of the ADF test. In fact, the ADF test in our modified procedure displays a power which is close to the nominal. Finally, estimating the cointegrating vector by QML described above leads to a substantial reduction in the small sample bias compared to the standard OLS estimator. In the light of these results, the failure of the standard E-G approach to test for long memory co-movements relies on the strict conditions, linearity and strict stationarity, that it imposes on the long-run relation among the variables. Since most of the macroeconomic data seem to be characterized by long memory and mean reversion, it is not surprising that the E-G approach rejects cointegration even when a long-run equilibrium seems economically plausible (as for the PPP-UIP theorem).

3.7 Conclusion

Standard cointegration analysis technique such as the E-G approach and the FIML imposes very strict restrictions of the data generating process of the data. In fact, in order to find an equilibrium relation among economic variables, this should be strictly linear and characterized by a fast convergence rate toward the equilibrium. It might be plausible that these conditions do not hold empirically and therefore it is very likely that cointegration is rejected even when economically plausible. In this work we show that once these two assumptions are relieved, we can find the existence of a long memory equilibrium among the variables. To this purpose we presented a methodology to test for the presence of cointegration when the variables are $I(1)$ and there exists a linear combination that is characterized as long memory process rather than as standard ARMA process. We showed that, differently from standard cointegration technique, our approach is able to detect the presence of a long-memory co-movements. Furthermore, we reported for this test both high size under the null hypothesis and high power under the alternative. Our approach also reduced the small sample bias in the estimation of the cointegrating vector by more than 41% compared to standard ordinary least square estimate. We applied our procedure to the data base for the UK purchasing-power parity and uncovered interest rate parity and showed that the null of no cointegration was rejected at 95% in contradiction with what was previously shown with standard single equation technique.

3.8 List of figures

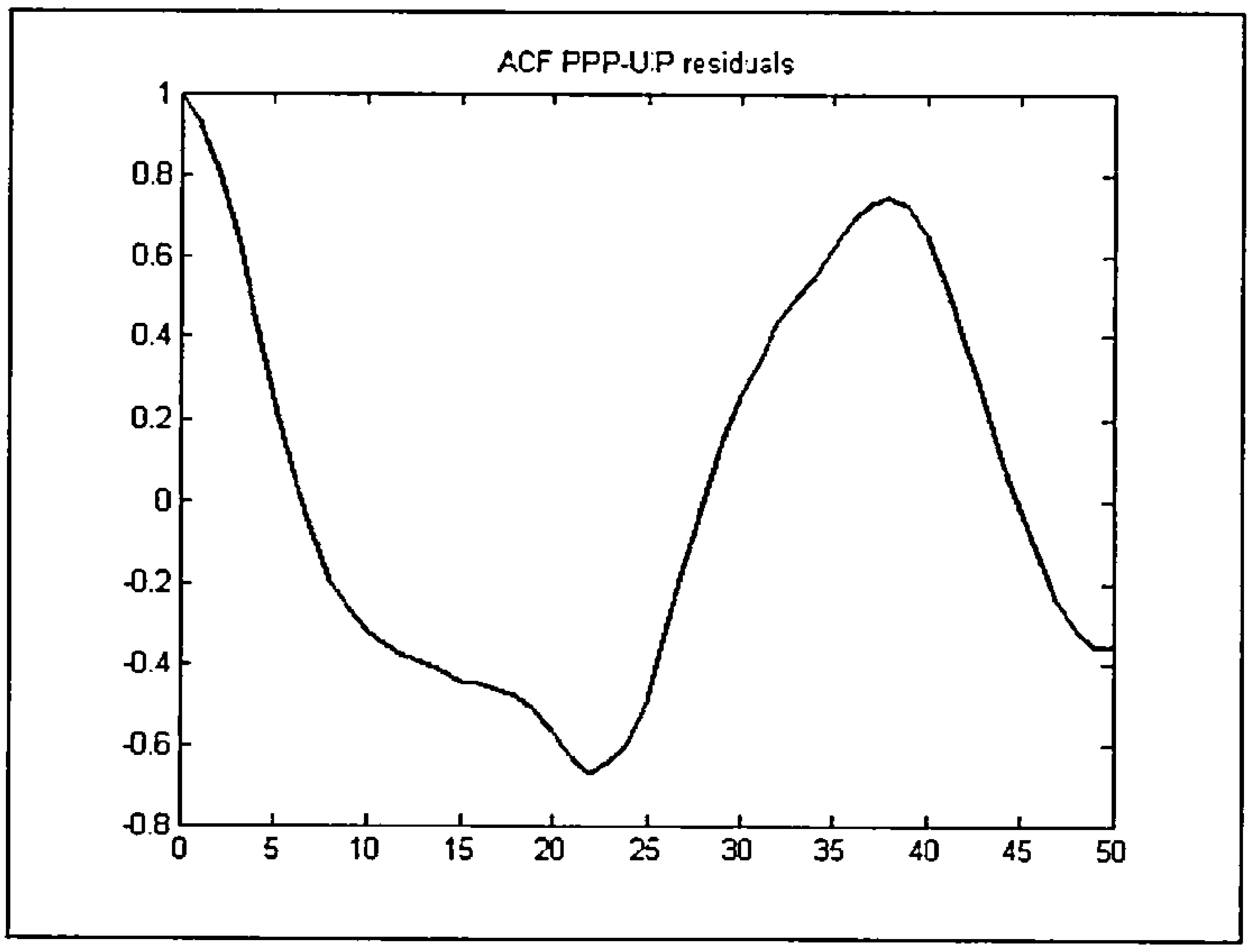

Figure 3.1: Autocorrelation function of the PPP-UIP regression residuals

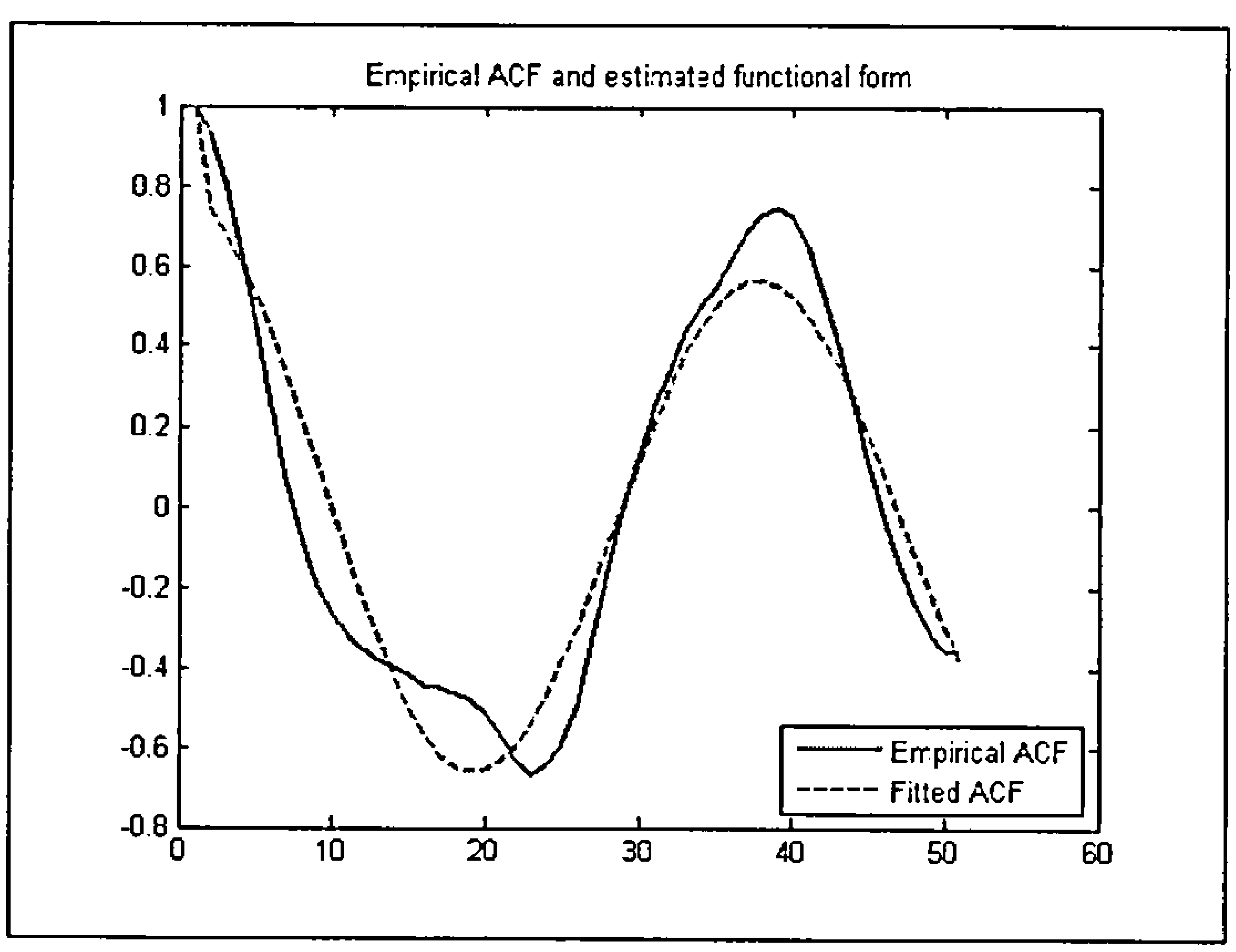

Figure 3.2: Autocorrelation function of the PPP-UIP regression residuals and estimated functional form

Chapter 4

Aggregation of almost unit root processes

4.1 Introduction

Aggregation is an inevitable aspect of both applied and theoretical macroeconomics. On one side, most of the observed macroeconomic time series are the result of summing a large number of heterogeneous units, such as commodities, firms or households. On the other, many macroeconomic theories are based on the aggregation of individual agents whose inter-temporal optimal decision rule is expressed as a linear model that contains some autoregressive component. For instance, the reduced form of many linearized microfounded models is represented by an ARMA model (see for instance King et al. (1988) and Michelacci and Zaffaroni (2000)).

Cross-sectional aggregation was first considered in the seminal works of Robinson (1978) and Granger (1980). A detailed reference with some economic applications can be found in Forni and Lippi (1997), Chambers (1998), Lewbel (1994) and Zaffaroni (2004).

The importance of aggregation in time series analysis derives from the fact that the statistical properties of an aggregate process can be dramatically different from those of its components. In particular, it is a well-known result[1] that averaging a large number of ARMA processes generates long memory in the aggregate process, even if its single components are strictly stationary.

Since the quoted works of Granger and Robinson, long memory and aggregation have gained much popularity both in theoretical and applied macroeconomics. On one side some researchers (see Diebold and Rudebusch (1989), Sowell (1992) and Abadir et.al. (2006) among others) have claimed that non-stationary long memory processes are a valid alternative to standard ARIMA processes to represent the dynamics of many economic time series. In particular, Diebold and Rudebusch (1989) and Abadir et al.(2006), using a long span of U.S. database, have produced strong evidence in support of long memory rather than difference stationarity in many economic variables. On the other, heterogeneity and aggregation have been recently employed to explain some "economic puzzles". Specifically, Altissimo, Mojon and Zaffaroni (2005) show that aggregation can help to reconcile slow macroeconomic adjustment with fast adjustment at microlevel in the euro area inflation; Abadir and Talmain (2005) use it to solve the "uncovered interest rate parity puzzle", and Imbs et al. (2005) to explain persistent deviations from the purchasing power parity.

The most popular case of aggregation in time series literature involves AR(1) processes. Specifically, given the AR(1) process

$$\ln x_{i,t} = \alpha_i \ln x_{i,t-1} + e_{i,t}; \qquad i = 1, .., N \tag{4.1}$$

$$e_{i,t} \sim IN\left(0, \omega_i^2\right) \tag{4.2}$$

where $x_{i,t}$ is the observed variables and the coefficients α_i and the shock $e_{i,t}$ vary across the N units, a linear aggregate process is defined as the arithmetic mean of the AR(1)'s just defined, i.e.

$$X_t = \frac{\sum_{i=1}^{N} \ln x_{i,t}}{N} \tag{4.3}$$

[1]See Robinson (1978), Granger (1980) and Lippi and Zaffaroni (1998).

The statistical properties of this process have been intensively studied. Apart from Robinson and Granger, further references are Forni and Lippi (1997) and Lippi and Zaffaroni (1998). However, they will be briefly recalled in the next section.

In a recent paper, Abadir and Talmain (2002), introducing firm heterogeneity in a general equilibrium model, derive another approach to aggregate data which arises from the microstructure of the model. Specifically, they consider the following aggregate process

$$\ln Z_t = \ln\left(\frac{\sum_{i=1}^{N} x_{i,t}}{N}\right) \tag{4.4}$$

We refer to this approach as CES or non-linear aggregation. The difference between these two approaches may seem subtle but, as we will see below, it has important implications for the dynamic properties of the aggregate process. In fact, Abadir and Talmain's approach differs from the linear aggregation in two aspects. First, it is coherent with micro-structure of the theoretical macro-model. Then, most importantly, it is consistent with the way economic data are aggregated in national accounts. In fact, the majority of the macroeconomic time series, available in the national statistical database, are obtained by adding up the levels, rather than the logarithm of disaggregated data. While X_t is the process usually studied in time series analysis, $\ln Z_t$ is the one that arises from the 'true' data generating process (DGP) in national accountancy. In this chapter we compare the statistical properties of these two different aggregate processes and derive the conditions under which their dynamics are closely related. For this purpose we evaluate the discrepancy between their autocovariance functions (ACvF) for different distributions of the variances of the shocks $e_{i,t}$. We show that if the $\omega's$ are dense in a neighborhood of zero, then the leading term of the ACvF of $\ln Z_t$ decays at the same rate as the ACvF of X_t. Under this assumption, X_t can represent closely the dynamics of $\ln Z_t$. Conversely, if this condition does not hold, the dynamics of the two processes can be dramatically different. In this case X_t is a misleading approximation to the true DGP since it tends to underestimate both its volatility and persistence.

Thus, although some authors, see for instance Forni and Lippi (1999), claim that X_t is a good approximation to $\ln Z_t$, choosing X_t to describe the dynamics of some macroeconomic variables could lead to misspecification due to the omission of the non-linearities that characterize $\ln Z_t$.

Furthermore, these non-linearities can also bias standard econometric techniques and give rise to spurious results. This point has been recently raised by Abadir and Talmain (2005) and is further discussed in the last chapter of the thesis.

The paper is organized as follows. In the next section, we present the properties of X_t. We first derive its properties for the case of the aggregation of independent series, then we extend to the case of dependent series. In section 3, we compare, theoretically and using simulations, the properties of the two aggregate processes. We also derive the conditions under which X_t approximate the leading term of the dynamics of $\ln Z_t$.

4.2 Linear aggregation of near unit root AR(1) processes

In this paragraph, we recall the general properties of the process X_t and then derive its ACvF when we allow for near-unit root processes in its components.

We start by making some basic assumptions about the structure of the AR(1) process, defined in eq. 4.1, that will be used throughout the chapter.

Assumption 1 We assume that the set of parameters $\theta_i = (\alpha_i, \omega_i)$ is defined on $\Theta = (0,1) \times \Re_{+}$, with α_i and ω_i mutually uncorrelated. We define $\Im = (\mathcal{F}, \mathcal{L})$ the set of the family $\mathcal{F}$ of absolutely continuous distribution defined on $(0,1)$, and $\mathcal{L}$ on $(0,\infty)$, with the distribution of α_i belonging to $\mathcal{F}$ and the distribution of ω_i belonging to $\mathcal{L}$.

This assumption states that the microstructure of the model is composed by strictly stationary AR(1) processes with near unit roots not being completely excluded. The set $\Theta = (0,1) \times \Re_{+}$, excludes the extreme cases where $\alpha_i = 0$ (*i.e.* $\ln x_{i,t} = \varepsilon_{i,t}$) and $\omega_i = 0$ (i.e. $\{x_{i,t}\}_{t=0}^{\infty}$ is not a stochastic sequence).

Assumption 2 Following Abadir and Talmain (2002) we assume that the $\alpha's$ are distributed according to the Beta density[2]

$$\beta(\alpha) = \begin{cases} \frac{\alpha^{g_a - 1}(1-\alpha)^{h_\alpha - 1}}{B(g_a, h_\alpha)} & 0 < \alpha < 1 \\ 0 & otherwise \end{cases} \tag{4.5}$$

where $h_\alpha \in (0,1]$ and $B(g_a, h_\alpha)$ is the Beta function.

Assumption 2 implies that almost unit roots are not completely ruled out. In fact, for $h_\alpha \in (0,1]$ the mode of the beta density lies in a neighborhood of one (i.e. $\beta(\alpha)$ goes to zero as α approaches 0). This hypothesis leads to non-stationarity of the aggregate process, in the sense of non-summable autocovariances and most of the results below rely on it.

To illustrate the properties of X_t let us recall eq. 4.1-4.3 and the autocovariance function $\gamma_{\ln x_{i,t}}$ for the AR(1) process $\ln x_{i,t}$,

$$\begin{aligned} X_t &= \frac{\sum_{i=1}^{N} \ln x_{i,t}}{N} && (4.6) \\ \ln x_{i,t} &= \alpha_i \ln x_{i,t-1} + e_{i,t} \qquad e_{i,t} \sim IN\left(0, \omega_i^2\right) \\ \gamma_{\ln x_{i,t}}(k) &= \frac{\alpha_i^k \left(1 - \alpha_i^{2t}\right)}{\left(1 - \alpha_i^2\right)} \omega_i^2 && (4.7) \end{aligned}$$

A few things are worth being noticed. First, it is well-known[3] that summing a finite number of ARMA process yields another ARMA process. For example, the sum of N distinct AR(1) process yields an ARMA(N,$N-1$). However, when N goes to infinity X_t does not belong to the class of ARMA process but, as Granger shows in his seminal

[2]For a detailed reference of the properties of the Beta density see Evans et. al. (2000).

[3]See Granger (1980).

paper, it is directly related to the class of long memory process so called fractional integrated ARMA process[4](FARIMA).

Then, while the statistical properties (conditioned on α_i and ω_i) of $\ln x_{i,t}$ are well defined, those of the aggregate processes X_t and Z_t are entirely defined by the distributions of the $\alpha's$ and the $\omega's$ as $N \to \infty$.

Finally, for values of the parameter α that are far away from one, the ACvF $\gamma_{\ln x_{i,t}}(k)$ can be approximated for large t by

$$\gamma_{\ln x_{i,t}}(k) \simeq \frac{\alpha_i^k}{(1-\alpha_i^2)}\omega_i^2 \tag{4.8}$$

This approximation also holds for distributions of the $\alpha's$ that are generated by the density $\beta(\alpha)$ with h_a strictly bigger than one.

Now, the ACvF of X_t can be recalled from Granger's results. He shows, in the frequency domain, that if the $\alpha's$ are distributed according to the modified Beta density $f(\alpha) = \frac{2}{B(h_a,g_a)}\alpha^{2g_a-1}(1-\alpha^2)^{h_a-1}$, and h_a is strictly bigger than one, then the ACvF of X_t, $\gamma_{X_t}(k)$, is given by

$$\gamma_{X_t}(k) = \frac{\Gamma(h_a-1)\Gamma\left(\frac{k}{2}-g_a\right)\omega^2}{B(g_a,h_\alpha)\Gamma\left(g_a+h_\alpha+\frac{k}{2}-1\right)}$$

It is easy to show that this is equivalent to evaluating, in a time domain, the mean of the ACvF of the AR(1) models $\ln x_{i,t}$, i.e. $\gamma_{X_t}(k) = N^{-1}\sum_{i=1}^{N} Cov(\ln x_{i.t+k}, \ln x_{i,t})$. In fact, recalling the ACvF of $\ln x_{i,t}$ in eq. 4.8 we have that

---

[4]They are defined as

$$(1-L)^d \Phi(L) X_t = \Theta(L) u_t$$

where $d \in (0,1)$, $\Phi(L)$ and $\Theta(L)$ are polynomial in the lag operator and $u_t \sim IN(0,\sigma_u^2)$.

An introduction to this kind of model can be found in Granger and Joyeux (1980), Hosking (1981) and more recently in Beran (1998). It is important here to recall that they require fractional differencing to achieve stationarity and do not have summable autocorrelations for $d > 0.5$.

$$
\begin{aligned}
\gamma_{X_t}(k) &= N^{-1}\sum_{i=1}^{N} Cov\left(\ln x_{i.t+k}, \ln x_{i,t}\right) = N^{-1}\sum_{i=1}^{N}\frac{\alpha_i^k\left(1-\alpha_i^{2t}\right)}{\left(1-\alpha_i^2\right)}\omega_i^2 \\
&\simeq N^{-1}\sum_{i=1}^{N}\frac{\alpha_i^k}{\left(1-\alpha_i^2\right)}\omega_i^2 \simeq E_i\left(\frac{\alpha_i^k}{\left(1-\alpha_i^2\right)}\sigma_i^2\right) = E\frac{\alpha_i^k}{\left(1-\alpha_i^2\right)}E\left(\omega_i^2\right) \quad (4.9) \\
&= \frac{2\omega^2}{B\left(g_\alpha, h_\alpha\right)}\int_0^1 \frac{\alpha_i^k}{\left(1-\alpha_i^2\right)}\alpha^{2g_a-1}(1-\alpha^2)^{h_a-1}d\alpha \quad (4.10) \\
&= \frac{\omega^2}{B\left(g_\alpha, h_\alpha\right)}\int_0^1 \alpha_i^{\frac{k}{2}+g_a-1}(1-\alpha)^{h_a-2}d\alpha \quad (4.11) \\
&= \frac{\omega^2\Gamma\left(h_a-1\right)\Gamma\left(\frac{k}{2}-g_a\right)}{B(g_a, h_\alpha)\Gamma\left(g_a+h_\alpha+\frac{k}{2}-1\right)} \quad (4.12)
\end{aligned}
$$

where we approximate for large t and large N in eq.4.9, replace α with $\alpha^{\frac{1}{2}}$ in eq. 4.10, and following Granger (1980), $E\left(\omega_i^2\right)$ with ω^2.

A few points are worth mentioning. First, the above formula holds only for h_a strictly larger than one since the Gamma function has a singularity[5] at zero. This means that eq. 4.12 is the ACvF of a stationary long memory process whose components $\ln x_{i,t}$ are strictly stationary (i.e. near unit root completely ruled out). This point is also confirmed by the fact that eq. 4.12 is the ACvF of a fractionally integrated process of order $d = (1-h_\alpha/2)$ which is stationary[6] as long as $d < 0.5$, which implies $h_\alpha > 1$. Then, the above formula has been derived under the assumption that $\varepsilon_{i,t}$ are independent. This excludes interesting cases such as the presence of common shocks. Generally, as shown below, the ACF of an aggregate process is different from the mean of the ACvF of its single component. In fact, the latter does not reflect the interactions between the components of the aggregate process. Conversely, when Abadir and Talmain analyze the properties of $\ln Z_t$, they explicitly take into account these interactions and show that they influence both the persistence and volatility of the aggregate process[7].

[5]See Gonzàles (1992) for a description of the properties and singularities of the gamma function.

[6]See Beran (1998) or Zaffaroni (2004)

[7]It has to be mentioned that also Granger considered the aggregation of dependent series but he obtained its results by dropping the interaction terms as a simplifying approximation.

Following this argument, it is possible now to extend these results to aggregate processes composed by almost unit root AR(1) processes, i.e. allow h_α to be less than one. In fact, in this case, the ACvF of X_t can be derived as

$$\begin{aligned}
\tilde{\gamma}_{X_t}(k) &= N^{-1}\sum_{i=1}^{N} Cov\left(\ln x_{i.t+k}, \ln x_{i,t}\right) \\
&= N^{-1}\sum_{i=1}^{N} \frac{\alpha_i^k\left(1-\alpha_i^{2t}\right)}{\left(1-\alpha_i^2\right)}\omega_i^2 \simeq E_i\frac{\alpha_i^k\left(1-\alpha_i^{2t}\right)}{\left(1-\alpha_i^2\right)}\omega_i^2 = E\frac{\alpha_i^k\left(1-\alpha_i^{2t}\right)}{\left(1-\alpha_i^2\right)}E\left(\omega_i^2\right) && (4.13) \\
&= \frac{2\omega^2}{B\left(g_\alpha,h_\alpha\right)}\int_0^1 \frac{\alpha_i^k\left(1-\alpha_i^{2t}\right)}{\left(1-\alpha_i^2\right)}\alpha^{2g_a-1}(1-\alpha^2)^{h_\alpha-1}d\alpha && (4.14) \\
&\simeq \frac{2\omega^2}{B\left(g_\alpha,h_\alpha\right)}\int_0^1 t\alpha_i^{k+2(t-1)+2g_a-1}(1-\alpha^2)^{h_\alpha-1}d\alpha && (4.15) \\
&= \frac{\omega^2}{B\left(g_\alpha,h_\alpha\right)}\int_0^1 t\alpha_i^{\frac{k}{2}+t+g_a-2}(1-\alpha)^{h_\alpha-1}d\alpha && (4.16) \\
&= \frac{\Gamma\left(h_a\right)\Gamma\left(\frac{k}{2}+t-g_a-1\right)\omega^2 t}{B(g_a,h_\alpha)\Gamma\left(g_a+h_\alpha+t+\frac{k}{2}-1\right)} && (4.17)
\end{aligned}$$

where we approximate for large N in eq.4.13, use l'Hopital's rule in eq. 4.14 and replace α with $\alpha^{\frac{1}{2}}$ in eq. 4.15.

It can be readily seen that, differently from $\gamma_{X_t}(k)$, $\tilde{\gamma}_{X_t}(k)$ is defined for $h_\alpha < 1$. It is possible now to compare the rate of decay of two ACvF for large k. In fact, using the Stirling's formula for $k \to \infty$, the rates of decay of the two ACvF's can be approximated by

$$\begin{aligned}
\gamma_{X_t}(k) &= \frac{\sigma^2\Gamma\left(h_a-1\right)\Gamma\left(\frac{k}{2}-g_a\right)}{B(g_a,h_\alpha)\Gamma\left(g_a+h_\alpha+\frac{k}{2}-1\right)} \\
&\simeq \frac{\Gamma\left(h_a-1\right)}{2^{1-h_\alpha}B(g_a,h_\alpha)}k^{1-h_\alpha} && (4.18) \\
\tilde{\gamma}_{X_t}(k) &= \frac{\Gamma\left(h_a\right)\Gamma\left(\frac{k}{2}+t-g_a-1\right)\sigma^2 t}{B(g_a,h_\alpha)\Gamma\left(g_a+h_\alpha+t+\frac{k}{2}-1\right)} \\
&\simeq \frac{\Gamma\left(h_a\right)\sigma^2 t}{2^{-h_\alpha}B(g_a,h_\alpha)}k^{-h_\alpha} && (4.19)
\end{aligned}$$

The above approximations show the importance of the shape of the beta density in determining the decay rate of the ACvF of X_t.

When $h_\alpha > 1$, the density $\beta(\alpha)$is not dense around one and $\gamma_{X_t}(k)$ tells us that X_t has finite variance and autocovariances decaying with an hyperbolic rate. Conversely, when $h_\alpha < 1$, a large mass of $\beta(\alpha)$ will be around the unit value and $\tilde{\gamma}_{X_t}(k)$ tells us that X_t has infinite variance but is still mean reverting. However, despite its non-stationarity it is still less persistent than a random walk. In fact, if we study the behavior of $\tilde{\gamma}_{X_t}(k)$ as h_α converges to zero, which implies that $\beta(\alpha)$ shrinks on the unit value, recalling that $B(g_a, h_\alpha) = \frac{\Gamma(g_a+h_\alpha)}{\Gamma(h_a)\Gamma(g_a)}$, we get

$$
\begin{aligned}
\tilde{\gamma}_{X_t}(k)\big|_{h_\alpha=0} &= \left.\frac{\Gamma(h_a)\,\omega^2 t}{2^{-h_\alpha}B(g_a, h_\alpha)}k^{-h_\alpha}\right|_{h_\alpha=0} \\
&= \left.\frac{\Gamma(g_a+h_\alpha)\Gamma(h_a)\,\omega^2 t}{2^{-h_\alpha}\Gamma(h_a)\,\Gamma(g_a)}k^{-h_\alpha}\right|_{h_\alpha=0} \\
&= \left.\frac{\Gamma(g_a+h_\alpha)\omega^2 t}{2^{-h_\alpha}\Gamma(g_a)}k^{-h_\alpha}\right|_{h_\alpha=0} \\
&= \omega^2 t
\end{aligned}
$$

which is the ACvF of a random walk.

This highlights the effects of heterogeneity and the distortion that is induced by approximating the dynamics of an aggregate process with a representative ARIMA process. For example, in this case, if we had fit the dynamics of X_t with a representative ARMA process we would have ended up with a random walk when the true DGP is not a random walk. In other words, we would have lost mean reversion. Finally, it has to be noticed that the parameter g_α does not influence the memory of the process but only its volatility.

To summarize, the aggregation of AR(1) process with occasional near unit root produces a non stationary long memory process whose ACvF is characterized by a hyperbolic rate of decay. Furthermore, it is mean reverting and has less memory than a random walk as long as heterogeneity is allowed, i.e. $h_\alpha > 0$.

We now relieve the hypothesis that $e_{i,t}$ are independent and generalize the results shown above. This is equivalent to having $e_{i,t} = \omega_i \varepsilon_t$ with $\varepsilon_t \sim IN(0,1)$. We also assume that the ω_i are distributed according to the gamma density

$$\gamma(\alpha,\beta) = \begin{cases} \frac{\alpha^{\beta}\omega^{\beta-1}e^{-\alpha\omega}}{\Gamma(\beta)} & 0 < \omega < \infty \\ 0 & otherwise \end{cases} \tag{4.20}$$

This is plausible distribution for values of the $\omega' s$ that are not far away from zero[8]. It can be then easily proven that the $\omega' s$ have mean $E(\omega)$ and variance $Var(\omega)$ respectively equal to $E(\omega) = \frac{\beta}{\alpha}$ and $Var(\omega) = \frac{\beta}{\alpha^2}$, see Evans et all (2000) for details on the moments of the gamma density. It is therefore possible with this density to increase the variance by keeping the mean constant. This property will be useful below when we analyze the effect of different values of the shock's variance for a given value of the mean.

In the following proposition we present the autocovariance function of X_t under the more general context where $\varepsilon_{i,t}$ is a dependent series.

Proposition 1 *Let X_t be the process defined in eq. 4.3 with*

$$\begin{cases} Cov(e_{i,t+k}, e_{j,t}) \neq 0 & k = 0 \\ 0 & otherwise \end{cases}$$

If assumption 1 and 2 hold and ω is distributed according to the density in eq. 4.20, then

$$Cov(X_t, X_{t+k}) \simeq \frac{\beta^2 t}{\alpha^2 B(g_\alpha, h_\alpha)^2} \frac{\Gamma(h_\alpha)^2 \Gamma(g_a + t - 1)}{\Gamma(g_a + t - 1 + h_\alpha)} \frac{\Gamma(g_a + t + k - 1)}{\Gamma(g_a + t + k - 1 + h_\alpha)} \tag{4.21}$$

Remark 2 *Using the Stirling's formula for large k the rate of decay of 4.21 can be approximated by*

$$Cov(X_t, X_{t+k}) \simeq \frac{\beta^2 \Gamma(g_a + h_\alpha)^2 \Gamma(g_a + t - 1)\, t}{\alpha^2 \Gamma(g_a)^2 \Gamma(g_a + t - 1 + h_\alpha)} k^{-h_\alpha} \tag{4.22}$$

[8] It has to be mentioned that this is a much simpler set up than Abadir and Talmain (2002) and the one assumed in the first chapter. However, it allows a clearer interpretation of the effects of the parameters α and β on the ACvF properties of the aggregate process.

while for large k and t it can by approximated by

$$Cov(X_t, X_{t+k}) \simeq \frac{\beta^2\Gamma(g_a+h_\alpha)^2}{\alpha^2\Gamma(g_a)^2}\frac{t}{[(t-1)(t+k-1)]^{h_\alpha}}$$

A few things are worth noticing. While the parameters α and β influence the volatility of the process, h_α is the only parameter that controls the rate of decay of the ACvF. In other words, in linear aggregation, the shape of the distribution of the $\omega's$ does not affect the memory of the process and h_α is the only parameter responsible for the degree of persistence of the aggregate process. This also shows that the introduction of dependency in the series does not play any effects on persistence either. In fact, both eq. 4.19 and eq. 4.22 are characterized by the same rate of decay. This is consistent with Granger's results for the aggregation of dependent series[9].

As we show below, the same is not generally true for $\ln Z_t$. In fact, not only the mean but also the spread of the distribution of the $\omega's$ affects its persistence. Particularly, if its distribution is not dense around zero, then the dependency of the series will play a determinant role on the persistence of the process $\ln Z_t$.

4.3 Non-linear aggregation of near unit root AR(1) processes

In this section we focus on the properties of $\ln Z_t$ and compare them with those of X_t. The statistical properties for this process have been analyzed by Abadir and Talmain (2002). In particular, they showed that it is characterized by long memory, non stationarity and mean reversion despite its persistence. Furthermore, as can be clearly seen, it is a highly non linear function of its geometric AR(1) components. As already mentioned, comparing the properties of these two processes is important in that while X_t is the one usually considered in time series analysis, Z_t is the process that arises from the data generating process in national accountancy.

[9] See section 3 in Granger (1980).

We start by showing in the next proposition that if the distribution of the ω's is dense[10] around zero, then the leading term of $\gamma_{\ln Z_t}(k)$ decays at the same rate as the ACvF of X_t.

Proposition 2 *Let Z_t be the process defined in eq. 4.4 and assumption 1 and 2 hold. If $\ln x_{i.t}$ and $\ln x_{j.t}$ are mutually correlated and ω is distributed according to the gamma density defined in eq. 4.20, then the leading term of the ACvF of $\ln Z_t$ is given by*

$$Cov(\ln Z_t, \ln Z_{t+k}) \simeq \frac{\beta^2 t}{\alpha^2 B(g_\alpha, h_\alpha)^2} \frac{\Gamma(h_\alpha)^2 \Gamma(g_a + t - 1)}{\Gamma(g_a + t - 1 + h_\alpha)} \frac{\Gamma(g_a + t + k - 1)}{\Gamma(g_a + t + k - 1 + h_\alpha)}$$

Corollary 1 *By the Stirling's formula for large k the rate of decay the ACvF of $\ln Z_t$ is given by*

$$Cov(\ln Z_t, \ln Z_{t+k}) \simeq \frac{\beta^2 \Gamma(g_a + h_\alpha)^2 \Gamma(g_a + t - 1)\, t}{\alpha^2 \Gamma(g_a)^2 \Gamma(g_a + t - 1 + h_\alpha)} k^{-h_\alpha}$$

Therefore, for values of the standard deviations close to zero[11], the leading term of $\gamma_{\ln Z_t}(k)$ decays at the same rate as $\gamma_{\ln Z_t}(k)$ derived in eq. 4.22. This means that X_t can represent a first order approximation of the dynamics of $\ln Z_t$. This result is consistent with what Forni and Lippi (1997) claim.

Conversely, if the mean of the ω's is far from zero or they assume occasionally large values, then the dynamics of the two processes become quite different. This result will be also confirmed by the simulations in the next section. This means that, as the ω's increase, X_t can no longer approximate the dynamics of $\ln Z_t$ in that also the higher terms of the leading term approximation start to affect the dynamics of the aggregate process. Since X_t represents the leading term of this expansion it is not able to capture this "higher order dynamics".

[10]This implies that the density has mean in a neighborhood of zero and its spread is not very large.
[11]It is straight forward to show that this result also holds the case of dependent series.

In the next proposition we extend the above results to more general aggregate processes.

Proposition 3 *Let* $\ln x_{i,t}$ *be the zero mean ARMA process*

$$\begin{aligned} \Phi(L) \ln x_{i,t} &= \Theta(L) u_{i,t} \\ u_{i,t} &\sim IN\left(0, \sigma_i^2\right) \end{aligned}$$

where $\Phi(L)$ *and* $\Theta(L)$ *are polynomials in the lag operator with roots inside the unit circle. If the distribution of the* σ *is dense around zero then, the leading term of* $\gamma_{X_t}(k)$ *decays at the same rate as* $\gamma_{\ln Z_t}(k)$, *i.e.*

$$\gamma_{\ln Z_t}(k) \simeq \gamma_{X_t}(k)$$

Thus, in general if the conditions of proposition 3 hold, then the leading term of the ACvF of a linear aggregate process coincides with the one of a non-linear aggregate process.

This result, together with proposition 2, shows the connection between the two aggregation approaches. Furthermore, it can also be seen as building up a link between the non-linear aggregate models and FARIMA models.

We evaluate now through simulation the goodness of this approximation. For this purpose we consider two cases. Firstly, we assume the $\omega's$ are uncorrelated, we generate a time series for each $\ln x_{i,t}$ and aggregate them both linearly and non-linearly in order to get time series data for $\ln Z_t$ and X_t. Finally, we evaluate the differences between their autocovariance functions for increasing values of the mean of the $\omega's$. The range of the mean considered goes from 0.005 to 0.1. It has to be mentioned that these values are not completely unrealistic. In fact, the mean of the standard deviations of many economic longitudinal data falls within this interval, see for instance the distributions of the TFP's that have been presented in chapter two.

We consider now the case when that $\omega's$ are correlated and distributed as the gamma density defined above and evaluate the approximation error between the two ACvF for different values of the mean and standard deviations.

Specifically, for each value of the mean in the interval defined above, we evaluate the effects of an increase in the spread of the gamma distribution. As mentioned above, for a gamma density it is always possible to increase its variance while keeping its mean constant. For the standard deviations we consider the interval from 0.0001 to 0.002. The simulation shows that the approximation error between the two ACvF increases not only as the mean of the $\omega's$ increases but also as their variance becomes larger. This implies that even if the distribution of the $\omega's$ has mean close to zero but is not concentrated around its mean, then X_t fails to be a good approximation of the dynamic properties of the "true" DGP. This means that the effects of isolated large variance processes build up instead of cancelling out, as N tends to infinity.

In figure Fig 4.1, we show the mean of the differences, between the ACvF's of X_t and $\ln Z$ for increasing values of $E(\omega)$.

figure 4.1

It can be clearly seen that for small values of ω the two ACvF are very close, while for values close to 0.1 the difference becomes substantial. Specifically, since the difference is positive, the process $\ln Z_t$ is characterized by a larger volatility and persistence compared to X_t.

This poor approximation for large values of the $\omega's$ is not surprising since $\gamma_{X_t}(k)$ represents the leading term of a small variance expansion of the autocovariance function of $\ln Z_t$. As ω increases, using the leading term of the approximation is not enough since the effects of the higher order terms are no longer negligible. As we show in the next graph, the same is true as the variance of the $\omega's$ becomes larger.

In the next figure we plot the difference of the two ACvF for different values of the mean and the variance of the $\omega's$.

figure 4.2

Figure 4.2 shows that the difference between $\gamma_{\ln Z_t}(k)$ and $\gamma_{X_t}(k)$ increases not only as the mean increases but also as the variance becomes larger. This suggests that the distribution of the standard deviations ω can affect the ACvF properties of the aggregate process $\ln Z_t$ both in terms of volatility and persistence. In fact, as the variance of the gamma distribution increases, occasionally larger values of the $\omega's$ are more likely to be observed. This has the effect to increase both the volatility and the persistence of the process $\ln Z_t$ as shown in fig. 4.2.

This kind of behavior can not be captured by linear aggregation since X_t tends to underestimate both the volatility and the persistence of $\ln Z_t$. In fact the difference between $\gamma_{\ln Z_t}(k)$ and $\gamma_{X_t}(k)$ becomes considerable for large values of the $\omega's$. This also shows that through non-linear aggregation the effect of these large variances builds up rather than canceling out as $N \to \infty$.

Hence, although some authors claim that X_t and $\ln Z_t$ are characterized by similar dynamics, if the micro component of the aggregate process are very volatile, then the dynamics of $\ln Z_t$ can not be fully captured by X_t which represent in this context a poor approximation of the true DGP. In this situation, the assumption of linearity of the aggregate process is not innocuous, not even as first order approximation, and could lead to dynamic misspecification due to the omitted non-linearities that characterize $\ln Z_t$.

4.4 Conclusion

In this chapter we compare the autocovariance properties of two different approaches to aggregate times series processes. The first is the standard linear aggregation proposed by Robinson (1978) and Granger (1980) that has become very popular in time series analysis over the last two decades. The other, is the non linear aggregation approach recently proposed by Abadir and Talmain (2002) which is, by construction, coherent with the way disaggregated data are put together in the national accounts database.

The importance of comparing these approaches relies on the fact that while the former has been widely studied in the literature, the latter is more consistent with the dynamics of the true "aggregate data generating process". We show that, unless certain conditions on the volatility of the microdata are satisfied, the two ways of aggregating data give rise to processes with very different dynamics. In particular, linear aggregation underestimates both the volatility and the persistence that would arise by the correct aggregation procedure.

4.5 List of figures

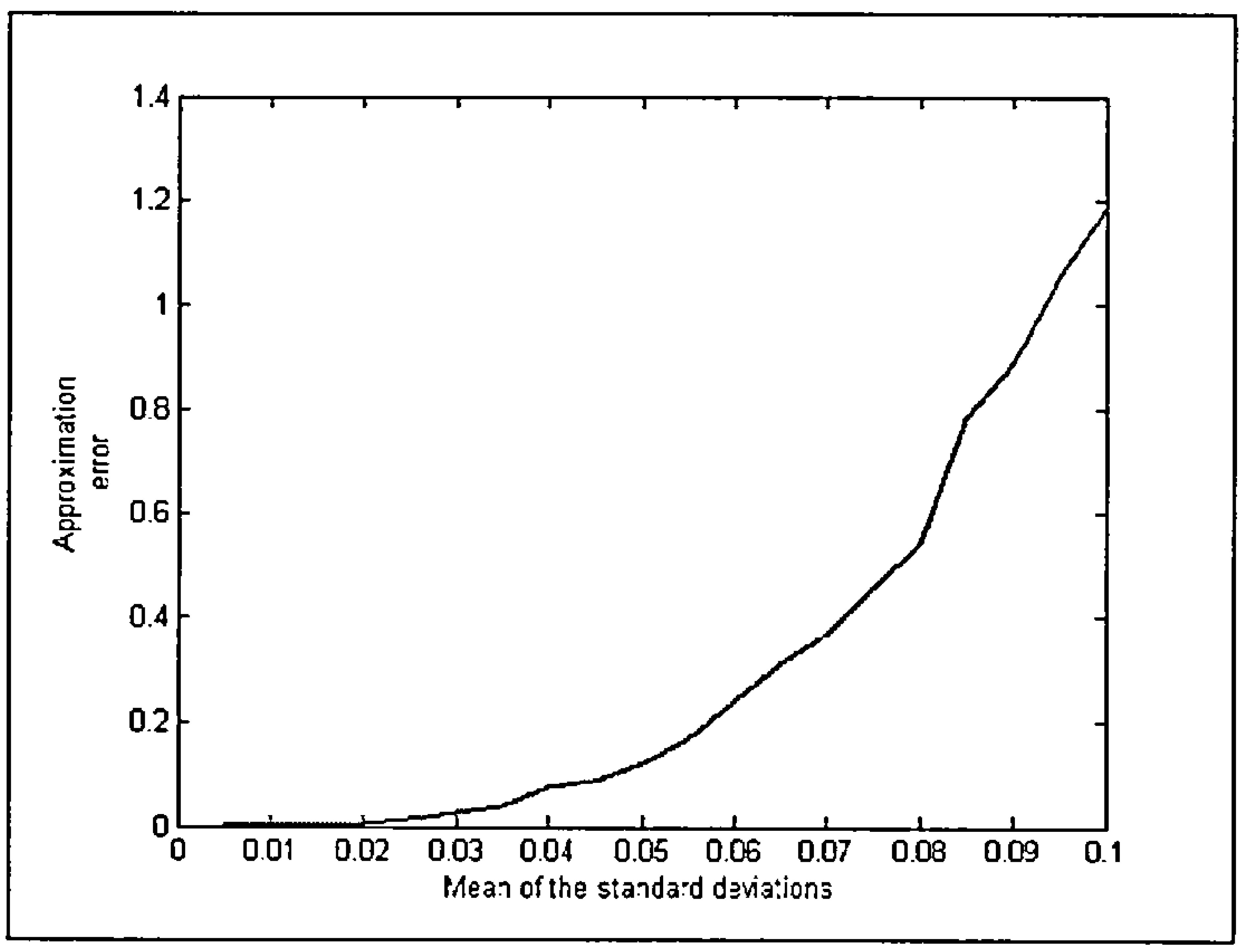

Figure 4.1: Difference between the autocovariance functions of lnZ and X for different mean of the omegas

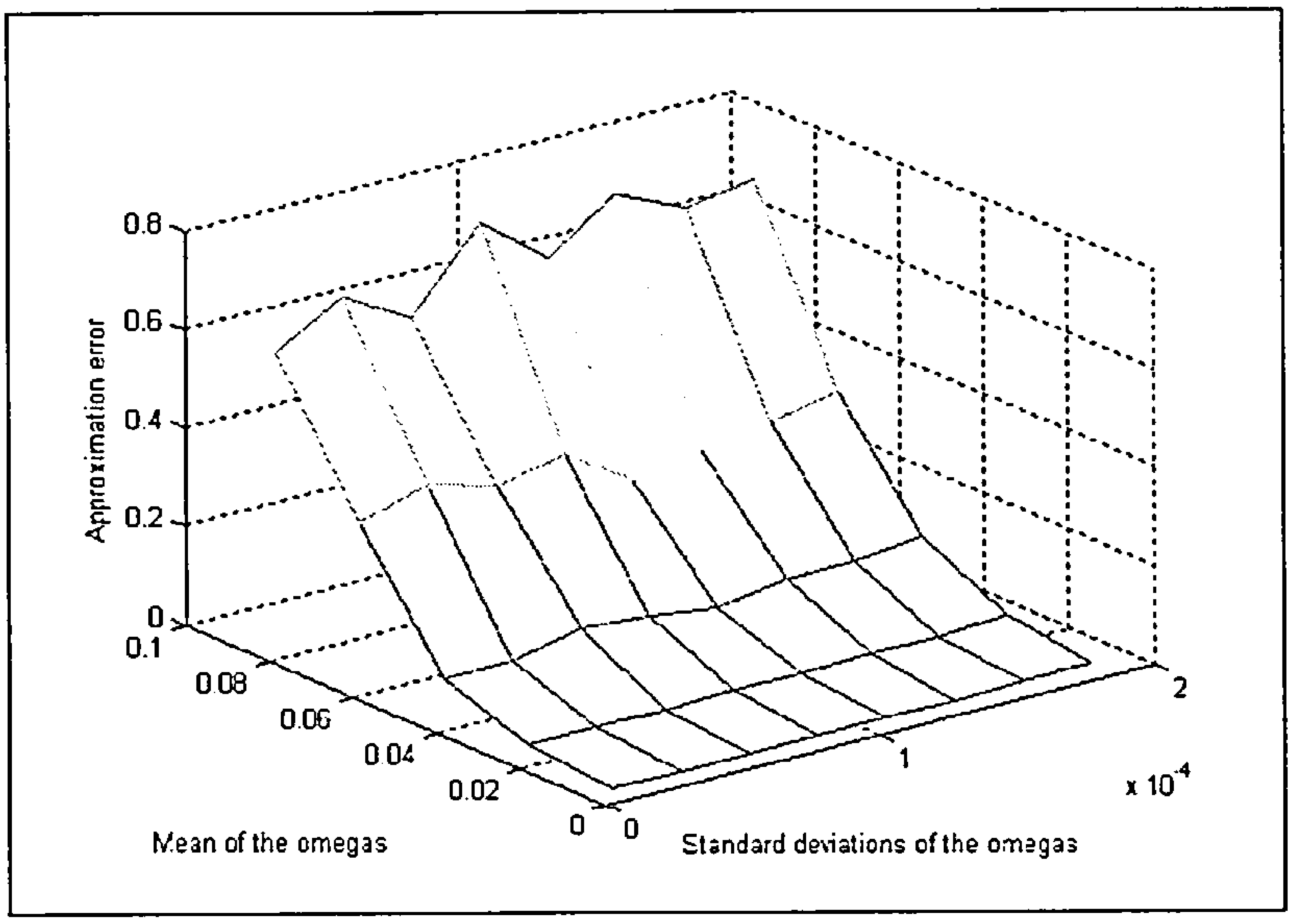

Figure 4.2: Difference between the two autocovariance functions for different values of the mean and standard deviations of the omega

4.6 Appendix

All the results presented in this appendix hold under the assumption stated above. In particular, the inclusion of almost unit roots in the aggregate process (i.e. the distribution $\beta(\alpha)$ is strictly increasing), stated in assumption 2, plays an important role. In fact, in many occasion we approximate the autocovariance function in a neighborhood of the mode of the distribution.

Proof of Proposition 1. Let us recall the aggregate process $X_t = (\sum_{i=1}^{N} \ln x_{i.t})$. Now, since $E(\ln x_{i.t}) = 0$ then $E(X_t) = 0$. Therefore,

$$\begin{aligned} Cov(X_t, X_{t+k}) &= E(X_t X_{t+k}) = \frac{1}{N^2} E\left[\left(\sum_{i=1}^{N} x_{i,t}\right)\left(\sum_{i=1}^{N} x_{i,t+k}\right)\right] \\ &= \frac{1}{N^2}\sum_{i=1}^{N} E\left[x_{i,t}x_{i,t+k}\right] + \frac{1}{N^2}\sum_{i=1}^{N}\sum_{s\neq i} E\left[x_{i,t}x_{s,t+k}\right] \simeq \frac{1}{N^2}\sum_{i=1}^{N}\sum_{s\neq i} E\left[x_{i,t}x_{s,t+k}\right] \end{aligned}$$

as $N \longrightarrow \infty$, the first term converges to zero, therefore

$$\begin{aligned} & Cov(X_t, X_{t+k}) \simeq \frac{1}{N^2}\sum_{i=1}^{N}\sum_{s\neq i} E\left[x_{i,t}x_{s,t+k}\right] \\ =\ & \frac{1}{N^2}\sum_{i=1}^{N}\sum_{s\neq i} E\left[\left(\sum_{j=0}^{t-1}\alpha_i^j \varepsilon_{i,t-j}\right)\left(\sum_{j=0}^{t+k-1}\alpha_i^j\varepsilon_{i,t+k-j}\right)\right] \\ =\ & \frac{1}{N^2}\sum_{i=1}^{N}\sum_{s\neq i}\frac{\alpha_s^k\left(1-(\alpha_i\alpha_s)^t\right)}{(1-\alpha_s\alpha_i)}\omega_i\omega_s \\ \simeq\ & \frac{1}{N^2}\sum_{i=1}^{N}\sum_{s\neq i}\left[t\alpha_s^k(\alpha_i\alpha_s)^{t-1}\omega_i\omega_s\right] \simeq E_i E_s\left[t\alpha_s^k(\alpha_i\alpha_s)^{t-1}\omega_i\omega_s\right] \\ \simeq\ & \int_0^1 \alpha_i^{g_a-1}(1-\alpha_i)^{h_a-1}\int_0^1 \alpha_s^{g_a-1}(1-\alpha_s)^{h_a-1} t\alpha_s^k(\alpha_i\alpha_s)^{t-1}\frac{d\alpha_i d\alpha_s}{B(g_\alpha,h_\alpha)^2}E(\omega_i)E(\omega_s) \\ =\ & \frac{\beta^2 t}{\alpha^2 B(g_\alpha,h_\alpha)^2} B(g_a+t-1,h_\alpha)B(g_a+t+k-1,h_\alpha) \\ =\ & \frac{\beta^2 t}{\alpha^2 B(g_\alpha,h_\alpha)^2}\frac{\Gamma(h_\alpha)^2\Gamma(g_a+t-1)}{\Gamma(g_a+t-1+h_\alpha)}\frac{\Gamma(g_a+t+k-1)}{\Gamma(g_a+t+k-1+h_\alpha)} \end{aligned} \tag{4.23}$$

where using l'Hopital's rule, we have replaced $\frac{\alpha_s^k\left(1-(\alpha_i\alpha_s)^t\right)}{(1-\alpha_s\alpha_i)}$ with $t\alpha_s^k(\alpha_i\alpha_s)^{t-1}$and the last approximations hold for large N. ■

We recall in the next lemma a result in Abadir and Talmain (2005) for logarithmic transformation, that will be often used below.

Lemma 3 *For a process $\{Z_t\}$ with elements $Z_t \in \mathbb{R}_+$, the autocovariance function of $\{\ln Z_t\}$ is obtained by letting $\lambda \to 0$ in the autocovariance function of $\{\lambda^{-1} Z_t^{\lambda}\}$, when they both exist. Additionally, the autocorrelation function of $\{\ln Z_t\}$ is obtained by letting $\lambda \to 0$ in the autocorrelation function of $\{Z_t^{\lambda}\}$.*

Proof. See pg. 231 of Abadir and Talmain (2005) ■

Proof of Proposition 2. Given the AR(1) processes $\ln x_{i,t} = \alpha_i \ln x_{i,t-1} + e_{i,t}$ defined above, it is quite straight forward to show that since $x_{i,t} = \exp(\sum_{j=0}^{t-1} \alpha_i^j e_{i,t-j})$ then

$$E_{t|i,s}\left(x_{i,t} x_{s,t+k}\right) = \exp\left(\frac{1-\alpha_i^{2t}}{2\left(1-\alpha_i^2\right)}\omega_i^2 + \frac{1-\alpha_s^{2t+k}}{2\left(1-\alpha_s^2\right)}\omega_s^2 + \frac{1-\alpha_i^t \alpha_s^t}{\left(1-\alpha_i \alpha_s\right)}\alpha_s^k \omega_s \omega_i\right) \tag{4.24}$$

where $E_{t|i,s}$ is the expectation with respect to time conditioned to the distribution of the parameters of x_i and x_s. This implies that

$$\begin{aligned}
Cov\left[Z_{t+k}, Z_t\right] &= \left[Cov\left(N^{-1}\sum_{i=1}^{N} x_{i,t+k}\right), \left(N^{-1}\sum_{i=1}^{N} x_{i,t}\right)\right] \\
&= E\left[\left(N^{-1}\sum_{i=1}^{N} x_{i,t+k}\right)\left(N^{-1}\sum_{i=1}^{N} x_{i,t}\right)\right] - E\left[\left(N^{-1}\sum_{i=1}^{N} x_{i,t+k}\right)\right] E\left[\left(N^{-1}\sum_{i=1}^{N} x_{i,t}\right)\right] \\
&= \frac{1}{N^2}\sum_{i=1}^{N} E\left[x_{i,t+k}x_{i,t} + \sum_{s\neq i}^{N} x_{s,t+k}x_{i,t}\right] - \frac{1}{N}\sum_{i=1}^{N} E\left[x_{i,t+k}\right]\frac{1}{N}\sum_{i=1}^{N} E\left[x_{i,t}\right] \\
&= \frac{1}{N^2}E\left[\sum_{i=1}^{N}\sum_{s\neq i}^{N} x_{s,t+k}x_{i,t}\right] - \frac{1}{N}\sum_{i=1}^{N} E\left[x_{i,t+k}\right]\frac{1}{N}\sum_{i=1}^{N} E\left[x_{i,t}\right]
\end{aligned}$$

where the last equality has been obtained using the fact that $N^{-2}\sum_{i=1}^{N} E\left[x_{i,t+k}x_{i,t}\right] \to 0$ for large N. By the laws of iterated expectation we have that $E\left[\sum_{i=1}^{N}\sum_{s\neq i}^{N} x_{s,t+k}x_{i,t}\right] =$

$E_{i,s}\left[E_{t|i,s}\left[\sum_{i=1}^{N}\sum_{s\neq i}^{N}x_{s,t+k}x_{i,t}\right]\right]$. Therefore,

$$\begin{aligned}
& Cov\left[Z_{t+k},Z_t\right] \\
= & E_{i,s}\left[\exp\left(\frac{1-\alpha_i^{2t}}{2\left(1-\alpha_i^2\right)}\omega_i^2+\frac{1-\alpha_s^{2t+k}}{2\left(1-\alpha_s^2\right)}\omega_s^2\right)\left(\exp\left(\frac{1-\alpha_i^t\alpha_s^t}{\left(1-\alpha_i\alpha_s\right)}\alpha_s^k\omega_s\omega_i\right)-1\right)\right] \\
= & \int_0^1\alpha_i^{g_\alpha-1}\left(1-\alpha_i\right)^{h_\alpha-1}\int_0^1\alpha_s^{g_\alpha-1}\left(1-\alpha_s\right)^{h_\alpha-1} \\
& E_{\omega_i}E_{\omega s}\left[\exp\left(\frac{1-\alpha_i^{2t}}{2\left(1-\alpha_i^2\right)}\omega_i^2+\frac{1-\alpha_s^{2t+k}}{2\left(1-\alpha_s^2\right)}\omega_s^2\right)\right. \\
& \left.\left(\exp\left(\frac{1-\alpha_i^t\alpha_s^t}{\left(1-\alpha_i\alpha_s\right)}\alpha_s^k\omega_s\omega_i\right)-1\right)\right]\frac{d\alpha_i d\alpha_s}{B\left(g_\alpha,h_\alpha\right)^2} \\
\simeq & \int_0^1\alpha_i^{g_\alpha-1}\left(1-\alpha_i\right)^{h_\alpha-1}\int_0^1\alpha_s^{g_\alpha-1}\left(1-\alpha_s\right)^{h_\alpha-1} \\
& E_{\omega_i}E_{\omega s}\left[\exp\left(\frac{t\alpha_i^{2(t-1)}}{2}\omega_i^2+\frac{\left(t+k\right)\alpha_s^{2(t+k-1)}}{2}\omega_s^2\right)\right. \\
& \left.\left(\exp\left(t\alpha_i^{t-1}\alpha_s^{t+k-1}\omega_s\omega_i\right)-1\right)\right]\frac{d\alpha_i d\alpha_s}{B\left(g_\alpha,h_\alpha\right)^2}
\end{aligned}$$

where we parameterized the $\alpha's$ using the beta density and the approximation has been obtained by l'Hopital's rule. Now, according to the results in Abadir and Talmain (2005), the leading term of the covariance function for $\ln Z_t$ can be obtained by a small variance expansion of the exponential term. The leading term of this expansion also characterizes the behavior of $Cov\left[\ln Z_{t+k},\ln Z_t\right]$ for values of the variance close to zero, which is what we are interested in. Therefore,

$$\begin{aligned}
& Cov\left[\ln Z_{t+k},\ln Z_t\right] \\
\simeq & \int_0^1\alpha_i^{g_\alpha-1}\left(1-\alpha_i\right)^{h_\alpha-1}\int_0^1\alpha_s^{g_\alpha-1}\left(1-\alpha_s\right)^{h_\alpha-1} \\
& E_{\omega_i}E_{\omega s}\left[t\alpha_i^{t-1}\alpha_s^{t+k-1}\omega_s\omega_i+o_p\left(\omega_i\omega_s\right)\right]\frac{d\alpha_i d\alpha_s}{B\left(g_\alpha,h_\alpha\right)^2} \\
= & t\int_0^1\alpha_i^{g_\alpha+t-2}\left(1-\alpha_i\right)^{h_\alpha-1}\int_0^1\alpha_s^{g_\alpha+t+k-2}\left(1-\alpha_s\right)^{h_\alpha-1} \\
& E_{\omega_i}E_{\omega s}\left[\omega_s\omega_i+o_p\left(\omega_i\omega_s\right)\right]\frac{d\alpha_i d\alpha_s}{B\left(g_\alpha,h_\alpha\right)^2} \\
= & \frac{\beta^2 t}{\alpha^2 B\left(g_\alpha,h_\alpha\right)^2}B\left(g_a+t-1,h_\alpha\right)B\left(g_a+t+k-1,h_\alpha\right)+o_p\left(\omega_i^2\omega_s^2\right)
\end{aligned}$$

which shows that the leading term of the RHS is the same quantity derived in eq. 4.23

■

Proof of proposition 3. Given the AR(1) processes defined above, we want to prove that

$$
\begin{aligned}
& Cov\left[\ln\left(N^{-1}\sum_{i=1}^{N}x_{i,t+k}\right), \ln\left(N^{-1}\sum_{i=1}^{N}x_{i,t}\right)\right] \\
= \; & Cov\left[\left(N^{-1}\sum_{i=1}^{N}\ln x_{i,t+k}\right), \left(N^{-1}\sum_{i=1}^{N}\ln x_{i,t}\right)\right]
\end{aligned}
$$

The key of the proof is to show that for small values of the variances ω_i

$$
\begin{aligned}
\lim_{\lambda\to 0}\lambda^{-2}Cov\left[Z_{t+k}^{\lambda}, Z_{t+k}^{\lambda}\right] &= \lim_{\lambda\to 0}\lambda^{-2}Cov\left[\left(N^{-1}\sum_{i=1}^{N}x_{i,t+k}\right)^{\lambda}, \left(N^{-1}\sum_{i=1}^{N}x_{i,t}\right)^{\lambda}\right] \\
&\simeq \lim_{\lambda\to 0}\lambda^{-2}Cov\left[\left(N^{-1}\sum_{i=1}^{N}x_{i,t+k}^{\lambda}\right), \left(N^{-1}\sum_{i=1}^{N}x_{i,t}^{\lambda}\right)\right] \qquad (4.25)
\end{aligned}
$$

Then, the result follows from Lemma 3. In fact, applied to the RHS of eq. 4.25 gives

$$
\begin{aligned}
& \lim_{\lambda\to 0}\lambda^{-2}Cov\left[\left(N^{-1}\sum_{i=1}^{N}x_{i,t+k}^{\lambda}\right), \left(N^{-1}\sum_{i=1}^{N}x_{i,t}^{\lambda}\right)\right] \\
= \; & Cov\left[\left(N^{-1}\sum_{i=1}^{N}\ln x_{i,t+k}\right), \left(N^{-1}\sum_{i=1}^{N}\ln x_{i,t}\right)\right]
\end{aligned}
$$

For space reasons we prove the theorem for the case of two variables. It can be easily extended to more than 2 variables by induction. However, the calculation becomes extremely cumbersome without changing the final result. Recalling lemma 3 we have that $Cov\left[\ln\left(Z_{t+k}\right), \ \ln\left(Z_t\right)\right]$ is given by

$$
Cov\left[\ln\left(Z_{t+k}\right), \ln\left(Z_t\right)\right] = \lim_{\lambda\to 0}\lambda^{-2}Cov\left[Z_{t+k}^{\lambda}, Z_{t+k}^{\lambda}\right]
$$

Now, for $N = 2$ the ACvF of Z_t^{ν} is equal to

$$
\begin{aligned}
& Cov\left(Z_{t+k}^{\nu}, Z_t^{\nu}\right) && (4.26) \\
= \; & E\left[Z_{t+k}^{\nu}Z_t^{\nu}\right] - E\left[Z_{t+k}^{\nu}\right]E\left[Z_t^{\nu}\right] && (4.27) \\
= \; & E\left[\left(\frac{x_{1,t+k}+x_{2,t}}{2}\right)^{\nu}\left(\frac{x_{1,t+k}+x_{2,t}}{2}\right)^{\nu}\right] - E\left[\left(\frac{x_{1,t}+x_{2,t}}{2}\right)^{\nu}\right]E\left[\left(\frac{x_{1,t}+x_{2,t}}{2}\right)^{\nu}\right] && (4.28)
\end{aligned}
$$

Using a binomial expansion[12] for $(\cdot)^{\nu}$ we can rewrite

$$\begin{aligned} Z_t^{\nu} &= \left(\frac{x_{1,t}+x_{2,t}}{2}\right)^{\nu} = \frac{1}{2}\left(\frac{x_{1,t}+x_{2,t}}{2}\right)^{\nu} + \frac{1}{2}\left(\frac{x_{1,t}+x_{2,t}}{2}\right)^{\nu} \\ &= \frac{1}{2^{1+\lambda}}\left(\sum_{n=0}^{\infty}\binom{\lambda}{n}x_{1,t}^{\lambda-n}x_{2,t}^{n}\right) + \frac{1}{2^{1+\lambda}}\left(\sum_{n=0}^{\infty}\binom{\lambda}{n}x_{2,t}^{\lambda-n}x_{1,t}^{n}\right) \end{aligned}$$

and $Cov\left(Z_{t+k}^{\nu}, Z_t^{\nu}\right)$ is given by

$$\begin{aligned} & Cov\left(Z_{t+k}^{\nu}, Z_t^{\nu}\right) \\ = & E\left[\left\{\frac{1}{2^{1+\lambda}}\left(\sum_{n=0}^{\infty}\binom{\lambda}{n}x_{1,t+k}^{\lambda-n}x_{2,t+k}^{n}\right) + \frac{1}{2^{1+\lambda}}\left(\sum_{n=0}^{\infty}\binom{\lambda}{n}x_{2,t+k}^{\lambda-n}x_{1,t+k}^{n}\right)\right\}\right. \\ & \left.\left\{\frac{1}{2^{1+\lambda}}\left(\sum_{n=0}^{\infty}\binom{\lambda}{n}x_{1,t}^{\lambda-n}x_{2,t}^{n}\right) + \frac{1}{2^{1+\lambda}}\left(\sum_{n=0}^{\infty}\binom{\lambda}{n}x_{2,t}^{\lambda-n}x_{1,t}^{n}\right)\right\}\right] \\ & -E\left[\left\{\frac{1}{2^{1+\lambda}}\left(\sum_{n=0}^{\infty}\binom{\lambda}{n}x_{1,t+k}^{\lambda-n}x_{2,t+k}^{n}\right) + \frac{1}{2^{1+\lambda}}\left(\sum_{n=0}^{\infty}\binom{\lambda}{n}x_{2,t+k}^{\lambda-n}x_{1,t+k}^{n}\right)\right\}\right] \\ & E\left[\left\{\frac{1}{2^{1+\lambda}}\left(\sum_{n=0}^{\infty}\binom{\lambda}{n}x_{1,t}^{\lambda-n}x_{2,t}^{n}\right) + \frac{1}{2^{1+\lambda}}\left(\sum_{n=0}^{\infty}\binom{\lambda}{n}x_{2,t}^{\lambda-n}x_{1,t}^{n}\right)\right\}\right] \end{aligned} \qquad (4.29)$$

It can be easily shown that $\lim_{\lambda\to 0}\left\{\lambda^{-2}Cov\left(Z_{t+k}^{\nu}, Z_t^{\nu}\right)\ \right\}$ is equal to zero for the terms of $Cov\left(Z_{t+k}^{\nu}, Z_t^{\nu}\right)$ that are higher than $O_p\left(\lambda^2\right)$. Therefore, multiplying and rearranging the term in eq. 4.29 we can focus only on the terms that are at most $O_p\left(\lambda^2\right)$

$$\begin{aligned} & Cov\left(Z_{t+k}^{\nu}, Z_t^{\nu}\right) \\ = & \frac{1}{2^{2(1+\lambda)}}\left\{\left[Cov\left(x_{1,t+k}^{\lambda},\ x_{1,t}^{\lambda}\right) + Cov\left(x_{1,t+k}^{\lambda},\ x_{2,t}^{\lambda}\right) + Cov\left(x_{1,t}^{\lambda},\ x_{2,t+k}^{\lambda}\right) + Cov\left(x_{2,t+k}^{\lambda},\ x_{2,t}^{\lambda}\right)\right]\right. \\ & +\lambda\left[Cov\left(x_{1,t+k}^{\lambda},\ x_{1,t}^{\lambda-1}x_{2,t}^{\lambda}\right) + Cov\left(x_{1,t+k}^{\lambda},\ x_{2,t}^{\lambda-1}x_{1,t}\right) + Cov\left(x_{2,t+k}^{\lambda},\ x_{1,t}^{\lambda-1}x_{2,t}\right)\right. \\ & \left. +Cov\left(x_{2,t+k}^{\lambda},\ x_{1,t}x_{2,t}^{\lambda-1}\right) + Cov\left(x_{1,t+k}^{\lambda-1}x_{2,t+k},\ x_{1,t}^{\lambda}\right) + Cov\left(x_{1,t+k}x_{2,t+k}^{\lambda-1},\ x_{1,t}^{\lambda}\right) + ...\right] \\ & +\frac{\lambda(\lambda-1)}{2}\left[... + Cov\left(x_{1,t+k}^{\lambda},\ x_{2,t+k}^{\lambda-2}x_{1,t}^{2}\right) ... + Cov\left(x_{1,t+k}^{\lambda},\ x_{2,t+k}^{2}x_{1,t}^{\lambda-2}\right)\right] + ... \\ & \left. +\lambda^2\left[... + Cov\left(x_{1,t+k}^{\lambda-1}x_{2,t},\ x_{1,t}^{1}x_{2,t+k}^{\lambda-1}\right) + Cov\left(x_{1,t+k}^{\lambda-1}x_{2,t},\ x_{1,t}^{\lambda-1}x_{2,t+k}^{1}\right) + ...\right] + ...\right\} \end{aligned} \qquad (4.30)$$

[12] With more than two variables we should have to use a multinomial expansion instead of the binomial expansion. Although it is not very complicated, it is very cumbersome.

Now, it is easy to show that by Lemma 3 all the terms in eq. 4.30 that are not $O_p(1)$ converge to zero as $\lambda \longrightarrow 0$. In fact, for any $m,\ n > 0$

$$
\begin{aligned}
& \lim_{\lambda\to 0}\left\{Cov\left(x_{i,t+k}^{\lambda-m}x_{j,t}^{m}\ ,\ x_{i,t}^{\lambda-n}x_{j,t}^{n}\right)+Cov\left(x_{i,t+k}^{\lambda-m}x_{j,t}^{m}\ ,\ x_{i,t}^{n}x_{j,t}^{\lambda-n}\right)\right\} \\
= \ & \lim_{\lambda\to 0}\{(\lambda-m)(\lambda-n)\,Cov(\ln x_{i,t+k},\ \ln x_{i,t})+(\lambda-m)\,nCov(\ln x_{i,t+k},\ \ln x_{j,t}) \\
& +m(\lambda-n)\,Cov(\ln x_{j,t},\ \ln x_{i,t})+mnCov(\ln x_{j,t},\ \ln x_{j,t}) \\
& +(\lambda-m)\,nCov(\ln x_{i,t+k},\ \ln x_{i,t})+(\lambda-m)(\lambda-n)\,Cov(\ln x_{i,t+k},\ \ln x_{j,t}) \\
& +mnCov(\ln x_{j,t},\ \ln x_{i,t})+m(\lambda-n)\,Cov(\ln x_{j,t},\ \ln x_{j,t})\} \\
= \ & 0
\end{aligned}
$$

This implies

$$
\lim_{\lambda\to 0}\lambda^{-2}Cov\left[\left(\frac{x_{1,t+k}+x_{2,t+k}}{2}\right)^{\lambda},\left(\frac{x_{1,t}+x_{2,t}}{2}\right)^{\lambda}\right] \tag{4.31}
$$

$$
\simeq \lim_{\lambda\to 0}\lambda^{-2}Cov\left[\left(\frac{x_{1,t+k}^{\lambda}+x_{2,t+k}^{\lambda}}{2}\right),\left(\frac{x_{1,t}^{\lambda}+x_{2,t}^{\lambda}}{2}\right)\right] \tag{4.32}
$$

Finally, the result follows by applying lemma 1 to eq. 4.32. ∎

Chapter 5

Concluding Remarks

In this last section we summarize the main findings of the thesis and discuss some further developments that could be potentially relevant in both macroeconomics and time series analysis.

The main aim of the thesis was to explore some of the implications of heterogeneity and structural aggregation on the dynamics of a standard real business cycle model as well as on the long run co-movements between macroeconomic variables. Our main findings can be summarized in the following points.

1. In the second chapter, we presented a heterogeneous RBC model where we allowed for cross sectional heterogeneity in the dynamics of the firm's productivities. We modelled this heterogeneity by fitting an autoregressive process whose parameters were estimated using micro data for 450 U.S. manufacturing firms. The heterogeneity allows the model to build up an internal propagation mechanism, that is actually missing in representative firm models, and generated dynamics in the simulated data that could mimic those of the U.S. data. In fact, using the approach in Cogley and Nason (2005) we failed to reject the hypothesis that both the autocorrelation and autocovariance function of the simulated data were significantly different from those of U.S. GDP. This suggests that heterogeneity and

structural aggregation could eventually be a solution to the"weak propagation mechanism puzzle" that affects standard RBC models.

2. In the third chapter, we presented a methodology to test for the presence long memory co-movements in the data when the variables are $I(1)$ and there exists a linear combination whose dynamics evolve as long memory process. This approach combines Engle and Granger (1987) methodology with the Maximum-Likelihood procedures for long memory process proposed by Abadir and Talmain (2005). We showed that, differently from standard cointegration technique, our approach was able to detect the presence of long-memory equilibrium relations. The test we proposed was, in fact, characterized by both high size under the null hypothesis and high power under the alternative. We reported also a significant reduction in the small sample estimation bias of the equilibrium relation compared to ordinary least square estimate. Finally, we applied our procedure to the data base for the UK purchasing-power parity and uncovered interest rate parity and showed that the null of no cointegration was rejected at 95% in contradiction with what was previously shown with standard single equation techniques.

3. In the fourth chapter, we compared standard linear aggregation approach, proposed by Robinson (1978) and Granger (1980) with the structural aggregation recently proposed by Abadir and Talmain (2002). We claimed that the importance of such comparison relies on fact that, while the former has been widely studied in the literature, the latter is more coherent with the way data are aggregated in national accounts. We showed, theoretically and through simulations, that unless certain conditions on the volatility of the microdata are satisfied, the two approaches of aggregating data give rise to processes with very different dynamics. In particular, linear aggregation tends to underestimate both the volatility and the persistence that would arise from the correct aggregation procedure.

We discuss briefly now some research topics that could have the potential for further development.

A first possibility for future research is to extend the model presented in chapter two to more general RBC models. Since in chapter two we were mainly interested in output dynamics and in showing that heterogeneity in the firm productivity can solve the "weak propagation mechanism puzzle", we did not consider many other interesting aspects. These involve monetary policy, fiscal policy, financial markets as well as labor markets. In this respect, heterogeneity has the potential to give a new insight into many controversial aspects of macroeconomics such as inflation or unemployment persistence. Furthermore, since differently with linear aggregation, this kind of microfounded model incorporates an aggregation structure, it can also help us to understand the micro relations responsible for generating nonlinear dynamics in the aggregate data.

A particularly interesting area of research is monetary policy. In fact, in the light of the result presented above, heterogeneity and aggregation are a good candidates in explaining persistence in monetary aggregates such as inflation and interests rates. The intuition is that once a strong propagation mechanism has been generated in the RBC model and if the monetary authority is reacting to innovations in productivity, this would lead to long lasting effects of the monetary policy. Furthermore, in this context, it would be also interesting to analyze the optimal response of a central bank to such movement in the aggregate output. A possible way to generate inflation persistence in DGE models could be through the construction of models with heterogeneous agents reacting with different degrees of persistence to nominal and real shocks. In chapter two, we showed that heterogeneity and aggregation succeeded to amplify significantly the effects of micro shock in a standard RBC model. In the light of these results, we believe that through the interaction of heterogeneous agents it is possible to improve the ability of monetary models their to reproduce inflation persistence.

With regards to econometrics methodology, the findings in chapter three suggest the need to develop new methods to deal with the high degree of persistence and non linearities that characterize actual data. In particular, the approach in chapter three can be developed further and extended to a multivariate case, in order to test for the presence of more than one "long memory equilibrium" relationship between the variables.

To conclude, it has to mentioned that many difficulties can arise from this approach. First, an immediate obstacle is the lack of micro data necessary to estimate the micro structure of a heterogeneous model. Although the number of surveys on individuals, such as families, consumers and firms, has increased significantly in the past few years, these database might not include all the information necessary to model the entire microstructure. Furthermore, most of them are not long enough to allow reliable estimates. Then, the high degree of non linearities that characterizes the microstructure of the model, could give rise to very complicated forms of the model that could not be either easily solved with standard technique or clearly interpreted. We have seen in chapter two, for instance, that the CES aggregation of the firm's productivities gave rise to a highly non linear aggregate productivity. This implied that the model had to be solved numerically since the aggregate process did not have a recursive closed form representation. Nevertheless, in the light of the results show in this thesis, we do not think that such difficulties should convince us to ignore the importance of heterogeneity and aggregation in explaining the complex dynamics that characterize many macroeconomic variables.

Bibliography

[1] Abadir, K. M. and Talmain, G. (2002), "Aggregation, persistence and volatility in a macromodel", Review of Economic Studies, 69, 749-779.

[2] Abadir, K. M. and Talmain, G. (2005), "Distilling co-movements from persistence macro and financial series", ECB working paper no. 525.

[3] Abadir, K. M., Caggiano, G. and Talmain, G. (2006)," Nelson and Plosser revised: the ACF approach", working paper 2005-7, University of Glasgow.

[4] Abramowitz, M. and Stegun, I. (1972),"Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables", Dover Publications.

[5] Abuaf, N. and Jorian, P. (1990), "Purchasing power parity in the long run", Journal of Finance, 45, 157-174.

[6] Adler, M. and Lehman, B. (1983), "Deviations for purchasing power parity in the long run", Journal of Finance, 38, 1471-1487.

[7] Altig, D., Christiano, L. J., Eichenbaum, M. and Linde, J. (2004), "Firm-specific capital, nominal rigidities and the business cycle", mimeo, Northwestern University.

[8] Altissimo, F., Mojon, B., Zaffaroni, P. (2005), "Fast micro and slow macro: can aggregation explain the persistence of inflation?", ECB Inflation Persistence Network Paper.

[9] Ambler, S., Cardia, E. and Zimmermann, C. (2004), "International business cycles: what are the facts?", Journal of Monetary Economics, 51, 257–276.

[10] Anderson, M.K. and Gredenhoff, M.P. (1999), "On the maximum likelihood cointegration procedure under fractional equilibrium error", Economic Letters, 65, 143-1.

[11] Backus, D. and Kehoe, P. (1992), "International evidence on the historical properties of business cycles", American Economic Review, 82, 864–888.

[12] Baillie, R.T and Bollerslev, T. (1994), "Cointegration, fractional cointegration and exchange rate dynamics", Journal of Finance, 49, 737-745.

[13] Baillie, R.T, Chung, C.F. and Tieslau, M.A. (1996), "Analyzing inflation by the fractionally integrated ARFIMA-GARCH model", Journal of Applied Econometrics, 11, 23-40.

[14] Bartelsman, E. J. and Dhrymes, P. J. (1998), "Productivity dynamics: U.S. manufacturing plants 1972-1986", Journal of Productivity Analysis, 9, 5-34.

[15] Bartelsman, E. J. and Gray, W. (1996), "The NBER manufacturing Database", NBER technical working paper 205.

[16] Basawa, I.V., Mallik, A.K., McCormick, W.P. and Taylor, R.L. (1996), "Bootstrapping unstable first order unit root processes", Annals of Statistics, 19, 1098-1101.

[17] Basu, S. (1998), "Technology and Business Cycles: How Well Do Standard Models Explain the Facts?", In: Fuhrer, J. C. and Schuh, S., eds., Beyond Shocks: What Causes Business Cycles, Conference Series 42, Federal Reserve Bank of Boston.

[18] Baxter, M. and King, R. G. (1993), "Fiscal policy in general equilibrium", American Economic Review, 83, 315-334.

[19] Baxter, M. and Crucini, M. (1993), "Explaining saving–investment correlations", American Economic Review, 83, 416–436.

[20] Beran, J. (1998), "Statistics for long-memory processes", Monograph on Statistics and Applied Probability, Chapman & Hall.

[21] Burnside, C., Eichenbaum, M. and Rebelo, S. (1993), "Labor hoarding and the business cycle shock", Journal of Political Economy, 101, 245-273.

[22] Burnside, C. and Eichenbaum, M. (1996), "Factor hoarding and the propagation of business cycle shock", American Economic Review, 86, 1154-1174.

[23] Carlstein, E. (1986), "The use of subseries values for estimating the variance of a general statistic from a stationary sequence", Annals of Statistics, 14, 1171-1179.

[24] Chambers, M. J. (1998), "Long memory and aggregation in macroeconomics", International Economic Review, 39, 1053-1072.

[25] Chari, V. V., Kehoe, P. J., and McGrattan, E. R. (2000), "Sticky price models of the business cycle: can the contract multiplier solve the persistence problem?", Econometrica, 68, 1151-1179.

[26] Cheung, Y. (1993), "Long memory in foreign-exchange rates", Journal of Business and Economic Statistics, 11, 93-101.

[27] Cheung, Y. and Lai, K.S. (1993), "A fractional cointegration analysis of purchasing power parity", Journal of Business and Economic Statistics, 11, 103-112.

[28] Cogley, T. and Nason, J. M. (1993), "Impulse dynamics and propagation mechanism in a real business cycle models", Economic Letters, 43, 77-81.

[29] Cogley, T. and Nason, J. M. (1995), "Output dynamics in Real-Business-Cycles models", American Economic Review, 85, 492-511.

[30] Cohen, A. C. and Whitten, B. J. (1988), "Parameter estimation in reliability and life span models". Marcel Dekker, New York.

[31] Crato, N. and Rothman, P. (1994), "Fractional integration analysis of long-run behavior for US macroeconomic time series", Economic Letters, 45, 287-291.

[32] Darby, M.R. (1983), "Movement in the purchasing power parity: the short and long runs", In: Darby, M.R., Lothian, J.R., Gandolfi, A.E., Schwartz, A.J. and Stockman, A.C.,The International Transmission of Inflation, University of Chicago Press, Chicago .

[33] Den Haan, W.J. and Marcet, A. (1994),"Solving the stochastic growth model by parameterizing expectations", Journal of Business and Economic Statistics, 8, 31-34.

[34] Den Haan, W., Ramey, G. and Watson, J. (2000), "Job destruction and propagation of shocks", American Economic Review, 90, 482–498.

[35] Diebold, F. X. and Rudebusch, G. D. (1989), "Long memory and persistence in aggregate output", Journal of Monetary Economics, 24, 189-209.

[36] Diebold, F.X. and Rudebusch, G. D. (1991a), "Is consumption too smooth? Long memory and the Deaton paradox", Review of Economics and Statistics, 73, 1-9.

[37] Diebold, F.X. and Rudebusch, G. D. (1991b), "On the power of the Dickey-Fuller tests against fractional alternatives", Economic Letters, 35, 155-160.

[38] Diebold, F.X., Husted, S., and Rush, M. (1991), "Real exchange rates under the gold standard", Journal of Political Economy; 99, 1252-1271.

[39] Diebold, F. X. and Senhadji, A. (1996), "The uncertain unit root in real GNP: Comment", American Economic Review, 86,1291-1298.

[40] Dolado, J. J., Gonzalo, J. and Mayoral, L. (2002), "A fractional Dickey-Fuller test for unit roots", Econometrica, 70, 1963-2006.

[41] Dotsey, M., King, R.G. and Wolman, A. L. (1999), "State-dependent pricing and the general equilibrium dynamics of money and output", Quarterly Journal of Economics, 114, 655-690.

[42] Efron, B. and Tibshirani, R. (1993), An introduction to the bootstrap, Chapman and Hall, London.

[43] Engle, R.F. and Granger, C.W.J. (1987), "Cointegration and error correction: representation, estimation and testing", Econometrica, 55, 251-276.

[44] Evans, M., Hasting N. and Peacock B. (2000), Statistical Distribution, Wiley Series in Probability and Statistics.

[45] Forni, M., and Lippi, M. (1997), "Aggregation and microfundations of dynamic economic" Oxford University Press.

[46] Forni, M., and Lippi, M. (1999), "Aggregation of linear dynamic microeconomic models", Journal of Mathematical Economics, 31, 131-158.

[47] Gil-Alana, L.A. (2003), "Testing of fractional cointegration in macroeconomic time series", Oxford Bulletin of Economics and Statistics 65, 517-529.

[48] Gonzàles, M. (1992), "Classical complex analysis", Marcel Dekker

[49] Gonzalo J. and Lee, T. (1998), "Pitfalls in testing for long run relationships", Journal of Econometrics, 86, 129-154.

[50] Gorman, W.M. (1953), "Community preference fields", Econometrica, 21, 63-80.

[51] Granger, C. W. J. (1980), "Long memory relationship and the aggregation of dynamic models", Journal of Econometrics, 14, 227-238.

[52] Granger, C.W.J. (1981), "Some properties of time series data and their use in econometric model specification", Journal of Econometrics, 16, 121-130.

[53] Granger, C. W. J. and Joyeux, R. (1980), "An introduction to long memory time series and fractional differencing", Journal of Time Series Analysis, 1, 15-30.

[54] Hall, P. (1992), The bootstrap and Edgeworth expansion, Springer, New York.

[55] Hansen, G. D. (1985), "Indivisible labor and the business cycle", Journal of Monetary Economics; 16, 309-327.

[56] Harris, R. (1995), Cointegration analysis in econometric modelling, Prentice Hall.

[57] Haslett, J. (1997), "On the sample variogram and the sample autocovariance for non-stationarity time series", Statistician, 46, 475-485.

[58] Hassler, U. and Wolters, J. (1994), "On the power of unit root test against fractional alternatives", Economics Letters, 45, 1-5.

[59] Hassler, U., and Wolters, J. (1995), "Long memory in inflation rates: International evidence", Journal of Business and Economic Statistics, 13, 37-45.

[60] Haubrich, J.G. and Lo, A.W. (1993), "The sources and nature of long-term memory in business cycle", working paper 9116, Federal Reserve Bank of Cleveland.

[61] Hosking, J. R. M. (1981), "Fractional differencing", Biometrica, 68, 165-176.

[62] Imbs, J., Muntaz, H., Ravn, M. and Rey H. (2005), "PPP strikes back: aggregation and the real exchange rate", Quarterly Journal of Economics, 120, 1-43.

[63] Keener, J. P. (1988), Principle of applied mathematics: transformation and approximation, Addison-Wesley.

[64] King, R.G., Plosser, C.I. and Rebelo, S.T. (1988), "Production, growth, and business cycles I: The basic neoclassical model", Journal of Monetary Economics, 21, 195-232.

[65] King, R.G. and Rebelo, S. (1999), "Resuscitating real business cycles", In: Taylor, J.B. and Woodford, M., eds. Handbook of macroeconomics, volume 1B, Amsterdam: Elsevier Science, North-Holland, 927-1007.

[66] Künsch, H.R. (1989), "The jackknife and the bootstrap for general stationary observations", Annals of Statistics, 17, 1217-1241.

[67] Johansen, S. (1995), Likelihood-based inference in cointegrated vector autoregressive models. Oxford University Press, Oxford.

[68] Johansen, S. (1988), "Statistical analysis of cointegrating vectors", Journal of Economic Dynamics and Control; 12, 231-254.

[69] Johansen, S. and Juselius (1992), "Testing structural hypotheses in a multivariate cointegration analysis of the PPP and the UIP for UK", Journal of Econometrics, 53, 211-244.

[70] Lee, D. and Schmidt, P. (1996), "On the power of KPSS test of stationarity against fractionally integrated alternatives", Journal of Econometrics, 73, 285-302.

[71] Lewbel, M. (1994), "Aggregation and simple dynamics", American Economic Review, 84, 905-918.

[72] Li, W.K. and McLeod, A.I. (1986), "Fractional time series modelling", Biometrica, 73, 217-221.

[73] Li, H. and Maddala, G.S. (1997), "Bootstrapping cointegrating regressions", Journal of Econometrics, 80, 297-318.

[74] Lippi, M., and Zaffaroni, P. (1998) "Aggregation of simple linear dynamics: exact asymptotic results", Econometric Discussion Paper, 350, STICERD-LSE.

[75] Merz, M. (1995), "Search in the labor market and the real business cycle", Journal of Monetary Economics, 36, 269–300.

[76] Micheal, P., Nobay, A. and Peel, D. (1997), "Transaction costs and nonlinear adjustments in real exchange rates: An empirical investigation", Journal of Political Economy, 105, 862-879.

[77] Michelacci, C. and Zaffaroni, P. (2000), "(Fractional) beta convergence", Journal of Monetary Economics, 45, 129-153.

[78] Muellbauer, J. (1976), "Aggregation, income distribution and consumer demand", Review of Economic Studies, 42, 525-543.

[79] Muellbauer, J. (1991), "Linear aggregation in neoclassical labour supply", Review of Economic Studies, 48, 21-36.

[80] Muellbauer, J. (1997), "The assessment: business cycles", Oxford Review of Economic Policy, 13, 1-18.

[81] Nelson, C. R. and Plosser, C. I. (1982), "Trends and random walks in macroeconomic time series: some evidence and implications". Journal of Monetary Economics 10, 139-162.

[82] Newbold, P. and Agiakloglou, C. (1993), "Bias in the sample auto-correlations of fractional noise", Biometrica, 80, 698-702.

[83] Obstfeld, M. and Rogoff, K. (2000), "The six major puzzles in international macroeconomics: is there a common cause?", in Bernanke, S. and Rogoff, K. (Eds), NBER Macroeconomics Annual, NBER, Cambridge, MA, pp.339-90.

[84] Olley, G. S. and Pakes, A. (1996)," The dynamics of productivity in the telecommunications equipment industry ", Econometrica, 64, 1263-1297.

[85] Pesaran, M.H. (2003), "Aggregation of linear dynamic models: an application to life-cycle consumption models under habit formation ", Economic Modelling, 20, 383-415.

[86] Politis, D.N. and Romano, J.P. (1994), "The stationary bootstrap", Journal of American Statistical Association, 89, 1303-1313.

[87] Rahbek, A. and Mosconi, R. (1999), "Cointegration rank inference with stationary regressors in VAR models", Econometrics Journal, 2, 76-91.

[88] Rebelo, S. (2005), "Real business cycle models: past, present and future", Scandinavian Journal of Economics, 107, 217-238.

[89] Robinson, P.M. (1978), "Statistical inference for a random coefficient autoregressive model", Scandinavian Journal of Statistics, 5, 163-168.

[90] Rotemberg, J. J. and Woodford, M. (1996), "Real-business-cycles models and the forecastable movements in output, hours, and consumption", American Economic Review, 86, 71-89.

[91] Rudebusch, G. D. (1989), "The uncertain unit root in real GNP", American Economic Review, 83, 264-272.

[92] Said, S., and Dickey, D.A. (1984), "Testing for unit roots in autoregressive moving average models of unknown order", Biometrika, 71, 599-608.

[93] Rudebusch, G. D. (1993), "The uncertain unit root in real GNP", American Economic Review, 83, 264-272.

[94] Shao, J. and Tu, D. (1995), The jackknife and bootstrap, Springer, New York.

[95] Smets, F. and Wouters, R. (2003), "An estimated dynamic stochastic general equilibrium model of the euro area", Journal of the European Economic Association, 1, 1123–1175.

[96] Sowell, F. (1992), "Modelling long-run behavior with the fractional ARIMA model", Journal of Monetary Economics, 29, 277-302.

[97] Stocker, T. M. (1984), "Completeness, distribution restrictions, and form of aggregate functions", Econometrica, 52, 887-907.

[98] Stocker, T. M. (1986), "Aggregation, efficiency, and cross-section regression", Econometrica, 54, 171-188.

[99] Theil, H. (1954), Linear aggregation of economic relations. Amsterdam: North-Holland.

[100] Taylor, A. (2001), "Potential pitfalls for the purchasing-power parity puzzle? Sampling and specification biases in mean-reversion tests of the law of one price", Econometrica, 69, 473-498.

[101] van Giersbergen, N.P.A. and Kiviet, J.F. (2002), "How to implement the bootstrap in static or stable dynamic regression models: test statistic versus confidence region approach", Journal of Econometrics, 108, 133-156.

[102] Watson, M.W. (1995), "Vector autoregression and cointegration", In: Engle, R.F., and McFadden, D., Handbook of Econometrics, vol.4, North-Holland, Amsterdam.

[103] Wei, W. S. (1990), Time series analysis: univariate and multivariate methods, Addison-Wesley.

[104] Zaffaroni, P. (2004),"Contemporaneous aggregation of linear dynamic models in large economies", Journal of Econometrics, 120, 75-102.