UNIVERSITY OF COPENHAGEN

Essays on Imperfect Knowledge Economics, Structural Change, and Persistence in the Cointegrated VAR Model

Tabor, Morten Nyboe

Publication date:
2014

Document version
Peer reviewed version

Citation for published version (APA):
Tabor, M. N. (2014). *Essays on Imperfect Knowledge Economics, Structural Change, and Persistence in the Cointegrated VAR Model*.

FACULTY OF SOCIAL SCIENCES
UNIVERSITY OF COPENHAGEN

PhD thesis

Morten Nyboe Tabor

Essays on Imperfect Knowledge Economics, Structural Change, and Persistence in the Cointegrated VAR Model

Academic advisor: Katarina Juselius

Submitted: 12/06/13

Essays on Imperfect Knowledge Economics, Structural Change, and Persistence in the Cointegrated VAR Model

Morten Nyboe Tabor

June 2013

PhD Thesis submitted to the

Department of Economics, University of Copenhagen

Supervisor: Katarina Juselius

Summary

This dissertation consists of three self-contained articles presented in three separate chapters. The overall aim is to provide a foundation for combining the new theoretical framework of Imperfect Knowledge Economics (IKE) developed by Frydman and Goldberg (2007, 2011) with the econometric methodology based on the cointegrated VAR model of Johansen (1996) developed at the University of Copenhagen. In the first chapter, I consider a simple general IKE asset pricing model and show how internal consistency can be fully incorporated in IKE models in a way that both allows for internal consistency compatible with individual rationality and accords individuals' expectations a partly autonomous role in driving aggregate outcomes. Moreover, I show how internal consistency conditions imply that the asset price and the exogenous variables in the model are cointegrated with stochastic cointegration parameters during different subperiods. Hence, I show how internal consistency is crucial for our ability to test empirical implications of IKE models based on the cointegrated VAR model and potential extensions which allow for stochastic cointegration parameters. In the second chapter, I simulate a simple model embedding key features of IKE and show that empirical regularities in the simulated data can be found using the cointegrated VAR model, despite bounded parameter-instability and stochastic cointegration in the data-generating process. Finally, the third chapter is a purely econometric article where I use simulations to show that the persistence frequently found in estimated cointegration relations—and corresponding low estimated adjustment parameters—can potentially be caused by stochastic cointegration parameters in the underlying data-generating process. Thereby the results in this thesis confirm the original intuition behind the attempt to combine IKE and the econometric approach based on the cointegrated VAR model, that the parameter-instability of IKE models could potentially be an important source of the persistence found empirically in estimated cointegrated VAR models for macroeconomic and financial time-series.

In the first chapter—*Combining Internal Consistency and Partly Autonomous Expectations in Imperfect Knowledge Economics Models*—I propose a set of conditions on the representation of expectations in terms of the parameters of the process underpinning aggregate outcomes in IKE models. The novelty of the conditions is that they allow for internal consistency compatible with individual rationality and yet accord expectations an autonomous role in driving aggregate outcomes.

I consider a simple general IKE asset pricing model with the key feature that it allows for nonrecurring structural change in the process underpinning aggregate outcomes. However, it is assumed that there are subperiods of varying length where these parameters are constant. As a consequence of nonrecurring structural change, individuals must base their expectations on contingent and inherently imperfect knowledge, and as a result expectations become an autonomous input to the

model which cannot be fully specified. However, until the recent paper by Frydman and Goldberg (2013 *a*), IKE models have not focused on internal consistency in specifications of expectations, so expectations have been accorded a completely autonomous role relative to the rest of the model.

I show how internal consistency can be incorporated in the IKE model by restricting the parameters in the representation of expectations to the class of stochastically trendless processes with unconditional means determined by the parameters of the process underpinning the asset price. However, I do not specify all changes over time in the parameters with a specific stochastic process within this class. The conditions imply that the forecasting errors and the gap between the asset price and its fundamental value in terms of the exogenous variables become stochastically trendless in each of the subperiods. Essentially this means that individuals' forecasting errors cannot deviate endlessly from zero over time and that the asset price cannot deviate endlessly from its fundamental value—although it does not imply that the forecasting errors will converge towards zero over time nor that the price will converge towards it fundamental value. Moreover, I show that the conditions imply that the asset price and the exogenous variables are cointegrated with stochastic cointegration parameters in each of the subperiods. Hence, I present a theoretical result which is crucial for the ability to test empirical implication of IKE models.

In the second chapter—*A Simulation Study of a Simple Imperfect Knowledge Economics Model of Stock Prices and Earnings with Cointegrated VAR Estimations*—I simulate outcomes from a simple IKE model of stock prices and earnings, which is based on the general IKE model of asset price swings and risk in Frydman and Goldberg (2013 *b*) and which satisfies the internal consistency conditions presented in Chapter 1. The aim is to address whether the cointegrated VAR model can serve as a valid statistical representation of the simulated data and whether the regularities in the simulated data can be found econometrically with the cointegrated VAR model as a first approximation.

The key features of the simple IKE model are: i) that there are streches of time where revisions of individuals' forecasting strategies are moderate, and ii) that fluctuations in the stock price around a benchmark price level determined by earnings are bounded. In modeling these key feature, earnings are assumed to fluctuate around a non-stationary long-run trend, with deviations caused by a bounded segmented trend process and a stationary component, and qualitative bounds are imposed on revisions of individuals' forecasting strategies, so that the causal parameter linking the stock price to earnings varies over time within specific bounds. Hence, the deviations between the stock price and the benchmark price determined by earnings are bounded and the variables are cointegrated as a linear relation between them is stochastically trendless.

The specification of the cointegrated VAR differs from the specification of the simulated data. Nonethelesss, the simulation results show that the cointegrated VAR model can serve as a statistically adequate representation of the simulated data. Moreover, the results show that, despite bounded instability in the time-varying cointegration parameters in the data-generating process of the simulated data, the cointegrated VAR model can provide a fairly precise estimate of the sample mean of the boundedly time-varying cointegration parameters. The results indicate that the cointegrated VAR model can serve as a good starting point for econometric analyses of IKE models, though more work is needed to fully establish this link.

In the third chapter—*Stochastic Parameters as a Source of Persistence in the Cointegrated VAR*

Model — A Simulation Study—I use simulations to show that persistence in estimated cointegration relations and slow adjustment can arise in the cointegrated VAR model as a consequence of stationary stochastic cointegration parameters in the underlying data-generating process. I simulate cointegrated data with stochastic cointegration parameters given by $\beta_t = \beta + B_t$, where B_t is a mean zero stationary autoregressive process simulated with different degrees of persistence and volatility. Hence, the linear relations $\beta_t' X_{t-1}$ and $\beta' X_{t-1}$ are stochastically trendless, and $\beta' X_{t-1}$ can be interpreted as the long-run average cointegration relations. The simulated data is analysed with the classic cointegrated VAR model—which has constant cointegration parameters β—using a general-to-specific modeling procedure, which first focuses on specification and testing of an unrestricted model as a valid statistical representation of the data, and second on testing for and estimating a reduced rank model with focus on the cointegration properties of the analysed data.

The results show that the estimated cointegrated VAR models appear to be fairly well-specified statistical representations of the simulated data, except from cases with high persistence and volatility in B_t, which result in non-normality of the estimated residuals, and in very long samples with $T = 1000$ observations. Moreover, the results show that the trace tests based on standard asymptotic inference on average suggest the correct reduced rank, except from the extreme cases mentioned, although the inference is sensitive to the misspecification caused by stochastic cointegration parameters. Finally, the results show that the cointegrated VAR model delivers a consistent and very precise estimate of the unconditional mean β of the stochastic cointegration parameters $\beta = \beta + B_t$, even in small samples. However, if there is persistence in the stochastic cointegration parameter, caused by persistence in B_t in the underlying data-generating process, it shows up in the estimated cointegrated VAR model as persistence in the estimated cointegration relations. As a result the estimated eigenvalues become small and the estimated adjustment coefficients skewed towards zero.

Thereby, the results show that stationary parameter-instability in the underlying data-generating process can potentially be a source of persistence in estimated cointegration relations and corresponding low estimated adjustment coefficients. Such persistence and slow adjustment is frequently found empirically in cointegrated VAR analyses of macroeconomic and financial data, and it has been a puzzle hard to explain for standard economic theory as it typically predicts a much faster adjustment to equilibrium and thereby less persistent deviations from the estimated equilibrium.

References

Frydman, R. and Goldberg, M. D. (2007), *Imperfect Knowledge Economics: Exchange Rates and Risk.*, Princeton University Press, Princeton.

Frydman, R. and Goldberg, M. D. (2011), *Beyond Mechanical Markets: Asset Price Swings, Risk, and the Role of the State*, Princeton University Press, Princeton.

Frydman, R. and Goldberg, M. D. (2013 *a*), The contingent expectations hypothesis: Rationality and contingent knowledge in macroeconomics and finance theory. Working paper prepared for the INET Annual Plenary Conference in Hong Kong, April 2013.

Frydman, R. and Goldberg, M. D. (2013 *b*), Opening Models of Asset Prices and Risk to Nonroutine Change, *in* R. Frydman and E. S. Phelps, eds, 'Rethinking Expectations: The Way Forward for Macroeconomics', Princeton University Press, Princeton, chapter 6, pp. 207–247.

Johansen, S. (1996), *Likelihood-Based Inference in Vector Autoregressive Models*, Oxford University Press, Oxford.

Resumé

Afhandlingen består af tre selvstændige artikler præsenteret i tre kapitler. Det overordnede mål for afhandlingen er at skabe et fundament for at kombinere det nye teoretiske framework Imperfect Knowledge Economics (IKE) udviklet af Frydman and Goldberg (2007, 2011) med den økonometriske metodologi, baseret på den kointegrerede VAR model af Johansen (1996), udviklet på Københavns Universitet. I det første kapitel viser jeg i en simpel IKE aktieprismodel, hvordan intern konsistens fuldt kan inkorporeres, således at modellen både er internt konsistent, og derved kompatibel med individuel rationalitet i modellen, samt tildeler individers forventningsdannelse en delvist autonom role i markedsudfald. Jeg viser, hvordan intern konsistens medfører, at aktieprisen og de eksogene variable i modellen er periodevis kointegrerede med stokastiske kointegrationsparametre. Jeg viser dermed, hvordan intern konsistens er afgørende for vores mulighed for at teste empiriske implikationer af IKE modeller og at dette kan gøres ved hjælp af den kointegrerede VAR model samt udvidelser med stokastiske parametre. I det andet kapitel simuleres en simpel model indeholdende hovedelementerne fra IKE og jeg viser, at empiriske regulariteter kan findes i en økonometrisk analyse baseret på den kointegrerede VAR model, til trods for begrænset ustabilitet i parametrene i den data-genererende process. Endeligt er det tredje kapitel en ren økonometrisk analyse, hvor jeg anvender simulationer til at vise, at den type persistens der regelmæssigt findes i estimerede kointegrationsrelationer, såvel som de tilsvarende lave estimerede tilpasningskoefficienter, potentielt kan være forårsaget af begrænset parameter-ustabilitet i den underliggende data-genererende process. Dermed bekræfter resultaterne den oprindelige intuition bag ønsket om at kombinere IKE teorien med den økonometrisk metode baseret på den kointegrerede VAR model, at den begrænset ustabilitet i IKE modeller potentielt kan være en væsentlig kilde til den persistens som regelmæssigt findes i empiriske analyser af makroøkonomiske og finansielle data baseret på den kointegrerede VAR model.

I det først kapitel—*Combining Internal Consistency and Partly Autonomous Expectations in Imperfect Knowledge Economics Models*—foreslår jeg et nyt sæt restriktioner på repræsentationen af forventningsdannelsen i forhold til parametrene i den process der understøtter aktieprisen i en IKE model. Det nye ved restriktionerne er, at de er internt konsistente, og dermed kompatible med individuel rationalitet i modellen, men samtidig tildeler forventningsdannelsen en autonom role i prissættelsen.

Jeg analyserer en simpel, generel IKE model for aktiepriser med det væsentlige hovedelement at den tillader ikke-gentagende strukturelle brud i processen der understøtter aktieprisen. Det antages dog, at der er del-perioder hvor disse parametre er konstante. Som en konsekvens af ikke-gentagende strukturelle brud må individer basere deres forventningsdannelse på kontingent og iboende ufuld-

kommen viden, og som følge deraf bliver forventningsdannelsen et autonomt input til modellen. Indtil den nylige artikel af Frydman and Goldberg (2013 *a*) har IKE modeller ikke fokuseret på intern konsistens i specifikationen af forventninger. Derved har forventningsdannelsen været tildelt en fuldstændig autonom rolle i forhold til resten af modellen.

Jeg viser, hvordan intern konsistens kan inkorporeres i IKE modellen ved at restriktere parametrene i repræsentationen af forventningsdannelsen til klassen af stokastisk trendløse processer med ubetinget middelværdi bestemt af parametrene i processen der understøtter aktieprisen. Dog specificerer jeg ikke ændringer i disse parametre over tid at følge en specifik stokastisk process inden for denne klasse. Restriktionerne medfører, at forventningsfejlene og afstanden imellem aktieprisen og dens fundamentale værdi i relation til de eksogene variable bliver stokastiske trendløse i hver del-periode. Dette betyder, at individers forventningsfejl ikke kan afvige vedvarende fra nul over tid, samt at aktieprisen ikke kan afvige vedvarende fra dens fundamentale værdi—dog betyder det ikke at forventningsfejlene konvergerer mod nul over tid eller at aktieprisen konvergerer mod sin fundamentale værdi over tid. Derudover viser jeg, at restriktionerne medfører, at aktieprisen og de eksogene variable er kointegrerede med stokastiske kointegrationsparametre i hver del-periode. Dermed præsenterer jeg et teoretisk resultat der er væsentligt for vores mulighed for at teste de empiriske implikationer af IKE modeller.

I det andet kapitel—*A Simulation Study of a Simple Imperfect Knowledge Economics Model of Stock Prices and Earnings with Cointegrated VAR Estimations*—simulerer jeg tidsserier fra en simpel IKE model for aktiepriser og virksomheders indtjening, som er baseret på den generelle IKE model for aktiepriser og risiko præsenteret i Frydman and Goldberg (2013 *b*) og som opfylder betingelserne for intern konsistens præsenteret i Kapitel 1. Formålet er, at addresere hvorvidt den kointegrerede VAR model i Johansen (1996) kan anvendes som en tilstrækkelig statistisk repræsentation af de simulerede data, samt hvorvidt regulariteterne i de simulerede data kan genfindes økonometrisk i den kointegrerede VAR model som en første approksimation.

De primære elementer i den simple IKE model er følgende: i) at der er tidsperioder, hvor revisioner af individers forventningsdannelse er moderate, samt ii) at udsving i aktieprisen omkring et benchmark prisniveau bestemt af indtjeningen er begrænsede. For at modelere disse primære elementer antages indtjeningen at flukturere omkring en ikke-stationær langsigtstrend, med afvigelserne forårsaget af en segmenteret trend og en stationær process. Desuden pålægges kvalitative restriktioner på revisioner af individers forventningsdannelse, således at den kausale parameter der forbinder aktieprisen til indtjeningen varierer over tid inden for specifikke grænser. Dermed bliver afvigelserne mellem aktieprisen og benchmarkprisniveauet, som er bestemt af den langsigtede trend i indtjeningen, begrænsede og variablene er kointegrerede, da en lineær relation imellem dem er stokastisk ikke-trendende.

Specifikationen af den kointegrerede VAR model er forskellig fra specifikationen af de simulerede data fra den simple IKE model. Alligevel viser simulationsresultaterne, at den kointegrerede VAR model kan anvendes som en tilstrækkelig statistisk repræsentation af de simulerede data. Derudover viser resultaterne, at den kointegrerede VAR model—til trods for begrænset tidsvariation i kointegrationsparameteren i den datagenererende process for de simulerede data—giver et relativt præcist estimat af gennemsnittet over tidsperioden af den begrænset tidsvarierende kointegrationsparameter. Samlet set indikerer resultaterne, at den kointegrerede VAR model kan anvendes som et godt udgangspunkt for økonometrisk analyse af IKE modeller, men mere arbejde er påkrævet for fuldt at etablere dette link.

I det tredje kapitel, *Stochastic Parameters as a Source of Persistence in the Cointegrated VAR Model — A Simulation Study*, anvender jeg simulationer til at vise, at persistens i estimerede kointegrationsrelationer og langsom tilpasning kan fremkomme i den kointegrerede VAR model, som en konsekvens af stationær parameter-ustabilitet i den underliggende data-genererende process. Jeg simulerer kointegrerede data med stokastiske kointegrationsparametre givet ved $\beta_t = \beta + B_t$, hvor B_t er en stationær autoregressiv process med ubetinget middelværdi nul, simuleret for forskellige grader af persistens og volatilitet, og hvor de lineære relationer $\beta' X_t$ er stokastisk trendløse. De simulerede data analyseres økonometrisk med den klassiske kointegrerede VAR model baseret på en general-til-specifik procedure. Først fokuseres på specifikation af modellen og test af en urestrikteret model som en valid statistisk repræsentation af data. Dernæst fokuseres på test for reduceret rang og en reduceret rang model estimeres med fokus på kointegrationsegenskaberne for de simulerede data.

Resultaterne viser, at den estimerede kointegrerede VAR model fremstår som en valid statistisk repræsentation af data, med undtagelse af tilfælde med høj persistens og volatilitet i B_t samt for meget lange tidsserier. Derudover viser resultaterne, at trace testet indikerer den korrekte reducerede rang, med undtagelse af de ekstreme tilfælde nævnt ovenfor. Endeligt viser resultaterne, at den kointegrerede VAR model giver et konsistent og meget præcist estimat af β, selv for korte tidsserier. Dog vil persistens i de stokastisk kointegrationsparametre B_t i den underliggende data-genererende process resultere i persistens i estimerede kointegrationsrelationer i den kointegrerede VAR model. Som følge heraf bliver de estimerede egenværdier meget små, i nogle tilfælge endda tæt på nul, og de estimerede tilpasningskoefficienter bliver skævvredet imod nul.

Dermed viser resultaterne, at stationær parameter-ustabilitet i den underliggende data- genererende process potentielt kan være en kilde til persistens i estimerede kointegrationsrelationer og dertil hørende lave estimerede tilpasningskoefficienter. Sådan persistens og langsom tilpasning findes regelmæssigt når makroøkonomiske og finansielle data analyseres empirisk med den kointegrerede VAR model og den har været svær at forklare for standard økonomisk teori der typisk forudsiger en langt hurtigere tilpasning til ligevægt, og derved mindre persistente afvigelser fra estimerede ligevægte.

References

Frydman, R. and Goldberg, M. D. (2007), *Imperfect Knowledge Economics: Exchange Rates and Risk.*, Princeton University Press, Princeton.

Frydman, R. and Goldberg, M. D. (2011), *Beyond Mechanical Markets: Asset Price Swings, Risk, and the Role of the State*, Princeton University Press, Princeton.

Frydman, R. and Goldberg, M. D. (2013 *a*), The contingent expectations hypothesis: Rationality and contingent knowledge in macroeconomics and finance theory. Working paper prepared for the INET Annual Plenary Conference in Hong Kong, April 2013.

Frydman, R. and Goldberg, M. D. (2013 *b*), Opening Models of Asset Prices and Risk to Nonroutine Change, *in* R. Frydman and E. S. Phelps, eds, 'Rethinking Expectations: The Way Forward for Macroeconomics', Princeton University Press, Princeton, chapter 6, pp. 207–247.

Johansen, S. (1996), *Likelihood-Based Inference in Vector Autoregressive Models*, Oxford University Press, Oxford.

Contents

Preface

This PhD thesis is the outcome of my PhD studies at the Department of Economics at the University of Copenhagen. The thesis marks the end of my PhD studies, but hopefully only the beginning of the ongoing project at the INET Center for Imperfect Knowledge Economics on combining Imperfect Knowledge Economics with the econometric approach developed in Copenhagen. Being part of the project has been extremely rewarding and inspiring, but also challenging and at times frustrating. I hope to continue working on this project and to further develop the ideas initiated in this thesis.

I would like to thank my supervisor Katarina Juselius for her support, patience, and guidance through my studies. Katarina's drive and enthusiasm have been, and still are, a great inspiration. I would like to thank Anders Rahbek, Søren Johansen, and Heino Bohn Nielsen for many great discussions, for always answering my questions, and for pointing me in the right direction. Moreover, I would like to thank fellow PhD students, in particular Andreas Hetland, and faculty members at the Department of Economics for many discussions from which I have learned a lot.

During my PhD studies I spent ten months as a visiting student at New York University from August 2010 to June 2011. I owe a special thanks to Roman Frydman for inviting me to New York University and introducing me to Imperfect Knowledge Economics. I am grateful for the faith Roman has shown in me, for his huge support, and for giving me opportunities beyond my imagination. Our discussions continue to inspire and I look forward to build on the collaboration. Furthermore, I thank Michael Goldberg for many inspiring discussions.

I am grateful for financial support from the Institute for New Economic Thinking (INET) and for inviting me to present part of the work in this thesis at the INET Annual Plenary Conference in Hong Kong in April 2013. Being part of the INET community, including the Young Scholars Initiative, has been extremely motivating for me and I am sure I will benefit even more from it in the future. Moreover, I am grateful for travel grants from Tuborgfondet and Danmark-Amerika Fondet-Rambøll, which made my studies in New York possible.

I would like to thank my family and friends for all their support. Finally, I owe the greatest thank to my girlfriend Signe for her loving support during ups and downs, for helping me back on track in frustrating times, and for listening to countless talks about my thesis.

Morten Nyboe Tabor

Copenhagen, June 2013

Chapter 1

Combining Internal Consistency and Partly Autonomous Expectations in Imperfect Knowledge Economics Models

Morten Nyboe Tabor[†]
June 12, 2013

Abstract

The promising feature of Imperfect Knowledge Economics (IKE) models, developed by Frydman and Goldberg (2007), is that by allowing for nonrecurring structural breaks in the process underpinning aggregate outcomes they accord individuals' expectations an autonomous role in driving aggregate outcomes. However, until recently IKE models have not focused on internal consistency in specifications of IKE models, and thereby expectations have been accorded a completely autonomous role relative to the rest of the model. In this paper, I consider a simple general IKE asset pricing model and I show how internal consistency can be incorporated by restricting the parameters in the representation of expectations to the class of stochastically trendless processes with unconditional means determined by the parameters of the process underpinning the asset price. The conditions imply that the representation of expectations becomes internally consistent with the process underpinning aggregate outcomes, yet they still accord individuals' expectations a partly autonomous role in driving aggregate outcomes. I show that the internal consistency conditions imply that in each subperiod with constant parameters in the process underpinning the asset price, the forecasting errors and the gap between the asset price and its fundamental value in terms of the exogenous variables become stochastically trendless. Essentially this means that individuals' forecasting errors cannot deviate endlessly from zero over time and that the asset price cannot deviate endlessly from its fundamental value. Moreover, I show that the conditions imply that in each subperiod the asset price is cointegrated with the exogenous variables with stochastic cointegration parameters. Hence, I provide a theoretical result which is crucial for the ability to test empirical implication of IKE models.

[†]Department of Economics, University of Copenhagen and INET Center for Imperfect Knowledge Economics at the University of Copenhagen. Email: morten.nyboe.tabor@econ.ku.dk.
Funding from the Institute for New Economic Thinking (INET) is gratefully acknowledged. I would like to thank Katarina Juselius, Roman Frydman, Kevin Hoover, Andreas Hetland, Søren Johansen, and Michael Goldberg for numerous discussions and comments to drafts of this paper.

1.1 Introduction

The key motivation behind the introduction of microfoundations in the pathbreaking Phelps et al. (1970) volume—which came to be known as the 'Phelps microfoundations volume' or just the 'Phelps volume'—was the "distinctive feature to accord market participants' expectations an autonomous role in economists' models of aggregate outcomes," Frydman and Phelps (2013, p. 1). Thus, individuals' expectations were seen as an important autonomous input to a theoretical model, which motivated Phelps' famous 'Island Model' where individuals form expectations independently on different 'islands'. In assigning individuals' expectations an autonomous role in driving aggregate outcomes the models in the Phelps volume relied on Adaptive Expectations as the representation of expectations. Hence, individuals revised their expectation of the aggregate outcome next period by a fixed proportion of their forecasting error in the current period.

However, as argued by Lucas (1996) such a representation of expectations is internally inconstistent with the structure of the model and incompatible with rational individual behavior within the model as individuals make systematic forecasting errors and forego obvious profit opportunities. To avoid such internal inconsistency between a model's representation of expectations and the process underpinning aggregate outcomes, influential papers such as Lucas (1976) relied on the Rational Expectations Hypothesis (REH) of Muth (1961) as the representation of expectations. Under REH, expectations correspond exactly to the time-invariant stochastic structure of a theoretical model. Implicitly it is assumed that there are no expectational coordination problems, so that a representative agent can be considered, and that the representative agent has complete knowledge of the structure and parameters of the model. Hence, the representative agent is assumed to have complete probabilistic knowledge of the structure underpinning aggregate outcomes at all future points in time. By basing expectations on this knowledge her expectations are internally consistent as they correspond exactly to the structure of the model and the representative agent does not make any systematic forecasting errors.

However, while expectations based on REH are internally consistent in time-invariant stochastic models—and therby compatible with rational individual decisionmaking given the assumptions of the model—they do not account individuals' expectations an autonomous role in driving aggregate outcomes as all changes in expectations over time are driven solely by exogenous shocks to the stochastic variables in the model, see Frydman and Phelps (2013) for a full discussion.

Imperfect Knowledge Economics (IKE) is a new theoretical framework for modeling aggregate macroeconomic and financial outcomes from the decision-making of rational individuals. The IKE framework has been developed and applied to the context of foreign exchange markets by Frydman and Goldberg (2007), while Frydman and Goldberg (2011) gives a nontechnical description and discussion of the implications of IKE. A crucial element of IKE models is that individuals must base their expectations on contingent and inherently imperfect knowledge as a consequence of nonrecurring structural breaks in the process underpinning aggregate outcomes. Hence, in the IKE models presented in Frydman and Goldberg (2007, 2013 *c*) individuals' expectations are totally autonomous compared to the rest of the model and revisions of expectations over time are restricted solely with qualitative conditions. However, recently Frydman and Goldberg (2013 *a*) focused on internal consistency as a core element in portraying rational individual decisionmaking in IKE models, but exclusively by restricting the parameters entering expectations to have the same sign as the parameters in the rest of the model.

In this paper, I show how internal consistency can be fully incorporated in IKE models in a way that both allow for internal consistency compatible with individual rationality and accord expectations a partly autonomous role within an IKE model.

I consider a simple general asset pricing model similar to the one discussed in Frydman and Goldberg (2013 *b*,). The key feature of the model is that it allows for nonrecurring structural breaks in the parameters of the stochatic processes of exogenous variables as well as the parameters linking them to the asset price. However, it is assumed that there are subperiods of varying length where these parameters are constant. Due to the nonrecurring structural breaks, individuals in the IKE model must base their expectations on contingent and inherently imperfect knowledge, which implies that revisions of their expectations over time cannot be fully specified in terms of a mechanical revision-rule for all points in time. Due to inherently imperfect knowledge of the current process underpinning the asset price and expectations of structural breaks in the future individuals make systematic forecasting errors over time, in the sense that the forecasting errors depend on the exogenous variables and thereby are correlated with information available at the time of the expectation formation. However, as IKE models so far have not focused on internal consistency between the representation of expectations and the process underpinning aggregate outcomes, the forecasting errors can be trending over time and the asset price can trend away from the asset's fundamental price, so that the gap between the asset price and its fundamental price can be trending.

The new internal consistency conditions I propose in this paper restrict the parameters entering the representation of expectations in each subperiod to belong to the class of stochastically trendless processes (Harris et al., 2002; McCabe et al., 2003) with unconditional means given by the non-expectational parameters of the model. Thereby, I establish a link between the parameters in the representation of expectations and the process underpinning the asset price, but I do not represent the forecasting parameters with a specific stochastic specification within this class of stochastic processes. The conditions imply that the forecasting errors, while depending on the non-stationary exogenous variables, become stochastically trendless. Moreover, the conditions imply that in each of the subperiods the gap between the asset price and its fundamental value based on the exogenous variables is stochastically trendless, though potentially quite persistent. Thereby, expectations based on contingent and inherently imperfect knowledge cause the asset price to fluctuate around the fundamental value, but the gap between them is restricted not to be trending endlessly in each subperiod. However, the conditions do not imply neither that the forecasting errors converge towards zero over time nor that the asset price converges towards its fundamental value.

Moreover, I show that the internal consistency conditions are crucial for our ability to test the empirical implications of IKE models. The internal consistency conditions imply that the asset price and the exogenous variables are cointegrated within each of the subperiods. This follows from the result that in each subperiod the gap between the asset price and the fundamental value based on the exogenous variables is stochastically trendless. Hence, a linear combination of the asset price and the exogenous variables becomes stochastically trendless and the variables are cointegrated, although with time-varying cointegration parameters with constant unconditional mean in each subperiod. Finally, I briefly discuss how a 'data-first approach' to econometrics, see Hoover (2006), based on the cointegrated VAR model of Johansen (1996) and extensions thereof can be used test these empirical implications. The challenge for testing empirical implications of IKE models is that they are partly open as they do not specify a complete stochastic specification which can be directly estimated. However, the 'data-first approach' and the underlying methodology based on a 'general-to-specific' modeling approach is suitable as it does not require that the theoretical model delivers a complete stochastic specification which can be directly estimated. Rather the 'data-first approach' starts with the specification of a stastical model as as valid statistical representation of the data, and thereby the complete stochastic specification is derived from the data for a specific sample rather than from economic theory. Although structural breaks cannot be specified in advance based on the theoretical model, they can nonetheless be identified in an econometric model for historical data. Moreover, although an IKE model does not impose a specific stochastic process for the time-varying parameters, the internal consistency conditions restrict the class of stochastic processes and within this class it can econometrically be addressed if a specific stochastic specification can adequately represent the data for a specific sample of data.

The rest of this paper is organized in the following way. In section 1.2, I briefly discuss internal model consistency as the core element in modeling rational individual decisionmaking in formal economic models. Section 1.3 present the general model considered throughout the paper. In Section 1.4 a time-invariant stochastic version of the model with REH is discussed, while Section 1.5 discusses the IKE version of the model. Section 1.6 briefly discusses how a 'data-first approach' to econometrics can be used to test implications of the IKE model. Section 1.7 concludes.

1.2 Individual Rationality as Internal Model Consistency

Like the majority of contemporary economics, IKE models rely on microfoundations in modeling aggregate macroeconomic and financial outcomes directly from formal modeling of rational decision-making of the individuals composing the economy. Thus, a core assumption of IKE—like most of economics in general—is that individuals can be modeled as rational economic agents. Rationality, in a broad sense, means that an individual is optimizing in pursuit of specific goals. In economic decisions rationality means that an individual optimally chooses the actions that satisfy her own desires, whatever they might be. However, the notion of what is optimal depends crucially on the context and setting faced by an individual.

In a formal economic model the *principle of internal consistency* (Frydman and Goldberg, 2013 *a*) is a core element of portraying rational individual decisionmaking: in a formal economic model an individual's actions must be consistent with her assumed goals and her understanding of aggregate outcomes—upon which her expectations are based—must be consistent with the model's representation of the aggregate outcomes, as influentially argued by Lucas (1976, 1996). Hence, it is the assumptions made in a specific model—not reality—which determine how to portray individual decisionmaking in a internally consistent way compatible with rational decisionmaking within the model. Whether the context faced by the individuals in a specific model corresponds to the context faced by individuals in the real world and whether individuals' model-consistent behavior corresponds to rational behavior in the real world are separate questions, which can essentially be reduced to a debate about the empirical relevance of the specific model in explaining real-world outcomes.

In an economic model the assumed goal of an individual is typically to maximize her utility as given by a specific utility function. Internal consistency requires that the individual consistently chooses the actions which maximize her utility, subject to the constraints she faces and given her expectations of the uncertain future outcomes of her actions. Choosing feasible actions which are not utility-maximizing would obviously be inconsistent with her assumed goal of utility maximization, and it would obviously be irrational for the individual as another feasible action could give her a greater utility. Hence, given specific assumptions about the individual's utility function, constraints, and expectations her rational decision-making can be formally modeled in an internally consistent way by solving a utility maximization problem.

However, most economic decisions—such as consumption, savings, and investment decisions—depend on uncertain future outcomes, so to derive the utility-maximizing behavior in a formal model a representation of the individual's expectations of the future outcomes of her actions is needed. The principle of internal consistency requires that the individual's understanding of the process underpinning aggregate outcomes must be consistent with the economic model's representation of this process, so that her expectations are consistent with the actual outcomes, see Lucas (1976, 1996) and in terms of IKE models the recent paper by Frydman and Goldberg (2013 *a*). To Lucas (1976, 1996), an important requirement for internal consistency is that a model's representation of expectations cannot lead individuals to make systematic forecasting errors. As argued by Lucas (1996) in discussion of REH, internal consistency requires that individuals do not make systematic forecasting errors, as they would thereby forego obvious profit opportunities which is incompatible with rational individual behavior within the model. However, this argument depends crucially on what kind of knowledge is assumed possible for individuals within a formal model, which is determined by core assumptions about representations of change over time in the model. Under REH a representative agent is assumed to have complete knowledge of the time-invariant stochastic model and obviously only using all this knowledge would be internally consistent and lead to unsystematic forecasting errors. However, if individuals are assumed not to have complete knowledge or if complete knowledge is impossible within a model by design, individuals will make systematic forecasting errors in the sense that the forecasting errors are correlated with information available at the time the forecasts were made. However, if this is the case it must still be a requirement for internal consistency that the representation of expectations is linked to the structure of the rest of the model, so that at a minimum expectations are not completely autonomous and diverging from the structure of the rest of the model.

In the next section, I use a simple asset pricing model to discuss how different core assumptions about the specification of change over time lead to different internally consistent specifications of expectations in contemporary economic models and in IKE models. In particular, different assumptions about changes over time lead to different degrees of knowledge possible and thereby to different internally consistent specification of the expectation formation of rational individuals.

1.3 A Simple Linear General Asset Pricing Model

Following Frydman and Goldberg (2013 b), consider the simple general linear asset pricing model in reduced form given by

$$P_t = a_t + b_t' X_t + c_t \widehat{P}_{t|t+1} + \epsilon_{p,t} \tag{1.1}$$

$$X_t = X_{t-1} + \mu_t + \epsilon_{x,t} \tag{1.2}$$

where an asset price P_t depends on the $(k \times 1)$ vector of exogenous variables in X_t and the aggregate forecast of next period's asset price $\widehat{P}_{t|t+1}$. The parameters b_t and μ_t are $(k \times 1)$ vectors, while a_t and c_t are scalars. It is assumed that $0 < c_t < 1$ for all t. The random shocks $\epsilon_{p,t}$ and $\epsilon_{x,t}$ are assumed mutually uncorrelated and identically and independently distributed mean zero Gaussian with variance σ_p^2 and covariance Σ_x, respectively.[1] The k variables in X_t are assumed exogenous relative to the asset price, so there is no feedback from the asset price to X_t, but in general the variables in X_t could be internally related.

Define the sets of parameters $\theta_{p,t} := \{a_t, b_t, c_t\}$ and $\theta_{x,t} := \{\mu_t\}$. Moreover, let $\theta_t := \{\theta_{p,t}, \theta_{x,t}\}$. The parameters in $\theta_{x,t}$ determine the specification of the exogenous variables X_t, while the parameters in $\theta_{p,t}$ are derived from the preferences of the individuals in the economy and determine how the exogenous variables load into the asset price at each point in time. Without any restrictions on changes over time in the parameters in θ_t and a representation of expectations $\widehat{P}_{t|t+1}$ the model is completely open and has no empirical implications.

For the representation of expectations to be internally consistent it must be consistent with the structure in equations (1.1) and (1.2). Hence, time t expectations of the asset price at time $t+1$ must take into account that P_{t+1} depends on X_{t+1}, $\widehat{P}_{t+1|t+2}$, and the parameters $\theta_{p,t+1}$. Likewise, the time t expectation of the time $t+1$ expectation of the asset price at time $t+2$ must take this structure into account. Iterating this argument forward $s > 0$ periods implies the following formulation for an internally consistent representation of expectations

$$\widehat{P}_{t|t+1} = f(\widehat{X}_{t|t+1}, ..., \widehat{X}_{t|t+s}; \widehat{\theta}_{p,t|t+1}, ..., \widehat{\theta}_{p,t|t+s}; \widehat{P}_{t|t+s}), \tag{1.3}$$

where $\widehat{(\cdot)}_{t|t+s}$ denotes the time t expectation of $(\cdot)_{t+s}$ and it is assumed that $\widehat{P}_{t|(t+i|t+i+1)} = \widehat{P}_{t|t+i+1}$. Equation (1.2) can be used to write $\widehat{X}_{t|t+\tau} = X_t + \sum_{i=1}^{\tau} \widehat{\mu}_{t|t+i}$ for $\tau > 0$ (assuming that $\widehat{\epsilon}_{t|t+\tau} = 0$) and without loss of generality it follows that

$$\widehat{P}_{t|t+1} = f(X_t; \widehat{\theta}_{t|t+1}, ..., \widehat{\theta}_{t|t+s}; \widehat{P}_{t|t+s}). \tag{1.4}$$

Hence, an internally consistent representation of expectations of the future asset price must be linked to the causal variables X_t and the parameters θ_t, and forming expectations about the future asset price within the model requires forming expectations about the future changes in the causal variables as well as expectations about their future impact on the asset price, i.e. forming expectations about the future values $\theta_{x,t+s}$ and $\theta_{p,t+s}$ for $s = 1, 2, ...$

[1] For simplicity it is assumed that the random shocks follow a time-invariant distribution, but in general that need not be the case for an IKE model. For example, the variance of the exogenous shocks can be assumed to vary over time by pre-multiplying the exogenous shocks with a time-varying factor. However, this would not change the conclusions in this section.

Without loss of generality consider the representation of expectations at time t given by

$$\widehat{P}_{t|t+1} = \alpha_t + \beta_t' X_t, \tag{1.5}$$

where α_t is a scalar and β_t is a $(k \times 1)$ vector. From equation (1.4) it follows that internal consistency requires that the parameters α_t and β_t are linked to the expected future values of the parameters in θ_t. However, to what extend knowledge at time t about future changes in the causal variables and their future impact on the asset price is possible within the model depends crucially on the assumptions made about changes over time in the parameters in θ_t. First, the assumptions made about changes over time in μ_t determine to what extend X_{t+i} can be forecasted at time t for $i > 0$. Second, given forecasts of X_{t+i}, the assumptions made about changes over time in $\theta_{p,t}$ determine to what extend P_{t+i} can be forecasted at time t.

The link between the forecasting parameters α_t and β_t and the parameters in θ_t defines the forecasting errors made by individuals over time. The forecasting error at time $t+1$, defined as $fe_{t+1} := P_{t+1} - \widehat{P}_{t|t+1}$, is given by

$$fe_{t+1} = \mathcal{A}_{t+1} + \mathcal{B}_{t+1} X_t + \upsilon_{t+1} \tag{1.6}$$

where

$$\mathcal{A}_{t+1} := (a_t + c_i \alpha_{t+1} - \alpha_t) + (b_t + c_t \beta_{t+1})' \mu_t \tag{1.7}$$

$$\mathcal{B}_{t+1} := (b_t + c_t \beta_{t+1} - \beta_t)' \tag{1.8}$$

$$\upsilon_{t+1} := \epsilon_{p,t+1} + (b_t + c_t \beta_{t+1})' \epsilon_{x,t+1}. \tag{1.9}$$

Hence, in general the forecasting error at time $t+1$ depends on the causal variables X_t and may be non-zero on average. To what extend individuals make systematic and non-zero forecasting errors on average over time depends on the link between the forecasting parameters α_t and β_t and the parameters in θ_t. Internal consistency requires as a minimum that the forecasting errors are not endlessly trending away from zero or positive or negative at all points in time, which could indeed be the case if the forecasting parameters α_t and β_t are totally autonomous relative to the parameters in θ_t.

The key difference between IKE models and contemporary economic models is the assumptions made about changes over time in the parameters in θ_t and given these assumptions the different following internally consistent representation of expectations. This is the focus of the next sections.

The core of contemporary economics assumes a time-invariant structure where $\theta_t = \theta$ for all t and relies on the Rational Expectations Hypothesis as an internally consistent representation of expectations. Implicitly it is assumed not only that $\theta_t = \theta$ for all t, but also that individuals know the values of θ and that there are no expectational coordination problems, so that a representative agent can be considered. Thus, the representative agent is implicitly assumed to have complete probabilistic knowledge of all future outcomes: at any point in time t all future values X_{t+s} can be forecasted based on the time-invariant stochastic process in equation (1.2) with $\theta_{x,t} = \theta_x$, and given these forecasts the asset price at all future times can be forecasted based on equation (1.1) with $\theta_{p,t} = \theta_p$. Thereby the representative agent does not make any systematic forecasting errors. The internally consistent general representation of expectations in equation (1.4) reduces to a time-invariant function of X_t and the known values of θ, so that $\alpha_t = \alpha$ and $\beta_t = \beta$ in the representation of expectations in equation (1.5) can be reduced to time-invariant functions of the parameters in θ for all t. Thereby the process underpinning individual and aggregate outcomes is represented with the exact same stochastic structure at all points in time and the asset price equals its fundamental value given by the present discounted value of all future values of the exogenous variables in X_t. All changes over time—including all changes in expectations of future outcomes—are driven solely by the stochastic shocks from a fixed probability distribution, so revisions of expectations play no autonomous role.

By contrast, IKE models assume that there are periods of varying length where θ_t is constant—so that $\theta_t = \theta_i$ for $t = T_{i-1} + 1, ..., T_i$ for $i = 1, 2, ...$ and with $T_{i-1} < T_i$—and importantly that θ_i changes in nonrecurring ways over time.[2] As a consequence of nonrecurring structural breaks in θ_t, individuals cannot, at any point in time t, gain complete knowledge of θ_{t+s} for all $s > 0$. Thereby, complete probabilistic knowledge of all future outcomes with only stochastic risk is, by design, impossible in IKE models. Hence, an internally consistent representation of expectations must take into account that individuals' expectations must be based on contingent and inherently imperfect knowledge. This implies that revisions of individuals' expectations formation cannot be fully specified with a mechanical revision-rule and that expectations may exhibit larger jumps as the contingent knowledge changes.[3] However, I show that conditions can still be imposed on the representation of expectations so that the representation of expectations becomes internally consistent and compatible with individual rationality, yet still allows for a partly autonomous role of individual expectations in driving aggregate outcomes.

First, I impose the qualitative conditions of 'guardedly moderate revisions' on changes in the forecasting parameters attached to the causal variables X_t, as suggested by Frydman and Goldberg (2007, 2013 *c*,), though I impose the conditions on each individual forecasting parameters rather than the vector of parameters. However, based on these conditions alone the representation of expectations is totally autonomous compared to the process underpinning market outcomes and essentially the forecasting parameters are allowed to diverge endlessly from the parameters θ_i over time.

Therefore, I propose an additional set of restrictions on the forecasting parameters which ensure internal consistency and still allow for a partly autonomous role for expectations. I propose to restrict the parameters of the representation of expectations in each subperiod i to the class of stochastic processes which are stochastically trendless with unconditional means given by values determined by θ_i. However, I do not specify all changes in the forecasting parameters within each subperiod to follow specific stochastic processes, only that the processes must have this general feature. I show how this restriction implies internal consistency between the representation of expectations and the structure of the IKE model, and, moreover, that in each subperiod i the forecasting errors as well as the gap between the asset price and its fundamental value determined by the exogenous variables become stochastically trendless. This just means that the representation of expectations is restricted not to be totally autonomous relative to the rest of the model, so that the deviation between the asset price and its fundamental value as well as the forecasting errors are not trending

[2] To be precise, I define nonrecurring structural changes as the assumption that the number of 'regimes' $i \to \infty$ as $t \to \infty$.

[3] See for example Dow (2012) for a full discussion about expectations under risk versus under genuine uncertainty.

over time. However, in each subperiod i the asset price can deviate persistently from its fundamental value and individuals can make persistent forecasting errors because they base their expectations on contingent and inherently imperfect knowledge, but the internal consistency conditions imply that the fundamental gap and forecasting errors are not endlessly trending, so that the average fundamental gap and the average forecasting errors converge in probability towards zero.

1.4 Contemporary Economics: A Time-Invariant Stochastic Structure

A core positive heuristic of contemporary economic models is to fully specify all changes in the causal structure of individual and aggregate outcomes over time, see for example Hoover (1991) and Dow (2012). This implies that all relevant causal variables and parameters characterizing individual and aggregate outcomes at all points in time are fully prespecified with deterministic or stochastic rules and procedures. First, a stochastic specification is assumed for all exogenous variables, so that all changes over time are driven by exogenous shocks drawn randomly from a specific probability distribution, which is typically the standard Gaussian distribution. Second, individuals' preferences and constraints are typically assumed constant over time and specified with constant parameters in terms of the set of exogenous variables. Finally, individuals' expectations are specified as functions of the same set of causal variables, where the parameters are either assumed constant over time or fully prespecified with a mechanical procedure determining how expectations change over time as a function of the shocks to the causal variables. Based on the assumed preferences, constraints, and expectations the rational individual decisionmaking is deduced by solving the individuals' utility maximization problem, and based on an aggregation procedure the aggregate outcomes are derived. Thereby, individual and aggregate outcomes at all points in time—past, present, and future—are specified with a time-invariant stochastic structure. All changes over time driven by random exogenous shocks from a specific probability distribution and the specification of preferences, constraints, and expectations determine how these shocks load through the system.

In the simple model considered above this implies that the parameters in θ_t are assumed constant over time, $\theta_t = \theta$ for all t.[4] First, by assuming $\theta_{x,t} = \theta_x$ for all t, the exogenous variables X_t are specified with the same stochastic process at all times. Second, assuming that individual preferences and constraints are constant over time implies that $\theta_{p,t} = \theta_p$ for all t. If a representative agent is assumed the parameters in θ_p are derived directly from the assumed preferences and constraints of the representative agent, but if the model allows for heterogeneity the aggregate parameters are derived by aggregating over the individuals in the economy.

Assuming $\theta_t = \theta$ for all t, and moreover that $0 < c < 1$, the model is given by

$$P_t = a + b' X_t + c\hat{P}_{t|t+1} + \epsilon_{p,t} \tag{1.10}$$

$$X_t = X_{t-1} + \mu + \epsilon_{x,t}, \tag{1.11}$$

[4] Alternatively changes in θ_t are represented with a specific stochastic process, but in that case the model still implies a time-invariant stochastic structure, where all changes over time are driven by stochastic shocks from a specific probability distribution. For example, the stochastic formulation of X_t might allow for switching between a fixed number of recurring states, such as μ_1 and μ_2, with fixed switching probabilities.

for all t. This implies that the process underpinning the asset price in equations (1.10) and (1.11) is fixed and identical at all points in time, and only a representation of the expected future price is needed to close the model. The process X_t is non-stationary with a drift given by the vector μ, which cumulates into a linear deterministic trend. All change over time is driven by the exogenous shocks $\epsilon_{p,t}$ and $\epsilon_{x,t}$, which are assumed drawn randomly from a specific probability distribution, and the constant parameters θ determine how the exogenous shocks load through the system. Now only a representation of the expectations of the future asset price is needed to finalize the model.

From the assumption $\theta_t = \theta$ for all t it follows that complete knowledge of all future outcomes as well as their likelihoods is possible within the model based on the time-invariant structure in equations (1.10) and (1.11). However, knowledge can be assumed incomplete as individuals can be assumed either to have limited knowledge about the parameters in θ or limited information in the form of limited access to the variables in X_t. Though, in both cases complete knowledge is possible. In the former case, individuals can learn about the parameter values over time by relying solely on standard statistical methods, and thereby they can gain complete knowledge of future outcomes, upon which to base their expectations.[5] In the latter case, complete knowledge is possible if only individuals gain access to all relevant information, i.e. access to all variables in X_t, though that might have a cost. Hence, when the process underpinning the aggregate market outcome is assumed fixed over time, knowledge of about future outcomes can be complete or incomplete, where complete knowledge is defined in a probabilistic sense and incomplete knowledge can be defined as cases where complete knowledge can be obtained by design.

1.4.1 The Rational Expectations Hypothesis

The core of contemporary economics relies on the Rational Expectations Hypothesis (REH) of Muth (1961) in representing expectations (Colander, 2006; Caballero, 2010; Dow, 2012). Under REH, the aggregate expectations equal the mathematical expectation of the economic model

$$\hat{P}^{REH}_{t|t+1} = E[P_{t+1}|I_t], \tag{1.12}$$

where I_t is all available information up to time t. The model is internally consistent as the representation of expectations of the aggregate outcome corresponds exactly to the model's representation of the aggregate outcome. By design, the representative agent[6] is assumed to have complete knowledge

[5] In learning models the stochastic specification of the exogenous variables is typically assumed time-invariant and known, i.e. $\theta_{x,t} = \theta_x$ for all t, so the only unknown parameters would be those corresponding to θ_p. However, because θ_p is assumed constant over time, individuals use standard statistical methods—typically ordinary least squares regressions—to learn about the parameters and revise their expectations over time, see for example Evans and Honkapohja (2001, 2013) for an overview of learning models. The key question to address is then if individuals' expectations converge towards the REH representation of expectation over time, so that ultimately they gain complete knowledge of the initially unknown parameters. As shown by Evans and Honkapohja (2013) this depends crucially on the initial values and the learning gain parameters, which determine how much the forecasting parameters are updated in each period, and in many cases the learning rules are diverging rather than converging towards complete knowledge.

[6] For a discussion about coordination of individual expectations in a Rational Expectations Equilibrium (REE), see Guesnerie (2005, 2013) and Frydman and Phelps (2013). Guesnerie (2013) considers an "eductive game" where individuals are assumed to known the full structure and parameters of a specific model, but in forming expectations must take the expectations of others into account. Guesnerie examines whether a mental process of forming expectations about other individuals' expectations eventually leads all individuals to base their expectations on Rational Expectations, so that there is expectational coordination with a Rational Expectations Equilibrium (REE). Guesnerie

of future outcomes in probabilistic terms as all time-invariant parameters in θ are assumed known and there is access to all the relevant information X_t. Hence, the decisionmaking of the representative agent takes the form of 'well-informed optimising choice in pursuit of specific goals', Dow (2012, p. 7), which is the center of the 'Walrasian macroeconomic research program' as described in Colander (2006). The representative agent can, at any point in time, take all potential outcomes and their likelihoods into account, and thereby her optimal decision plan for all points in time can be deduced by solving a single utility maximization problem.

By plugging in for P_{t+1} in equation (1.12), iterating forward, using the law of iterated expectations $E[E[X_{t+s+1}|I_{t+s}]|I_t] = E[X_{t+s+1}|I_t]$ for all $s > 0$, and applying a transversality condition the time t forecast of the asset price at time $t+1$ is given by

$$\begin{aligned}\hat{P}^{REH}_{t|t+1} &= \tilde{a} + b'E[X_{t+1}|I_t] + cb'E[X_{t+2}|I_t] + c^2b'E[X_{t+3}|I_t] + \dots \\ &= \tilde{a} + \sum_{i=0}^{\infty} c^i b' E[X_{t+1+i}|I_t] \end{aligned} \tag{1.13}$$

where $\tilde{a} := a/(1-c)$. This is a formalization of the general expression in equation (1.3). Hence, the asset price forecast at time t depends on the present discounted value of all future X_t and a constant term determined by the preference parameters a and c. At any point in time, all future realizations of the causal variables, X_{t+s} for $s > 0$, can be forecasted up to a random error term from the current observations X_t. Based on the stochastic specification in equation (1.11) the s-period ahead forecast of X_t is given by

$$E[X_{t+s}|I_t] = X_t + s\mu, \tag{1.14}$$

and as the effect of the causal variables on the asset price is fixed, b, all future realizations of the asset price can be forecasted up to a random error term with a fixed probability distribution. Hence, the time t forecast is given by

$$\hat{P}^{REH}_{t|t+1} = \alpha^{REH} + \beta^{REH\prime} X_t \tag{1.15}$$

where

$$\alpha^{REH} := \frac{a}{1-c} + \frac{b'\mu}{(1-c)^2} \quad \text{and} \quad \beta^{REH} := \frac{b}{1-c}. \tag{1.16}$$

Thus, the asset price forecast is a time-invariant function of the causal variables X_t and the constant parameters α^{REH} and β^{REH} are direct functions of the parameters in θ. It is worth noting that the weight attached to X_t, i.e. β^{REH}, depends only on the parameters b and c, while the constant drift in X_t given by the vector μ only enters the constant term α^{REH}. Over time all changes in expectations are driven by the changes in X_t, and as α^{REH} and β^{REH} are constant revisions of expectation formation play no autonomous role in driving the aggregate outcomes, see also Frydman and Phelps (2013) for a discussion.

Under REH the asset price is given by

$$P^{REH}_t = \tilde{\alpha}^{REH} + \beta^{REH\prime} X_t + \epsilon_{p,t}, \tag{1.17}$$

where

$$\tilde{\alpha}^{REH} := \frac{a}{1-c} + \frac{cb'\mu}{(1-c)^2}. \tag{1.18}$$

finds that in general this is not the case, though there are parameter values for which REE occurs, see also Frydman and Phelps (2013).

Hence, at every point in time the asset price equals what can be defined as its fundamenatal value $P_t^{\star REH}$ determined by the present discounted value of all future X_t, as given by

$$P_t^{\star REH} := \tilde{\alpha}^{REH} + \beta^{REH\prime} X_t, \tag{1.19}$$

plus a random error term $\epsilon_{p,t}$, so that all deviations from the asset price's fundamental value over time are random. An important insight from this is that the fundamental value of the asset price depends not only on the current X_t and the preference-parameters in θ_p, but also on the drift term μ as it determines the future deterministic growth rate in X_t.

Moreover, the forecasting error at time $t+1$ becomes

$$fe_{t+1}^{REH} = \epsilon_{p,t+1} + \beta^{REH\prime} \epsilon_{x,t+1}, \tag{1.20}$$

as $\mathcal{A}_{t+1} = \mathcal{B}_{t+1} = 0$. Hence, the forecasting error is uncorrelated with the information at time t. Thereby, by design of REH the representative agent does not make any systematic forecasting errors, and the internally consistent representation of expectations is compatible with rational behavior of the representative agent. By contrast, any deviation from the REH representation of expectations in equation (1.15) would lead the representative agent to make systematic forecasting errors and thereby forego obvious profit opportunities, which would be incompatible with rational individual behavior as argued by Lucas (1996). Hence, unless specific assumptions about limited knowledge or limited information are assumed, REH follows as the only representation of expectations internally consistent with the model in equations (1.10) and (1.11) and the asset price must equal its fundamental value determined by the present discounted value of all future X_t. However, these conclusions follow from the assumption that $\theta_t = \theta$ for all t—so that the stochastic structure underpinning the aggregate outcome is time-invariant—combined with the assumption that the representative agent knows θ. Obviously, if the underlying structure is assumed fixed and known, only expectations based on this structure are rational for the representative agent, and consequently the asset price will equal its fundamental value and the forecasting errors will be random.

1.5 Imperfect Knowledge Economics

The IKE framework is based on Popper's (1990) fundamental insight that "quite apart from the fact that we do not know the future, the future is objectively not fixed. The future is open: objectively open", see Frydman and Goldberg (2007, 2011, 2013 *a*). Hence, IKE models are partly open; they allow for nonrecurring structural change in the both parameters of the stochastic processes characterizing exogenous variables and those linking them to the endogenous variables, i.e. they allow for structural change in the parameters in θ_t.

However, IKE assumes that there are periods of varying length where the parameters in θ_t can be represented as constant, as given by

$$\theta_t = \theta_i \quad \text{for} \quad t = T_{i-1} + 1, ..., T_i \text{ and } i = 1, 2, ..., \tag{1.21}$$

and where $T_{i-1} < T_i$ for all i, so that

$$\begin{aligned}
\theta_t &= \theta_1 \quad \text{for } t = 1, ..., T_1 \\
\theta_t &= \theta_2 \quad \text{for } t = T_1 + 1, ..., T_2 \\
&\vdots
\end{aligned}$$

and for each subperiod i for $t = T_{i-1} + 1, ..., T_i$ and $i = 1, 2, ...$, the model is given by

$$P_t = a_i + b_i' X_t + c_i \widehat{P}_{t|t+1} + \epsilon_{p,t} \tag{1.22}$$

$$X_t = X_{t-1} + \mu_i + \epsilon_{x,t} \tag{1.23}$$

Hence, the process underpinning aggregate market outcomes is contingent: it changes at unanticipated times and in nonrecurring ways over time. The crucial element here is the nonrecurring parameter-values over time as these imply that at any point in time t complete knowledge of the parameters at all future points in time $t + s$ for $s > 0$ is impossible by design. Whether or not the breakpoints where the structural break in θ_t occur can be anticipated or not is of secondary relevance, because even if they can be anticipated or assigned a probability the new parameters after a breakpoint cannot be known in advance due to the nonrecurring structure. Moreover, it is important to note that the nonrecurring structure does not rule out some repititive features over time—it just does not assume that θ_i swicthes between a fixed number of values, say θ_1 and θ_2, as this would imply that complete knowledge of the underlying structure becomes possible. Furthermore, the nonrecurring structural breaks might involve different compositions of the variables in X_t if $b_{m,i} = 0$ for some $m = 1, 2, ..., k$ and some i. Finally, for simplicity I here consider the case where the breaks in all parameters in θ_t occur at the same points in time, but in general the breaks can be allowed to occur at different points in time in an IKE model.

Due to the nonrecurring structural breaks in θ_i, complete knowledge of the process underpinning the asset price at all future points in time is impossible by design. In particular, there is no way individuals can gain complete knowledge of all future values θ_{i+s} for all $s > 0$ solely by relying on statistical methods based on historical data up to a specific point in time. Internal consistency requires that individuals in the model must be aware of the occurence of these nonrecurring structural breaks. Hence, they must base their expectations on contingent and inherently imperfect knowledge.[7] Knowledge is inherently imperfect in the sense that complete probabilistic knowledge of all future outcomes is impossible; even if an individual is assumed to know the values of θ_i at a specific point in time t she can never know the values of θ_{i+s} for $s > 0$. Thereby it might indeed not even be optimal to base her expectations on her knowledge of θ_i as she would be aware that the current structure might cease to be relevant at some point in time in the future. Moreover, she would have to take into account the expectation formation of others, including expectations of other individuals' expectations regarding structural breaks in θ_i. Finally, the contingency of knowledge implies that knowledge is fallible, see Soros (1987), and thereby it might be subject to sudden changes and shifts.

[7] For a full discussion of expectations under quantifiable risk and unquantifiable 'Knigthian uncertainty' of Knight (1921), see chapter Dow (2012). A key element of expectations under genuine uncertainty is that rational individuals rely on calculations based on fundamental factors as well as psychological and social considerations and conventions, see also Chapters 7 and 9 in Frydman and Goldberg (2007).

Because expectations are based on contingent and inherently imperfect knowledge, the parameters α_t^{IKE} and β_t^{IKE} in the representation of expectations, given by[8]

$$P_{t|t+1}^{IKE} = \alpha_t^{IKE} + \beta_t^{IKE\prime} X_t, \tag{1.24}$$

cannot be explicitly linked to the parameters in θ_i, although internal consistency requires as a minimum that $sign(\beta_t^{IKE}) = sign(b_i)$ for $t = T_{i-1}+1, ..., T_i$ as recently argued by Frydman and Goldberg (2013 *a*). Moreover, changes in the forecasting parameters α_t^{IKE} and β_t^{IKE} cannot be fully specified with a mechanical revision-rule as there is simply no way to select a specific rule which can adequately describe how rational individuals would revise these when basing their expectations on contingent and inherently imperfect knowledge. Indeed, individuals can have very different expectations of future outcomes even when they have access to exactly the same information as they interpret it in different ways, and there is simply no way to judge a priori exactly how individuals rationally process new information over time. For example, if there is an upward current trend in $X_{1,i}$ with $\mu_{1,i} > 0$ for some t in subperiod i, one individual might rationally expect this trend to continue for a prolonged period, while another individual might expect a break in the trend in the near future. Due to nonrecurring structural breaks and the following inherently imperfect knowledge, both individuals' expectations can be based on fully rational considerations even though they are considering the exact same information. The key element is that under contigent and inherently imperfect knowledge there is no single rational way for individuals to form expectations, so there is no dualistic separation between rationality and irrationality with respect to individuals' expectation formation, which is in contrast to time-invariant models relying on REH, see Dow (2012).

However, instead of linking expectations explicitly to the parameters in θ_t and restricting all revisions of expectations to follow a specific revision-rule, IKE models impose enough structure on expectations for the model to have empirical implications. In the next sections, I first discuss the conditions of guardedly moderate revisions of Frydman and Goldberg (2007, 2013 *c*), which impose qualitative conditions on revisions of the parameters in β_t^{IKE} over time, so that expectations play a completely autonomous role in the model. But as these conditions alone does not imply an internally consistent representation of expectations, I propose a second set of conditions which link the forecasting parameters to the parameters in θ_i in a way that ensures internal consistency and yet maintains a partly autonomous role for expectations. Moreover, I show how these conditions are crucial for our ability to test the empirical implications of the IKE model, even though the conditions do not impose a specific revision-rule.

1.5.1 Guardedly Moderate Revisions

Frydman and Goldberg (2007, 2013 *a*,) impose the qualitative conditions of guardedly moderate revisions on changes in the parameters in β_t^{IKE}. Consider the total change in expectations from period $t-1$ to t given by

$$P_{t|t+1}^{IKE} - P_{t-1|t}^{IKE} = \Delta\alpha_t^{IKE} + \beta_{t-1}^{IKE\prime}\Delta X_t + \Delta\beta_t^{IKE\prime} X_t. \tag{1.25}$$

[8]In the following, I discuss the representation of the aggregate expectations, as in Frydman and Goldberg (2013 *a*), rather than a single individual. However, the conditions of 'guardedly moderate revisions' are imposed on each individual's revisions of expectations in e.g. Frydman and Goldberg (2007) and considering only the aggregate expectations is strictly speaking inconsistent as the inherently imperfect knowledge stems not only from nonrecurring changes in θ_i, but also from inherently imperfect knowledge of other individuals' expectation formation.

The first term is the effect of changes in α_t^{IKE}, while the second and third terms are the effects of changes in the exogenous variables and the effects of revisions on β_t^{IKE}, respectively.

For each of the k forecasting parameters in β_t^{IKE}, the condition of guardedly moderate revisions is given by[9]

$$|\Delta\beta_{m,t}^{IKE}| < |\beta_{m,t-1}^{IKE}\Delta X_{m,t}X_{m,t}^{-1}| \quad \text{for each } m = 1, 2, ..., k. \tag{1.28}$$

The qualitative condition ensures that individuals do not revise the individual weights attached to each of the k variables in X_t in ways so that the effect of revisions, $\Delta\beta_{m,t}^{IKE}X_{m,t}$, outweights the effect of changes in the causal variables, $\beta_{m,t-1}\Delta X_{m,t}$. An important feature of the guardedly moderate revisions conditions is that they allow individuals to revise the forecasting parameters β_t^{IKE} in ways that either impede or reinforce the changes in X_t over time. Moreover, the conditions are compatible with experimental findings from behavioral economics that actual market participants tend to revise their expectations only gradually and conservatively, see for example Edwards (1968) and Shleifer (2000).

However, the guardedly moderate revisions conditions are not imposed in every single period as the contingency of knowledge implies that expectations might be subject to sudden shifts in the expectation formation which does not satisfy these qualitative conditions.

It is important to note that the guardedly moderate revisions only restrict the changes in the forecasting parameters β_t^{IKE} from one period to the next. As they do not relate the parameters α_t^{IKE} and β_t^{IKE} to the parameters in θ_i at any point in time, the representation of expectations is completely autonomous relative to the rest of the model and hence the conditions alone do not imply internal consistency in the IKE model. In particular, the conditions allow the representation of expectations to diverge totally from the structure of the model. Hence, additional conditions on the forecasting parameters are needed in order to ensure that the representation of expectations is internally consistent with the process underpinning the asset price within the model.

1.5.2 Internal Consistency Conditions

In addition to the guardedly moderate revisions condition, I propose a new set of conditions on the representation of expectations which restrict the forecasting parameters in each subperiod i to the

[9] In Frydman and Goldberg (2007, 2013 *c*) the conditions in equation (1.26) are imposed on the $(k+1 \times 1)$ vector $\tilde{\beta}_t^{IKE} := (\alpha_t^{IKE}; \beta_t^{IKE})$, but here I impose the condition on each individual variable in β_t^{IKE}, as imposing the qualitative conditions on the vector allows for large changes in e.g. $\beta_{1,t}^{IKE}$ with no effect on the change in the price forecast when only the effect is offset by a large change in the opposite direction in e.g. $\beta_{2,t}^{IKE}$. Intuitively imposing the conditions on each $\beta_{m,t}^{IKE}$ seems more natural, although an additional qualitative condition on changes in α_t^{IKE} is then needed.

Moreover, Frydman and Goldberg (2007, 2013 *c*) allow for a larger set of variables to enter the representation of expectations, so that

$$\hat{P}_{t|t+1}^{IKE} = \alpha_t^{IKE} + \beta_t^{IKE\prime}X_t + \gamma_t^{IKE\prime}Z_t, \tag{1.26}$$

where Z_t is a k_z vector of observable variables. However, including Z_t in the representation of expectations is internally inconsistent with the model in equations (1.22) and (1.23), unless it is explicitly assumed that X_t depend on Z_t. This could for example be done by replacing the specification of X_t with the assumption that

$$X_t = \lambda_t' Z_t + \epsilon_{x,t}, \tag{1.27}$$

and thereby including both X_t and Z_t in the expectation formation becomes internally consistent.

class of stochastic processes which are stochastically trendless with unconditional mean given by values determined by the parameters in θ_i.[10]

Hence, in contrast to Frydman and Goldberg (2007, 2013 *c*), I assume that the forecasting parameters α_t^{IKE} and β_t^{IKE} in each subperiod i can be characterized as a stochastic process. However, I do not represent the parameters with a specific stochastic process, such as an $AR(1)$ process, as there is no theoretical basis for selecting a specific one. I only restrict the processes for α_t^{IKE} and β_t^{IKE} to a broad class of stochastic processes with general features which ensure that the representation of expectations is internally consistency with the structure of the model, and moreover are crucial for our ability to confront the IKE model with empirical evidence.

For each subperiod i, I impose the conditions that the process β_t^{IKE} belongs to the class of stochastically trendless processes with unconditional mean

$$E[\beta_t] = B_i, \tag{1.30}$$

and that the process α_t^{IKE} belongs to the class of stochastically trendless processes with unconditional mean

$$E[\alpha_t] = A_i, \tag{1.31}$$

where[11]

$$A_i := \frac{a_i}{1-c_i} + \frac{b_i'\mu_i}{(1-c_i)^2} \tag{1.32}$$

$$B_i := \frac{b_i}{1-c_i}, \tag{1.33}$$

for $t = T_{i-1}+1, ..., T_i$ and $i = 1, 2, ...$ with $T_i > T_{i-1}$. The class of stochastic processes satisfying these conditions is broad and not mutually exclusive with neither the guardedly moderate revisions conditions nor larger jumps in the processes within certain ranges.

Consider the forecasting error $fe_{t+1}^{IKE} := P_{t+1}^{IKE} - P_{t|t+1}^{IKE}$ for subperiod i, as given by

$$fe_{t+1}^{IKE} = \mathcal{A}_{t+1}^{IKE} + \mathcal{B}_{t+1}^{IKE} X_t + \upsilon_{t+1}^{IKE} \tag{1.34}$$

where

$$\mathcal{A}_{t+1}^{IKE} := (a_i + c_i\alpha_{t+1}^{IKE} - \alpha_t^{IKE}) + (b_i + c_i\beta_{t+1}^{IKE})'\mu_i \tag{1.35}$$

$$\mathcal{B}_{t+1}^{IKE} := (b_i + c_i\beta_{t+1}^{IKE} - \beta_t^{IKE})' \tag{1.36}$$

$$\upsilon_{t+1}^{IKE} := \epsilon_{p,t+1} + (b_i + c_i\beta_{t+1}^{IKE})'\epsilon_{x,t+1}. \tag{1.37}$$

---

[10]The process U_t is stochastically trendless if for $s \to \infty$ (for fixed t)

$$E_t[U_{t+s}] - E[U_{t+s}] \xrightarrow{P} 0, \tag{1.29}$$

so that the s-step ahead forecast converges towards the unconditional mean for $s \to \infty$, see Harris et al. (2002) and McCabe et al. (2003). The class of stochastically trendless processes includes weakly stationary processes, such as a simple $AR(1)$ process, while stochastically trending processes for example include a random walk process.

Harris et al. (2002) show that the process $Z_t := V_t W_t$, where V_t is a stochastically trendless process with mean zero and W_t is a stochastically trending process, is a stochastically trendless process. Hence, the stochastically trendless property of V_t dominates the multiplicative process Z_t asymptotically.

[11]Note the exact correspondence between the parameters A_i and B_i and the REH parameters α^{REH} and β^{REH} defined in equations (1.16).

From the conditions above it follows that in each subperiod i for $t = T_{i-1}+1, ..., T_i$ the forecasting error becomes stochastically trendless as all three terms in equation (1.34) are stochastically trendless, even though the second term is a multiplicative process including the stochastically trending variables X_t. To see that this is the case, consider first the second term. From equation (1.30) it follows that $E[\mathcal{B}_{t+1}^{IKE}] = 0$ for $t = T_{i-1}+1, ..., T_i - 1$, so $\mathcal{B}_{t+1}^{IKE}$ is a mean zero stochastically trendless process and the process $\mathcal{B}_{t+1}^{IKE\prime} X_t$ becomes a stochastically trendless process as it is a multiplicative process of a mean zero stochastically trendless process and a stochastically trending process. Hence, the stochastically trendless property of $\mathcal{B}_{t+1}^{IKE}$ dominates the multiplicative process asymptotically, so that it resembles a mean zero stationary process, although it might be very persistent and heteroskedastic. Consider next the term $\mathcal{A}_{t+1}^{IKE}$. From equation (1.31) it follows that $E[\mathcal{A}_{t+1}^{IKE}] = 0$ for $t = T_{i-1}+1, ..., T_i - 1$, so $\mathcal{A}_{t+1}^{IKE}$ is a zero mean stochastically trendless process. Finally, consider the term υ_{t+1}^{IKE}. From equation (1.30) it follows that $E[b_i + c_i \beta_{t+1}^{IKE}] = B_i$ for $t = T_{i-1}+1, ..., T_i - 1$, so υ_{t+1}^{IKE} is the sum of a Gaussian error term and a multiplicative process of a stochastically trendless process and a Gaussian error term, which is a stochastically trendless process. Hence, under the assumptions that α_t^{IKE} and β_t^{IKE} belong to the class of stochastically trendless properties with unconditional means A_i and B_i, respectively, the forecasting errors have a mean zero stochastically trendless representation in each subperiod i, even though they are systematic in the sense that they depend on X_t. Another way of interpreting the conditions is that the forecasting parameters α_t^{IKE} and β_t^{IKE} must fluctuate around A_i and B_i, respectively, to prevent a trend in the forecasting errors, which would be incompatible with rational individual decisionmaking. Thus, the conditions ensure internal consistency between the IKE model's representation of expectations and the process underpinning the asset price and yet it accords expectations a partly autonomous role.

Moreover, the internal consistency conditions imply that the asset price fluctuates around its fundamental value, which itself changes over time as a consequence of nonrecurring structural breaks in θ_i. Consider the asset price in subperiod i for $t = T_{i-1}+1, ..., T_i$, as given by

$$P_t^{IKE} = (a_i + c_i \alpha_t^{IKE}) + (b_i + c_i \beta_t^{IKE})' X_t + \epsilon_{p,t}. \tag{1.38}$$

By adding and subtracting A_i and B_i, as given in equation (1.32) and (1.33), re-arranging terms, and using that $a_i + c_i A_i = A_i$ and $b_i + c_i B_i = B_i$, the asset price in subperiod i can be written as

$$P_t^{IKE} = (A_i + c_i(\alpha_t^{IKE} - A_i)) + (B_i + c_i(\beta_t^{IKE} - B_i))' X_t + \epsilon_{p,t}, \tag{1.39}$$

Moreover, define the fundamental price in subperiod i for $t = T_{i-1}+1, ..., T_i$ as

$$P_t^{\star IKE} := A_i + B_i' X_t, \tag{1.40}$$

which is the fundamental value of the asset price in subperiod i as determined by the present discounted value of all future X_t if the current structure given by θ_i was to continue indefinitely. An important insight from this expression is that the fundamental value is by definition forward-looking; it is the discounted future values of X_t which determines the fundamental value—not the past values of X_t. Due to nonrecurring structural breaks in θ_i, the fundamental value changes over time not only due to changes in X_t, but also due to changes in θ_t.

The gap between the current asset price and its current fundamental value in subperiod i is given by

$$P_t^{IKE} - P_t^{\star IKE} = c_i(\alpha_t^{IKE} - A_i) + c_i(\beta_t^{IKE} - B_i)' X_t + \epsilon_{p,t}, \tag{1.41}$$

Hence, only for $\alpha_t^{IKE} = A_i$ and $\beta_t^{IKE} = B_i$—which would correspond to imposing REH in each subperiod i, though that would only be internally consistent if the contingent changes in θ_i were replaced with time-invariant θ for all t—would the deviations between the asset price and its current fundamental value be random and zero on average. But in the IKE model individuals must base their expectations on contingent and inherently imperfect knowledge so their expectations do not correspond to those values; first, individuals are not assumed to know the parameters θ_i, and second, individuals take into account that θ_i will be subject to structural break in the future. The latter part implies that even if individuals at a given point in time t were assumed to know the current values of θ_i, it might not even be optimal for them to base their expectations solely on these if they expect future structural breaks in θ_i.

If the forecasting parameters α_t^{IKE} and β_t^{IKE} are totally autonomous relative to the parameters A_i and B_i the asset price can deviate endlessly from its fundamental value $P_t^{\star IKE}$, so the gap between the two can be endlessly trending. For example, if $\alpha_t^{IKE} - A_i > 0$ for all t the asset price will systematically be above its fundamental value for all t (disregarding the second term in equation (1.41)) and if furthermore α_t^{IKE} is allowed to continue increasing over time, the asset price can diverge endlessly from its current fundamental value.

However, the internal consistency conditions imply that the gap between the asset price and its current fundamental value is stochastically trendless in each subperiod i. First, from equation (1.31) it follows that the first term is a stochastically trendless process with $E[\alpha_t^{IKE} - A_i] = 0$. Likewise, the second term is stochastically trendless as it follows from equation (1.30) that $(\beta_t^{IKE} - B_i)$ is a stochastically trendless process with $E[\beta_t^{IKE} - B_i] = 0$, so that the multiplicative process becomes a stochastically trendless process. Hence, under the internal consistency conditions the gap between the asset price and its current fundamental value becomes stochastically trendless in each subperiod i. Essentially this mean that the gap resembles a mean zero stationary—though potentially persistent and heteroskedastic—process in each subperiod. But due to nonrecurring structural breaks in θ_i the asset's fundamental value itself changes over time, so there are continuous fluctuations of the asset price around its changing fundamental values and the asset price does not converge towards its fundamental value over time. Moreover, expectations are still allowed to play a partly autonomous role; if for example individuals at time t expect a future increase in μ_i, so that the slope of the deterministic trends in X_t is expected to increase, they will increase α_t^{IKE} above A_i and the asset price will rise above its current fundamental value. Subsequently one of two things can happen: either their expectations turn out correct and $\mu_{i+1} > \mu_i$ for some $t+s$ with $s > 0$, or their expectations turn out incorrect so μ_i remains the same for $t+s$. In the former case, the increase in μ_{i+1} will imply an increase in A_i, so that $\alpha_t^{IKE} - A_i$ falls. In the latter case, individuals must eventually revise their unfullfilled expectations by lowering α_t^{IKE}. Hence, revisions in the forecasting parameters are inputs to the model driving the aggregate outcomes, but the expectations are only allowed a partly autonomous role relative to the rest of the model.

1.6 Testing Empirical Implications of the IKE Model

The challenge for testing empirical implications of IKE models is that they are contingent and partly open along two dimensions; they allow for nonrecurring structural breaks in the the parameters in the process underpinning aggregate outcomes and they do not specify all changes over time in the forecasting parameters of the model with specific stochastic processes. Thereby, IKE models do not specify a complete probabilistic structure which can be directly estimated using a 'theory-first approach' to econometrics, see Hoover (2006) and Spanos (2009).

However, partly open IKE models can be confronted with empirical evidence based on the methodological framework of the LSE-Oxford-Copenhagen approach which relies on a 'data-first approach' to econometrics, see Hoover (2006) or Spanos (2009) for an overview.

A key element in the 'data-first approach' is that the stochastic specification of the econometric model, though guided by economic theory, is chosen based on the data. The approach is based on general-to-specific modeling, where first a general unrestricted model is specified as a valid statistical representation of the data over the considered sample period, and thereafter restrictions are imposed and tested on the general model with the aim of reducing the general model to a more parsimoniously specific model that accounts for the information of the general model. The first step focuses on testing the statistical adequacy of the model as a representation of the data for the sample period considered and is concerned with what Spanos (2010) calls 'statistical testing'. Once a valid statistical representation of the data is found, the empirical validity of potentially conflicting hypotheses from economic theory can be imposed and tested as restrictions on the general model, which is concerned with what Spanos (2010) calls 'substantive testing.'

The 'data-first approach' is suitable for empirically testing IKE models as it searches for stochastic specifications as valid statistical representations of the data. First, IKE models do not specify a stochastic formulation for the time-varying parameters, but a specific stochastic process can be selected in an econometric model because the statistical validity of the econometric model can be tested against the data. The simplest case to consider is the case of piecewise constant parameters, while the internal consistency conditions proposed in this paper suggest that a stochastic process for the time-varying parameters must belong to the class of stochastically trendless processes. Within this class of processes, a simple $AR(1)$ representation might be a good first approximation of the time-varying parameters, although estimation of such time-varying parameter models is not straightforward. The key point is that while an IKE model does not assume a specific stochastic representation of the time-varying parameters *for theoretical reasons*, doing so for *econometric reasons* in an econometric model is possible because it is testable based on the specific sample of data. Second, while IKE models allow for structural breaks in both the parameters of the stochastic processes and the parameters linking the variables without specifying exactly when they occur, statistical tests can be used in an econometric model to test for parameter-constancy. Thereby, subperiods where a specific statistical model can adequately represent the data can be found and tested. If a statistical representation can be found which is both compatible with the IKE model and an adequate statistical representation of the data for a specific sample period it can be concluded that the IKE model cannot be falsified based on the data for the specific time period.

An important implication of imposing the internal consistency conditions proposed in this paper is that in each subperiod i the gap between the asset price and its fundamental value determined by the exogenous variables X_t becomes stochastically trendless, c.f. equation (1.41). As a linear combination of the variables becomes stochastically trendless the variables are cointegrated with stochastic cointegration parameters in each subperiod i, see also Chapter 3 in this thesis. In Chapter

2, I show that the classic cointegrated VAR model of Johansen (1996) can be used as a first-hand approximation to estimate long-run relations between time-series simulated from a simple IKE model. These simulations build in the internal consistency conditions proposed in this paper, and I show that restricting the parameters in the representation of expectations in relation to process underpinning the asset price is crucial for getting good approximate results from the econometric analysis. Hence, the internal consistency conditions are crucial for our ability to test the empirical implications of IKE models. Moreover, I show in Chapter 3 that if data generated from a data-generating process with stationary stochastic cointegration parameters are analyzed econometrically with the classic cointegrated VAR model, the underlying parameter-instability can show up in the estimated model as persistent deviations from estimated cointegration relations and low estimated adjustment coefficients. Such findings are frequent in analyzes of macroeconomic and financial data and the results indicate they can potentially stem from time-varying cointegration parameters in the underlying data-generating process.

1.7 Conclusion

The promising feature of IKE models is that by allowing for nonrecurring structural breaks in the process underpinning individual and aggregate outcomes they accord individuals' expectations an autonomous role in driving aggregate outcomes. However, until recently IKE models have not focused on internal consistency as an important element in specifications of IKE models, and thereby expectations have been accorded a completely autonomous role relative to the rest of the model.

In this paper, I have shown how internal consistency can be fully incorporated in IKE models by restricting the parameters in the representation of expectations to the class of stochastic processes which are stochastically trendless and with unconditional means determined by the parameters of stochastic specification of the exogenous variables and the parameters linking these to the endogenous variables. While the conditions imply that the representation of expectations becomes internally consistent they still accord individuals' expectations a partly autonomous role in driving aggregate outcomes. Moreover, the conditions are useful in terms of extending existing models to IKE models, and the implication that the asset price and the exogenous variables become cointegrated in each subperiod is crucial for testing empirical implications of IKE models.

By allowing for nonrecurring structural breaks in the process underpinning aggregate outcomes IKE models add a degree of openness compared to time-invariant stochastic models relying on REH. Due to nonrecurring structural change, IKE models allow the future to be partly open relative to the past instead of directly linked over all points in time through a time-invariant stochastic process. More importantly, the nonrecurring structural change implies that IKE models portray rational individual decisionmaking in a setting of contingent and inherently imperfect knowledge rather than complete probabilistic knowledge. As a consequence, individuals in IKE models make systematic forecasting errors, but by incorporating internal consistency the forecasting errors in an IKE model become stochastically trendless, whereas they are *i.i.d.* in REH models. Likewise the gap between the asset price and its current fundamental value becomes stochastically trendless in an IKE model, whereas the gap is *i.i.d.* in REH models. Hence, the subperiod cointegration relations between the asset price and the exogenous variables in X_t are stochastically trendless, and hence potentially persistent and heteroskedastic, whereas they are assumed *i.i.d.* in REH models.

These features are promising in terms of explaining empirical puzzles in macroeconomics and finance. Many of these puzzles share the feature that the degree of persistence in the deviations from estimated equilibria—typically in terms of estimated cointegration relations—are more persistent than standard economic theory predicts. For example, the seminal paper by Shiller (1981) showed that stock prices fluctuate much more than their fundamental value based on the present discounted value of future dividends can account for. The IKE framework appears promising as such persistent deviations can arise from partly autonomous expectations based on contingent and inherently imperfect knowledge.

1.8 References

Caballero, R. J. (2010), 'Macroeconomics after the crisis: Time to deal with the pretense-of-knowledge syndrome', *Journal of Economic Perspectives* **24**(4), 85–102.

Colander, D. (2006), Introduction, *in* D. Colander, ed., 'Post Walrasian Macroeconomics', Cambridge University Press, New York, pp. 1–23.

Dow, S. C. (2012), *Foundations for New Economic Thinking*, Palgrave Macmillan, Basingstoke.

Edwards, W. (1968), Conservatism in human information processing, *in* B. Kleinmuth, ed., 'Formal Representation of Human Judgement', John Wiley and Sons, New York.

Evans, G. W. and Honkapohja, S. (2001), *Learning and Expectations in Macroeconomics*, Princeton University Press, Princeton, NJ.

Evans, G. W. and Honkapohja, S. (2013), Learning as a Rational Foundation of Macroeconomics and Finance, *in* R. Frydman and E. S. Phelps, eds, 'Rethinking Expectations: The Way Forward for Macroeconomics', Princeton University Press, Princeton, pp. 130–165.

Frydman, R. and Goldberg, M. D. (2007), *Imperfect Knowledge Economics: Exchange Rates and Risk*, Princeton University Press, Princeton.

Frydman, R. and Goldberg, M. D. (2011), *Beyond Mechanical Markets: Asset Price Swings, Risk, and the Role of the State*, Princeton University Press, Princeton.

Frydman, R. and Goldberg, M. D. (2013 *a*), The contingent expectations hypothesis: Rationality and contingent knowledge in macroeconomics and finance theory. working paper prepared for the INET Annual Plenary Conference in Hong Kong, April 2013.

Frydman, R. and Goldberg, M. D. (2013 *b*), The imperfect knowledge imperative in modern macroeconomics and finance theory, *in* R. Frydman and E. S. Phelps, eds, 'Rethinking Expectations: The Way Forward for Macroeconomics', Princeton University Press, Princeton, chapter 6, pp. 130–167.

Frydman, R. and Goldberg, M. D. (2013 *c*), Opening models of asset prices and risk to nonroutine change, *in* R. Frydman and E. S. Phelps, eds, 'Rethinking Expectations: The Way Forward for Macroeconomics', Princeton University Press, Princeton, chapter 6, pp. 207–247.

Frydman, R. and Phelps, E. S. (2013), Which Way Forward for Macroeconomics and Policy Analysis?, *in* R. Frydman and E. S. Phelps, eds, 'Rethinking Expectations: The Way Forward for Macroeconomics', Princeton University Press, Princeton, pp. 1–46.

Guesnerie, R. (2005), *Assessing Rational Expectations: "Eductive" Stability in Economics*, MIT Press, Cambdridge, MA.

Guesnerie, R. (2013), Expectational Coordination Failures and Market Volatility, *in* R. Frydman and E. S. Phelps, eds, 'Rethinking Expectations: The Way Forward for Macroeconomics', Princeton University Press, Princeton, pp. 49–66.

Harris, D., McCabe, B. and Leybourne, S. (2002), 'Stochastic cointegration: estimation and inference', *Journal of Econometrics* **111**(2), 363–384.

Hoover, K. D. (1991), Scientific research program or tribe? a joint appraisal of lakatos and the new classical macroeconomics, *in* N. de Marchi and M. Blaug, eds, 'Appraising Economic Theories: Studies in the Methodology of Research Programs', Elgar, Aldershot, U.K., and Brookfield, VT, pp. 364–394.

Hoover, K. D. (2006), The Past as the Future: The Marshallian Approach to Post Walrasian Econometrics, *in* D. Colander, ed., 'Post Walrasian Macroeconomics: Beyond the Dynamic Stochastic General Equilibrium Model', Cambridge University Press, Cambridge, chapter 12, pp. 239–257.

Johansen, S. (1996), *Likelihood-Based Inference in Vector Autoregressive Models*, Oxford University Press, Oxford.

Knight, F. H. (1921), *Risk, Uncertainty and Profit*, Houghton Mifflin, Boston.

Lucas, R. E. (1976), 'Econometric policy evaluation: A critique', *Carnegie-Rochester Conference Series on Public Policy* **1**(1), 19–46. Reprinted in: R. E. Lucas (1981), *Studies in Business-Cycle Theory.* Cambridge, MA, MIT Press, p. 104–130.

Lucas, R. E. (1996), 'Nobel lecture: Monetary neutrality', *Journal of Political Economy* **104**(4), 661–82.

McCabe, B., Leybourne, S. and Harris, D. (2003), Testing for Stochastic Cointegration and Evidence for Present Value Models. Working Paper.

Muth, J. F. (1961), 'Rational Expectations and the Theory of Price Movements', *Econometrica* **29**, 315–335.

Phelps, E. S., Archibald, G. C. and Alchian, A. A., eds (1970), *Microeconomic Foundations of Employment and Inflation Theory*, W. W. Norton, New York.

Popper, K. R. (1990), *A World of Propensities*, Thoemmes Antiquarian Books, Bristol.

Shiller, R. E. (1981), 'Do stock prices fluctuate too much to be justified by subsequent dividends?', *American Economic Review* **71**(3).

Shleifer, A. (2000), *Inefficient Markets*, Oxford University Press, Oxford.

Soros, G. (1987), *The Alchemy of Finance*, Wiley, New York.

Spanos, A. (2009), 'The Pre-Eminence of Theory versus the European CVAR Perspective in Macroeconometric Modeling', *Economics: The Open-Access, Open-Assessment E-Journal* **3**(10).

Spanos, A. (2010), 'Statistical Adequacy and the Trustworthiness of Empirical Evidence: Statistical vs. Substantive Information', *Economic Modelling* **27**(27), 1436–1452.

Chapter 2

A Simulation Study of a Simple Imperfect Knowledge Economics Model of Stock Prices and Earnings with Cointegrated VAR Estimations

Morten Nyboe Tabor[†]

June 12, 2013

Abstract

In this paper, I simulate outcomes from a simple Imperfect Knowledge Economics (IKE) model of stock prices and earnings, based on Frydman and Goldberg (2013 *b*), with the aim is to address whether the cointegrated VAR model of Johansen (1996) can serve as a valid statistical representation of the simulated data and whether the regularities in the simulated data can be found econometrically with the cointegrated VAR model as a first approximation. The key features of the simple IKE model are: i) that there are streches of time where forecasting strategies are revised moderately, and ii) fluctuations in the stock price around a benchmark price level determined by earnings are bounded. In modeling these key features, earnings are assumed to fluctuate around a non-stationary long-run trend, with deviations caused by a bounded segmented trend process and a stationary component, and qualitative bounds are imposed on revisions of individuals' forecasting strategies, so that the causal parameter linking the stock price to earnings varies over time within specific bounds. Hence, deviation between the stock price and the benchmark price determined by earnings are bounded and the variables are cointegrated as a linear relation between them is stochastically trendless. The simulation results show that the cointegrated VAR model can serve as a statistically adequate representation of the simulated data, though the specification of the cointegrated VAR differs from the specification of the simulated data. Moreover, the results show that, despite bounded instability in the time-varying cointegration parameters in the data-generating process of the simulated data, the cointegrated VAR model can provide a fairly precise estimate of the sample mean of the boundedly time-varying cointegration parameters. The results indicate that the cointegrated VAR model can serve as a good starting point for econometric analyses of IKE models.

[†]Department of Economics, University of Copenhagen and INET Center for Imperfect Knowledge Economics at the University of Copenhagen. Email: morten.nyboe.tabor@econ.ku.dk. Funding from the Institute for New Economic Thinking (INET) is gratefully acknowledged. I would like to thank Roman Frydman, Michael Goldberg, Katarina Juselius, Søren Johansen, and Andreas Hetland for numerous discussions and comments to drafts of this paper. A draft of this paper was presented at the INET Annual Plenary Conference in Hong Kong, April 2013.

2.1 Introduction

A core premise of contemporary economic models is that researchers can adequately specify in probabilistic terms how individuals alter the way they make decisions and how the processes underpinning market outcomes unfold over time. Based on this core premise individual and aggregate outcomes at all points in time are represented with a time-invariant stochastic structure. Following Frydman and Goldberg (2013 *a*), I refer to such models as determinate. To confront determinate models with empirical evidence a 'theory-first approach' to econometrics is typically used, see Hoover (2006 *b*) and Spanos (2009). In the 'theory-first approach' the theoretical model delivers a complete stochastic specification that relates aggregate outcomes to a set of explanatory variables, and the role of econometrics is solely to quantify the theoretical parameters of interest and test their statistical significance using regression or other statistical techniques (Spanos, 2006).

By contrast, the Imperfect Knowledge Economics (IKE) approach (Frydman and Goldberg, 2007, 2011, 2013 *a*) recognizes that the process underpinning market outcomes is contingent: it changes at times and in ways that no one can fully anticipate. Hence, theoretical IKE models are by design contingent and partly open: they allow for nonrecurring changes in the causal structure. As a consequence individuals' expectations of future market outcomes, upon which to base their utility-maximizing behavior, is based on contingent and inherently imperfect knowledge. Confronting IKE models with empirical evidence is a challenge due to the contingency, as they imply different causal structures during different periods of time. Hence, they do not imply a time-invariant causal structure for aggregate outcomes that can be directly estimated and tested using standard econometric tools.

This paper presents an initial attempt to address this challenge. I show how the methodological framework of the LSE-Oxford-Copenhagen approach can serve as a basis for an econometric methodology for theoretical IKE model, with the cointegrated VAR model of Johansen (1996) as the starting point.[1] The LSE-Oxford-Copenhagen methodology and the cointegrated VAR model is based on a 'data-first approach' to econometrics, where the stochastic specification of the econometric model is derived from the data based on statistical testing, rather than imposed from the outset based on *a priori* assumptions of a determinate theoretical model. Although it is important to point out that *a priori* considerations based on economic theory is used as an invaluable guide in the variable selection, specification, and testing of the econometric model.

The essence of the 'data-first approach' to econometrics is the general-to-specific approach, which first seeks a general unrestricted model as a valid statistical representation of the data for the sample period considered, and then tests restrictions on the general model with the aim of finding a specific model that accounts for the information of the general model more parsimoniously. The first step involves what Spanos (2010) calls 'statistical testing,' which focuses on testing the statistical

[1] The LSE-Oxford-Copenhagen methodology originated from the work of Dennis Sargan at London School of Economics, but is today mainly associated with the work of David Hendry and co-workers at Oxford University and Søren Johansen and Katarina Juselius at the University of Copenhagen. For a broad introduction and discussion of the main econometric methodologies, see Hoover (2006 *a*), and for a discussion of the 'data-first' and 'theory-first' approaches, see Hoover (2006 *b*) and Spanos (2009) .

For a broad introduction to the theory and application of the cointegrated VAR model see Juselius and Johansen (2006), Johansen (1996), and Juselius (2006). Hendry (1995) provides a broad introduction to econometric modeling based on a general-to-specific approach, see also Mizon (1995) for a survey.

adequacy of the model as a representation of the data for the sample period considered. Once a valid statistical representation of the data is found, the empirical validity of potentially conflicting hypotheses from economic theory can be imposed and tested as restrictions and reductions of the general model, which involves what Spanos (2010) calls 'substantive testing.'

The 'data-first approach' is suitable for empirically testing IKE models as it allows for contingency in the underlying data-generating process by searching for stochastic specifications as statistically valid local representations of the data. Because a theoretical IKE model does not specify exactly when and how structural breaks occur the specification of an IKE econometric model has to be based upon and tested against the data. The 'data-first' methodology of the cointegrated VAR model allows for structural change to be identified *ex post* in the historical data without an *ex ante* probabilistic specification of exactly when and how the structural breaks occur. The key point here is that while IKE acknowledges that an economist cannot fully specify the occurence of structural breaks in an economic model *ex ante*, an econometrician using the 'data-first approach' can test for and identify structural breaks in the historical economic data *ex post*. However, it should be noted that IKE does not imply that there are no empirical relations that are stable at the aggregate level over time; IKE just don't start from an *a priori* assumption that all empirical relations are indeed stable at all points in time.

The cointegrated VAR model's system approach and its distilling of time series according to their degree of persistence has proven to be extremely useful for representing and modeling non-stationary macroeconomic and financial data. In practice, the specification of a statistically well-specified cointegrated VAR model requires selecting a suitable lag-length, including level shifts and dummy variables, and potentially splitting the sample into subsamples with different cointegrated VAR models for each subsample. With good econometric modeling skills, and a sense of the context under study, an econometrician can identify samples of historical data in which a specific cointegrated VAR model adequately represents the data. Inference in the cointegrated VAR model is then valid and testable hypotheses based on an IKE model can be tested as restrictions on the general model.

Although the 'data-first approach' of the cointegrated VAR provides a suitable way to empirically estimate and test IKE models, there are important challenges in bridging the empirics and theory. First, IKE models allow for contingent change in the stochastic representation of exogenous variables and the causal parameters linking the variables in IKE models are boundedly unstable over time, potentially with both frequent changes within qualitative ranges and less frequent large jumps. By contrast, the classic cointegrated VAR model has constant parameters. Important questions here are under what conditions can structural breaks in both the stochastic representations of the variables and the causal parameters linking them be identified using standard statistical procedures and residual misspecification tests; to what extent and under what conditions can we determine whether seperate subsample analyses are preferred over a full-sample analysis; or, alternatively, under which conditions can the time-varying parameters of IKE models be represented stochastically and estimated using extensions of the cointegrated VAR model with stochastic parameters. Moreover, many IKE models imply that markets are boundedly unstable: wide price swings away from benchmark values are eventually reversed and sustained movements back towards these values occur.[2] This implication suggests that there may be a connection between the boundedness of the

[2]For example, in the Frydman and Goldberg (2007, 2013 *b*) model of asset price swings and risk, persistent trends

market process and our ability to estimate cointegration relationships using the cointegrated VAR model. For example, is it the case that a greater tendency for reversals in the market leads to a greater chance that the system will be characterized by cointegration relationships?

To analyze these and other questions, this paper simulates outcomes from a simple IKE model of stock prices and earnings and analyze the simulated data econometrically with the cointegrated VAR model. The aim is to analyze if the cointegrated VAR model can serve as a approximation of the simulated data and if the regularities in the simulated data can be found econometrically, despite bounded parameter-instability in the data-generating process of the simulated data. There are two key features of this model that underpin our results: i) there are stretches of time in which market participants either maintain their forecasting strategies or revise them only moderately, and ii) price swings away from the benchmark value are bounded. In modeling these features, the simulated earnings fluctuate boundedly around a non-stationary long-run trend, with the bounded fluctuations being caused by a segmented trend specification, and the parameters linking earnings to the stock price varies over time within specific bounds. Hence, the simulated stock price and earnings are assumed to fluctuate persistently, but boundedly so, around a common long-run trend in earnings, and there exists a linear relation between the two which is stochastically trendless, so the variables are cointegrated.

The simulations show that even though the specification of the cointegrated VAR model is 'wrong' compared to the specification of the data-generating process used to simulate the data from the simple IKE model, it can nonetheless be used as a statistically adequate representation of the simulated data with an adequate lag structure. Furthermore, I show that the bounded instability of the relationship between the simulated asset prices and earnings plays a key role in our ability to understand and interpret the estimates of the cointegrated VAR model. Cointegration between the simulated time-series can be found during periods where the time-varying cointegration parameters of the simulated series are bounded, which implies that the variables are stochastically cointegrated because the linear relations $\beta' X_t$ are stochastically trendless. The results show that despite bounded instability in the time-varying cointegration parameters, with both frequent small changes and infrequent large jumps within a specific range, the cointegrated VAR model can provide an estimate of the unconditional sample mean. This extends the results in Chapter 3 of this thesis, where I show that the cointegrated VAR model can be used as an approximation which provides a consistent estimate of the unconditional mean of stationary autoregressive stochastic cointegration parameters.

The rest of the paper is structured in the following way. In section 2.2 a simple IKE model of stock prices and earnings is presented and the link between boundedness in the model and cointegration is discussed. Section 2.3 presents the simulation setup for the simple IKE model and shows an illustration of the simulated data. Section 2.4 introduces the cointegrated VAR model used for the econometric analysis of the simulated data, and the results from the estimations are presented in Section 2.5. Finally, Section 2.6 concludes.

---

in fundamental variables and the influence of psychological and social factors can lead market participants to bid asset prices persistently away from benchmark values over a stretch of time. But, this instability is bounded: if departures from benchmark values continued to grow, they would eventully lead market participants to revise their forecasting strategies in ways that resulted in a sustained countermovement back toward benchmark values.

2.2 A Simple IKE Model of Stock Prices and Earnings

In this paper I consider a simple version of an IKE model of stock prices and earnings. The model is a simplified version of the general IKE model of asset price swings and risk presented in Frydman and Goldberg (2013 *b*). The simple model considered here captures some—but not all—of the main ideas of an IKE asset pricing model in a simple way that mimics some of the key features of the stock market. The simple model allows me to simulate potential outcomes from an IKE model which can be econometrically analysed with a cointegrated VAR model in a fairly simple setup. For a full presentation and discussion of the general IKE model of asset price swings and risk, see Frydman and Goldberg (2013 *b*).

The general IKE model of long swings in asset prices can be written in reduced form at the aggregate level as

$$p_t = \widehat{p}_{t|t+1} - \widehat{up}_t + \varepsilon_{p,t}, \tag{2.1}$$

where p_t is the asset price at time t, $\widehat{p}_{t|t+1}$ is a representation of the aggregate forecast of the future asset price, $\widehat{up}_t$ is an uncertainty premium, and $\varepsilon_{p,t}$ is an *i.i.d.* Gaussian error term with variance σ_p^2.[3]

A key element of the IKE asset pricing model is the assumption that the uncertainty premium covaries positively over time with the gap between the asset price and a historical benchmark level. Defining this gap as

$$gap_t = p_t - p_t^{BM}, \tag{2.2}$$

the uncertainty premium can be represented as

$$\widehat{up}_t = \sigma \cdot gap_t = \sigma(p_t - p_t^{BM}), \tag{2.3}$$

where p_t^{BM} is the benchmark level for the asset price. The parameter σ determines the effect of the gap on the asset price, and here we assume for simplicity that the parameter is constant.

Given the specification of the uncertainty premium the asset price can be written as

$$p_t = \widehat{p}_{t|t+1} - \sigma(p_t - p_t^{BM}) + \varepsilon_t^p, \tag{2.4}$$

which is equivalent to

$$p_t = \lambda \widehat{p}_{t|t+1} + (1-\lambda) p_t^{BM}, \tag{2.5}$$

or

$$p_t = p_t^{BM} + \lambda(\widehat{p}_{t|t+1} - p_t^{BM}). \tag{2.6}$$

Equation (2.5) shows that in each period the asset price is represented as a weighted average of the price forecast and the benchmark price with weights given by $\lambda := 1/(1+\sigma)$ and $1-\lambda$, respectively. Equation (2.6) shows that the asset price can also be represented as the benchmark price plus a multiple of the deviation between the forecasted price and the benchmark price.

The overall idea of the IKE model of asset price swings and risk is that market participants base their forecasting strategies of the future price on a combination of fundamental, psychological, and social factors, and that persistent trends in the fundamental variables and the influence of

[3]Compared to the model in Frydman and Goldberg (2013 *b*), I have added the *i.i.d.* error term.

psychological and social factors can lead market participants to bid asset prices persistently away from the benchmark price over a strech of time. However, this instability is bounded: if departures from benchmark values continued to grow, they would eventually lead market participants to revise their forecasting strategies in ways that resulted in a sustained countermovement back towards benchmark values.

The bounded instability implies that the price forecast $\widehat{p}_{t|t+1}$ is allowed to move persistently away from the benchmark price level p_t^{BM}, but ultimately such movements are bounded. Hence, from equation (2.6) it follows that the asset price p_t moves persistently, but boundedly so, around the benchmark price p_t^{BM}.

2.2.1 A Representation of the Benchmark Price and Price Forecast

I now depart from the general IKE model of asset price swings and risk and consider a simple model of stock prices and earnings. In this simple model both the benchmark price and the representation of the aggregate price forecast depend only on corporate earnings, and thereby the stock price is assumed to depend only on corporate earnings. I assume that earnings has a long-run non-stationary trend and a short-run component fluctuating persistently around the long-run trend. The benchmark price depends on the long-run trend in earnings, while the price forecast for simplicity is represented only in terms of currently observed earnings.

First, assume that there is a non-stationary long-run trend in earnings, which can be represented as a random walk with a drift given by

$$\overline{x}_t = \overline{x}_{t-1} + \mu_{\overline{x}} + \varepsilon_{\overline{x},t} = \overline{x}_0 + \sum_{i=1}^{t}(\mu_{\overline{x}} + \varepsilon_{\overline{x},i}), \tag{2.7}$$

where $\mu_{\overline{x}} > 0$ is a constant positive drift term and $\varepsilon_{\overline{x},t}$ is an *i.i.d.* Gaussian error with variance $\sigma^2_{\overline{x}}$.

Assume next that current earnings x_t fluctuate persistently around the long-run trend $\overline{x}_t$, and that the fluctuations can be represented by a segmented trend specification. The segmented trend push current earnings persistently away from the long-run trend given by $\overline{x}_t$, but eventually a reversal in the segmented trend occurs, thereby causing a countermovement of current earnings back towards the long-run trend. As current earnings reach the long-run trend level they are allowed to continue away from the long-run trend in the opposite direction, but eventually another reversal will cause another countermovement back towards the long-run trend. Hence, the idea is that the short-run fluctuations in earnings are bounded around the long-run trend, so that current earnings has a non-stationary long-run trend and a bounded short-run trend represented with a segmented trend specification.

To capture this idea, assume that current earnings can be represented as

$$x_t = \Psi_t + \chi_t \tag{2.8}$$

where Ψ_t is a segmented trend and χ_t is a stationary process. First, the stationary process χ_t is represented as a first-order autoregressive process given by

$$\chi_t = \rho\chi_{t-1} + \varepsilon_{x,t}, \tag{2.9}$$

where $\varepsilon_{x,t}$ is a standard *i.i.d.* Guassian error term with variance σ_x^2 and with $0 < \rho < 1$. Hence, χ_0 can be given an initial distribution so that the process χ_t is stationary and has the representation

$$\chi_t = \sum_{i=0}^{\infty} \rho^i \varepsilon_{x,t-i}. \tag{2.10}$$

Moreover, it is assumed that the variance of the shocks to χ_t is greater than the variance of the long-run trend in earnings $\overline{x}_t$, i.e. $\sigma_x^2 > \sigma_{\overline{x}}^2$.

The segmented trend Ψ_t has a number, n, of long swings for $t = 1, 2, ..., T$ and we let $0 = T_0^* < T_1^* < T_2^* < ... < T_n^* = T$ denote the points in time at which the segmented trend changes direction, so the length of the i'th swing is given by $T_i = T_i^* - T_{i-1}^*$. The segmented trend process is given by

$$\Psi = \Psi_{t-1} + \mu_t = \Psi_0 + \sum_{i=1}^{t} \mu_t, \tag{2.11}$$

where

$$\mu_t = \mu_i \qquad \text{for } t = T_{i-1}^*, ..., T_i^* - 1 \tag{2.12}$$

and where μ_i is restricted to take on values with opposite signs in subsequent segments, so that

$$sign(\mu_i) \neq sign(\mu_{i-1}). \tag{2.13}$$

The IKE model does not specify when the switches in μ_i occur and what values it can take on with a probability distribution. Though, it is here assume that the probability of a switch in the direction of the segmented trend increases with the deviation between the segmented trend process and the long-run trend in earnings, $\Psi_t - \overline{x}_t$.[4] Hence, as the gap $\Psi_t - \overline{x}_t$ increases the probability of a reversal in the segmented trend increases, and eventually a shift in the direction of the segmented trend occurs, so that the deviation $\Psi_t - \overline{x}_t$ is assumed bounded.

Given these specifications, current earnings are given by

$$x_t = \Psi_0 + \sum_{i=1}^{t} \mu_t + \sum_{i=0}^{\infty} \rho^i \varepsilon_{x,t-i}, \tag{2.14}$$

so x_t can be represented as a combination of segmented trend and a stationary process. Alternatively, by adding and subtracting $\overline{x}_t$, the current earnings can be represented as

$$x_t = \overline{x}_t + (\Psi_t - \overline{x}_t) + \chi_t, \tag{2.15}$$

which shows that current earnings can be represented as a combination of a non-stationary long-run component determined by $\overline{x}_t$, the deviation between the segmented trend and the long-run trend $(\Psi_t - \overline{x}_t)$, which by design is a bounded process, and finally a stationary component. The important implication of the assumed specification is that it follows that the deviation between current earnings and their long-run trend, $x_t - \overline{x}_t$, is bounded.

Next, the uncertainty premium depends on the gap between current stock price and the benchmark price, which it is assumed can be represented as a multiple of the long-run trend in earnings, as given by

$$p_t^{BM} = B' \overline{x}_t, \tag{2.16}$$

[4] In the simulations presented below the probabilities of a switch in the segmented trend are necessarily specified probabilistically. A simple logistic function is used.

where B is a parameter assumed to be constant for the sample considered, which however need not be the case in a general IKE model.

Finally, the aggregate forecast of the future price is represented as

$$\widehat{p}_{t|t+1} = b_t' x_t, \tag{2.17}$$

where, for simplicity, it is assumed that the aggregate forecasts of the future stock price can be represented only in terms of current earnings x_t, so that b_t is a scalar representing the weight attached to earnings in the forecasting strategy at time t.

Movements in the price forecast depend on two factors, movements in earnings and revisions of the forecasting strategies, as given by

$$\Delta \widehat{p}_{t|t+1} = b_{t-1}' \Delta x_t + \Delta b_t' x_t. \tag{2.18}$$

In modeling revisions of the forecasting strategies, a slightly modified version of the 'guardedly moderate revisions' conditions of Frydman and Goldberg (2007, 2013 *b*) is imposed. The guardedly moderate revisions conditions restrict the changes in the forecasting weights b_t from one period to the next, so that the impact of revisions on the total change in the price forecast is smaller than the impact of the segmented trend in earnings[5], as given by

$$\left|\Delta b_t' x_t\right| < \left|b_{t-1}' \mu_t\right|, \tag{2.19}$$

where $|\cdot|$ denotes an absolute value and $\left|b_{t-1}' \mu_t\right|$ represents the 'baseline trend' in the price forecast which would occur on average over the period from T_i^* to T_{i-1}^* if the forecasting strategies were not revised, i.e. if $\Delta b_t = 0$, over the period. The condition embodies the idea that if individuals revise their forecasting strategies, they are reluctant to do so in ways that would outweight the effect of the baseline drift caused by the movements in x_t.

As the long-run and segmented trends in earnings unfold over time, the way they feed into the price forecast and the stock price depend on the revisions of the forecasting strategies. During streches of time where the guardedly moderate revisions hold, the revisions can either reinforce or impede the trends in earnings. Thus, one can think of the revisions of the forecasting weight to current earnings as representing how market participants project current trends in earnings into the future; if, for example, market participants forecast that an upward current trend in earnings is unsustainable, so that they expect a reversal some time in the near future, they might revise their forecasting strategies in impeding ways, so that the impact of their revisions counteract the current trend in earnings. Likewise, market participants forecasting that a downward current trend will soon be reversed might revise their forecasting strategies in reinforcing ways.

Based on this interpretation of the simple IKE model of stock prices and earnings, it can be interpreted as equivalent to the present-value model of Barsky and De Long (1993), with earnings taking the role of dividends in their model. In the Barsky and DeLong present value model a small part of the shocks to dividends in each period feeds into the future growth rate of dividends, which changes the present value of the future dividends that determines the stock price. However, while the

[5] I impose only the first of the two conditions specified in Frydman and Goldberg (2007, 2013 *b*), as the second becomes redundant in the univariate case considering here.

shocks to dividends feed into both current dividends and the growth rate the time-invariant model in Barsky and De Long (1993) assumes that the impact of such shocks on both the trend in dividends and the stock price is constant over time. By contrast, this simple IKE model acknowledges that the way the expected trends in earnings feed into the stock price depends on how market participants revise their forecasting strategies, and thereby it varies over time within specific bounds. Thereby, the simple IKE model allows for periods of both over- and underreactions to news about earnings—in form of exogenous shocks and changes in the segmented trend—within the qualitative ranges. Such over- and underreaction to news have been found important empirically in behavioral studies, see for example the influential paper by Barberis et al. (1998), and in the simple IKE model it is a natural consequence of rational individual behavior under contingent and imperfect knowledge.

Moreover, the contingency of an IKE model allows for non-moderate revisions of the forecasting strategies at points in time that cannot be specified with a probability distribution. Thus, at points in time that cannot be anticipated the revisions of the forecasting strategies are allowed not to fall within the the qualitative range specified by the condition in equation (2.19). However, I do impose the additional condition that whenever a non-moderate revision of the forecasting strategies occurs, the new forecasting weight b_t^{NM} falls within a qualitatively range which is symmetrically bounded around the parameter B, as given by the inequalities

$$\underline{b} < b_t^{NM} < \overline{b}, \tag{2.20}$$

where $\underline{b} = B - \tau_b$ and $\overline{b} = B + \tau_b$ represent the upper and lower bounds. This additional condition of non-moderate revisions within the bounds given in equation (2.20) implies that the forecasting weight b_t is symmetrically bounded around the parameter B. Though the guardedly moderate revisions alone do not imply that the forecasting weight b_t is bounded, this additional condition implies that even if b_t is pushed outside the range from $\underline{b}$ to $\overline{b}$, eventually a non-moderate revision will occur and thereby force b_t back within this range.

2.2.2 Bounded Instability and Cointegration Between Stock Prices and Earnings

Based on equation (2.6) and the above representations of earnings, the benchmark price, and the price forecast, the stock price can be written as

$$\begin{aligned} p_t &= p_t^{BM} + \lambda \left(\widehat{p}_{t|t+1} - p_t^{BM} \right) + \varepsilon_{p,t} \\ &= B' \overline{x}_t + \lambda \left(b_t' x_t - B' \overline{x}_t \right) + \varepsilon_{p,t}. \end{aligned} \tag{2.21}$$

For the stock price to fluctuate boundedly around the benchmark level consistent with the long-run trend in earnings, the deviation between the price forecast and the benchmark price must be bounded. The simple representation considered here allows me to decompose this deviation into two components as follows

$$\begin{aligned} \widehat{p}_{t|t+1} - p_t^{BM} &= b_t' x_t - B' x_t \\ &= B' \left(x_t - \overline{x}_t \right) + \left(b_t - B \right)' \overline{x}_t, \end{aligned} \tag{2.22}$$

and boundedness can be considered for each of the two components individually.

First, the representation of earnings implies that the deviation between current earnings and the long-run trend in earnings, $x_t - \overline{x}_t$, is bounded, c.f. equation (2.15). The segmented trend cause current earnings to fluctuate persistently around the long-run trend $\overline{x}_t$, so even though the long-run trend $\overline{x}_t$ is non-stationary—and hence not bounded—the deviation between the two is bounded, as $\Psi_t - \overline{x}_t$ is bounded by design and χ_t is a standard stationary process.

The second term in equation (2.22) is a product of $(b_t - B)$ and the non-stationary long-run trend in earnings $\overline{x}_t$. However, despite that $\overline{x}_t$ itself is not bounded, boundedness of $(b_t - B)' \, \overline{x}_t$ over time requires only that $b_t - B$ is bounded with mean zero over time. In that case the product of a mean zero bounded process and a non-bounded process will become bounded, as also confirmed by the simulation presented below.[6] The guardedly moderate revisions condition in equation (2.19) does not imply boundedness of b_t, so to achieve boundedness of $b_t - B$ we assume that the probability of a non-moderate revision increases with the deviation $b_t - B$, and by restricting the non-moderate revision to the symmetric range around B, c.f. equation (2.20).

Based on the above specifications, each of the two terms $x_t - \overline{x}_t$ and $b_t - B$ are bounded

$$x_t - \overline{x}_t \sim \text{bounded}, \tag{2.23}$$

$$b_t - B \sim \text{bounded}, \tag{2.24}$$

which implies that

$$(b_t - B)' \, \overline{x}_t \sim \text{bounded}, \tag{2.25}$$

so that the deviation between the price forecast and the benchmark price, as given in equation (2.22), is bounded

$$\widehat{p}_{t|t+1} - p_t^{BM} \sim \text{bounded}. \tag{2.26}$$

It follows from equation (2.21) that fluctuations of the stock price are bounded around the benchmark price over time. Hence, the stock price can move persistently away from the benchmark price consistent with the long-run trend in earnings, as the segmented trend pushes current earnings away from their long-run trend, or as the forecasting strategies are revised in reinforcing ways so that b_t move away from B. Moreover, the two effects might impact the stock price in the same direction during some streches of time, while they might outweigh each other during other streches of time. However, movements in the stock price away from the benchmark price consistent with the long-run trend in earnings are ultimately bounded as a reversal in the segmented trend push current earnings back towards the long-run trend, or as market participants revise their forecasting strategies in non-moderate ways causing a reversal of the price forecast—and hence the stock price—back towards the benchmark price.

The important implication of the two boundedness conditions in equations (2.23) and (2.24) is that, despite the underlying bounded instability, the stock price and current earnings fluctuate around a common trend given by the non-stationary long-run trend in earnings.

Because the stock price and earnings share a common trend we can think of them both as being cointegrated with the long-run trend in earnings—as well as with each other—though the specification of boundedness does not fully correspond to a standard cointegration relation. First, the

[6] In a standard stochastic specification the multiplicative process of a stationary mean zero process and a non-stationary process becomes 'stochastically trendless', which means that the stochastically trendless property of the stationary process dominates the multiplicative process, see Harris et al. (2002) and McCabe et al. (2003).

boundedness between current earnings and their long-run trend can be thought of as a cointegration relation with a time-varying adjustment coefficient: when the segmented trend push earnings away from the long-run trend it corresponds to an equilibrium-increasing adjustment coefficient; and when the segmented trend push earnings towards the long-run trend it corresponds to an adjustment coefficient which is equilibrium-adjusting. Second, as the stock price fluctuates boundedly around the long-run trend in earnings we can think of the two as being cointegrated, with persistent deviations from the cointegration relation driven by the two bounded terms in equation (2.22).

As the stock price and current earnings fluctuate boundedly around the same common stochastic trend given by the non-stationary long-run trend in earnings, the deviation between them is bounded and we can think of the stock price and current earnings as being cointegrated. Though, the boundedness of the IKE model is based on a time-varying specification, which differs from a standard stochastic specification of cointegrated relations.

To see that the deviation between the stock price and current earnings is indeed bounded, rewrite equation (2.21) to

$$p_t = b_t' x_t - (1-\lambda)\left(b_t' x_t - B'\overline{x}_t\right) + \varepsilon_{p,t}, \tag{2.27}$$

and re-arrange terms to get

$$p_t - b_t' x_t = -(1-\lambda)\left(b_t' x_t - B'\overline{x}_t\right) + \varepsilon_{p,t}. \tag{2.28}$$

We know from above that the term $b_t' x_t - B'\overline{x}_t$ is bounded given the assumptions, so the right-hand side of equation (2.28) is bounded. Hence, the deviation between the stock price and the price forecast on the left-hand side is bounded. Furthermore, the assumption that $(b_t - B)$ is bounded with mean zero implies that the stock price p_t and the multiple of current earnings $B' x_t$ are bounded, and we can think of the stock price and current earnings as being stochastically cointegrated.[7] By rewriting equation (2.28) as

$$p_t - B' x_t = (b_t - B)' x_t - (1-\lambda)\left(b_t' x_t - B'\overline{x}_t\right) + \varepsilon_{p,t}. \tag{2.29}$$

it can be seen that all terms on the right-hand side are bounded, once again because the term $(b_t - B)' x_t$ can be thought of as being stochastically trendless. Thus, $p_t - B' x_t$ becomes bounded and can loosely be given an interpretation as a cointegration relation.

2.3 Simulations of the IKE Model of Stock Prices and Earnings

I use simulations to examine the link from the bounded instability of an IKE model to estimation of cointegration relations in a cointegrated VAR model. Thus, I simulate outcomes from the simple IKE model of stock prices and earnings and analyze the simulated data econometrically with the cointegrated VAR model.

[7] In a standard stochastic specification, stochastic cointegration between two variables requires that the time-varying cointegration parameter β_t can be represented as $\beta_t = \beta + \tilde{\beta}_t$, where $\tilde{\beta}_t$ is mean zero stationary. Under this assumption the cointegration relation with time-varying cointegration parameter β_t can be written as

$$\beta_t' X_t = \beta' X_t + \tilde{\beta}_t' X_t,$$

where the second term becomes stochastically trendless, see Harris et al. (2002) and McCabe et al. (2003) as well as Chapter 3 in this thesis.

The cointegrated VAR model is 'wrong' compared to the specification of the IKE model used to simulate the data: the specification of boundedness in the IKE model does not correspond to the specification of cointegration relations in a cointegrated VAR model, and the parameters of the IKE model are boundedly time-varying, while the cointegrated VAR model has constant parameters. The results in Chapter 3, however, suggest that the cointegrated VAR model is a quite robust model to use even though there is an underlying bounded parameter instability in the data. Though, it is unclear if and under what conditions the cointegrated VAR model can be used to estimate empirical relations between variables based on the specification of a the simple IKE model.

I address this question by using simulations. I show that as long as the deviations of $\Psi_t - \overline{x}_t$ and $b_t - B$ are not 'too large' the cointegrated VAR model can be used as a statistically valid representation of the data along the key dimensions, and I do find the simulated stock price and current earnings to be cointegrated—and hence sharing a common stochastic trend—with the estimated cointegration coefficients close to the coefficient B as we would expect based on equation (2.29).

2.3.1 The Simulation Design

An IKE model acknowledges contingent changes that cannot be specified in advance with a probability distribution. By contrast, computer simulations requires a deterministic or probabilistic specification of both when and how the contingent structural breaks occur. Though, while an IKE model itself cannot be simulated because it is contingent by design, we can simulate outcomes that are consistent with an IKE model, and using simulations we can easily check the robustness of a specific specification.

In the simulations presented here a standard logistic function is used to simulate the probabilities of a switch in the direction of the segmented trend and the probabilities of non-moderate revision of the forecasting strategies at each point in time. The logistic function has the form

$$P\left(\text{break}_{i,t}\right) = \left[1 + \exp\left(-g_i\left(z_{i,t-1} - c_i\right)\right)\right]^{-1}, \tag{2.30}$$

where c_i is a threshold value where the probability of a break is one, g_i determines the curvature, and $z_{i,t-1} - c_i$ determines the distance that the probabilities depend on for $i = \Psi, b$, which represent the probabilities of a break in the segmented trend and the forecasting weights, respectively.

I let the probability of a break in the segmented trend depend on the absolute deviation $\Psi_{t-1} - \overline{x}_{t-1}$, so I set $z_{\Psi,t} = |\Psi_{t-1} - \overline{x}_{t-1}|$, and let the probability of a non-moderate revision of the forecasting strategies depend on the absolute deviation $b_{t-1} - B$, so I set $z_{b,t} = |b_{t-1} - B|$. Thus, as $|\Psi_{t-1} - \overline{x}_{t-1}|$ increases, the probability of a switch in the direction of the segmented trend increases, and eventually as $|\Psi_{t-1} - \overline{x}_{t-1}| \geq c_\Psi$ the probability of a switch reaches one. Moreover, after a switch in the direction of the segmented trend cause a countermovement in Ψ_t towards $\overline{x}_t$, I set the probability of a switch in the direction to zero until Ψ_t has crossed $\overline{x}_t$. Likewise, as $|b_{t-1} - B|$ increases, the probability of a non-moderate revision increases, and eventually the probability of a non-moderate revision reaches one as $|b_{t-1} - B| \geq c_b$.

At each point in time, we make two random draws from a standard uniform distribution at each point in time, and if they exceed the simulated probabilities we draw a new μ_t or b_t, respectively. The new μ_t is uniformly drawn within a specified range from $\underline{\mu}$ to $\overline{\mu}$, and with opposite sign compared to μ_{t-1}, while the new b_t^{NM} is drawn uniformly within the range from $B - \tau_b$ to $B + \tau_b$.

The curvature parameters $g_\Psi = g_b = 1.0$ and the threshold value for non-moderate revisions $c_b = 4.0$ are all fixed, and simulate the IKE model for different values of the threshold parameter c_Ψ and the range for non-moderate revisions τ_b. These two parameters are crucial determinants of the degree of bounded instability in the simulated system, and hence they are the parameters of greatest interest. The greater the threshold parameter c_Ψ, the greater deviation between current earnings and the long-run trend is allowed before a reversal eventually occurs. The parameter τ_b determines the symmetrical range around B within which a non-moderate revision is randomly drawn. The guardedly moderate revisions are symmetrical and by themselves do not ensure boundedness of b_t around B, so this boundedness occurs solely through the non-moderate revisions in this specification: if b_{t-1} is far above B, but below $B + \tau_b$, the probability of drawing a new b_t^{NM} below b_{t-1} is large.

The rest of the simulation setup follows the IKE model of stock prices and earnings as described above. The following parameter values are fixed for all simulations presented here

$$\begin{aligned}
B &= 2.0, && (2.31)\\
b_0 &= 2.0, && (2.32)\\
\lambda &= 0.5, && (2.33)\\
[\sigma_p, \sigma_x, \sigma_{\overline{x}}] &= [0.5, 0.5, 0.1], && (2.34)\\
[p_0, x_0, \overline{x}_0] &= 5.0, && (2.35)\\
\rho &= 0.5, && (2.36)\\
\mu_{\overline{x}} &= 0.01, && (2.37)\\
[\underline{\mu}, \overline{\mu}] &= [0.02, 0.15] && (2.38)\\
[g_\Psi, g_b] &= 1.0, && (2.39)\\
c_b &= 4.0. && (2.40)
\end{aligned}$$

I simulate time-series for p_t, x_t, and $\overline{x}_t$ for $i = 1, 2, ..., N$ different data-generating processes based on the $N = 16$ combinations of the parameters

$$\tau_b^i \in \{0.25; 0.50; 1.00; 1.50\}, \tag{2.41}$$

and

$$c_\Psi^i \in \{2.0; 3.0; 4.0; 5.0\},$$

where the upper limits for the two parameters are selected as the upper limits where cointegration appears to be found among the variables. For each data-generating process i, $S = 10,000$ replications of time-series are simulated with the different sample lenghts t as given by

$$t \in \{200; 400; 1000\}, \tag{2.42}$$

so that in total $480,000$ time-series are simulated. All simulations and estimations have been done in Ox 6.20, see Doornik (2007), with a random seed set to $1,000$ and reset for each new i, so that the sequences of random shocks are the same across the different data-generating processes i.

A cointegrated VAR model is estimated for each of the simulated time-series for the stock price p_t and earnings x_t, and averages of the results over the $S = 10,000$ replications are reported for each data-generating process i for each of the different sample lenghts considered.

2.3.2 An Illustration of the Simulated Series

The simulated outcomes for the specification $i = 7$, where $\tau_b = 0.5$ and $c_\Psi = 4.0$, are shown in Figures 2.1 to 2.5.

Figure 2.1 shows the simulated series x_t and $\overline{x}_t$ in the upper panel, the gap between the two in the middle panel, and finally the simulated probabilities of a reversal in the segmented trend in the lower panel. From the upper panel the segmented trend specification of x_t around $\overline{x}_t$ is evident. Moreover, it should be noted that after each reversal in the segmented trend, a new value for μ_t is randomly drawn within a range, so the segmented trends have different slopes for the different segments. From the middle panel it can be seen how the gap between current earnings and their long-run trend is bounded over time, and in the lower panel it can be seen that the probability of a reversal increases as the segmented trend drives current earnings away from the long-run trend. The blue squares in the lower panel indicate the 23 reversals in the segmented trend over the sample.

Figure 2.2 shows the simulated weights attached to current earnings in the forecasting strategies. The upper panel shows the simulated weights b_t, along with B and indicators for non-moderate revisions. It is clear from the graphs that the simulated parameter b_t is boundedly unstable over time, and that the non-moderate revisions imply a number of large jumps in the forecasting weights. The middle panel shows the qualitative ranges for the guardedly moderate revisions, along with the simulated revisions of the forecasting strategies within these ranges. The lower panel shows the probabilities of a non-moderate revision of the forecasting strategies over time. On average this probability is around 2.5 percent, and over the long sample of $1,000$ observations in the illustration, 22 revisions are simulated as non-moderate as indicated by the blue squares.

Figure 2.3 shows the simulated stock price, the benchmark price, and the price forecast in the upper panel. As the stock price is represented as a weighted average of the benchmark price and the price forecasts it lies between the two over the entire sample period. The middle panel shows the deviation between the price forecast and the benchmark price, which is equivalent to a scaled version of the gap between the stock price and the benchmark price. The lower panel shows the decomposition of the deviation between the price forecast and the benchmark price as specified in equation (2.29). From the graphs it can be seen that each of the two components are bounded over the sample.

In Figure 2.4 the simulated stock price and current earnings are shown in the upper panel along with the simulated long-run trend in earnings. The middle panel shows the deviation $p_t - B'x_t$, while the lower panel show the deviation $p_t - B'\overline{x}_t$.

Finally, Figure 2.5 shows the first-differences of the stock price and current earnings in the two upper panels, and it can be seen that there are a few very large outliers, which are caused by the jumps in b_t due to non-moderate revisions. The lower graph displays the estimated cointegration relation, which looks almost identical with the graph in the lower panel in Figure 2.4. Despite the bounded instability in the relation between the stock price and earnings, a cointegrated VAR model for the series illustrated in Figures 2.1-2.4 finds the two variables to be cointegrated with the estimated cointegration relation given by $\hat{\beta}'X_t = p_t - 2.38x_t$.

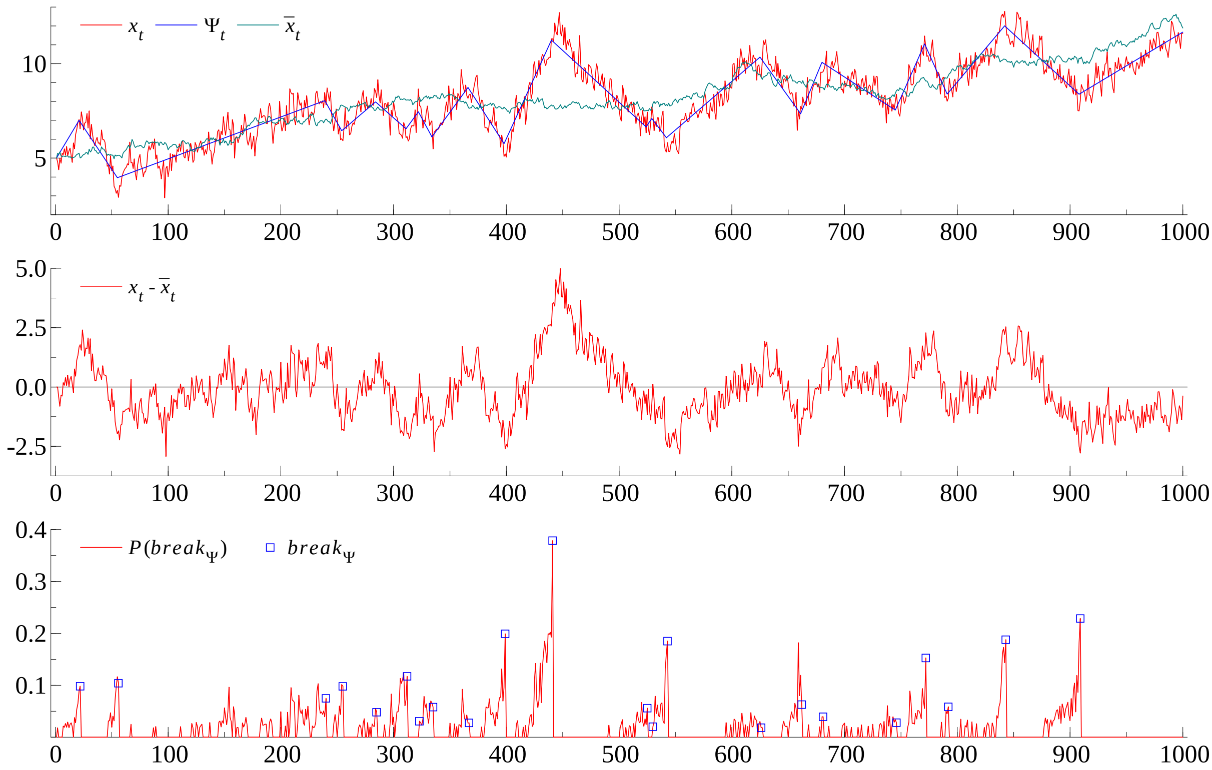

Figure 2.1: Current earnings, the segmented trend, and the long-run trend in earnings. The upper panel shows the long-run trend in earnings $\bar{x}_t$ (green line), with the segmented trend Ψ_t (blue line) and current earnings x_t (red line). The middle panel shows the deviation between current earnings and their long-run trend, $\bar{x}_t - x_t$, which determines the simulated probability of a break in the segmented trend. The simulated probabilities are shown in the lower panel, where the blue squares indicate a reversal in the direction of the segmented trend.

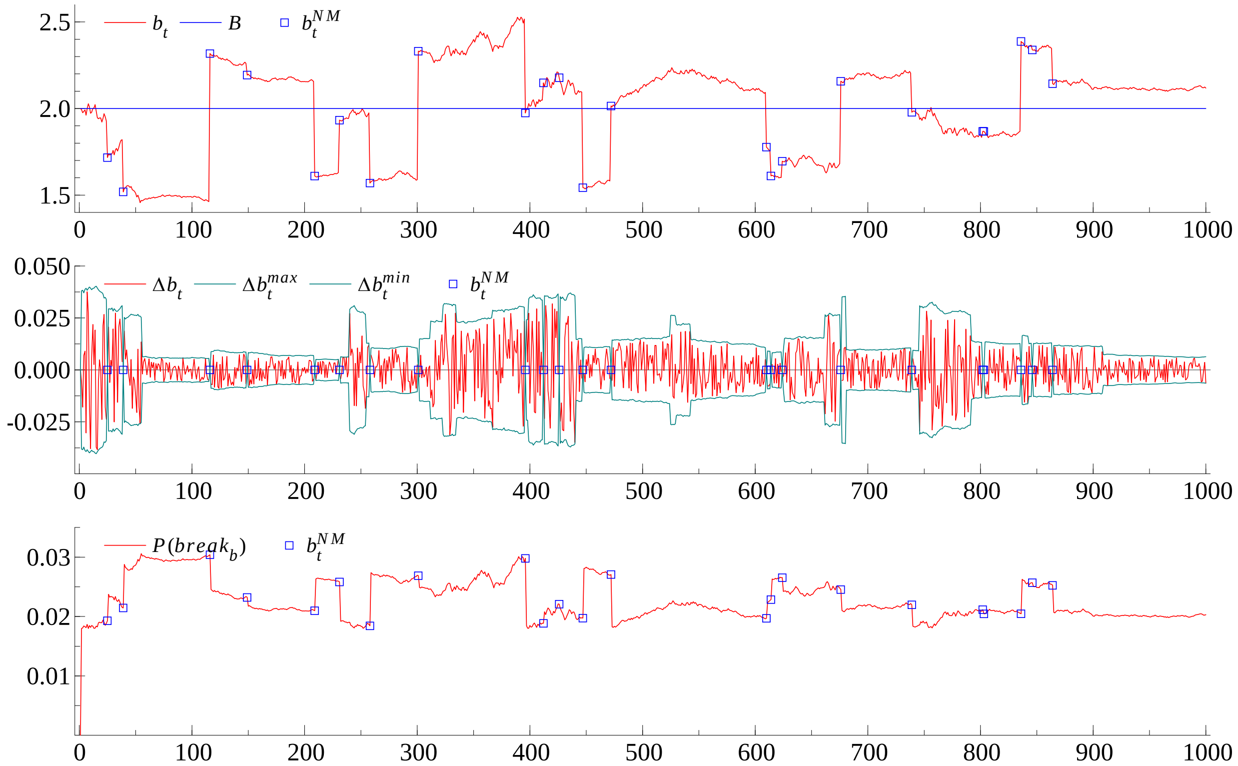

Figure 2.2: Revision of forecasting strategies. The upper panel shows the weights attached to current earnings in the forecasting strategies over time (red line) along with non-moderate revisions the forecasting strategies (blue squares) and the parameter B (green line). The middle panel shows the qualitative ranges imposed on the revisions of the forecasting strategies by the guardedly moderate revisions condition (green lines), the simulated revisions of the forecasting strategies (red line), and the non-moderate revisions the forecasting strategies (blue squares). The lower panel shows the simulated probability of a non-moderate revision (red line) and the non-moderate revisions the forecasting strategies (blue squares).

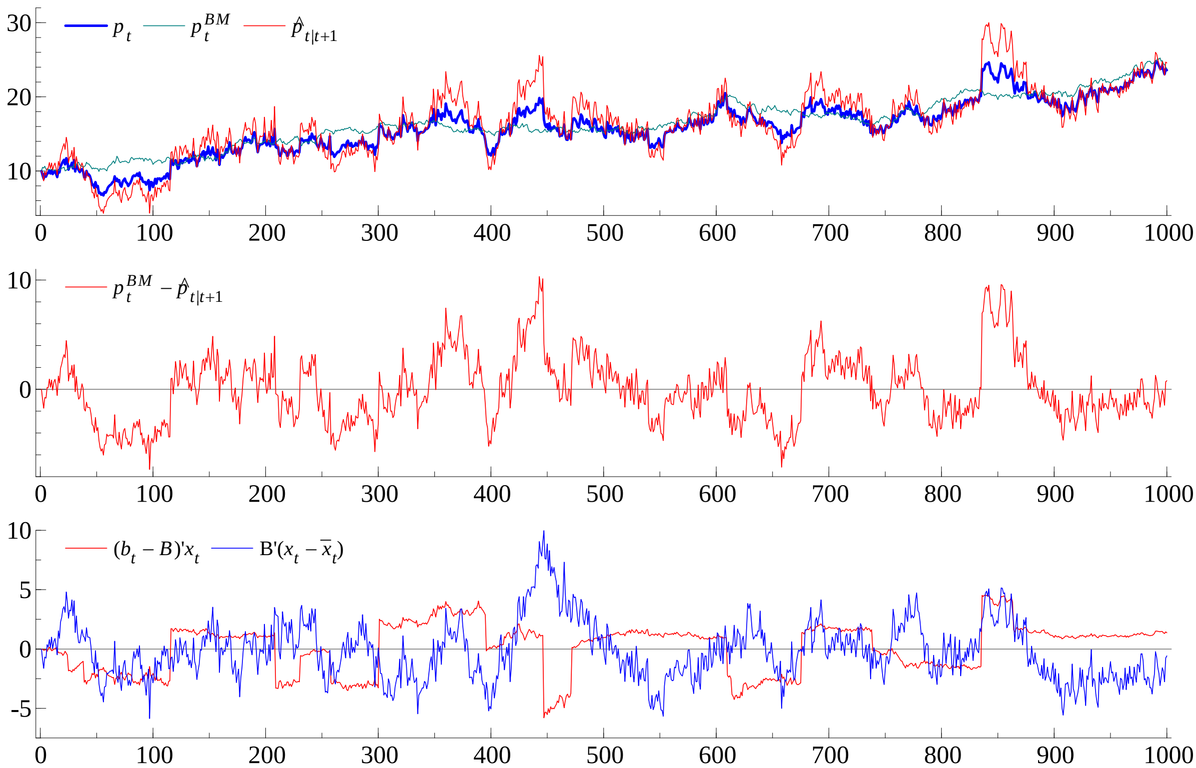

Figure 2.3: The stock price, benchmark price, price forecast. The upper panel shows the simulated stock price (blue line), the benchmark price (green line), and the price forecast (red line). The middle panel shows the deviation between the price forecast and the benchmark price. The lower panel shows the decomposition of the deviation between the price forecast and the benchmark price into two terms, which are individually bounded: $(b_t - B)'z_t$ and B$'(x_t - \bar{x}_t)$.

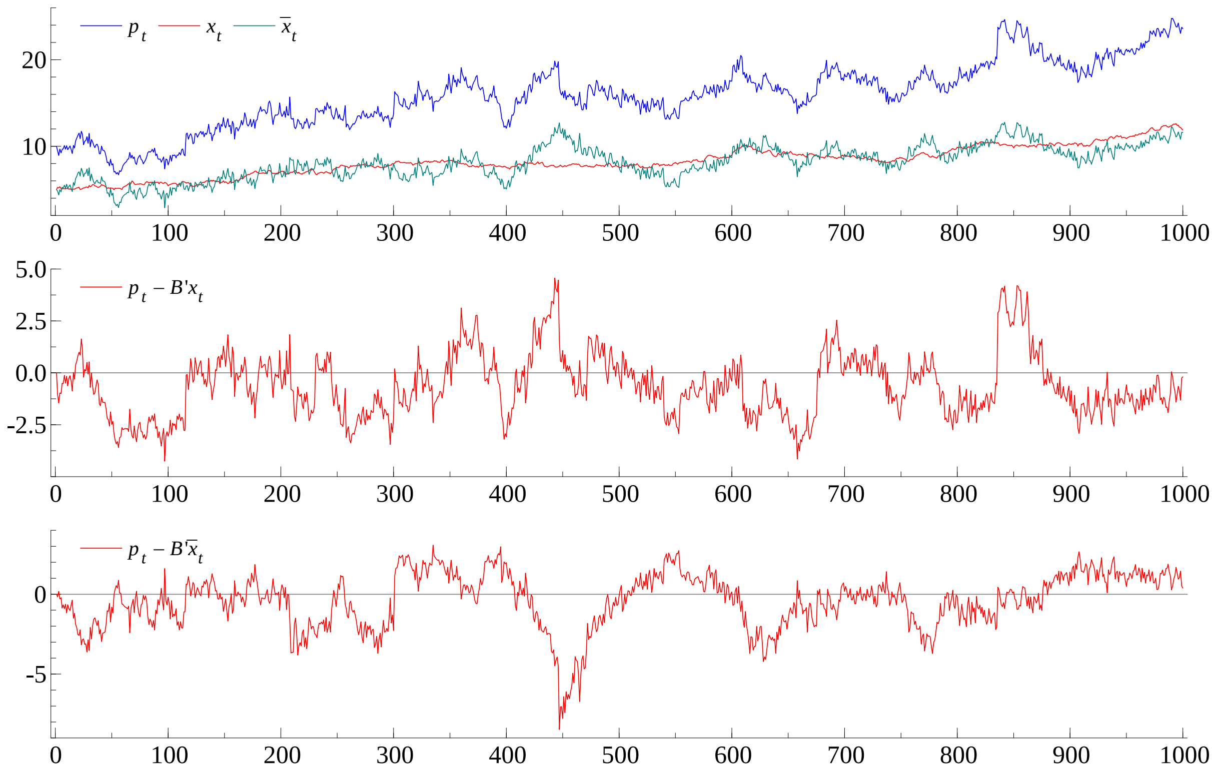

Figure 2.4: The stock price, benchmark price, price forecast. The upper panel shows the simulated stock price (blue line), the benchmark price (green line), and the price forecast (red line). The middle panel shows the deviation between the price forecast and the benchmark price. The lower panel shows the decomposition of the deviation between the price forecast and the benchmark price into two terms, which are individually bounded: $(b_t - B)'z_t$ and B$'(x_t - \bar{x}_t)$.

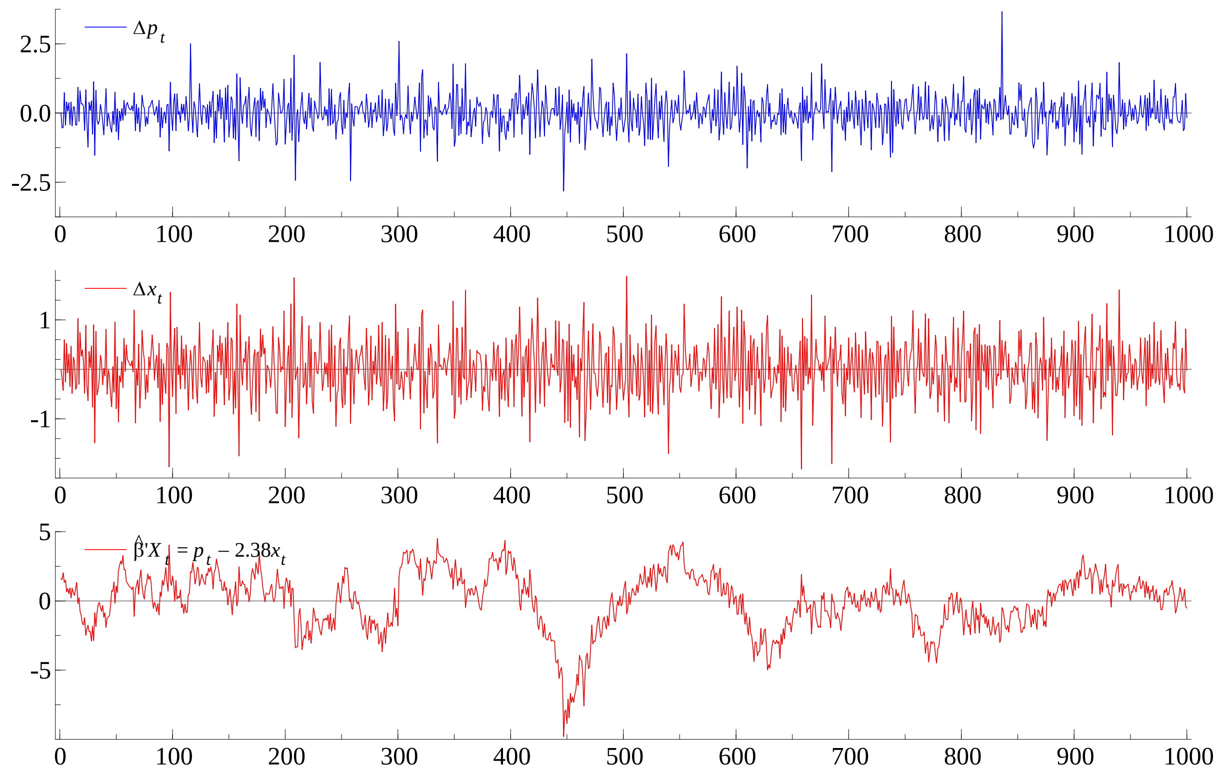

Figure 2.5: Estimated cointegration relation and first-differences. The upper and middle panels show the first-differences of the stock price and earnings data, which are analyzed econometrically with a cointegrated VAR model. The lower panel shows the estimated cointegration relation $\hat{\beta}' X_t = p_t - 2.38 x_t$.

2.4 The Cointegrated VAR Model

The p-dimensional vector autoregressive (VAR) model with k lags in error-correction form is given by

$$\Delta X_t = \Pi X_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} + \mu_0 + \mu_1 t + \Phi D_t + \varepsilon_t, \tag{2.43}$$

for $t = 1, 2, ..., T$ and where $X_{-k-1}, ..., X_0$ are fixed. The error terms ε_t are assumed to be independent and Gaussian with mean zero and covariance Σ. The parameters Π and Γ_i for $i = 1, ..., k-1$ are of dimension $(p \times p)$, the parameters μ_0 and μ_1 of dimension $(p \times 1)$. Dummy variables and mean shifts can be included in D_t, which has dimension $\left(p^D \times 1\right)$, and the parameters Φ has dimensions $\left(p \times p^D\right)$.

Assumption 1 *Assume that the autoregressive polynomial $A(z) = I - Az$ has exactly $p - r$ unit roots at $z = 1$ and the remaining roots are larger than one in absolute value, $|z| > 0$.*

Johansen (1996, Theorem 4.2) shows that under Assumption 1 the matrix Π has reduced rank r and can be represented as

$$\Pi = \alpha\beta', \tag{2.44}$$

where the $(p \times r)$ matrices α and β have full column rank, so that the cointegrated VAR model is given by

$$\Delta X_t = \alpha\beta' X_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} + \mu_0 + \mu_1 t + \Phi D_t + \varepsilon_t. \tag{2.45}$$

The levels X_t are nonstationary while the r cointegration relations $\beta' X_t$ are stationary. Hence, while the levels X_t are integrated of order one, $X_t \sim I(1)$, the linear combinations $\beta' X_t$ are integrated of order zero, $\beta' X_t \sim I(0)$, so that the process ΔX_t is stationary, $\Delta X_t \sim I(0)$. The cointegration relations determine the deviations from the long-run relations between the variables, while the α-coefficients measures the rate of adjustments to deviations from the long-run cointegration relations. For a full introduction to the theory and application of the cointegrated VAR model, see Johansen (1996) and Juselius (2006).

The notion of cointegration can be interpreted in the following way: there is a stationary long-run equilibrium relation between the non-stationary variables, and whenever the variables move away from this long-run relation at least one of the system variables is adjusting, so that the deviations from the long-run relations are stationary. Hence, even though the non-stationarity makes the individual system variables path-dependent, the cointegration relations ensure that the deviations between them are bounded.

Despite that the cointegrated VAR model has constant parameters, it has shown able to estimate the unconditional mean of time-varying cointegration parameters (at least in small systems) given by $\beta_t = \beta + \tilde{\beta}_t$, where $\tilde{\beta}_t$ is a mean zero stationary autoregressive process, see Chapter 3. In this case the linear relations $\beta' X_{t-1}$ and $\beta_t' X_{t-1}$ are stochastically trendless, see Harris et al. (2002) and McCabe et al. (2003), and the linear relations $\beta' X_{t-1}$ can be interpreted as the long-run average cointegration relations. Disregarding deterministic terms and lagged first-differences, a stochastically cointegrated system can be written as

$$\Delta X_t = \alpha\beta_t' X_{t-1} + \varepsilon_t, \tag{2.46}$$

or equivalently as

$$\Delta X_t = \alpha\beta' X_{t-1} + \alpha\tilde{\beta}_t' X_{t-1} + \varepsilon_t. \tag{2.47}$$

As $\tilde{\beta}_t$ is a mean zero stationary process, the product $\tilde{\beta}_t' X_{t-1}$ becomes stochastically trendless, and the system becomes stochastically cointegrated, see McCabe et al. (2003), Harris et al. (2002). Based on simulations, I show in Chapter 3 that the cointegrated VAR model gives a consistent estimate of the unconditional mean of the time-varying cointegration parameter in a bivariate system, i.e. a consistent estimate of β, when $\tilde{\beta}_t$ is specified as a stationary first-order autoregressive process with mean zero. However, if there is a high degree of (stationary) persistence in the time-varying parameter the underlying parameter instability shows up as an additional degree of persistence in the estimated cointegrated VAR model, which causes the estimated adjustment coefficients to be skewed towards zero.

2.5 Estimation Results

For each of the simulated time-series of the stock price p_t and current earnings x_t, a cointegrated VAR model is estimated based on an automated procedure. The automatic procedure first selects a lag-length k for the unrestricted VAR model in equation (2.43), where the lag-length is selected as the lowest number where the multivariate test of no first-order autocorrelation cannot be rejected at a 5-percent significance level. Given the lag-length, the automated procedure tests for univariate and multivariate autocorrelation, normality, and ARCH in the estimated residuals. Next, the rank test for reduced rank is performed and the largest roots of the companion matrix are calculated. Finally, the automatic procedure estimates the reduced rank cointegrated VAR model with a rank of $r = 1$ imposed, irrespective of the conclusion of the rank test. For details on the estimation and testing procedures, see Johansen (1996) and Juselius (2006), and references therein.

Tables 2.1-2.9 show the average results over the $S = 10,000$ replications for each of the simulated series and estimations, and for the three different sample lengths considered. It should be noted that the samples of $T = 1,000$ observations are included with the purpose to show the asymptotic results based on a long sample.

2.5.1 Breakpoints in the Simulated Data

Table 2.1 shows the average simulated probabilities of a break in the segmented trend or a non-moderate revision of the forecasting strategies for each of the specification, as well as the number of breaks occuring in each of the two per 100 observations. It can be seen that as the range for drawing non-moderate revisions increases, the average simulated probability of a non-moderate revision increases from 2 to 3 percent. Hence, the number of simulated non-moderate revisions increases from 2.0 to 3.4 on average per 100 observations. As the threshold parameter c^i_Ψ used in the logistic function to simulate the probabilities of a reversal in the trend increases, the average simulated probability of a reversal decreases almost exponentially from 8 to 2 percent as c^i_Ψ falls from 2.0 to 4.0. As the threshold value increases, fewer simulated reversals in the segmented trend occur. For $c^i_\Psi = 2.0$ an average of 8 reversals occur per 100 observations, meaning that the swings in earnings around the long-run trend are not very long and persistent. However, for $c^i_\Psi = 4.0$

Table 2.1: Breakpoints in the Simulated Data

i	τ_b^i	c_Ψ^i	T	$P(break_b)$[a]	$breaks_b$[b]	$P(break_\Psi)$[a]	$breaks_\Psi$[b]
1	0.25	2.00	200	0.02	2.0	0.07	7.0
			400	0.02	2.0	0.07	6.8
			1000	0.02	2.0	0.07	6.7
2	0.25	3.00	200	0.02	2.0	0.04	4.1
			400	0.02	2.0	0.04	3.9
			1000	0.02	2.0	0.04	3.8
3	0.25	4.00	200	0.02	2.0	0.03	2.5
			400	0.02	2.0	0.02	2.4
			1000	0.02	2.0	0.02	2.3
4	0.25	5.00	200	0.02	2.0	0.02	1.7
			400	0.02	2.0	0.02	1.6
			1000	0.02	2.0	0.02	1.5
5	0.50	2.00	200	0.02	2.2	0.07	7.0
			400	0.02	2.2	0.07	6.8
			1000	0.02	2.3	0.07	6.7
6	0.50	3.00	200	0.02	2.2	0.04	4.1
			400	0.02	2.2	0.04	3.9
			1000	0.02	2.3	0.04	3.8
7	0.50	4.00	200	0.02	2.2	0.03	2.5
			400	0.02	2.2	0.02	2.4
			1000	0.02	2.3	0.02	2.3
8	0.50	5.00	200	0.02	2.2	0.02	1.7
			400	0.02	2.2	0.02	1.6
			1000	0.02	2.3	0.02	1.5
9	1.00	2.00	200	0.03	2.6	0.07	7.0
			400	0.03	2.7	0.07	6.8
			1000	0.03	2.8	0.07	6.7
10	1.00	3.00	200	0.03	2.6	0.04	4.1
			400	0.03	2.7	0.04	3.9
			1000	0.03	2.8	0.04	3.8
11	1.00	4.00	200	0.03	2.6	0.03	2.5
			400	0.03	2.7	0.02	2.4
			1000	0.03	2.8	0.02	2.3
12	1.00	5.00	200	0.03	2.6	0.02	1.7
			400	0.03	2.7	0.02	1.6
			1000	0.03	2.8	0.02	1.5
13	1.50	2.00	200	0.03	3.1	0.07	7.0
			400	0.03	3.3	0.07	6.8
			1000	0.03	3.4	0.07	6.7
14	1.50	3.00	200	0.03	3.1	0.04	4.1
			400	0.03	3.3	0.04	3.9
			1000	0.03	3.4	0.04	3.8
15	1.50	4.00	200	0.03	3.1	0.03	2.5
			400	0.03	3.3	0.02	2.4
			1000	0.03	3.4	0.02	2.3
16	1.50	5.00	200	0.03	3.1	0.02	1.7
			400	0.03	3.3	0.02	1.6
			1000	0.03	3.4	0.02	1.5

All reported values are averages over $S = 10,000$ replications.

[a] Average simulated probability of a non-moderate revision of the forecasting strategies and a reversal in the segmented trend, respectively.

[b] Average number of breakpoints per 100 observations.

the number of simulated reversals decreses to 1.6 per 100 observations, which implies that the movements in earnings away from the long-run trend, caused by the segmented trend, are very long and persistent.

2.5.2 Specification and Residual Misspecification Tests

Table 2.2 shows the chosen lag-lengths in the estimated unrestricted VAR model based on the selection procedure described above. From the table it can be seen that as the threshold value for the breaks in the segmented trend, c^i_Ψ, increases, the number of lags needed in the unrestricted model to be able to not reject no autocorrelation increases. The same holds for an increase in the range for non-moderate revisions, although the effect appears to be smaller. Furthermore, it can be seen that the number of lags needed in the model increases with the sample size.

Table 2.3 shows the misspecification tests for no autocorrelation in the estimated residuals. It is of interest that it is possible to get non-autocorrelated residuals by choosing an appropriate lag-length in all cases. This is important as standard asymptotic inference in the cointegrated VAR model is extremely sensitive to autocorrelation in the residuals, and it shows that the flexibility of the lag structure can capture the persistent deviations from the estimated long-run structure caused by the underlying bounded instability.

The misspecification tests for normality of the estimated residuals are presented in Table 2.4. The results show that only with low values for both τ^i_b and c^i_Ψ, combined with a sample of $T = 200$, can normality of the residuals not be rejected based on the multivariate test, and even in these cases the results are really borderline. In all other cases normality is rejected based on the multivariate test. However, in most cases univariate normality of the residuals in the equation for the stock price cannot be rejected for samples of $T = 200$. By contrast, univariate normality of the residuals in the equation for current earnings cannot be rejected for all specifications and sample lengths.

By looking at Table 2.5, which shows the univariate skewness and excess kurtosis of the standardized estimated residuals,[8] it can be concluded that the rejection of univariate normality in the stock price equation and the rejection of multivariate normality is caused by a very large degree of excess kurtosis. Thus, the densities of the residuals have 'fat tails', which appears to be primarily associated with a few large outliers due to large non-moderate revisions in b_t. These outliers can easily be spotted based on a graphical inspection of the data—see for example the first-differences Δp_t in Figure 2.5 in the illustration above—and a careful empirical analysis and modeling of a single time-series would—and should—capture the outliers by including a few dummy variables in the model. Though, it is worth mentioning that the cointegrated VAR is quite robust to excess kurtosis, see Juselius (2006). By contrast, skewness is more problematic for inference in the cointegrated VAR model, but the results in Table 2.5 show that skewness is not a problem.

Table 2.6 shows the final misspecification test, and the results show that no ARCH cannot be rejected for all specifications with high p-values. This is not surprising as the variance of the random shocks was assumed constant in the simulations. However, ARCH-effects in the residuals might also arise from time-varying parameters, but there do not appear to be any noticable volatility clustering in the residuals.

[8] The skewness of the standardized normally distributed residuals is 0.0, and the kurtosis is 3.0.

Table 2.2: Selected Lag-Lengths in the Unrestricted Model

i	τ_b^i	c_Ψ^i	T	$Av(k)$[a]	$k=1$[b]	$k=2$[b]	$k=3$[b]	$k\geq 4$[b]
1	0.25	2.00	200	1.22	0.80	0.18	0.02	0.00
			400	1.78	0.44	0.38	0.15	0.04
			1000	3.80	0.01	0.11	0.29	0.59
2	0.25	3.00	200	1.36	0.69	0.26	0.04	0.00
			400	2.11	0.27	0.43	0.24	0.06
			1000	3.92	0.00	0.06	0.29	0.65
3	0.25	4.00	200	1.53	0.57	0.35	0.08	0.01
			400	2.46	0.13	0.41	0.34	0.11
			1000	4.06	0.00	0.03	0.24	0.72
4	0.25	5.00	200	1.68	0.46	0.40	0.12	0.01
			400	2.67	0.08	0.37	0.39	0.16
			1000	4.19	0.00	0.02	0.21	0.77
5	0.50	2.00	200	1.27	0.77	0.20	0.03	0.00
			400	2.05	0.32	0.40	0.22	0.07
			1000	4.40	0.00	0.03	0.18	0.79
6	0.50	3.00	200	1.39	0.66	0.28	0.05	0.00
			400	2.27	0.21	0.42	0.28	0.09
			1000	4.23	0.00	0.03	0.20	0.77
7	0.50	4.00	200	1.55	0.55	0.35	0.08	0.01
			400	2.51	0.12	0.41	0.35	0.13
			1000	4.21	0.00	0.02	0.20	0.78
8	0.50	5.00	200	1.69	0.46	0.41	0.12	0.02
			400	2.67	0.07	0.37	0.40	0.16
			1000	4.26	0.00	0.01	0.19	0.80
9	1.00	2.00	200	1.34	0.71	0.24	0.04	0.01
			400	2.34	0.21	0.39	0.28	0.12
			1000	4.77	0.00	0.01	0.11	0.88
10	1.00	3.00	200	1.44	0.63	0.30	0.06	0.01
			400	2.44	0.15	0.40	0.32	0.12
			1000	4.49	0.00	0.01	0.14	0.85
11	1.00	4.00	200	1.57	0.54	0.36	0.09	0.01
			400	2.59	0.10	0.39	0.37	0.15
			1000	4.39	0.00	0.01	0.16	0.83
12	1.00	5.00	200	1.71	0.45	0.41	0.12	0.02
			400	2.72	0.07	0.36	0.39	0.18
			1000	4.38	0.00	0.01	0.16	0.83
13	1.50	2.00	200	1.39	0.68	0.26	0.05	0.01
			400	2.49	0.17	0.37	0.30	0.16
			1000	4.88	0.00	0.01	0.09	0.90
14	1.50	3.00	200	1.49	0.61	0.31	0.07	0.01
			400	2.54	0.13	0.38	0.33	0.15
			1000	4.59	0.00	0.01	0.12	0.87
15	1.50	4.00	200	1.61	0.52	0.37	0.10	0.02
			400	2.67	0.09	0.37	0.37	0.17
			1000	4.47	0.00	0.01	0.14	0.85
16	1.50	5.00	200	1.73	0.44	0.41	0.12	0.02
			400	2.78	0.06	0.35	0.39	0.20
			1000	4.43	0.00	0.01	0.15	0.84

All reported values are averages over $S = 10,000$ replications.
[a] Average lag-length k.
[b] Percentage with lag length $k = 1, 2, 3$ and $k \geq 4$, respectively.

Table 2.3: Misspecification Tests Part 1: Autocorrelation

				Vector test no autocorr. order 1[a]		Vector test no autocorr. order 1-2[b]		Univar. test no autocorr. order 1 in $\hat{\varepsilon}_{1t}$[b]		Univ. test no autocorr. order 1 in $\hat{\varepsilon}_{2t}$[b]	
i	τ_b^i	c_Ψ^i	T	$\chi^2(4)$	$p-val$	$\chi^2(8)$	$p-val$	$\chi^2(4)$	$p-val$	$\chi^2(4)$	$p-val$
1	0.25	2.00	200	4.55	0.41	9.95	0.79	4.26	0.64	4.28	0.66
			400	5.41	0.32	12.43	0.65	6.48	0.42	6.82	0.43
			1000	5.76	0.29	13.97	0.39	8.00	0.15	8.44	0.15
2	0.25	3.00	200	4.99	0.36	10.94	0.76	5.15	0.57	5.49	0.58
			400	5.56	0.31	12.82	0.58	6.74	0.31	7.34	0.31
			1000	5.59	0.31	12.97	0.44	7.18	0.20	7.64	0.19
3	0.25	4.00	200	5.31	0.33	11.77	0.70	5.85	0.48	6.44	0.47
			400	5.61	0.30	13.02	0.52	6.77	0.25	7.53	0.23
			1000	5.53	0.31	12.50	0.46	6.80	0.23	7.32	0.22
4	0.25	5.00	200	5.44	0.32	12.24	0.65	6.19	0.41	6.92	0.38
			400	5.64	0.30	12.97	0.49	6.68	0.24	7.53	0.21
			1000	5.51	0.31	12.23	0.47	6.55	0.25	7.08	0.23
5	0.50	2.00	200	4.65	0.40	10.26	0.78	4.21	0.62	4.73	0.63
			400	5.50	0.31	12.97	0.58	6.23	0.35	7.54	0.32
			1000	5.80	0.29	14.32	0.36	7.40	0.18	8.98	0.13
6	0.50	3.00	200	4.99	0.36	11.02	0.75	4.85	0.56	5.67	0.55
			400	5.55	0.31	12.88	0.55	6.16	0.30	7.59	0.26
			1000	5.57	0.31	13.04	0.43	6.48	0.23	7.82	0.18
7	0.50	4.00	200	5.26	0.33	11.75	0.70	5.37	0.49	6.48	0.46
			400	5.57	0.31	12.92	0.51	6.10	0.27	7.64	0.22
			1000	5.49	0.32	12.49	0.46	6.05	0.26	7.39	0.21
8	0.50	5.00	200	5.40	0.32	12.15	0.65	5.65	0.43	6.96	0.37
			400	5.57	0.31	12.87	0.48	6.02	0.26	7.64	0.20
			1000	5.41	0.33	12.23	0.47	5.81	0.28	7.15	0.23
9	1.00	2.00	200	4.76	0.39	10.68	0.75	3.49	0.63	5.17	0.59
			400	5.58	0.31	13.48	0.51	4.63	0.41	8.05	0.25
			1000	5.85	0.28	14.64	0.35	4.84	0.34	9.07	0.13
10	1.00	3.00	200	5.01	0.36	11.20	0.73	3.82	0.59	5.87	0.53
			400	5.57	0.31	13.13	0.51	4.54	0.41	7.82	0.23
			1000	5.59	0.31	13.26	0.42	4.36	0.38	7.78	0.19
11	1.00	4.00	200	5.26	0.34	11.88	0.68	4.17	0.55	6.58	0.44
			400	5.58	0.31	13.04	0.49	4.46	0.40	7.75	0.21
			1000	5.49	0.32	12.70	0.45	4.10	0.41	7.27	0.22
12	1.00	5.00	200	5.38	0.32	12.17	0.64	4.32	0.51	6.97	0.37
			400	5.53	0.31	12.93	0.48	4.39	0.40	7.71	0.20
			1000	5.40	0.33	12.42	0.47	3.95	0.43	7.05	0.23
13	1.50	2.00	200	4.85	0.38	10.94	0.74	2.87	0.66	5.33	0.57
			400	5.61	0.30	13.65	0.49	3.44	0.53	8.02	0.23
			1000	5.94	0.28	15.04	0.34	3.37	0.50	9.05	0.13
14	1.50	3.00	200	5.05	0.36	11.37	0.72	3.05	0.64	5.93	0.52
			400	5.59	0.31	13.33	0.49	3.40	0.54	7.78	0.22
			1000	5.64	0.30	13.55	0.42	3.15	0.52	7.69	0.19
15	1.50	4.00	200	5.28	0.34	11.90	0.68	3.27	0.62	6.54	0.44
			400	5.59	0.31	13.20	0.48	3.38	0.54	7.65	0.21
			1000	5.52	0.31	13.00	0.45	3.01	0.54	7.14	0.23
16	1.50	5.00	200	5.40	0.32	12.29	0.64	3.40	0.59	6.93	0.37
			400	5.53	0.31	12.98	0.48	3.30	0.54	7.60	0.21
			1000	5.49	0.32	12.81	0.47	2.93	0.55	6.94	0.24

All reported values are averages over $S = 10,000$ replications.

[a] Multivariate test for no autocorrelation of order 1 and order $1-2$, respectively, in the estimated residuals, see Godfrey (1988). The first columns report the test values, while the second report the corresponding p-values.

[b] Univariate tests for no autocorrelation of order 1 in the estimated residuals, see Godfrey (1988). The first columns report the test values, while the second report the corresponding p-values.

Table 2.4: Misspecification Tests Part 2: Normality

				Vector test normality[b]		Univar. test normality of $\hat{\varepsilon}_{1t}$[b]		Univar. test normality of $\hat{\varepsilon}_{2t}$[b]	
i	τ_b^i	c_Ψ^i	T	$\chi^2(4)$	$p-val$	$\chi^2(2)$	$p-val$	$\chi^2(2)$	$p-val$
1	0.25	2.00	200	82.84	0.06	3.01	0.45	1.98	0.51
			400	352.36	0.01	5.97	0.36	1.98	0.51
			1000	2729.66	0.00	54.37	0.09	2.03	0.50
2	0.25	3.00	200	81.22	0.07	2.85	0.46	2.00	0.51
			400	348.04	0.01	5.54	0.37	1.99	0.51
			1000	2716.42	0.00	51.41	0.10	2.02	0.50
3	0.25	4.00	200	81.75	0.07	2.89	0.46	2.00	0.51
			400	347.58	0.01	5.40	0.38	1.98	0.51
			1000	2727.83	0.00	48.94	0.11	2.01	0.50
4	0.25	5.00	200	84.03	0.08	3.06	0.46	2.00	0.51
			400	355.81	0.01	5.54	0.38	1.97	0.51
			1000	2756.61	0.00	49.01	0.12	2.03	0.50
5	0.50	2.00	200	210.56	0.02	12.99	0.28	1.99	0.51
			400	870.60	0.00	59.43	0.10	2.00	0.51
			1000	5085.27	0.00	734.24	0.00	2.03	0.50
6	0.50	3.00	200	206.60	0.03	12.03	0.29	2.00	0.51
			400	864.28	0.00	56.90	0.11	1.99	0.51
			1000	5085.60	0.00	723.64	0.00	2.02	0.50
7	0.50	4.00	200	205.02	0.03	12.21	0.29	2.00	0.51
			400	855.95	0.00	56.36	0.11	1.99	0.51
			1000	5108.99	0.00	719.21	0.00	2.00	0.50
8	0.50	5.00	200	206.04	0.03	13.31	0.29	2.00	0.51
			400	857.49	0.00	58.34	0.12	1.97	0.51
			1000	5146.05	0.00	733.15	0.00	2.04	0.50
9	1.00	2.00	200	397.42	0.01	94.32	0.08	1.98	0.51
			400	1396.03	0.00	427.84	0.01	1.99	0.51
			1000	6543.87	0.00	3302.14	0.00	2.03	0.50
10	1.00	3.00	200	387.15	0.01	90.43	0.09	2.01	0.51
			400	1382.05	0.00	424.21	0.01	2.00	0.51
			1000	6558.70	0.00	3301.01	0.00	2.01	0.50
11	1.00	4.00	200	388.84	0.01	90.69	0.09	2.02	0.51
			400	1369.90	0.00	421.63	0.01	1.99	0.51
			1000	6592.66	0.00	3314.13	0.00	2.00	0.50
12	1.00	5.00	200	392.25	0.02	95.31	0.09	2.00	0.51
			400	1376.70	0.00	430.00	0.01	1.97	0.51
			1000	6669.62	0.00	3368.25	0.00	2.04	0.50
13	1.50	2.00	200	504.76	0.01	211.03	0.04	1.99	0.51
			400	1579.83	0.00	833.54	0.00	1.98	0.51
			1000	6556.85	0.00	4764.30	0.00	2.03	0.50
14	1.50	3.00	200	494.16	0.01	204.70	0.04	2.01	0.51
			400	1573.99	0.00	829.04	0.00	1.99	0.51
			1000	6591.70	0.00	4776.29	0.00	2.01	0.50
15	1.50	4.00	200	493.83	0.01	205.75	0.04	2.00	0.51
			400	1564.44	0.00	828.07	0.00	1.99	0.51
			1000	6648.16	0.00	4815.04	0.00	2.00	0.50
16	1.50	5.00	200	503.71	0.01	209.60	0.04	2.00	0.51
			400	1580.12	0.00	842.45	0.00	1.97	0.51
			1000	6716.39	0.00	4884.33	0.00	2.04	0.50

All reported values are averages over $S = 10,000$ replications.

[a] Multivariate test for normality of the estimated residuals, see Doornik and Hansen (1994).

[b] Univariate tests for normality of the estimated residuals, see Doornik and Hansen (1994).

Table 2.5: Misspecification Tests Part 3: Skewness, Kurtosis, and Standard Deviation

				Skewness[a]		Kurtosis[a]		Std.dev.[a]	
i	τ_b^i	c_Ψ^i	T	$\hat{\varepsilon}_{1t}$	$\hat{\varepsilon}_{2t}$	$\hat{\varepsilon}_{1t}$	$\hat{\varepsilon}_{2t}$	$\hat{\varepsilon}_{1t}$	$\hat{\varepsilon}_{2t}$
1	0.25	2.00	200	-0.02	-0.00	3.15	2.98	0.56	0.53
			400	-0.03	0.00	3.31	2.98	0.56	0.54
			1000	-0.03	-0.00	4.09	3.00	0.58	0.54
2	0.25	3.00	200	-0.00	0.00	3.12	2.98	0.57	0.54
			400	-0.01	-0.00	3.28	2.99	0.57	0.55
			1000	-0.02	-0.00	4.05	3.00	0.58	0.54
3	0.25	4.00	200	0.00	-0.00	3.12	2.98	0.57	0.55
			400	-0.00	0.00	3.27	2.98	0.57	0.55
			1000	-0.01	-0.00	4.01	2.99	0.58	0.55
4	0.25	5.00	200	0.00	-0.00	3.14	2.97	0.58	0.56
			400	0.00	0.00	3.27	2.98	0.58	0.55
			1000	-0.00	-0.00	4.01	3.00	0.58	0.55
5	0.50	2.00	200	-0.01	-0.00	3.99	2.98	0.58	0.54
			400	-0.01	-0.00	5.20	2.98	0.60	0.54
			1000	-0.00	-0.00	11.11	3.00	0.65	0.54
6	0.50	3.00	200	0.01	0.00	3.92	2.98	0.59	0.55
			400	0.01	-0.00	5.12	2.99	0.60	0.55
			1000	0.00	-0.00	11.02	3.00	0.65	0.55
7	0.50	4.00	200	0.02	-0.00	3.94	2.98	0.59	0.55
			400	0.02	0.00	5.11	2.99	0.61	0.55
			1000	0.01	-0.00	10.99	2.99	0.65	0.55
8	0.50	5.00	200	0.02	-0.00	4.00	2.98	0.60	0.56
			400	0.03	-0.00	5.17	2.98	0.61	0.55
			1000	0.02	-0.00	11.16	3.00	0.65	0.55
9	1.00	2.00	200	0.07	-0.00	9.19	2.98	0.68	0.54
			400	0.08	-0.00	15.33	2.98	0.74	0.55
			1000	0.06	-0.00	35.76	2.99	0.90	0.55
10	1.00	3.00	200	0.07	0.00	8.99	2.98	0.68	0.55
			400	0.10	-0.00	15.32	2.99	0.74	0.55
			1000	0.06	-0.00	35.87	2.99	0.91	0.55
11	1.00	4.00	200	0.08	-0.00	9.11	2.98	0.69	0.55
			400	0.10	-0.00	15.49	2.99	0.74	0.55
			1000	0.06	-0.00	36.09	2.99	0.91	0.55
12	1.00	5.00	200	0.07	-0.00	9.52	2.98	0.69	0.56
			400	0.10	-0.00	15.98	2.98	0.74	0.55
			1000	0.06	-0.00	37.04	2.99	0.91	0.55
13	1.50	2.00	200	0.10	-0.00	16.26	2.98	0.83	0.54
			400	0.15	-0.00	26.56	2.98	0.95	0.55
			1000	0.08	-0.00	49.79	2.99	1.28	0.55
14	1.50	3.00	200	0.10	0.00	16.01	2.98	0.83	0.55
			400	0.16	-0.00	26.66	2.99	0.95	0.55
			1000	0.08	-0.00	50.14	2.99	1.28	0.55
15	1.50	4.00	200	0.11	-0.00	16.28	2.98	0.84	0.55
			400	0.15	-0.00	27.09	2.99	0.95	0.55
			1000	0.08	-0.00	50.61	2.99	1.28	0.55
16	1.50	5.00	200	0.09	-0.00	16.86	2.98	0.85	0.56
			400	0.15	-0.00	27.86	2.98	0.96	0.55
			1000	0.08	-0.00	51.91	2.99	1.29	0.55

All reported values are averages over $S = 10,000$ replications.

[a] The skewness of the estimated residuals is calculated as $skewness_i = T^{-1}\sum_{t=1}^{T}(\hat{\varepsilon}_{it}/\hat{\sigma}_i)^3$ and $kurtosis_i = T^{-1}\sum_{t=1}^{T}(\hat{\varepsilon}_{it}/\hat{\sigma}_i)^4$, where $\hat{\varepsilon}_{it}$ are the estimated system residuals and $\hat{\sigma}_i$ their standard deviations for $i = 1, 2$ as reported in the final column, see Juselius (2006, p. 75).

Table 2.6: Misspecification Tests Part 4: ARCH

				Vector test no. ARCH order 1[a]		Vector test no ARCH order 1-2[a]		Univar. test no ARCH order 1 in $\hat{\varepsilon}_{1t}$[b]		Univ. test no ARCH order 1 in $\hat{\varepsilon}_{2t}$[b]	
i	τ_b^i	c_Ψ^i	T	$\chi^2(9)$	$p-val$	$\chi^2(18)$	$p-val$	$\chi^2(1)$	$p-val$	$\chi^2(1)$	$p-val$
1	0.25	2.00	200	31.43	0.40	40.00	0.41	0.99	0.51	0.95	0.51
			400	19.84	0.51	28.76	0.52	1.03	0.51	0.98	0.50
			1000	9.78	0.62	19.38	0.60	1.01	0.53	1.00	0.50
2	0.25	3.00	200	28.16	0.43	36.77	0.44	1.00	0.51	0.97	0.51
			400	15.46	0.56	24.35	0.56	1.00	0.51	0.97	0.51
			1000	9.66	0.62	19.32	0.60	1.01	0.53	0.99	0.50
3	0.25	4.00	200	23.84	0.48	32.40	0.48	1.00	0.51	0.98	0.50
			400	11.76	0.60	20.77	0.59	0.99	0.51	0.97	0.50
			1000	9.73	0.61	19.35	0.60	1.02	0.53	1.00	0.50
4	0.25	5.00	200	20.37	0.52	29.03	0.51	1.00	0.51	0.98	0.51
			400	10.12	0.62	19.01	0.61	1.01	0.51	0.99	0.50
			1000	9.77	0.61	19.46	0.60	1.01	0.52	1.02	0.50
5	0.50	2.00	200	20.43	0.46	29.97	0.46	0.95	0.53	0.96	0.51
			400	13.22	0.53	23.49	0.52	0.94	0.56	0.96	0.51
			1000	11.39	0.59	22.21	0.57	1.18	0.63	1.01	0.50
6	0.50	3.00	200	18.83	0.49	28.35	0.49	0.95	0.53	0.97	0.51
			400	12.00	0.55	22.18	0.54	0.93	0.56	0.96	0.51
			1000	11.40	0.59	22.30	0.57	1.19	0.63	0.99	0.50
7	0.50	4.00	200	16.69	0.52	26.13	0.51	0.95	0.53	0.98	0.50
			400	11.04	0.57	21.41	0.55	0.93	0.56	0.97	0.50
			1000	11.38	0.58	22.22	0.57	1.16	0.63	1.00	0.50
8	0.50	5.00	200	15.12	0.54	24.50	0.53	0.95	0.53	0.98	0.51
			400	10.55	0.57	20.72	0.56	0.94	0.56	0.98	0.50
			1000	11.40	0.58	22.35	0.56	1.20	0.63	1.02	0.50
9	1.00	2.00	200	15.29	0.51	26.01	0.49	0.86	0.62	0.96	0.51
			400	12.00	0.55	23.10	0.53	1.00	0.67	0.97	0.50
			1000	11.98	0.61	23.24	0.59	1.89	0.69	1.01	0.50
10	1.00	3.00	200	14.33	0.52	24.97	0.50	0.84	0.62	0.97	0.51
			400	11.73	0.56	22.80	0.54	0.99	0.67	0.96	0.51
			1000	11.92	0.61	23.29	0.59	1.85	0.69	0.99	0.50
11	1.00	4.00	200	13.40	0.54	23.88	0.51	0.86	0.62	0.98	0.50
			400	11.49	0.56	22.57	0.54	0.97	0.67	0.97	0.50
			1000	11.85	0.61	23.27	0.59	1.86	0.69	1.00	0.50
12	1.00	5.00	200	12.66	0.55	23.07	0.53	0.87	0.62	0.98	0.51
			400	11.34	0.57	22.41	0.54	1.00	0.67	0.98	0.50
			1000	12.07	0.61	23.54	0.58	1.89	0.70	1.02	0.50
13	1.50	2.00	200	14.00	0.54	25.15	0.51	0.94	0.68	0.96	0.51
			400	11.67	0.58	22.94	0.55	1.37	0.69	0.97	0.50
			1000	12.65	0.62	24.77	0.58	2.68	0.65	1.00	0.50
14	1.50	3.00	200	13.23	0.55	24.15	0.52	0.92	0.68	0.97	0.51
			400	11.53	0.59	22.87	0.56	1.35	0.69	0.97	0.51
			1000	12.71	0.62	24.90	0.58	2.68	0.65	0.99	0.50
15	1.50	4.00	200	12.73	0.56	23.58	0.52	0.97	0.67	0.98	0.51
			400	11.44	0.59	22.78	0.56	1.37	0.70	0.97	0.50
			1000	12.78	0.62	25.17	0.58	2.68	0.65	1.00	0.50
16	1.50	5.00	200	12.03	0.57	22.92	0.54	0.94	0.68	0.98	0.50
			400	11.47	0.59	22.87	0.56	1.42	0.70	0.98	0.50
			1000	13.05	0.62	25.43	0.58	2.78	0.65	1.02	0.50

All reported values are averages over $S = 10,000$ replications.

[a] Multivariate test for no ARCH of order 1 and order $1-2$, respectively, in the estimated residuals, see Lötkepohl and Krätzig (2004).

[b] Univariate tests for no ARCH of order 1 in the estimated residuals, see Lötkepohl and Krätzig (2004).

To conclude on the results from the misspecification tests, it appears that the estimated unrestricted VAR models are fairly good statistical representations of the simulated data when an adequate lag-length is selected, despite the misspecification of the econometric model compared to the data-generating process used to simulate the data. In particular, it appears that the stochastic autoregressive specification of the VAR model is so flexible that it can be used as a valid statistical representation of x_t, which was simulated as a combination of a segmented trend around a non-stationary trend and a standard stationary autoregressive process. This holds even when the bounds on the segmented trend are quite wide, so that the swings caused by the segmented trend are quite long and persistent. Moreover, it appears that the jumps in b_t cause a number of 'outliers' in the equation for p_t, and as a consequence excess kurtosis are found in the estimated residuals in that equation and normality of the estimated residuals is rejeceted, even in small samples. However, overall the estimated VAR models appear to be fairly well-specified based on the residual misspecifiation tests.

2.5.3 Reduced Rank Tests and Estimates

The reduced rank tests are reported in Table 2.7. The reduced rank tests test the model with a rank of $r = 0$ and $r = 1$, respectively, against the unrestricted model with full rank $r = p$. A rank of $r = 0$ corresponds to no cointegration in the system, while a rank of $r = 1$ corresponds to one cointegration relation and $p - r = 1$ common stochastic trend in the system. For all specifications a rank of $r = 1$ cannot be rejected on average over all repititions, with p-values well over 0.05 in most cases. However, only for low values of τ_b^i and c_Ψ^i can a rank of $r = 0$ be rejected, which would lead us to choose a reduced rank of $r = 1$. It must be pointed out, though, that in general setting the rank is a difficult choice which should not be based solely on the trace test, but by a combination of different indices—such as the number of near-unit roots in the system—as suggested by Juselius (2006).

It is clear from Table 2.7, that increasing τ_b^i does not appear to have a large impact on the rank tests, while an increase in c_Ψ^i has a large impact leading the test for a rank of $r = 0$ to not be rejected, so that the rank test indicates a preferred rank of $r = 0$. This indicates that the larger fluctuations of current earnings around the long-run trend, the harder it is to find cointegration between the stock price and earnings based on the multivariate rank test. This result is not surprising as an increase in c_Ψ^i makes the deviations between current earnings and the long-run trend in earnings longer and more persistent as the segmented trend is allowed to move further away from $\overline{x}_t$ before a reversal occurs. The greater persistence in the relation between the current earnings and the long-run trend in earnings implies that deviations the estimated cointegration relation in the estimated cointegrated VAR model becomes more persistent, and consequently the largest estimated eigenvalues decreases. Moreover, the greater persistence implies that a second near-unit root is found in the estimated model, as seen from the columns with the roots of the companion matrix in Table 2.7. Simultaneously, the estimated cointegration adjustment coefficients $\widehat{\alpha}_i$ for $i = p, x$ in the reduced rank model with $r = 1$ decrease, as the estimated adjustment to the cointegration relations becomes slower, see Table 2.9.

Tables 2.8 and 2.9 present the estimated cointegration coefficients $\widehat{\beta}$ and the adjustment coeffi-

Table 2.7: Reduced Rank Determination: Rank Test

i	τ_b^i	c_Ψ^i	T	Reduced rank tests $\mathcal{H}(r)$ against $\mathcal{H}(p)$[a]				Roots of comp. matrix[b]		
				$\mathcal{H}(0)$	$p-val$	$\mathcal{H}(1)$	$p-val$	$\hat{v}_{1,r=2}$	$\hat{v}_{2,r=2}$	$\hat{v}_{2,r=1}$
1	0.25	2.00	200	34.04	0.01	3.13	0.24	0.97	0.72	0.72
			400	45.43	0.00	2.37	0.30	0.99	0.80	0.80
			1000	48.58	0.00	1.35	0.40	1.00	0.90	0.90
2	0.25	3.00	200	24.37	0.04	3.01	0.24	0.97	0.81	0.81
			400	29.53	0.02	2.33	0.30	0.99	0.87	0.87
			1000	34.52	0.00	1.34	0.40	1.00	0.93	0.93
3	0.25	4.00	200	16.97	0.13	2.77	0.24	0.97	0.88	0.87
			400	18.08	0.09	2.25	0.30	0.99	0.93	0.93
			1000	22.90	0.02	1.34	0.40	1.00	0.96	0.96
4	0.25	5.00	200	12.78	0.27	2.39	0.25	0.98	0.92	0.91
			400	12.20	0.26	2.12	0.30	0.99	0.96	0.96
			1000	15.18	0.11	1.34	0.40	1.00	0.97	0.97
5	0.50	2.00	200	31.42	0.01	3.46	0.20	0.97	0.75	0.75
			400	35.70	0.01	3.04	0.22	0.99	0.84	0.84
			1000	31.34	0.01	1.86	0.30	1.00	0.94	0.94
6	0.50	3.00	200	23.45	0.04	3.34	0.20	0.97	0.82	0.82
			400	25.82	0.03	2.99	0.22	0.99	0.89	0.89
			1000	26.89	0.01	1.88	0.30	1.00	0.95	0.95
7	0.50	4.00	200	17.03	0.13	3.09	0.20	0.97	0.88	0.88
			400	17.69	0.10	2.87	0.22	0.99	0.93	0.93
			1000	21.30	0.03	1.88	0.30	1.00	0.96	0.96
8	0.50	5.00	200	13.19	0.24	2.68	0.21	0.97	0.92	0.91
			400	12.88	0.23	2.65	0.23	0.99	0.96	0.95
			1000	16.33	0.09	1.86	0.30	1.00	0.97	0.97
9	1.00	2.00	200	29.13	0.02	4.24	0.14	0.96	0.77	0.77
			400	29.22	0.03	4.60	0.12	0.98	0.88	0.88
			1000	27.02	0.01	2.71	0.21	0.99	0.95	0.95
10	1.00	3.00	200	22.84	0.05	4.04	0.14	0.96	0.83	0.83
			400	23.60	0.05	4.48	0.12	0.98	0.91	0.91
			1000	25.84	0.01	2.76	0.20	0.99	0.95	0.95
11	1.00	4.00	200	17.47	0.12	3.68	0.15	0.96	0.88	0.87
			400	18.41	0.09	4.12	0.13	0.98	0.93	0.93
			1000	23.58	0.02	2.71	0.20	0.99	0.96	0.96
12	1.00	5.00	200	13.98	0.20	3.13	0.17	0.97	0.91	0.90
			400	14.89	0.16	3.60	0.14	0.98	0.95	0.95
			1000	20.92	0.04	2.55	0.21	0.99	0.96	0.96
13	1.50	2.00	200	28.91	0.02	5.20	0.10	0.95	0.78	0.78
			400	28.92	0.03	5.99	0.08	0.97	0.89	0.88
			1000	30.97	0.01	3.13	0.18	0.99	0.94	0.94
14	1.50	3.00	200	23.25	0.05	4.89	0.10	0.95	0.83	0.83
			400	24.69	0.04	5.74	0.08	0.97	0.91	0.90
			1000	30.41	0.01	3.19	0.17	0.99	0.94	0.94
15	1.50	4.00	200	18.27	0.10	4.34	0.11	0.95	0.88	0.87
			400	20.44	0.06	5.07	0.09	0.97	0.93	0.92
			1000	28.94	0.01	3.06	0.18	0.99	0.94	0.94
16	1.50	5.00	200	15.01	0.17	3.56	0.14	0.96	0.90	0.90
			400	17.34	0.11	4.16	0.11	0.98	0.94	0.93
			1000	27.04	0.01	2.79	0.19	0.99	0.95	0.95

All reported values are averages over $S = 10,000$ replications.

[a] Likelihood ratio test of rank r against the unrestricted model with $r = p$, see Johansen (1996).

[b] $\hat{v}_{i,r=j}$ refers to the modulus of the i'th largest root of the companion matrix for the model with rank $r = j$. Thus, the first two columns are the two largest unrestricted roots of the companion matrix for the unrestricted model, while the final column is the largest unrestricted root in the reduced rank model with $r = 1$ (where the largest root is restricted to a unit root).

Table 2.8: Reduced Rank Estimations with $r = 1$: Cointegration Coefficients β

i	τ_b^i	c_Ψ^i	T	$\hat{\beta}_2$	$\hat{\beta}_2^*$	$se_{\hat{\beta}_2^*}$ [a]	$\tau_{\hat{\beta}_2^*}$ [b]	$\bar{b}_t$ [c]	$\hat{\beta}_2^* - \bar{b}_t$ [d]
1	0.25	2.00	200	-2.57	-2.93	0.59	-12.62	-2.00	-0.93
			400	-2.33	-2.33	0.16	-21.47	-2.00	-0.33
			1000	-2.05	-2.05	0.05	-42.44	-2.00	-0.06
2	0.25	3.00	200	-2.02	-2.46	0.82	-9.20	-2.00	-0.46
			400	-2.44	-2.44	0.30	-15.10	-2.00	-0.44
			1000	-2.05	-2.05	0.07	-31.21	-2.00	-0.05
3	0.25	4.00	200	-1.50	-1.50	1.09	-6.87	-2.00	0.50
			400	-2.22	-2.32	0.53	-10.02	-2.00	-0.32
			1000	-2.04	-2.04	0.11	-21.26	-2.00	-0.04
4	0.25	5.00	200	-1.54	-1.73	1.23	-6.17	-2.00	0.27
			400	-2.14	-2.14	0.88	-7.16	-2.00	-0.14
			1000	-2.04	-2.04	0.17	-14.39	-2.00	-0.04
5	0.50	2.00	200	-2.66	-2.90	0.86	-9.87	-2.00	-0.90
			400	-2.83	-2.83	0.33	-14.79	-2.00	-0.83
			1000	-2.11	-2.11	0.10	-23.30	-2.00	-0.11
6	0.50	3.00	200	-2.06	-2.26	1.12	-7.50	-2.00	-0.26
			400	-2.86	-2.74	0.49	-11.43	-2.00	-0.74
			1000	-2.10	-2.10	0.12	-20.05	-2.00	-0.10
7	0.50	4.00	200	0.90	-1.91	1.29	-5.87	-2.00	0.08
			400	-4.02	-2.54	0.66	-8.37	-2.00	-0.54
			1000	-2.07	-2.07	0.15	-16.08	-2.00	-0.07
8	0.50	5.00	200	6.86	-1.47	1.31	-5.39	-2.00	0.53
			400	-1.11	-2.17	0.89	-6.48	-2.00	-0.18
			1000	-1.99	-1.99	0.19	-12.46	-2.00	0.01
9	1.00	2.00	200	-1.73	-3.35	1.94	-6.52	-2.00	-1.35
			400	-5.63	-3.03	0.80	-8.74	-2.00	-1.03
			1000	-2.15	-2.15	0.19	-12.86	-2.00	-0.16
10	1.00	3.00	200	-4.50	-2.27	1.89	-5.20	-2.00	-0.28
			400	-10.33	-2.84	0.89	-7.36	-2.00	-0.84
			1000	-2.14	-2.14	0.20	-12.20	-2.00	-0.14
11	1.00	4.00	200	-6.03	-1.70	1.80	-4.28	-2.00	0.30
			400	-0.57	-2.58	1.04	-6.05	-2.00	-0.58
			1000	-2.03	-2.03	0.21	-11.11	-2.00	-0.03
12	1.00	5.00	200	-2.23	-1.19	1.27	-4.09	-2.00	0.81
			400	-0.23	-1.60	0.96	-5.19	-2.00	0.40
			1000	-1.94	-1.94	0.22	-9.83	-2.00	0.06
13	1.50	2.00	200	1.31	-3.00	2.45	-4.98	-2.00	-1.00
			400	-1.17	-3.06	1.10	-6.56	-2.00	-1.06
			1000	-2.13	-2.13	0.23	-10.09	-2.00	-0.13
14	1.50	3.00	200	-6.03	-2.69	2.53	-4.08	-2.00	-0.69
			400	-2.49	-2.98	1.13	-5.77	-2.00	-0.98
			1000	-2.12	-2.12	0.24	-9.81	-2.00	-0.12
15	1.50	4.00	200	-1.24	-2.00	1.98	-3.45	-2.00	-0.00
			400	0.41	-2.00	0.97	-4.95	-2.00	-0.00
			1000	-2.00	-2.00	0.24	-9.27	-2.00	0.00
16	1.50	5.00	200	-2.62	-1.19	1.43	-3.38	-2.00	0.81
			400	-1.23	-2.01	0.81	-4.46	-2.00	-0.01
			1000	-1.87	-1.87	0.24	-8.62	-2.00	0.13

The column for $\hat{\beta}_2$ reports averages over $S = 10,000$ replications. However, the columns for $\hat{\beta}_2^*$ report averages where a total of 178 out of the $480,000$ estimates are excluded due to extreme estimates, where $|\hat{\beta}_2| > 1.000$.

[a] Standard error of $\hat{\beta}_2$.

[b] T-value of $\hat{\beta}_2$.

[c] Sample average of the parameter b_t in the simulations.

[d] Difference between the estimated parameter $\hat{\beta}_2$ and the sample average of b_t.

Table 2.9: Reduced Rank Estimations with $r = 1$: Adjustment Coefficients α

i	τ_b^i	c_Ψ^i	T	$\hat{\alpha}_1$	$se_{\hat{\alpha}_1}$ [a]	$\tau_{\hat{\alpha}_1}$ [b]	$\hat{\alpha}_2$	$se_{\hat{\alpha}_2}$ [a]	$\tau_{\hat{\alpha}_2}$ [b]
1	0.25	2.00	200	0.18	0.04	4.65	0.19	0.04	5.02
			400	0.17	0.03	5.86	0.18	0.03	6.40
			1000	0.10	0.02	5.30	0.11	0.02	6.40
2	0.25	3.00	200	0.11	0.03	3.11	0.12	0.03	3.41
			400	0.11	0.02	4.41	0.12	0.02	4.85
			1000	0.07	0.02	4.41	0.08	0.01	5.34
3	0.25	4.00	200	0.05	0.04	1.51	0.06	0.03	1.81
			400	0.06	0.02	2.97	0.07	0.02	3.33
			1000	0.05	0.01	3.47	0.05	0.01	4.23
4	0.25	5.00	200	0.00	0.04	0.28	0.02	0.04	0.60
			400	0.03	0.02	1.80	0.04	0.02	2.12
			1000	0.03	0.01	2.67	0.03	0.01	3.30
5	0.50	2.00	200	0.12	0.03	3.68	0.13	0.03	4.28
			400	0.10	0.02	4.18	0.11	0.02	5.22
			1000	0.03	0.01	2.04	0.05	0.01	4.26
6	0.50	3.00	200	0.08	0.03	2.41	0.09	0.03	2.93
			400	0.07	0.02	3.22	0.08	0.02	4.12
			1000	0.03	0.01	1.89	0.04	0.01	3.94
7	0.50	4.00	200	0.03	0.03	1.06	0.05	0.03	1.55
			400	0.04	0.02	2.12	0.05	0.02	2.90
			1000	0.02	0.01	1.55	0.03	0.01	3.39
8	0.50	5.00	200	-0.01	0.04	-0.02	0.02	0.04	0.49
			400	0.02	0.02	1.14	0.03	0.02	1.85
			1000	0.01	0.01	1.15	0.02	0.01	2.77
9	1.00	2.00	200	0.05	0.02	1.98	0.07	0.02	3.12
			400	0.02	0.02	1.59	0.05	0.01	3.68
			1000	-0.02	0.01	-1.47	0.02	0.01	2.55
10	1.00	3.00	200	0.03	0.02	1.14	0.05	0.02	2.15
			400	0.01	0.02	1.02	0.04	0.01	2.96
			1000	-0.02	0.01	-1.42	0.02	0.01	2.50
11	1.00	4.00	200	-0.00	0.03	0.18	0.03	0.02	1.16
			400	0.00	0.02	0.33	0.03	0.01	2.11
			1000	-0.02	0.01	-1.49	0.02	0.01	2.24
12	1.00	5.00	200	-0.03	0.03	-0.63	0.01	0.03	0.38
			400	-0.01	0.02	-0.38	0.02	0.01	1.32
			1000	-0.02	0.01	-1.60	0.01	0.01	1.90
13	1.50	2.00	200	0.02	0.02	0.85	0.04	0.01	2.45
			400	-0.01	0.02	-0.13	0.03	0.01	2.86
			1000	-0.04	0.01	-3.38	0.01	0.01	1.84
14	1.50	3.00	200	0.00	0.02	0.23	0.03	0.02	1.71
			400	-0.02	0.02	-0.54	0.02	0.01	2.30
			1000	-0.04	0.01	-3.34	0.01	0.01	1.82
15	1.50	4.00	200	-0.02	0.03	-0.57	0.02	0.02	0.89
			400	-0.03	0.02	-1.10	0.02	0.01	1.59
			1000	-0.04	0.01	-3.36	0.01	0.01	1.65
16	1.50	5.00	200	-0.05	0.03	-1.22	0.01	0.02	0.31
			400	-0.04	0.02	-1.60	0.01	0.01	0.98
			1000	-0.04	0.01	-3.38	0.01	0.01	1.41

All reported values are averages over $S = 10,000$ replications.
[a] Standard error of $\hat{\alpha}_1$ and $\hat{\alpha}_2$, respectively.
[b] T-value of $\hat{\alpha}_1$ and $\hat{\alpha}_2$, respectively

cients $\widehat{\alpha}$. The cointegration relations are normalized on $\widehat{\beta}_1$, which is the coefficient to the stock price, so the cointegration relations are given by $p_t - \widehat{\beta}_2 x_t$. Hence, Table 2.8 presents the average estimates of $\widehat{\beta}_2$ over all $S = 10,000$ replications, as well as an average over all replications excluding a total of 178 out of $480,000$ very influential estimates, where the estimated coefficient $\left|\widehat{\beta}_2\right| > 1,000$.[9] For transparency, the average estimates with and without the 178 very influential estimates are shown, but standard errors and t-values are only shown for the latter. Finally, Table 2.8 reports the averages over all replications of the individual sample averages of b_t, as well as the average difference between the estimate $\widehat{\beta}_2$ and the sample average of b_t.

The results show that the estimated coefficients are fairly close to the sample averages of b_t (which are very close to B as expected) when the sample size is long. For a sample size of $T = 200$ the estimated coefficients are in many cases far from the sample averages of b_t. Though, we must point out that even after excluding the 178 most influential replications, the averages of the estimated coefficients are still very influenced by a few number of extreme estimates, which a careful econometric analysis would not get.

From Table 2.9 it can be seen that the estimated adjustment coefficient for the earnings variable is found to be equilibrium adjusting in all specifications (i.e. $\widehat{\alpha}_2 > 0$), while the stock price is generally found to be equilibrium-increasing for low values of τ_b^i and c_Ψ^i (i.e. $\widehat{\alpha}_1 > 0$), and equilibrium-adjusting for higher values of τ_b^i and c_Ψ^i (i.e. $\widehat{\alpha}_1 < 0$). Moreover, in the former case the estimated adjustment coefficients are on average significant, but as τ_b^i and c_Ψ^i increase the significance decreases, and eventually both adjustment coefficients become insignificant. As mentioned above, these results can be understood from the fact that increasing τ_b^i and c_Ψ^i allows for a greater degree of persistence in the fluctuations of p_t and x_t around the common long-run trend in earnings, and hence a greater degree of persistence in $p_t - B'x_t$.

2.6 Conclusion

To conclude on the simulations, the results from the estimations indicate that the cointegrated VAR model—with its system approach, lag structure, and decomposition of the data according to its degree of persistence—can be used as a surprisingly good statistical representation of the simulated data with an adequate lag structure. Importantly, the estimations also show that in many cases a 'correct' reduced rank of $r = 1$ is found, and the estimated cointegration coefficients are close to the corresponding parameters in the simulations. Finally, a large degree of persistence is found in the estimated cointegrated VAR model, which indicates that the underlying bounded instability in the individual processes and parameters shows up as persistence in the cointegrated VAR model.

The results are surprising, in particular when one takes into account that the specification of the simulations do not correspond to the specification of the cointegrated VAR model, and that the simulations have bounded instability in the parameters, while the cointegrated VAR model has constant parameters.

[9] The exclusion criteria used here is somewhat arbitrary, but for transparency results are shown with and without the excluded estimates.

It is important to note that the inclusion of lagged first-differences in the cointegrated VAR model appears to be an extremely important element in the specification of a general unrestricted VAR model as a statistically valid representation of the simulated data. It appears that the underlying bounded instability in the stochastic processes and parameters can be fairly well captured by the lagged first-differences in the short-run structure—so that the estimated residuals are fairly well-specified, though with 'fat tails' in the density of the estimated residuals—while the cointegration relations capture the stable long-run relations in the data.

However, it is worth pointing out that the simulations were based on bounded instability in the short run, but with stability in the causal structure in the long run. On that basis the results here might not be very surprising, and it will be interesting to expand the fairly simple simulations considered here. In future work on bridging IKE models with the cointegrated VAR model and extensions thereof—and more generally the 'data-first approach' to econometrics—I intend to allow for more variables to enter the forecasting strategies in the simulations; allow for contingent change that is not bounded within a narrow range, so that there are contingent changes in the long-run structure of the model; focus directly on testing for structural change in the cointegrated VAR model; and, finally, focus directly on developing and using extensions of the cointegrated VAR model with stochastic parameters to model the parameter-instability directly.

However, the results presented in this paper show that regularities in the simulated outcomes from a simple model embedding key features of IKE can indeed be found econometrically in the frequently used classic cointegrated VAR model. This suggests that the cointegrated VAR model may serve as a good starting point for empirically testing IKE models.

2.7 References

Barberis, N., Shleifer, A. and Vishny, R. W. (1998), 'A Model of Investor Sentiment', *Journal of Financial Economics* **49**, 307–343.

Barsky, R. B. and De Long, J. B. (1993), 'Why Does the Stock Market Fluctuate?', *The Quarterly Journal of ...* **CVIII**(2), 291–311.

Doornik, J. A. (2007), *Object-Oriented Matrix Programming Using Ox*, 3 edn, Timberlake Consultants Press.

Doornik, J. A. and Hansen, H. (1994), An omnibus test for univariate and multivariate normality. Working Paper.

Frydman, R. and Goldberg, M. D. (2007), *Imperfect Knowledge Economics: Exchange Rates and Risk.*, Princeton University Press, Princeton.

Frydman, R. and Goldberg, M. D. (2011), *Beyond Mechanical Markets: Asset Price Swings, Risk, and the Role of the State*, Princeton University Press, Princeton.

Frydman, R. and Goldberg, M. D. (2013 *a*), The contingent expectations hypothesis: Rationality and contingent knowledge in macroeconomics and finance theory. working paper prepared for the INET Annual Plenary Conference in Hong Kong, April 2013.

Frydman, R. and Goldberg, M. D. (2013 *b*), Opening Models of Asset Prices and Risk to Nonroutine Change, *in* R. Frydman and E. S. Phelps, eds, 'Rethinking Expectations: The Way Forward for Macroeconomics', Princeton University Press, Princeton, chapter 6, pp. 207–247.

Godfrey, L. G. (1988), *Misspecification Tests in Econometrics*, Cambridge University Press, Cambridge.

Harris, D., McCabe, B. and Leybourne, S. (2002), 'Stochastic cointegration: estimation and inference', *Journal of Econometrics* **111**(2), 363–384.

Hendry, D. (1995), *Dynamic Econometrics*, Oxford University Press, Oxford.

Hoover (2006 *a*), The Methodology of Econometrics, *in* T. C. Mills and K. Patterson, eds, 'New Palgrave Handbook of Econometrics', Macmillan, London, chapter 2, pp. 61–87.

Hoover, K. D. (2006 *b*), The Past as the Future: The Marshallian Approach to Post Walrasian Econometrics, *in* D. Colander, ed., 'Post Walrasian Macroeconomics: Beyond the Dynamic Stochastic General Equilibrium Model', Cambridge University Press, Cambridge, chapter 12, pp. 239–257.

Johansen, S. (1996), *Likelihood-Based Inference in Vector Autoregressive Models*, Oxford University Press, Oxford.

Juselius, K. (2006), *The Cointegrated VAR Model: Methodology and Applications*, Oxford University Press, Oxford.

Juselius, K. and Johansen, S. r. (2006), Extracting Information from the Data: A European View on Empirical Macro, *in* D. Colander, ed., 'Post Walrasian Macroeconomics: Beyond the Dynamic Stochastic General Equilibrium Model', Cambridge University Press, Cambridge, chapter 16, pp. 301–346.

Lötkepohl, H. and Krätzig, M. (2004), *Applied Time Series Econometrics*, Cambridge University Press, Cambridge.

McCabe, B., Leybourne, S. and Harris, D. (2003), Testing for Stochastic Cointegration and Evidence for Present Value Models. Working Paper.

Mizon, G. E. (1995), Progressive modelling of economic time series: The LSE methodology, *in* 'Macroeconometrics: Developments, Tensions, and Prospects', Kluwer, Dordrecht, pp. 107–70.

Spanos, A. (2006), Econometrics in Retrospect and Prospect, *in* T. C. Mills and K. Patterson, eds, 'New Palgrave Handbook of Econometrics', Macmillan, London, chapter 1, pp. 3–58.

Spanos, A. (2009), 'The Pre-Eminence of Theory versus the European CVAR Perspective in Macroeconometric Modeling', *Economics: The Open-Access, Open-Assessment E-Journal* **3**(10).

Spanos, A. (2010), 'Statistical Adequacy and the Trustworthiness of Empirical Evidence: Statistical vs. Substantive Information', *Economic Modelling* **27**(27), 1436–1452.

Chapter 3

Stochastic Cointegration Parameters as a Source of Persistence in the Cointegrated VAR Model

– A Simulation Study

Morten Nyboe Tabor[†]

June 12, 2013

Abstract

In this paper, I simulate cointegrated data with stochastic cointegration parameters given by $\beta_t = \beta + B_t$, where B_t is a mean zero stationary autoregressive process with different degrees of persistence and volatility, and the linear relations $\beta' X_t$ are asymptotically stationary. The simulated data is analysed econometrically with the classic cointegrated VAR model of Johansen (1996). The results show that the cointegrated VAR model appears to be a fairly well-specified statistical representation of the simulated data, except from in extremely long samples or with near non-stationary persistence and high volatility in B_t. Moreover, the results show that the trace tests suggest the correct reduced rank, except from the extreme cases mentioned. Finally, the results show that the cointegrated VAR model delivers a consistent and very precise estimate of β, even in small samples. However, if there is persistence in the stochastic cointegration parameter B_t in the underlying data-generating process it shows up in the estimated cointegrated VAR model as persistence in the estimated cointegration relations. As a result the estimated eigenvalues become very small and the estimated adjustment coefficients are skewed towards zero. Thereby, the results show that bounded parameter-instability in the underlying data-generating process can potentially be a source of persistence in estimated cointegration relations and corresponding low estimated adjustment coefficients. Such persistence and slow adjustment is frequently found empirically in macroeconomic and financial data, and it has been hard to explain for standard economic theory maintaining the assumption of constant parameters.

[†]Department of Economics, University of Copenhagen and INET Center for Imperfect Knowledge Economics at the University of Copenhagen. Email: morten.nyboe.tabor@econ.ku.dk.
Funding from the Institute for New Economic Thinking (INET) is gratefully acknowledged. I would like to thank Katarina Juselius, Anders Rahbek, Søren Johansen, Andreas Hetland, and Heino Bohn Nielsen for numerous discussions and comments to drafts of this paper. An early draft of this paper was presented at the launch conference of the INET Center for Imperfect Knowledge Economics at the Department of Economics, University of Copenhagen in April 2013.

3.1 Introduction

The cointegrated vector autoregressive (VAR) model of Johansen (1996) has proven extremely useful for representing and modeling of macroeconomic and financial data due to non-stationarity of such data. The estimated cointegration relations can be given a natural interpretation as corresponding to equilibrium relations of economic theories, and specific economic theories can be tested empirically by formulating the theoretical implications as testable restrictions on the parameters of a cointegrated VAR model, see Juselius (2006) and Møller (2008).

However, it is an empirical regularity that deviations from estimated cointegration relations are found to be very persistent, with deviations frequently lasting many years, and consequently the estimated adjustment coefficients are found to be very low. For examples of such persistence found in cointegrated VAR analyses, see for example Juselius (1995, 2009 *a*), Juselius and MacDonald (2004), and Johansen et al. (2010) for analyses of 'purchasing power parity' and uncovered interest parity in foreign exchange markets; Johansen and Juselius (1990) and Juselius (2006) for analyses of domestic money demand and inflation dynamics; and Juselius and Toro (2005) for international monetary transmission effects.

The persistent deviations and slow adjustment to estimated long-run equilibria found empirically have been puzzling for standard economic theory, because it typically predicts a much faster equilibrium adjustment. The 'purchasing power parity (PPP) puzzle' provides a notable example of a 'persistence puzzle' in international finance, with consensus estimates of equilibrium adjustments around 15 percent per year and long swings in exchange rates around the PPP level lasting years, see Rogoff (1996) for an overview. In response to such empirical 'persistence puzzles' economic theory has broadly focused on either adding highly persistent exogenous shocks, caused for example by changes in tastes and technology, causing persistent deviations from equilibrium; or on adding various market imperfections, such as sticky prices or limited capital mobility, causing slow adjustment towards equilibrium. See for example Stockman (1980), Helpman (1981), and Svensson (1985) for theoretical responses to the 'PPP puzzle' of the former kind, and Obstfeld and Rogoff (1995, 2000) for the latter kind.

In this paper, I use simulations to show that persistent deviations from estimated cointegration relations and slow adjustment can arise in the cointegrated VAR model as a consequence of boundedly time-varying cointegration parameters in the underlying data-generating process—even with instant adjustment and without persistent exogenous shocks to the variables.

The simulations consider cointegrated data with stochastic cointegration parameters given by $\beta_t = \beta + B_t$, where B_t is a mean zero stationary autoregressive process. Importantly, the multiplicative process $B_t' X_{t-1}$ is stochastically trendless, see Harris et al. (2002) and McCabe et al. (2003), so the linear relations $\beta_t' X_{t-1}$ and $\beta' X_{t-1}$ are stochastically trendless and the simulated process X_t is a cointegrated process with average cointegration relations $\beta' X_{t-1}$. Simulated data from this class of data-generating processes are analysed econometrically with the cointegrated VAR model. I show that the cointegrated VAR provides a consistent estimate of the unconditional mean of the underlying stochastic cointegration parameters as given by β, and thereby the estimated stationary cointegration relation $\hat{\beta}' X_t$ can be interpreted as the long-run average equilibrium relation. However, persistence in B_t in the underlying data-generating process cause persistence in the estimated cointegration relations $\hat{\beta}'X_t$ in a cointegrated VAR model and correspondingly skews the estimated adjustment coefficients $\hat{\alpha}$ towards zero. Thereby, the results offer a novel potential understanding of the empirical 'persistence puzzle' in cointegrated VAR analyses as a result of bounded parameter-instability in the underlying cointegration relations.

I simulate a simple bivariate system where $\Pi_t = \alpha\beta_t'$ is stochastic and with reduced rank $r = 1$. One variable is a weakly exogenous standard random walk, while the other is instantly adjusting to the cointegration relation $\beta_t'X_{t-1}$. The stochastic cointegration parameters are specified as $\beta_t = \beta + B_t$, where B_t is simulated as a first-order autoregressive process. This simple representation allows me to consider three general cases with respect to B_t and the cointegration properties of the simulated system: i) $B_t = 0$, so the parameters are constant and the system is cointegrated in the standard sense; ii) B_t is stationary with unconditional mean zero, so that $\beta'X_{t-1}$ is stochastically trendless and the simulated process X_t is cointegrated; and, iii) B_t is non-stationary, so that $\beta'X_t$ is non-stationary and the system is not cointegrated. In all three cases the system is simulated for different volatilities of the stochastic parameters and for different sample lengths. Moreover, in the case of stationary parameters I consider both *i.i.d.*, medium persistence, and near non-stationary persistence around a constant unconditional mean.

Using the simulated data, I mimic the modeling of an econometrician (assumed unaware of the data-generating processes used in the simulations) analysing the data with the cointegrated VAR model based on the general-to-specific procedure described in Johansen (1996) and Juselius (2006). Therefore, the econometric analysis is based on standard asymptotic inference. The modeling procedure focuses first on testing the statistical adequacy of an unrestricted model, and second on the cointegration properties of the system by testing for and estimating a reduced rank model.

Due to stochastic cointegration parameters in the data-generating process of the simulated data the cointegrated VAR model is misspecified. However, the simulation study shows that the estimated cointegrated VAR models appear statistically well-specified based on various residual misspecification tests, except from extreme cases with both high persistence and high volatility in the stochastic parameters and in very long samples. In particular, the results show that autocorrelation and persistence in the estimated cointegration relation caused by persistence in B_t can be captured through the short-run structure of the cointegrated VAR model by selecting an adequate lag length, so that the estimated residuals are found not to be autocorrelated.

The results show that the trace tests for reduced rank, on average, suggest the correct rank, though persistence in B_t skews the test distributions of the rank test towards zero. Moreover, the rank test is found to correctly reject cointegration when the stochastic cointegration parameters are non-stationary, in which case the variables are not cointegrated.

Finally, the simulation results show that under stochastic cointegration in the data-generating process, the cointegrated VAR model produces a consistent estimate of the unconditional mean β of the stochastic cointegration parameters β_t, and that the estimator is very precise even in small samples. This is an encouraging result as it shows that the cointegrated VAR model can be used as an approximation of a more complex underlying data-generating process with bounded parameters-instability if the primary focus is to estimate the average long-run relations. Though inference is sensitive to the misspecification of this approximation, the results show that the conclusions with respect to reduced rank and significance of the cointegration and adjustment coefficients are qualitatively correct.

However, if B_t is persistent, but stationary with mean zero, the estimated cointegration relation $\hat{\beta}'X_t$ becomes persistent, but stationary, with the persistence caused solely by the persistence in B_t. As a result the largest estimated eigenvalue falls significantly and the estimated adjustment coefficients $\hat{\alpha}$ get skewed towards zero—though inference on their significance is qualitatively correct. Moreover, when B_t becomes near non-stationary the modulus of the largest unrestricted characteristic root in the system is found to be almost one, and thereby an additional near unit root is found in the estimated cointegrated VAR model. These are the key findings which provide a novel understanding of the source of the 'persistence puzzle' as a result of bounded parameter-instability in the underlying data-generating process.

The rest of the paper is organised as follows. In Section 3.2 the cointegrated VAR model and extensions with stochastic parameters are described. Section 3.3 presents the simulation design and Section 3.4 presents the results from the estimations based on the simulated data. Section 3.5 concludes.

3.2 The Cointegrated VAR Model and Extensions with Stochastic Parameters

Consider the p-dimensional VAR model with one lag and no deterministic terms written in error-correction form

$$\Delta X_t = \Pi X_{t-1} + \varepsilon_t, \tag{3.1}$$

for $t = 1, 2, ..., T$, where Π has dimension $(p \times p)$ and X_0 is fixed. The error terms ε_t are assumed independent and Gaussian with mean zero and covariance Σ.

Assumption 2 *Assume that the autoregressive polynomial $A(z) = I_p(1-z) - \Pi z$ has exactly $p-r$ unit roots at $z = 1$ and the remaining roots are larger than one in absolute value, $|z| > 0$.*

Johansen (1996, Theorem 4.2) shows that under Assumption 2 the matrix Π has reduced rank r and can be represented as

$$\Pi = \alpha\beta', \tag{3.2}$$

where the $(p \times r)$ matrices α and β have full column rank, and X_0 can be given an initial distribution so that X_t has the representation

$$X_t = C_\Sigma X_0 + C_\Sigma \sum_{i=1}^{t} \varepsilon_i + C_S S_t, \tag{3.3}$$

where $S_t := \beta' X_t$, $C_\Sigma := \beta_\perp(\alpha'_\perp\beta_\perp)^{-1}\alpha'_\perp$, $C_S := \alpha(\beta'\alpha)^{-1}$, and $\alpha_\perp$ and $\beta_\perp$ are othogonal matrices of full rank and dimension $(p \times p - r)$, such that $\alpha'_\perp\alpha = 0$ and $\beta'_\perp\beta = 0$. This is an instance of Granger's representation theorem (Engle and Granger, 1987) and it shows that the process X_t can be represented in terms of a stochastic trend, a stationary process, and the initial values. The process X_t is a cointegrated $I(1)$ process with $p-r$ common stochastic trends given by $\alpha'_\perp X_t$ and r cointegration relations given by $\beta' X_t$. The first-difference ΔX_t and the cointegration relations $\beta' X_t$

are $I(0)$. The cointegration relations measure the deviations from the long-run equilibrium relations between the variables, while the α-coefficients determine the adjustments to such deviations.

For the derivations in the next sections it is useful to derive the Granger representation. First, under Assumption 2 the cointegrated VAR model is given by

$$\Delta X_t = \alpha\beta' X_{t-1} + \varepsilon_t. \tag{3.4}$$

Pre-multiplying equation (3.4) with β' and re-arranging, the r linear relations $S_t := \beta' X_t$ are given by

$$S_t = (I_r + \beta'\alpha)S_{t-1} + \beta'\varepsilon_t, \tag{3.5}$$

which is asymptotically stable if $(I_r + \beta'\alpha)$ has eigenvalues inside the unit circle. Under Assumption 2 this holds and X_0 can be given an initial distribution, so that the r-dimensional process S_t is stationary and can be represented as

$$S_t = \sum_{i=0}^{\infty}(I_r - \beta'\alpha)^i \beta' \varepsilon_{t-i}. \tag{3.6}$$

Pre-multiplying equation (3.4) with $\alpha_\perp'$, re-arranging, and cumulating over t, the $(p-r)$-dimensional process $\alpha_\perp' X_t$ is given by

$$\alpha_\perp' X_t = \alpha_\perp' X_0 + \alpha_\perp' \sum_{i=1}^{t} \varepsilon_i, \tag{3.7}$$

which are the $p-r$ common stochastic trends in X_t.

Using the following skew-projection from Johansen (1996),

$$I_p = \beta_\perp(\alpha_\perp'\beta_\perp)^{-1}\alpha_\perp' + \alpha(\beta'\alpha)^{-1}\beta', \tag{3.8}$$

X_t can be decomposed into

$$X_t = \beta_\perp(\alpha_\perp'\beta_\perp)^{-1}\alpha_\perp' X_t + \alpha(\beta'\alpha)^{-1}\beta' X_t, \tag{3.9}$$

and by plugging in for $\alpha_\perp' X_t$ from equation (3.7) and $S_t := \beta' X_t$ from equation (3.6) it follows that X_t has the representation in equation (3.3).

In the next sections, two extensions of the standard cointegration framework with stochastic parameters are considered. In the first case α_t is stochastic and β constant, and in the second case α is constant while β_t is stochastic. Like the case with constant parameters we want to show that under specific assumptions the linear combinations $\beta' X_t$ are stationary and find a representation for X_t.

3.2.1 Stochastic Adjustment Parameters α_t

In the first case the stochastic adjustment parameter is given by

$$\alpha_t = \alpha + A_t, \tag{3.10}$$

where A_t is *i.i.d.* Gaussian with mean zero and of dimension $(p \times r)$. The cointegration parameters β are constant and the model is then given by

$$\Delta X_t = \alpha_t \beta' X_{t-1} + \varepsilon_t. \tag{3.11}$$

As in the case with constant parameters above, consider the r-dimensional process $S_t := \beta' X_t$, which is given by

$$S_t = (I_r + \beta'\alpha_t)S_{t-1} + \beta'\varepsilon_t, \tag{3.12}$$

where $(I_r + \beta'\alpha_t)$ is *i.i.d.* Gaussian with unconditional mean $(I_r + \beta'\alpha)$. The process S_t is a random coefficient autoregressive process with stochastic autoregressive parameter $(I_r + \beta'\alpha_t)$, which under regularity conditions specified in e.g. Theorem 2 in Rahbek and Nielsen (2012) is geometrically ergodic and has a stationary representation.

Next, multiply equation (3.11) with the orthogonal matrix $\alpha_\perp$ (where $\alpha_\perp'\alpha = 0$ as above, so that $\alpha_\perp'\alpha_t = \alpha_\perp'\alpha + \alpha_\perp' A_t = \alpha_\perp' A_t$) to get

$$\alpha_\perp'\Delta X_t = \alpha_\perp' A_t S_{t-1} + \alpha_\perp'\varepsilon_t \tag{3.13}$$

where $\alpha_\perp' A_t$ is *i.i.d.* mean zero Gaussian. Re-arranging and cumulating over t, the linear process $\alpha_\perp' X_t$ is given by

$$\alpha_\perp' X_t = \alpha_\perp' X_0 + \alpha_\perp' \sum_{i=1}^{t} (\varepsilon_i + A_i S_{i-1}). \tag{3.14}$$

Using the skew-projection in equation (3.9) again, with S_t and $\alpha_\perp' X_t$ given in equations (3.12) and (3.14), respectively, it follows that X_t has the represention

$$X_t = C_\Sigma X_0 + C_\Sigma \sum_{i=1}^{t} (\varepsilon_i + A_i S_{i-1}) + C_S^{-1} S_t, \tag{3.15}$$

where $C_\Sigma := \beta_\perp(\alpha_\perp'\beta_\perp)^{-1}\alpha_\perp'$ and $C_S := \alpha(\beta'\alpha)^{-1}$, as above. Thus, the process X_t can be represented in terms of a stochastic trend, a stationary process, and the initial values. The stochastic trend component $C\sum_{i=1}^{t}(\varepsilon_i + A_i S_{i-1})$ is asymptotically equivalent to a random walk process, as it satisfies a functional central limit theorem if normalized correctly, see also Kristensen and Rahbek (2010, 2013). To sum up, for $\alpha_t = \alpha + A_t$ with A_t *i.i.d.* mean zero Gaussian, the process X_t is asymptotically $I(1)$ and the cointegration relations $S_t := \beta' X_t$ are stationary under regularity conditions.

Extensions of the classic cointegrated VAR model within the class of models with stochastic adjustment parameters, α_t, has recently been an active area of research. The asymptotic theory has been developed and estimation techniques for this class of models have already been developed. For example, Bec and Rahbek (2004) consider regime switching in the adjustment coefficients; Bec et al. (2008) and Paruolo et al. (2013) consider mixture models with smooth transition in the adjustment parameters; and Kristensen and Rahbek (2010, 2013) consider nonlinear and asymmetric adjustments parameters.

3.2.2 Stochastic Cointegration Parameters β_t

Consider next the case of stochastic cointegration parameters as given by

$$\beta_t = \beta + B_t, \tag{3.16}$$

where B_t is *i.i.d.* Gaussian with mean zero and of dimension $(p \times r)$. The adjustment parameters α are constant and the model is then given by

$$\Delta X_t = \alpha\beta_t' X_{t-1} + \varepsilon_t. \tag{3.17}$$

Consider again the r linear relations $S_t := \beta' X_t$, which are now given by

$$\begin{aligned} S_t &= (I_r + \beta'\alpha)S_{t-1} + \beta'\alpha B_t' X_{t-1} + \beta'\varepsilon_t \\ &= \sum_{i=0}^{t-1}(I_r + \beta'\alpha)^i \beta'\varepsilon_t + \sum_{i=0}^{t-1}(I_r + \beta'\alpha)^i \beta'\alpha B_{t-i}' X_{t-1-i} \end{aligned} \tag{3.18}$$

where $S_0 = 0$ is assumed for simplicity. The first term is recognized from the standard case with constant parameters and is stationary if $(I_r + \beta'\alpha)$ has eigenvalues inside the unit circle, which is the case under Assumption 2. The second term is a sum of the multiplicative process $B_{t-i}' X_{t-1-i}$ for $i = 1, ..., t-1$ with exponentially decreasing coefficients. This multiplicative process is stochastically trendless, as defined by Harris et al. (2002) and McCabe et al. (2003), with $E[B_t' X_{t-1}] = 0$. This holds because the stochastically trendless property of B_t dominates the multiplicative process, even though X_t has a stochastic trend. Thus, the process S_t is asymptotically stationary despite the presence of $B_t' X_{t-1}$.

Next, pre-multiply equation (3.17) by $\alpha_\perp'$, re-arrange, and cumulate over t to get

$$\alpha_\perp' X_t = \alpha_\perp' X_0 + \alpha_\perp' \sum_{i=1}^{t} \varepsilon_t, \tag{3.19}$$

as $\alpha_\perp' \alpha = 0$. Hence, the $p - r$ common stochastic trends $\alpha_\perp' X_t$ are identical to the standard case with constant parameters in equation (3.7), and hence they are $I(1)$.

Finally, using the skew-projection in equation (3.9) again, with S_t and $\alpha_\perp' X_t$ given in equations (3.18) and (3.19), respectively, it follows that X_t has the represention

Finally, by using the skew-projection in equation (3.9) again it follows that X_t has the representation

$$X_t = C_\Sigma X_0 + C_\Sigma \sum_{i=1}^{t} \varepsilon_i + C_S S_t, \tag{3.20}$$

where, once again, $C_\Sigma := \beta_\perp(\alpha_\perp' \beta_\perp)^{-1}\alpha_\perp'$ and $C_S := \alpha(\beta'\alpha)^{-1}$. Now the process X_t can be represented in terms of a stochastic trend, a stochastically trendless process, and the initial values. The common stochastic trends are $I(1)$ and identical to the case with constant parameters. The component $\alpha(\beta'\alpha)^{-1} S_t$ is stochastically trendless as the multiplicative process $B_t' X_{t-1}$ is stochastically trendless. However, the theory for stochastic cointegration parameters β_t is not yet developed and, in particular, we miss a full theory for the impact of the last term in the interpretation of cointegration. To sum up, for $\beta_t = \beta + B_t$ with B_t *i.i.d.* mean zero Gaussian, the process X_t is $I(1)$ and the cointegration relations $S_t := \beta' X_t$ are stochastically trendless.

The simulations presented next consider a more general case of stochastic cointegration parameters $\beta_t = \beta + B_t$ than considered above, as B_t is simulated as a mean zero first-order autoregressive process rather than a mean zero Gaussian process. However, the stochastically trendless property of the multiplicative process $B_t' X_{t-1}$ holds in the more general case where B_t is a mean zero stationary autoregressive process, see Harris et al. (2002) and McCabe et al. (2003), and thereby $\beta' X_t$ is still stochastically trendless. The purpose is to use simulations to address to what extent the cointegrated VAR model can be used as a approximation to estimate β when the data-generating process has a stochastic $\beta_t = \beta + B_t$. Moreover, it is of interest to what extent the misspecification caused by the presence of $B_t \neq 0$ in the data-generating process can be identified using misspecification tests and how it affects the estimations and inference.

Note that both cases with stochastic parameters considered nest standard cointegration as the special case where $A_t = 0$ or $B_t = 0$. In these cases the representations of X_t in equations (3.15) and (3.20) are reduced to the representation of X_t in equation (3.9).

3.3 The Simulation Design

The data-generating process for the simulated time series $X_t^{(i)}$ for $i = 0, 1, ..., 16$ is the simple bivariate system given by

$$\Delta X_t^{(i)} = \alpha \beta_t^{(i)\prime} X_{t-1}^{(i)} + \varepsilon_t, \tag{3.21}$$

for $t = 1, 2, ..., T$, where $X_t^{(i)}$ has dimension $p = 2$. The system is initialized at $X_{10} = X_{20} = 0$. The random shocks ε_t are *i.i.d.* standard Gaussian and mutually uncorrelated.

The parameters α and $\beta_t^{(i)}$ of dimension $(p \times r)$ are given by

$$\alpha = \begin{bmatrix} -1 \\ 0 \end{bmatrix} \quad \text{and} \quad \beta_t^{(i)} = \beta + B_t^{(i)} = \begin{bmatrix} 1 \\ -b \end{bmatrix} + \begin{bmatrix} 0 \\ -b_t^{(i)} \end{bmatrix}, \tag{3.22}$$

where $\beta = (1, -b)'$ and $B_t^{(i)} = (0, -b_t^{(i)})'$, so that $\Pi_t^{(i)} = \alpha \beta_t^{(i)\prime}$ has reduced rank $r = 1$. The stochastic parameter $b_t^{(i)}$ has dimension (1×1) and is simulated as an $AR(1)$ process given by

$$b_t^{(i)} = \rho^{(i)} b_{t-1}^{(i)} + \sigma^{(i)} \eta_t, \tag{3.23}$$

where the random shocks η_t are *i.i.d.* standard Gaussian and uncorrelated with ε_t. For $-1 < \rho^{(i)} < 1$ the process $b_t^{(i)}$ is stationary with unconditional mean zero, and hence the process $b + b_t^{(i)}$ is stationary with unconditional mean b. The system is simulated for different combinations of the parameters $(\rho^{(i)}, \sigma^{(i)})$, while $b = 1$ is fixed and $b_t^{(i)}$ is initialized at $b_0 = 0$. Hence, the differences between the simulated time-series stems solely from differences in the parameters $\rho^{(i)}$ and $\sigma^{(i)}$.

The cointegration properties of $X_t^{(i)}$ depend on whether the stochastic parameter $b_t^{(i)}$ is stationary, which is determined by the autoregressive parameter $\rho^{(i)}$. The parameter $\sigma^{(i)}$ determines the standard deviation of shocks to $b_t^{(i)}$ relative to the normalized shocks to the system variables $X_t^{(i)}$.

As $\Delta X_{2t} = \varepsilon_{2t}$ and $X_{20} = 0$ the variable X_{2t} is a weakly exogenous random walk given by

$$X_{2t} = \sum_{i=1}^{t} \varepsilon_{2i}, \tag{3.24}$$

and hence it is identical for all i and integrated of first order, $X_{2t} \sim I(1)$, in the standard sense.

The level of $X_{1t}^{(i)}$ is given by

$$X_{1t}^{(i)} = (1 + b_t^{(i)}) X_{2t-1} + \varepsilon_{1t} = \sum_{i=1}^{t-1} \varepsilon_{2i} + b_t^{(i)} \sum_{i=1}^{t-1} \varepsilon_{2i} + \varepsilon_{1t}, \tag{3.25}$$

Thus, $X_{1t}^{(i)}$ contains a standard $I(1)$ stochastic trend given by the cumulated shocks to the weakly exogenous variable X_{2t}, a multiplicative process of $b_t^{(i)}$ and the stochastic trend, and an *i.i.d.* Gaussian error term. For $-1 < \rho^{(i)} < 1$ the stochastic parameter $b_t^{(i)}$ is stationary with unconditional mean zero and the multiplicative process $b_t^{(i)} X_{2t-1}$ becomes stochastically trendless—as the stochastically trendless property of $b_t^{(i)}$ dominates the multiplicative process—and heteroskedastic.

Finally, the r linear relations $\beta' X_t^{(i)} = X_{1t}^{(i)} - X_{2t}$ are given by

$$\beta' X_t^{(i)} = \varepsilon_{1t} - \varepsilon_{2t} + b_t^{(i)} \sum_{i=1}^{t-1} \varepsilon_{2i}, \tag{3.26}$$

which is a combination of the *i.i.d.* Guassian error terms and the multiplicative process $b_t^{(i)} X_{2t-1}$, as the common stochastic trend $\sum_i^{t-1} \varepsilon_{2t-1}$ cancels out. Once again, for $-1 < \rho^{(i)} < 1$ the multiplicative process $b_t^{(i)} X_{2t-1} = b_t^{(i)} \sum_{i=1}^{t-1} \varepsilon_{2i}$ is stochastically trendless, with $E[b_t^{(i)} X_{2t-1}] = 0$, and heteroskedastic. Hence, the linear relation $\beta' X_t^{(i)}$ becomes stochastically trendless and heteroskedastic, and $X_t^{(i)}$ is a cointegrated process with average cointegration relations $\beta' X_t^{(i)} = X_{1t}^{(i)} - X_{2t}$.

In total $N = 17$ data-generating processes $i = 0, 1, 2, \ldots, N-1$ are simulated. The first simulated system is the benchmark case of constant parameters, so for $i = 0$ the parameters are $(\rho^{(0)}, \sigma^{(0)}) = (0, 0)$. In the 16 remaining data-generating processes $(\rho^{(i)}, \sigma^{(i)})$ are given by all combinations of the following values

$$\rho^{(i)} \in \{0.0; 0.5; 0.95; 1.0\} \tag{3.27}$$

$$\sigma^{(i)} \in \{0.1; 0.2; 0.5; 1.0\}, \tag{3.28}$$

see also Table 3.1, which allows me to consider the following general cases:

- $i = 0$: *Constant parameters and standard cointegration.*
  The benchmark case of constant parameters as $(\rho^{(0)}, \sigma^{(0)}) = (0, 0)$. As $b_t^{(0)} = 0$, the process $X_t^{(0)}$ reduces to a standard cointegrated $I(1)$ process, where X_{2t} is weakly exogenous and $X_{1t}^{(0)}$ is purely adjusting, and the cointegration relation $\beta' X_t^{(0)} = X_{1t}^{(0)} - X_{2t}$ becomes *i.i.d.* mean zero.

- $i = 1$ *to* 12: *Stationary parameters and cointegration.*
  As $0 < \rho^{(i)} < 1$ for $i = 1, 2, ..., 12$, $b_t^{(i)}$ is mean zero stationary, so the linear relation $\beta' X_t^{(i)} = X_{1t}^{(i)} - X_{2t}$ is stochastically trendless and the process $X_t^{(i)}$ is cointegrated. Different degrees of persistence in $b_t^{(i)}$ are considered. First, $\rho^{(i)} = 0.0$ for $i = 1, ..., 4$, so the parameter $b_t^{(i)}$ is *i.i.d.* mean zero Gaussian, so $\beta' X_t^{(i)}$ is mean zero heteroskedastic due to the non-linear effect of shocks to $b_t^{(i)}$ in the multiplicative process $b_t^{(i)} X_{2t-1}$. Second, $\rho^{(i)} = 0.5$ for $i = 5, ..., 8$, so there is medium persistence in the stationary parameter $b_t^{(i)}$ and $\beta' X_t^{(i)}$ becomes heteroskedastic and persistent, but still stochastically trendless. Finally, $\rho^{(i)} = 0.95$ for $i = 9, ..., 12$, so the parameter $b_t^{(i)}$ is almost non-stationary and $\beta' X_t^{(i)}$ becomes heteroskedastic and stochastically trendless, though very persistent.

- $i = 13$ *to* 16: *Non-stationary parameters and no cointegration.*
  As $\rho^{(i)} = 1.0$ for $i = 13, ..., 16$, $b_t^{(i)}$ is non-stationary, so the linear relations $\beta' X_t^{(i)}$ are non-stationary as the multiplicative process $b_t^{(i)} X_{2t-1}$ is not stochastically trendless. Hence, the system is not cointegrated.

Moreover, all cases are considered for different values of $\sigma^{(i)}$, which determines the variance of the shocks to $b_t^{(i)}$ compared to the normalized variances of the shocks to the levels of the variables, and the time series are simulated with sample lengths of $T = \{100, 200, 400, 1000\}$ observations, respectively. The shorter samples are most interestingly from an economic point of view as these are typical sample lengths for macroeconomic and (low frequency) financial data, while $T = 1000$ is included to study the large sample properties of the cointegrated VAR model. For each data generating process $S = 10,000$ replications of the series are simulated for each sample length, so a total of $680,000$ systems of time series are simulated. The same sequences of random shocks ε_t and η_t for $t = 1, 2, ..., T$ are used for simulations for the different specifications, so the only variation between the simulated time-series stem from the different values of $(\rho^{(i)}, \sigma^{(i)})$. All simulations and estimations have been done in Ox 6.20, see Doornik (2007). Simulated outcomes from the different cases are presented in section 3.3.2.

3.3.1 Econometric Modeling and Estimation

For each of the simulated bivariate time-series a cointegrated VAR model is estimated based on the automated procedure described in details below. The idea is to mimic—to the extent possible when estimating $680,000$ models—the procedure of an econometrician, unaware of the actual data-generating process used in the simulations, using the cointegrated VAR model based on the general-to-specific modeling procedure described in Johansen (1996) and Juselius (2006). Therefore, standard asymptotic inference derived based on the assumption of *i.i.d.* Gaussian residuals and constant parameters is used in the entire modeling and testing process.

The econometric approach is based on a 'data-first' approach to econometrics, see Hoover (2006), Hoover et al. (2008), and Juselius (2009 *b*). Based on this approach, the first step of the modeling procedure focuses on specification of an unrestricted VAR model as a statistically adequate description of the data, while the second part focuses on testing and imposing reduced rank restrictions on the model.

First, the unrestricted CVAR model given by

$$\Delta X_t^{(i)} = \Pi X_{t-1}^{(i)} + \sum_{j=1}^{k-1} \Gamma_j \Delta X_{t-j}^{(i)} + \mu_0 + \mu_1 t + \varepsilon_t, \tag{3.29}$$

is estimated for $t = 1, 2, ..., T$. The initial values $X_{-k+1}^{(i)}, ..., X_0^{(i)}$ are fixed and the residuals are assumed independent and Guassian with mean zero and covariance matrix Ω.

The lag length k is chosen as the minimum k needed so that the multivariate test for no first-order autocorrelation of the residuals is not rejected at the 5 percent significance level.[1] Thus, first the model in equation (3.29) is estimated for $k = 1$, and the test for no multivariate first-order autocorrelation is performed based on the estimated residuals. If the null of no autocorrelation is rejected at the 5 percent significance level the model is re-estimated with an additional lag included, i.e. $k = 2$. This specific-to-general process for the lag length specification continues until the null of no autocorrelation in the estimated residuals cannot be rejected.

Given the selected lag lenght, univariate and multivariate misspecification tests for autocorrelation, normality, and ARCH in the estimated residuals of the unrestricted model are performed to

[1] Results similar to those presented are found if the lag length selection is set based on no second-order multivariate autocorrelation or no combined first-second order autocorrelation

check if the estimated model is a statistically adequate representation of the data for the sample considered, see Juselius (2006, chapter 3), Dennis et al. (2006), and references therein for descriptions of the various misspecification tests.

Moreover, the two eigenvalues of the unrestricted model as well as the modulus of the two largest unrestricted characteristic root of the companion matrix for the unrestricted are reported.

Second, the reduced rank model

$$\Delta X_t^{(i)} = \alpha\beta' X_{t-1}^{(i)} + \sum_{j=1}^{k-1} \Gamma_j \Delta X_{t-j}^{(i)} + \mu_0 + \varepsilon_t \tag{3.30}$$

under the hypothesis $\mathcal{H}(r) : \Pi = \alpha\beta'$ is considered, where the constant term μ_0 is included unrestricted. To determine the cointegration rank r, the trace test for the reduced rank $\mathcal{H}(r)$ against $\mathcal{H}(p)$ of Johansen (1996) is performed. A top-down procedure—where $\mathcal{H}(0)$ is tested first, and if rejected $\mathcal{H}(1)$ is tested, etc. until $\mathcal{H}(r)$ cannot be rejected—is used as suggested by Johansen (1996) and Juselius (2006). Next, the reduced rank model for $r = 1$ is estimated for all simulated series, including those where the rank test indicates a reduced rank of $r = 0$.

The maximum likelihood estimator of β solves the eigenvalue problem

$$|\lambda S_{11} - S_{10} S_{00} {-} 1 S_{01}| = 0, \tag{3.31}$$

for the eigenvalues $1 > \hat{\lambda}_1 > ...\hat{\lambda}_p > 0$ and the eigenvectors $\hat{V} = (\hat{v}_1, ..., \hat{v}_p)$, which are normalized by $\hat{V}' S_{11} \hat{V} = I$, see Johansen (1996, Theorem 6.1). The cointegration parameters β are given by the first r eigenvectors

$$\hat{\beta} = (\hat{v}_1, ..., \hat{v}_r), \tag{3.32}$$

where the eigenvectors are ordered according to the eigenvalues. Note the correspondence between the estimated eigenvalues, adjustment parameters, and cointegration parameters as given by

$$\mathrm{diag}(\hat{\lambda}_1, ..., \hat{\lambda}_r) = \hat{\alpha} S_{00}^{-1} \hat{\alpha} = \hat{\beta}' S_{10} S_{00}^{-1} S_{01} \hat{\beta}. \tag{3.33}$$

Each eigenvalue $\hat{\lambda}_i$ can be interpreted as a measure of the 'stationarity' of cointegration relation i; the greater the eigenvalue the 'more stationary' the cointegration relation (Juselius, 2006, p. 119).

The reported estimates of $\hat{\beta}$ are normalized on β_1, and as the β-vector is identified standard errors and corresponding t-values for $\hat{\beta}$ are calculated based on equation (12.13) in Juselius (2006, p. 215). Given cointegration, $\hat{\alpha}$ is estimated by OLS for a given $\hat{\beta}$, with corresponsing standard errors and t-values.

3.3.2 Illustration

Figures 3.1 and 3.2 illustrate a simulated outcome for nine of the data-generating processes. The graphs in the first columns show the simulated variables $X_{1t}^{(i)}$ (red line) and X_{2t} (blue line). Note, that because the same random shocks ε_t are used for all i, the weakly exogenous variables X_{2t} are identical for all i (though the scales of the axes differ in the graphs), while the simulated $X_{1t}^{(i)}$ differ due to the different simulated $b_t^{(i)}$. The simulated outcomes of the stochastic variable $b + b_t^{(i)}$ are shown as the blue line in the graphs in the second columns along with the constant $b = 1$ (red

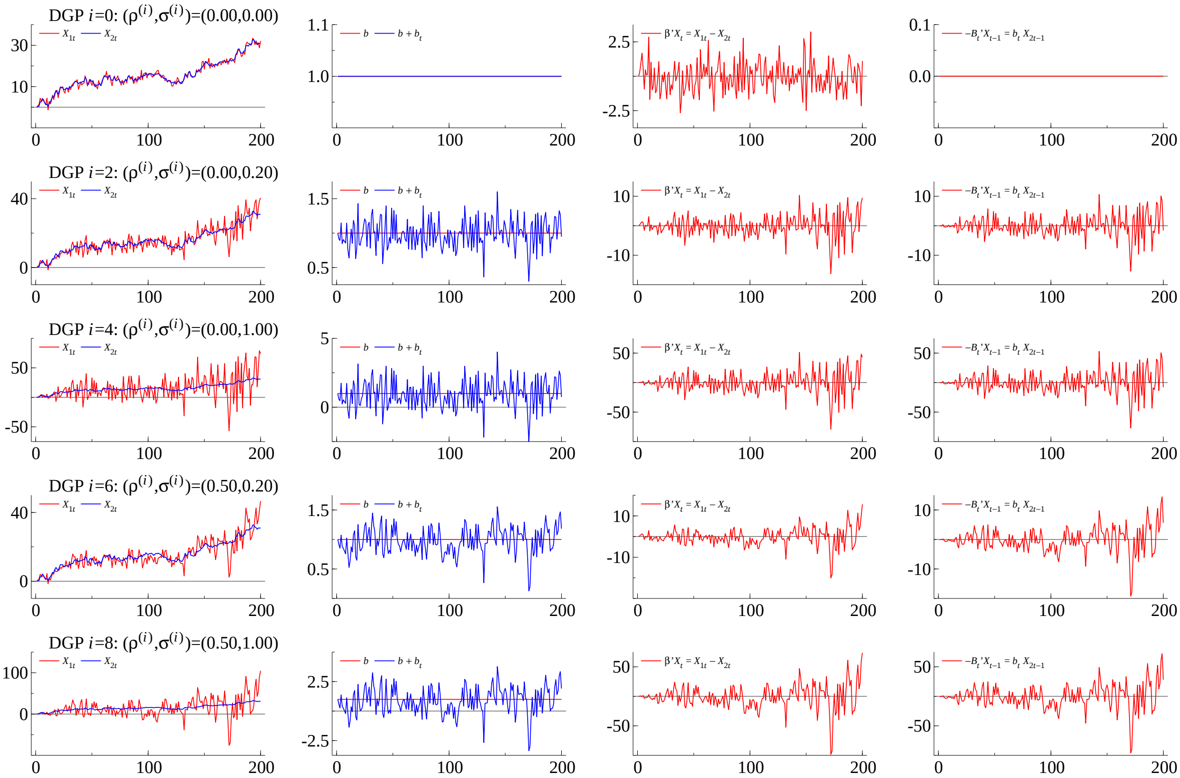

Figure 3.1: Illustration of the simulated data, part 1. The graph shows the simulated series for different data-generating processes i. The first column shows the levels of $X_{1t}^{(i)}$ (red line) and X_{2t} (blue line). The second column shows the stochastic parameter $b + b_t^{(i)}$ (red line) and the constant $b = 1$ (blue line), so the difference between the two equals $b_t^{(i)}$. The third column shows the linear relation $\beta' X_t^{(i)} = X_{1t}^{(i)} - X_{2t} = \varepsilon_{1t} - \varepsilon_{2t} + b_t^{(i)} X_{2t-1}$, and the final column shows the term $-B_t^{(i)\prime} X_{t-1}^{(i)} = b_t^{(i)} X_{2t-1}$.

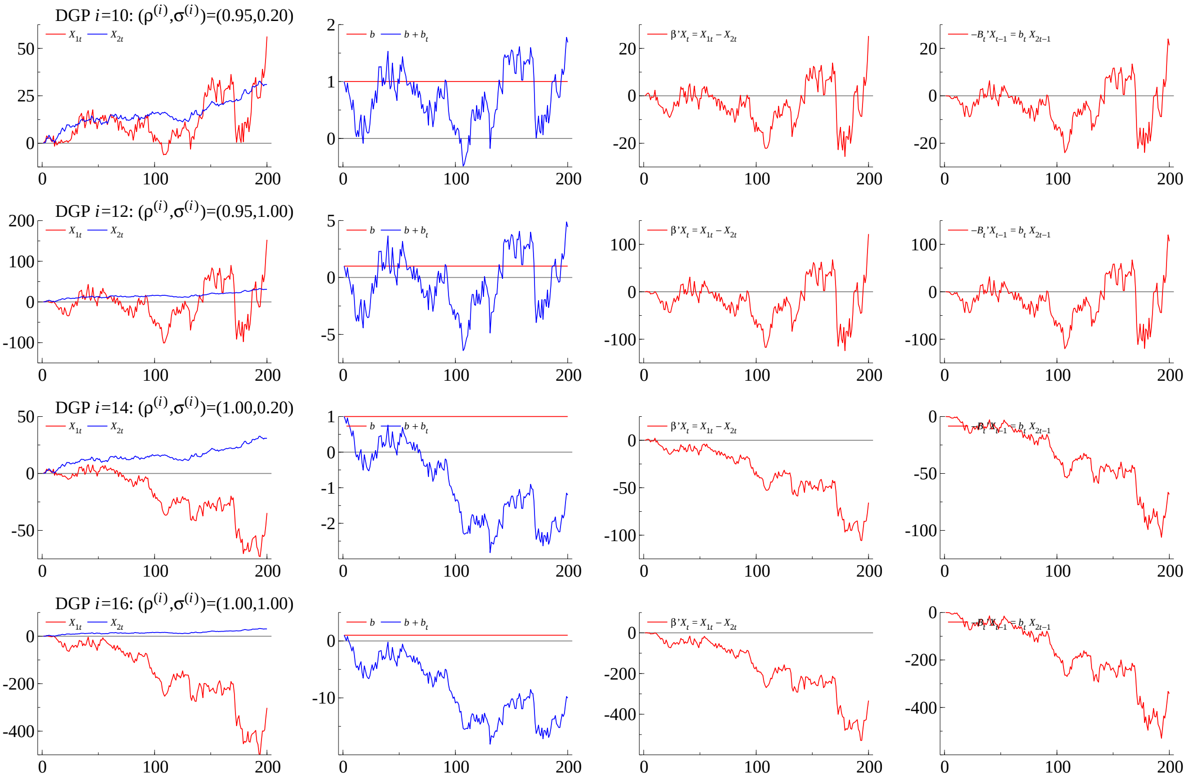

Figure 3.2: Illustration of the simulated data, part 2. The graph shows the simulated series for different data-generating processes i. The first column shows the levels of $X_{1t}^{(i)}$ (red line) and X_{2t} (blue line). The second column shows the stochastic parameter $b + b_t^{(i)}$ (red line) and the constant $b = 1$ (blue line), so the difference between the two equals $b_t^{(i)}$. The third column shows the linear relation $\beta' X_t^{(i)} = X_{1t}^{(i)} - X_{2t} = \varepsilon_{1t} - \varepsilon_{2t} + b_t^{(i)} X_{2t-1}$, and the final column shows the term $-B_t^{(i)\prime} X_{t-1}^{(i)} = b_t^{(i)} X_{2t-1}$.

line). The graphs in the two last columns show the linear combinations $\beta' X_t^{(i)} = X_{1t}^{(i)} - X_{2t} = \varepsilon_{1t} - \varepsilon_{2t} + b_t^{(i)} X_{2t-1}$ and $-B_t^{(i)\prime} X_{t-1} = b_t^{(i)} X_{2t-1}$, respectively.

In the cases of cointegration with $b_t^{(i)} \neq 0$ (i.e. for $i = 1, ..., 12$) it can be seen that $X_{1t}^{(i)}$ fluctuates around X_{2t}. While these fluctuations can be heteroskedastic as well as persistent, it is clear that $X_{1t}^{(i)}$ is mean-reverting to the level of X_{2t}, so that the deviation between the levels of the two variables is bounded. Hence, the linear combination $\beta' X_t^{(i)} = X_{1t}^{(i)} - X_{2t}$ is stationary, though heteroskedastic and potentially persistent. The graphs clearly show that an increase in $\sigma^{(i)}$ increases the heteroskedasticity in $\beta' X_t^{(i)}$ and $-B_t^{(i)\prime} X_{t-1}^{(i)}$ as the volatility of $b_t^{(i)}$ increases. Thereby, the heteroskedasticity in the fluctuations of $X_{1t}^{(i)}$ around X_{2t} increases. Moreover, an increase in $\rho^{(i)}$ for $0 < \rho^{(i)} < 1$ adds persistence in the stochastic parameter $b_t^{(i)}$. As a consequence, $-B_t^{(i)\prime} X_{t-1}^{(i)}$ becomes persistent and thereby $\beta' X_t^{(i)}$ becomes persistent as $X_{1t}^{(i)}$ fluctuates heteroskedastically and persistently around X_{2t}.

For non-stationary $b_t^{(i)}$ (for $i = 13, ..., 16$ as $\rho^{(i)} = 1$) it can clearly be seen from the graphs in the last row of Figure 3.2 that the variables $X_{1t}^{(i)}$ and X_{2t} are not cointegrated. Due to non-stationarity of $b_t^{(i)}$ the linear relations $\beta' X_{t-1}^{(i)}$ and $-B_t^{(i)\prime} X_{t-1}^{(i)}$ are not stochastically trendless and the system is not cointegrated, as clearly evident from the graphs.

3.4 Results

The results from the cointegrated VAR estimations are divided into two parts. First, section 3.4.1 presents the specification of the lag length and tests for misspecification of the estimated residuals for the different data-generating processes. Second, Section 3.4.2 presents the tests for reduced rank and the estimates of the reduced rank model with $r = 1$.

Tables 3.1-3.9 present the average results over the $S = 10,000$ replications for each of the $N = 17$ data-generating processes $i = 0, 1, 2, ..., N-1$ and for time-series of lengths $T = 100, 200, 400, 1000$. All inference presented is based on standard asymptotic inference of the cointegrated VAR model.

Figures 3.3-3.8 present the estimated kernel densities of the simulated distributions of the reduced rank tests for $r = 0$ and $r = 1$, the estimated modulus of the largest unrestricted characteristic root of the companion matrix for the estimated reduced rank model with $r = 1$, and, finally, the $\hat{\beta}$ and $\hat{\alpha}$ coefficients.

3.4.1 Lag Length Specification and Misspecification Tests

This section first presents the lag length specification and misspecification tests for no autocorrelation in the estimated residuals, and thereafter results from residual misspecification tests for normality, skewness and kurtosis, and finally residual misspecification tests for no ARCH.

Lag Length Specification and No Autocorrelation

Table 3.1 shows the average selected lag lengths for each of the data-generating processes as well as the fraction of estimated models with a lag length of $k = 1, 2, 3$ and $k \geq 4$, respectively. Table 3.2 shows the multivariate tests of no first-order autocorrelation (Godfrey, 1988), which was used to select the lag length as described above, as well as the combined first- and second-order test

Table 3.1: Selected Lag-Lengths in the Unrestricted Model

i	$\rho^{(i)}$	$\sigma^{(i)}$	T	$Av(k)$[a]	$k=1$[b]	$k=2$[b]	$k=3$[b]	$k\geq 4$[b]
0	0.00	0.00	100	1.06	0.94	0.05	0.00	0.00
			200	1.05	0.95	0.05	0.00	0.00
			400	1.05	0.95	0.05	0.00	0.00
			1000	1.05	0.95	0.05	0.00	0.00
1	0.00	0.10	100	1.07	0.94	0.06	0.00	0.00
			200	1.07	0.94	0.06	0.00	0.00
			400	1.06	0.94	0.06	0.00	0.00
			1000	1.08	0.93	0.07	0.00	0.00
2	0.00	0.20	100	1.07	0.93	0.07	0.00	0.00
			200	1.09	0.92	0.07	0.00	0.00
			400	1.09	0.92	0.08	0.01	0.00
			1000	1.11	0.90	0.09	0.01	0.00
3	0.00	0.50	100	1.09	0.91	0.08	0.01	0.00
			200	1.11	0.90	0.09	0.01	0.00
			400	1.12	0.90	0.09	0.01	0.00
			1000	1.13	0.89	0.10	0.01	0.00
4	0.00	1.00	100	1.11	0.90	0.09	0.01	0.00
			200	1.12	0.89	0.10	0.01	0.00
			400	1.13	0.89	0.10	0.01	0.00
			1000	1.13	0.88	0.10	0.01	0.00
5	0.50	0.10	100	1.06	0.94	0.06	0.00	0.00
			200	1.08	0.93	0.07	0.00	0.00
			400	1.12	0.89	0.11	0.01	0.00
			1000	1.22	0.79	0.19	0.01	0.00
6	0.50	0.20	100	1.07	0.93	0.07	0.00	0.00
			200	1.09	0.91	0.08	0.00	0.00
			400	1.12	0.89	0.10	0.01	0.00
			1000	1.15	0.87	0.12	0.01	0.00
7	0.50	0.50	100	1.09	0.92	0.08	0.00	0.00
			200	1.10	0.90	0.09	0.01	0.00
			400	1.12	0.89	0.10	0.01	0.00
			1000	1.13	0.88	0.11	0.01	0.00
8	0.50	1.00	100	1.10	0.91	0.09	0.01	0.00
			200	1.12	0.89	0.10	0.01	0.00
			400	1.13	0.88	0.11	0.01	0.00
			1000	1.14	0.88	0.11	0.01	0.00
9	0.95	0.10	100	1.34	0.70	0.28	0.03	0.00
			200	1.86	0.28	0.58	0.12	0.01
			400	2.15	0.15	0.59	0.23	0.03
			1000	2.14	0.17	0.57	0.22	0.04
10	0.95	0.20	100	1.25	0.77	0.21	0.02	0.00
			200	1.47	0.57	0.39	0.04	0.00
			400	1.54	0.52	0.43	0.05	0.00
			1000	1.49	0.56	0.40	0.04	0.00
11	0.95	0.50	100	1.10	0.91	0.09	0.01	0.00
			200	1.12	0.89	0.10	0.01	0.00
			400	1.14	0.87	0.12	0.01	0.00
			1000	1.14	0.87	0.12	0.01	0.00
12	0.95	1.00	100	1.10	0.91	0.08	0.01	0.00
			200	1.10	0.91	0.09	0.01	0.00
			400	1.12	0.89	0.10	0.01	0.00
			1000	1.12	0.89	0.10	0.01	0.00
13	1.00	0.10	100	1.49	0.56	0.38	0.05	0.00
			200	2.00	0.19	0.64	0.16	0.02
			400	2.10	0.14	0.64	0.19	0.02
			1000	1.97	0.20	0.65	0.14	0.01
14	1.00	0.20	100	1.28	0.74	0.24	0.02	0.00
			200	1.43	0.61	0.36	0.03	0.00
			400	1.43	0.61	0.35	0.03	0.00
			1000	1.34	0.69	0.28	0.03	0.00
15	1.00	0.50	100	1.10	0.91	0.08	0.01	0.00
			200	1.11	0.90	0.09	0.01	0.00
			400	1.12	0.89	0.10	0.01	0.00
			1000	1.11	0.90	0.10	0.01	0.00
16	1.00	1.00	100	1.10	0.91	0.08	0.01	0.00
			200	1.10	0.90	0.09	0.01	0.00
			400	1.10	0.91	0.08	0.01	0.00
			1000	1.11	0.90	0.09	0.01	0.00

All reported values are averages over $S = 10,000$ replications.
[a] Average lag-length k.
[b] Percentage with lag length $k = 1, 2, 3$ and $k \geq 4$, respectively.

for no multivariate autocorrelation and the univariate test for no first-order autocorrelation in the estimated residuals in each of the two equations.

Importantly, the results show that for all cases considered the lag length can be set so that no multivariate first-order autocorrelation in the estimated residuals cannot be rejected with p-values of roughly 40 to 50 percent. This holds both with respect to univariate and multivariate tests for no first-order autocorrelation and the multivariate test for no combined first and second order autocorrelation. More importantly, it holds for all sample lengths considered. The p-values for the univariate tests for no first-order autocorrelation are found to be approximately 60 to 70 percent, and for the multivariate test for no combined first and second order autocorrelation the p-values are found to be approximately 80 percent.

Generally, the lag length necessary to not reject no autocorrelation in the estimated residuals increases with the persistence in the stochastic parameter $b_t^{(i)}$. This suggests that potential autocorrelation and persistence in the data caused by the term $b_t^{(i)} X_{2t-1}$ can be captured through the inclusion of lagged first-differences in the short-run structure of the cointegrated VAR model, so that the estimated residuals are not found to be autocorrelated. This holds even when $\rho^{(i)} = 1$, so that $X_t^{(i)}$ is not a cointegrated process. This is an important finding as it is well-known that standard asymptotic inference in the cointegrated VAR model, based on the assumption of independent residuals, is very sensitive to residual autocorrelation (Juselius, 2006, p. 74).

In the constant parameter case ($i = 0$) the average selected lag length is just above $k = 1$, with only around 5 percent of the estimated models having a lag length above one, which reflects the 5 percent significance level used for the tests for no autocorrelation in determining the lag length. The univariate tests for no autocorrelation in $\hat{\varepsilon}_{1t}$ and $\hat{\varepsilon}_{2t}$ are, on average, not rejected with average test sizes of roughly $\chi^2(4) = 1.9$ and corresponding average p-values of 0.7 in both cases.

With stochastic $b_t^{(i)}$ in the data-generating process ($i = 1, \ldots, 16$) the selected lag length increases with the autoregressive parameter $\rho^{(i)}$ and with the sample length. Increasing $\sigma^{(i)}$ does not have a large effect on the selected lag length, though the lag length increases slightly with $\sigma^{(i)}$ for $\rho^{(i)} = (0.0; 0.5)$ and decreases slightly with $\sigma^{(i)}$ for $\rho^{(i)} = (0.95; 1.0)$.

As $\rho^{(i)}$ increases, the average test size of the multivariate test for no first-order autocorrelation decreases (even as the selected lag length on average increases), though the average p-values remain above 39 percent. Likewise the average univariate test for no first-order autocorrelation in $\hat{\varepsilon}_{1t}$ decreases with $\rho^{(i)}$, though the average p-values are above 59 percent. The test for no first-order univariate autocorrelation in $\hat{\varepsilon}_{2t}$ remains unaffected, reflecting that X_{2t} is simulated as a weakly exogenous random walk.

Normality, Skewness, and Excess Kurtosis

Table 3.3 shows the average results of multivariate and univariate test for normality of the estimated residuals (Doornik and Hansen, 1994), while Table 3.4 shows the skewness, kurtosis, and standard deviations of the estimated residuals.

In the case with constant parameters in the data-generating process ($i = 0$) the null of normality of the estimated residuals can clearly not be rejected for both the multivariate and univariate tests and for all sample lengths. Moreover, there are no signs of skewness or excess kurtosis.

Table 3.2: Misspecification Tests Part 1: Autocorrelation

i	$\rho^{(i)}$	$\sigma^{(i)}$	T	Vector test no autocorr.(1)[a] $\chi^2(4)$	$p-val$	Vector test no autocorr.(1-2)[a] $\chi^2(8)$	$p-val$	Univ. test no autocorr.(1) in $\hat{\varepsilon}_{1t}$[a] $\chi^2(4)$	$p-val$	Univ. test no autocorr.(1) in $\hat{\varepsilon}_{2t}$[a] $\chi^2(4)$	$p-val$
0	0.00	0.00	100	3.61	0.52	7.74	0.82	1.89	0.71	1.94	0.70
			200	3.64	0.52	7.64	0.83	1.86	0.72	1.90	0.71
			400	3.58	0.53	7.59	0.83	1.84	0.71	1.89	0.71
			1000	3.58	0.53	7.63	0.83	1.84	0.71	1.89	0.71
1	0.00	0.10	100	3.60	0.52	7.61	0.82	1.91	0.70	1.93	0.70
			200	3.65	0.52	7.58	0.82	1.90	0.71	1.89	0.71
			400	3.72	0.51	7.66	0.81	1.93	0.69	1.88	0.71
			1000	3.82	0.50	7.77	0.80	2.03	0.68	1.88	0.71
2	0.00	0.20	100	3.68	0.51	7.68	0.81	1.98	0.69	1.92	0.70
			200	3.76	0.51	7.70	0.81	2.00	0.69	1.89	0.71
			400	3.82	0.50	7.89	0.80	2.08	0.67	1.87	0.71
			1000	3.93	0.49	8.04	0.79	2.18	0.65	1.88	0.71
3	0.00	0.50	100	3.80	0.50	7.91	0.80	2.18	0.66	1.93	0.70
			200	3.86	0.49	7.98	0.79	2.21	0.65	1.88	0.71
			400	3.90	0.49	8.10	0.79	2.26	0.64	1.87	0.71
			1000	3.96	0.48	8.19	0.79	2.34	0.63	1.86	0.71
4	0.00	1.00	100	3.84	0.50	8.02	0.79	2.31	0.65	1.93	0.70
			200	3.89	0.49	8.09	0.79	2.33	0.63	1.88	0.71
			400	3.91	0.49	8.14	0.79	2.36	0.63	1.87	0.71
			1000	3.97	0.48	8.23	0.79	2.41	0.62	1.86	0.71
5	0.50	0.10	100	3.62	0.52	7.66	0.82	1.88	0.71	1.93	0.70
			200	3.81	0.50	7.94	0.81	1.95	0.69	1.89	0.71
			400	4.01	0.47	8.33	0.79	2.08	0.67	1.86	0.71
			1000	4.23	0.45	8.68	0.77	2.31	0.64	1.83	0.71
6	0.50	0.20	100	3.69	0.51	7.86	0.82	1.96	0.70	1.93	0.70
			200	3.83	0.50	8.22	0.80	2.08	0.68	1.88	0.71
			400	3.93	0.48	8.52	0.79	2.27	0.66	1.86	0.71
			1000	4.00	0.48	8.69	0.79	2.52	0.64	1.85	0.71
7	0.50	0.50	100	3.83	0.50	8.35	0.80	2.27	0.67	1.93	0.70
			200	3.86	0.49	8.57	0.79	2.48	0.65	1.88	0.71
			400	3.94	0.48	8.75	0.78	2.76	0.64	1.85	0.71
			1000	3.96	0.48	8.88	0.78	3.02	0.62	1.86	0.71
8	0.50	1.00	100	3.90	0.49	8.57	0.79	2.57	0.65	1.92	0.70
			200	3.88	0.49	8.71	0.79	2.82	0.64	1.88	0.71
			400	3.96	0.48	8.85	0.78	3.09	0.62	1.85	0.71
			1000	3.97	0.48	8.94	0.78	3.25	0.62	1.86	0.71
9	0.95	0.10	100	4.66	0.40	9.23	0.79	2.36	0.67	1.87	0.70
			200	4.84	0.39	9.53	0.74	2.76	0.64	1.79	0.72
			400	4.62	0.41	9.14	0.71	2.94	0.60	1.83	0.72
			1000	4.41	0.43	8.82	0.74	2.95	0.59	1.84	0.71
10	0.95	0.20	100	4.48	0.42	8.81	0.80	2.52	0.67	1.88	0.70
			200	4.51	0.42	8.87	0.79	2.79	0.66	1.83	0.71
			400	4.35	0.44	8.76	0.78	2.92	0.64	1.84	0.71
			1000	4.23	0.45	8.76	0.77	3.04	0.62	1.84	0.71
11	0.95	0.50	100	3.87	0.49	8.31	0.79	2.53	0.65	1.93	0.70
			200	3.95	0.48	8.53	0.77	2.81	0.63	1.87	0.71
			400	3.99	0.48	8.61	0.77	2.97	0.62	1.86	0.71
			1000	4.00	0.48	8.82	0.76	3.19	0.60	1.85	0.71
12	0.95	1.00	100	3.76	0.51	8.29	0.78	2.60	0.64	1.92	0.70
			200	3.84	0.50	8.53	0.77	2.86	0.62	1.87	0.71
			400	3.90	0.49	8.61	0.76	3.02	0.61	1.86	0.71
			1000	3.95	0.48	8.84	0.75	3.23	0.59	1.85	0.71
13	1.00	0.10	100	4.81	0.39	9.57	0.78	2.70	0.65	1.84	0.71
			200	4.76	0.39	9.38	0.73	2.96	0.60	1.77	0.72
			400	4.56	0.42	9.00	0.74	2.87	0.60	1.82	0.72
			1000	4.32	0.44	8.67	0.77	2.69	0.62	1.79	0.72
14	1.00	0.20	100	4.57	0.41	8.94	0.80	2.73	0.67	1.89	0.70
			200	4.51	0.42	8.88	0.80	2.86	0.65	1.82	0.71
			400	4.34	0.44	8.78	0.79	2.83	0.64	1.84	0.71
			1000	4.23	0.45	8.71	0.79	2.81	0.64	1.82	0.71
15	1.00	0.50	100	3.81	0.50	8.26	0.79	2.60	0.65	1.94	0.70
			200	3.83	0.50	8.32	0.78	2.75	0.63	1.88	0.71
			400	3.79	0.50	8.29	0.78	2.78	0.63	1.85	0.71
			1000	3.78	0.50	8.37	0.77	2.81	0.62	1.84	0.71
16	1.00	1.00	100	3.68	0.52	8.21	0.78	2.68	0.63	1.92	0.70
			200	3.72	0.51	8.27	0.78	2.77	0.63	1.86	0.71
			400	3.75	0.51	8.30	0.78	2.81	0.62	1.86	0.71
			1000	3.75	0.51	8.35	0.77	2.83	0.62	1.83	0.71

All reported values are averages over $S = 10,000$ replications.

[a] Multivariate and univariate tests for no autocorrelation of order 1 or order $1-2$, respectively, in the estimated residuals, see Godfrey (1988).

Table 3.3: Misspecification Tests Part 2: Normality

i	$\rho^{(i)}$	$\sigma^{(i)}$	T	Vector test normality[a] $\chi^2(4)$	$p-val$	Univar. test normality of $\hat{\varepsilon}_{1t}$[b] $\chi^2(2)$	$p-val$	Univar. test normality of $\hat{\varepsilon}_{2t}$[b] $\chi^2(2)$	$p-val$
0	0.00	0.00	100	3.82	0.53	1.92	0.52	1.90	0.52
			200	3.91	0.52	1.94	0.51	1.93	0.52
			400	3.96	0.51	1.98	0.51	1.97	0.51
			1000	3.95	0.51	2.00	0.50	1.98	0.51
1	0.00	0.10	100	4.52	0.48	2.28	0.49	1.91	0.52
			200	6.78	0.39	3.48	0.42	1.93	0.52
			400	16.80	0.21	9.32	0.26	1.97	0.51
			1000	75.34	0.03	48.86	0.06	1.98	0.51
2	0.00	0.20	100	7.03	0.37	3.98	0.38	1.92	0.52
			200	15.39	0.19	9.61	0.22	1.93	0.52
			400	43.07	0.04	30.26	0.06	1.97	0.51
			1000	157.72	0.00	126.98	0.00	1.98	0.51
3	0.00	0.50	100	14.47	0.16	10.52	0.16	1.91	0.52
			200	33.18	0.05	26.87	0.05	1.94	0.52
			400	79.14	0.01	69.68	0.01	1.96	0.51
			1000	226.75	0.00	213.54	0.00	1.98	0.51
4	0.00	1.00	100	19.40	0.10	16.24	0.08	1.91	0.52
			200	41.57	0.03	37.57	0.02	1.94	0.52
			400	90.95	0.01	86.25	0.00	1.96	0.51
			1000	242.79	0.00	237.51	0.00	1.98	0.51
5	0.50	0.10	100	4.51	0.48	2.31	0.48	1.91	0.52
			200	6.65	0.39	3.45	0.42	1.93	0.52
			400	15.93	0.21	8.95	0.26	1.97	0.51
			1000	70.00	0.03	46.04	0.05	1.98	0.51
6	0.50	0.20	100	6.83	0.37	3.92	0.38	1.92	0.52
			200	14.57	0.20	9.24	0.22	1.93	0.52
			400	40.35	0.05	28.71	0.07	1.97	0.51
			1000	149.18	0.00	121.14	0.00	1.99	0.50
7	0.50	0.50	100	13.77	0.17	10.11	0.17	1.91	0.52
			200	31.54	0.05	25.74	0.05	1.94	0.52
			400	76.18	0.01	67.35	0.01	1.96	0.51
			1000	221.75	0.00	209.12	0.00	1.99	0.50
8	0.50	1.00	100	18.58	0.11	15.59	0.09	1.91	0.52
			200	40.12	0.03	36.31	0.03	1.94	0.52
			400	88.97	0.01	84.44	0.00	1.97	0.51
			1000	239.98	0.00	234.80	0.00	1.99	0.50
9	0.95	0.10	100	4.52	0.48	2.48	0.46	1.91	0.52
			200	6.10	0.40	3.54	0.40	1.94	0.52
			400	12.82	0.24	8.05	0.25	1.98	0.51
			1000	54.53	0.04	38.70	0.05	1.98	0.51
10	0.95	0.20	100	6.27	0.38	3.99	0.36	1.91	0.52
			200	11.72	0.23	8.36	0.22	1.94	0.52
			400	31.04	0.07	24.11	0.07	1.97	0.51
			1000	121.11	0.00	102.83	0.00	1.98	0.51
11	0.95	0.50	100	11.26	0.22	8.76	0.19	1.91	0.52
			200	24.09	0.08	20.59	0.07	1.93	0.52
			400	59.62	0.01	54.13	0.01	1.97	0.51
			1000	189.66	0.00	180.62	0.00	1.98	0.51
12	0.95	1.00	100	14.48	0.15	11.95	0.12	1.91	0.52
			200	30.12	0.05	27.08	0.04	1.93	0.52
			400	70.10	0.01	66.29	0.01	1.96	0.51
			1000	208.13	0.00	203.55	0.00	1.98	0.51
13	1.00	0.10	100	4.73	0.46	2.66	0.45	1.92	0.52
			200	6.93	0.36	4.42	0.36	1.94	0.52
			400	15.61	0.18	11.79	0.18	1.97	0.51
			1000	61.11	0.02	50.72	0.02	1.98	0.51
14	1.00	0.20	100	6.91	0.35	4.56	0.34	1.91	0.52
			200	13.65	0.20	10.21	0.19	1.94	0.52
			400	33.62	0.05	27.05	0.05	1.97	0.51
			1000	108.82	0.00	89.41	0.00	1.98	0.51
15	1.00	0.50	100	12.39	0.19	9.03	0.18	1.91	0.52
			200	25.20	0.08	19.55	0.07	1.94	0.52
			400	53.63	0.02	42.76	0.02	1.96	0.51
			1000	141.15	0.00	113.69	0.00	1.98	0.51
16	1.00	1.00	100	15.26	0.14	11.24	0.13	1.92	0.52
			200	29.35	0.06	22.62	0.06	1.94	0.52
			400	59.00	0.01	46.77	0.01	1.96	0.51
			1000	147.58	0.00	118.31	0.00	1.98	0.51

All reported values are averages over $S = 10,000$ replications.
[a] Multivariate test for normality of the estimated residuals, see Doornik and Hansen (1994).
[b] Univariate tests for normality of the estimated residuals.

Table 3.4: Misspecification Tests Part 3: Skewness, Kurtosis, and Standard Deviation

				Skewness[a]		Kurtosis[a]		Std.dev.[a]	
i	$\rho^{(i)}$	$\sigma^{(i)}$	T	$\hat{\varepsilon}_{1t}$	$\hat{\varepsilon}_{2t}$	$\hat{\varepsilon}_{1t}$	$\hat{\varepsilon}_{2t}$	$\hat{\sigma}_1$	$\hat{\sigma}_2$
0	0.00	0.00	100	-0.00	-0.00	2.94	2.94	1.38	0.97
			200	-0.00	-0.00	2.97	2.97	1.40	0.99
			400	-0.00	-0.00	2.99	2.99	1.41	0.99
			1000	0.00	0.00	3.00	2.99	1.41	1.00
1	0.00	0.10	100	0.00	-0.00	3.04	2.94	1.54	0.97
			200	-0.00	-0.00	3.21	2.97	1.69	0.99
			400	0.00	-0.00	3.51	2.99	1.93	0.99
			1000	-0.00	0.00	4.08	2.99	2.49	1.00
2	0.00	0.20	100	0.00	-0.00	3.38	2.94	1.90	0.97
			200	-0.00	-0.00	3.81	2.97	2.30	0.99
			400	0.00	-0.00	4.39	2.99	2.92	0.99
			1000	-0.00	0.00	5.14	2.99	4.23	1.00
3	0.00	0.50	100	0.00	-0.00	4.38	2.94	3.42	0.97
			200	-0.00	-0.00	5.02	2.97	4.63	0.98
			400	0.00	-0.00	5.61	2.99	6.41	0.99
			1000	-0.00	0.00	6.08	2.99	9.95	1.00
4	0.00	1.00	100	-0.00	-0.00	5.08	2.94	6.31	0.97
			200	-0.00	-0.00	5.63	2.97	8.86	0.98
			400	0.00	-0.00	6.05	2.99	12.53	0.99
			1000	-0.00	0.00	6.31	2.99	19.71	1.00
5	0.50	0.10	100	0.00	-0.00	3.04	2.94	1.56	0.97
			200	-0.00	-0.00	3.21	2.97	1.72	0.99
			400	-0.00	-0.00	3.50	2.99	1.97	0.99
			1000	-0.00	0.00	4.04	2.99	2.52	1.00
6	0.50	0.20	100	0.00	-0.00	3.38	2.94	1.94	0.97
			200	-0.00	-0.00	3.79	2.97	2.34	0.99
			400	0.00	-0.00	4.34	2.99	2.96	0.99
			1000	-0.00	0.00	5.07	2.99	4.26	1.00
7	0.50	0.50	100	0.00	-0.00	4.33	2.94	3.46	0.97
			200	-0.00	-0.00	4.95	2.97	4.67	0.98
			400	0.00	-0.00	5.54	2.99	6.44	0.99
			1000	-0.00	0.00	6.03	2.99	9.97	1.00
8	0.50	1.00	100	-0.00	-0.00	5.01	2.94	6.36	0.97
			200	-0.00	-0.00	5.56	2.97	8.89	0.98
			400	0.00	-0.00	6.00	2.99	12.56	0.99
			1000	-0.00	0.00	6.29	2.99	19.73	1.00
9	0.95	0.10	100	0.00	-0.00	3.08	2.94	1.66	0.97
			200	-0.00	-0.00	3.24	2.97	1.84	0.98
			400	-0.00	-0.00	3.48	2.99	2.09	0.99
			1000	0.00	0.00	3.94	2.99	2.64	1.00
10	0.95	0.20	100	0.00	-0.00	3.40	2.94	2.10	0.97
			200	-0.01	-0.00	3.74	2.97	2.51	0.98
			400	-0.00	-0.00	4.19	2.99	3.12	0.99
			1000	0.00	0.00	4.85	2.99	4.39	1.00
11	0.95	0.50	100	0.00	-0.00	4.17	2.94	3.80	0.97
			200	-0.00	-0.00	4.65	2.97	4.98	0.98
			400	-0.00	-0.00	5.17	2.99	6.70	0.99
			1000	0.00	0.00	5.74	2.99	10.15	1.00
12	0.95	1.00	100	0.00	-0.00	4.59	2.94	7.02	0.97
			200	-0.00	-0.00	5.06	2.97	9.46	0.98
			400	-0.00	-0.00	5.52	2.99	13.01	0.99
			1000	0.00	0.00	5.98	2.99	20.03	1.00
13	1.00	0.10	100	0.00	-0.00	3.13	2.94	1.76	0.96
			200	-0.00	-0.00	3.35	2.97	2.05	0.98
			400	-0.00	-0.00	3.69	2.99	2.48	0.99
			1000	0.00	0.00	4.15	2.99	3.39	1.00
14	1.00	0.20	100	0.00	-0.00	3.51	2.94	2.40	0.97
			200	-0.00	-0.00	3.90	2.97	3.08	0.98
			400	-0.00	-0.00	4.31	2.99	4.08	0.99
			1000	0.00	0.00	4.69	2.99	6.12	1.00
15	1.00	0.50	100	0.00	-0.00	4.21	2.94	4.78	0.97
			200	-0.00	-0.00	4.58	2.97	6.69	0.98
			400	-0.00	-0.00	4.84	2.99	9.41	0.99
			1000	0.00	0.00	4.99	2.99	14.76	1.00
16	1.00	1.00	100	0.00	-0.00	4.51	2.94	9.12	0.97
			200	-0.00	-0.00	4.78	2.97	13.04	0.98
			400	-0.00	-0.00	4.96	2.99	18.57	0.99
			1000	0.00	0.00	5.05	2.99	29.35	1.00

All reported values are averages over $S = 10,000$ replications.

[a] The skewness of the estimated residuals is calculated as $skewness_i = T^{-1}\sum_{t=1}^{T}(\hat{\varepsilon}_{it}/\hat{\sigma}_i)^3$ and $kurtosis_i = T^{-1}\sum_{t=1}^{T}(\hat{\varepsilon}_{it}/\hat{\sigma}_i)^4$, where $\hat{\varepsilon}_{it}$ are the estimated system residuals and $\hat{\sigma}_i$ their standard deviations for $i = 1, 2$ as reported in the final column, see Juselius (2006, p. 75).

Under cointegration with stochastic parameters in the data-generating process the tests reveal problems with non-normality of the estimated residuals, but only for long samples and more pronounced for high volatility in the stochastic parameter $b_t^{(i)}$ as determined by $\sigma_i^{(i)}$. In samples with $T = 100$ or $T = 200$ observations, which are common samples length for typical macroeconomic data, the multivariate test for non-normality can, on average, not be rejected as long as the volatility of $b_t^{(i)}$ is small, i.e. when $\sigma^{(i)}$ is small. Though, the test becomes borderline rejected in the extreme cases of both high persistence and high volatility in $b_t^{(i)}$ ($i = 11$ and $i = 12$). For long samples the multivariate test reveals the non-normality of the estimated residuals. The univariate tests reveal that when multivariate normality is rejected it is caused by non-normality of $\hat{\varepsilon}_{1t}$, as expected, while the null of normality cannot be rejected for $\hat{\varepsilon}_{2t}$ in all cases considered. Looking at the skewness and kurtosis it can be seen that normality of $\hat{\varepsilon}_{1t}$ is rejected due to excess kurtosis in $\hat{\varepsilon}_{1t}$, while there are no problems with skewness on average.

The normality tests reveal that an increase in $\sigma^{(i)}$ (for fixed $\rho^{(i)}$) significantly increases the average test sizes and lowers the average p-values of the normality tests, while an increase in $\rho^{(i)}$ does not have a big impact on the normality tests for a fixed $\sigma^{(i)}$. This reflects that as $\sigma^{(i)}$ increases, the volatility of the heteroskedastic term $b_t^{(i)} X_{2t-1}$ in the simulated data increases, causing excess kurtosis and fat tails in the estimated residuals $\hat{\varepsilon}_{1t}$ in the estimated cointegrated VAR model. However, it requires a longer sample than is typically available for macroeconomic data to identify this non-normality in the estimated residuals.

It is worth noting that the average standard deviations of the residuals clearly reveal the non-linear effect of the random shocks to the stochastic parameter $b_t^{(i)}$ on the level of $X_{1t}^{(i)}$. As $\rho^{(i)}$ increases, the standard deviation of $\hat{\varepsilon}_{1t}$ increases dramatically and becomes up to 20 times larger than the standard deviation of $\hat{\varepsilon}_{2t}$ on average in the case of stochastic cointegration, and up to 30 times larger than the standard deviation of $\hat{\varepsilon}_{2t}$ for $\rho^{(i)} = 1.0$.[2]

No ARCH

Table 3.5 presents the results of the univariate and multivariate tests for no ARCH in the estimated residuals (Lötkepohl and Krätzig, 2004). Based on the multivariate test the null of no ARCH cannot be rejected on average, except from the cases with high $\sigma^{(i)}$ in very long samples with $T = 1000$. As was the case for the normality tests, increasing $\sigma^{(i)}$ has some effect on the tests for no ARCH, while increasing $\rho^{(i)}$ has much less effect. As the sample size increases, the test values increase and the p-values decrease. However, in samples with $T = 1000$ observations the null of no multivariate first-order ARCH can only be boarderline rejected, while the univariate test for no ARCH in $\hat{\varepsilon}_{1t}$ can be boarderline rejected in samples with $T = 400$ observations for $\sigma^{(i)} \geq 0.5$. When the null of no multivariate ARCH can be rejected it can be seen from the univariate tests that it is due to ARCH effects in $\hat{\varepsilon}_{1t}$ as expected.

[2] It is also worth noting that in the case of constant parameters in the data-generating process, the average standard deviation of $\hat{\varepsilon}_{1t}$ is well above 1.0. This seems to reflect that the X_{2t} is weakly exogenous while $X_{1t}^{(i)}$ is purely adjusting. Hence, only shocks to X_{2t} cumulate into a common stochastic trend, which is very well estimated in the cointegrated VAR model, while the shocks to $X_{1t}^{(i)}$ do not have a long-run effect and thereby are less precisely estimated. However, by imposing reduced rank and weak exogeneity of X_{2t}, the standard deviation of $\hat{\varepsilon}_{1t}$ eventually becomes 1.0 on average in this case.

Table 3.5: Misspecification Tests Part 4: ARCH

i	$\rho^{(i)}$	$\sigma^{(i)}$	T	Vector test no ARCH(1)[a] $\chi^2(9)$	$p-val$	Vector test no ARCH(1-2)[a] $\chi^2(18)$	$p-val$	Univ. test no ARCH(1) in $\hat{\varepsilon}_{1t}$[a] $\chi^2(1)$	$p-val$	Univ. test no ARCH(1)in $\hat{\varepsilon}_{2t}$[a] $\chi^2(1)$	$p-val$
0	0.00	0.00	100	8.46	0.59	17.60	0.56	0.89	0.51	0.90	0.51
			200	8.69	0.58	17.71	0.57	0.92	0.51	0.97	0.50
			400	8.66	0.59	17.82	0.57	0.97	0.50	0.98	0.50
			1000	8.83	0.57	17.84	0.57	0.99	0.51	0.98	0.50
1	0.00	0.10	100	8.36	0.59	17.40	0.57	0.98	0.50	0.89	0.51
			200	9.50	0.54	19.04	0.52	1.43	0.47	0.97	0.50
			400	13.29	0.40	25.73	0.36	3.33	0.35	0.98	0.51
			1000	30.79	0.12	56.00	0.07	14.92	0.11	0.99	0.50
2	0.00	0.20	100	9.08	0.56	18.76	0.52	1.40	0.47	0.89	0.51
			200	12.11	0.44	23.62	0.39	3.12	0.36	0.97	0.50
			400	19.71	0.24	36.49	0.18	8.44	0.17	0.98	0.51
			1000	43.88	0.04	77.33	0.02	29.39	0.01	0.99	0.50
3	0.00	0.50	100	10.93	0.48	21.90	0.43	2.76	0.38	0.89	0.51
			200	15.66	0.33	29.51	0.27	6.43	0.20	0.97	0.50
			400	24.85	0.16	44.78	0.11	14.71	0.06	0.98	0.51
			1000	50.31	0.03	87.40	0.01	39.42	0.00	0.99	0.50
4	0.00	1.00	100	11.76	0.45	23.23	0.40	3.63	0.33	0.90	0.51
			200	16.74	0.31	31.24	0.24	7.82	0.15	0.97	0.50
			400	25.94	0.15	46.47	0.10	16.43	0.04	0.98	0.51
			1000	51.39	0.03	89.04	0.01	41.37	0.00	0.99	0.50
5	0.50	0.10	100	8.36	0.59	17.36	0.57	0.98	0.51	0.90	0.51
			200	9.54	0.54	18.98	0.52	1.46	0.46	0.97	0.50
			400	13.55	0.39	25.73	0.35	3.40	0.35	0.98	0.51
			1000	30.98	0.11	55.18	0.07	14.96	0.10	0.99	0.50
6	0.50	0.20	100	9.00	0.56	18.54	0.53	1.40	0.47	0.90	0.51
			200	12.05	0.44	23.24	0.40	3.12	0.36	0.97	0.50
			400	19.52	0.24	35.74	0.19	8.36	0.17	0.98	0.51
			1000	43.31	0.04	75.81	0.02	29.06	0.01	0.98	0.50
7	0.50	0.50	100	10.73	0.49	21.41	0.45	2.75	0.38	0.90	0.51
			200	15.49	0.33	28.94	0.28	6.43	0.20	0.97	0.50
			400	24.70	0.16	44.24	0.11	14.67	0.06	0.98	0.51
			1000	50.20	0.03	86.87	0.01	39.35	0.00	0.99	0.50
8	0.50	1.00	100	11.60	0.46	22.79	0.41	3.65	0.32	0.90	0.51
			200	16.68	0.30	30.86	0.25	7.86	0.15	0.97	0.50
			400	26.00	0.15	46.27	0.10	16.52	0.04	0.98	0.51
			1000	51.51	0.03	88.90	0.01	41.49	0.00	0.99	0.50
9	0.95	0.10	100	8.26	0.60	17.10	0.58	1.04	0.50	0.89	0.51
			200	9.15	0.55	18.29	0.54	1.53	0.46	0.96	0.50
			400	12.82	0.41	24.27	0.38	3.43	0.34	0.98	0.50
			1000	28.52	0.13	50.10	0.08	14.21	0.09	0.99	0.50
10	0.95	0.20	100	8.92	0.57	18.08	0.54	1.52	0.47	0.89	0.51
			200	11.57	0.45	22.25	0.42	3.14	0.35	0.97	0.50
			400	18.35	0.26	33.64	0.20	7.82	0.17	0.98	0.50
			1000	40.49	0.05	71.05	0.02	26.96	0.02	0.98	0.50
11	0.95	0.50	100	10.92	0.48	21.17	0.45	2.71	0.38	0.89	0.51
			200	15.66	0.32	28.90	0.26	6.01	0.21	0.97	0.50
			400	24.71	0.15	44.02	0.10	13.48	0.07	0.98	0.51
			1000	49.46	0.03	85.88	0.01	37.21	0.00	0.98	0.50
12	0.95	1.00	100	12.05	0.43	22.84	0.40	3.37	0.34	0.90	0.51
			200	17.30	0.28	31.40	0.23	7.11	0.17	0.97	0.50
			400	26.63	0.13	47.00	0.09	15.17	0.05	0.98	0.51
			1000	51.47	0.02	89.02	0.01	39.54	0.00	0.98	0.50
13	1.00	0.10	100	8.59	0.59	17.72	0.56	1.09	0.50	0.89	0.51
			200	10.66	0.49	21.18	0.46	1.92	0.43	0.96	0.50
			400	17.77	0.26	33.17	0.22	4.68	0.28	0.98	0.50
			1000	47.74	0.03	83.11	0.01	16.95	0.06	0.98	0.50
14	1.00	0.20	100	9.89	0.53	19.80	0.50	1.69	0.45	0.89	0.51
			200	14.69	0.35	27.71	0.30	3.69	0.31	0.97	0.50
			400	26.61	0.13	47.35	0.09	8.49	0.15	0.98	0.51
			1000	66.30	0.01	112.75	0.01	24.34	0.02	0.98	0.50
15	1.00	0.50	100	12.63	0.42	24.00	0.38	2.79	0.38	0.89	0.51
			200	19.85	0.23	35.72	0.18	5.80	0.21	0.97	0.50
			400	34.48	0.08	59.17	0.05	11.59	0.08	0.98	0.51
			1000	77.29	0.01	129.41	0.00	28.25	0.01	0.98	0.50
16	1.00	1.00	100	13.90	0.38	25.81	0.34	3.25	0.34	0.89	0.51
			200	21.42	0.20	38.04	0.16	6.38	0.19	0.97	0.50
			400	36.29	0.07	61.93	0.04	12.26	0.08	0.98	0.51
			1000	79.17	0.01	132.33	0.00	28.91	0.01	0.98	0.50

All reported values are averages over $S = 10,000$ replications.

[a] Multivariate and univariate tests for no ARCH of order 1 or order $1-2$, respectively, in the estimated residuals, see Lötkepohl and Krätzig (2004).

To conclude on the results from the misspecification tests, the results reveal that in samples of $T = 100$ or $T = 200$ observations the estimated unrestricted VAR models appear to be fairly good statistical representations of the simulated data based on the various misspecification tests considered. This holds regardless of the cointegration properties of the underlying data-generating process. However, in cases of both high persistence and volatility in $b_t^{(i)}$ and in very long samples the misspecification can be identified based on the misspecification tests.

The results show that when the parameter $\rho^{(i)}$ increases, the persistence it creates in the simulated data through the term $b_t^{(i)} X_{2t-1}$ can to some degree be captured in the unrestricted VAR model by including lagged first-differences. Hence, the null of no autocorrelation in the estimated residuals cannot be rejected. Moreover, when the parameter $\sigma^{(i)}$ increases, the heteroskedasticity in $b_t^{(i)} X_{2t-1}$ increases the heteroskedasticity in the simulated data, which cause some excess kurtosis and non-normality in the estimated residuals. However, standard asymptotic inference in the cointegrated VAR model is less sensitive to misspecification due to excess kurtosis and non-normality than to skewness and autocorrelation (Juselius, 2006, p. 77), and most likely an econometrician would continue the econometric analysis despite signs of such misspecification—except in cases of both high persistence and high volatility of the stochastic cointegration parameter, i.e. high $\rho^{(i)}$ and high $\sigma^{(i)}$, or with very long samples. Next section considers the tests for reduced rank and estimation of the cointegration relations.

3.4.2 Reduced Rank Tests and Estimates

The main focus of the econometric analysis is on the cointegration properties, in particular to what extent the cointegrated VAR model can be used as an approximation to estimate β when the data-generating process has a stochastic $\beta_t^{(i)} = \beta + B_t^{(i)}$, where $B_t^{(i)}$ is a mean zero stationary autoregressive process. Thus, it is first of interest whether a reduced rank of $r = 1$ is found in the estimated cointegrated VAR model based on the trace test of Johansen (1996); second, whether the estimated cointegration parameters are consistent estimates of $\beta = (1, -1)'$ as used in the data-generating process of the simulated data; and finally, how the misspecification caused by the stochastic cointegration parameters $\beta_t^{(i)} = \beta + B_t^{(i)}$ affects the inference of the trace tests and the reduced rank estimators. Results from tests for reduced rank are considered first, and next the estimates from the reduced rank model with $r = 1$ imposed.

Reduced Rank Tests

The average results from the maximum likelihood test for reduced rank of Johansen (1996), known as the trace test, are presented in Table 3.6 along with p-values and rejection frequencies. The tests for $\mathcal{H}(r)$ against $\mathcal{H}(p)$ are reported for $r = 0$ and $r = 1$, respectively, with p-values based on the Gamma approximation (Doornik, 1998) of the asymptotic distributions derived in Johansen (1996). The quantiles of this approximation are reprinted in Juselius (2006) and Dennis et al. (2006). The asymptotic distribution depends on the number of of unit roots in the system, $p - r$, and the deterministic specification. For $\mathcal{H}(0)$ there are $p - r = 2$ unit roots (and no cointegration) and in the case of an unrestricted constant the 95 percent quantile is 15.41. For $\mathcal{H}(1)$ with $p - r = 1$ unit root (and one cointegration relation) the 95 percent quantile is 3.84. Figures 3.3 and 3.4 show the estimated kernel densities of the estimated trace tests for $\mathcal{H}(0)$ and $\mathcal{H}(1)$ against $\mathcal{H}(p)$, respectively. The 95 percent quantiles are shown as the vertical black lines.

The rank of $\Pi_t^{(i)} = \alpha\beta_t^{(i)}$ in the data-generating process of the simulated data was constant $r = 1$ for all cases with constant or stationary parameter $b_t^{(i)}$ $(i = 0, ..., 12)$. Using a top-down testing procedure as suggested by Johansen (1996) and Juselius (2006)—where $\mathcal{H}(0)$ is first tested against $\mathcal{H}(p)$, and if rejected, $\mathcal{H}(1)$ is tested against $\mathcal{H}(p)$ etc. until the lowest rank which cannot be rejected is found—we would hope that $\mathcal{H}(0)$ is rejected and $\mathcal{H}(1)$ not rejected, so that the trace test suggests a cointegration rank of $r = 1$ for $i = 0, ..., 12$. In the case of non-stationary stochastic parameters $(i = 13, ..., 16)$ we would hope that the trace test rejects cointegration, which corresponds to the null of rank $r = 0$ for $\mathcal{H}(1)$ not being rejected.

In the case of constant parameters in the underlying data-generating process $(i = 0)$ the null of $r = 0$ is clearly rejected in all cases, with a rejection rate of 1.0 suggesting a high power of the test in rejecting the false null.[3] From Figure 3.3 it is clear how increasing the sample skews the test sizes further away from the 95 percent quantile leading to a clearer rejection of the null. The tests for $\mathcal{H}(1)$ show that on average the null of $r = 1$, which is correct, is not rejected with average p-values of 20 percent. However, the results reveal some size distortion as the null of $r = 1$ is falsily rejected approximately 30 percent of the time, independently of the sample length, which is also evident from Figure 3.4. This illustrates the well-known problems with both size and power distortions of the trace test based on standard asymptotic inference, which has been documented in numerous simulation studies, see Juselius (2006, ch. 8) and references therein. Therefore, Juselius (2006) suggests that the rank selection in a cointegrated VAR model must be based on trace test as well as other indices, such as the α-estimates in the unrestricted model (not presented) and the characteristic roots of the companion matrix (presented in Table 3.7).

Consider next the cases of stochastic cointegration in the underlying data-generating process and with no or low persistence in $b_t^{(i)}$ (for $i = 1, ..., 8$). With no persistence in $b_t^{(i)}$ there is almost no effect on the average sizes of the trace tests compared to the case with constant parameters. A rank of $r = 0$ is rejected with a rejection rate of 1.0, while a rank of $r = 1$ cannot on average be rejected with p-values of approximately 20 percent as in the case with constant parameters in the data-generating process. However, the trace test of $r = 1$ is still rejected in almost one third of the cases, which is also clear from Figure 3.4 where approximately a third of the estimated kernel density lies in the rejection area to the right of the critical value. Increasing $\rho^{(i)}$ to 0.5 skews the test size of $\mathcal{H}(0)$ to the left, while the test size for $\mathcal{H}(1)$ is almost not affected as evident from Figures 3.3 and 3.4, respectively. However, the conclusions with respect to the rank remains unchanged; the trace tests clearly reject a reduced rank of $r = 0$ and in most cases a reduced rank of $r = 1$ cannot be rejected as evident from Table 3.6 for all cases with $\rho^{(i)} \leq 0.5$. Thus, the results show that standard asymptotic inference of the trace test is not very sensitive to misspecification caused by stochastic cointegration parameters in the underlying data-generating process as long as there is no or low persistence in the stochastic coitegration parameters.

By contrast, near non-stationary or non-stationary persistence in $b_t^{(i)}$ can be seen to have a big

[3]Here the special case with full adjustment to the cointegration relation in each period is considered as $\alpha = (-1, 0)'$. This might be the reason why the trace test always reject the false null of $r = 0$, which is in contrast to power distortions of the trace test found in many simulation studies.

Table 3.6: Reduced Rank Determination: Rank Test

				Reduced rank tests $\mathcal{H}(r)$ against $\mathcal{H}(p)$[a]					
i	$\rho^{(i)}$	$\sigma^{(i)}$	T	$\mathcal{H}(0)$	$p-val$	$Reject$	$\mathcal{H}(1)$	$p-val$	$Reject$
0	0.00	0.00	100	73.42	0.00	1.00	3.06	0.22	0.30
			200	141.44	0.00	1.00	3.06	0.21	0.30
			400	276.72	0.00	1.00	3.04	0.22	0.30
			1000	685.23	0.00	1.00	3.06	0.22	0.31
1	0.00	0.10	100	73.27	0.00	1.00	3.06	0.22	0.30
			200	140.75	0.00	1.00	3.05	0.21	0.30
			400	276.18	0.00	1.00	3.04	0.22	0.30
			1000	678.39	0.00	1.00	3.06	0.22	0.30
2	0.00	0.20	100	73.36	0.00	1.00	3.06	0.22	0.30
			200	140.14	0.00	1.00	3.06	0.21	0.30
			400	274.02	0.00	1.00	3.04	0.22	0.30
			1000	671.41	0.00	1.00	3.06	0.22	0.30
3	0.00	0.50	100	73.42	0.00	1.00	3.06	0.22	0.30
			200	139.33	0.00	1.00	3.06	0.21	0.30
			400	271.86	0.00	1.00	3.04	0.22	0.30
			1000	666.44	0.00	1.00	3.06	0.22	0.30
4	0.00	1.00	100	73.35	0.00	1.00	3.06	0.22	0.30
			200	139.04	0.00	1.00	3.06	0.21	0.30
			400	271.13	0.00	1.00	3.04	0.22	0.30
			1000	666.01	0.00	1.00	3.06	0.22	0.30
5	0.50	0.10	100	60.09	0.00	1.00	3.06	0.22	0.30
			200	102.48	0.00	1.00	3.06	0.21	0.30
			400	173.39	0.00	1.00	3.05	0.22	0.30
			1000	350.27	0.00	1.00	3.06	0.22	0.31
6	0.50	0.20	100	50.05	0.00	1.00	3.06	0.21	0.30
			200	82.73	0.00	1.00	3.06	0.21	0.30
			400	141.96	0.00	1.00	3.05	0.22	0.30
			1000	311.00	0.00	1.00	3.06	0.22	0.31
7	0.50	0.50	100	40.68	0.00	0.99	3.07	0.21	0.31
			200	69.06	0.00	1.00	3.07	0.21	0.30
			400	125.28	0.00	1.00	3.05	0.22	0.30
			1000	293.04	0.00	1.00	3.06	0.22	0.30
8	0.50	1.00	100	37.99	0.00	0.99	3.07	0.21	0.31
			200	65.90	0.00	1.00	3.07	0.21	0.30
			400	121.96	0.00	1.00	3.05	0.22	0.30
			1000	289.65	0.00	1.00	3.06	0.22	0.30
9	0.95	0.10	100	31.57	0.08	0.75	2.89	0.22	0.28
			200	27.26	0.09	0.68	2.95	0.21	0.29
			400	25.47	0.04	0.82	3.01	0.22	0.31
			1000	37.94	0.00	0.99	3.09	0.21	0.31
10	0.95	0.20	100	20.43	0.14	0.56	2.77	0.23	0.26
			200	19.18	0.12	0.56	2.87	0.22	0.28
			400	22.22	0.05	0.77	3.00	0.22	0.31
			1000	36.51	0.00	0.99	3.09	0.21	0.31
11	0.95	0.50	100	14.53	0.21	0.37	2.63	0.23	0.24
			200	16.02	0.16	0.46	2.83	0.22	0.28
			400	20.60	0.06	0.74	3.00	0.21	0.30
			1000	35.38	0.00	0.99	3.09	0.21	0.31
12	0.95	1.00	100	13.26	0.24	0.30	2.58	0.24	0.23
			200	15.11	0.17	0.42	2.81	0.22	0.28
			400	19.95	0.06	0.72	3.00	0.21	0.31
			1000	34.86	0.00	0.99	3.09	0.21	0.31
13	1.00	0.10	100	23.81	0.15	0.57	2.60	0.25	0.24
			200	15.81	0.24	0.36	2.41	0.26	0.21
			400	12.61	0.29	0.25	2.28	0.27	0.18
			1000	11.56	0.32	0.20	2.23	0.27	0.17
14	1.00	0.20	100	16.01	0.22	0.39	2.40	0.26	0.20
			200	12.94	0.28	0.27	2.30	0.27	0.19
			400	11.85	0.30	0.22	2.23	0.27	0.18
			1000	11.38	0.32	0.19	2.22	0.27	0.17
15	1.00	0.50	100	12.24	0.30	0.24	2.23	0.27	0.17
			200	11.58	0.32	0.21	2.22	0.27	0.17
			400	11.36	0.32	0.19	2.19	0.28	0.17
			1000	11.17	0.33	0.18	2.20	0.27	0.17
16	1.00	1.00	100	11.49	0.32	0.20	2.17	0.28	0.16
			200	11.22	0.33	0.19	2.20	0.27	0.17
			400	11.18	0.33	0.18	2.18	0.28	0.17
			1000	11.10	0.33	0.18	2.19	0.27	0.16

All reported values are averages over $S = 10,000$ replications.

[a] LR-test of rank r against the unrestricted model with full rank p, see Johansen (1996). The first column reports the average test sizes, the second the p-value based on the asymptotic distributions in Doornik (1998), and the third column reports the rejection frequency.

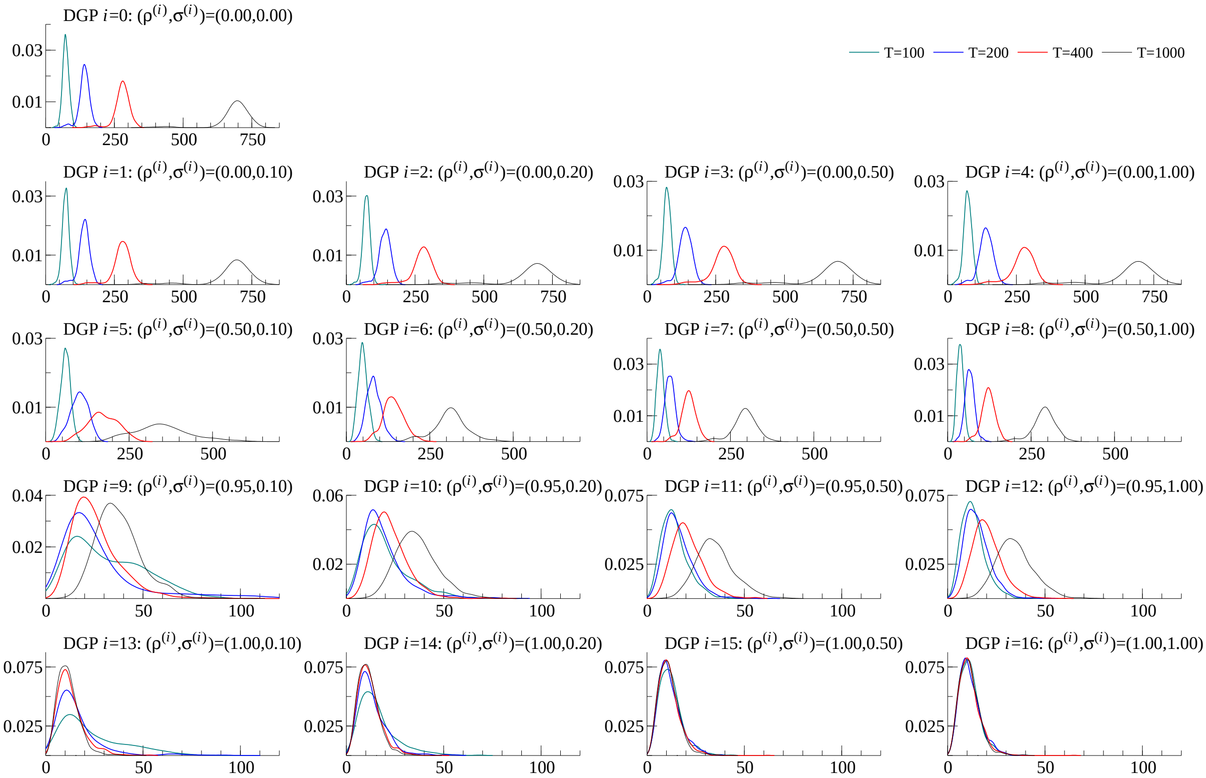

Figure 3.3: Estimated kernel densities of the reduced rank tests $\mathcal{H}(0)$ against $\mathcal{H}(p)$. The solid black vertical lines indicate the 95 quancentile of the the asymptotic distribution, derived in Johansen (1996), based on the Gamma approximation in Doornik (1998).

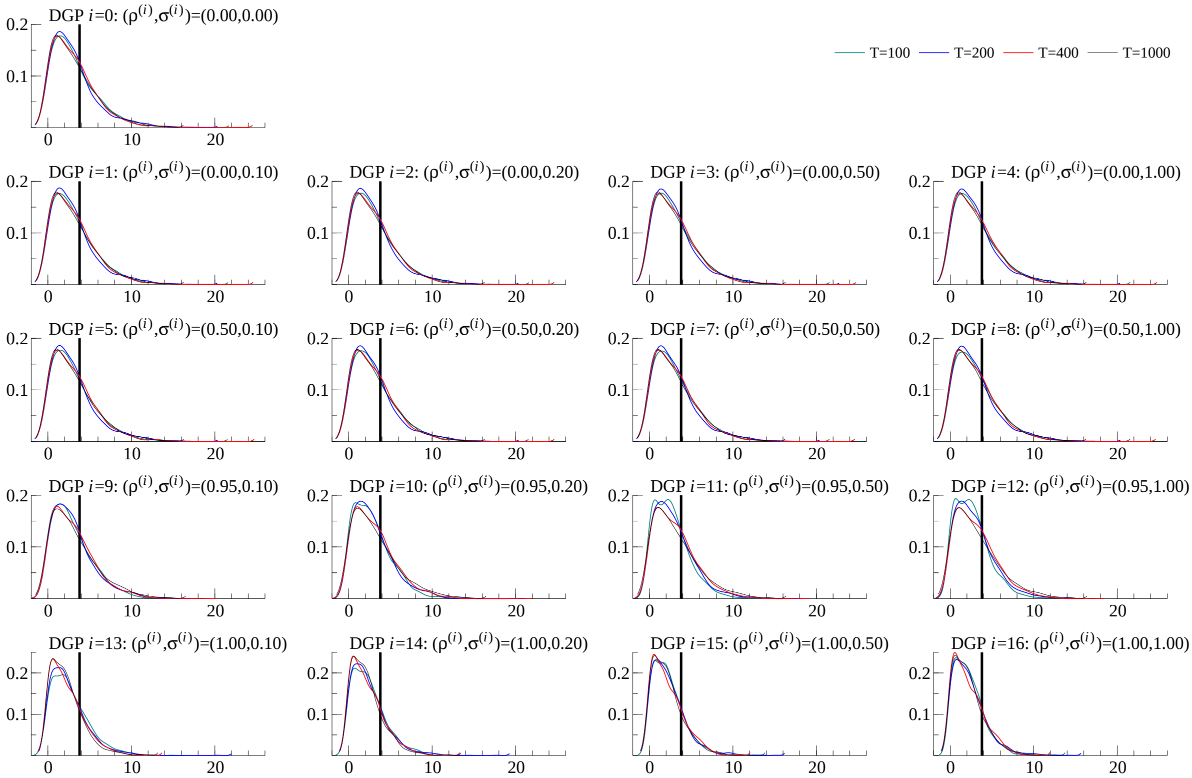

Figure 3.4: Estimated kernel densities of the reduced rank tests $\mathcal{H}(1)$ against $\mathcal{H}(p)$. The solid black vertical lines indicate the 95 quancentile of the the asymptotic distribution, derived in Johansen (1996), based on the Gamma approximation in Doornik (1998).

impact on the trace test for $r = 0$, while the trace test for $r = 1$ is still almost unaffected. In the case of near non-stationary $b_t^{(i)}$ (i.e. $\rho^{(i)} = 0.95$ in $i = 9, ..., 12$) the trace test for $r = 0$ is skewed dramatically towards zero, in particular this is evident for the very long samples with $T = 1000$ compared to the cases with less persistence in $b_t^{(i)}$. As the test size gets skewed towards zero the probability of not rejecting the null of a rank of $r = 0$ increases. This is clearly evident from the estimated kernel densities of the trace test for $r = 0$ in Figure 3.3, and from Table 3.6 it can be seen that the rejection rate of $r = 0$ falls as $\rho^{(i)}$ increases. For example, in samples of $T = 100$ observations the rank $r = 0$ is only rejected in 75 percent of the models for $i = 9$ and 56 percent for $i = 10$, so for a significant part of the simulated data a rank of $r = 0$ cannot be rejected. However, in very long samples with $T = 1000$ observations the rejection rate of $\mathcal{H}(0)$ increases to 99 percent. In the extreme case with both close to non-stationary persistence and high volatility in $b_t^{(i)}$ $(i = 12)$ the null of $r = 0$ is only rejected for 30 percent of the simulated series for $T = 100$ observations, and 42 percent for $T = 200$ observations. However, in long samples with $T = 1000$ observations the null of $r = 0$ can be rejected in 99 percent of the simulated series for all $i = 9, ..., 12$. The trace test for $r = 1$ is almost unaffected by the persistence in $b_t^{(i)}$, though a decrease in the rejection rate can be seen in small samples with high persistence and volatility in $b_t^{(i)}$.

For non-stationary $b_t^{(i)}$ $(i = 13, ..., 16)$ the null of $r = 0$ is rejected with average p-values well above the standard 5 percent significance level for all $i = 13, ..., 16$ and for all sample lenghts considered. However, for small samples the null of $r = 0$ is rejected in up to 59 percent of the series ($i = 13$ with $T = 100$), which shows that when the variance of the shocks to $b_t^{(i)}$ is small it requires a fairly long sample to reject $\mathcal{H}(0)$ even though $b_t^{(i)}$ is simulated as non-stationary. As the sample size increases, the rejection frequency decreases, though the null of $r = 0$ is still rejected in approximately 20 percent of the series in samples with $T = 1000$ observations, irrespective of $\sigma^{(i)}$.

To conclude, the results show that the trace test based on the approximations of the standard asymptotic distributions are not very sensitive to misspecification caused by stochastic cointegration parameters in the underlying data-generating process, as long as there is no or low persistence in the stochastic cointegration parameters in the underlying data-generating process. Though the distribution of test sizes for $\mathcal{H}(0)$ gets skewed toward zero the null of $r = 0$ is clearly rejected based on standard asymptotic inference. Hence, despite that standard asymptotic inference is invalid the conclusions based on such inference appear to be qualitatively correct as the trace tests on average suggest a reduced rank of $r = 1$. However, when $b_t^{(i)}$ is simulated as near non-stationary the distribution of test sizes for $\mathcal{H}(0)$ get so skewed towards zero that the null of $r = 0$ cannot be rejected in a significant proportion of the simulated series unless the sample is very long.

Eigenvalues and Characteristic Roots

Consider next the characteristic roots of the companion matrix and the eigenvalues of $\hat{\Pi}$, shown in Table 3.7. The former are informative about the persistence in the estimated model as measured by the number of unit roots, while the latter are informative about the degree of persistence in the linear combinations $\beta' X_t^{(i)}$.

First, the last columns in Table 3.7 show the modulus of the two largest characteristic roots in the system for the unrestricted model with full rank, and the modulus of the largest unrestricted

Table 3.7: Eigenvalues and Characteristic Roots

				Eigenvalues[a]		Characteristic Roots[b]		
i	$\rho^{(i)}$	$\sigma^{(i)}$	T	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\nu}_{1,r=2}$	$\hat{\nu}_{2,r=2}$	$\hat{\nu}_{2,r=1}$
0	0.00	0.00	100	0.51	0.03	0.95	0.10	0.10
			200	0.50	0.02	0.97	0.07	0.07
			400	0.49	0.01	0.99	0.05	0.05
			1000	0.49	0.00	0.99	0.04	0.04
1	0.00	0.10	100	0.50	0.03	0.95	0.11	0.11
			200	0.50	0.02	0.97	0.08	0.08
			400	0.49	0.01	0.99	0.06	0.06
			1000	0.49	0.00	0.99	0.05	0.05
2	0.00	0.20	100	0.50	0.03	0.95	0.12	0.12
			200	0.49	0.02	0.97	0.09	0.09
			400	0.49	0.01	0.99	0.08	0.08
			1000	0.49	0.00	0.99	0.06	0.06
3	0.00	0.50	100	0.50	0.03	0.95	0.14	0.14
			200	0.49	0.02	0.97	0.11	0.11
			400	0.49	0.01	0.99	0.09	0.09
			1000	0.48	0.00	0.99	0.07	0.07
4	0.00	1.00	100	0.50	0.03	0.95	0.15	0.15
			200	0.49	0.02	0.97	0.12	0.12
			400	0.49	0.01	0.99	0.09	0.09
			1000	0.48	0.00	0.99	0.07	0.07
5	0.50	0.10	100	0.43	0.03	0.95	0.17	0.17
			200	0.39	0.02	0.97	0.23	0.23
			400	0.34	0.01	0.99	0.31	0.31
			1000	0.29	0.00	0.99	0.40	0.40
6	0.50	0.20	100	0.37	0.03	0.95	0.27	0.27
			200	0.33	0.02	0.97	0.35	0.35
			400	0.29	0.01	0.99	0.41	0.41
			1000	0.26	0.00	0.99	0.46	0.46
7	0.50	0.50	100	0.31	0.03	0.95	0.39	0.39
			200	0.28	0.02	0.97	0.44	0.44
			400	0.26	0.01	0.99	0.47	0.47
			1000	0.25	0.00	0.99	0.48	0.48
8	0.50	1.00	100	0.29	0.03	0.95	0.43	0.43
			200	0.27	0.02	0.97	0.46	0.46
			400	0.26	0.01	0.99	0.48	0.48
			1000	0.25	0.00	0.99	0.49	0.49
9	0.95	0.10	100	0.24	0.03	0.95	0.54	0.54
			200	0.11	0.01	0.97	0.78	0.78
			400	0.05	0.01	0.99	0.89	0.89
			1000	0.03	0.00	0.99	0.93	0.93
10	0.95	0.20	100	0.16	0.03	0.95	0.70	0.70
			200	0.08	0.01	0.97	0.85	0.85
			400	0.05	0.01	0.99	0.91	0.91
			1000	0.03	0.00	0.99	0.93	0.93
11	0.95	0.50	100	0.11	0.03	0.95	0.80	0.80
			200	0.06	0.01	0.97	0.88	0.88
			400	0.04	0.01	0.99	0.92	0.92
			1000	0.03	0.00	0.99	0.94	0.94
12	0.95	1.00	100	0.10	0.03	0.95	0.82	0.82
			200	0.06	0.01	0.97	0.89	0.89
			400	0.04	0.01	0.99	0.92	0.92
			1000	0.03	0.00	0.99	0.94	0.94
13	1.00	0.10	100	0.18	0.03	0.95	0.66	0.65
			200	0.06	0.01	0.98	0.89	0.88
			400	0.03	0.01	0.99	0.96	0.96
			1000	0.01	0.00	1.00	0.99	0.98
14	1.00	0.20	100	0.12	0.02	0.95	0.78	0.77
			200	0.05	0.01	0.98	0.91	0.91
			400	0.02	0.01	0.99	0.96	0.96
			1000	0.01	0.00	1.00	0.99	0.98
15	1.00	0.50	100	0.09	0.02	0.96	0.85	0.84
			200	0.05	0.01	0.98	0.93	0.92
			400	0.02	0.01	0.99	0.96	0.96
			1000	0.01	0.00	1.00	0.99	0.99
16	1.00	1.00	100	0.09	0.02	0.96	0.86	0.85
			200	0.04	0.01	0.98	0.93	0.93
			400	0.02	0.01	0.99	0.97	0.96
			1000	0.01	0.00	1.00	0.99	0.99

All reported values are averages over $S = 10,000$ replications.

[a] λ_i is the i'th largest solution to the eigenvalue problem $|\lambda S_{11} - S_{10} S_{00}^{-1} S_{01}| = 0$, see Johansen (1996, Theorem 6.1).

[b] $\hat{\nu}_{i,r=j}$ refers to the modulus of the i'th largest root of the companion matrix for the model with rank $r = j$.

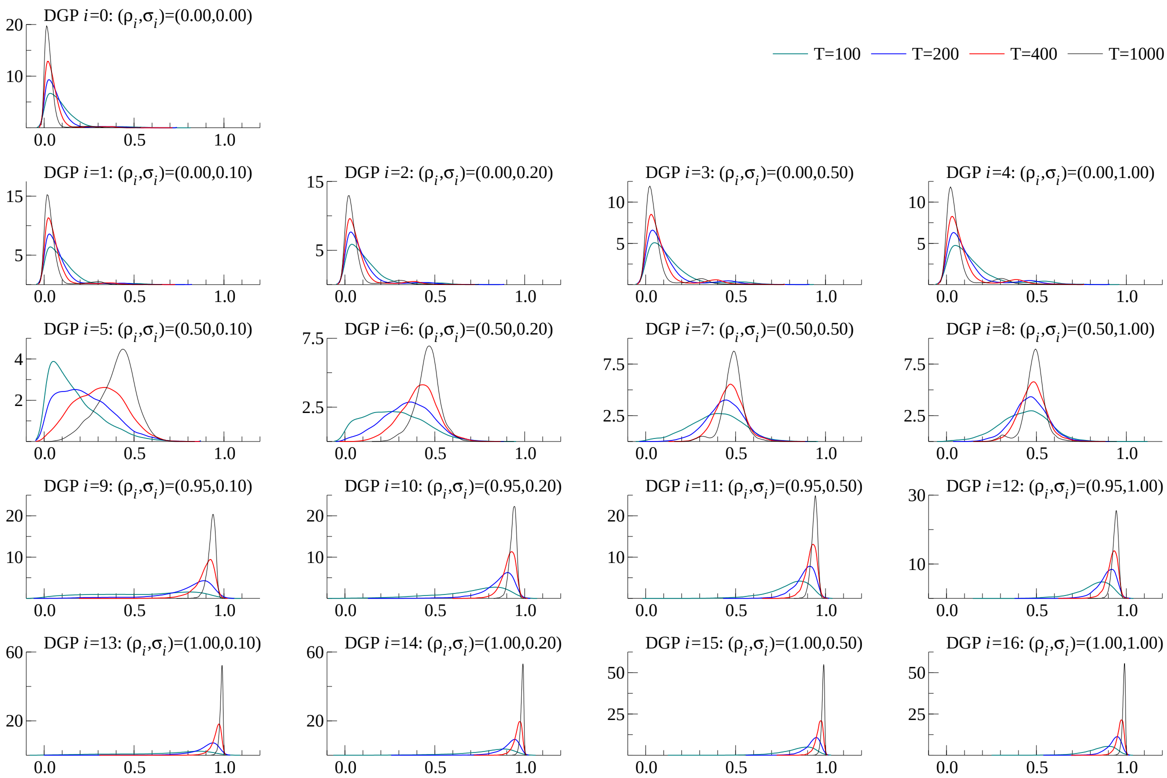

Figure 3.5: Modulus of the largest unrestricted root of the companion matrix for the reduced rank model with $r = 1$.

characteristic root of the reduced rank model with $r = 1$, where a unit root is imposed on the largest characteristic root. Moreover, Figure 3.5 shows the estimated kernel densities of the largest unrestricted characteristic root in the reduced rank model is shown.

In the case with $b_t^{(i)} = 0$ (for $i = 0$) in the data-generating process of the simulated data a root indistinguishably close to a unit root is found, while the second largest characteristic root is low and close to zero. For example, with $T = 100$ observations the largest root is found to be 0.95, on average, which for the short sample is indistinguishable from a unit root, as shown by e.g. Johansen (2006). As the number of observations increases the largest characteristic root converges towards one, while the second largest characteristic root converges towards zero. By imposing a reduced rank $r = 1$ the largest characteristic root is restricted to be a unit root, see Assumption 2, and the second largest characteristic root is left unrestricted.

For stationary $b_t^{(i)}$ (for $i = 1, ..., 12$) in the data-generating process of the simulated data the largest characteristic root remains indistinguishably close to a unit root, as in the case with constant cointegration parameters. However, an increase in $\rho^{(i)}$ increases the persistence in $b_t^{(i)}$, and as a result the second largest characteristic root of the unrestricted system increases, and it increases with the sample length for fixed $(\rho^{(i)}, \sigma^{(i)})$. In the cases with both high $\rho^{(i)}$ and high $\sigma^{(i)}$ the second largest characteristic root is even found to be close to a unit root. As can be seen from the final column in Table 3.7, imposing a reduced rank of $r = 1$, and thereby restricting the largest characteristic root to a unit root, leaves a large unrestricted characteristic root in the estimated reduced rank model. As is evident from Figure 3.5 the estimated kernel densities of the largest unrestricted characteristic root in the reduced rank model increases with the size of $\rho^{(i)}$. For $\rho^{(i)} = 0.95$ the largest unrestricted characteristic root is very close to a unit root, which in the standard cointegrated VAR model typically would be interpreted as $I(2)$-type persistence. This shows that persistence in the process $X_t^{(i)}$, caused by a high degree of persistence in $b_t^{(i)}$, shows up as an 'extra' degree of persistence in the estimated cointegrated VAR model.

Consider next the estimated eigenvalues presented in Table 3.7. Recall that the eigenvalues $1 > \hat{\lambda}_1 > \hat{\lambda}_2 > 0$ can be interpreted as a measure of the 'stationarity' of the cointegration relation i for $i = 1, 2$. The greater the eigenvalue, the more 'stationary' is the linear relation and the cointegration rank is determined by the number of non-zero eigenvalues. With an eigenvalue $\hat{\lambda}_i = 0$ the linear combination $\hat{\beta}' X_t^{(i)}$ is non-stationary, and consequently $\hat{\alpha}_i = 0$, see equation (3.33).

With constant parameters the largest eigenvalue is, on average, found to be 0.5, while the second eigenvalue is, on average, found to be close to zero. This clearly indicates that the bivariate system has one stationary cointegration relation to which at least one variable is equilibrium adjusting, as also indicated by the rank test.

With $\rho^{(i)} = 0$ for $i = 1, ..., 4$ the eigenvalues are, on average, identical to those found in the case with constant parameters in the data-generating process of the simulated data. As $b_t^{(i)}$ is *i.i.d.*, $b_t^{(i)} X_{2t-1}$ is a heteroskedastic term with no persistence, so the estimated cointegration relation $\hat{\beta}' X_t^{(i)}$ is heteroskedastic, but with no persistence, so the eigenvalues are (almost) not affected by $b_t^{(i)}$.

With $0 < \rho^{(i)} < 1$ for $i = 5, ..., 12$, the largest estimated eigenvalues decrease with $\rho^{(i)}$ and with $\sigma^{(i)}$, and moreover they decrease significantly with the sample size. Hence, the greater the persistence and volatility in $b_t^{(i)}$, the lower the largest estimated eigenvalue and the 'less stationary' is the cointegration relation $\hat{\beta}' X_t^{(i)}$. When there is both near non-stationary persistence and high volatility in $b_t^{(i)}$ the largest estimated eigenvalue becomes extremely close to zero. For example, for $i = 12$ the largest eigenvalue is, on average, $\hat{\lambda}_1 = 0.1$ for $T = 100$ observations, $\hat{\lambda}_1 = 0.06$ for $T = 200$ observations, and $\hat{\lambda}_1 = 0.03$ for $T = 1000$ observations. These low eigenvalues illustrate why the trace test in these cases could not reject a reduced rank of $r = 0$, except from in extremely long samples, as this corresponds to $\hat{\lambda}_1 = 0$.

These results show that persistence in the stochastic cointegration parameter of the underlying data-generating process, through the stochastically trendless term $b_t^{(i)} X_{2t-1}$, results in persistence in the estimated cointegration relation $\hat{\beta}' X_t^{(i)}$ in the cointegrated VAR model. Thereby, the largest estimated eigenvalue becomes very small, and if the persistence in $b_t^{(i)}$ is close to non-stationary it can even be found to be very close to zero.

Reduced Rank Estimation of β

As the estimated cointegration vector $\hat{\beta}$ is normalized on $\hat{\beta}_1$, so that $\hat{\beta} = (1, \hat{\beta}_2)$, Table 3.8 shows only the average estimates of $\hat{\beta}_2$ over the $S = 10,000$ replications. The estimated cointegration parameter $\hat{\beta}_2$ is directly comparable to $-b = -1$ for $0 \leq \rho_i < 1.0$, in which case $-b$ is the unconditional mean of the stochastic cointegration parameter given by $-(b + b_t^{(i)})$. As the cointegration rank is $r = 1$, the estimated cointegration vector is identified, and hence standard errors based on equation (12.13) in Juselius (2006) and corresponding t-ratios are presented in Table 3.8. Figure 3.6 shows the estimated kernel densities of the estimated cointegration parameter $\hat{\beta}_2$.

In the case of constant parameters in the data-generating process of the simulated data ($i = 0$) $\hat{\beta}_2$ is a superconsistent estimate of $-b$, see Johansen (1996). This is evident from Figure 3.6, where the distribution of the estimator collapses rapidly around the true value $-b = -1$ as T increases. From Table 3.8 it can be seen that the estimates are very precise and statistically significant, with very small standard errors and very high t-values on average.

In cases of stationary $b_t^{(i)}$ in the data-generating process ($i = 1, ..., 12$) the results show that the reduced rank estimator is a consistent estimator of the unconditional mean $-b$ of the stochastic cointegration parameters $-(b + b_t^{(i)})$, though the rate of convergence is slower than in the case with constant parameters. From Figure 3.6 it is evident that as the sample length increase the estimated kernel density of $\hat{\beta}$ collapses around $-b = -1$. An increase in either $\rho^{(i)}$ or $\sigma^{(i)}$ (holding the other fixed) slows the convergence and increases the estimated standard errors based on standard asymptotic inference, as can be seen from Table 3.8. For example, keeping $\rho^{(i)} = 0.0$ fixed, an increase in $\sigma^{(i)}$ from 0.1 to 1.0 increases the average standard errors by a multiple of roughly 5, with an increase from an average standard error of 0.03 for $T = 100$ to 0.16 for $T = 1000$. However, the estimates of $\hat{\beta}_2$ are still found to be clearly significant on average, except from the extreme case with high $\rho^{(i)}$ and high $\sigma^{(i)}$ for $i = 12$. In the latter case the estimator still converges towards $-b$, but very slowly, and for small samples the distributions are very dispersed.

In cases with non-stationary $b_t^{(i)}$ the estimator does not converge, which is expected as the variables are not found to be cointegrated.

To conclude, the results show that the cointegrated VAR model provides a consistent and very precise estimate of the unconditional mean of the stochastic cointegration parameters in the underlying data-generating process, except from the case of near non-stationarity and high volatility

Table 3.8: Reduced Rank Estimations with $r = 1$: Cointegration Coefficients β

i	$\rho^{(i)}$	$\sigma^{(i)}$	T	$\hat{\beta}_2^*$	$se_{\hat{\beta}_2^*}$ [a]	$\tau_{\hat{\beta}_2^*}$ [1]
0	0.00	0.00	100	-1.00	0.03	-39.97
			200	-1.00	0.02	-77.65
			400	-1.00	0.01	-153.78
			1000	-1.00	0.00	-380.13
1	0.00	0.10	100	-1.00	0.03	-32.31
			200	-1.00	0.02	-55.75
			400	-1.00	0.01	-93.47
			1000	-1.00	0.01	-174.30
2	0.00	0.20	100	-1.00	0.04	-24.33
			200	-1.00	0.03	-38.56
			400	-1.00	0.02	-59.24
			1000	-1.00	0.01	-100.46
3	0.00	0.50	100	-1.00	0.09	-13.03
			200	-1.00	0.06	-18.78
			400	-1.00	0.04	-26.73
			1000	-1.00	0.03	-42.60
4	0.00	1.00	100	-1.00	0.16	-7.06
			200	-1.00	0.11	-9.84
			400	-1.00	0.08	-13.68
			1000	-1.00	0.05	-21.50
5	0.50	0.10	100	-1.00	0.04	-26.55
			200	-1.00	0.03	-41.89
			400	-1.00	0.02	-63.66
			1000	-1.00	0.01	-105.17
6	0.50	0.20	100	-1.00	0.06	-17.33
			200	-1.00	0.05	-24.77
			400	-1.00	0.03	-34.79
			1000	-1.00	0.02	-54.38
7	0.50	0.50	100	-1.00	0.15	-7.89
			200	-1.00	0.11	-10.50
			400	-1.00	0.08	-14.18
			1000	-1.00	0.05	-21.86
8	0.50	1.00	100	-1.00	0.30	-4.03
			200	-1.01	0.22	-5.30
			400	-1.00	0.16	-7.10
			1000	-1.00	0.10	-10.93
9	0.95	0.10	100	-1.01	0.15	-13.13
			200	-0.96	0.18	-12.30
			400	-0.99	0.13	-11.81
			1000	-1.00	0.09	-13.81
10	0.95	0.20	100	-1.02	0.43	-6.38
			200	-0.94	0.48	-5.62
			400	-1.04	0.29	-5.58
			1000	-1.01	0.18	-6.76
11	0.95	0.50	100	-0.59	1.53	-2.28
			200	-0.47	1.17	-2.10
			400	-1.04	0.69	-2.14
			1000	-1.00	0.46	-2.65
12	0.95	1.00	100	-0.71	3.29	-1.08
			200	-2.09	2.59	-1.04
			400	-1.12	1.42	-1.05
			1000	-1.01	0.93	-1.32
13	1.00	0.10	100	-1.80	0.70	-9.33
			200	-1.24	0.91	-6.13
			400	-0.57	2.72	-3.96
			1000	-2.38	2.82	-2.32
14	1.00	0.20	100	-1.09	1.51	-4.28
			200	8.06	4.71	-2.77
			400	-7.15	6.50	-1.92
			1000	4.70	7.76	-1.16
15	1.00	0.50	100	-1.54	3.58	-1.46
			200	-3.14	8.35	-1.03
			400	-26.91	16.55	-0.79
			1000	-0.29	14.53	-0.47
16	1.00	1.00	100	-1.00	8.73	-0.66
			200	-15.54	24.07	-0.51
			400	27.04	28.45	-0.43
			1000	-8.30	26.19	-0.25

All reported values are averages over $S = 10,000$ replications.

[a] Maximum likelihood estimated of $\hat{\beta}_2$ (Johansen, 1996, Theorem 6.1), along with standard errors and t-values as given by equation (12.13) in Juselius (2006).

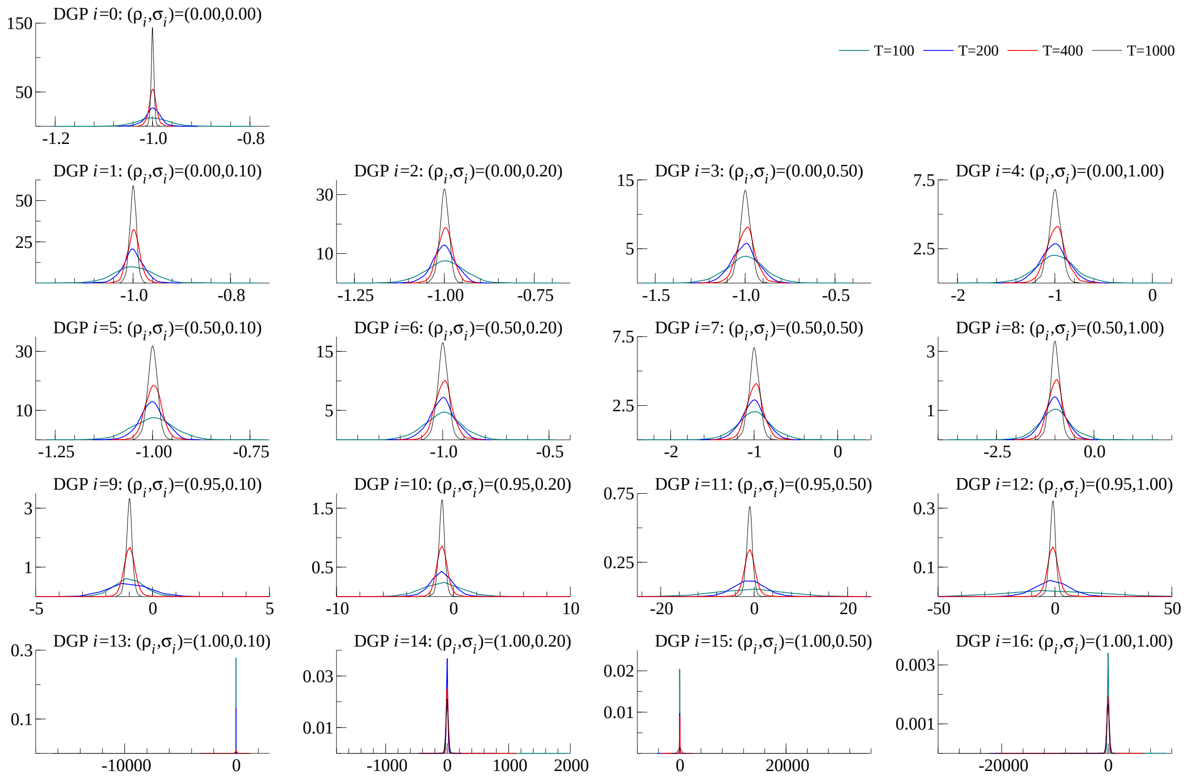

Figure 3.6: Estimated kernel density of the estimated cointegration parameter $\hat{\beta}_2$. The estimated β vector is normalized on $\hat{\beta}_1$, so $\hat{\beta} = (1, \hat{\beta}_2)$ and $\hat{\beta}_2$ is comparable to $-b$ in the data-generating process of the simulated data.

in $b_t^{(i)}$. Though the rate of convergence is slower than in the case of constant parameters in the data-generating process, the estimates of $\hat{\beta}_2$ are found to be clearly significant based on standard asymptotic inference.

Reduced Rank Estimation of α

Finally, consider the estimated adjustment coefficients $\hat{\alpha}$ presented in Table 3.9, while Figures 3.7 and 3.8 present the estimated kernel densities of $\hat{\alpha}_1$ and $\hat{\alpha}_2$, respectively. In the simulations $\alpha = (-1, 0)'$ was used, so, for $0 \leq \rho^{(i)} < 1$, X_{2t} was a weakly exogenous random walk and $X_{1t}^{(i)}$ was purely, and instantly, adjusting to the cointegration relation with stochastic cointegration parameters given by $\beta_t^{(i)\prime} X_{t-1}^{(i)} = (\beta + B_t^{(i)})' X_{t-1}^{(i)}$. However, in the estimated cointegrated VAR model $\hat{\alpha}$ measures the adjustment to the estimated cointegration relation $\hat{\beta}' X_{t-1}^{(i)}$, rather than $\beta_t^{(i)\prime} X_{t-1}^{(i)}$ and therefore we would not expect $\hat{\alpha}$ to be a consistent estimate of α for $0 < \rho^{(i)} < 1$.

In the constant parameter case, $b_t^{(i)} = 0$ for $i = 0$, $\hat{\alpha}_1 = -1$ and clearly significant, while $\hat{\alpha}_2 = 0$ and clearly insignificant, on average, for all sample lengths. As the sample length increases, the average standard errors of both $\hat{\alpha}_1$ and $\hat{\alpha}_2$ decreases.

For $\rho^{(i)} = 0$ and *i.i.d.* stochastic parameter $b_t^{(i)}$, the same results are found; $\hat{\alpha}_1 = -1$ and clearly significant, while $\hat{\alpha}_2 = 0$ and clearly insignificant, on average. In fact, the standard errors decrease on average compared to the case with constant cointegration parameters, and, in particular, an increase in $\sigma^{(i)}$ for fixed $\rho^{(i)} = 0$ decreased the average estimated standard errors.

However, adding persistence in $b_t^{(i)}$ has a significant impact on the estimated adjustment coefficients. From Figures 3.7 and 3.8 it can be seen that as $\rho^{(i)}$ increases (for $0 < \rho^{(i)} < 1$) the estimated kernel density of $\hat{\alpha}_1$ gets skewed towards zero, while the estimated kernel density of $\hat{\alpha}_2$ collapses around zero. Moreover, as $\rho^{(i)}$ and $\sigma^{(i)}$ increase, the estimated kernel density of $\hat{\alpha}_1$ converges faster towards zero as the number of observations T increases, and the estimated kernel density of $\hat{\alpha}_2$ collapses faster around zero as T increases. These results are also evident from table 3.9, where it can be seen that on average $\hat{\alpha}_1$ and $\hat{\alpha}_2$ become smaller as $\rho^{(i)}$ increases. However, the average standard errors are also decreasing, although at a slower rate than the parameter estimates, so on average $\hat{\alpha}_1$ is found to be significant, while $\hat{\alpha}_2$ is clearly found to be insignificant on average. Thus, despite the low estimated adjustment coefficients the estimates correctly finds that $X_{1t}^{(i)}$ is adjusting to the estimated cointegration relation, while X_{2t} is found to be weakly exogenous.

The finding that persistence in $b_t^{(i)}$ skews the estimated kernel density towards zero (for $0 < \rho^{(i)} < 1$) is a direct result of persistence in the estimated cointegration relation $\hat{\beta}' X_t^{(i)}$ caused by the stochastically trendless term $b_t^{(i)} X_{2t-1}$ in the underlying data-generating process. As a result of persistence in $\hat{\beta}' X_t^{(i)}$, the estimated eigenvalues gets skewed towards zero, whereby the estimated adjustment coefficients get skewed towards zero, c.f. the relation between the estimated eigenvalues and adjustment parameters in equation (3.33).

To summarize the main findings for the reduced rank estimations, the results show that the cointegrated VAR model delivers a consistent estimate of the unconditional mean of the stochastic cointegration parameters of the underlying data-generating process. However, if there is persistence in the stochastic cointegration parameters of the data-generating process, this persistence implies persistence in the estimated cointegration relation. As a consequence, the largest estimated

Table 3.9: Reduced Rank Estimations with $r = 1$: Adjustment Coefficients α

i	$\rho^{(i)}$	$\sigma^{(i)}$	T	$\hat{\alpha}_1$	$se_{\hat{\alpha}_1}$ [a]	$\tau_{\hat{\alpha}_1}$ [b]	$\hat{\alpha}_2$	$se_{\hat{\alpha}_2}$ [a]	$\tau_{\hat{\alpha}_2}$ [b]
0	0.00	0.00	100	-1.02	0.15	-7.07	-0.00	0.10	-0.00
			200	-1.01	0.10	-9.95	-0.00	0.07	-0.01
			400	-1.01	0.07	-13.97	0.00	0.05	0.00
			1000	-1.00	0.05	-22.08	0.00	0.03	0.01
1	0.00	0.10	100	-1.02	0.14	-7.64	-0.00	0.09	0.00
			200	-1.01	0.09	-11.14	-0.00	0.06	-0.00
			400	-1.01	0.06	-16.45	-0.00	0.03	0.00
			1000	-1.00	0.04	-27.47	0.00	0.02	0.01
2	0.00	0.20	100	-1.03	0.13	-8.33	0.00	0.07	0.01
			200	-1.02	0.08	-12.19	0.00	0.04	0.00
			400	-1.01	0.06	-17.90	-0.00	0.02	-0.00
			1000	-1.00	0.03	-29.29	0.00	0.01	0.01
3	0.00	0.50	100	-1.04	0.11	-9.33	0.00	0.04	0.01
			200	-1.02	0.08	-13.32	0.00	0.02	0.00
			400	-1.01	0.05	-19.06	-0.00	0.01	0.00
			1000	-1.00	0.03	-30.31	0.00	0.00	0.01
4	0.00	1.00	100	-1.04	0.11	-9.76	0.00	0.02	0.01
			200	-1.02	0.08	-13.70	0.00	0.01	0.00
			400	-1.01	0.05	-19.35	-0.00	0.01	0.00
			1000	-1.00	0.03	-30.53	0.00	0.00	0.01
5	0.50	0.10	100	-0.87	0.13	-6.62	0.00	0.08	0.01
			200	-0.79	0.09	-8.94	-0.00	0.05	-0.00
			400	-0.71	0.06	-12.04	-0.00	0.03	0.00
			1000	-0.62	0.03	-17.94	0.00	0.02	0.01
6	0.50	0.20	100	-0.75	0.12	-6.32	0.00	0.06	0.01
			200	-0.66	0.08	-8.52	0.00	0.04	0.00
			400	-0.60	0.05	-11.65	-0.00	0.02	-0.00
			1000	-0.55	0.03	-17.99	0.00	0.01	0.01
7	0.50	0.50	100	-0.63	0.10	-6.14	0.00	0.03	0.02
			200	-0.57	0.07	-8.35	0.00	0.02	0.01
			400	-0.54	0.05	-11.59	-0.00	0.01	-0.00
			1000	-0.52	0.03	-18.09	0.00	0.00	0.01
8	0.50	1.00	100	-0.59	0.10	-6.14	0.00	0.02	0.01
			200	-0.55	0.07	-8.36	0.00	0.01	0.01
			400	-0.52	0.05	-11.60	-0.00	0.00	-0.00
			1000	-0.51	0.03	-18.10	0.00	0.00	0.01
9	0.95	0.10	100	-0.48	0.11	-4.14	-0.00	0.07	0.00
			200	-0.26	0.06	-3.69	-0.00	0.04	-0.01
			400	-0.14	0.04	-3.83	0.00	0.02	-0.00
			1000	-0.09	0.02	-5.22	-0.00	0.01	-0.01
10	0.95	0.20	100	-0.29	0.08	-3.22	0.00	0.04	0.01
			200	-0.16	0.04	-3.23	-0.00	0.02	-0.01
			400	-0.10	0.02	-3.81	0.00	0.01	0.00
			1000	-0.07	0.01	-5.42	-0.00	0.00	-0.00
11	0.95	0.50	100	-0.18	0.06	-2.66	0.00	0.02	0.01
			200	-0.11	0.03	-3.06	0.00	0.01	0.00
			400	-0.08	0.02	-3.84	0.00	0.00	0.00
			1000	-0.06	0.01	-5.51	-0.00	0.00	-0.01
12	0.95	1.00	100	-0.15	0.05	-2.50	0.00	0.01	0.01
			200	-0.10	0.03	-2.98	0.00	0.00	0.00
			400	-0.08	0.02	-3.82	0.00	0.00	0.00
			1000	-0.06	0.01	-5.50	-0.00	0.00	-0.01
13	1.00	0.10	100	-0.34	0.09	-3.23	0.00	0.05	0.01
			200	-0.12	0.04	-2.25	-0.00	0.02	-0.02
			400	-0.04	0.02	-1.84	0.00	0.01	0.02
			1000	-0.01	0.01	-1.66	-0.00	0.00	-0.00
14	1.00	0.20	100	-0.20	0.07	-2.37	0.00	0.03	0.00
			200	-0.07	0.03	-1.91	-0.00	0.01	-0.02
			400	-0.03	0.01	-1.70	0.00	0.00	0.01
			1000	-0.01	0.01	-1.60	-0.00	0.00	-0.00
15	1.00	0.50	100	-0.11	0.05	-1.77	0.00	0.01	0.01
			200	-0.05	0.03	-1.64	-0.00	0.00	-0.01
			400	-0.02	0.01	-1.57	0.00	0.00	0.02
			1000	-0.01	0.01	-1.54	-0.00	0.00	-0.01
16	1.00	1.00	100	-0.10	0.05	-1.60	0.00	0.01	0.00
			200	-0.05	0.03	-1.55	0.00	0.00	-0.01
			400	-0.02	0.01	-1.53	0.00	0.00	0.02
			1000	-0.01	0.01	-1.52	-0.00	0.00	-0.01

All reported values are averages over $S = 10,000$ replications.
[a] Standard errors of $\hat{\alpha}_1$ and $\hat{\alpha}_2$, respectively, as given by equation (12.14) in Juselius (2006).
[b] Corresponding t-values of $\hat{\alpha}_1$ and $\hat{\alpha}_2$, respectively

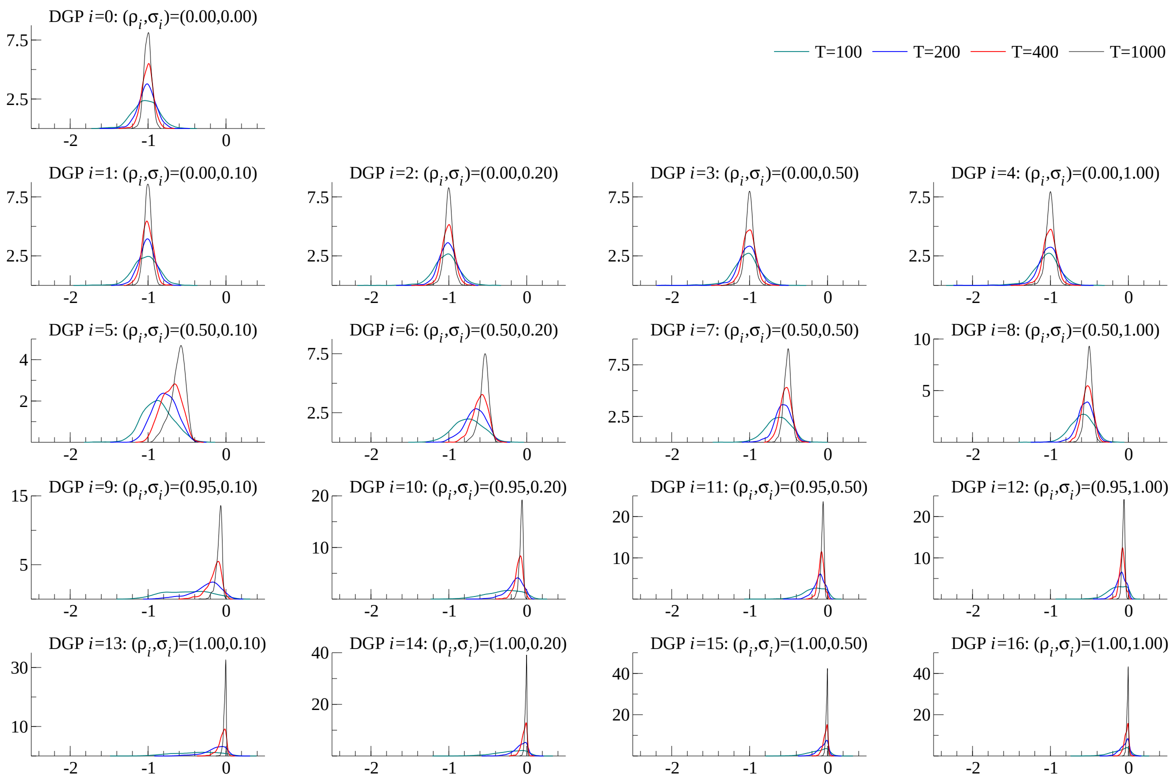

Figure 3.7: Estimated cointegration adjustment parameter $\hat{\alpha}_1$.

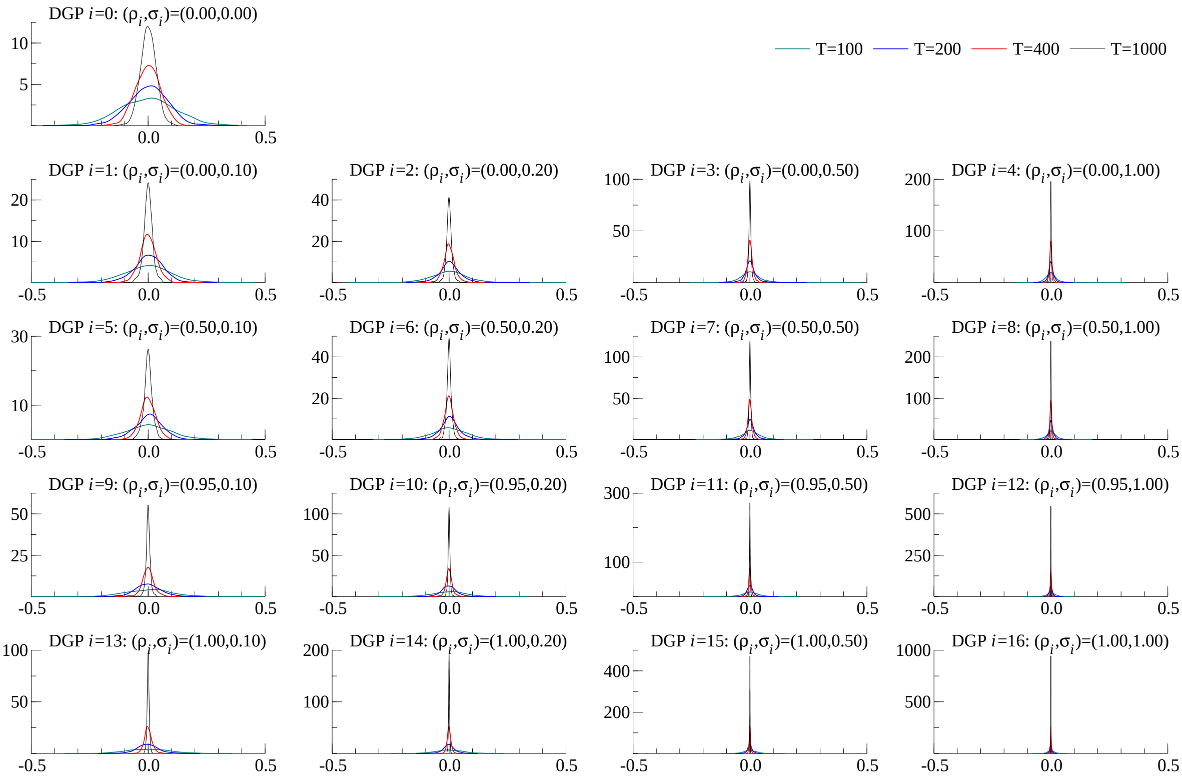

Figure 3.8: Estimated cointegration adjustment parameter $\hat{\alpha}_2$.

eigenvalue decreases significantly and the estimated adjustment parameters get skewed significantly towards zero. Though standard asymptotic inference is found to be sensitive to the misspecification caused by the stochastic cointegration parameters in the data-generating process, the conclusions reached based on the standard inference appears to be qualitatively correct, in the sense that in most cases considered a reduced rank of $r = 1$ is found, the estimated cointegration parameters are clearly significant, and the correct variable is found to be significantly adjusting to the estimated cointegration relation.

3.5 Conclusion

The purpose of the simulation study presented in this paper is to address to what extent the classic cointegrated VAR model can be used as an approximation to estimate the unconditional mean of stochastic cointegration parameters—given by $\beta_t = \beta + B_t$, where B_t is a mean zero stationary process so that β is the unconditional mean of β_t—and how the misspecification caused by B_t not being captured by the constant parameter cointegrated VAR model affects the results and inference.

First, the results show that the estimated cointegrated VAR models appear statistically well-specified in shorter samples typical for macroeconomic data (except from the extreme cases with both very persistent and very volatile stochastic parameters), despite the misspecification of the cointegrated VAR model compared to the data-generating process of the simulated data. Importantly, it is found that persistence caused by the stochastic parameters of the underlying data-generating process can be captured through the short-run structure in the cointegrated VAR model, so that the estimated residuals are found not to be autocorrelated. Though, heteroskedasticity caused by the stochastic parameters cause problems with excess kurtosis and non-normality in the estimated residuals of the cointegrated VAR model, but identifying this misspecification requires a fairly long sample, except if the variance of the shocks to the stochastic cointegration parameters is of the same magnitude as the variance of the shocks to the levels of the variables.

Second, the results show that though the trace test based on standard asymptotic inference is sensitive to the misspecification caused by the stochastic cointegration parameters in the data-generating process, the trace test correctly suggests a reduced rank of $r = 1$ and thereby it appears to be qualitatively correct, except from cases with near non-stationary persistence and high volatility in the stochastic cointegration parameters.

Third, the results show that the cointegrated VAR model provides a consistent and very precise estimate of the constant unconditional mean of the stochastic cointegration parameters. Persistence in the underlying stochastic parameters shows up in the estimated cointegrated VAR model as persistent deviations from the estimated cointegrated relations, low estimated eigenvalues, and adjustment coefficients skewed towards zero. Moreover, it shows up in the short-run structure as inclusion of lagged first-differences is required to remove autocorrelation in the estimated residuals, and it results in excess kurtosis and non-normality of the estimated residuals.

A limitation of the presented simulation study is that recursive tests for parameter instability are not considered, and such tests might identify the parameter instability. The results presented in this paper suggest that bounded underlying parameter-instability is most likely identified as persistence and in the adjustment coefficients, the short-run structure, and the residuals in the cointegrated

VAR model, rather than in the estimated cointegration parameters. Though, recursive tests for parameter non-constancy are interesting potential extensions left for future work.

However, the point of this study is not to show how underlying bounded parameter-instability might be identified and modeled directly, which would require other tests and another econometric model than the classic cointegrated VAR model considered in this paper. Rather the point is to show that if misspecification caused by bounded parameter-instability in the underlying data-generating process is not identified, it will show up in an estimated cointegrated VAR model as persistent deviations from the estimated cointegration relations and correspondingly low estimates of the adjustment coefficients. In this case, the estimated cointegration relations can still be interpreted as long-run equilibrium relations, but they are defined by the unconditional mean of the stationary stochastic cointegration parameters rather than constant parameters, and the persistent deviations and slow estimated adjustment are consequences of the underlying bounded parameter-instability.

Thereby, the findings in this paper provide a new potential explanation for the 'persistence puzzle', which is a frequent puzzle when standard macroeconomic and financial theories are estimated and tested empirically with the cointegrated VAR model. The novelty consists in the result that persistent deviations from the estimated long-run equilibrium relations and slow estimated adjustment can potentially be caused by persistent, but stationary, parameter-instability and stochastic cointegration in the underlying data-generating process—even when the adjustment to the stochastic equilibrium takes place instantly and the exogenous shocks to the levels of the variables are not persistent. The results presented suggest that if such underlying bounded parameter-instability is present, it is not captured and hard to identify in the cointegrated VAR model, thus leading to the conclusion of an empirical 'persistence puzzle' as long as the assumption of constant parameters is maintained.

An interesting extension for future work is to address how persistence caused by boundedly time-varying cointegration parameters can be distinguished from persistence caused by slow adjustment or persistent exogenous shocks. It might indeed be possible to distinguish between these sources of persistence within the cointegrated VAR model, for example through recursive tests for parameter constancy, but more likely it requires development of new econometric methods to directly estimate extensions of the cointegrated VAR model with stochastic cointegration parameters.

3.6 References

Bec, F. and Rahbek, A. (2004), 'Vector equilibrium correction models with non-linear discontinuous adjustments', *The Econometrics Journal* **7**(2), 628–651.

Bec, F., Rahbek, A. and Shephard, N. (2008), 'The ACR Model: A Multivariate Dynamic Mixture Autoregression', *Oxford Bulletin of Economics and Statistics* **70**(5), 583–618.

Dennis, J. G., Juselius, K., Johansen, S. and Hansen, H. (2006), *CATS in RATS: Cointegration analysis of time series*, Estima, Evanston, IL.

Doornik, J. (1998), 'Approximations to the asymptotic distribution of cointegration tests', *Journal of the American Statistical Association* **74**, 427–431.

Doornik, J. A. (2007), *Object-Oriented Matrix Programming Using Ox*, 3 edn, Timberlake Consultants Press.

Doornik, J. A. and Hansen, H. (1994), An omnibus test for univariate and multivariate normality. Working Paper.

Engle, R. F. and Granger, C. (1987), 'Co-integration and error correction: Representation, estimation, and testing', *Econometrica* **55**(2), 251–276.

Godfrey, L. G. (1988), *Misspecification Tests in Econometrics*, Cambridge University Press, Cambridge.

Harris, D., McCabe, B. and Leybourne, S. (2002), 'Stochastic cointegration: estimation and inference', *Journal of Econometrics* **111**(2), 363–384.

Helpman, E. (1981), 'An exploration in the theory of exchange-rate regimes', *Journal of Political Economy* **89**(5), 865–890.

Hoover, K. D. (2006), The Past as the Future: The Marshallian Approach to Post Walrasian Econometrics, *in* D. Colander, ed., 'Post Walrasian Macroeconomics: Beyond the Dynamic Stochastic General Equilibrium Model', Cambridge University Press, Cambridge, chapter 12, pp. 239–257.

Hoover, K. D., Johansen, S. and Juselius, K. (2008), 'Allowing the data to speak freely: The macroeconometrics of the cointegrated vector autoregression', *The American Economic Review* **98**(No. 2, Papers and Proceedings of the One Hundred Twentieth Annual Meeting of the American Economic Association), 251–255.

Johansen, S. (1996), *Likelihood-Based Inference in Vector Autoregressive Models*, Oxford University Press, Oxford.

Johansen, S. (2006), Confronting the economic model with the data, *in* D. Colander, ed., 'Post Walrasian Macroeconomics: Beyond the Dynamic Stochastic General Equilibrium Model', Cambridge University Press, Cambridge, chapter 15, pp. 287–300.

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

Johansen, S., Juselius, K., Frydman, R. and Goldberg, M. D. (2010), 'Testing Hypotheses in an I(2) Model with Application to the Persistent Long Swings in the Dmk/$ Rate', *Journal of Econometrics* **158**, 117–129.

Juselius, K. (1995), 'Do purchasing power parity and uncovered interest parity hold in the long run? an example of likelihood inference in a multivariate time-series model', *Journal of Econometrics* **69**, 211–240.

Juselius, K. (2006), *The Cointegrated VAR Model: Methodology and Applications*, Oxford University Press, Oxford.

Juselius, K. (2009 *a*), The long swings puzzle: What the data tell when allowed to speak freely, *in* K. Patterson and T. C. Mills, eds, 'The Palgrave Handbook of Empirical Econometrics', Macmillan, pp. 367–384.

Juselius, K. (2009 *b*), 'Special issue on using econometrics for assessing economic models—an introduction', *Economics: The Open-Access, Open-Assessment E-Journal* **3**(2009-28).

Juselius, K. and MacDonald, R. (2004), 'International parity relationships between the usa and japan', *Japan and the World Economy* **16**(1), 17–34.

Juselius, K. and Toro, J. (2005), 'The effect of joining the ems. monetary transmission mechanisms in spain', *Journal of International Money and Finance* **24**, 509–531.

Kristensen, D. and Rahbek, A. (2010), 'Likelihood-based inference for cointegration with nonlinear error-correction', *Journal of Econometrics* **158**(1), 78–94.

Kristensen, D. and Rahbek, A. (2013), 'Testing and Inference in Nonlinear Cointegrating Vector Error Correction Models', *Econometric Theory* **forthcoming**.

Lötkepohl, H. and Krätzig, M. (2004), *Applied Time Series Econometrics*, Cambridge University Press, Cambridge.

McCabe, B., Leybourne, S. and Harris, D. (2003), Testing for Stochastic Cointegration and Evidence for Present Value Models. Working Paper.

Møller, N. F. (2008), 'Bridging economic theory models and the cointegrated vector autoregressive model', *Economics: The Open-Access, Open-Assessment E-Journal* **2**(2008-36).

Obstfeld, M. and Rogoff, K. (1995), 'Exchange rate dynamics redux', *Journal of Political Economy* **103**, 624–660.

Obstfeld, M. and Rogoff, K. (2000), 'New directions for stochastic open economy models', *Journal of International Economics* **50**(1), 117–153.

Paruolo, P., Nejstgaard, E. and Rahbek, A. (2013), Likelihood-based inference in the acr cointegrated model. Working Paper, Department of Economics, University of Copenhagen.

Rahbek, A. and Nielsen, H. B. (2012), Unit root vector autoregression with volatility induced stationarity. Working Paper, Department of Economics, University of Copenhagen.

Rogoff, K. (1996), 'The purchasing power parity puzzle', *Journal of Economic Literature* **34**(2), 647–668.

Stockman, A. C. (1980), 'A theory of exchange rate determination', *Journal of Political Economy* **88**, 673–698.

Svensson, L. E. (1985), 'Currency prices, terms of trade, and interest rates: A general equilibrium asset-pricing cash-in-advance approach', *Journal of International Economics* **18**(1-2), 17–41.