

VU Research Portal

On the stability of stochastic dynamic systems and their use in econometrics

Nientker, M.H.C.

2019

document version

Publisher's PDF, also known as Version of record

[Link to publication in VU Research Portal](#)

citation for published version (APA)

Nientker, M. H. C. (2019). *On the stability of stochastic dynamic systems and their use in econometrics*. [PhD-Thesis - Research and graduation internal, Vrije Universiteit Amsterdam].

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:

vuresearchportal.ub@vu.nl

On the stability of stochastic dynamic systems and their use in econometrics

Marc Nientker



**On the stability of stochastic dynamic systems
and their use in econometrics**

Marc Nientker

ISBN: 978 90 3610 577 4

Cover design: Crasborn Graphic Designers bno, Valkenburg a.d. Geul

© Marc Nientker, 2019

All rights reserved. Save exceptions stated by the law, no part of this publication may be reproduced, stored in a retrieval system of any nature, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, included a complete or partial transcription, without the prior written permission of the author, application for which should be addressed to the author.

This book is no. 749 of the Tinbergen Institute Research Series, established through cooperation between Thela Thesis and the Tinbergen Institute. A list of books which already appeared in the series can be found in the back.

VRIJE UNIVERSITEIT

On the stability of stochastic dynamic systems and their use in econometrics

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor of Philosophy
aan de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. V. Subramaniam,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de School of Business and Economics
op woensdag 9 oktober 2019 om 13.45 uur
in de aula van de universiteit,
De Boelelaan 1105

door

Marc Nientker

geboren te Amsterdam

promotoren: dr. F. Blasques Albergaria Amaral
prof.dr. S.J. Koopman

Acknowledgements

Throughout the writing of my PhD dissertation I have received a great deal of support and assistance. First and foremost I would like to thank my supervisors Francisco Blasques and Siem Jan Koopman for their invaluable guidance that already began far before my PhD period. They have taught me the core of econometrics while I was a student, have supported me via advice and reference letters with my ambitions to study abroad and afterwards allowed me to come back to work on any topic I found interesting. During my PhD they have not only helped with my research and writing, but also provided constant career advice and countless opportunities to learn new things both in and outside of academics. For all of that and more, I am extremely grateful.

I thank my colleagues at the VU University Amsterdam, in particular, Agnieszka, Artem, Charles, Ilka, Julia, Mengheng and Rutger, for their help with my research and teaching, but also for the many fun hours outside of work. Special thanks go to my long time office mates Hande and Paolo for enduring me for all those years, answering all the possible questions I could come up with and filling each work day with laughter.

Finally, I would like to thank my parents Peter and Dominique, my brothers Wouter and Eric and my girlfriend Jessica. My PhD years have been the best of my life so far, and having been able to share them with you has made them that much more meaningful.

Amsterdam, August 19, 2019.

Marc Nientker

Contents

Contents	vii
1 Introduction	1
1.1 Time-series literature: a selective review	1
1.2 Contributions of the thesis	4
2 Stationarity, Ergodicity and Mixing of Resetting Time Series	7
2.1 Introduction	7
2.2 Stability results	10
2.3 Application to heteroscedastic volatility modelling.	17
2.3.1 Examples	22
2.3.2 Moments	24
2.4 Conclusion	28
3 A Time-Varying Parameter Model for Local Explosions	29
3.1 Introduction	29
3.2 Model for Local Explosions	31
3.2.1 Bubble variety	34
3.3 Probabilistic and statistical analysis	37
3.3.1 The model as a data generating process	38
3.3.2 The model as a filter	40
3.3.3 The likelihood	42
3.3.4 Asymptotic results	42
3.4 Illustrations	45
3.4.1 Simulation study	46

CONTENTS

3.4.2	The BTC/USD exchange rate	48
3.5	Conclusion	50
3.6	Appendix: Proofs	51
4	Transformed Perturbation Solutions for Dynamic Stochastic General Equi-	
	librium Models	65
4.1	Introduction	65
4.2	Transformed Perturbation	69
4.2.1	The state space	69
4.2.2	Function approximation methods	70
4.2.3	Perturbation	71
4.2.4	The transformed perturbation method	72
4.3	Probabilistic analysis of the solutions	74
4.4	The plug-in tau	78
4.4.1	Ensuring stability	78
4.4.2	Preserving nonlinear dynamics	80
4.4.3	Choice for tau	82
4.5	Accuracy	83
4.5.1	Theoretical results	83
4.5.2	Applications	86
4.6	Conclusion	99
4.7	Appendix: Proofs	100
	Bibliography	113
	Summary	121
	Samenvatting	123

Chapter 1

Introduction

1.1 Time-series literature: a selective review

A time series is a collection of observations that are indexed in the order of time. Such data structures are found in many fields of science (such as economics, finance, demographics, etc) and business (such as risk assessment, inventory management, logistics optimization, etc). The main motivation for studying time series is the assumption that the past and the future behaviour of a process contain similarities, and thus that one can build predictions for the future based on past observations. The connection between past and future is quite a delicate one, if the dependence is very weak then prediction is fruitless, while a too complex dependence makes prediction very complicated. Statistical analysis and modelling tries to describe the dependence by designing a collection of possible probabilistic descriptions of the data, called a model, and then using the observed past to decide on the most appropriate specification within the model through the construction of an estimator.

The popularisation of time series models started with linear models such as the autoregressive moving average model (ARMA) of Box et al. (1970). The fundamental idea was to let dependence change over time and instead fix the conditional dependence. That is, given yesterday we expect different dependence for today and tomorrow, but the way today depends on yesterday and tomorrow depends on today is the same. The linearity of the models makes it possible to get closed forms for most entities of interest and thus makes the analysis of the models very tractable. ARMA models are able to describe a large class of time series dynamics, in fact, the renowned Wold decomposition theorem

(Wold, 1938) shows that any covariance-stationary time series has a representation as the sum of a deterministic time series and an infinite moving average. Therefore such time series can theoretically be approximated by taking a large number of lags in the ARMA specification.

Nevertheless, a statistical model is only as strong as the closest description to the actual data generating process that it contains. Therefore statisticians have proceeded by enlarging models to increase the likeliness of containing such a close description. Non-linear statistical models have become increasingly popular with models such as threshold models (Tong and Lim, 1980; Zakoian, 1994) or Markov regime switching models (Hamilton, 1989) that allow different ranges of the state space to have distinct effects on the dependence structure. A general way to extend a given model to a larger collection of distributions is to make it dynamic by choosing a parameter and making it time-varying. The most famous application of this approach can be found in heteroskedastic volatility models. The pioneering examples of such models are the ARCH model of (Engle, 1982), that has led to a large portion of literature on extended specifications such as the GARCH (Bollerslev, 1986) and the EGARCH (Nelson, 1991) model, and the stochastic volatility model of Taylor (2008) that has lead to many extensions both specification (Harvey and Shephard, 1996) and estimation (Jacquier et al., 2002) wise.

There are two general classes of time-varying parameter models. The first class, called parameter-driven, specifies the dynamics of the time-varying parameter as a new stochastic data generating process with its own disturbance process. The second class, denoted observation-driven, describes the time varying parameter at each point of time as a function of a, potentially infinitely long, sequence of past observations of the data. Both strands of time-varying parameter time series models have their advantages and disadvantages. Parameter driven models typically satisfy desired statistical stability properties, however the likelihood usually does not have a closed form so that computationally intensive simulation techniques are needed to find the best fitting specification within the model. Observation driven models do have a closed form for the likelihood and thus are generally easier to estimate. However, showing that the model is statistically stable typically requires careful mathematical analysis.

Stability features of both the data generating process behind the observations and the unobserved time varying parameter are very helpful. Statistical properties such as stationarity, ergodicity and mixing provide the basis for the analysis of limiting estimator behavior as the amount of observations goes to infinity. Specifically they imply versions of the law of large numbers and the central limit theorem that can then be applied to show consistency (convergence to the true unknown model) and asymptotic normality (convergence rate and asymptotic distribution). The stability properties for linear models can be fully characterised using Lyapunov exponents as is done in Bougerol and Picard (1992a,b). Nonlinear models require a lot more work and have traditionally been studied using Markov chain theory to obtain geometric ergodicity as in (Meyn and Tweedie, 2012). Verifying the underlying assumptions regarding proper commuting behavior of the Markov chain over the state space can be difficult, but once those are satisfied, ergodicity follows very generally from moderate “Foster-Lyapunov” drift criteria that essentially ensure that the Markov chain never wanders too far off.

Time-varying parameter models pose a new challenge, because the parameter process itself is an unobserved component of the model. As a solution statisticians construct an approximation for the process by choosing a starting point and then recursively filtering the time-varying parameter. Deriving limiting behaviour of this process then requires that the approximation converges to a stable solution, a property that has been denoted invertibility in Straumann (2005). Bougerol (1993) and Straumann (2005) propose a method to obtain invertibility that is based on stochastic recurrence equations (SREs) satisfying contraction conditions. Their result is a stochastic variant of Banach’s fixed point theorem and requires the stochastic recurrence equation (SRE) to be uniformly contracting over the whole state space, on average. Although the condition is fairly elementary to write down for a given model, it introduces some complications. The one step contraction condition is typically possible to derive analytically, but imposes a very strict restriction on the possible parameter values within the model. Typically, SREs exhibit non contractive areas over small parts of the state space, which by the uniform condition imply that the stability parameter region becomes impractically small. Resorting to higher fold contraction conditions leads to larger stability regions, however these regions cannot be analytically determined as they depend on the unknown data generating process. Alternatively, one

can compute empirical stability regions as in Wintenberger (2013) and Blasques et al. (2018a).

1.2 Contributions of the thesis

This thesis contains three chapters of research that explore the stability of time series models in a purely theoretic, a financial and a macroeconomic setting. Chapter 2 and Chapter 3 are somewhat linked and explore invertibility for a class of observation driven time-varying parameter models. Chapter 4 is fully self contained and discusses nonlinear macro economic time series models without time-varying parameters. I provide a short description of each chapter below, a more detailed explanation including relevant references in the literature can be found in the introduction of each specific chapter.

Chapter 2 is joint work with Francisco Blasques and discusses a novel invertibility condition that provides a stability region for a large collection of models that is typically impossible to analyze using the contraction condition of Bougerol (1993). Specifically the condition allows for discontinuities and explosive, non-contracting or chaotic behavior of the SRE over parts of the state space. In return for relaxing the contraction condition we impose that the SRE satisfies a resetting requirement that involves there being a positive probability of the SRE updating to a value that is independent of the past of the process. The proof that this implies invertibility can be summarised in one sentence: it is not important how far two paths diverge from one another as long as they eventually collapse to the same value. That means that between any two resetting times, any imaginable sample path behavior is allowed as the reset ensures that it returns to a fixed value. The resetting condition seems strong, but is typically satisfied in time series that exhibit bubble collapses, in which case the collapse itself can be chosen as the resetting moment. Many time series contain these collapsing dynamics such as volatility bubbles studied in Saïdi (2003) and Saïdi and Zakoian (2006), financial bubbles studied in Gouriéroux and Zakoian (2013, 2017), Blasques et al. (2018b) or Chapter 3 of this thesis, and time series for overshooting predator-prey populations such as the famous Canadian lynx-hare and wolf-moose datasets studied in Tsay (1989) or Teräsvirta (1994). Additionally, the framework lends itself naturally to regime switching models, where we then make one

regime independent of the past so that it enforces a reset and in return get complete freedom for the other regimes. We illustrate the generality of the theory and how to apply it by deriving the specific parameter region for the volatility bubble model in Saïdi and Zakoian (2006).

Chapter 3 is joint work with Francisco Blasques and Siem Jan Koopman and studies speculative bubbles in time series of financial asset prices. The rational expectations literature on asset pricing models the asset price process as the sum of a fundamental value process and a locally explosive process that describes a burst followed by sharp mean-reverting dynamics. We mimic this approach within an observation driven time-varying parameter model, where we split the level in a time varying fundamental value process and a bubble specification. We provide a general framework that encompasses a wide possibility of bubble dynamics as the interplay between the fundamental value and bubble process allows for various impulse response functions. We estimate the model using a classical maximum likelihood approach. All the stability properties such as stationarity, ergodicity, mixing and invertibility are proven using the theoretical results from Chapter 2. We illustrate the flexibility of the model by filtering the bitcoin / US dollar exchange rate around its biggest relative bubble and show that in sample predictions anticipate the bubble collapse in advance. We give some insights into the advantages of observation driven models and the ease at which they allow for the derivation of quantities such as bubble burst probabilities, bubble emergence probabilities or expected bubble life times.

Chapter 4 is joint work with Francisco Blasques and provides econometric foundations for perturbation, a well known method to approximate solutions of dynamic stochastic equilibrium models. Many approximation methods exist and there are some properties to keep in mind when selecting one, mainly: accuracy, speed, stability and accessibility. Arbitrarily accurate global approximation methods exist such as value function iteration (Bertsekas, 1987), projection (Judd, 1992) or machine learning Norets (2012). Typically these methods are less accessible to the practitioner as they are more complex to code. Moreover, traditionally these methods were too slow for an estimation setting and thus require additional techniques such as parallel computing and entering parameters as pseudo-states in the model. Perturbation (Judd and Guu, 1997; Schmitt-Grohé and Uribe, 2004) is the most commonly used method and focusses on speed and accessibility by

approximating the solution via a Taylor expansion around the deterministic steady state. It is well known, however, that a higher-order polynomial, and thus perturbation, defines an unstable dynamic system which produces explosive paths. This means that none of the typical stability properties hold and thus that existence of relevant moments and consistency or asymptotic normality of estimators cannot be derived. In order to deal with the unstable dynamics of higher-order perturbation solutions, Kim et al. (2008) proposed the pruning method. The pruning method has been successfully implemented in software packages and effectively solves the problem of explosive dynamics. New complications are introduced however, pruning is a simulation-based approximation and hence does not provide a policy function. Moreover, the method has to sacrifice local accuracy of the approximation to obtain stability. Our paper introduces a simple correction to perturbation solutions that is designed to enrich perturbation solutions with all the desirable stochastic properties needed for parameter estimation and statistical inference. Our correction transforms the standard perturbation approximation by replacing higher order monomials in the Taylor expansion with transformed ones that are based on the transformed polynomials introduced in Blasques et al. (2014). These transformed monomials force sample paths that move far away from the deterministic steady state into linear dynamics, which makes the resulting dynamic system a prime candidate to be analysed within a Markov chain setting. We prove that transformed perturbation produces non explosive paths and that solutions are stationary and ergodic with bounded moments to which sample moments of the process converge. Finally we demonstrate that our method is very accurate within the setting of fast and accessible solution methods. We provide a detailed analysis and comparison with both first order perturbation and pruning for two nonlinear DSGE models in which higher order perturbation is infeasible.

Chapter 2

Stationarity, Ergodicity and Mixing of Resetting Time Series

2.1 Introduction

Since the popularisation of linear time series models such as the autoregressive moving average model (Box et al., 1970) for level modelling and the autoregressive conditional heteroskedasticity (Engle, 1982) model for volatility modelling much innovation has been made. Present-day, fitting time series with nonlinear models has become increasingly common. A selection of such nonlinear methods that can be applied to both level and volatility modelling are regime switching models (Hamilton, 1989), threshold models (Tong and Lim, 1980; Zakoian, 1994) and score driven models (Creal et al., 2013; Harvey, 2013).

Stability properties of both data generated by a model and unobserved parameters when filtering data are very useful. Knowing when a time series is stationary ergodic with mixing properties allows one to apply limit theorems to obtain consistency and asymptotic normality of estimators. However, ensuring stability of econometric models becomes increasingly harder as the models get more complicated. Nonlinear dynamics imply that the theory on Lyapunov exponents as developed in Bougerol and Picard (1992a,b) cannot be used. Therefore one has to resort to more involved methods such as Markov chain theory and geometric ergodicity (Meyn and Tweedie, 2012) or stochastic recurrence equation theory as developed in Bougerol (1993) and Straumann (2005). These methods ensure

CHAPTER 2. STATIONARITY, ERGODICITY AND MIXING OF RESETTING TIME SERIES

stability by imposing that the updating function satisfies certain contraction or bounded growth or drift properties. See for example Cline and Pu (1999) and Saïdi and Zakoian (2006) or Blasques et al. (2014) and Straumann (2005) for the application of these conditions to various models.

This paper derives general stability conditions for a large collection of models that are typically either infeasible or less efficient (in the sense of a smaller parameter space) to analyse with the existing methods. This collection of models is characterised by a property that we denote resetting, which requires that the model has a positive probability to update to a fixed, but possibly stochastic, state, irrespective of its past values. The resetting condition allows for very wild sample path behaviour between resetting times, as the reset ensures that the sample path will return to a stable base line. That means that we can include typically unstable dynamics such as explosive or very discontinuous updates in the time series. The framework lends itself naturally to regime switching models, where we are then free to make all but one regime as unstable as we want as long as we ensure that the last regime enforces a reset.

The resetting condition might appear to be restrictive at first, but is often satisfied in time series where sudden drops or increases are observed. Typical examples of such time series are stocks exhibiting financial bubbles, where the crash of the bubble is the moment where the time series resets. See for example the model in Blasques et al. (2018b) and Chapter 3 of this thesis that is developed to describe the Bitcoin/USD exchange rate studied in Hencic and Gouriéroux (2015). There the exchange rate X_t is modelled as the sum of a stationary ergodic process μ_t and a nonnegative bubble process b_t , where

$$b_t = (\omega + \alpha b_{t-1}) \mathbb{1}\{b_{t-1} < k(\mu_t - c)\}$$

with $\omega, \alpha > 0$ and $k, c \in \mathbb{R}$. This model consists out of two regimes: one autoregressive regime $b_t = \omega + \alpha b_{t-1}$ and one collapsing regime $b_t = 0$. The bubble process b_{t-1} is nonnegative, so if the innovation $\mu_t < c$, then the indicator function does not hold for any possible value of b_{t-1} and hence the bubble process will collapse/reset regardless of its past values. Note that the stability conditions allow the autoregressive parameter α to be greater than one in this model, in fact this is encouraged to describe bubble behaviour.

This is something that is normally associated with unstable behaviour in autoregressive processes.

An example of collapsing, and thus resetting, volatility dynamics can be found in a model used by Saïdi and Zakoian (2006) to study the real financial time series discussed in Saïdi (2003). They define the dynamics of a heteroskedastic time series $(\epsilon_t)_{t \in \mathbb{Z}}$ as

$$\begin{aligned}\epsilon_t &= \sigma_t \eta_t, \\ \sigma_t^2 &= \omega + \alpha \epsilon_{t-1}^2 \mathbb{1} \{ \epsilon_{t-1}^2 > k \epsilon_{t-2}^2 \},\end{aligned}\tag{2.1}$$

where $(\eta_t)_{t \in \mathbb{Z}}$ is a strictly stationary and ergodic sequence of random variables, the parameters α and k are nonnegative and ω is positive. Similarly to the previous example the parameter α is allowed to be greater than one and model (2.1) consists of two regimes. One regime is the traditional ARCH(1) update $\sigma_t^2 = \omega + \alpha \epsilon_{t-1}^2$ and the other is the conditional homoskedastic model $\sigma_t^2 = \omega$. The model changes from the constant volatility regime to the ARCH(1) specification when the relative variation $\epsilon_{t-1}^2 / \epsilon_{t-2}^2$ becomes large, indicating a setting in which it is more likely for the volatility to be time varying. The collapse condition is harder to discern in this model, but occurs when two consecutive η 's are much smaller than their predecessors. Saïdi and Zakoian (2006) analyse the stability of their model (2.1) using Markov chain theory. They show the existence of a stationary and β -mixing solution, under the condition that the distribution of the underlying process $(\eta_t)_{t \in \mathbb{Z}}$ is independent and identically distributed (iid), has strict positive density and fixed moments $\mathbb{E}(\eta_t) = 0$ and $\mathbb{E}(\eta_t^2) = 1$. Using our approach we can show the existence of a unique, stationary and ergodic or φ -mixing solution, to which any sample path converges. The assumptions needed to get these results are less strict than those imposed in Saïdi and Zakoian (2006). Our method also allows for extensions of the model with minimum additional theoretical work.

The rest of the paper is structured as follows. Section 2.2 discusses stability conditions for random functions on separable Banach spaces and states our results in their most general form. Section 2.3 illustrates how to apply the theory to a practical model by considering a generalisation of model (2.1) and deriving the stability conditions for various distributional assumptions. Moreover, we showcase the ease of application by deriving the conditions for various practical examples including leverage effects and robust news

impact curves and also show how the method can be used to derive moment bounds.

2.2 Stability results

In this section we prove our main results for resetting time series. Theorem 2.2.1 shows the existence of a stationary ergodic solution to which all sample paths converge and Theorem 2.2.4 discusses the existence of a solution that is φ -mixing at geometric rate. We base our treatment of stability on *stochastic recurrence equations* (SREs) as is done in Straumann (2005) and Straumann and Mikosch (2006). The main advantage of stochastic recurrence equation (SRE) techniques is that they are very general. For example, proposition 7.6 in Kallenberg (2002) proves that any homogeneous Markov chain can be seen as a solution to a SRE. We refer the reader to Diaconis and Freedman (1999) for a thorough overview of SREs.

We will work with random elements on Banach spaces. This allows us to describe time varying variables in econometric models as functions of the model parameters, which can be used to obtain stronger inference results as is done in Straumann and Mikosch (2006). Let S be a closed subset of a separable Banach space equipped with a norm $\|\cdot\|$ and Borel sigma-algebra $\mathcal{B}(S)$ and let (E, \mathcal{E}) be a measurable space. Let $(\eta_t)_{t \in \mathbb{Z}}$ be a sequence of stochastic elements taking values in E and let $\phi : S \times E \rightarrow S$ be a measurable map. Then we can define a sequence of random functions $(\phi_t)_{t \in \mathbb{Z}}$ by setting $\phi_t := \phi(\cdot, \eta_t)$. Let T be either \mathbb{Z} or \mathbb{N} . A stochastic process $(X_t)_{t \in T}$ taking values in S that satisfies

$$X_{t+1} = \phi_t(X_t) \quad \forall t \in T \quad (2.2)$$

is said to be a *solution* to the SRE associated with $(\phi_t)_{t \in \mathbb{Z}}$ if $T = \mathbb{Z}$, and a *partial solution* if $T = \mathbb{N}$. We now construct a specific possible solution $(Y_t)_{t \in \mathbb{Z}}$ to (2.2) by using the backward iterates defined as $\phi_t^{(0)} = \text{Id}_S$ and

$$\phi_t^{(m)} = \phi_t \circ \phi_{t-1} \cdots \circ \phi_{t-m+1}, \quad m \in \mathbb{N}.$$

Let $x \in S$ be an element such that

$$Y_{t+1} := \lim_{m \rightarrow \infty} \phi_t^{(m)}(x) \quad (2.3)$$

exists almost surely for all $t \in \mathbb{Z}$. Bougerol (1993) and Straumann and Mikosch (2006) show that this is the case under appropriate regularity conditions involving the contracting behavior of each ϕ_t and the distribution of $(\phi_t)_{t \in \mathbb{Z}}$. Moreover, they show that the sequence of limits $(Y_t)_{t \in \mathbb{Z}}$ is then the unique ergodic solution to (2.2) and that any partial solution converges to this unique one at a geometric rate as $t \rightarrow \infty$. In this article we pursue a similar approach, we also focus on the limit of the backward iterates in (2.3), show that it is well defined and that the resulting sequence $(Y_t)_{t \in \mathbb{Z}}$ possesses the right properties. However, we rely on considerably different conditions and replace the contraction condition in Bougerol (1993) with a new resetting condition.

Assumption A. The sequence $(\phi_t)_{t \in \mathbb{Z}}$ satisfies the following conditions:

- A1. The function ϕ is $\mathcal{B}(S) \times \mathcal{E}/\mathcal{B}(S)$ measurable.
- A2. The sequence $(\eta_t)_{t \in \mathbb{Z}}$ is strictly stationary ergodic.
- A3. There exists an $M \in \mathbb{N}$ and an event $A \in \mathcal{E}^M$ such that $(\eta_t, \eta_{t-1}, \dots, \eta_{t-M+1}) \in A$ with positive probability and

$$(\eta_t, \eta_{t-1}, \dots, \eta_{t-M+1}) \in A \quad \Rightarrow \quad \phi_t^{(M)}(x) = \phi_t^{(M)}(y) \quad \forall x, y \in S.$$

Condition A1 is rather weak and designed to ensure that backward iterates of $(\phi_t)_{t \in \mathbb{Z}}$ evaluated at any point $x \in S$ are proper random variables in S . Condition A2 is common in the literature on SREs, note that it is less strict than assuming that the sequence $(\eta_t)_{t \in \mathbb{Z}}$ is independent and identically distributed. An in depth discussion on stationarity and ergodicity can, for example, be found in chapter one of Krengel (1985). Condition A3 is the resetting condition and states that there exists an event over M periods that guarantees that the corresponding M 'th iterate is constant over S , but not necessarily over E^M . This implies that $\phi_t^{(M)}$ is constant for a given realisation of $(\eta_t)_{t \in \mathbb{Z}}$ in A and thus resets to one

CHAPTER 2. STATIONARITY, ERGODICITY AND MIXING OF RESETTING TIME SERIES

constant value, irrespective of its argument and hence the past values of a solution to the SRE.

Under Assumption A we can prove that the limit of the backward iterates (2.3) exists, by showing that the sequence of backward iterates $(\phi_t^{(m)}(x))_{m \in \mathbb{N}}$ is almost surely eventually constant. The proof relies on the fact that events of positive probability occur infinitely often over time in a strictly stationary ergodic sequence. Therefore the event $(\eta_t, \eta_{t-1}, \dots, \eta_{t-M+1}) \in A$ occurs for infinitely many $t \in \mathbb{Z}$ and thus the limit of the backward iterates trivially exists. Uniqueness and convergence of paths follow from the same observation, since any two paths in model (2.2) will coincide at all such t , and therefore must be the same (eventually).

Theorem 2.2.1. *Let Assumption A hold and $x \in S$. Then the sequence $(\phi_t^{(m)}(x))_{m \in \mathbb{N}}$ is almost surely eventually constant for all $t \in \mathbb{Z}$. Consequently, $(Y_t)_{t \in \mathbb{Z}}$ is well defined, strictly stationary ergodic and the unique solution to (2.2). Moreover, for any partial solution $(\tilde{Y}_t)_{t \in \mathbb{N}}$ and function $f : \mathbb{N} \rightarrow \mathbb{R}$ we have $\lim_{t \rightarrow \infty} f(t) \|Y_t - \tilde{Y}_t\| = 0$.*

We have to discuss some preliminary results on strictly stationary ergodic (SE) sequences before we can prove Theorem 2.2.1. One reason that SE sequences play a big role in time series analysis is that they satisfy the conditions needed for Birkhoff's ergodic theorem, Birkhoff (1931). This theorem applied to an SE sequence of real valued random variables $(X_t)_{t \in \mathbb{N}}$ states that if $\mathbb{E}|X_1| < \infty$, then almost surely

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{t=1}^n X_t = E(X_1).$$

SE sequences are also easy to manipulate to create new SE sequences. We provide two results from Straumann (2005).

Lemma 2.2.2. *Let (E, \mathcal{E}) and $(\tilde{E}, \tilde{\mathcal{E}})$ be two measurable spaces, let $(X_t)_{t \in \mathbb{Z}}$ be an SE sequence of E -valued random elements and let $f : E^{\mathbb{N}} \rightarrow \tilde{E}$ be a $\mathcal{E}^{\mathbb{N}}/\tilde{\mathcal{E}}$ measurable function. Then the sequence of \tilde{E} -valued random elements $(\tilde{X}_t)_{t \in \mathbb{Z}}$ defined as $\tilde{X}_t = f(X_t, X_{t-1}, \dots)$ is SE.*

PROOF. See proposition 2.1.1 in Straumann (2005). ■

Lemma 2.2.3. *Let (E, \mathcal{E}) be a measurable space and let $(S, \mathcal{B}(S))$ be a closed subset of a separable Banach space endowed with its Borel sigma-algebra. Let $(X_t)_{t \in \mathbb{Z}}$ be a SE sequence of E -valued random elements and let $(f_m)_{m \in \mathbb{N}}$ be a sequence of functions $E^{\mathbb{N}} \rightarrow S$ that are $\mathcal{E}^{\mathbb{N}}/\mathcal{B}(S)$ measurable. Suppose that there exists a $t \in \mathbb{Z}$ such that*

$$\lim_{m \rightarrow \infty} f_m(X_t, X_{t-1}, \dots)$$

exists almost surely. Then there exists a function $f : E^{\mathbb{N}} \rightarrow S$ that is $\mathcal{E}^{\mathbb{N}}/\mathcal{B}(S)$ measurable and satisfies

$$\tilde{X}_t := \lim_{m \rightarrow \infty} f_m(X_t, X_{t-1}, \dots) = f(X_t, X_{t-1}, \dots)$$

for all $t \in \mathbb{Z}$. Moreover, the sequence of S -valued random elements $(\tilde{X}_t)_{t \in \mathbb{Z}}$ is SE.

PROOF. See corollary 2.1.3 in Straumann (2005). ■

PROOF OF THEOREM 2.2.1. Fix a $t \in \mathbb{Z}$. We begin by proving that $(\phi_t^{(m)}(x))_{m \in \mathbb{N}}$ is almost surely eventually constant. Define for $s \geq 0$,

$$I_s = \mathbb{1}\{(\eta_{t-s}, \eta_{t-s-1}, \dots, \eta_{t-s-M+1}) \in A\}.$$

The sequence $(I_s)_{s \geq 0}$ is SE by Lemma 2.2.2. Then, by Birkhoff's ergodic theorem, almost surely

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{s=0}^{n-1} I_s = E(I_0) = \mathbb{P}((\eta_t, \eta_{t-1}, \dots, \eta_{t-M+1}) \in A) > 0.$$

This implies that the event $I_s = 1$ occurs almost surely for infinitely many $s \geq 0$. Therefore we can choose the smallest such s , note that it is a random variable, and conclude that

$$\phi_t^{(m)}(x) = \phi_t^{(s)}\left(\phi_{t-s}^{(m-s)}(x)\right) = \phi_t^{(s)}\left(\phi_{t-s}^{(M)}\left(\phi_{t-s-M}^{(m-s-M)}(x)\right)\right) = \phi_t^{(s)}\left(\phi_{t-s}^{(M)}(x)\right)$$

for all $m \geq s + M$. It follows by Lemma 2.2.3 that the sequence $(Y_t)_{t \in \mathbb{Z}}$ is well defined

CHAPTER 2. STATIONARITY, ERGODICITY AND MIXING OF RESETTING TIME SERIES

and SE. Moreover, for $s = 0$ we get $Y_{t+1} = \phi_t^{(M)}(x) = \phi_t^{(M)}(Y_{t-M+1}) = \phi_t(Y_t)$, while for $s \geq 1$ we have

$$\begin{aligned} Y_{t+1} &= \lim_{m \rightarrow \infty} \phi_t^{(m)}(x) = \phi_t^{(s)}\left(\phi_{t-s}^{(M)}(x)\right) = \phi_t\left(\phi_{t-1}^{(s-1)}\left(\phi_{t-s}^{(M)}(x)\right)\right) \\ &= \phi_t\left(\lim_{m \rightarrow \infty} \phi_{t-1}^{(m)}(x)\right) = \phi_t(Y_t). \end{aligned}$$

Therefore $(Y_t)_{t \in \mathbb{Z}}$ is a solution to (2.2). If $(X_t)_{t \in \mathbb{Z}}$ is any other solution to (2.2), then

$$X_{t+1} = \phi_t^{(s)}\left(\phi_{t-s}^{(M)}(X_{t-s-M+1})\right) = \phi_t^{(s)}\left(\phi_{t-s}^{(M)}(Y_{t-s-M+1})\right) = Y_{t+1},$$

and hence it is identical to $(Y_t)_{t \in \mathbb{Z}}$.

It remains to prove the final statement. Similarly as before, we can almost surely find an $s > M - 1$ such that $(\eta_s, \eta_{s-1}, \dots, \eta_{s-M+1}) \in A$ and thus

$$Y_{t+1} = \phi_t^{(t-s)}\left(\phi_s^{(M)}(Y_{t-s-M+1})\right) = \phi_t^{(t-s)}\left(\phi_s^{(M)}(\tilde{Y}_{t-s-M+1})\right) = \tilde{Y}_{t+1}$$

for all $t \geq s$. We conclude that

$$\lim_{t \rightarrow \infty} f(t) \|Y_t - \tilde{Y}_t\| = 0,$$

irrespective of the function f , because $\|Y_t - \tilde{Y}_t\|$ is almost surely eventually zero. \blacksquare

A consequence of Theorem 2.2.1 is that we can derive sufficient conditions for the process $(Y_t)_{t \in \mathbb{Z}}$ to be φ -mixing. Let $(X_t)_{t \in \mathbb{Z}}$ be a stationary process and let \mathcal{F}_s^t , for $-\infty \leq s < t \leq \infty$, denote the sigma algebra generated by $(X_s, X_{s+1}, \dots, X_t)$. Then the φ -mixing coefficients for $(X_t)_{t \in \mathbb{Z}}$ are given by

$$\varphi_X(t) = \sup_{C \in \mathcal{F}_{-\infty}^0, D \in \mathcal{F}_t^\infty, \mathbb{P}(C) > 0} |\mathbb{P}(D|C) - \mathbb{P}(D)|$$

and the process is called φ -mixing if $\varphi_X(t) \rightarrow 0$ as $t \rightarrow \infty$.

Theorem 2.2.4. *Suppose Assumption A holds and that additionally $(\eta_t)_{t \in \mathbb{Z}}$ is φ -mixing with geometric rate. Then $(Y_t, \eta_t)_{t \in \mathbb{Z}}$ is φ -mixing with geometric rate.*

The proof will depend on Theorem 2.2.1 as follows. Usually Y_{t+1} depends on the entire past $(\eta_t, \eta_{t-1}, \dots)$. However, if the event $(\eta_{t-s}, \eta_{t-s-1}, \dots, \eta_{t-s-M+1}) \in A$ occurs for some $s \geq 0$, then $Y_{t+1} = \phi_t^{(t-s)}(\phi_s^{(M)}(x))$ and thus Y_{t+1} depends only on $(\eta_t, \dots, \eta_{t-s-M+1})$. Therefore it will be enough to show that the probability that s is large vanishes at a geometric rate. To show this we need the following two lemma's.

Lemma 2.2.5. *Let (E, \mathcal{E}) and $(\tilde{E}, \tilde{\mathcal{E}})$ be two measurable spaces, let $(X_t)_{t \in \mathbb{Z}}$ be a sequence of E -valued random elements that is φ -mixing (with geometric rate). For a $m \in \mathbb{N}$ we denote $f : E^m \rightarrow \tilde{E}$ to be a $\mathcal{E}^m / \tilde{\mathcal{E}}$ measurable function. Then the sequence of \tilde{E} -valued random elements $(\tilde{X}_t)_{t \in \mathbb{Z}}$ defined as*

$$\tilde{X}_t = f(X_t, \dots, X_{t-m})$$

is φ -mixing (with geometric rate).

PROOF. The sigma-algebra generated by $(\dots, \tilde{X}_{-1}, \tilde{X}_0)$ is contained in the sigma-algebra generated by (\dots, X_{-1}, X_0) . Similarly, the sigma-algebra generated by $(\tilde{X}_t, \tilde{X}_{t+1}, \dots)$ is contained in the sigma-algebra generated by $(X_{t-m}, X_{t-m+1}, \dots)$. Therefore $\varphi_{\tilde{X}}(t) \leq \varphi_X(t-m)$ for all $t \geq m$. \blacksquare

Lemma 2.2.6. *Let (E, \mathcal{E}) be a measurable space and let $(X_i)_{i \in \mathbb{Z}}$ be a strictly stationary sequence of E -valued random elements that is φ -mixing. Then for any $B \in \mathcal{E}$ such that $\mathbb{P}(X_1 \notin B) < 1$, we have*

$$\mathbb{P}\left(\bigcap_{i=1}^t \{X_i \notin B\}\right) \rightarrow 0 \quad \text{as} \quad t \rightarrow \infty$$

at a geometric rate.

PROOF. For a real number $z \in \mathbb{R}$ we write $\lfloor z \rfloor$ to denote the largest integer that is not larger than z . Also, we use the $;$ symbol to denote joint probabilities. For any integer

$k \leq t$ we have

$$\begin{aligned}
 & \mathbb{P}(X_t \notin B; \dots; X_1 \notin B) \\
 &= \prod_{i=0}^{\lfloor t/k \rfloor - 1} \mathbb{P}(X_{t-ik} \notin B; \dots; X_{t-(i+1)k+1} \notin B \mid X_{t-(i+1)k} \notin B; \dots; X_1 \notin B) \\
 &\leq \prod_{i=0}^{\lfloor t/k \rfloor - 1} \mathbb{P}(X_{t-ik} \notin B \mid X_{t-(i+1)k} \notin B; \dots; X_1 \notin B) \\
 &\leq \prod_{i=0}^{\lfloor t/k \rfloor - 1} \mathbb{P}(X_{t-ik} \notin B) + \varphi_X(k) \\
 &= (\mathbb{P}(X_1 \notin B) + \varphi_X(k))^{\lfloor t/k \rfloor - 1}.
 \end{aligned}$$

Choose k big enough such that $\mathbb{P}(X_1 \notin B) + \varphi_X(k) < 1$, which can be done since $\varphi_X(k) \rightarrow 0$ as $k \rightarrow \infty$. Note that if any of the events that we conditioned on has probability zero, then the lemma follows immediately. \blacksquare

PROOF OF THEOREM 2.2.4. For $-\infty \leq s < t \leq \infty$ we write \mathcal{G}_s^t to denote the sigma algebra generated by $(\eta_s, \eta_{s+1}, \dots, \eta_t)$ and \mathcal{H}_s^t to denote the sigma algebra generated by $((Y_s, \eta_s), (Y_{s+1}, \eta_{s+1}), \dots, (Y_t, \eta_t))$. Fix $t \in \mathbb{N}$ and let $s \geq 0$ again be the random variable that denotes the smallest number such that $(\eta_{t-s}, \eta_{t-s-1}, \dots, \eta_{t-s-M+1}) \in A$ occurs. Then we have $Y_{t+1} = \lim_{m \rightarrow \infty} \phi_t^{(m)}(x) = \phi_t^{(t-s)}(\phi_s^{(M)}(x))$. Therefore, for a $B \in \mathcal{B}(S)$ and a $k \geq 0$ the event $\{Y_{t+1} \in B; s \leq k\} \in \mathcal{G}_{t-k-M+1}^t$, since $\{s \leq k\} \in \mathcal{G}_{t-k-M+1}^t$. Similarly, for any $D \in \mathcal{H}_t^\infty$ the event $D \cap \{s \leq t/2 - M + 1\} \in \mathcal{G}_{\lceil t/2 \rceil}^\infty$, where we write $\lceil z \rceil$ to denote the smallest integer that is not smaller than z . It follows for $C \in \mathcal{H}_{-\infty}^0 \subseteq \mathcal{G}_{-\infty}^0$, by partitioning on $s \leq t/2 - M + 1$ and its complement, that

$$|\mathbb{P}(D|C) - \mathbb{P}(D)| \leq \varphi_\eta(\lceil t/2 \rceil) + \mathbb{P}(D; s > t/2 - M + 1|C) + \mathbb{P}(D; s > t/2 - M + 1).$$

Since $\{s > t/2 - M + 1\} \in \mathcal{G}_{\lceil t/2 \rceil}^t$ we get

$$\mathbb{P}(D; s > t/2 - M + 1|C) \leq \mathbb{P}(s > t/2 - M + 1|C) \leq \mathbb{P}(s > t/2 - M + 1) + \varphi_\eta(\lceil t/2 \rceil).$$

It follows that

$$\varphi_{(Y,\eta)}(t) \leq 2\varphi_\eta(\lceil t/2 \rceil) + 2\mathbb{P}(s > t/2 - M + 1).$$

The first term goes geometrically fast to zero by assumption. For the second part we define $X_t = (\eta_t, \eta_{t-1}, \dots, \eta_{t-M+1})$. Then $(X_t)_{t \in \mathbb{Z}}$ is φ -mixing by Lemma 2.2.5. Therefore, by Lemma 2.2.6, and the fact that $\mathbb{P}(X_t \in A) > 0$, we have

$$\mathbb{P}(s > t/2 - M + 1) = \mathbb{P}\left(\bigcap_{i=0}^{\lceil t/2 \rceil} \{X_{t-i} \notin A\}\right) \rightarrow 0$$

geometrically fast as $t \rightarrow \infty$. ■

2.3 Application to heteroscedastic volatility modelling.

We now introduce a general nonlinear ARCH model that contains the model of Saïdi and Zakoian (2006) and illustrate how to apply our main results of Section 2.2. Let $u : \mathbb{R}^2 \rightarrow [0, \infty)$ be a nonnegative Borel measurable function that possibly depends on a vector of parameters θ that lie in a parameter space Θ . The general model of interest is given by

$$\begin{aligned} \epsilon_t &= \sigma_t \eta_t, \\ \sigma_t^2 &= \omega + u(\epsilon_{t-1}, \sigma_{t-1}^2; \theta) \mathbb{1} \{ \epsilon_{t-1}^2 > k \epsilon_{t-2}^2 \}, \end{aligned} \tag{2.4}$$

where ω and k are strictly positive. The generalisation compared to (2.1) is that we replace the term $\alpha \epsilon_{t-1}^2$ with a general updating function u . We discuss model (2.1) and other examples in Section 2.3.1.

We start by analysing the dynamics concerning the time varying volatility. Given that u is nonnegative we immediately see that any possible solution to (2.4) must satisfy $\sigma_t^2 \in I := [\omega, \infty)$. Assuming that the model is well specified, we get

$$\sigma_t^2 = \omega + \tilde{u}(\sigma_{t-1}^2, \eta_{t-1}; \theta) \mathbb{1} \{ \sigma_{t-1}^2 \eta_{t-1}^2 > k \sigma_{t-2}^2 \eta_{t-2}^2 \}, \tag{2.5}$$

CHAPTER 2. STATIONARITY, ERGODICITY AND MIXING OF RESETTING TIME SERIES

where $\tilde{u}(\sigma_{t-1}^2, \eta_{t-1}; \theta) = u(\epsilon_{t-1}(\sigma_{t-1}, \eta_{t-1}), \sigma_{t-1}^2; \theta)$. Our analysis will focus on this model, since any solution to (2.5) can be used to create a solution to (2.4). Note that σ_t^2 depends both on σ_{t-1}^2 and σ_{t-2}^2 . The random functions $(\phi_t)_{t \in \mathbb{Z}}$ associated with (2.5) will therefore be defined on I^2 and are given by

$$\phi_{t-1}(x, y) = \phi(x, y, \eta_{t-1}, \eta_{t-2}) = \omega + \tilde{u}(x, \eta_{t-1}; \theta) \mathbb{1} \{x\eta_{t-1}^2 > ky\eta_{t-2}^2\}.$$

Unfortunately these are not in the framework of the SRE theory in Section 2.2, since $\phi : I^2 \times \mathbb{R}^2 \rightarrow I$. Therefore we will look at the two dimensional model

$$(\sigma_t^2, \sigma_{t-1}^2) = (\phi_{t-1}(\sigma_{t-1}^2, \sigma_{t-2}^2), \sigma_{t-1}^2), \quad (2.6)$$

which has state space $S := I^2$. The random functions associated with (2.6) are given by

$$\psi_{t-1}(x, y) = \psi(x, y, \eta_{t-1}, \eta_{t-2}) = (\phi_{t-1}(x, y), x).$$

Define $\phi_t^{(-1)}(x, y) = y$ and $\phi_t^{(0)}(x) = x$, then the backward iterates for $m \in \mathbb{N}$ are given by

$$\begin{aligned} \phi_t^{(m)}(x, y) &= \phi_t \left(\phi_{t-1}^{(m-1)}(x, y), \phi_{t-2}^{(m-2)}(x, y) \right), \\ \psi_t^{(m)}(x, y) &= \left(\phi_t^{(m)}(x, y), \phi_{t-1}^{(m-1)}(x, y) \right). \end{aligned}$$

We now state the weakest assumption for our nonlinear ARCH model that ensures we satisfy Assumption A and therefore obtain the results from Theorems 2.2.1 and 2.2.4. This result is derived in Theorem 2.3.2.

Assumption B.

B1. The sequence $(\eta_t)_{t \in \mathbb{Z}}$ is SE.

B2. The following event has positive probability of occurring:

$$\eta_t^2 \leq \inf_{x \in I} \frac{kx\eta_{t-1}^2}{\omega + \tilde{u}(x, \eta_{t-1}^2; \theta)} \quad \text{and} \quad \eta_{t-1}^2 \leq \inf_{x \in I} \frac{kx\eta_{t-2}^2}{\omega + \tilde{u}(x, \eta_{t-2}^2; \theta)}. \quad (2.7)$$

2.3. APPLICATION TO HETEROSCEDASTIC VOLATILITY MODELLING.

Assumption B is very general, but quite complex and thus hard to interpret. It is a restriction on the joint probability law of $(\eta_t, \eta_{t-1}, \eta_{t-2})$ that confines η_t and η_{t-1} with positive probability to an area described by the functions in (2.7). This area can be abstract and depends on the parameters k and θ . In what follows we derive a condition that is easier to verify than Assumption B2 by only focussing on this area close to the origin. Note that if η_t and η_{t-1} given η_{t-2} can be arbitrarily small with positive probability, then Assumption B2 is satisfied if the infima are nonzero. To that end we define the function

$$g(\eta; \theta) := \sup_{x \in I} \frac{\tilde{u}(x, \eta; \theta)}{x}.$$

Assumption C.

C1. For all $\eta \in \mathbb{R}$ and $\theta \in \Theta$ we have $g(\eta; \theta) < \infty$.

C2. The sequence $(\eta_t)_{t \in \mathbb{Z}}$ is SE.

C3. There exist a $N \in \mathbb{N}$ such that $\mathbb{P}(|\eta_t| < 1/n; |\eta_{t-1}| < 1/m \mid \eta_{t-2}) > 0$ almost surely for all $n, m \geq N$. Also the probability that $\eta_t = 0$ is zero.

Assumption C1 is an assumption on the updating function u of model (2.4). The condition is of a similar nature as those found in theory on geometric ergodicity of nonlinear time series, see Cline and Pu (1999). It implies that the function \tilde{u} as a function of x is bounded on any closed interval, and asymptotically as $x \rightarrow \infty$ is bounded by a linear function. These two facts ensure that the infima in (2.7) are nonzero.

The other conditions are purely on the distribution of $(\eta_t)_{t \in \mathbb{Z}}$. Assumption C3 entails that η_t and η_{t-1} have positive probability of being arbitrarily small, independent of the value of η_{t-2} . An example on how Assumption C3 can be derived is if $(\eta_t)_{t \in \mathbb{Z}}$ is obtained as a SE solution from another model. For example, suppose that $(\eta_t)_{t \in \mathbb{Z}}$ is given by a SE solution to an autoregressive process of order one

$$\eta_{t+1} = \beta\eta_t + \zeta_t.$$

Then a sufficient condition would be that $(\zeta_t)_{t \in \mathbb{Z}}$ is iid, that ζ_t is absolutely continuous with respect to the Lebesgue measure on \mathbb{R} and that ζ_t has a strictly positive probability

CHAPTER 2. STATIONARITY, ERGODICITY AND MIXING OF RESETTING TIME SERIES

density function. Note that these conditions imply that any set in $B \in \mathcal{B}(\mathbb{R} \times \mathbb{R} \times \mathbb{R})$ has $\mathbb{P}((\eta_t, \eta_{t-1}, \eta_{t-2}) \in B) > 0$, so in particular Assumption C3 is implied.

If we can assume that the sequence $(\eta_t)_{t \in \mathbb{Z}}$ is independent, then Assumption C simplifies as follows:

Assumption D.

D1. For all $\eta \in \mathbb{R}$ and $\theta \in \Theta$ we have $g(\eta; \theta) < \infty$.

D2. The sequence $(\eta_t)_{t \in \mathbb{Z}}$ is iid.

D3. There exist a $N \in \mathbb{N}$ such that $\mathbb{P}(|\eta_t| < 1/n) > 0$ for all $n \geq N$. Also the probability that $\eta_t = 0$ is zero.

Assumption D3 implies Assumption C3 if $(\eta_t)_{t \in \mathbb{Z}}$ is iid and describes that η_t being arbitrarily small has positive probability. This, for example, is implied if η_t is absolutely continuous with respect to the Lebesgue measure on \mathbb{R} and the probability density function of η_t is strictly positive on an open interval around zero. Common distributions such as the normal and student- t distribution satisfy this condition.

Lemma 2.3.1. *Assumption C implies Assumption B.*

PROOF. We need to check whether Assumption B2 is satisfied. Assumption C1 ensures that the random variable

$$\inf_{x \in I} \frac{kx\eta_{t-1}^2}{\omega + \tilde{u}(x, \eta_{t-1}^2; \theta)}$$

is equal to zero if and only if $\eta_{t-1} = 0$, since

$$\inf_{x \in I} \frac{kx\eta_{t-1}^2}{\omega + \tilde{u}(x, \eta_{t-1}^2; \theta)} \geq \inf_{x \in I} \frac{kx\eta_{t-1}^2}{\omega + g(\eta_{t-1}; \theta)x} = \frac{k\eta_{t-1}^2}{1 + g(\eta_{t-1}; \theta)}.$$

Assumption C3 therefore implies that $k\eta_{t-1}^2/(1 + g(\eta_{t-1}; \theta))$ is nonzero with probability one. Therefore, the probability that

$$\eta_t^2 \leq \frac{k\eta_{t-1}^2}{1 + g(\eta_{t-1}; \theta)} \quad \text{and} \quad \eta_{t-1}^2 \leq \frac{k\eta_{t-2}^2}{1 + g(\eta_{t-2}; \theta)}$$

2.3. APPLICATION TO HETEROSCEDASTIC VOLATILITY MODELLING.

is greater than zero. This follows, since the infima are nonzero, due to the fact that η_t and η_{t-1} can be arbitrarily small with positive probability so in particular, they have positive probability to be smaller than these upper bounds. ■

Theorem 2.3.2. *If Assumption B holds, then there exists a solution $((\epsilon_t, \sigma_t^2))_{t \in \mathbb{Z}}$ to (2.4) given by*

$$\begin{aligned}\sigma_{t+1}^2 &= \lim_{m \rightarrow \infty} \phi_t^{(m)}(x, y), \\ \epsilon_{t+1} &= \sqrt{\lim_{m \rightarrow \infty} \phi_t^{(m)}(x, y) \eta_{t+1}}.\end{aligned}\tag{2.8}$$

This solution is stationary ergodic, unique and any partial solution converges to it at any rate. Moreover, if additionally $(\eta_t)_{t \in \mathbb{Z}}$ is φ -mixing with geometric rate, then $((\epsilon_t, \sigma_t^2))_{t \in \mathbb{Z}}$ is φ -mixing with geometric rate.

PROOF. We will start by verifying that assumptions A are all satisfied, so that Theorem 2.2.1 implies that

$$\left(\lim_{m \rightarrow \infty} \psi_t^{(m)}(x, y) \right)_{t \in \mathbb{Z}}$$

is a SE and unique solution to (2.6) such that all partial solutions converge to it. Assumption A1 is satisfied by Borel-measurability of u . Assumption A2 requires the sequence $((\eta_t, \eta_{t-1}))_{t \in \mathbb{Z}}$ to be SE, which is implied by B1 and Lemma 2.2.2. Finally, we will show that (2.7) implies that $\psi_t^{(3)}(x, y) = (\omega, \omega)$ for all $(x, y) \in S$ and therefore implies Assumption A3. Note that

$$\phi_t^{(2)}(x, y) = \phi_t(\phi_{t-1}(x, y), x) = \omega + \tilde{u}(\phi_{t-1}(x, y), \eta_t; \theta) \mathbb{1} \{ \phi_{t-1}(x, y) \eta_t^2 > kx \eta_{t-1}^2 \},$$

so that $\phi_t^{(2)}(x, y) = \omega$ for all $(x, y) \in S$ iff $\eta_t^2 \leq \frac{kx \eta_{t-1}^2}{\phi_{t-1}(x, y)}$ for all $(x, y) \in S$, which is implied by

$$\eta_t^2 \leq \inf_{x \in I} \frac{kx \eta_{t-1}^2}{\omega + \tilde{u}(x, \eta_{t-1}^2; \theta)}.$$

The first part of the proof is concluded by noting that

$$\psi_t^{(3)}(x, y) = (\phi_t^2(\phi_{t-2}(x, y), x), \phi_{t-1}^2(x, y)) = (\phi_t^2(\tilde{x}, \tilde{y}), \phi_{t-1}^2(x, y)).$$

Next, a unique and SE solution to (2.6) to which all partial solutions converge to implies the existence of a solution to (2.5) with the same properties, by projecting on the first coordinate. The found solution is given by

$$\lim_{m \rightarrow \infty} \phi_t^{(m)}(x, y),$$

which is a measurable function of $(\eta_{t-1}, \eta_{t-2}, \dots)$. Therefore $\epsilon_t = \sigma_t \eta_t$ is a measurable function of $(\eta_t, \eta_{t-1}, \dots)$ and thus (2.8) is a SE solution to (2.4) by Lemma 2.2.2. Uniqueness and convergence of partial solutions transfer directly from those properties for (2.5).

Finally, suppose $(\eta_t)_{t \in \mathbb{Z}}$ is φ -mixing with geometric rate. Then $((\eta_t, \eta_{t-1}))_{t \in \mathbb{Z}}$ is φ -mixing with geometric rate by Lemma 2.2.5 and thus

$$\left(\lim_{m \rightarrow \infty} \phi_t^{(m)}(x, y), \eta_{t+1} \right)_{t \in \mathbb{Z}}$$

is φ -mixing with geometric rate by applying Theorem 2.2.4 and Lemma 2.2.5 again. Applying Lemma 2.2.5 once more shows that (2.8) is φ -mixing with geometric rate. ■

2.3.1 Examples

This section discusses a couple of specifications of the updating function u in model (2.4). We assume that the sequence $(\eta_t)_{t \in \mathbb{Z}}$ is φ -mixing at a geometric rate and satisfies the distributional conditions of either Assumption C or Assumption D. We then display how quickly our theory can be applied by checking whether Assumption C1/D1 holds for these examples.

Example 1 (Saïdi and Zakoian (2006)). First, we consider model (2.1). We repeat it here

for readability.

$$\begin{aligned}\epsilon_t &= \sigma_t \eta_t, \\ \sigma_t^2 &= \omega + \alpha \epsilon_{t-1}^2 \mathbb{1} \{ \epsilon_{t-1}^2 > k \epsilon_{t-2}^2 \},\end{aligned}$$

where α is nonnegative. We have $u(\epsilon_{t-1}, \sigma_{t-1}^2; \alpha) = \alpha \epsilon_{t-1}^2$, which is a measurable and nonnegative function. Moreover, the function $g(\eta_t; \alpha) = \alpha < \infty$, so Assumption C1 respective D1 is immediately satisfied. Therefore there exists a strictly stationary and φ -mixing at geometric rate solution to which all partial solutions converge almost surely. Saïdi and Zakoian (2006) assume that $(\eta_t)_{t \in \mathbb{Z}}$ is iid. They then add the assumptions that η_t is absolutely continuous with respect to the Lebesgue measure on \mathbb{R} and η_t has a strictly positive probability density function. Note that this assumption is stronger than our Assumption D3. Finally, Saïdi and Zakoian (2006) assume that $\mathbb{E}\eta_t = 0$ and $\mathbb{E}\eta_t^2 = 1$, while we don't have any moment conditions at all.

Example 2 (Asymmetric news impact curve). Second, we consider a model that allows the update function to be asymmetric in ϵ_{t-1} rather than using the quadratic update ϵ_{t-1}^2 considered above. In particular, we follow Engle and Ng (1993) in using the asymmetric news impact curve $u(\epsilon_{t-1}, \sigma_{t-1}^2) = \alpha(\epsilon_{t-1} + \delta \sigma_{t-1})^2$ and obtain the following model

$$\begin{aligned}\epsilon_t &= \sigma_t \eta_t, \\ \sigma_t^2 &= \omega + \alpha(\epsilon_{t-1} + \delta \sigma_{t-1})^2 \mathbb{1} \{ \epsilon_{t-1}^2 > k \epsilon_{t-2}^2 \},\end{aligned}$$

where α is nonnegative and $\delta \in \mathbb{R}$. Notice how for $\delta < 0$, negative returns ϵ_t have greater impact on future volatility σ_{t+1}^2 than positive returns of the same magnitude, thus capturing the empirical regularity known as the *leverage effect*. In this example we have $\tilde{u}(x, \eta_t; \alpha) = \alpha x(\eta_t + \delta)^2$ and thus $g(\eta_t; \alpha) = \alpha(\eta_t + \delta)^2 < \infty$. Therefore, Assumption C1/D1 is satisfied again and thus there exists a strictly stationary and φ -mixing at geometric rate solution to which all partial solutions converge almost surely.

Example 3 (Robust volatility update). Finally, we consider a robust nonlinear ARCH model by adopting an update function that is bounded in ϵ_{t-1} rather than quadratic. In particular, we study a model which embodies the news impact curve of the student- t score volatility model introduced in Creal et al. (2011, 2013) and the beta- t EGARCH

model proposed by Harvey (2013),

$$\begin{aligned}\epsilon_t &= \sigma_t \eta_t, \quad \eta_t \sim t(\lambda) \\ \sigma_t^2 &= \omega + \alpha \frac{\epsilon_{t-1}^2}{1 + \lambda^{-1} \epsilon_{t-1}^2} \mathbb{1} \left\{ \epsilon_{t-1}^2 > k \epsilon_{t-2}^2 \right\},\end{aligned}$$

where α and λ are nonnegative. Notice that the innovations η_t are allowed to be fat tailed. In particular, they belong to the family of student's- t distributed random variables with λ degrees of freedom. The updating function of this model becomes more robust (with a lower upper bound) as $\lambda \rightarrow 0$ so that the innovations η_t become fatter tailed and outliers become more frequent. In contrast, as we approach the Gaussian case by letting $\lambda \rightarrow \infty$, then the updating function reverts back to that of the nonlinear ARCH model considered in Saïdi and Zakoian (2006). We now have $\tilde{u}(x, \eta_t; \alpha, \lambda) = \alpha \frac{x \eta_t^2}{1 + x \eta_t^2 / \lambda} \leq \alpha \lambda$, thus $g(\eta_t; \alpha, \lambda) \leq \alpha \lambda / \omega < \infty$ and Assumption C1/D1 is satisfied again. Hence, there exists a strictly stationary and φ -mixing at geometric rate solution to which all partial solutions converge almost surely.

2.3.2 Moments

Moment conditions for model (2.4) can be obtained by showing that the moments of the backward iterates have a converging subsequence. To state our result we define

$$h(\eta; \theta) = \limsup_{x \rightarrow \infty} \frac{\tilde{u}(x, \eta; \theta)}{x}.$$

Theorem 2.3.3. *Let Assumption D hold. Let $p \geq 1$ and $\tilde{\Theta} \subseteq \Theta$ be such that $\mathbb{E}|\eta_t|^{2p} < \infty$ and $\mathbb{E}g(\eta_t; \theta)^p < \infty$ and*

$$\mathbb{E} \left(h(\eta_t; \theta) h(\eta_{t-1}; \theta) \mathbb{1} \left\{ \eta_t^2 > \frac{k \eta_{t-1}^2}{h(\eta_{t-1}; \theta)} \right\} \right)^p < 1 \quad (2.9)$$

for all $\theta \in \tilde{\Theta}$. Then the unique solution to (2.8) has finite absolute $2p$ 'th moment, that is $\mathbb{E}|\epsilon_t|^{2p} < \infty$ and $\mathbb{E}\sigma_t^{2p} < \infty$.

Theorem 2.3.3 is a generalisation of Theorem 3.3 in Saïdi and Zakoian (2006), their assumption to ensure moments in model (2.1) follows as a specific case from our result.

The expectation in condition (2.9) can be hard to calculate, because of the indicator function.

Corollary 2.3.4. *Condition (2.9) is implied by*

$$\mathbb{E}h(\eta_t; \theta)^p < 1. \quad (2.10)$$

PROOF. This follows directly from Assumption D2 and the fact that the indicator function is bounded by one. ■

Condition (2.10) is much easier to calculate, but sacrifices flexibility by ignoring the indicator function. Saïdi and Zakoian (2006) show that (2.9) delivers more flexible bounds for model (2.1) than (2.10) when $\eta_t \sim N(0, 1)$. We will discuss the examples of Section 2.3.1 to illustrate how both conditions can be useful.

PROOF OF THEOREM 2.3.3. By Assumption D2 we have $\mathbb{E}|\epsilon_t|^{2p} = \mathbb{E}|\eta_t|^{2p}\mathbb{E}\sigma_t^{2p}$, so we only have to show $\mathbb{E}\sigma_t^{2p} < \infty$. We know by theorem 2.3.2 that

$$\sigma_t^2 = \lim_{m \rightarrow \infty} \phi_t^{(m)}(x, y),$$

so by continuity of the norm and Fatou's lemma we have $\mathbb{E}\sigma_t^{2p} < \infty$ if

$$\liminf_{m \rightarrow \infty} \mathbb{E} \left| \phi_t^{(m)}(x, y) \right|^p < \infty. \quad (2.11)$$

We will prove inequality (2.11). To ease notation we will write $\phi_t^m = \phi_t^{(m)}(x, y)$ and suppress the dependence of the functions g and h on θ . We have

$$\begin{aligned} \phi_t^m &= \omega + \tilde{u}(\phi_{t-1}^{m-1}, \eta_{t-1}) \mathbb{1} \left\{ \phi_{t-1}^{m-1} \eta_{t-1}^2 > k \phi_{t-2}^{m-2} \eta_{t-2}^2 \right\} \\ &\leq \omega + g(\eta_{t-1}) \phi_{t-1}^{m-1} \\ &\leq \omega + g(\eta_{t-1})(\omega + g(\eta_{t-2}) \phi_{t-2}^{m-2}) \end{aligned}$$

Let $n \in \mathbb{N}$ be any integer. We separate the problem into three scenarios. Suppose $\phi_{t-1}^{m-1} \leq$

n , then ϕ_t^m is bounded by

$$\omega + g(\eta_{t-1})n. \quad (2.12)$$

If $\phi_{t-2}^{m-2} \leq n$, then ϕ_t^m is bounded by

$$\omega + g(\eta_{t-1})(\omega + g(\eta_{t-2})n). \quad (2.13)$$

Finally, suppose $\phi_{t-1}^{m-1}, \phi_{t-2}^{m-2} \geq n$. Define

$$h_n(\eta) = \sup_{x \geq n} \frac{\tilde{u}(x, \eta; \theta)}{x}.$$

Then, for $n \geq \omega$, we have $h_n(\eta) \leq g(\eta)$ and thus

$$\begin{aligned} \phi_t^m &\leq \omega + h_n(\eta_{t-1})\phi_{t-1}^{m-1} \mathbb{1} \left\{ \eta_{t-1}^2 > \frac{k\phi_{t-2}^{m-2}\eta_{t-2}^2}{\omega + \tilde{u}(\phi_{t-2}^{m-2}, \eta_{t-2})} \right\} \\ &\leq (1 + g(\eta_{t-1}))\omega + h_n(\eta_{t-1})h_n(\eta_{t-2})\phi_{t-2}^{m-2} \mathbb{1} \left\{ \eta_{t-1}^2 > \frac{k\eta_{t-2}^2}{\omega/n + h_n(\eta_{t-2})} \right\}. \end{aligned} \quad (2.14)$$

It follows that ϕ_t^m is bounded by the sum of (2.12)-(2.14) and therefore by independence of ϕ_{t-2}^{m-2} with η_{t-1} and η_{t-2} we get by Minkowski's inequality that

$$[\mathbb{E}(\phi_t^m)^p]^{\frac{1}{p}} \leq C(n) + [\mathbb{E}f_n(\eta_{t-1}, \eta_{t-2})^p]^{\frac{1}{p}} [\mathbb{E}(\phi_{t-2}^{m-2})^p]^{\frac{1}{p}},$$

where $C(n)$ is a finite constant depending on n and

$$f_n(\eta_{t-1}, \eta_{t-2}) = h_n(\eta_{t-1})h_n(\eta_{t-2}) \mathbb{1} \left\{ \eta_{t-1}^2 > \frac{k\eta_{t-2}^2}{\omega/n + h_n(\eta_{t-2})} \right\}.$$

A sufficient condition for (2.11) is to find an appropriate $n \in \mathbb{N}$ such that the expectation $\mathbb{E}f_n(\eta_{t-1}, \eta_{t-2})^p < 1$. This happens for any choice of n that is large enough, as implied by (2.9) and the dominated convergence theorem, because $f_n(\eta_{t-1}, \eta_{t-2})$ is bounded by $g(\eta_{t-1})g(\eta_{t-2})$ for large enough n and as $n \rightarrow \infty$ it converges pointwise to

$$h(\eta_{t-1})h(\eta_{t-2})\mathbb{1}\left\{\eta_{t-1}^2 > \frac{k\eta_{t-2}^2}{h(\eta_{t-2})}\right\}.$$

■

Example 1 (Saïdi and Zakoian (2006) continued). Using Theorem 2.3.3 we can follow the approach of Saïdi and Zakoian (2006) and find the same conditions for model (2.1) that ensure $\mathbb{E}|\epsilon_t|^{2p} < \infty$ and $\mathbb{E}\sigma_t^{2p} < \infty$. We need $\mu_{2p} := \mathbb{E}|\eta_t|^{2p} < \infty$ and note that it implies $\mathbb{E}g(\eta_t; \alpha)^p = \alpha^p \mu_{2p} < \infty$. In this example condition (2.9) boils down to

$$\mathbb{E}\left(\alpha^2 \eta_t^2 \eta_{t-1}^2 \mathbb{1}\left\{\eta_t^2 > \frac{k}{\alpha}\right\}\right)^p < 1. \quad (2.15)$$

Using Hölder's and Markov's inequalities we get for any $m \in \mathbb{N}$ that the expectation in (2.15) is bounded by

$$\begin{aligned} \alpha^{2p} \mu_{2p} \mathbb{E}\left(\eta_t^2 \mathbb{1}\left\{\eta_t^2 > \frac{k}{\alpha}\right\}\right)^p &\leq \alpha^{2p} \mu_{2p} \mu_{2pm}^{1/m} \mathbb{P}\left(\eta_t^{2m} > \left(\frac{k}{\alpha}\right)^m\right)^{\frac{m-1}{m}} \\ &\leq \alpha^{2p} \mu_{2p} \mu_{2pm}^{1/m} \mu_{2m}^{(m-1)/m} \left(\frac{\alpha}{k}\right)^{m-1} \end{aligned}$$

Therefore a sufficient condition for (2.9) is

$$\alpha < \max_{m \in \mathbb{N}} \left(\frac{k^{m-1}}{\mu_{2p} \mu_{2pm}^{1/m} \mu_{2m}^{(m-1)/m}} \right)^{1/(2p+m-1)}.$$

Example 2 (Asymmetric news impact curve continued). The model with leverage effects requires again $\mu_{2p} < \infty$, which implies $\mathbb{E}g(\eta_t; \alpha)^p = \alpha^p \mathbb{E}(\eta_t + \delta)^{2p} \leq 2^{2p-1} \alpha^p (\mu_{2p} + |\delta|^{2p}) < \infty$. This model provides an example where the expectation in (2.9) is hard to calculate. The condition here leads to

$$\mathbb{E}\left(\alpha^2 (\eta_t + \delta)^2 (\eta_{t-1} + \delta)^2 \mathbb{1}\left\{\eta_t^2 > \frac{k\eta_{t-1}^2}{\alpha(\eta_{t-1} + \delta)^2}\right\}\right)^p < 1,$$

but we cannot easily use the Markov inequality to bound the indicator function, since this would lead to moments of the reciprocal of η_t . Instead we use (2.10) and get the sufficient

condition $\alpha < [\mathbb{E}(\eta_t + \delta)^{2p}]^{-1/p}$ to obtain $\mathbb{E}|\epsilon_t|^{2p} < \infty$ and $\mathbb{E}\sigma_t^{2p} < \infty$.

Example 3 (Robust volatility update continued). The robust model has a bounded updating function for the volatility, so therefore we immediately know that $\mu_{2p} < \infty$ is the only condition we need $\mathbb{E}|\epsilon_t|^{2p} < \infty$ and $\mathbb{E}\sigma_t^{2p} < \infty$. This result also follows from Theorem 2.3.3, since $g(\eta; \alpha, \lambda) \leq \alpha\lambda$ and $h(\eta; \alpha, \lambda) = 0$ for all $\eta \in \mathbb{R}$.

2.4 Conclusion

This chapter has introduced a new set of conditions that ensure the existence of a unique stationary, ergodic and φ -mixing solution for time series models. Moreover, sample paths are guaranteed to converge to this solution over time. The assumptions are different from existing conditions as they do not impose Lipschitz, bounded growth or drift restrictions. Instead we require that the time series contains resetting dynamics, where a reset implies that the model has a positive probability to update to a value that does not depend on the past. These dynamics are present in time series with sudden changes, such as stock prices with financial bubbles. We have demonstrated the value of our results and illustrated how to apply them by examining a generalisation of the nonlinear ARCH model studied in Saïdi and Zakoian (2006).

Chapter 3

A Time-Varying Parameter Model for Local Explosions

3.1 Introduction

Many financial and economic time series display phases of locally explosive behaviour that is followed by a burst or sharp mean-reverting dynamics. This stochastic behaviour is especially prevalent in financial asset prices, stock indices and exchange rates. The literature on rational expectations models for asset pricing typically describe the asset price process as the sum of a fundamental value process and aforementioned locally explosive process. The second process is then defined as a *speculative bubble*, see for instance Blanchard and Watson (1982), West (1987), Diba and Grossman (1988) and more. The bubble is considered to be an explosive nonstationary process and its presence is tested via unit root and cointegration tests. However, Evans (1991) noted that periodically collapsing bubbles can cause the bubble paths to look more like a stationary process, making it difficult for regular tests to detect the existence of bubbles. Using recursive testing techniques, evidence for the existence of a speculative bubble has been found by for example Phillips et al. (2011) and Homm and Breitung (2012) in the Nasdaq real price index and Phillips and Yu (2011) in the U.S. house price index, the price of crude oil and the spread between Baa and Aaa bond rates.

A different approach has been proposed by Gouriéroux and Zakoïan (2013) who describe speculative bubble dynamics using a noncausal autoregressive process of order

CHAPTER 3. A TIME-VARYING PARAMETER MODEL FOR LOCAL EXPLOSIONS

one with Cauchy innovations. This specification is able to model speculative bubbles as in reverse time the model is a causal autoregressive process of order one with a fat tailed innovation distribution, and thus produces large spikes followed by mean reversion. From the calendar time perspective such dynamics are observed as exponential explosions followed by sudden collapse. The noncausal approach to bubble modelling has been extended to stable distributed innovations in Gouriéroux and Zakoïan (2017) and to higher order mixed causal and noncausal linear models in Fries and Zakoïan (2017). A difference with the rational expectations approach is that noncausal models work within a stationary framework, which allows for the derivation of many theoretical results. Gouriéroux and Zakoïan (2013) show that the sample autocorrelation converges to a number smaller than one in absolute value. It demonstrates that a unit root test generally rejects the unit root hypothesis and thus will be unable to identify the presence of speculative bubbles. Moreover, they indicate the possibility of calculating and predicting future bubble behaviour and show the existence of moments. A major disadvantage of the noncausal approach is its computational challenge. Distinguishing causal and noncausal components is based on extreme value clustering, see the discussions in Fries and Zakoïan (2017). The prediction of these components depends on computational methods such as Metropolis-Hasting or sampling/importance resampling, see Gouriéroux and Jasiak (2016). In addition, the models are unable to distinguish the potential speculative bubble from the fundamental value. The noncausal models allow for only one generic type of bubble baseline path for a given set of parameters.

We introduce an observation driven model with time varying parameters designed as a new approach to modelling multiple speculative bubbles. As in the literature on rational expectations, our proposed model splits the asset price into a sum of two processes. The first process represents the fundamental value and can be modelled by any contracting or mean reverting process, while the second process represents the bubble effect characterised by the typical exponential increase followed by a burst. We provide various bubble burst conditions and discuss their respective merits and shortcomings. The advantage of using such a specification is that we can filter data into its fundamental value and a potential speculative bubble. Furthermore, the sum of the two processes is very flexible due to the joint dynamics of the individual components and can describe various baseline

paths for the same set of parameters. Finally, the model has a conventional observation driven specification, which implies that parameter estimation can rely on the method of maximum likelihood where the likelihood function is obtained via the prediction error decomposition. It further implies that point predictions, confidence intervals, bubble burst probabilities, bubble emergence probabilities, expected bubble life times, and more, can be derived straightforwardly.

Similar to the noncausal literature our model describes locally explosive behaviour in a strictly stationary framework. Due to earlier work in Blasques and Nientker (2017) we can immediately show that the model admits a stationary ergodic and ϕ -mixing solution under very mild conditions. Additionally, in this paper, we prove that the model as a filter also admits a stationary ergodic and mixing solution and that any initiated sample path converges to this solution. The derivations of these results are nonstandard because the filter contains a discontinuity, rendering classical contraction results such as those in Bougerol (1993) infeasible. The results are then used to obtain consistency and asymptotic normality for our maximum likelihood estimator on the parameters that enter continuously in the likelihood. In a simulation exercise we show that other parameters are well behaved.

The rest of the paper is structured as follows. Section 3.1 introduces our modelling framework for local explosions. In Section 4.3 we study probabilistic and statistical properties of the model. Evidence from simulations and a real time series are provided in Section 3.4. Concluding remarks are in Section 3.5. The proofs are presented in the Appendix.

3.2 Model for Local Explosions

Our model decomposes the asset price X_t into a sum of three elements

$$X_t = \mu_t + b_t + \varepsilon_t, \tag{3.1}$$

where μ_t is the fundamental value of the asset price, b_t is the value of a potential speculative bubble and ε_t is an error term. The error ε_t an element of an independent and

identically normal distributed sequence

$$(\varepsilon_t)_{t \in \mathbb{Z}} \sim \text{NID}(0, \sigma^2), \quad (3.2)$$

where σ is a strictly positive constant. The fundamental value μ_t is defined as the value of the asset price if no speculative bubbles were to exist. The main focus of this paper is on describing bubble dynamics. Hence we consider a basic observation driven updating equation for the fundamental value, that is

$$\mu_t = \delta + \beta\mu_{t-1} + \gamma(X_{t-1} - \mu_{t-1} - b_{t-1}), \quad (3.3)$$

where δ , β and γ are fixed unknown parameters. The dynamics for the fundamental value are mean reverting if $|\beta| < 1$, but partially correct by a factor γ for the past error $\varepsilon_{t-1} = (X_{t-1} - \mu_{t-1} - b_{t-1})$. This updating equation can be interpreted as an observation driven analogue of the parameter driven local level model in Chapter 2 of Durbin and Koopman (2012) and can also be obtained when using a score updating rule for the mean as in Creal et al. (2013). Many other dynamic processes for the fundamental value can also be considered. The model specification (3.3) can be augmented with more lags of the μ and ε processes, similar to a stationary autoregressive moving average (ARMA) process. Also, we can adopt a completely exogenous stationary process for μ_t which is potentially based on economic or financial reasoning. We will maintain the stationary framework explored in the non-causal literature. In practice, this means that one would have to add a nonstationary component when the objective is to model bubbles in non-stationary time-series, such as asset prices which are typically nonstationary. Alternatively, one could allow the fundamental value to be non-stationary, such as a random walk. This is however outside the scope of this paper.

The speculative bubble process is nonnegative and defined according to the following updating equation

$$b_t = (\omega + \alpha b_{t-1}) \mathbb{1}\{\text{survival condition}\}. \quad (3.4)$$

To ensure nonnegativity of b_t we impose $\omega > 0$, while α can be any nonnegative num-

ber, but typically is thought of as a parameter that is greater than one. This implies that the bubble process satisfies an exponential increase, as is commonly observed in locally explosive time series. The bubble b_t then diverges to infinity, if not for the indicator function, which forces the bubble to collapse down to zero if the survival condition is no longer satisfied. As with the fundamental value process, many options are available for the survival condition. Let $F_t = (X_t, \mu_t, b_t)$ be the information obtained at time t . Then a general survival condition that encompasses a variety of useful model choices is given by thresholded functions

$$\mathbb{1}\{g(F_{t-1}) < 0\}, \quad (3.5)$$

where g is some real-valued function, which we will call the *survival function*. A few example choices for the survival function are given by

- E1** $g(F_{t-1}) = X_{t-1} - c$, for some $c \in \mathbb{R}$.
- E2** $g(F_{t-1}) = X_{t-1} - \mu_{t-1} - c$, for some $c \geq 0$.
- E3** $g(F_{t-1}) = b_{t-1} - kX_{t-1}$, for some $k \in [0, 1]$.
- E4** $g(F_{t-1}) = b_{t-1} - k(\mu_t - c)$, for some $k \geq 0$ and $c \in \mathbb{R}$.

The simplest survival function **E1** lets the bubble grow until the asset price reaches a fixed level c . This allows for various bubble sizes b_t , as X_{t-1} also depends on the fundamental value μ_{t-1} and the shock process ε_{t-1} , but does describe time series in which the asset price always drops from approximately the same critical level. To allow for varying levels one can use a survival function such as **E2**. This function allows the bubble to grow as long as the difference between the asset price and the fundamental value is not too large, which leaves flexibility for the actual critical level. Examples **E1-E2** have less control for the emergence rate of bubbles. If a bubble just collapsed, then X_{t-1} is equal to its fundamental value in expectation, which has dynamics that are potentially likely to immediately initiate another bubble. In example **E2** for instance, if bubbles are very large, then c will be relatively large with respect to the dynamics of the fundamental value. When the bubble collapses at time $t - 1$, then $g(F_{t-1}) = \varepsilon_{t-1} - c$, which means that there is a very high probability of a new bubble being created.

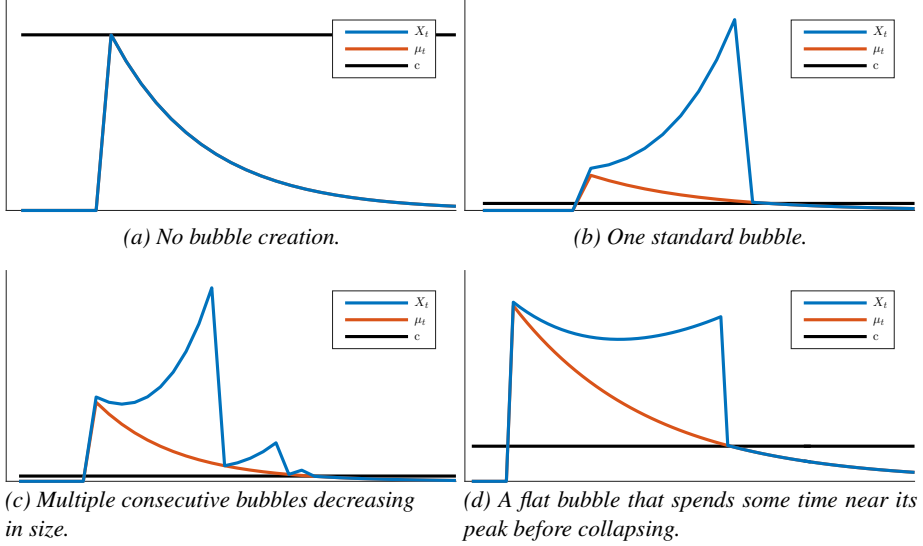
To gain more control of bubble emergence dynamics one can use more involved survival functions such as example **E3**. Here a bubble collapses if it makes out more than a fraction k of the total asset price. This allows for various bubble sizes and critical levels of the asset price as X_{t-1} depends on the fundamental value and the shock process. In fact, a higher fundamental value allows for larger bubbles, a result that can be argued to be appealing as a high fundamental price can be one of the driving reasons for the existence of the bubble. Example **E3** can control for the emergence of bubbles, as $X_{t-1} < 0$ implies that the break condition is not satisfied for any possible value of b_{t-1} and thus no bubble is created. A period of negative asset value thus ensures no bubble is created during that time, hence **E3** can be used well to describe time series which contain explosive and non-explosive windows. Finally, example **E4** captures the same effects as **E3**, but elaborates on the connection between the fundamental value and the bubble process. The bubble size and collapse and emergence times are now all directly related to the fundamental value. If the fundamental value is below the threshold c , then no bubbles are created. If the fundamental value goes above c , then a bubble is created which grows until its size is larger than k times the difference between the fundamental value and c . There are two general driving forces that cause the bubble to burst. Firstly, the fundamental value process can stay above c for an extended period of time, but as it is mean reverting while the bubble is exponentially increasing, the bubble process grows much faster and thus we eventually observe that $b_{t-1} \geq k(\mu_t - c)$. Secondly, the fundamental value can fall quickly below c again, which immediately makes the bubble collapse. Combinations between these two collapse reasons are also possible, allowing for a wide variety of bubble sizes and overall asset price dynamics.

3.2.1 Bubble variety

The bubble process described in (3.4) might appear to be rather restricted at first sight, as conditional on the current value of the bubble the updating equation allows for only two possible values, all of which are within a countable space. However, the joint dynamics between the fundamental value and the bubble process can cause the asset price to be very flexible in describing various bubble sizes, shapes and frequencies. This is especially true for examples **E3** and **E4**, which we demonstrate in Figure 3.2.1 by examining some of the

possible impulse response functions (IRFs) for the model described in equations (3.1)-(3.5) with survival condition **E4**. Figure 3.2.1a illustrates how a small impulse that does

Figure 3.2.1: Several impulse response functions for the bubble model as described in equations (3.1)-(3.5) with survival condition **E4**.



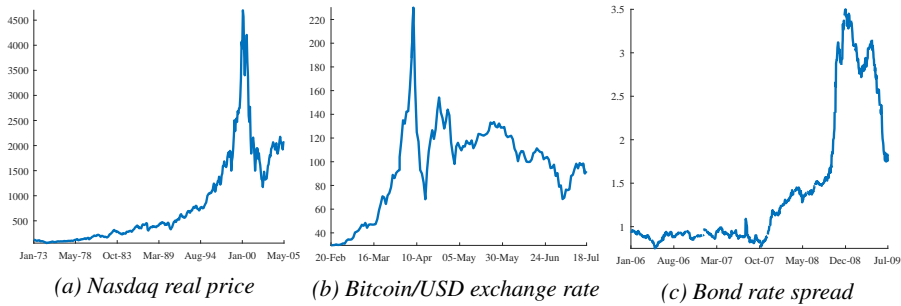
not push the fundamental process above the threshold c creates no speculative bubble. The resulting dynamics in the asset price are therefore just the mean reverting ones from the fundamental value process. In Figure 3.2.1b we have increased the size of the impulse, which results in a typical unique bubble characterised by its exponential increase followed by a sudden collapse. The collapse is caused by the fundamental process reverting back to its mean lower than c . If we further increase the size of the impulse as in Figure 3.2.1c, then we obtain a similar initial scenario, but now the bubble collapses even though the fundamental value is still above the threshold, because its size is larger than $k(\mu_{t-1} - c)$. This results in another smaller bubble immediately created once the first bubble has collapsed. Finally, Figure 3.2.1d illustrates the effect of an impulse size that causes the mean reverting fundamental process dynamics to approximately cancel out the explosive bubble dynamics. The resulting joint dynamics for the asset price show a bubble that spends some time at its peak level before collapsing.

The different possible joint dynamics in the asset price as illustrated in Figure 3.2.1

CHAPTER 3. A TIME-VARYING PARAMETER MODEL FOR LOCAL EXPLOSIONS

are often encountered in financial time series. Figure 3.2.2 exhibits some time series for which evidence for the existence of a speculative bubble has been found. The bubble shapes in each time series are remarkably different. Figure 3.2.2a plots the monthly Nas-

Figure 3.2.2: Several time series with evidence for the existence of a speculative bubble. Panel a is the monthly Nasdaq real price from January 1973 to May 2005, Panel b is the daily Bitcoin/USD exchange rate from February 20, 2013 to July 18, 2013 and Panel c is the daily spread between US Baa bond rates and Aaa bond rates from January 3, 2006 to July 2, 2009.



daq real price from January 1973 to May 2005, studied in Phillips et al. (2011). This time series contains a single bubble where the exponential increase is followed by an immediate burst and no new explosive behaviour, see the similarity with the impulse response function (IRF) in Figure 3.2.1b. Figure 3.2.2b depicts the daily Bitcoin/US dollar exchange rate from February 20 to July 18 in 2013, studied in Hencic and Gouriéroux (2015). This time series contains a classic bubble, which collapses on April 10. However, different to Figure 3.2.2a, it is followed immediately by a new smaller exponential increase and downwards burst, analogous to the IRF of Figure 3.2.1c. Figure 3.2.2c shows the daily spread between US Baa bond rates and Aaa bond rates from January 3, 2006 to July 2, 2009, studied in Phillips and Yu (2011). Here we observe a speculative bubble that increases exponentially, but then spends some time around its peak before collapsing, as in the IRF of Figure 3.2.1d. All time series contain windows where no speculative bubble is apparent, as in the IRF of Figure 3.2.1a.

3.3 Probabilistic and statistical analysis

In this section we study the probabilistic properties of our model as defined in equations (3.1)–(3.5). The bubble model contains several irregular components making the results in this section nonstandard. Firstly, the parameter α is allowed to be greater than one, which means that the bubble model is locally explosive on its sample space. Secondly, the updating equation for the bubble process (3.4) contains a discontinuity. These aspects imply that typical stability properties necessary for almost everywhere contraction conditions or smoothness assumptions do not hold, which means that it is not possible to employ standard stability theory results as developed in Bougerol (1993) or Meyn and Tweedie (1993). Instead we rely on previous work in Blasques and Nientker (2017) that provides stability results for resetting dynamic systems. Such a system is defined by an updating function that sometimes resets to a fixed, possibly random, value regardless of the past. These dynamics are present in the bubble process when the bubble collapses back to zero.

We split the parameter vector in two sub-vectors (θ, λ) which belong to the parameter space $\Theta \times \Lambda$. Here θ contains all the parameters that enter continuously in (μ_t, b_t) and λ contains the remaining parameters. Among the parameters $(\sigma^2, \delta, \beta, \gamma, \omega, \alpha)$, it is clear that σ^2 is always an element of θ , while the remaining parameters may be elements of θ or λ depending on the chosen survival function g . If the survival function depends on the bubble process, then (α, ω) belong to λ and if g is nonconstant in the fundamental value process, then (δ, β, γ) belong to λ . We examine the examples from Section 3.2 as an illustration.

- | | |
|---|--|
| E1 $g(F_{t-1}) = X_{t-1} - c,$ | then $\theta = (\sigma^2, \delta, \beta, \gamma, \omega, \alpha)$ and $\lambda = c.$ |
| E2 $g(F_{t-1}) = X_{t-1} - \mu_{t-1} - c,$ | then $\theta = (\sigma^2, \omega, \alpha)$ and $\lambda = (\delta, \beta, \gamma, c).$ |
| E3 $g(F_{t-1}) = b_{t-1} - kX_{t-1},$ | then $\theta = (\sigma^2, \delta, \beta, \gamma)$ and $\lambda = (\omega, \alpha, k).$ |
| E4 $g(F_{t-1}) = b_{t-1} - k(\mu_t - c),$ | then $\theta = \sigma^2$ and $\lambda = (\delta, \beta, \gamma, \omega, \alpha, k).$ |

Deriving consistency and asymptotic normality for λ is generally difficult. Therefore we approach the problem by deriving these results for θ conditionally on a calibrated value

of λ . This means that we will work with functions f in the Banach space $L^\infty(\Theta, \mathbb{R})$, we write $\|f\|_\Theta$ for the supremum norm. The parameter space Θ is assumed to be compact throughout this section.

We show in Section 3.3.1 that the model admits stable solutions under lenient restrictions on the parameters and survival function. We then continue to analyse the model as a filter in Section 3.3.2 and show that filter paths converge to a stable solution. We derive the likelihood in Section 3.3.3 and provide consistency and asymptotic normality for a maximum likelihood (ML) estimator in Section 3.3.4. All the proofs can be found in Appendix 3.6.

3.3.1 The model as a data generating process

This section provides results that guarantee that our model generates (strictly) stationary ergodic data with finite moments. Moreover, we show that partial solutions converge to the stationary sequence. These results will be required later on to show consistency and asymptotic normality for the ML estimator in a correctly specified model.

Data generated by our model partially adheres to very standard dynamics, as equation (3.1) holds for such data and thus (3.3) simplifies to

$$\mu_t = \delta + \beta\mu_{t-1} + \gamma\varepsilon_{t-1}. \quad (3.6)$$

The fundamental value process therefore is an autoregressive process of order one with Gaussian errors, a specification that is well studied and known to have stable solutions. We need the following assumptions to ensure the stability results:

DGP 1. *The parameter space satisfies $|\beta| < 1$.*

DGP 2. *Let $b \geq 0$ and $\mu, \varepsilon \in \mathbb{R}$. There exists a set $S \subset \mathbb{R} \times \mathbb{R}$ of positive Lebesgue measure such that the survival function satisfies*

$$\tilde{g}(\varepsilon, \mu) := \inf_{b \geq 0} g(\mu + b + \varepsilon, \mu, b) \geq 0 \quad \text{for all } (\varepsilon, \mu) \in S.$$

Assumption DGP 1 is standard in the literature on autoregressive processes. Condition DGP 2 seems complicated but essentially requires that the bubble process always has

positive probability to collapse next period, regardless of its current and past values. If this were not the case, then there are scenarios in which the bubble is guaranteed to continue growing, something that can be considered unnatural. Assumption DGP 2 is usually easy to verify.

Lemma 3.3.1. *Assumption DGP 2 holds if \tilde{g} is a continuous and surjective function.*

PROOF. The set $[0, \infty)$ contains an open subset, say O . Since \tilde{g} is surjective $\tilde{g}^{-1}(O)$ is nonempty and by continuity it is open. Any nonempty open subset in Euclidean space is of positive Lebesgue measure. ■

We verify condition DGP 2 on our examples **E1–E4** as an illustration. If our survival function is given by **E1** then $\tilde{g}(\varepsilon, \mu) = \mu + \varepsilon - c$, if our survival function is given by **E2** then $\tilde{g}(\varepsilon, \mu) = \varepsilon - c$, if our survival function is given by **E3** then $\tilde{g}(\varepsilon, \mu) = k(\mu + \varepsilon)$ as $k \leq 1$ and finally if our survival function is given by **E4** then $\tilde{g}(\varepsilon, \mu) = -k(\delta + \beta\mu + \gamma\varepsilon - c)$. All of these functions are continuous and surjective and thus condition DGP 2 is satisfied for all our examples.

Theorem 3.3.2. *Suppose that assumptions DGP 1–2 hold. Then there exists a unique causal stationary ergodic solution $((X_t, \mu_t, b_t))_{t \in \mathbb{Z}}$ to model (3.1)–(3.5). Moreover, any other solution $((\hat{X}_t, \hat{\mu}_t, \hat{b}_t))_{t \in \mathbb{N}}$ initialised at $(\hat{X}_1, \hat{\mu}_1, \hat{b}_1)$ almost surely converges exponentially fast to the stationary ergodic one, that is*

$$\left\| (\mu_1, b_1) - (\hat{\mu}_1, \hat{b}_1) \right\|_{\Theta} \xrightarrow{eas} 0 \quad \text{as } t \rightarrow \infty.$$

Theorem 3.3.6 establishes the existence of a unique causal solution to our model for each choice of θ that satisfies assumptions DGP 1–2. One can simulate an arbitrary close approximation of this solution by using any initialisation of choice and discarding the first portion of the time series.

We finalize this section by providing a result on the existence of moments for the solution found in Theorem 3.3.2. Showing such existence is dependent on the chosen survival function. We will use example **E4** in our application, so we derive the result for this survival function.

Corollary 3.3.3. *Suppose that assumptions DGP 1 holds and that the survival function is given as in E4. Then the unique stationary ergodic solution has a uniform n 'th moment for all $n \in \mathbb{N}$, i.e. $\mathbb{E}\|(X_t, \mu_t, b_t)\|_{\Theta}^n < \infty$.*

3.3.2 The model as a filter

This section focusses on the model as a filter for general data $(X_t)_{t \in \mathbb{Z}}$. Such a filter will always have to be initialised at some values $\hat{\mu}_1$ and \hat{b}_1 , as the fundamental respective bubble processes are unobserved. We impose conditions that ensure that our filtered model admits a unique stationary ergodic solution that is twice continuously differentiable, has bounded moments and that any initialised process converges to. The first set of conditions assume structure on the dependence between the (X_t) . We write $\log^+(x) = \max\{0, \log x\}$ and \mathcal{F}_s^t for the σ -algebra of (X_s, \dots, X_t) for any $s \leq t \in [-\infty, \infty]$.

FLT 1. *The data sequence $(X_t)_{t \in \mathbb{Z}}$ is stationary ergodic and has a finite log moment, that is $\mathbb{E} \log^+ |X_t| < \infty$.*

FLT 2. *Each X_t is absolutely continuous with full support pdf. If $A \in \mathcal{F}_{-\infty}^{-1}$ and $B \in \mathcal{F}_0^{\infty}$ are events of positive probability, then $\mathbb{P}(A \text{ and } B) > 0$.*

FLT 3. *The conditional distributions $X_t \mid X_{t-1}, \dots, X_{t-n}$ are absolutely continuous and of bounded density uniformly over $n \in \mathbb{N}$ and almost all possible past values with respect to Lebesgue measure.*

Condition FLT 1 is standard and necessary, one cannot expect to obtain stationary ergodic filter paths if the original data sequence is not so. A log moment is implied by the existence of any regular moment by Jensen's inequality. Assumption FLT 2 is less common, but has an intuitive interpretation. It requires the sequence $(X_t)_{t \geq 0}$ to be non exclusive, that is, conditional on the past some future events are more likely than others, however, anything that was possible unconditionally can still happen. Assumption FLT 3 is a technical one that is satisfied for most reasonable distributions. We realise conditions FLT 2 and FLT 3 are unusual in the literature. We provide the following result to illustrate that many stochastic processes satisfy these conditions.

Proposition 3.3.4. *Suppose $(X_t)_{t \in \mathbb{Z}}$ is a real valued stationary ergodic solution of a Markov chain $X_t = f(X_{t-1}, \zeta_t)$. If $f(x, \cdot)$ is a continuously differentiable function for all $x \in \mathbb{R}$ with derivative bounded away from zero for almost all $x \in \mathbb{R}$ with respect to Lebesgue measure, and $(\zeta_t)_{t \in \mathbb{Z}}$ is a sequence of independent, identically distributed and absolutely continuous random variables such that $f(x, \zeta_t)$ has full support for all $x \in \mathbb{R}$. Then conditions FLT 2 and FLT 3 are satisfied.*

Proposition 3.3.4 implies that typical processes such as general AR(1) given by $X_t = h(X_{t-1}) + \varepsilon_t$, or multiplicative specifications of the type $X_t = h(X_{t-1})\varepsilon_t$ usually satisfy conditions FLT 2 and FLT 3. The proposition can also be extended to multivariate processes where the data is one of the entries in the vector. This implies processes such as ARMA or GARCH satisfy our conditions.

As mentioned before, the dynamics of our model rely heavily on the survival function chosen, specifically whether g is nonconstant in any of its arguments. We provide the desired results for the most complex case in which g is nonconstant in any of its variables. We then need the following additional parameter restrictions.

FLT 4. *The function g is Lipschitz with derivative bounded away from zero almost everywhere, it is monotone in its first argument, decreasing and continuous in its second argument and increasing in its third argument. Moreover, the probability $\mathbb{P}(g(X_t, \mu, 0) \geq 0)$ is positive for all $\mu \in \mathbb{R}$ and the inverse of g in its third argument is L -Lipschitz.*

FLT 5. *The parameters satisfy $r := |\beta - \gamma| < 1$ and the polynomial $p(x) = 1 - rx + \gamma\alpha Lx^2$ has roots outside of the unit circle.*

Assumption 4 contains quite some restrictions of the survival function. It can be easily checked however that these all hold for example **E4**.

Theorem 3.3.5. *Suppose that assumptions FLT 1–4 hold. Then there exists a unique causal stationary ergodic solution $((\mu_t^*, b_t^*))_{t \in \mathbb{Z}}$ to model (3.3)–(3.5) that is twice continuously differentiable over Θ . Moreover, any other solution $((\hat{\mu}_t, \hat{b}_t))_{t \in \mathbb{N}}$ initialised at $(\hat{\mu}_1, \hat{b}_1)$ almost surely converges exponentially fast to the stationary ergodic one, that is,*

$$\left\| (\mu_t^*, b_t^*) - (\hat{\mu}_t, \hat{b}_t) \right\|_{\Theta} \xrightarrow{eas} 0 \quad \text{as } t \rightarrow \infty.$$

Finally, if X_t has an n 'th moment for some $n \in \mathbb{N}$, then (μ_t^*, b_t^*) has an n 'th moment too.

3.3.3 The likelihood

As mentioned in the beginning of Section 4.3 we derive our asymptotic results for θ conditionally on a calibrated value of λ . The likelihood evaluated at some $\theta \in \Theta$ for a sequence (X_1, \dots, X_T) is the joint density implied by (3.1)–(3.5). The fundamental value and bubble processes are unobserved, so we choose initialised values $\hat{\mu}_1$ and \hat{b}_1 which deliver filtered sequences $(\hat{\mu}_t(\theta, \lambda))_{t=2}^T$ and $(\hat{b}_t(\theta, \lambda))_{t=2}^T$ according to (3.3)–(3.5). It follows that

$$X_t | X_1, \dots, X_{t-1} = X_t | \hat{\mu}_t(\theta, \lambda), \hat{b}_t(\theta, \lambda) \sim N(\hat{\mu}_t(\theta, \lambda) + \hat{b}_t(\theta, \lambda), \sigma^2)$$

and thus prediction error decomposition delivers the average log likelihood as a function $\Theta \rightarrow \mathbb{R}$ given by

$$\begin{aligned} \hat{L}_T(\theta) &\propto \frac{1}{T} \sum_{t=2}^T \ell(X_t, \hat{\mu}_t(\theta, \lambda), \hat{b}_t(\theta, \lambda), \sigma^2), \\ \ell(X_t, \hat{\mu}_t(\theta, \lambda), \hat{b}_t(\theta, \lambda), \sigma^2) &:= -\frac{1}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} (X_t - \hat{\mu}_t(\theta, \lambda) - \hat{b}_t(\theta, \lambda))^2. \end{aligned}$$

From here on out we will omit mentioning that the plug in processes depend on θ and λ to keep notation clear.

3.3.4 Asymptotic results

Consistency

The ML estimator of θ is defined as

$$\hat{\theta}_T = \arg \max_{\theta \in \Theta} \hat{L}_T(\theta).$$

We need the following conditions to obtain consistency.

CS 1. $(X_t)_{t \in \mathbb{Z}}$ is stationary and ergodic with bounded second moment: $\mathbb{E}|X_t|^2 < \infty$.

CS 2. *The filter vector $((\hat{\mu}_t, \hat{b}_t))_{t \in \mathbb{N}}$ is invertible and converges to a limit process $((\mu_t^*, b_t^*))_{t \in \mathbb{Z}}$ uniformly over Θ with two uniform bounded moments. That is,*

$$\left\| (\hat{\mu}_t, \hat{b}_t) - (\mu_t^*, b_t^*) \right\|_{\Theta} \xrightarrow{a.s.} 0 \quad \text{as } t \rightarrow \infty \quad \text{and} \quad \mathbb{E} \|(\mu_t^*, b_t^*)\|_{\Theta}^2 < \infty,$$

Moreover, the joint process $((X_t, \mu_t^, b_t^*))_{t \in \mathbb{Z}}$ is strictly stationary and ergodic.*

CS 3. *There exists a unique maximizer θ_0 of the limit log likelihood, that is, for every $\theta \in \Theta$ that is unequal to θ_0 we have*

$$\mathbb{E} \ell(X_t, \mu_t^*(\theta, \lambda), b_t^*(\theta, \lambda), \sigma^2) < \mathbb{E} \ell(X_t, \mu_t^*(\theta_0, \lambda), b_t^*(\theta_0, \lambda), \sigma_0^2).$$

The assumptions CS 1–3 are typical conditions used in the theory of M -estimators. Assumptions CS 1 and CS 2 assume stochastic properties of our model that ensure that a law of large numbers can be applied. Note that they are both implied by assumption FLT 1 – FLT 5 and Theorem 3.3.5. Assumption CS 3 ensures that the limit log likelihood is maximised at a unique point θ_0 , given the fixed parameter λ . Note that the expectations exist by the moment assumptions in CS 1–2. When the model is assumed to be well specified and λ is fixed at its true value λ_0 , then it is often easy to show that this assumption holds and that the parameter of interest θ_0 is the true parameter, that is, the parameter that corresponds to the data generating process for $\{X_t\}_{t \in \mathbb{Z}}$. If the model is misspecified, or λ is set at some arbitrary value $\lambda \neq \lambda_0$, then the uniqueness of the parameter of interest θ_0 is harder to establish.¹ In this case, the limit parameter θ_0 is a ‘pseudo-true parameter’, i.e. a parameter that minimizes a Kullback-Leibler divergence between the true conditional density of the data and the model-implied conditional density, see Section 2.3 of White (1994).

Theorem 3.3.6 (Consistency). *If assumptions CS 1–3 hold, then $\hat{\theta}_T \xrightarrow{a.s.} \theta_0$.*

Theorem 1 establishes the a.s. convergence of the ML estimator $\hat{\theta}_T$ to the pseudo-true parameter θ_0 which is the unique maximizer of the limit log likelihood or any given value of $\lambda \in \Lambda$. In this sense, the θ_0 provides the best Kullback-Leibler approximation to the

¹When the uniqueness assumption fails, set consistency can be easily established, thus ensuring that the ML estimator converges to the limit argmin set; see Lemma 4.2 in Pötscher and Prucha (1997) for standard conditions that apply here.

CHAPTER 3. A TIME-VARYING PARAMETER MODEL FOR LOCAL EXPLOSIONS

true unknown distribution of the data, for the given value of λ . Naturally, if the model is correctly specified and λ is calibrated at its true value, then θ_0 corresponds to the true parameter.

Corollary 3.3.7. *Suppose that the model is correctly specified and that λ is calibrated at its true value. If the results of Theorem 3.3.2 can be applied for all $\theta \in \Theta$, and*

$$\mathbb{E} \|(\mu_t, b_t)\|_{\Theta}^2 < \infty.$$

Then the ML estimator $\hat{\theta}_T$ converges a.s. to the true parameter θ_0 .

Asymptotic normality

In what follows, we establish the asymptotic normality of the ML estimator $\hat{\theta}_T$ as $T \rightarrow \infty$. Theorem 3.3.8 focuses on the case of a well specified and correctly calibrated model, and Theorem 3.3.9 obtains asymptotic normality for a misspecified or incorrectly calibrated model where $\lambda \neq \lambda_0$. We need the following standard assumptions.

AN 1. *The conditions CS 1–3 hold and θ_0 belongs to the interior of Θ .*

AN 2. *The limit process (μ_t^*, b_t^*) is twice continuously differentiable on Θ for all $t \in \mathbb{Z}$.*

AN 3. *$\{X_t\}_{t \in \mathbb{Z}}$ has four bounded moments $\mathbb{E}|X_t|^4 < \infty$.*

AN 4. *The Fisher information matrix is invertible.*

The following assumption ensures that the filter derivatives converge almost surely and exponentially fast to limit strictly stationary and ergodic sequences. The exponential rate for the filter was established in Section 3.3.2. It is clear that the same argument applies to the derivative processes $\{\frac{\partial \hat{\mu}_t}{\partial \theta}\}$ and $\{\frac{\partial \hat{b}_t}{\partial \theta}\}$. In particular, we have again that $\{\frac{\partial \hat{b}_t}{\partial \theta}\}$ converges at any rate since it is reset to zero in a finite number of steps with positive probability, and $\{\frac{\partial \hat{\mu}_t}{\partial \theta}\}$ converges exponentially fast due to its autoregressive nature.

AN 5. *The derivative processes are invertible at exponential rates and feature four bounded moments,*

$$\left\| (\nabla^{0:2} \hat{\mu}_t, \nabla^{0:2} \hat{\beta}_t) - (\nabla^{0:2} \mu_t, \nabla^{0:2} \beta_t) \right\|_{\Theta} \xrightarrow{a.s.} 0 \quad \text{as } t \rightarrow \infty,$$

$$\text{and } \mathbb{E} \left\| (\nabla^{0:2} \mu_t, \nabla^{0:2} \beta_t) \right\|_{\Theta}^4 < \infty.$$

Theorem 3.3.8. (*Asymptotic Normality: Correct specification*) Let assumptions AN 1–5 hold. Suppose further that the model is well specified and that $\lambda = \lambda_0$. Then

$$\sqrt{T}(\hat{\theta}_T - \theta_0) \xrightarrow{d} N(\mathbf{0}, \mathcal{I}^{-1}(\theta_0)) \text{ as } T \rightarrow \infty,$$

where $\mathcal{I}^{-1}(\theta_0)$ denotes the inverse information matrix.

We need the following additional assumption to obtain asymptotic normality under a misspecified model.

AN 6. $\{x_t\}$ is near epoch dependent of size -1 on a ϕ -mixing sequence of size $-r/(r-1)$ for some $r > 2$.

Theorem 3.3.9. (*Asymptotic Normality: Incorrect specification/calibration*) Let assumptions AN 1–6 hold. Then

$$\sqrt{T}(\hat{\theta}_T^1 - \theta_0^1) \xrightarrow{d} N(\mathbf{0}, \Sigma(\theta_0^1, \theta^2)) \text{ as } T \rightarrow \infty,$$

where

$$\Sigma(\theta_0^1, \theta^2) = (\mathbb{E}\hat{\ell}_t''(\theta_0^1, \theta^2))^{-1} (\mathbb{E}\hat{\ell}_t'(\theta_0^1, \theta^2) \mathbb{E}\hat{\ell}_t'(\theta_0^1, \theta^2)^\top) (\mathbb{E}\hat{\ell}_t''(\theta_0^1, \theta^2))^{-1}.$$

3.4 Illustrations

In this section we test the descriptive capability of our model and illustrate the accessibility and ease of further analysis after estimation. We will use our model with survival function **E4** and estimate it on a part of the Bitcoin/US dollar exchange rate. We realise that the chosen survival function implies that most of our parameters belong to λ , the vector of parameters that enter discontinuously into the likelihood. Therefore we add a short simulation study in Section 3.4.1 in which we examine the distribution of our ML estimator for a representative choice of parameters. Section 3.4.2 contains the estimation results and further analysis.

3.4.1 Simulation study

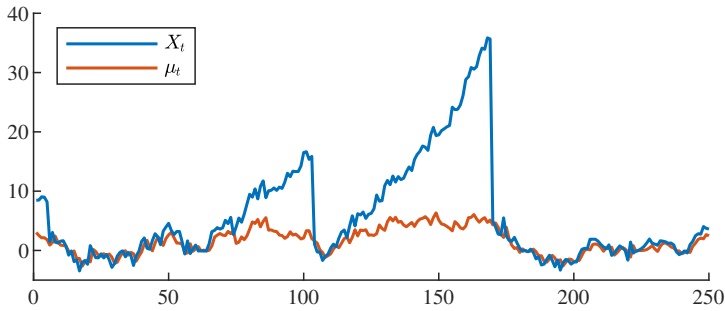
We examine the distribution of the ML estimator for a given parametrization, stated in Table 3.4.1. This choice represents a medium amount of bubbles of size relative to the

σ	δ	β	γ	ω	α	k	c
1.0	0.1	0.95	0.7	0.2	1.03	7	-0.1

Table 3.4.1: Parametrization used for simulation study.

fundamental value process. A typical simulation for the implied model can be found in Figure 3.4.1. Note that there are windows of locally explosive behaviour, but also times

Figure 3.4.1: An example simulated path for the parametrization of the model described in Table 3.4.1.



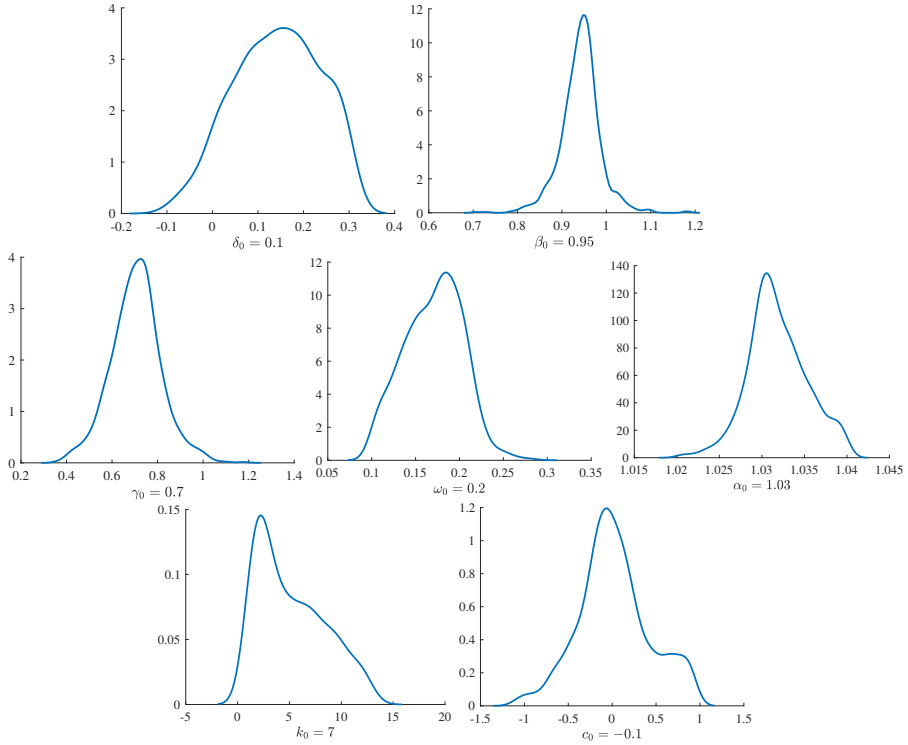
at which no bubbles seem to form. The size of the bubbles is substantially larger than that of the fundamental value, but not so far as to render its value insignificant compared to the magnitude of the bubble process.

We estimate the model parameters by maximising the likelihood over an area centered around the true values. The likelihood however is discontinuous and nondifferentiable, which means that gradient based optimizing algorithms cannot be applied. Instead we implement a procedure based on the genetic algorithm in Matlab, which is inspired by natural selection observed in biological evolution. The algorithm generates a population of points and then successively selects a partially random subpopulation to be parents to the next population. We use a total of fifty generations to get near to the optimal point and then use that as a starting value for a gradient based optimizer to get quicker to the maximiser. An important observation about our procedure is that the resulting optimizing algorithm is stochastic. This implies that found parameter values cannot be reproduced,

however, the algorithm works sufficiently well such that successive estimations on the same data produce very similar results.

For the simulation we calculate one thousand estimate values, each of which is based on a sample path of length one thousand. The resulting estimated densities for the ML estimator are portrayed in Figure 3.4.2. Here we see that all densities, except the one

Figure 3.4.2: Estimated kernel densities for the ML estimator based on one thousand estimate values, each of which is based on a sample path of length one thousand.

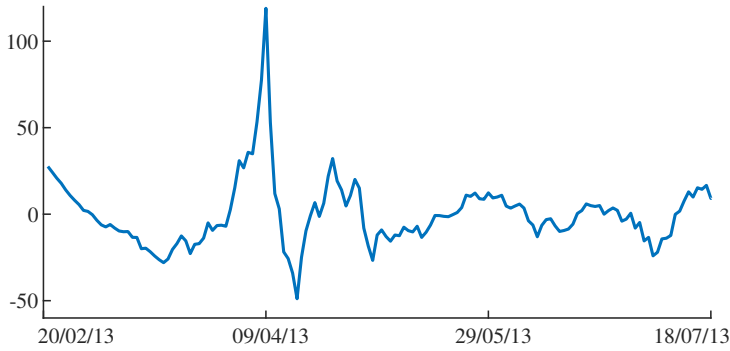


for k , are close to symmetric with their peak at the true value. The estimator for k has more inaccuracy than the others, because most bubbles in this parametrization collapse due to the fundamental process dropping below the threshold c . A path of one thousand observations contains approximately ten bubbles, so therefore there is relatively little data to estimate k .

3.4.2 The BTC/USD exchange rate

The data set that we use is equivalent to the one studied in Hencic and Gouriéroux (2015). We take the detrended daily Bitcoin/US dollar exchange rate from February 20 to July 18 in 2013, given in Figure 3.4.3. There appears to be a big bubble that collapses on April 10, 2013. Moreover, it is potentially followed by a second smaller bubble. Afterwards it tends to behave as a standard stable and mean reverting process.

Figure 3.4.3: Detrended daily Bitcoin/USD exchange rate, taken from Hencic and Gouriéroux (2015).



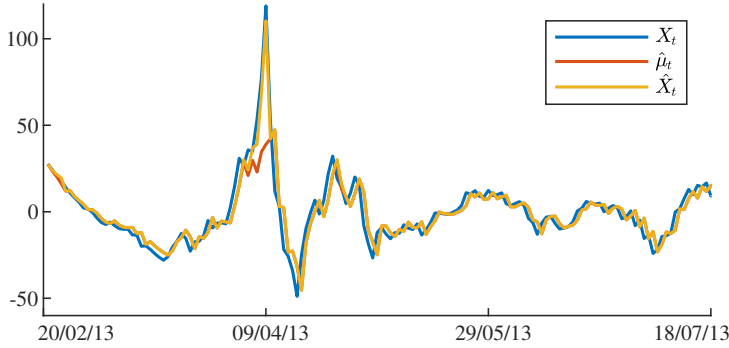
We estimate the model parameters as discussed in the simulation study. The results are given in Table 3.4.2. The estimate of α is relatively large, which means that any

σ	δ	β	γ	ω	α	k	c
2.05	-0.25	0.86	1.04	0.44	2.10	2.39	12.29

Table 3.4.2: Parametrization used for simulation study.

potential bubble is highly explosive. Moreover, the value of c implies that the potential smaller second bubble is mostly identified as the fundamental value moving away from its mean. These observations are substantiated when we look at the filtered time series in Figure 3.4.4, note that these are also the in sample one step ahead predictions. Our model describes only one significant bubble, which is preceded by an increase in the fundamental value. It then collapses due to the k parameter restriction on bubble size. Afterwards the fundamental value stays below the threshold c and hence the rest of the time series is filtered as an autoregressive process. Our model performs well as it predicts the burst of the bubble on April 10 correctly. It does however underestimate the additional decrease

Figure 3.4.4: The filtered daily Bitcoin/USD exchange rate, taken from Hencic and Gouriéroux (2015).



in fundamental value after the bubble burst.

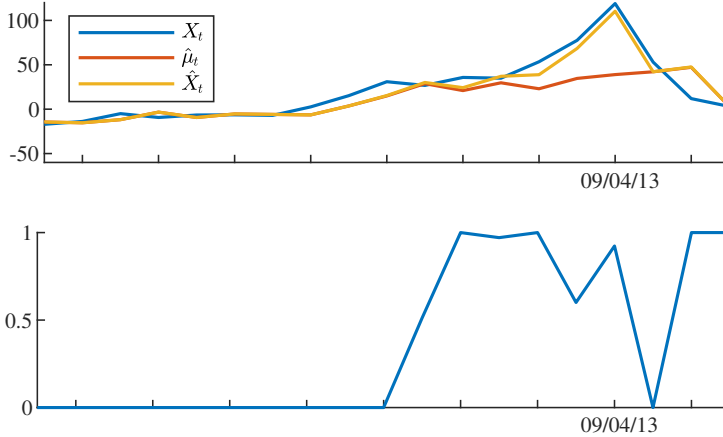
We compare our bubble model to the nested simpler model in which we set the bubble always to zero. In that case we have only four parameters left out of eight. The resulting Akaike information criteria are 646 for the full bubble model and 726 for the simpler model. Therefore we conclude that including the bubble process adds descriptive power and thus we prefer that model.

Observation driven parameter varying time series models have two main advantages. The first one is that they are easy to estimate as the likelihood is accessible through prediction error decomposition as discussed in Section 4.3. The second advantage is that further analysis is straightforward once the model has been estimated, as we have closed form formulas for the filtered time series. For example, we can calculate the probability that the bubble condition in the next period holds. Figure 3.4.5 plots the filtered time series and these probabilities for some period centered around the bubble. Here we see that the probability of nonzero bubble values before the bubble start are virtually equal to zero. As the bubble starts, the probability of a nonzero bubble is almost one. The probability dips a little for April 8, as the fundamental value on April 7 has gone down a little. The fundamental value then increases on April 8 however and thus so does the probability for April 9. When we get to April 10 the probability is again almost zero as the large exponential growth of the bubble has outgrown the fundamental value process on April 9 by far too much.

The estimated probabilities above are just an example of many possible features that

CHAPTER 3. A TIME-VARYING PARAMETER MODEL FOR LOCAL EXPLOSIONS

Figure 3.4.5: The filtered daily Bitcoin/USD exchange rate around the bubble in the top frame, and the probability that the bubble condition holds in the next period shifted one period to the right in the second frame.



can be predicted. For example, once the model parameters are estimated, one can predict expected remaining bubble life, the probability that a bubble will emerge at a given time, or the expected maximal bubble size.

3.5 Conclusion

We have introduced a new observation driven time varying parameter model to describe locally explosive behaviour in a stationary setting. We do so by splitting the asset price into the sum of its fundamental value and a speculative bubble. For the sake of simplicity we have assumed an AR(1) process for the fundamental value and a collapsing AR(1) for the bubble process. However, these are not binding assumptions and many extensions and variations are possible. Of course any mean reverting stationary process for the fundamental value fits exactly in the theoretic domain presented in the paper. Dynamics can be changed by using a nonstationary process such as a random walk for the fundamental value. Most asset pricing data is nonstationary, so this would mean that one does not have to detrend the data, which makes out of sample forecasting possible. Other possibilities that could be explored are extensions to the break condition. One could for example add external stochastics allowing for more structural models that include specific financial or

economic variables that can help in predicting bubble collapses. Another extension can be made by changing the sudden collapse in a more smooth exponential decrease. The dynamics of this model would then be very close to those described in the mixed causal and noncausal literature.

3.6 Appendix: Proofs

3.6.1 Proof of Theorem 3.3.2

It is straightforward to verify the assumptions of Theorem 3.1 in Bougerol (1993) for the system defined in equations (3.2) and (3.6). It then immediately follows that there exists a unique causal stationary ergodic sequence $(\mu_t)_{t \in \mathbb{Z}}$ that satisfies (3.6) and that any other solution $(\hat{\mu}_t)_{t \in \mathbb{N}}$ initialised at $\hat{\mu}_1$ satisfies

$$\|\mu_t - \hat{\mu}_t\|_{\Theta} \xrightarrow{eas} 0 \quad \text{as } t \rightarrow \infty.$$

Next, we substitute equation (3.1) into the survival function to obtain the bubble updating function

$$b_{t+1} = \phi_t(b_t), \quad \text{where} \quad \phi_t(b) = (\omega + \alpha b) \mathbb{1} \{g(\mu_t + b + \varepsilon_t, \mu_t, b) < 0\}.$$

We then check Assumption A for Theorem 2.1 in Blasques and Nientker (2017). Assumption A1 is trivial and A2 is satisfied by Krengel's Lemma: a measurable function of a stationary ergodic sequence produces a stationary ergodic sequence, see Proposition 4.3 in Krengel (1985). For the final assumption A3 we note that $\phi_t(b) = 0$ for all $b \in [0, \infty)$, if

$$g(\mu_t + b + \varepsilon_t, \mu_t, b) \geq \inf_{b \geq 0} g(\mu_t + b + \varepsilon_t, \mu_t, b) \geq 0.$$

Therefore Assumption A3 is satisfied if

$$\mathbb{P} \left(\inf_{b \geq 0} g(\mu_t + b + \varepsilon_t, \mu_t, b) \geq 0 \right) > 0.$$

CHAPTER 3. A TIME-VARYING PARAMETER MODEL FOR LOCAL EXPLOSIONS

This is implied by condition DGP 2, because (3.2) implies that μ_t and ε_t are independent and both absolutely continuous on the real line. Therefore the joint random variable (μ_t, ε_t) is absolutely continuous on \mathbb{R}^2 and thus $\mathbb{P}((\mu_t, \varepsilon_t) \in S) > 0$. We conclude that there exists a unique causal stationary ergodic sequence $(b_t)_{t \in \mathbb{Z}}$ that satisfies the bubble updating process defined in (3.1)–(3.5) and that any other solution $(\hat{b}_t)_{t \in \mathbb{N}}$ initialised at \hat{b}_1 satisfies

$$\left\| b_t - \hat{b}_t \right\|_{\Theta} \xrightarrow{eas} 0 \quad \text{as } t \rightarrow \infty.$$

The final conclusion follows again by Krengel's Lemma.

3.6.2 Proof of Corollary 3.3.3

The solution for the fundamental process found in Theorem 3.3.2 is given by

$$\mu_t = \sum_{i=0}^{\infty} \beta \varepsilon_{t-i-1},$$

so all moments of μ_t are finite over Θ as it is a compact space so that $|\beta|$ is bounded and the ε_t are Gaussian and thus have all finite moments. For the bubble process we have

$$b_t = (\omega + \alpha b_{t-1}) \mathbb{1}_{\{b_{t-1} < k(\mu_t - c)\}} \leq \omega + \max\{\alpha k(\mu_t - c), 0\}$$

and hence moment existence follows from those of the fundamental value process.

3.6.3 Proof of Proposition 3.3.4

We fix some $\zeta \in \mathbb{R}$ and let $\xi = f(x, \zeta)$. The fact that $f(x, \cdot)$ is continuously differentiable implies by the inverse function theorem that it is invertible on a neighbourhood O around ζ and that the inverse $f^{-1}(x, \cdot)$ is also continuously differentiable. Moreover, by assumption, for Lebesgue almost all $x \in \mathbb{R}$ there exists an $L > 0$ such that $|f'(x, \zeta)| \geq L$ and thus $f^{-1}(x, \cdot)$ is Lipschitz as

$$\frac{d}{d\xi} f^{-1}(x, \xi) = \frac{1}{f'(x, \zeta)} \leq \frac{1}{L}.$$

The real line is separable, hence we can choose a countable number of disjoint compact neighbourhoods $\{O_k\}_{k \in \mathbb{N}}$ whose union is equal to \mathbb{R} and $f(x, \cdot)$ is invertible on each neighbourhood as above. A continuously differentiable function on a compact set is absolutely continuous, which in turn implies that it has the Luzin property, that is, sets of measure zero are mapped to sets of measure zero.

We now prove that X_t is absolutely continuous. Let $E \subset \mathbb{R}$ be a set of Lebesgue measure zero and let F denote the distribution function of X_t , then by independence of the ζ_t we have

$$\begin{aligned} \mathbb{P}(X_t \in E) &= \mathbb{P}(f(X_{t-1}, \zeta_t) \in E) \\ &= \int_{\mathbb{R}} \mathbb{P}(f(x, \zeta_t) \in E) F(dx) = \sum_{k=1}^{\infty} \int_{\mathbb{R}} \mathbb{P}(f(x, \zeta_t) \in E \cap O_k) F(dx) \\ &= \sum_{k=1}^{\infty} \int_{\mathbb{R}} \mathbb{P}(\zeta_t \in f^{-1}(x, E \cap O_k)) F(dx) = \sum_{k=1}^{\infty} \int_{\mathbb{R}} 0 F(dx) = 0, \end{aligned}$$

where we used that $E \cap O_k$ has Lebesgue measure zero, $f^{-1}(x, \cdot)$ has the Luzin property on each O_k and ζ_t is absolutely continuous. The absolute continuity of the conditional distributions follows similarly as by the independence of the ζ_t and the Markov property

$$\begin{aligned} \mathbb{P}(X_t \in E \mid X_{t-1} = x_1, \dots, X_{t-n} = x_n) &= \mathbb{P}(X_t \in E \mid X_{t-1} = x_1) \\ &= \mathbb{P}(f(x_1, \zeta_t) \in E) = 0. \end{aligned}$$

Next we show that the conditional densities are uniformly bounded. By assumption, we know that the density of ζ_t is bounded by some $B > 0$. For some $\xi \in \mathbb{R}$ and $\eta > 0$ we then have

$$\begin{aligned} \mathbb{P}(\xi < X_t \leq \xi + \eta \mid X_{t-1} = x_1, \dots, X_{t-n} = x_n) &= \mathbb{P}(\xi < f(x_1, \zeta_t) \leq \xi + \eta) \\ &= \sum_{k=1}^{\infty} \mathbb{P}(\zeta_t \in f^{-1}(x_1, (\xi, \xi + \eta] \cap O_k)) \leq \frac{B}{L} \eta, \end{aligned}$$

where we used that $f^{-1}(x, \cdot)$ is $\frac{1}{L}$ -Lipschitz on each O_k and the density of ζ_t is bounded by B . Taking the limit of $\eta \rightarrow 0$ shows that the conditional densities are all bounded by $\frac{B}{L}$.

CHAPTER 3. A TIME-VARYING PARAMETER MODEL FOR LOCAL EXPLOSIONS

Finally we note that the full support of X_t follows directly from the fact that $f(x, \zeta_t)$ has full support for all $x \in \mathbb{R}$ and ζ_t is absolutely continuous, and show the non exclusive property in FLT 2. This follows from the Markov chain setup. Let $A_0 \subset \mathbb{R}$ be all points x such that $\mathbb{P}(A \mid X_0 = x) > 0$. Then A_0 has positive Lebesgue measure as X_0 is absolutely continuous and

$$P(X_0 \in A_0) \geq \mathbb{P}(A \text{ and } X_0 \in A_0) = P(A) > 0.$$

It follows that

$$\begin{aligned} \mathbb{P}(A \text{ and } B) &= \int_{\mathbb{R}} \mathbb{P}(A \text{ and } B \mid X_0 = x) F(dx) \\ &= \int_{\mathbb{R}} \mathbb{P}(A \mid X_0 = x) \mathbb{P}(B \mid X_0 = x) F(dx) > 0, \end{aligned}$$

where we used that $\mathbb{P}(A \mid X_0 = x)$ is greater than zero on a set of positive Lebesgue measure and $\mathbb{P}(B \mid X_0 = x)$ is greater than zero for all $x \in \mathbb{R}$ as $f(x, \zeta_t)$ has full support for all $x \in \mathbb{R}$ and ζ_t is absolutely continuous.

3.6.4 Proof of Theorem 3.3.5

The existence of a stationary ergodic solution

We follow the approach used in Theorem 3.1 of Bougerol (1993) where we expand the model equations backwards and show that this converges to a stationary ergodic solution. We define the joint updating equation $(\mu_t, b_t) = \Phi_{t-1}(\mu_{t-1}, b_{t-1})$, where for $u \in \mathbb{R}$ and $v \geq 0$ we have

$$\begin{aligned} \Phi_{t-1}(u, v) &= (\phi_{t-1}(u, v), \psi_{t-1}(u, v)), \\ \phi_{t-1}(u, v) &= \delta + ru + \gamma X_{t-1} - \gamma v, \\ \psi_{t-1}(u, v) &= (\omega + \alpha v) \mathbb{1}\{g(X_{t-1}, u, v) < 0\}. \end{aligned}$$

To ease notation we write $\mu_t^{(0)} = u$ and $b_t^{(0)} = v$ and then define the backward iterates recursively for $m \in \mathbb{N}$ as

$$\mu_t^{(m)} = \phi_{t-1} \left(\mu_{t-1}^{(m-1)}, b_{t-1}^{(m-1)} \right) \quad \text{and} \quad b_t^{(m)} = \psi_{t-1} \left(\mu_{t-1}^{(m-1)}, b_{t-1}^{(m-1)} \right).$$

The goal will be to show that $b_t^{(m)}$ is almost surely eventually constant as $m \rightarrow \infty$ and that $\lim_{m \rightarrow \infty} \mu_t^{(m)}$ exists. The stationary ergodic solution is then given by

$$\left(\lim_{m \rightarrow \infty} \mu_t^{(m)}, \lim_{m \rightarrow \infty} b_t^{(m)} \right)_{t \in \mathbb{Z}}.$$

It is stationary ergodic by Corollary 2.1.3 of Straumann and Mikosch (2006) and it is a solution since

$$\lim_{m \rightarrow \infty} \mu_t^{(m)} = \lim_{m \rightarrow \infty} \phi_{t-1} \left(\mu_{t-1}^{(m-1)}, b_{t-1}^{(m-1)} \right) = \phi_{t-1} \left(\lim_{m \rightarrow \infty} \mu_{t-1}^{(m-1)}, \lim_{m \rightarrow \infty} b_{t-1}^{(m-1)} \right),$$

where we are allowed to swap the limit in because the second argument is eventually constant and ϕ_t is continuous in its first argument. Similarly

$$\lim_{m \rightarrow \infty} b_t^{(m)} = \lim_{m \rightarrow \infty} \psi_{t-1} \left(\mu_{t-1}^{(m-1)}, b_{t-1}^{(m-1)} \right) = \psi_{t-1} \left(\lim_{m \rightarrow \infty} \mu_{t-1}^{(m-1)}, \lim_{m \rightarrow \infty} b_{t-1}^{(m-1)} \right),$$

where we are allowed to swap the limit in because the second argument is eventually constant, the random variable

$$g \left(X_{t-1}, \lim_{m \rightarrow \infty} \mu_{t-1}^{(m-1)}, \lim_{m \rightarrow \infty} b_{t-1}^{(m-1)} \right)$$

is absolutely continuous by assumption FLT 3 and the fact that g is monotone in its first argument, and finally because g is continuous in its second argument by assumption FLT 4.

Lemma 3.6.1. *The sequence $b_0^{(m)}$ is eventually constant as $m \rightarrow \infty$ and $\lim_{m \rightarrow \infty} \mu_0^{(m)}$ converges almost surely.*

CHAPTER 3. A TIME-VARYING PARAMETER MODEL FOR LOCAL EXPLOSIONS

PROOF. Define

$$s_t = \limsup_{m \rightarrow \infty} \mu_t^{(m)} \quad \text{and} \quad i_t = \liminf_{m \rightarrow \infty} \mu_t^{(m)}.$$

The proof that the backward iterate limits above exist consists of two steps that we will show later:

- (i) We show that for every $\eta > 0$ there exists an event $A_\eta \in \mathcal{F}_{-\infty}^{-1}$ of positive probability, such that conditional on A_η we have almost surely $s_0 - i_0 < \eta$ and $b_0^{(m)}$ is constant for sufficiently large m .
- (ii) We show that there exists an event $B_\eta \in \mathcal{F}_{-\infty}^\infty$ that contains A_η , such that conditional on B_η we have $s_t - i_t \leq r^t(s_0 - i_0)$ for all $t \in \mathbb{N}$. Moreover $b_t^{(m)}$ is eventually constant for all $t \in \mathbb{N}$.

Since $(X_t)_{t \in \mathbb{Z}}$ is stationary ergodic there almost surely are infinitely many $0 > -t_1 > -t_2 > \dots$ for which the event B_η shifted by t_k to the right occurs. If it occurs for such a $-t_k$, then

$$s_0 - i_0 \leq r^{t_k}(s_{-t_k} - i_{-t_k}) \leq r^{t_k}\eta.$$

Taking the limit of $k \rightarrow \infty$ then delivers $s_0 = i_0$ and thus the limit $\lim_{m \rightarrow \infty} \mu_0^{(m)}$ converges almost surely. The fact that $\left(b_0^{(m)}\right)_{m \in \mathbb{N}}$ is eventually constant follows immediately from part (ii) and the same argument that the event B_η occurs for some $-t < 0$. ■

Lemma 3.6.2. *Claim (i) holds.*

PROOF. We start out by showing that almost surely $s_t - i_t < \infty$. This follows by a series of upper bounds. Firstly we have

$$\mu_t^{(m)} = \delta + r\mu_{t-1}^{(m-1)} + \gamma X_{t-1} - \gamma b_{t-1}^{(m-1)} \leq \delta + r\mu_{t-1}^{(m-1)} + \gamma X_{t-1}$$

Assumptions FLT 5 and FLT 1 together with Lemma 2.1 in Straumann and Mikosch (2006) ensure that expanding backwards and taking the limit converges, hence $s_t < \infty$.

The infimum requires more work. Note that the Lipschitz condition of the inverse of g in its third argument as stated in assumption FLT 4 ensure the existence of two constants $L, K > 0$ such that

$$\begin{aligned} b_t^{(m)} &= \left(\omega + \alpha b_{t-1}^{(m-1)} \right) \mathbb{1} \left\{ g \left(X_{t-1}, \mu_{t-1}^{(m-1)}, b_{t-1}^{(m-1)} \right) < 0 \right\} \\ &= \left(\omega + \alpha b_{t-1}^{(m-1)} \right) \mathbb{1} \left\{ b_{t-1}^{(m-1)} < g^{-1} \left(X_{t-1}, \mu_{t-1}^{(m-1)}, 0 \right) \right\} \\ &\leq \omega + \alpha \max \left\{ g^{-1} \left(X_{t-1}, \mu_{t-1}^{(m-1)}, 0 \right), 0 \right\} \\ &\leq \omega + \alpha \max \left\{ K + L \left(\mu_{t-1}^{(m-1)} + X_{t-1} \right), 0 \right\}. \end{aligned}$$

It follows that

$$\begin{aligned} \mu_t^{(m)} &= \delta + r \mu_{t-1}^{(m-1)} + \gamma X_{t-1} - \gamma b_{t-1}^{(m-1)} \\ &\geq (\delta - \gamma(\omega + \alpha K)) + r \mu_{t-1}^{(m-1)} - \gamma \alpha L \mu_{t-2}^{(m-2)} + \gamma (X_{t-1} - \alpha L X_{t-2}). \end{aligned}$$

Again assumptions FLT 5 and FLT 1 together with Lemma 2.1 in Straumann and Mikosch (2006) ensure that expanding backwards converges, hence $i_t > -\infty$. We conclude that $s_t - i_t < \infty$. Note that these bounds immediately prove the moment statement in Theorem 3.3.5.

Next, we choose an $M > 0$ such that $\mathbb{P}(s_t - i_t < M) > 0$ and let $t = \left\lceil \frac{\log(\eta/m)}{\log r} \right\rceil$, where $\lceil x \rceil$ is the smallest integer larger than x . Continuity of g in its second argument, the positive probability condition in assumption FLT 4 and assumption FLT 2 guarantee that by conditioning on the past we can show for each $0 \leq v < t$ that

$$\mathbb{P} \left(\limsup_{m \rightarrow \infty} b_{-v}^{(m)} = 0 \right) \geq \mathbb{P} (g(X_{-v-1}, s_{-v-1}, 0) \geq 0) > 0.$$

It follows by Assumption FLT 2 that there exists an event $A_\eta \in \mathcal{F}_{-\infty}^{-1}$ of positive probability such that

$$s_t - i_t < M \quad \text{and} \quad \limsup_{m \rightarrow \infty} b_{-v}^{(m)} = 0 \quad \text{for all } 0 \leq v < t.$$

This then implies that

$$\begin{aligned} s_0 - i_0 &= r(s_{-1} - i_{-1}) + \gamma \left(\limsup_{m \rightarrow \infty} b_{-1}^{(m)} - \liminf_{m \rightarrow \infty} b_{-1}^{(m-1)} \right) \\ &= r^t(s_{-t} - i_{-t}) < r^t M \leq \eta, \end{aligned}$$

which concludes the proof of part (i). ■

Lemma 3.6.3. *Claim (ii) holds.*

PROOF. The argument will be a recursive one, conditional on A_η . Suppose that $s_t - i_t < r^t \eta$ and $b_t^{(m)}$ is eventually constant, then

$$s_{t+1} - i_{t+1} = r(s_t - i_t) < r^{t+1} \eta.$$

Next we show that there exists an event such that $b_{t+1}^{(m+1)}$ is eventually constant. Note that this holds if and only if

$$\text{sign}(g(X_t, i_t, b_t)) = \text{sign}(g(X_t, s_t, b_t)), \quad (3.7)$$

where $b_t = \lim_{m \rightarrow \infty} b_t^{(m)}$. The Lipschitz condition in assumption FLT 4 implies that there exists a $K > 0$ such that

$$|g(X_t, s_t, b_t) - g(X_t, i_t, b_t)| \leq K(s_t - i_t) < K r^t \eta.$$

Moreover, the derivative being bounded away from zero by at least some $B > 0$ and the monotonicity of g in its second argument implied by assumption FLT 4 then ensure that (3.7) follows from

$$|g(X_t, s_t, b_t)| > B K r^t \eta.$$

We conclude that statement (ii) follows if $|g(X_t, s_t, b_t)| > B K r^t \eta$ for all $t \in \mathbb{N}$.

Next we determine the probability of this event. Let $I_{2K r^t \eta}(s_t, b_t)$ be a stochastic interval of length $K r^t \eta$ such that if $|g(X_t, s_t, b_t)| \leq B K r^t \eta$ then $X_t \in I_{2K r^t \eta}(s_t, b_t)$.

Then by assumption FLT 3 there exists a $U > 0$ such that

$$\begin{aligned} \mathbb{P}(|g(X_t, s_t, b_t)| \leq BKr^t\eta | A_\eta) &\leq \mathbb{P}(X_t \in I_{2K^t\eta}(s_t, b_t) | A_\eta) \\ &= \int \mathbb{P}(X_t \in I_{2K^t\eta}(s_t, b_t) | s_t, b_t, A_\eta) d\mathbb{P}(s_t, b_t) \\ &\leq \int 2UKr^t\eta d\mathbb{P}(s_t, b_t) = 2UKr^t\eta. \end{aligned}$$

It follows that

$$\begin{aligned} \mathbb{P}(|g(X_t, s_t, b_t)| > BKr^t\eta, \forall t \in \mathbb{N} | A_\eta) \\ &= 1 - \mathbb{P}(|g(X_t, s_t, b_t)| \leq BKr^t\eta, \exists t \in \mathbb{N} | A_\eta) \\ &\geq 1 - \sum_{t=1}^{\infty} \mathbb{P}(|g(X_t, s_t, b_t)| \leq BKr^t\eta | A_\eta) \\ &\geq 1 - \sum_{t=1}^{\infty} 2UKr^t\eta \geq 1 - \frac{2UK\eta r}{1-r}. \end{aligned}$$

This last number can be made larger than zero by choosing η sufficiently small. ■

Partial solutions and continuous differentiability

The convergence of partial solutions to the true ones is essentially almost the same as the one for the existence of a stationary ergodic solution. We can use the same bounds as in statement (i) to show that $|\mu_t^*|$ and $|\hat{\mu}_t|$ are bounded by some η with positive probability and that their respective bubble processes are zero. It then follows by the same derivation as in part (ii) that they converge with positive probability. As $(X_t)_{t \in \mathbb{Z}}$ is stationary ergodic this event happens with probability one at some point in time and thus we get the convergence.

Continuous differentiability follows by the same way as in Straumann and Mikosch (2006). The stochastic recurrence equations for the derivatives of the fundamental and bubble processes are either linear or standard resetting systems. Therefore their respective backward iterations converge to stationary ergodic solutions. This then implies the continuous differentiability by a standard analysis result.

3.6.5 Proof of Theorem 3.3.6

We follow the usual consistency proof for M -estimators which involves showing firstly the uniform convergence of the sample average log likelihood to the limit log likelihood and secondly the identifiable uniqueness of the parameter of interest; see e.g. Theorem 3.4 in White (1994) or Lemma 3.1 in Pötscher and Prucha (1997). To ease notation we define the following functions $\Theta \rightarrow \mathbb{R}$:

$$\begin{aligned}\hat{\ell}_t &= \ell(X_t, \hat{\mu}_t(\cdot, \lambda), \hat{b}_t(\cdot, \lambda), \sigma^2), \\ \ell_t^* &= \ell(X_t, \mu_t^*(\cdot, \lambda), b_t^*(\cdot, \lambda), \sigma^2).\end{aligned}$$

Lemma 3.6.4. *The sample average log likelihood almost surely converges uniformly to the limit log likelihood, i.e.*

$$\left\| \hat{L}_T - \mathbb{E} \ell_t^* \right\|_{\Theta} \xrightarrow{a.s.} 0 \quad \text{as } T \rightarrow \infty.$$

PROOF. We have

$$\left\| \hat{L}_T - \mathbb{E} \ell_t^* \right\|_{\Theta} = \left\| \frac{1}{T} \sum_{t=2}^T \hat{\ell}_t - \mathbb{E} \ell_t^* \right\|_{\Theta} \leq \frac{1}{T} \sum_{t=2}^T \left\| \hat{\ell}_t - \ell_t^* \right\|_{\Theta} + \left\| \frac{1}{T} \sum_{t=2}^T \ell_t^* - \mathbb{E} \ell_t^* \right\|_{\Theta}. \quad (3.8)$$

We will show that the two rightmost terms in (3.8) go to zero as $T \rightarrow \infty$. For the first term note that ℓ is a differentiable function, we write

$$\ell_f(\tilde{\mu}, \tilde{b}) = \left. \frac{\partial \ell(X_t, \mu, b, \sigma^2)}{\partial (\mu, b)} \right|_{(\tilde{\mu}, \tilde{b})}.$$

We then invoke the mean value theorem to obtain the existence of some $(\tilde{\mu}_t, \tilde{b}_t)$ between $(\hat{\mu}_t, \hat{b}_t)$ and (μ_t^*, b_t^*) that satisfies

$$\begin{aligned}\left\| \hat{\ell}_t - \ell_t^* \right\|_{\Theta} &\leq \left\| \ell_f(\tilde{\mu}_t, \tilde{b}_t) \right\|_{\Theta} \left\| (\hat{\mu}_t, \hat{b}_t) - (\mu_t^*, b_t^*) \right\|_{\Theta} \\ &\leq \left\| \ell_f(\tilde{\mu}_t, \tilde{b}_t) - \ell_f(\mu_t^*, b_t^*) \right\|_{\Theta} \left\| (\hat{\mu}_t, \hat{b}_t) - (\mu_t^*, b_t^*) \right\|_{\Theta} \\ &\quad + \left\| \ell_f(\mu_t^*, b_t^*) \right\|_{\Theta} \left\| (\hat{\mu}_t, \hat{b}_t) - (\mu_t^*, b_t^*) \right\|_{\Theta}\end{aligned} \quad (3.9)$$

The function ℓ_f is linear in its arguments and thus is a K -Lipschitz function for some $K > 0$. Therefore assumption CS 2 guarantees that

$$\begin{aligned} \left\| \ell_f(\tilde{\mu}_t, \tilde{b}_t) - \ell_f(\mu_t^*, b_t^*) \right\|_{\Theta} &\leq K \left\| (\tilde{\mu}_t, \tilde{b}_t) - (\mu_t^*, b_t^*) \right\|_{\Theta} \\ &\leq K \left\| (\hat{\mu}_t, \hat{b}_t) - (\mu_t^*, b_t^*) \right\|_{\Theta} \xrightarrow{a.s.} 0 \quad \text{as } t \rightarrow \infty, \end{aligned}$$

hence (3.9) almost surely goes to zero exponentially fast by assumption CS 2 and thus we have

$$\frac{1}{T} \sum_{t=2}^T \left\| \ell_f(\tilde{\mu}_t, \tilde{b}_t) - \ell_f(\mu_t^*, b_t^*) \right\|_{\Theta} \left\| (\hat{\mu}_t, \hat{b}_t) - (\mu_t^*, b_t^*) \right\|_{\Theta} \xrightarrow{a.s.} 0 \quad \text{as } T \rightarrow \infty.$$

Next, note that $(\|\ell_f(\mu_t^*, b_t^*)\|_{\Theta})_{t \in \mathbb{Z}}$ is a stationary sequence by assumption CS 2 and Proposition 4.3 in Krengel (1985). Therefore

$$\frac{1}{T} \sum_{t=2}^T \|\ell_f(\mu_t^*, b_t^*)\|_{\Theta} \left\| (\hat{\mu}_t, \hat{b}_t) - (\mu_t^*, b_t^*) \right\|_{\Theta} \xrightarrow{a.s.} 0 \quad \text{as } T \rightarrow \infty$$

if $\|\ell_f(\mu_t^*, b_t^*)\|_{\Theta}$ has a log moment by assumption CS 2 and Lemma 2.1 in Straumann and Mikosch (2006). Let $\log^+(x) = \max\{0, \log x\}$. The log moment follows from the fact that

$$\mathbb{E} \log^+ \|\ell_f(\mu_t^*, b_t^*)\|_{\Theta} = \frac{1}{\sigma^2} \mathbb{E} \log^+ \left\| \begin{array}{c} X_t - \mu_t^* - b_t^* \\ X_t - \mu_t^* - b_t^* \end{array} \right\|_{\Theta},$$

the finiteness of which is implied by the moment conditions in assumptions CS 1 and CS 2. We conclude that the first term in (3.8) converges to zero almost surely.

Finally, we discuss the second term in (3.8). We show that

$$\left\| \frac{1}{T} \sum_{t=2}^T \ell_t^* - \mathbb{E} \ell_t^* \right\|_{\Theta} \xrightarrow{a.s.} 0 \quad \text{as } T \rightarrow \infty$$

by application of the uniform law of large numbers, Theorem 2.7 in Straumann and Mikosch (2006). The law of large numbers holds since $(\ell_t^*)_{t \in \mathbb{N}}$ is strictly stationary and

ergodic by assumption CS 2 and Proposition 4.3 in Krengel (1985), and because

$$\begin{aligned}\mathbb{E} \|\ell_t^*\|_{\Theta} &= \mathbb{E} \left\| \frac{1}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} (X_t - \mu_t^* - \beta_t^*)^2 \right\|_{\Theta} \\ &\leq \left\| \frac{1}{2} \log(2\pi\sigma^2) \right\|_{\Theta} + c \left\| \frac{1}{2\sigma^2} \right\|_{\Theta} (\mathbb{E} X_t^2 + \mathbb{E} \|\mu_t^*\|_{\Theta}^2 + \mathbb{E} \|b_t^*\|_{\Theta}^2),\end{aligned}$$

for some $c > 0$. This upper bound is finite by assumption CS 2, because Θ is compact and the fact that $\sigma^2 > 0$. ■

Lemma 3.6.5. *The parameter θ_0 is identifiable unique on Θ .*

PROOF. The identifiable uniqueness of $\theta_0 \in \Theta$ is implied by the uniqueness assumption CS 3, the continuity of $\mathbb{E}\ell_t^*$ and the compactness of Θ , see Chapter 3 in Pötscher and Prucha (1997). The continuity of $\mathbb{E}\ell_t^*$ follows directly from the fact that the sample likelihood, which is continuous, converges uniformly to $\mathbb{E}\ell_t^*$. ■

3.6.6 Proof of Corollary 3.3.7

Theorem 3.3.6 ensures that $(\mu_t^*, b_t^*)_{t \in \mathbb{Z}} = (\mu_t, b_t)_{t \in \mathbb{Z}}$, so that condition CS 2 follows. The maximiser of the limit log likelihood is equal to the minimiser of the Kullback-Leibler divergence between the true conditional density of the data and the model-implied conditional density, see for instance Section 2.3 of White (1994). Therefore it follows by the Gibbs inequality that the limit log likelihood is uniquely maximised at the true parameter θ_0 and thus condition CS 3 holds for the true parameter.

3.6.7 Proof of Theorem 3.3.8

This proof is identical to Section 7 of Straumann and Mikosch (2006).

3.6.8 Proof of Theorem 3.3.9

The desired result follows by the same argument as used above for proving asymptotic normality under correct specification, with the exception that the score is not granted to

be a martingale difference sequence. However, by assumptions AN 5 and AN 6, we have that the score sequence is near epoch dependent of size -1 on a ϕ -mixing sequence of size $-r/(r-1)$ for some $r > 2$. Given the moment bounds, we can thus appeal to the central limit theorem for near epoch dependent sequences in Potscher and Prucha (1997, Theorem 10.2).

Chapter 4

Transformed Perturbation Solutions for Dynamic Stochastic General Equilibrium Models

4.1 Introduction

Since the seminal paper of Kydland and Prescott (1982) many different methods have been proposed to approximate the solution of Dynamic Stochastic General Equilibrium (DSGE) models, see for example Taylor and Uhlig (1990), Christiano and Fisher (2000) and Aruoba et al. (2006) for comparison studies. It is well known that, in most cases, closed form analytical solutions do not exist, and hence we need numerical solution methods.

When selecting solution methods, two properties are of main interest: *speed and accuracy*. On the one hand, arbitrarily accurate solution algorithms such as value function iteration (Bertsekas, 1987) and projection methods (Judd, 1992) have existed for a long time. However, such methods need long computing times. This is problematic, especially when one is interested in estimating a DSGE model, since then the solution will have to be computed for a range of different parameter values. On the other hand, very fast solution methods such as linearization (Blanchard and Kahn, 1980) and higher-order perturbation methods (Judd and Guu, 1997; Schmitt-Grohé and Uribe, 2004) are available. These methods approximate the solution by taking a Taylor series expansion around the

CHAPTER 4. TRANSFORMED PERTURBATION SOLUTIONS FOR DYNAMIC STOCHASTIC GENERAL EQUILIBRIUM MODELS

deterministic steady state. Unfortunately, despite being very fast, perturbation methods also have important limitations.

Linearization, or first order perturbation, can be very inaccurate and is often too simplistic from an economic perspective. For example, linear solutions are certainly equivalent and therefore miss potential volatility dynamics in the innovations. That means that one needs higher order perturbation methods for risk to matter, which affects a multitude of topics. For instance, this is a relevant limitation when attempting to model time varying risk premia as in Fernández-Villaverde et al. (2011); Rudebusch and Swanson (2012); Fernández-Villaverde et al. (2015) and requires a perturbation approximation of at least third order to be solved. Similarly, linearization is highly inaccurate when comparing welfare across different environments and can lead to paradoxical results (Tesar, 1995). Kim and Kim (2003b) show that a welfare comparison based on a linear approximation of the policy function may yield spurious results in a two-agent economy and that perturbation approximations of at least second order are required. Some welfare studies that use higher order perturbation approximations can be found in Kollmann (2002), Kim and Kim (2003a) and Bergin et al. (2007).¹ Finally, Van Binsbergen et al. (2012) discuss the need for higher order perturbation solutions to study consumer risk aversion.

The speed of perturbation methods and their ability to locally capture important non-linear dynamics renders high-order perturbation a popular solution method. However, higher-order perturbation is an unattractive approximation method from a global perspective as it defines an unstable dynamic system which produces explosive paths. In fact, one can commonly show that sample paths generated using higher-order perturbations diverge to infinity almost surely, even if the true policy function implies stable dynamics with nonexplosive paths. This problem is outlined in Aruoba et al. (2006) and Den Haan and De Wind (2010) and encountered in Fahr and Smets (2010) and Den Haan and De Wind (2012), among others. See Section 3.3.2 and Section 5 in Den Haan and De Wind (2010) for extensively discussed examples.

In order to deal with the unstable dynamics of higher-order perturbation solutions, Kim et al. (2008) proposed the pruning method. The pruning method has been successfully implemented in software packages and effectively solves the problem of explosive

¹Woodford (2002) discusses a set of assumptions that ensure first order approximations are sufficient.

dynamics; see also Andreasen et al. (2017) for recent results on the stability and stationarity of pruned solutions. However, pruned solutions must sacrifice local approximation accuracy for stability. Den Haan and De Wind (2010) show that pruning “*creates large systematic distortions*”. Furthermore, pruning is a simulation-based approximation and hence does not provide a policy function. In fact, approximations based on the pruning procedure contain different updates for identical values of the model’s original state variables. This means that “*the implied policy rule is not even a function of the model’s state variables*” (Den Haan and De Wind, 2010).

Our paper introduces a new transformed perturbation solution method for DSGE models that is designed to avoid explosive paths produced by higher-order perturbation solutions. Transformed perturbation is as fast as standard perturbation methods and can be easily implemented in existing software packages like *Dynare* as it is obtained directly as a transformation of existing perturbation solutions. The new method transforms the standard perturbation approximation by replacing higher order monomials in the Taylor expansion with transformed ones that are based on the transformed polynomials introduced in Blasques et al. (2014). Transformed polynomial functions share the same fundamental approximation properties as polynomial functions. Blasques et al. (2014) shows that transformed polynomials are dense in the space of continuous functions and attain the same rates of convergence as polynomials in Sobolev spaces of n times continuously differentiable functions. Additionally, in this paper, transformed perturbation is shown to converge on analytic function domains and to have the same excellent local properties as the standard perturbation method for continuously differentiable functions of appropriate order. From a global perspective however, transformed perturbation performs infinitely better than regular perturbation, because it provides a way of scaling down the higher order perturbation terms that cause explosive behavior when the solution path moves far away from the steady state. That way, transformed perturbation can be guaranteed to not create additional fixed points if the Blanchard-Kahn conditions are satisfied. Moreover, unlike pruning, the new solution method does not need to sacrifice accuracy by ignoring higher order effects. Additionally, transformed perturbation is guaranteed to be always more accurate than standard perturbation methods, which is not the case for pruned solutions. Finally, in contrast to pruning, transformed perturbation also has the advantage of

delivering a policy function from which the simulations are drawn.

In this paper, we prove that transformed perturbation produces non explosive paths and that solutions are stable and strictly stationary ergodic with bounded moments. Additionally, we show that solution paths exhibit fading memory (i.e. geometric ergodicity and absolute regularity or β -mixing) and that sample moments of the process converge exponentially fast to the moments of the solution. These are crucial properties for conducting simulation-based estimation of parameters and simulation-based analysis of the DSGE model. Overall, this renders the transformed polynomial solution attractive from both a practical and theoretical stand-point.

We demonstrate the accuracy of the transformed perturbation method extensively for two nonlinear DSGE models in which higher order perturbation is infeasible. We compare second order transformed perturbation to first order perturbation and second order pruning. The first model is a partial equilibrium model in which agents face idiosyncratic income risk, introduced in Deaton Angus (1991) and Den Haan and De Wind (2012). For this model we find that sample path errors of our method are less than half of those of pruning and up to six times less than those for first order perturbation. This then results into sample moments of the transformed perturbation method being up to ten times more accurate than pruning and one-hundred times more accurate than perturbation. The second DSGE model we study is a matching model from Den Haan and De Wind (2012). Here transformed perturbation outperforms pruning up to a factor ten on path errors and a factor thirty for sample moments. Moreover perturbation has path errors that are up to twenty-five times larger and sample moment errors that are up to one-hundred times larger compared to transformed perturbation.

The paper is structured as follows. We start by stating the definition of the transformed perturbation method in Section 4.2. Section 4.3 analyses the statistical properties of the transformed perturbation system and provides lenient and accessible conditions that ensure paths are nonexplosive and laws of large numbers can be applied. Section 4.4 provides a theoretical foundation and motivation for the transformed perturbation approximation method. Finally Section 4.5 discusses the accuracy of the new method. We provide theoretical results that show that transformed perturbation, accuracy wise, matches regular perturbation locally and strongly outperforms it globally. Moreover, we demonstrate

for two example models that transformed perturbation outperforms pruning and regular perturbation on numerous common criteria.

4.2 Transformed Perturbation

4.2.1 The state space

Let $\bar{\mathbf{y}}_t$ be an n_y -dimensional vector of control variables, let $\bar{\mathbf{x}}_t$ be an n_x -dimensional vector of endogenous state variables and let \mathbf{z}_t be an n_z -dimensional vector of exogenous state variables. We study the general class of DSGE models characterized by a set of first-order dynamic optimality conditions that can be written as

$$0 = \mathbb{E}_t(f(\bar{\mathbf{y}}_{t+1}, \bar{\mathbf{y}}_t, \bar{\mathbf{x}}_{t+1}, \bar{\mathbf{x}}_t, \mathbf{z}_{t+1}, \mathbf{z}_t)), \quad (4.1)$$

$$\mathbf{z}_{t+1} = \mathbf{\Lambda} \mathbf{z}_t + \sigma \boldsymbol{\eta} \varepsilon_{t+1}. \quad (4.2)$$

Here \mathbb{E}_t denotes the expectation operator conditional on the information at time t , and $f : \mathbb{R}^{2(n_x+n_y+n_z)} \rightarrow \mathbb{R}^{n_y+n_x}$ is a real function. The matrix $\mathbf{\Lambda}$ is assumed to be invertible with spectral radius smaller than one. Finally σ is the auxiliary perturbation parameter and ε_{t+1} is a n_z -dimensional vector of exogenous innovations with mean zero and finite second moment that takes values in $\mathcal{E} \subseteq \mathbb{R}^{n_z}$. Throughout the paper, we will assume that $(\varepsilon_t)_{t \in \mathbb{N}}$ is an independent and identically distributed (iid) stochastic process.

We define the deterministic steady states \mathbf{y}_{ss} and \mathbf{x}_{ss} of $\bar{\mathbf{y}}_t$ and $\bar{\mathbf{x}}_t$ respectively such that

$$f(\mathbf{y}_{ss}, \mathbf{y}_{ss}, \mathbf{x}_{ss}, \mathbf{x}_{ss}, \mathbf{0}_{n_z}, \mathbf{0}_{n_z}) = 0.$$

Furthermore, let $\mathbf{y}_t = \bar{\mathbf{y}}_t - \mathbf{y}_{ss}$ and $\mathbf{x}_t = \bar{\mathbf{x}}_t - \mathbf{x}_{ss}$ denote the random variables in deviations from the steady-state, where \mathbf{y}_t takes values in $\mathcal{Y} \subseteq \mathbb{R}^{n_y}$ and \mathbf{x}_t takes values in $\mathcal{X} \subseteq \mathbb{R}^{n_x}$. We write $\mathcal{Z} \subseteq \mathbb{R}^{n_z}$ for the domain of \mathbf{z}_t . Following Den Haan and De Wind (2012), the

solution to the model given in equation (4.1) is of the form

$$\mathbf{y}_{t+1} = g(\mathbf{x}_t, \mathbf{z}_{t+1}, \sigma), \quad (4.3)$$

$$\mathbf{x}_{t+1} = h(\mathbf{x}_t, \mathbf{z}_{t+1}, \sigma). \quad (4.4)$$

We refer to (4.3) and (4.4) as the observation and state equations respectively. It follows from our setup that $g(\mathbf{0}_{n_x}, \mathbf{0}_{n_z}, 0) = \mathbf{0}_{n_y}$ and $h(\mathbf{0}_{n_x}, \mathbf{0}_{n_z}, 0) = \mathbf{0}_{n_x}$.

Both functions g and h , known as *policy functions*, are unknown functions that must be approximated. If the function g in the observation equation is measurable, then the stability of the solution of a DSGE model depends entirely on the state equation. For this reason we will focus on approximating the function h in (4.4).

4.2.2 Function approximation methods

A wide range of techniques have been proposed in the literature to approximate the unknown policy function h . In most cases, the approximate policy function is obtained as an element of a vector space spanned by a set of basis functions $\{\phi_1, \dots, \phi_m\}$:

$$h(\mathbf{x}, \mathbf{z}, \sigma) \approx \sum_{i=1}^m A_i \phi_i(\mathbf{x}, \mathbf{z}, \sigma),$$

where A_1, \dots, A_m are matrices of coefficients that weight the basis functions ϕ_1, \dots, ϕ_m .

There exist a multitude of popular sets of basis functions and weight matrix calculation methods that have been proposed in the function approximation literature. Well known classes of basis functions include *power monomials*, which are used with great success in Taylor expansions, *sigmoid trigonometric functions*, that are prominently featured in Fourier approximations, *Chebyshev polynomials*, that play an important role on orthogonal polynomial function approximation, *Legendre polynomials*, which are often used for approximating density functions, and *logistic functions*, comprehensively explored in artificial neural network approximations. Popular methods for calculating the weight matrices include *Taylor's method* which obtains the matrices as weighted derivatives at a given expansion point and minimizes the so called Taylor semi-norm (Apostol, 1967), *function collocation methods*, which minimize a discrete distance between the true

and approximate policy function at a finite number of points and *spectral approximation methods*, that minimize a continuous distance between the two functions. See e.g. Powell (1981) for an overview of approximation literature and Judd (1998) for an application of these methods to approximating policy functions of dynamic stochastic models.

4.2.3 Perturbation

Perturbation is a method that approximates the unknown policy function h by using power monomials as basis functions in combination with Taylor's method to find the weighting matrices. This method is of particular interest in approximating policy functions of DSGE models as it provides a fast and analytically tractable way of obtaining the weighting matrices. The expansion point used in Taylor's method is the deterministic steady state $(\mathbf{0}_{n_x}, \mathbf{0}_{n_z}, 0)$. Choose $\mathbf{x} \in \mathcal{X}$ and $\mathbf{z} \in \mathcal{Z}$ and define $\mathbf{v} = (\mathbf{x}, \mathbf{z})$ and $\bigotimes_i \mathbf{v} = \underbrace{\mathbf{v} \otimes \cdots \otimes \mathbf{v}}_{i \text{ times}}$, where the empty Kronecker product is set to one. Then the m 'th order perturbation approximation of h evaluated at $(\mathbf{x}, \mathbf{z}, \sigma)$ can be expressed as

$$h_p(\mathbf{x}, \mathbf{z}, \sigma) := H_0 + H_{\mathbf{x}}\mathbf{x} + H_{\mathbf{z}}\mathbf{z} + \sum_{i=2}^m H_i \bigotimes_i \mathbf{v}, \quad (4.5)$$

where we grouped all terms of \mathbf{v} of the same power, regarding σ as a constant. That is,

$$\begin{aligned} H_0 &= \sum_{j=0}^m \frac{1}{j!} \frac{\partial^j}{\partial \sigma^j} h(\mathbf{0}_{n_x}, \mathbf{0}_{n_z}, 0) \sigma^j & H_{\mathbf{x}} &= \sum_{j=0}^{m-1} \frac{1}{j!} \frac{\partial^{j+1}}{\partial \sigma^j \partial \mathbf{x}} h(\mathbf{0}_{n_x}, \mathbf{0}_{n_z}, 0) \sigma^j \\ H_{\mathbf{z}} &= \sum_{j=0}^{m-1} \frac{1}{j!} \frac{\partial^{j+1}}{\partial \sigma^j \partial \mathbf{z}} h(\mathbf{0}_{n_x}, \mathbf{0}_{n_z}, 0) \sigma^j & H_i &= \sum_{j=0}^{m-i} \frac{1}{i!j!} \frac{\partial^{j+i}}{\partial \sigma^j \partial \mathbf{v}^i} h(\mathbf{0}_{n_x}, \mathbf{0}_{n_z}, 0) \sigma^j \end{aligned}$$

Thus, H_0 is an $n_x \times 1$ vector that is the sum of all the derivatives of h with respect to powers of σ . The matrix $H_{\mathbf{x}}$ is an $n_x \times n_x$ matrix that is the sum of all the derivatives of h with respect to \mathbf{x} and powers of σ . The matrix $H_{\mathbf{z}}$ is an $n_x \times n_z$ matrix that is the sum of all the derivatives of h with respect to \mathbf{z} and powers of σ . Finally, the matrices H_i are of dimension $n_x \times (n_x + n_z)^i$ and given by the sum of all the derivatives of h with respect to \mathbf{v}^i and powers of σ .

4.2.4 The transformed perturbation method

A disadvantage of the power monomial set of basis functions, and therefore of perturbation, is that the derivative of the approximation function tends to infinity away from the steady state if $m > 1$. This creates highly explosive regions in the state space which in practice means that sample paths eventually diverge to infinity with probability one. The transformed perturbation method solves this problem by using another set of basis functions called the transformed power monomials. This set of basis functions satisfies all the advantageous properties that classical power monomials do. Blasques et al. (2014) shows that transformed polynomials with unrestricted weighting matrices can be used to approximate continuous functions with arbitrary accuracy, in the same way as classical polynomials, by application of the *Stone-Weierstrass Theorem* (Stone, 1937, 1948). Additionally, Blasques et al. (2014) characterizes the convergence rates of transformed polynomials on Sobolev spaces of smooth n -times continuously differentiable functions with n th derivative bounded in L_p norm, through the application of Plesniak's extension of *Jackson's Theorem* (Plesniak, 1990).

The set of transformed power monomials is obtained by multiplying the monomials of order greater than one with an exponentially fast decaying function $\Phi_\tau : \mathcal{X} \rightarrow \mathbb{R}$ that is a multivariate adaptation of the transformed function of Blasques et al. (2014) and is defined as

$$\Phi_\tau(\mathbf{x}) = e^{-\tau \|\mathbf{x}\|_e^2}, \quad (4.6)$$

where $\|\mathbf{x}\|_e$ denotes the Euclidean norm of \mathbf{x} . Figure 4.2.1 plots the second and third order one dimensional transformed monomials for varying values of τ . Note that the case τ is zero sets the transformed monomials equal to the regular monomial basis functions. The figure shows that the transformed monomials are almost identical to regular monomials close to the steady state at zero. However, the derivatives of transformed monomials vanish away from the steady state, which implies that no explosive regions are created in the state space. In Section 4.5 we will further show that transformed perturbation has the same local approximation properties as classical perturbation. In particular, local approximation rates are the same as for classical perturbation, and transformed perturbation

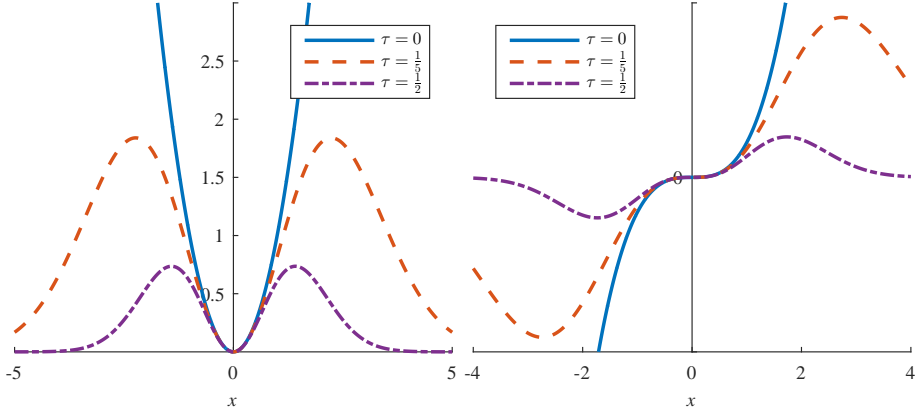


Figure 4.2.1: Plots of the second, respective third, order one dimensional transformed monomial in the left, respective right, panel for values of $\tau \in \{0, 0.2, 0.5\}$.

approximations converge uniformly on compact analytic domains, just like perturbation methods do. A large number of additional advantages of transformed perturbation over classical perturbation and pruning methods are documented in Section 4.3 and Section 4.5.

State variables can be of different orders in size, so we replace the vector \mathbf{x} in (4.6) by the relative differences from the steady state $\tilde{\mathbf{x}} = \mathbf{x}/\mathbf{x}_{ss}$, where dividing is done entry wise, to ensure all variables have equal effect. This definition works poorly if an entry of \mathbf{x}_{ss} is close to zero. For such an entry we take the simple transformation $x \mapsto e^x \approx 1 + x$, which is almost linear close to zero, and define $\tilde{x} = (e^{x+x_{ss}} - e^{x_{ss}})/e^{x_{ss}}$. The m 'th order transformed perturbation approximation of h evaluated at $(\mathbf{x}, \mathbf{z}, \sigma)$ is then defined as

$$h_{tp}(\mathbf{x}, \mathbf{z}, \sigma) = H_0 + H_{\mathbf{x}}\mathbf{x} + H_{\mathbf{z}}\mathbf{z} + \left(\sum_{i=2}^m H_i \bigotimes_i \mathbf{v} \right) \Phi_{\tau}(\tilde{\mathbf{x}}), \quad (4.7)$$

where all the H matrices are obtained using Taylor's method and thus they are identical to those in the regular perturbation function (4.5).

The constant τ determines the speed at which the higher order terms in (4.7) are going to zero when moving away from the origin. Its value influences the shape of the resulting policy function, and thus requires careful consideration. We offer two methods to set τ . The first method is to find the optimal τ , denoted τ^* , by minimizing some criterion

function. In this paper we chose to minimise the maximum Euler errors on a relevant set around the steady state. The advantage of this method is that we get the best possible value for τ , according to the criterion function. The disadvantage is that minimizing the criterion function potentially is time-consuming. In an estimation setting we fix the optimal τ^* at the start and then estimate the remaining parameters while τ^* remains fixed. This means that the possibly time consuming task of finding τ^* has to be executed only once, making the method almost as fast as perturbation, still viable for estimation and very accurate if the optimal τ^* does not vary too much with the parameters. The second method is designed to avoid the optimization completely and is characterised by a plug-in τ , denoted $\hat{\tau}$, which is less precise, but found immediately. The plug-in value is given by

$$\hat{\tau} = \frac{1}{c} \log \left(\frac{1}{1 - \rho(H_{\mathbf{x}})} \right), \quad (4.8)$$

where $\rho(H_{\mathbf{x}})$ is the spectral value of the autoregressive part of the regular perturbation solution and c is an approximation of the average range that the state variables take place in. This range could be set according to prior knowledge on the variables, or approximated by another solution method. In our case we used linear perturbation to simulate a series and find the approximate range of our variables. In an estimation setting we update $\hat{\tau}$ as the parameters are updated, since its calculation is very fast. See Section 4.4 for a detailed discussion on the choice for our plug-in value.

4.3 Probabilistic analysis of the solutions

Throughout this paper we work with norms $\|\cdot\|$ on Euclidean space and their induced matrix norms, which we will denote with the same notation $\|\cdot\|$ as there should be no confusion in their use. Note that all matrix norms are equivalent, so that our statements will work for any chosen norm.

Let $\mathbf{x}_0 \in \mathcal{X}$ and $\mathbf{z}_0 \in \mathcal{Z}$ be fixed and define the exogenous sample paths $(\mathbf{z}_t)_{t \geq 0}$ and

the transformed perturbation sample paths $(\mathbf{x}_t)_{t \geq 0}$ recursively by

$$\begin{aligned}\mathbf{z}_{t+1} &= \mathbf{\Lambda} \mathbf{z}_t + \sigma \boldsymbol{\eta} \varepsilon_{t+1}, \\ \mathbf{x}_{t+1} &= h_{tp}(\mathbf{x}_t, \mathbf{z}_{t+1}, \sigma).\end{aligned}$$

In this section we analyse the dynamics of the transformed perturbation system and provide two results on the stability of sample paths. To do so we split the perturbation updating equation (4.7) into the sum of its linear part $H_0 + H_{\mathbf{x}}\mathbf{x} + H_{\mathbf{z}}\mathbf{z}$ and its nonlinear part

$$D(\mathbf{x}, \mathbf{z}) := \left(\sum_{i=2}^m H_i \bigotimes_i \mathbf{v} \right) \Phi_{\tau}(\tilde{\mathbf{x}}). \quad (4.9)$$

Our results are based on the observation that the transformed perturbation policy function (4.7) is asymptotically equal to its linear part as $\|\mathbf{x}\| \rightarrow \infty$. This follows because an exponential function decays at greater speed than a polynomial, see Figure 4.2.1, and thus for any $0 \leq i \leq m$ we have

$$\lim_{\|\mathbf{x}\| \rightarrow \infty} \left(\bigotimes_i \mathbf{x} \right) \Phi_{\tau}(\tilde{\mathbf{x}}) = \mathbf{0}_{n_x^i}.$$

We therefore study the transformed perturbation method as its asymptotic linear process plus a deviation (4.9). Linear autoregressive processes and their stability have been extensively studied. They are much easier to analyse compared to their nonlinear counterparts, because we get an analytical closed form when we expand the expressions for \mathbf{x}_t and \mathbf{z}_t back in time. It can be shown that if backwards expanding converges, then the limit is a stationary ergodic solution to the system. See Theorem 3.1 in Bougerol (1993) for a general result on the stability of contracting systems that uses this approach. In our first result we closely mimic this technique by bounding the deviation from the linear process. We require the following assumptions.

Assumption A.

A1. The spectral radius $\rho(\mathbf{\Lambda}) < 1$.

A2. The spectral radius $\rho(H_{\mathbf{x}}) < 1$.

A3. There exists an $r > 0$ such that $\mathbb{E}\|\varepsilon_t\|^{rm} < \infty$.

Our first result shows that solution paths generated by the transformed perturbation solution are non-explosive almost surely if Assumption A holds. The conditions in Assumption A are very lenient. Assumption A2 is close to being both sufficient and necessary. The spectral radius is a measure for the maximal scale at which H_x can stretch a vector. Therefore, if $\rho(H_x) > 1$, then an eigenvector belonging to the eigenvalue that is greater than one in absolute value is expanded by H_x . If the space spanned by this vector is reachable from the exogenous variable space \mathcal{Z} , then expanding backwards will explode and thus diverge. Assumption A3 is satisfied for any r if, for example, the ε_t have finite support, or are normally distributed, or have sub-exponential tails. Additionally, for fat tailed distributions, the moments of x_t and z_t are a fraction of those of the innovations.

Theorem 4.3.1 (Non explosive paths). *Suppose that Assumption A holds. Then the dynamic system defined in (4.2) and (4.4), featuring the transformed perturbation policy function given in (4.7), produces sample paths that are non explosive almost surely, i.e. the paths $(z_t)_{t \in \mathbb{N}}$ and $(x_t)_{t \in \mathbb{N}}$ satisfy*

$$\liminf_{t \rightarrow \infty} \|z_t\| < \infty \quad \text{and} \quad \liminf_{t \rightarrow \infty} \|x_t\| < \infty \quad a.s.$$

Theorem 4.3.1 shows that the transformed perturbation method does not produce explosive paths, unlike regular perturbation sample paths. However, we can show much more. Our stability results are based on Markov chain theory as developed in Meyn and Tweedie (1993). We are in a Markov chain setting, because we have assumed that $(\varepsilon_t)_{t \in \mathbb{N}}$ is an iid sequence. We provide two sets of assumptions, the first of which is more general and harder to verify, while the second set imposes additional constraints that are straightforward to verify.

A point $x^* \in \mathcal{X}$ is called *reachable* if for every open set $x^* \in O \subseteq \mathcal{X}$ and starting value $x_0 \in \mathcal{X}$ there exists a $t \in \mathbb{N}$ such that $\mathbb{P}(x_t \in O) > 0$. A subset of \mathcal{X} is called *reachable* if all the points in it are reachable. We will need the following additional assumptions.

Assumption B.

- B1. \mathcal{X} has an open reachable subset.
- B2. The innovation ε_t is absolutely continuous with respect to the Lebesgue measure on \mathcal{E} with strictly positive density on a connected subset of \mathcal{E} .

Our second result establishes the stationarity and ergodicity of the transformed perturbation solution. Additionally, it shows that the solution paths have fading memory in the sense of geometric ergodicity and absolutely regularity (or β -mixing) of the process. Finally, it also shows that the solution paths have finite r -th moment. Stationarity, fading memory, and bounded moments are all important ingredients in the statistical analysis of DSGE models, from estimation to probabilistic analysis.

Theorem 4.3.2. (Stationarity, fading-memory and bounded moments) *Suppose that Assumptions A and B hold. Then there exists a unique stationary ergodic solution $(\mathbf{x}_t^*, \mathbf{z}_t^*)_{t \geq 0}$ to the dynamic system defined in (4.2) and (4.4), featuring the transformed perturbation policy function given in (4.7). Additionally,*

- (i) *the solution has fading memory, i.e. it is geometrically ergodic and absolutely regular (or β -mixing);*
- (ii) *the solution has finite moments $\mu_r := \mathbb{E}\|\mathbf{x}_t^*\|^r$ and $\nu_{rm} = \mathbb{E}\|\mathbf{z}_t^*\|^{rm}$;*
- (iii) *laws of large numbers apply to the sample paths, that is, almost surely*

$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \|\mathbf{x}_t\|^r = \mu_r \quad \text{and} \quad \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \|\mathbf{z}_t\|^{rm} = \nu_{rm}.$$

Assumption B imposes additional conditions on our state space system. Assumption B2 is quite weak and is satisfied for all distributions that are used in practice. The stronger, and also harder to check, condition is Assumption B1. We present Assumption B, because simplifying Assumption B1 will require us to assume that the innovations have full support. This is not always the case, as we might, for example, have strictly positive innovations. If we can make the assumption of full support, then we get an easier set of conditions.

Assumption C.

- C1. There exists an integer $t \geq 1$ such that the matrix $\begin{bmatrix} H_x^{t-1} H_z & \cdots & H_x H_z & H_z \end{bmatrix}$ has rank n_x .
- C2. The matrix H_x is invertible.
- C3. The innovation ε_t is absolutely continuous with respect to the Lebesgue measure on \mathbb{R}^{n_z} with strict positive density on the whole space \mathbb{R}^{n_z} .

Proposition 4.3.3. *Assumption C implies Assumption B.*

Assumption C2 ensures that the transformed perturbation policy function does not move to lower dimensional subspaces of \mathcal{X} . Condition C1 implies that the effect of the innovations is not contained in a lower dimensional subspace. This means that, together with Assumption C3, they make sure that the transformed perturbation policy function can reach any point in \mathcal{X} and thus Assumption B1 is satisfied.

4.4 The plug-in tau

In this section we motivate our choice for $\hat{\tau}$, the plug in value of τ , as defined in (4.8). As mentioned in Section 4.2, its value influences the shape of the transformed perturbation policy function and thus has an effect on sample path behaviour in the resulting transformed perturbation dynamic system. We want to ensure two important properties for this dynamic system. Firstly, we want sample paths to be stable and non locally explosive. In Section 4.4.1 we argue that this requires relatively large values of τ . Secondly, nonlinear dynamics must be preserved, which needs τ to take on somewhat small values, see Section 4.4.2. Together these two conditions specify a rather narrow collection of available functions, resulting in (4.8), as derived in Section 4.4.3.

4.4.1 Ensuring stability

The transformed perturbation method guarantees stable and nonexplosive paths regardless of the choice of τ , as proved in Section 4.3. However, picking τ very small can create locally explosive dynamics. Locally explosive dynamics originate when the jacobian of

the policy function with respect to \mathbf{x} has expected spectral radius greater than one on a large enough subset of \mathcal{X} . A spectral radius greater than one implies that the policy function expands on some subspace, which can create multiple fixed points, as happens with the regular perturbation policy function. Sample paths then typically move around one fixed point, until a large innovation pushes it to another fixed point after which the path moves around the new one. These jumps can locally look very similar to explosive sample paths, even though the dynamic system is stable. We illustrate this effect with the following example updating equation

$$x_{t+1} = 0.3x_t + z_{t+1} + 2x_t^3 e^{-0.5x_t^2}, \quad (4.10)$$

where the $(z_t)_{t \in \mathbb{N}}$ are updated as in (4.2). Note that this is a univariate example of (4.7) with $\tau = 0.5$. Figure 4.4.1a plots the expected value $\mathbb{E}(x_{t+1} \mid x_t)$ as a function of x_t . This function has large intervals on which its absolute derivative exceeds one, which has

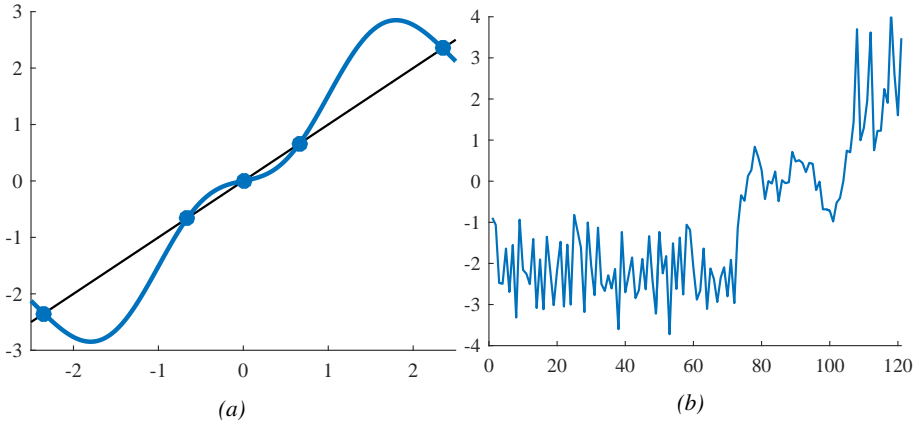


Figure 4.4.1: The expected policy function (left panel) and an example sample path (right panel) for the updating equations defined in (4.10) and (4.2).

resulted in a total of five fixed points. The smallest one at -2.35, the middle one at zero and the largest one at 2.35 are attractors while the other two are repellers. A sample path produced while using (4.10) will jump between the neighbourhoods around the three attractors. Figure 4.4.1b plots an example sample paths that first spends some time around -2.35, then jumps to a neighbourhood of the origin and then quickly moves on to the area

CHAPTER 4. TRANSFORMED PERTURBATION SOLUTIONS FOR DYNAMIC STOCHASTIC GENERAL EQUILIBRIUM MODELS

around the largest attractor. Notice the similarity with an explosive sample path, even though this path will almost surely eventually come down to the lowest attractor again.

We wish to keep the spectral value of the Jacobian of the transformed perturbation policy function typically below one (in expectation) to avoid locally explosive behaviour. This Jacobian is of the form

$$J = H_{\mathbf{x}} + P(\mathbf{x}, \mathbf{z})\Phi_{\tau}(\tilde{\mathbf{x}}),$$

where P is a m 'th order multivariate polynomial function. We can only control the non-linear part of the derivative, i.e. the second part of the summation, with our choice for τ . Any norm of P goes to infinity as $\|\mathbf{x}\|$ goes to infinity. Hence, if we choose τ too small, then $P(\mathbf{x}, \mathbf{z})\Phi_{\tau}(\tilde{\mathbf{x}})$ creates large areas on the state space with expected spectral radius greater than one. If we were only concerned with ensuring stability, then ideally we would choose $\tau = \infty$, so that Assumption A2 ensures that $\rho(J) < 1$ on the entire state space. Doing so, however, cancels all nonlinear effects making the transformed perturbation method equal to linear perturbation, which as discussed in the introduction has many flaws. Therefore we conclude that we would like to make τ as large as possible, while preserving as much nonlinear dynamics as possible close to the steady state. If we choose τ unequal to infinity, then its size generally must depend on $\rho(H_{\mathbf{x}})$. The closer $\rho(H_{\mathbf{x}})$ is to one, the less room remains available for $P(\mathbf{x}, \mathbf{z})\Phi_{\tau}(\tilde{\mathbf{x}})$. Accordingly we have to impose that τ goes to infinity as $\rho(H_{\mathbf{x}})$ gets closer to one. Therefore we must find a function $f : (0, 1) \rightarrow [0, \infty)$ such that $\tau = f(\rho(H_{\mathbf{x}}))$ and

$$\lim_{\rho(H_{\mathbf{x}}) \rightarrow 1} f(\rho(H_{\mathbf{x}})) = \infty. \quad (4.11)$$

4.4.2 Preserving nonlinear dynamics

We have concluded that we want to choose large τ to avoid locally explosive behaviour, but not so large as to destroy relevant nonlinear dynamics. In this section we formalise what we mean with preserving nonlinear dynamics. To do so we expand \mathbf{x}_t back in time,

mimic the proof of Theorem 4.3.1 and use Proposition 4.7.1 to find the upper bound

$$\|\mathbf{x}_t\| \leq \tilde{c} + \sum_{k=0}^{\infty} \|H_{\mathbf{x}}\|^k \|H_{\mathbf{z}}\| \|\mathbf{z}_{t-k}^*\| + c \sum_{j=0}^m \sum_{k=0}^{\infty} \|H_{\mathbf{x}}\|^k \tau^{-j/2} \left(\sum_{i=0}^{m-j} \|\mathbf{z}_{t-k}^*\|^i \right).$$

for some constants $c, \tilde{c} > 0$. The first, respective second, summation here is the approximate total effect over time of the linear, respective nonlinear, terms in (4.7). The first summation

$$\sum_{k=0}^{\infty} \|H_{\mathbf{x}}\|^k \|H_{\mathbf{z}}\| \|\mathbf{z}_{t-k}^*\|,$$

is the familiar term that arises in autoregressive processes.² The autoregressive part $H_{\mathbf{x}}\mathbf{x}$ of the policy function (4.7) introduces memory into the system, so that past innovations $\|\mathbf{z}_{t-k}\|$ influence the value of $\|\mathbf{x}_t\|$. The strength of the memory depends on the size of $\rho(H_{\mathbf{x}})$. If it is close to zero, then memory fades away fast and past innovations are of little weight to \mathbf{x}_t . As $\rho(H_{\mathbf{x}})$ increases, past innovations matter more up to the limit case $\rho(H_{\mathbf{x}}) = 1$, where memory does not fade anymore, at which point every past innovation is equally important and the sum diverges for all matrix norms.

We would like the impact of past innovations through the nonlinear terms of the transformed perturbation policy function to be of the same magnitude as those of the linear effect, so that both the linear and nonlinear dynamics are present in the solution paths. Specifically, we want the rate at which τ goes to infinity to be restricted such that the series

$$\sum_{k=0}^{\infty} \|H_{\mathbf{x}}\|^k \tau^{-j/2} \left(\sum_{i=0}^{m-j} \|\mathbf{z}_{t-k}^*\|^i \right)$$

diverge for all $0 \leq j \leq m$ as $\rho(H_{\mathbf{x}}) \rightarrow 1$. If this were not the case, then they would converge and thus we would restrict some nonlinear effects so much that the linear effect is infinitely stronger as $\rho(H_{\mathbf{x}})$ increases. To ease notation we define $\delta_t = \sum_{i=0}^{m-j} \|\mathbf{z}_t^*\|^i$.

²Note that it converges by Assumption A, Proposition 2.5.1 of Straumann (2005) and Proposition 4.3 of Krengel (1985).

CHAPTER 4. TRANSFORMED PERTURBATION SOLUTIONS FOR DYNAMIC STOCHASTIC GENERAL EQUILIBRIUM MODELS

The argument above then amounts to the following desired result: for all $j \in \mathbb{N}$ we have

$$\lim_{\rho(H_{\mathbf{x}}) \rightarrow 1} \sum_{k=0}^{\infty} \rho(H_{\mathbf{x}})^k \tau^{-j/2} \delta_{t-k} = \lim_{\rho(H_{\mathbf{x}}) \rightarrow 1} f(\rho(H_{\mathbf{x}}))^{-j/2} \sum_{k=0}^{\infty} \rho(H_{\mathbf{x}})^k \delta_{t-k} = \infty. \quad (4.12)$$

It is not immediately clear what divergence rates for $f(\rho(H_{\mathbf{x}}))$ satisfy (4.12). Therefore we include the following result to simplify the expression.

Lemma 4.4.1. *Suppose that $\mathbb{E}\|\varepsilon_t\|^m < \infty$. Then the limit*

$$\lim_{\rho(H_{\mathbf{x}}) \rightarrow 1} (1 - \rho(H_{\mathbf{x}})) \sum_{k=0}^{\infty} \rho(H_{\mathbf{x}})^k \delta_{t-k}$$

converges to a finite and nonzero value.

It now follows from Lemma 4.4.1 that (4.12) is equivalent to

$$\lim_{\rho(H_{\mathbf{x}}) \rightarrow 1} f(\rho(H_{\mathbf{x}}))^{j/2} (1 - \rho(H_{\mathbf{x}})) = 0. \quad (4.13)$$

4.4.3 Choice for tau

We need a function $f : (0, 1) \rightarrow (0, \infty)$ that satisfies both (4.11) and (4.13). To simplify these equations further we define $\tilde{f} : (0, \infty) \rightarrow [0, \infty)$ as $\tilde{f}(\frac{1}{1-\rho(H_{\mathbf{x}})}) = f(\rho(H_{\mathbf{x}}))$ and substitute $y = \frac{1}{1-\rho(H_{\mathbf{x}})}$. Equations (4.11) and (4.13) then can be rewritten as

$$\lim_{y \rightarrow \infty} \tilde{f}(y) = \infty \quad \text{and} \quad \lim_{y \rightarrow \infty} \frac{\tilde{f}(y)^{j/2}}{y} = 0.$$

These two equations together specify a fairly small collection of functions. To find the function that diverges fastest we consider families of familiar functions in decreasing order of rate of divergence. Note that any exponential, polynomial or radical function diverges to infinity too fast to satisfy the rightmost limit for all $j \in \mathbb{N}$. The next natural candidate in line for the rate of divergence is the logarithmic function, which leads to the specification

$$f(\rho(H_{\mathbf{x}})) = \log \left(\frac{1}{1 - \rho(H_{\mathbf{x}})} \right).$$

This is the function we used for our choice in (4.8).

The constant τ should also depend on the size of the range on which the state variables take place. Suppose that we increase the scale of our dynamic system while keeping the exact same dynamics. Then τ should become smaller as regions farther away from the steady-state are visited more often. Therefore we include the c parameter to make sure that as we make the scale larger, τ becomes smaller. Many of the other elements involved in the perturbation updating function, such as σ or H_i for $i \geq 2$ seem to be omitted in calculating the plug in τ . However, these elements have an effect on the range of the state variables and thus are implicitly included via c .

4.5 Accuracy

In this section we evaluate the accuracy of the transformed perturbation solution. In Section 4.5.1 we prove theoretic results on both global and local accuracy. We show that the optimal transformed perturbation solution is always at least as accurate as regular perturbation and demonstrate that transformed polynomials, like regular polynomials can perfectly approximate the real policy function h as we let the approximation order m go to infinity. Moreover, we prove that transformed perturbation is locally as accurate as standard perturbation and present common situations in which transformed perturbation globally outperforms regular perturbation. Section 4.5.2 discusses two DSGE models from Den Haan and De Wind (2012) and compares all discussed solution methods according to several criteria such as path errors, euler errors and produced moments. It shows that transformed perturbation outperforms pruning and regular perturbation for both the optimal τ^* and the plug in $\hat{\tau}$.

4.5.1 Theoretical results

In order to analyse the accuracy of our approximation method we define the pointwise approximation errors attained by the perturbation and transformed perturbation methods

respectively, at $(\mathbf{x}, \mathbf{z}, \sigma) \in \mathcal{X} \times \mathcal{Z} \times \mathbb{R}_{\geq 0}$ as

$$\begin{aligned} E_p(\mathbf{x}, \mathbf{z}, \sigma) &:= \|h_p(\mathbf{x}, \mathbf{z}, \sigma) - h(\mathbf{x}, \mathbf{z}, \sigma)\|, \\ E_{tp}(\mathbf{x}, \mathbf{z}, \sigma) &:= \|h_{tp}(\mathbf{x}, \mathbf{z}, \sigma) - h(\mathbf{x}, \mathbf{z}, \sigma)\|. \end{aligned}$$

We begin by showing that the function approximation by transformed perturbation converges on analytic domains, like the standard perturbation approximation. This result implies that we can arbitrarily accurately approximate the true policy function by increasing the order m .

Proposition 4.5.1. *Suppose that the true policy function is analytic over a compact set $S \subseteq \mathcal{X} \times \mathcal{Z} \times \mathbb{R}_{\geq 0}$. Then m -order transformed perturbation errors vanish uniformly over S for any sequence $\tau \rightarrow 0$ as the perturbation order diverges to infinity. That is,*

$$\lim_{m \rightarrow \infty, \tau \rightarrow 0} \sup_{(\mathbf{x}, \mathbf{z}, \sigma) \in S} E_{tp}^{(m)}(\mathbf{x}, \mathbf{z}, \sigma) := \|h_{tp}^{(m)}(\mathbf{x}, \mathbf{z}, \sigma) - h(\mathbf{x}, \mathbf{z}, \sigma)\| = 0.$$

Next, we prove that transformed perturbation is always able to outperform regular perturbation.

Proposition 4.5.2. *For any policy function h there exists a $\tau \geq 0$ such that $E_{tp}(\mathbf{x}, \mathbf{z}, \sigma) \leq E_p(\mathbf{x}, \mathbf{z}, \sigma)$ for all possible values of \mathbf{x} , \mathbf{z} and σ .*

Note that this result makes no assumptions on the true policy function and implies that using the optimal τ^* for the transformation guarantees an equal or better approximation compared to regular perturbation. This result is true even when regular perturbation sample paths do not seem to explode. Therefore, it may be argued that transformed perturbation should always be used over regular perturbation.

We proceed by studying the accuracy properties of the transformed perturbation method for arbitrary values of the constant τ . First we show that locally the transformed polynomials inherit the excellent approximation qualities of perturbation methods. This follows because the exponential function $\Phi_\tau(\tilde{\mathbf{x}})$ is asymptotically quadratic as $\|\mathbf{x}\|$ goes to zero. A consequence of the proposition below is that, close to the steady state, errors between the transformed perturbation paths and the true paths are of the same magnitude as the errors between the regular perturbation paths and the true paths for $m = 2, 3$.

Proposition 4.5.3. *Suppose that $\mathbf{x}_0 = \mathbf{0}_{n_x}$ and $\mathbf{z}_0 = \mathbf{0}_{n_z}$. Let $(\mathbf{x}_t)_{t \geq 0}$ be the path generated by the true policy function (4.4) and let $(\hat{\mathbf{x}}_t)_{t \geq 0}$ be the path generated by the m 'th order transformed perturbation policy function, both initialised at these same starting values. Then it holds for all $t \in \mathbb{N}$ that*

$$\|\hat{\mathbf{x}}_t - \mathbf{x}_t\| = \begin{cases} O(\sigma^3) & \text{if } m = 2 \\ O(\sigma^4) & \text{if } m > 2 \end{cases} \quad \text{as } \sigma \rightarrow 0.$$

Transformed perturbation has the same local properties as regular perturbation, but on a global scale it is almost guaranteed to perform much better. Clearly if the true policy function produces explosive sample paths, then our method, which does not, cannot be assured to work well. The next result exhibits a very general set up in which the true policy function is ensured to produce nonexplosive sample paths making transformed perturbation infinitely more accurate in the tails than regular perturbation.

Proposition 4.5.4. *Suppose that Assumptions A1, A3 and C3 hold and that the true policy function h satisfies*

$$\limsup_{\|\mathbf{x}\| \rightarrow \infty} \frac{\mathbb{E}(\|h(\mathbf{x}, \mathbf{z}_1, \sigma)\| \mid \mathbf{z}_0 = \mathbf{z})}{\|\mathbf{x}\|} < 1 \quad (4.14)$$

for all possible values of \mathbf{z} and σ . Then the true policy function almost surely produces nonexplosive sample paths and if $h_p(\mathbf{x}, \mathbf{z}, \sigma)$ contains a nonzero higher order monomial in \mathbf{x} , then

$$\lim_{\|\mathbf{x}\| \rightarrow \infty} \frac{E_{tp}(\mathbf{x}, \mathbf{z}, \sigma)}{E_p(\mathbf{x}, \mathbf{z}, \sigma)} = 0 \quad (4.15)$$

for all possible values of \mathbf{z} and σ outside of a set of Lebesgue measure zero.

Moreover, condition (4.14) is implied by each of the following common conditions that are found in the literature on stable stochastic dynamic systems. The true policy function h

(i) *is eventually bounded by the 45 degree line for all possible values of \mathbf{z} and σ . That*

is,

$$\limsup_{\|\mathbf{x}\| \rightarrow \infty} \frac{\|h(\mathbf{x}, \mathbf{z}, \sigma)\|}{\|\mathbf{x}\|} < 1.$$

(ii) is uniformly contracting for all possible values of \mathbf{z} and σ . That is,

$$\sup_{\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{X}} \frac{\|h(\mathbf{x}_1, \mathbf{z}, \sigma) - h(\mathbf{x}_2, \mathbf{z}, \sigma)\|}{\|\mathbf{x}_1 - \mathbf{x}_2\|} < 1.$$

(iii) is slowly varying at infinity for all possible values of \mathbf{z} and σ . That is, for all $a > 0$ we have

$$\lim_{\|\mathbf{x}\| \rightarrow \infty} \frac{h(a\mathbf{x}, \mathbf{z}, \sigma)}{h(\mathbf{x}, \mathbf{z}, \sigma)} = 1.$$

4.5.2 Applications

In this section, we revisit two DSGE models used in Den Haan and De Wind (2012) to compare transformed perturbation to pruning and other solution methods. Below, we will show that the transformed perturbation approximation significantly outperforms both the regular perturbation approximation and the pruning method. For the purpose of comparing the performance of different solution methods, the true policy function will be approximated to an arbitrary level of accuracy on a relevant set using techniques such as projection methods or value function iteration, see Aruoba et al. (2006). We can then compare the solution methods by analysing sample paths between the “true” solutions and the approximated ones. The length of our time paths are $T = 10^4$, with a burn in period of 500 observations.

We compare sample paths according to three different criteria. The first one measures the distance between a period t variable generated by an approximation versus the one generated by the true policy function as in Den Haan and De Wind (2012). Let x_t be a generalisation of a univariate variable according to the true solution, let \hat{x}_t be generated according to some approximation and let M be the mean of the path $(x_t)_{t=1}^T$. Then we

define the error at time t as

$$\min \left\{ \left| \frac{\dot{x}_t - x_t}{x_t} \right|, \left| \frac{\dot{x}_t - x_t}{M} \right| \right\},$$

that is, we take the minimum of the absolute percentage error and the absolute error relative to the mean of the true solution path. The minimum between these two is chosen because the percentage error inflates the error when x_t is close to zero, while the error scaled by the mean overestimates inaccuracy when variables take on values far away from their mean.

The second criteria that we use are Euler errors. The equilibrium condition (4.1) is typically unequal to zero when we use an approximation method instead of the true solution. Its size is an indication for accuracy, because the size of the difference in supremum norm on a compact set between an approximate policy function and the true solution is of the same magnitude as the Euler error, see Theorem 3.3 of Santos (2000). We report the non normalized sample Euler error. We don't normalize our Euler errors, because we are only interested in relative accuracy.

Finally we compare sample moments generated by the approximated paths versus the true ones. DSGE models are often estimated using moment based approaches such as the (simulated) method of moments or indirect inference. Therefore the accuracy of the moments will have an impact on the estimated parameters. Let x_t and y_t be univariate variables, then we compare the sample moments

$$\mu^k(x_t) = \frac{1}{T} \sum_{t=1}^T x_t^k$$

and cross moments $\mu(x_t^i y_t^j)$.

The Deaton model

The first model we consider is a partial equilibrium model in which agents face idiosyncratic income risk. The original model was proposed in Deaton Angus (1991), however, we use the modified penalty function that was introduced in Den Haan and De Wind (2012) to compare pruned and non-pruned perturbation solution methods. The model

CHAPTER 4. TRANSFORMED PERTURBATION SOLUTIONS FOR DYNAMIC STOCHASTIC GENERAL EQUILIBRIUM MODELS

is therefore identical to Model 3 in Den Haan and De Wind (2012). The optimization problem is given by

$$\begin{aligned} \max_{(c_t, a_t)_{t=1}^{\infty}} \quad & \mathbb{E}_1 \sum_{t=1}^{\infty} \beta^{t-1} \left(\frac{c_t^{1-\gamma} - 1}{1-\gamma} - P(a_t) \right), \\ \text{s.t.} \quad & c_t + a_t(1+r) = a_{t-1} + e^{z_t}, \\ & z_t = \bar{z} + \varepsilon_t, \\ & \varepsilon_t \sim N(0, \sigma^2), \\ & a_0 \text{ given,} \end{aligned}$$

where c_t stands for the agents consumption, e^{z_t} represents exogenous and random income and r is the exogenous interest rate. The variable a_t denotes the amount of chosen assets in period t , we assume that a_0 is given. The amount of assets is allowed to be negative, so the agent can borrow. The function P is given by

$$P(a_t) = \frac{\eta_1}{\eta_0} e^{-\eta_0 a_t} + \eta_2 a_t.$$

Note that it is decreasing in its argument and thus penalizes utility when the agent decides to borrow. We write $x_t = a_{t-1} + e^{z_t}$ for the amount of cash on hand at time t . Note that this DSGE model has a univariate state equation in x_t , because the z_t are independent.

Our calibration is copied from the original paper and given in Table 4.5.1. The value

r	γ	\bar{z}	σ	β	η_0	η_1	η_2
0.03	3	0.4	0.1	0.9	20	0.04464	0.00352

Table 4.5.1: The choice of parameter values for the Deaton model.

of β is low to make agents impatient and ensure that borrowing constraints have sufficient effect on the decision process. The value of σ is chosen large, because the model describes single agent/household behavior and thus works with idiosyncratic uncertainty. The values of η_1 and η_2 are chosen such that a_t has the same moments as in Deaton Angus (1991). We refer to Den Haan and De Wind (2012) for a more detailed discussion on the model and choice of parameters.

We use a second order perturbation approximation to obtain

$$x_{t+1} - x_{ss} = 0.01 + 0.42(x_t - x_{ss}) + 1.02(x_t - x_{ss})^2 + e^{z_{t+1}}$$

and values for τ given by $\tau^* = 1.08$ and $\hat{\tau} = 0.98$. The innovations in the model are strictly positive, so we cannot use Assumption C to ensure stability of transformed perturbation sample paths. Instead we use Assumption B, which is easy to check in univariate cases. Note that all parameters are positive and the autoregressive parameter is smaller than one. It immediately follows that the transformed perturbation approximation is able to reach any sufficiently large point and thus we have an open interval of reachable points and Assumption B1 is satisfied. All the other Assumptions in A and B are easily checked. Therefore we obtain all the desired stability results from Theorem 4.3.2.

To compare the approximate policy functions we plot in Figure 4.5.1a the expected value of next-period's cash on hand $\mathbb{E}(x_{t+1} \mid x_t)$, because this directly reveals whether the dynamics are stable or not. The true policy function has a single stable fixed point (an attractor). In contrast, the second order perturbation policy function has a second fixed point (a repeller). This second intersection with the $y = x$ line is located above the true steady state. Sample paths produced by the second order perturbation function eventually reach the state space to the right of the repeller, after which they are expected to diverge, and eventually do with probability one. Since the second fixed point is relatively close to the true steady state this also frequently occurs in our finite time simulated paths, making second order perturbation infeasible. The transformed perturbation policy function solves the problem as it negates the second order monomial fast enough to ensure that no second fixed-point is created. The optimal and plug in values for τ , while irrelevant for stability, therefore create a policy function that generates very similar dynamics as the true policy function. Figure 4.5.1c displays the same functions as in Figure 4.5.1a, but focussed on the relevant part of the state space when using stable methods. In addition we have added a scatter plot of the pruning sample path. From this plot it becomes immediately apparent that pruning does not deliver a policy function on the original state space, as we have different updates for the same starting value. Moreover, it can be seen that pruning on average is less accurate than both the transformed perturbation methods. The policy

CHAPTER 4. TRANSFORMED PERTURBATION SOLUTIONS FOR DYNAMIC STOCHASTIC GENERAL EQUILIBRIUM MODELS

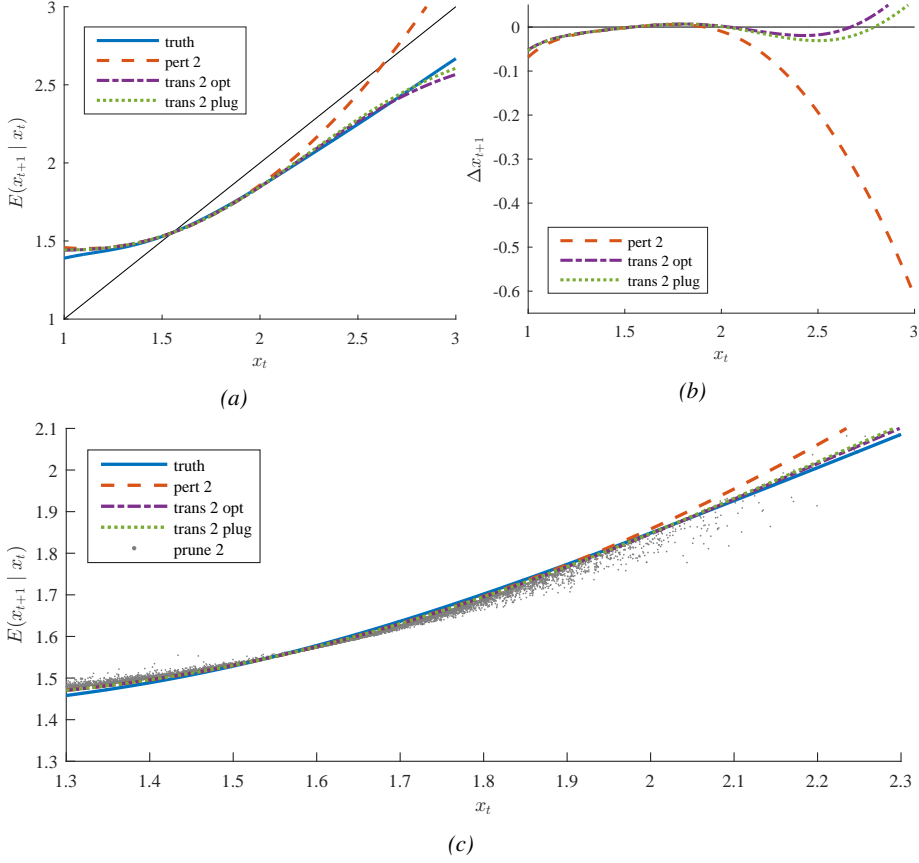


Figure 4.5.1: Expected policy functions for x_t in the Deaton model generated by a second order perturbation approximation and the transformed perturbation method for both the optimal τ^* and the plug in $\hat{\tau}$. Figure 4.5.1a shows the actual policy functions, Figure 4.5.1b shows the pointwise errors with respect to a close approximation of the true policy function and Figure 4.5.1c zooms in on the relevant part of the state space to compare the previous methods to pruning.

function corresponding to the optimal τ^* can be seen to be slightly more accurate than the plug in $\hat{\tau}$. This is extra apparent when we look at the pointwise errors between the true path and the perturbation respective transformed perturbation approximations in Figure 4.5.1b.

It's not surprising that the resulting transformed perturbation sample paths are very close to the true ones. The sample path accuracy results are summarised in Table 4.5.2, where we report maximum and mean absolute path errors in addition to Euler errors. Here we see that second order perturbation explodes, so sample paths created by this

	Path errors				Euler errors
	a_t		c_t		
	max	mean	max	mean	
Perturbation 1	132	38.4	10.7	1.06	3.11
Perturbation 2	∞	∞	∞	∞	-
Transformed 2 optimal	53.0	6.54	3.29	0.31	0.28
Transformed 2 plug-in	54.4	6.50	3.40	0.31	0.29
Pruning 2	123	13.6	6.42	0.69	0.51

Table 4.5.2: Absolute sample path and Euler errors for the Deaton model. Path errors are compared to a projection approximation and given in percentages. Euler errors are also scaled by 10^2 . The results are based on a time path of 10^4 observation with a burn in time of 500 observations.

approximation are unusable. Therefore we need a stable approximation approach. The transformed perturbation approximation performs better than pruning and much better than linear approximation on all criteria. Note that the maximum and mean path errors for the transformed perturbation are about half of those for the pruning approximation, in both the asset and consumption paths.

The difference in accuracy is extra apparent when we look at the cumulative path errors, see Figure 4.5.2, which are significantly smaller for our method. This accumulation of inaccuracy then leads to larger errors when we compute some of the sample moments, which can be found in Table 4.5.3. Here we see that first order perturbation performs a

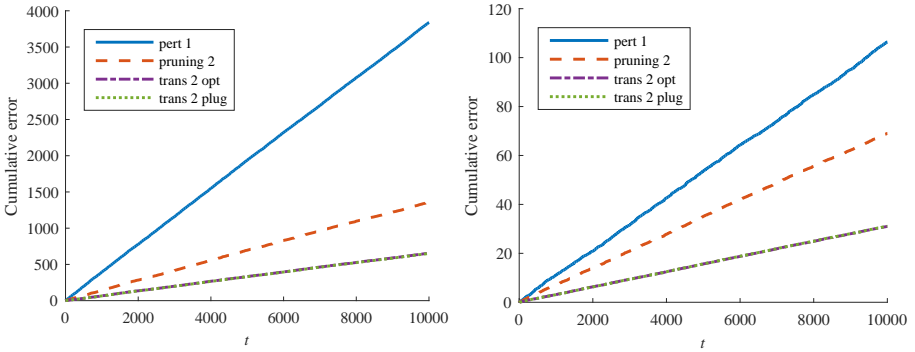


Figure 4.5.2: Cumulative paths errors for the number of assets in the left panel and consumption in the right panel. Errors are calculated by a close approximation of the true policy function.

lot worse than the other methods on the asset moments, which was to be expected, as it missed the nonlinear effects. Transformed perturbation is more accurate than pruning for

CHAPTER 4. TRANSFORMED PERTURBATION SOLUTIONS FOR DYNAMIC STOCHASTIC GENERAL EQUILIBRIUM MODELS

	Sample moments							
	$\mu(a_t)$	$\mu^2(a_t)$	$\mu^3(a_t)$	$\mu^4(a_t)$	$\mu(c_t)$	$\mu^2(c_t)$	$\mu^3(c_t)$	$\mu^4(c_t)$
True	0.083	0.016	0.004	0.001	1.502	2.264	3.423	5.192
Perturbation 1	58.1	62.2	84.2	89.7	0.09	0.11	0.03	0.16
Transformed 2 optimal	0.70	7.58	7.83	6.79	0.00	0.03	0.09	0.17
Transformed 2 plug-in	0.43	6.89	6.25	3.77	0.00	0.03	0.09	0.17
Pruning 2	5.88	23.50	31.66	38.27	0.01	0.06	0.20	0.40
	Cross moments							
	$\mu(a_t c_t)$	$\mu(a_t c_t^2)$	$\mu(a_t c_t^3)$	$\mu(a_t^2 c_t)$	$\mu(a_t^2 c_t^2)$	$\mu(a_t^3 c_t)$		
True	0.13	0.20	0.31	0.03	0.04	0.006		
Perturbation 1	57.3	56.5	55.7	63.0	63.6	84.1		
Transformed 2 optimal	0.96	1.18	1.35	7.45	7.30	7.65		
Transformed 2 plug-in	0.69	0.90	1.07	6.73	6.57	6.03		
Pruning 2	6.30	6.58	6.75	23.2	22.8	31.1		

Table 4.5.3: Sample and cross moments up to fourth order for the Deaton model. The true row presents the moments given by a close approximation. The other moments are given as absolute percentage differences from the true ones. The results are based on a time path of 10^4 observation with a burn in time of 500 observations.

all moments, especially for ones concerning the assets where we see improvement up to a factor ten. Surprising is that the plug-in $\hat{\tau}$ transformed policy function performs better on the moments than the optimal τ^* transformed policy function.

Performance in a parameter estimation scenario

When researchers are interested in estimating parameters, it is important to ensure that the employed approximation method is accurate across a wide range of parameter values. It is thus important to investigate what happens to the accuracy of our approximation methods when we move the parameters away from an initial calibrated parameter value. Figure 4.5.3 plots the expected Euler errors for varying values of β and γ . Note that, as described in Section 4.2.4, for the optimal transformed perturbation method we have kept the initial calculated optimal τ^* , while the plug in transformed perturbation method updates $\hat{\tau}$ along with the parameters. We see in Figure 4.5.3 that the expected Euler errors for both the transformed perturbation methods are smaller than those for the pruning method on a significant area around the calibration. This implies that each transformed perturbation method outperforms the pruning method in an estimation setting when the initial parameters have been set sufficiently close to the true ones. The two transformed perturbation methods have such similar Euler errors, because the plug-in $\hat{\tau}$ does not vary much as we change the parameters and stays especially close to the optimal τ^* .

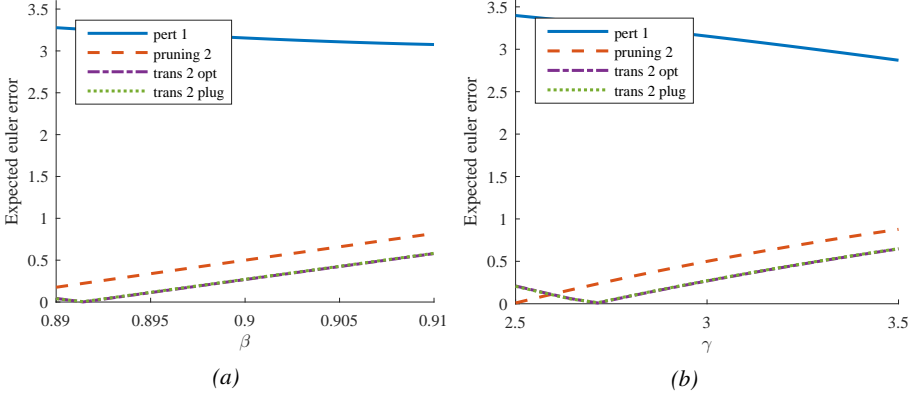


Figure 4.5.3: Expected Euler errors for the Deaton model on an area around the calibrated parameter values. Figure 4.5.3a portrays the results when changing β and Figure 4.5.3b when changing γ .

The Matching model

The second model we examine is a matching model also featured in Den Haan and De Wind (2012). The model has two types of agents, workers and entrepreneurs, both of which are members of the same representative household. The household earns wages and firm profits from its members at the end of each period. These are then distributed among the members for consumption.

Firms: The main decision is made by a representative entrepreneur who tries to maximise future discounted firm profits. The maximisation problem is given by

$$\begin{aligned}
 \max_{(n_t, v_t)_{t=1}^{\infty}} \quad & \mathbb{E}_1 \sum_{t=1}^{\infty} \beta^{t-1} \left(\frac{c_t}{c_1} \right)^{-\gamma} ((e^{z_t} - w) n_{t-1} - \psi v_t), \\
 \text{s.t.} \quad & n_t = (1 - \rho_n) n_{t-1} + p_{f,t} v_t, \\
 & z_t = \begin{cases} z_{t-1} & \text{with probability } \rho_z \\ -z_{t-1} & \text{with probability } 1 - \rho_z \end{cases}, \\
 & n_0, z_1 \text{ given.}
 \end{aligned}$$

Here c_t is the consumption level of the household, n_t is the number of employees at the end of period t , v_t is the number of vacancies set by the firm, $p_{f,t}$ is the number of matches per vacancy, w is the wage rate, ψ is the cost of placing a vacancy and ρ_n is the

CHAPTER 4. TRANSFORMED PERTURBATION SOLUTIONS FOR DYNAMIC STOCHASTIC GENERAL EQUILIBRIUM MODELS

exogenous separation rate. Each worker produces e^{z_t} , which means that the profit per worker is given by $e^{z_t} - w$. The random variable z_t can only take on two values, which we denote $-\zeta$ and $+\zeta$. This is an artificial simplification introduced in Den Haan and De Wind (2012) enabling us to easily analyse the approximation methods to the model in a graphical manner. Alternatively, one can use a standard autoregressive updating function for z_t . Finally, the firm takes the number of matches $p_{f,t}$ as given.

Consumers: The household consumes the whole income earned by its members. That is,

$$c_t = wn_{t-1} + (e^{z_t} - w)n_{t-1} - \psi v_t = e^{z_t}n_{t-1} - \psi v_t.$$

Matching market: The number of hires per vacancy is determined on a matching market where the firms and $1 - n_{t-1}$ unemployed workers search for a match. The total number of matches is given by

$$m_t = \phi_0(1 - n_{t-1})^\phi v_t^{1-\phi}.$$

This implies that the total number of matches per vacancy is given by

$$p_{f,t} = \frac{m_t}{v_t} = \phi_0 \left(\frac{1 - n_{t-1}}{v_t} \right)^\phi.$$

The model requires some restrictions on the parameters to ensure that a solution in the interior of the domain exists and thus that the policy function is smooth. Our choice of parameter values is again taken from Den Haan and De Wind (2012) and given in Table 4.5.4. See the original paper for a detailed discussion on the matching model, the parameter values and further references.

γ	w	ψ	ρ_n	ρ_z	ζ	σ	β	ϕ_0	ϕ
4.5	0.973	0.4026	0.0368	0.975	0.0224	0.007	0.99	0.7	0.5

Table 4.5.4: The choice of parameter values for the Matching model.

A second order perturbation approximation of the state equation delivers

$$\begin{aligned} n_{t+1} - n_{ss} = & 0.95 + 0.46(n_t - n_{ss}) + 0.52z_{t+1} \\ & - 2.92(n_t - n_{ss})^2 - 6.57(n_t - n_{ss})z_{t+1} - 1.01z_{t+1}^2 \end{aligned}$$

and we find

$$\tau^* = 26.1 \quad \text{and} \quad \hat{\tau} = 13.6.$$

The updating equation for the exogenous state variable z_t is not of the type (4.2). One can extend the theory in a rather straightforward way to also apply to general Markov chain updating equations for the exogenous state variables, but we chose not to do this to keep the assumptions and proofs relatively clear and concise. Note that if we would have chosen a standard autoregressive process of order one for $(z_t)_{t \geq 0}$, then Assumptions A and C can easily be seen to be satisfied as we have a univariate system. Therefore, in that case, we would have obtained all the desired stability results from Theorem 4.3.2.

The control variables can be explicitly calculated once the path for the single state variable, the number of employees, is known. We therefore compare the approximation methods according to their best performance: either calculating the control variables directly, or approximating the observation equation. We compare the transformed perturbation and regular perturbation approximation in Figure 4.5.4. Figure 4.5.4a shows the policy functions for the number of employees in the two possible scenarios for z_t . The case $z_t = -\zeta$ is the crucial one here, as the regular perturbation approximation stays below the $y = x$ line and therefore does not intersect it. This implies that the second order perturbation sample paths for n_t tend to minus infinity if z_t is equal to $-\zeta$ for many consecutive times. The case $z_t = +\zeta$ goes to minus infinity for values of n_t much smaller than portrayed in the figure. Hence, once n_t has become small enough it has no chance of recovering and thus sample paths diverge to minus infinity with probability one. As in the previous example, the explosive behaviour is encountered in our finite time sample paths with a high frequency, rendering regular perturbation infeasible. The transformed perturbation policy function avoids the problem described for both values of τ as they both scale down the second order monomial fast enough to ensure that the policy func-

CHAPTER 4. TRANSFORMED PERTURBATION SOLUTIONS FOR DYNAMIC STOCHASTIC GENERAL EQUILIBRIUM MODELS

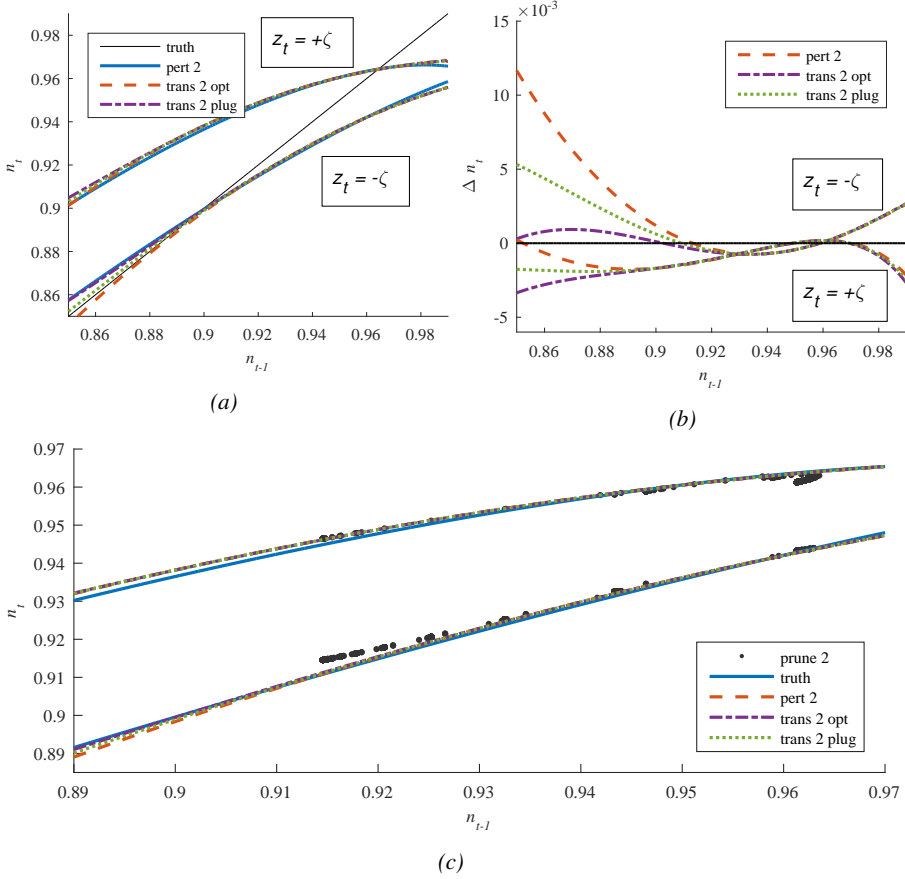


Figure 4.5.4: Policy functions for n_t in the matching model generated by a second order perturbation approximation and the transformed perturbation method. Figure 4.5.4a shows the actual policy functions, Figure 4.5.4b shows the pointwise errors with respect to a close approximation of the true policy function and Figure 4.5.1c zooms in on the relevant part of the state space to compare the previous methods to pruning.

tions cross the $y = x$ line at a unique point, like the true policy function. The dynamics of our approximated systems therefore closely mimic the true dynamics for n_t . Figure 4.5.4c again zooms in on the relevant part of the state space when using the stable solution methods and includes a scatter plot of the pruning sample path. Again, we are reminded that pruning does not provide a policy function on the original state space. Moreover, pruning provides less accurate updates, especially for large value of n_t in the case $z_t = +\zeta$ and small values in the case $z_t = -\zeta$. The policy function corresponding to the optimal τ^* is

clearly the most accurate method in our comparison, which is extra clear when we look at the pointwise errors between the true path and the perturbation respective transformed perturbation approximations in Figure 4.5.4b.

The graphical results are strengthened by studying the sample path errors in Table 4.5.5. Here we see that the transformed perturbation approximation is both in extreme

	Path errors				Euler errors
	n_t		c_t		
	max	mean	max	mean	
Perturbation 1	3.20	1.89	3.53	1.80	0.26
Perturbation 2	∞	∞	∞	∞	-
Transformed 2 optimal	0.26	0.07	0.64	0.32	0.08
Transformed 2 plug-in	0.71	0.25	0.97	0.23	0.04
Pruning 2	1.79	0.95	1.76	0.95	0.10

Table 4.5.5: Absolute sample path and Euler errors for the matching model. Path errors are compared to a close approximation of the truth and given in percentages. Euler errors are also scaled by 10^2 . The results are based on a time path of 10^4 observation with a burn in time of 500 observations.

cases and on average performing better than both perturbation and pruning. The improvement compared to perturbation is not surprising given the nonlinearity of the plots in Figure 4.5.4. This time the optimal transformed perturbation method performs better than the plug in approximation. It is also more than a factor ten times better on average than pruning for the number of employees and more than a factor three times better on average than pruning on consumption paths.

We emphasize the gravity of the difference in accuracy by plotting the cumulative path errors in Figure 4.5.5. This total difference in accuracy then again leads to a large difference in sample moment accuracy, which is summarised in Table 4.5.6. Like before we see that the transformed perturbation method, especially the optimal one, is best at mimicking the dynamics of the sample paths. Note that both the optimal and transformed perturbation method outperform pruning on all moments, especially for the higher order moments, where pruning loses relatively more accuracy by ignoring higher order effects. Optimal transformed perturbation outperforms pruning up to a factor forty for the fourth order moments of consumption, while plug-in transformed perturbation outperforms pruning by a factor three for most moments.

CHAPTER 4. TRANSFORMED PERTURBATION SOLUTIONS FOR DYNAMIC STOCHASTIC GENERAL EQUILIBRIUM MODELS

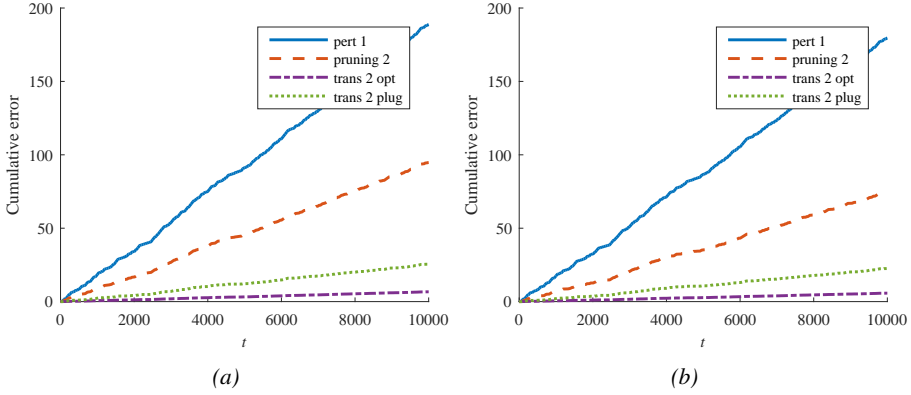


Figure 4.5.5: Cumulative paths errors for the number of employees in the left panel and consumption in the right panel. Errors are calculated by a close approximation of the true policy function.

	Sample moments							
	$\mu(n_t)$	$\mu^2(n_t)$	$\mu^3(n_t)$	$\mu^4(n_t)$	$\mu(c_t)$	$\mu^2(c_t)$	$\mu^3(c_t)$	$\mu^4(c_t)$
True	0.93	0.87	0.81	0.76	0.91	0.83	0.76	0.70
Perturbation 1	1.90	3.77	5.62	7.43	1.81	3.57	5.27	6.91
Transformed 2 optimal	0.03	0.05	0.07	0.10	0.02	0.04	0.05	0.06
Transformed 2 plug-in	0.23	0.45	0.65	0.83	0.20	0.39	0.55	0.70
Pruning 2	0.68	1.30	1.88	2.40	0.65	1.23	1.73	2.16
	Cross moments							
	$\mu(n_t c_t)$	$\mu(n_t c_t^2)$	$\mu(n_t c_t^3)$	$\mu(n_t^2 c_t)$	$\mu(n_t^2 c_t^2)$	$\mu(n_t^3 c_t)$		
True	0.85	0.78	0.71	0.79	0.73	0.74		
Perturbation 1	3.67	5.39	7.04	5.50	7.17	7.30		
Transformed 2 optimal	0.04	0.06	0.07	0.07	0.08	0.09		
Transformed 2 plug-in	0.42	0.59	0.73	0.62	0.77	0.80		
Pruning 2	1.27	1.78	2.22	1.83	2.28	2.34		

Table 4.5.6: Sample and cross moments up to fourth order for the matching model. The true row presents the moments given by a close approximation. The other moments are given as absolute percentage differences from the true ones. The results are based on a time path of 10^4 observation with a burn in time of 500 observations.

Performance in a parameter estimation scenario

Once more we investigate the accuracy of the discussed methods in an area around the calibrated parameter values. Figure 4.5.6 plots the expected Euler errors for varying values of β and γ while keeping the steady state values for the number of employees, the number of matches per unemployed worker and the number of matches per vacancy equal. As in the previous example we fix the optimal τ^* at the initial derived value at the calibrated parameters, while the plug in $\hat{\tau}$ is updated along with the parameters. Figure 4.5.6 shows

us that the expected Euler errors for each transformed perturbation method is smaller than those for the pruning method on a relevant area around the calibration. Therefore, we again conclude that an estimation procedure using the transformed perturbation method improves accuracy over using either linear perturbation or pruning when the starting values are decently close to the true parameters.

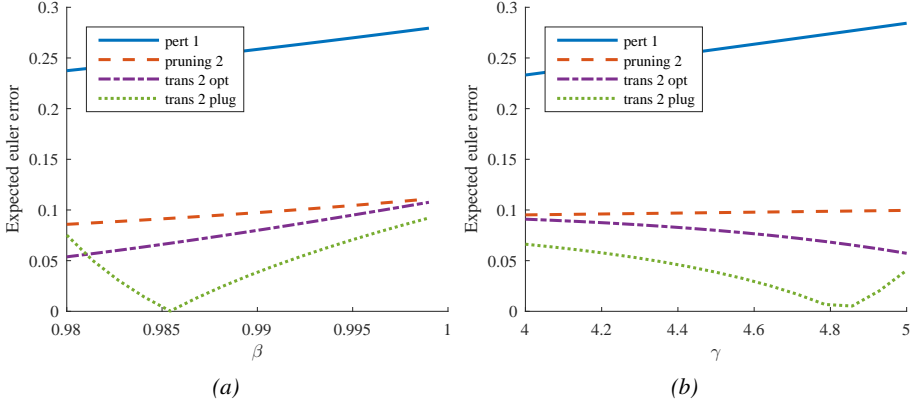


Figure 4.5.6: Expected Euler errors for the matching model on an area around the calibrated parameter values. Figure 4.5.6a portrays the results when changing β and Figure 4.5.6b when changing γ while keeping the steady state values for the number of employees, the number of matches per unemployed worker and the number of matches per vacancy equal.

4.6 Conclusion

This paper introduces a new solution method for DSGE models that produces non explosive paths. The proposed solution method is as fast as standard perturbation methods and can be easily implemented in existing software packages like *Dynare* as it is obtained directly as a transformation of existing perturbation solutions proposed by Judd and Guu (1997) and Schmitt-Grohe and Uribe (2004), among others. The transformed perturbation method shares the same advantageous function approximation properties as standard higher order perturbation methods and, in contrast to those methods, generates stable sample paths that are stationary, geometrically ergodic and absolutely regular. Additionally, moments are shown to be bounded. The method is an alternative to the pruning method as proposed in Kim et al. (2008). The advantages of our approach are that, unlike pruning,

it does not need to sacrifice accuracy around the steady state by ignoring higher order effects and it delivers a policy function. Moreover, the newly proposed solution is always more accurate globally than standard perturbation methods and has proven to have superior accuracy compared to regular perturbation and pruning for two example nonlinear DSGE models.

4.7 Appendix: Proofs

4.7.1 Proofs of Section 4.3

We study the transformed perturbation method, as indicated in Section 4.3, as its asymptotic linear process plus a deviation (4.9). The deviation is bounded in \mathbf{x} as $\Phi_\tau(\tilde{\mathbf{x}})$ dominates the function far away from the origin. The following result gives a uniform upper bound to the size of the deviation over \mathcal{X} .

Proposition 4.7.1. *There exists a constant $c \geq 0$ that does not depend on τ , such that*

$$\sup_{\mathbf{x} \in \mathcal{X}} \|D(\mathbf{x}, \mathbf{z})\| \leq c \sum_{j=0}^m \tau^{-j/2} \left(\sum_{i=0}^{m-j} \|\mathbf{z}\|^i \right).$$

PROOF. In this proof we specifically choose $\|\cdot\|$ equal to the Euclidean matrix norm $\|\cdot\|_e$. This matrix norm is a crossnorm, i.e. it is multiplicative on Kronecker products, see for example Lancaster and Farahat (1972). This implies, together with sub-additivity and sub-multiplicativity, that

$$\begin{aligned} \|D(\mathbf{x}, \mathbf{z})\| &\leq \sum_{i=2}^m \|H_i\| \|\mathbf{v}\|^i \Phi_\tau(\tilde{\mathbf{x}}) \\ &\leq \left(\max_{2 \leq i \leq m} \|H_i\| \right) \sum_{i=2}^m \sum_{j=0}^i \|\mathbf{x}\|^j \|\mathbf{z}\|^{i-j} \Phi_\tau(\tilde{\mathbf{x}}) \\ &\leq \left(\max_{2 \leq i \leq m} \|H_i\| \right) \sum_{j=0}^m \|\mathbf{x}\|^j \Phi_\tau(\tilde{\mathbf{x}}) \left(\sum_{i=0}^{m-j} \|\mathbf{z}\|^i \right). \end{aligned}$$

Next, note that

$$\|\mathbf{x}\|^j \Phi_\tau(\tilde{\mathbf{x}}) \leq \|\mathbf{x}\|^j e^{-\tau \|\mathbf{x}\|_e^2 / \max\{\mathbf{x}_{ss}\}},$$

which is a univariate function in $\|\mathbf{x}\|_e$, since we chose $\|\cdot\|$ equal to $\|\cdot\|_e$. It is straightforward to verify that this function is maximised at $\|\mathbf{x}\|_e^2 = \frac{j \max\{\mathbf{x}_{ss}\}}{2\tau}$ and thus there exists a constant \tilde{c} that does not depend on τ or \mathbf{x} such that

$$\sup_{\mathbf{x} \in \mathcal{X}} \|\mathbf{x}\|^j \Phi_\tau(\tilde{\mathbf{x}}) \leq \tilde{c} \tau^{-j/2} \quad \text{for all } 0 \leq j \leq m.$$

■

Proof of Theorem 4.3.1

Assumptions A1 and A3 imply by Theorem 3.1 in Bougerol (1993) and the monotone convergence theorem that there exists a unique stationary ergodic solution $(\mathbf{z}_t^*)_{t \in \mathbb{N}}$ to (4.2) with $\mathbb{E} \|\mathbf{z}_t^*\|^{rm} < \infty$. Moreover, $\|\mathbf{z}_t - \mathbf{z}_t^*\|$ converges exponentially almost surely to zero as $t \rightarrow \infty$, which implies that

$$\liminf_{t \rightarrow \infty} \|\mathbf{z}_t\| < \infty \quad \text{a.s.}$$

and that, for every realisation, there exists a constant $d > 0$ such that $\|\mathbf{z}_t\|^i \leq \|\mathbf{z}_t^*\|^i + d$ for all $t \geq 0$ and $0 \leq i \leq m$.

Next, we repeatedly expand the term $H_{\mathbf{x}} \mathbf{x}$ in (4.7) to obtain the following expression for the transformed perturbation path:

$$\mathbf{x}_t = H_{\mathbf{x}}^t \mathbf{x}_0 + \sum_{k=0}^{t-1} H_{\mathbf{x}}^k (H_0 + H_{\mathbf{z}} \mathbf{z}_{t-k} + D(\mathbf{x}_{t-1-k}, \mathbf{z}_{t-k})).$$

We now use Proposition 4.7.1 to bound the deviation terms and then use the bounds on

the path $(\mathbf{z}_t)_{t \geq 0}$ to obtain

$$\begin{aligned}
 & \|\mathbf{x}_t\| - \|H_{\mathbf{x}}\|^t \|\mathbf{x}_0\| \\
 & \leq \sum_{k=0}^{t-1} \|H_{\mathbf{x}}\|^k \left(\|H_0\| + \|H_{\mathbf{z}}\| \|\mathbf{z}_{t-k}\| + c \sum_{j=0}^m \tau^{-j/2} \left(\sum_{i=0}^{m-j} \|\mathbf{z}_{t-k}\|^i \right) \right) \\
 & \leq \sum_{k=0}^{t-1} \|H_{\mathbf{x}}\|^k \left(\|H_0\| + \|H_{\mathbf{z}}\| (\|\mathbf{z}_{t-k}^*\| + d) + c \sum_{j=0}^m \tau^{-j/2} \left(\sum_{i=0}^{m-j} \|\mathbf{z}_{t-k}^*\|^i + d \right) \right)
 \end{aligned} \tag{4.16}$$

Next we artificially extend $(\mathbf{z}_t^*)_{t \geq 0}$ to a stationary ergodic sequence $(\mathbf{z}_t^*)_{t \in \mathbb{Z}}$ and then note that (4.16) is bounded by

$$Y_t := \sum_{k=0}^{\infty} \|H_{\mathbf{x}}\|^k \left(\|H_0\| + \|H_{\mathbf{z}}\| (\|\mathbf{z}_{t-k}^*\| + d) + c \sum_{j=0}^m \tau^{-j/2} \left(\sum_{i=0}^{m-j} \|\mathbf{z}_{t-k}^*\|^i + d \right) \right).$$

The term within the brackets is stationary ergodic by Krengel's lemma, see Proposition 4.3 in Krengel (1985), and the fact that $(\mathbf{z}_t^*)_{t \in \mathbb{Z}}$ is stationary ergodic. Moreover it has a finite log moment since $\mathbb{E} \|\mathbf{z}_t^*\|^{rm} < \infty$. Next, we can choose a matrix norm such that $\|H_{\mathbf{x}}\| < 1$ by Assumption A2. Therefore, the infinite sum converges almost surely by Proposition 2.5.1 of Straumann (2005). Again, the sequence $(Y_t)_{t \in \mathbb{Z}}$ is stationary ergodic by Krengel's lemma and thus there almost surely exists an $M > 0$ such that $\{Y_t \leq M\}$ occurs for infinitely many $t > 0$. We conclude that

$$\liminf_{t \rightarrow \infty} \|\mathbf{x}_t\| \leq M < \infty.$$

Proof of Theorem 4.3.2

We study the processes $(\mathbf{z}_t)_{t \geq 0}$ and $(\mathbf{x}_t)_{t \geq 0}$ as a joint Markov process. This section will make extensive use of Meyn and Tweedie (1993). We will first assume that $(\mathbf{z}_t, \mathbf{x}_t)_{t \geq 0}$ is a ψ -irreducible and aperiodic T -chain. See sections 4.2, 5.4 and 6.2 of Meyn and Tweedie (1993) for a detailed discussion on these properties.

Proposition 4.7.2. *Suppose $(\mathbf{z}_t, \mathbf{x}_t)_{t \geq 0}$ is a ψ -irreducible and aperiodic T -chain and let Assumptions A and B hold. Then all the results of Theorem 4.3.2 hold.*

PROOF. We will check the drift condition for t -step transitions, which is described in condition (iii) of Theorem 1 in Saïdi and Zakoian (2006), adapted from Theorem 19.1.3 in Meyn and Tweedie (1993) and originally suggested by Tjøstheim (1990). The condition states that we need to find a non-negative function $V : \mathcal{X} \times \mathcal{Z} \rightarrow \mathbb{R}$ and a $t \in \mathbb{N}$ such that

$$\frac{\mathbb{E}(V(\mathbf{x}_t, \mathbf{z}_t) \mid \mathbf{x}_0 = \mathbf{x}, \mathbf{z}_0 = \mathbf{z})}{V(\mathbf{x}, \mathbf{z})} \quad (4.17)$$

is finite on a compact set $C \subseteq \mathcal{X} \times \mathcal{Z}$ and smaller than one outside of C . Note that the set C actually has to be petite, but all compact sets are petite in a ψ -irreducible T -chain, Theorem 6.2.5 in Meyn and Tweedie (1993). It then follows by Theorem 1 in Saïdi and Zakoian (2006) that there exists a unique stationary ergodic solution $(\mathbf{x}_t^*, \mathbf{z}_t^*)_{t \geq 0}$ that is geometrically ergodic and has the required moments, given our choice for V . Absolute regularity follows from Theorem 1 in Davydov (1974) and the laws of large numbers follow from Theorem 17.0.1 in Meyn and Tweedie (1993). The reason that we resort to t -step, instead of 1-step, transitions is that Assumption A1 and Assumption A2 do not guarantee that there exists a matrix norm such that both $\|\mathbf{A}\| < 1$ and $\|H_{\mathbf{x}}\| < 1$. Assumption A1 can ensure that there exists a matrix norm such that $\|\mathbf{A}\| < 1$, but then Assumption A2 only provides the existence of a $t \in \mathbb{N}$ such that $\|H_{\mathbf{x}}^t\| < 1$ by Gelfand's formula.

We adopt the ideas of Cline and Pu (1999) and use the test function

$$V(\mathbf{x}, \mathbf{z}) = 1 + (\|\mathbf{x}\| + \omega\|\mathbf{z}\|^m)^r,$$

where we will choose $\omega > 0$ sufficiently large. If $r \leq 1$, then $(\|\mathbf{x}\| + \omega\|\mathbf{z}\|^m)^r \leq \|\mathbf{x}\|^r + \omega^r\|\mathbf{z}\|^{rm}$. We prove the theorem for the case $r \geq 1$, as it is the harder case. In that case Minkowski's inequality provides the upper bound

$$\mathbb{E}((\|\mathbf{x}_t\| + \omega\|\mathbf{z}_t\|^m)^r \mid \mathbf{x}_0, \mathbf{z}_0) \leq \left(\mathbb{E}(\|\mathbf{x}_t\|^r \mid \mathbf{x}_0, \mathbf{z}_0)^{\frac{1}{r}} + \omega \mathbb{E}(\|\mathbf{z}_t\|^{rm} \mid \mathbf{x}_0, \mathbf{z}_0)^{\frac{1}{r}} \right)^r.$$

We start by bounding the second expectation. Note that the expectations $\mathbb{E}\|\sigma\boldsymbol{\eta}\boldsymbol{\varepsilon}_s\|^{rm}$ are bounded for all $s \in \{1, \dots, t\}$ by Assumption A3. Expanding backwards and working

out brackets then gives

$$\begin{aligned} \mathbb{E}(\|\mathbf{z}_t\|^{rm} \mid \mathbf{z}_0) &\leq \mathbb{E}((\|\Lambda^t \mathbf{z}_0\| + \|\Lambda^{t-1} \varepsilon_1\| + \dots + \|\varepsilon_t\|)^{rm} \mid \mathbf{z}_0) \\ &\leq \|\Lambda\|^{trm} \|\mathbf{z}_0\|^{rm} + o(\|\mathbf{z}_0\|^{rm}) \quad \text{as } \|\mathbf{z}_0\| \rightarrow \infty. \end{aligned} \quad (4.18)$$

Next, by Proposition 4.7.1 there exist constants $c_1, c_2 > 0$ such that $\|D(\mathbf{x}_{s-1}, \mathbf{z}_s)\| < c_1 + c_2(1 + \|\mathbf{z}_s\|^m)$ for all $s \in \{1, \dots, t\}$. It then follows again by backwards expansion and the fact that $\|\mathbf{z}_s\| \leq 1 + \|\mathbf{z}_s\|^m$ that there exist constants $d_1, d_2 > 0$ such that

$$\begin{aligned} \mathbb{E}(\|\mathbf{x}_t\|^r \mid \mathbf{x}_0, \mathbf{z}_0) &\leq \mathbb{E}\left(\left\|H_{\mathbf{x}}^t \mathbf{x}_0 + \sum_{k=0}^{t-1} H_{\mathbf{x}}^k (H_0 + H_{\mathbf{z}} \mathbf{z}_{t-k} + D(\mathbf{x}_{t-1-k}, \mathbf{z}_{t-k}))\right\|^r \mid \mathbf{x}_0, \mathbf{z}_0\right) \\ &\leq \mathbb{E}\left(\left(\|H_{\mathbf{x}}^t\| \|\mathbf{x}_0\| + d_1 + d_2 \sum_{k=0}^{t-1} \|\mathbf{z}_{t-k}\|^m\right)^r \mid \mathbf{x}_0, \mathbf{z}_0\right) \\ &\leq (\|H_{\mathbf{x}}^t\| \|\mathbf{x}_0\| + O(\|\mathbf{z}_0\|^m))^r \quad \text{as } \|\mathbf{z}_0\| \rightarrow \infty. \end{aligned}$$

The last inequality follows by repeated application of Minkowski's inequality in combination with the same calculations as in (4.18). Filling everything in then upper bounds (4.17) by

$$\frac{1 + (\|H_{\mathbf{x}}^t\| \|\mathbf{x}\| + (\|\Lambda\|^{tm} + \omega^{-1} O(1)) \omega \|\mathbf{z}\|^m + o(\|\mathbf{z}\|^m))^r}{1 + (\|\mathbf{x}\| + \omega \|\mathbf{z}\|^m)^r} \quad \text{as } \|\mathbf{z}\| \rightarrow \infty.$$

Recall that $\|H_{\mathbf{x}}^t\| < 1$ and $\|\Lambda\| < 1$ and choose ω large enough such that $\|\Lambda\|^{tm} + \omega^{-1} O(1) < 1$ as $\|\mathbf{z}\| \rightarrow \infty$. Then we can make the fraction smaller than one if we choose $\|\mathbf{x}\|, \|\mathbf{z}\| > M$ for a sufficiently large M . Let $C = \{(\mathbf{x}, \mathbf{z}) \in \mathcal{X} \times \mathcal{Z} \mid \|\mathbf{x}\|, \|\mathbf{z}\| \leq M\}$, then (4.17) is bounded over C and smaller than one outside of C . \blacksquare

It remains to be proven that $(\mathbf{z}_t, \mathbf{x}_t)_{t \geq 0}$ is a ψ -irreducible and aperiodic T -chain, which follows from the results of sections 6.0 - 1 of Meyn and Tweedie (1993). We have, similarly to Proposition 6.1.2 and 6.1.3, that Assumption B2 ensures that the Markov chain is strong Feller. It then follows by Proposition 6.1.5 and Assumption B1 that the Markov chain is ψ -irreducible. Finally, we conclude that $(\mathbf{x}_t)_{t \geq 0}$ is an aperiodic T -chain by Lemma 6.1.4 and part (iii) of Theorem 6.0.1.

Proof of Proposition 4.3.3

It is clear that Assumption C3 implies Assumption B2, so it remains to prove Assumption C also implies Assumption B1. We will prove a stronger statement: Fix any $\mathbf{x}^* \in \mathcal{X}$ then that point is reachable. Let t be the smallest integer such that assumption C1 holds. The approach will be to show that we can find values for $\mathbf{z}_1, \dots, \mathbf{z}_t$ that bring \mathbf{x}_t arbitrarily close to \mathbf{x}^* . It then follows by Assumption C3 that we have positive probability of \mathbf{x}_t being arbitrarily close to \mathbf{x}^* .

To find the values for the exogenous state variables, we start by expanding \mathbf{x}_t back in time as

$$\begin{aligned} \mathbf{x}_t &= \sum_{k=0}^{t-1} H_{\mathbf{x}}^k H_0 + H_{\mathbf{x}}^t \mathbf{x}_0 \\ &\quad + \begin{bmatrix} H_{\mathbf{x}}^{t-1} H_{\mathbf{z}} & \cdots & H_{\mathbf{x}} H_{\mathbf{z}} & H_{\mathbf{z}} \end{bmatrix} \begin{bmatrix} \mathbf{z}'_1 & \mathbf{z}'_2 & \cdots & \mathbf{z}'_t \end{bmatrix}' \\ &\quad + \sum_{k=0}^{t-1} H_{\mathbf{x}}^k D(\mathbf{x}_{t-1-k}, \mathbf{z}_{t-k}). \end{aligned}$$

Assumption C1 ensures that we can select n_x linearly independent columns from the matrix $\begin{bmatrix} H_{\mathbf{x}}^{t-1} H_{\mathbf{z}} & \cdots & H_{\mathbf{x}} H_{\mathbf{z}} & H_{\mathbf{z}} \end{bmatrix}$, which we denote $\mathbf{a}_1, \dots, \mathbf{a}_{n_x}$. Let $A = \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_{n_x} \end{bmatrix}$ and let $\boldsymbol{\delta} = \begin{bmatrix} \delta_1 & \cdots & \delta_{n_x} \end{bmatrix}'$ be the vector consisting of the univariate stochastic variables inside $\begin{bmatrix} \mathbf{z}'_1 & \mathbf{z}'_2 & \cdots & \mathbf{z}'_t \end{bmatrix}'$ that correspond to the columns $\mathbf{a}_1, \dots, \mathbf{a}_{n_x}$. Then, by setting the random variables corresponding to the other columns equal to zero, we get

$$\mathbf{x}_t = \sum_{k=0}^{t-1} H_{\mathbf{x}}^k H_0 + H_{\mathbf{x}}^t \mathbf{x}_0 + A \boldsymbol{\delta} + \sum_{k=0}^{t-1} H_{\mathbf{x}}^k D(\mathbf{x}_{t-1-k}, \mathbf{z}_{t-k}). \quad (4.19)$$

Suppose all the deviations are zero, then we immediately obtain that we need to choose

$$\boldsymbol{\delta} = A^{-1} \left(\mathbf{x}^* - \sum_{k=0}^{t-1} H_{\mathbf{x}}^k H_0 - H_{\mathbf{x}}^t \mathbf{x}_0 \right). \quad (4.20)$$

Generally, the deviations are nonzero, so that the choice (4.20) does not guarantee that \mathbf{x}_t

CHAPTER 4. TRANSFORMED PERTURBATION SOLUTIONS FOR DYNAMIC STOCHASTIC GENERAL EQUILIBRIUM MODELS

is close to \mathbf{x}^* . In fact we would obtain

$$\mathbf{x}_t = \mathbf{x}^* + \sum_{k=0}^{t-1} H_{\mathbf{x}}^k D(\mathbf{x}_{t-1-k}, \mathbf{z}_{t-k}). \quad (4.21)$$

The idea is then as follows. We show that sample paths can reach arbitrarily large values, and then take such a large value to be our starting point \mathbf{x}_0 . We then show that as the starting point gets larger our choice for δ will get larger according to (4.20) and the whole path from \mathbf{x}_0 to \mathbf{x}_t will be arbitrarily large. Since deviations converge to zero away from the steady state we conclude that we can get \mathbf{x}_t arbitrarily close to \mathbf{x}^* .

Formally, the deviations in (4.19) are nonlinear, which together with Assumption C2 and the fact that A is invertible means that we can for any starting point \mathbf{x}_0 reach a point $\mathbf{x}_t \in \mathcal{X}$ such that $H_{\mathbf{x}}^t \mathbf{x}_t = \sum_{i=1}^{n_x} \lambda_i \mathbf{a}_i$ has all $\lambda_i \in \mathbb{R}$ arbitrarily large. Therefore we can assume the same for our starting point \mathbf{x}_0 , that is, for all $d > 0$ we can choose \mathbf{x}_0 such that $H_{\mathbf{x}}^t \mathbf{x}_0 = \sum_{i=1}^{n_x} \lambda_i \mathbf{a}_i$ with $|\lambda_i| > d$ for all $1 \leq i \leq n_x$. It immediately follows from (4.20) that each $|\delta_i|$ goes to infinity linearly in d as we increase d .

Next, we show that increasing d ensures that each $\|\mathbf{x}_{t-j}\|$ for $0 < j < t$ becomes arbitrarily large. Let $A^{(j)}$ and $\delta^{(j)}$ be the sub-matrix respective sub-vector of A and δ such that for partially expanding \mathbf{x}_t we have

$$\mathbf{x}_t = \sum_{k=0}^{j-1} H_{\mathbf{x}}^k H_0 + H_{\mathbf{x}}^j \mathbf{x}_{t-j} + A^{(j)} \delta^{(j)} + \sum_{k=0}^{j-1} H_{\mathbf{x}}^k D(\mathbf{x}_{t-1-k}, \mathbf{z}_{t-k}).$$

Note that $A^{(j)}$ and $\delta^{(j)}$ are nonempty since we chose t as small as possible. Combining this with (4.21) gives

$$H_{\mathbf{x}}^j \mathbf{x}_{t-j} = \mathbf{x}^* - \sum_{k=0}^{j-1} H_{\mathbf{x}}^k H_0 - A^{(j)} \delta^{(j)} + \sum_{k=j}^{t-1} H_{\mathbf{x}}^k D(\mathbf{x}_{t-1-k}, \mathbf{z}_{t-k}).$$

It then follows, since $\|\mathbf{x}_{t-j}\| \geq \|H_{\mathbf{x}}^j\|^{-1} \|H_{\mathbf{x}}^j \mathbf{x}_{t-j}\|$, that we get

$$\|\mathbf{x}_{t-j}\| \geq \|H_{\mathbf{x}}^j\|^{-1} \left(\|A^{(j)} \delta^{(j)}\| - \|\mathbf{x}^*\| - \left\| \sum_{k=0}^{j-1} H_{\mathbf{x}}^k H_0 \right\| - \left\| \sum_{k=j}^{t-1} H_{\mathbf{x}}^k D(\mathbf{x}_{t-1-k}, \mathbf{z}_{t-k}) \right\| \right). \quad (4.22)$$

The remaining part of the proof is a recursive argument. We start at $j = t - 1$, in which case (4.22) gives

$$\|\mathbf{x}_1\| \geq d_1 \left(\|A^{(t-1)}\boldsymbol{\delta}^{(t-1)}\| - \|D(\mathbf{x}_0, \mathbf{z}_1)\| \right) + d_2.$$

This goes to infinity linearly in d as we increase d , as the first norm increases linearly with d while

$$\lim_{d \rightarrow \infty} D(\mathbf{x}_0, \mathbf{z}_1) = 0,$$

because the deviation is exponentially fast decreasing in its first argument and increasing at only a polynomial rate in its second argument. Next, since $\|\mathbf{x}_1\|$ goes to infinity linearly in d , it follows by a similar argument

$$\|\mathbf{x}_2\| \geq d_2 \left(\|A^{(t-2)}\boldsymbol{\delta}^{(t-2)}\| - \|D(\mathbf{x}_1, \mathbf{z}_2) + H_{\mathbf{x}}D(\mathbf{x}_0, \mathbf{z}_1)\| \right) + d_3$$

goes to infinity linearly in d as we increase d . Iterate until \mathbf{x}_{t-1} to conclude that each $\|\mathbf{x}_{t-j}\|$ for $0 < j \leq t$ increases linearly with d to infinity and thus we can always choose d large enough to ensure that the deviations in (4.21) are arbitrarily close to zero.

4.7.2 Proofs of Section 4.4

Proof of Lemma 4.4.1

We can rewrite

$$\sum_{k=0}^{\infty} \rho(H_{\mathbf{x}})^k \delta_{t-k} = (1 - \rho(H_{\mathbf{x}})) \sum_{k=0}^{\infty} \delta_{t-k} \sum_{j=k}^{\infty} \rho(H_{\mathbf{x}})^j = (1 - \rho(H_{\mathbf{x}})) \sum_{j=0}^{\infty} \rho(H_{\mathbf{x}})^j \sum_{k=0}^j \delta_{t-k}.$$

Next, $(\delta_t)_{t \in \mathbb{Z}}$ is a stationary ergodic sequence by Krengel's lemma, Proposition 4.3 in Krengel (1985), and $\mathbb{E}\delta_{t-k} < \infty$ by the assumption that $\mathbb{E}\|\boldsymbol{\varepsilon}_t\|^m < \infty$ and part (ii) of

Theorem 4.3.2. Therefore a law of large numbers holds and thus

$$\begin{aligned}
 \lim_{\rho(H_{\mathbf{x}}) \rightarrow 1} (1 - \rho(H_{\mathbf{x}})) \sum_{k=0}^{\infty} \rho(H_{\mathbf{x}})^k \delta_{t-k} &= \lim_{\rho(H_{\mathbf{x}}) \rightarrow 1} (1 - \rho(H_{\mathbf{x}}))^2 \sum_{j=0}^{\infty} \rho(H_{\mathbf{x}})^j \sum_{k=0}^j \delta_{t-k} \\
 &= \lim_{\rho(H_{\mathbf{x}}) \rightarrow 1} (1 - \rho(H_{\mathbf{x}}))^2 \sum_{j=0}^{\infty} \rho(H_{\mathbf{x}})^j (j+1) \mathbb{E} \delta_{t-k} \\
 &= \mathbb{E} \delta_0.
 \end{aligned}$$

4.7.3 Proofs of Section 4.5

Proof of Proposition 4.5.1

Note that

$$\begin{aligned}
 &\|h_{tp}^{(m)}(\mathbf{x}, \mathbf{z}, \sigma) - h(\mathbf{x}, \mathbf{z}, \sigma)\| \\
 &\leq \|h_{tp}^{(m)}(\mathbf{x}, \mathbf{z}, \sigma) - h_p^{(m)}(\mathbf{x}, \mathbf{z}, \sigma)\| + \|h_p^{(m)}(\mathbf{x}, \mathbf{z}, \sigma) - h(\mathbf{x}, \mathbf{z}, \sigma)\|.
 \end{aligned}$$

Now,

$$\lim_{m \rightarrow \infty} \sup_{(\mathbf{x}, \mathbf{z}, \sigma) \in S} \|h_{tp}^{(m)}(\mathbf{x}, \mathbf{z}, \sigma) - h_p^{(m)}(\mathbf{x}, \mathbf{z}, \sigma)\| = 0,$$

because S is compact and $\tau \rightarrow 0$ as $m \rightarrow \infty$ and

$$\lim_{m \rightarrow \infty} \sup_{(\mathbf{x}, \mathbf{z}, \sigma) \in S} \|h_p^{(m)}(\mathbf{x}, \mathbf{z}, \sigma) - h(\mathbf{x}, \mathbf{z}, \sigma)\| = 0,$$

by the assumptions that the true policy function is analytic over a compact set S and the Weierstrass M-test.

Proof of Proposition 4.5.2

This result follows immediately by noticing that setting $\tau = 0$ makes the transformed polynomials equal to the regular polynomials. Therefore we can always find a τ for which transformed perturbation performs equally or better than regular perturbation.

Proof of Proposition 4.5.3

Let $(\bar{\mathbf{x}}_t)_{t \geq 0}$ be the path generated by the m 'th order perturbation policy function, also initialised at the origin. Additionally, let $\mathbf{v}_t = (\mathbf{x}_{t-1}, \mathbf{z}_t)$ and $\bar{\mathbf{v}}_t = (\bar{\mathbf{x}}_{t-1}, \mathbf{z}_t)$. Throughout this proof we let $\|\cdot\|$ be the infinity norm, or maximum norm.

It follows from the exogenous variable updating function in (4.2) and the fact that $\mathbf{z}_0 = \mathbf{0}_{n_z}$ that

$$\begin{aligned} \|\mathbf{z}_t\| &\leq \|\mathbf{A}\| \|\mathbf{z}_{t-1}\| + \sigma \|\boldsymbol{\eta} \boldsymbol{\varepsilon}_t\| = \|\mathbf{A}\| \|\mathbf{z}_{t-1}\| + O(\sigma) \\ &= \|\mathbf{A}\|^t \|\mathbf{z}_0\| + O(\sigma) = O(\sigma), \quad \forall t \in \mathbb{N}. \end{aligned}$$

Next, we proof by induction that $\|\bar{\mathbf{x}}_t\| = O(\sigma)$ for all $t \in \mathbb{N}$. It is true for $t = 1$, since $\mathbf{x}_0 = \mathbf{0}_{n_x}$ and thus

$$\begin{aligned} \|\bar{\mathbf{x}}_1\| &\leq \|H_0\| + \|H_z\| \|\mathbf{z}_1\| + \sum_{i=2}^m \|H_i\| \|\bar{\mathbf{v}}_1\|^i \\ &= O(\sigma) + O(\sigma) + \sum_{i=2}^m \|H_i\| \|\mathbf{z}_1\|^i = O(\sigma), \end{aligned}$$

where we used that $\|\mathbf{z}_1\| = O(\sigma)$ by the previous derivation and $\|H_0\| = O(\sigma)$ by the definition of H_0 . Similarly, if $\|\bar{\mathbf{x}}_{t-1}\| = O(\sigma)$, then

$$\|\bar{\mathbf{x}}_t\| \leq \|H_0\| + \|H_x\| \|\bar{\mathbf{x}}_{t-1}\| + \|H_z\| \|\mathbf{z}_t\| + \sum_{i=2}^m \|H_i\| \|\bar{\mathbf{v}}_t\|^i = O(\sigma).$$

We proceed by showing via induction that $\|\bar{\mathbf{x}}_t - \mathbf{x}_t\| = O(\sigma^{m+1})$ and $\|\mathbf{x}_t\| = O(\sigma)$ for all $t \in \mathbb{N}$. This is true for $t = 1$, since by the reverse triangle inequality and the properties of a Taylor approximation we have that

$$\begin{aligned} \|\|\bar{\mathbf{x}}_1\| - \|\mathbf{x}_1\|\| &\leq \|\bar{\mathbf{x}}_1 - \mathbf{x}_1\| = \|h_p(\mathbf{v}_1, \sigma) - h(\mathbf{v}_1, \sigma)\| \\ &= O(\|\mathbf{v}_1, \sigma\|^{m+1}) = O(\|\mathbf{z}_1, \sigma\|^{m+1}) = O(\sigma^{m+1}). \end{aligned}$$

CHAPTER 4. TRANSFORMED PERTURBATION SOLUTIONS FOR DYNAMIC STOCHASTIC GENERAL EQUILIBRIUM MODELS

If the statement hold for $t - 1$, then likewise

$$\begin{aligned} ||\bar{\mathbf{x}}_t|| - ||\mathbf{x}_t|| &\leq ||\bar{\mathbf{x}}_t - \mathbf{x}_t|| = ||h_p(\bar{\mathbf{v}}_t, \sigma) - h(\mathbf{v}_t, \sigma)|| \\ &\leq ||h_p(\bar{\mathbf{v}}_t, \sigma) - h_p(\mathbf{v}_t, \sigma)|| + ||h_p(\mathbf{v}_t, \sigma) - h(\mathbf{v}_t, \sigma)|| \end{aligned}$$

The second term is of $O(\sigma^{m+1})$ by the same argument as before. The first term requires a bit more work

$$||h_p(\bar{\mathbf{v}}_t, \sigma) - h_p(\mathbf{v}_t, \sigma)|| \leq ||H_{\mathbf{x}}|| ||\bar{\mathbf{x}}_{t-1} - \mathbf{x}_{t-1}|| + \sum_{i=2}^m ||H_i|| \left\| \bigotimes_i \bar{\mathbf{v}}_t - \bigotimes_i \mathbf{v}_t \right\|,$$

which is $O(\sigma^{m+1})$ since

$$\begin{aligned} \left\| \bigotimes_i \bar{\mathbf{v}}_t - \bigotimes_i \mathbf{v}_t \right\| &\leq i ||\bar{\mathbf{v}}_t - \mathbf{v}_t|| \max\{||\bar{\mathbf{v}}_t||, ||\mathbf{v}_t||\}^{i-1} \\ &= i ||\bar{\mathbf{x}}_{t-1} - \mathbf{x}_{t-1}|| ||(\bar{\mathbf{v}}_t, \mathbf{v}_t)||^{i-1} = O(\sigma^{m+1}). \end{aligned}$$

The next step is to show that $||\bar{\mathbf{x}}_t - \hat{\mathbf{x}}_t|| = O(\sigma^{\min\{m+1, 4\}})$ and $||\hat{\mathbf{x}}_t|| = O(\sigma)$ for all $t \in \mathbb{N}$. Since $\mathbf{x}_0 = \mathbf{0}_{n_x}$ we have $||\bar{\mathbf{x}}_1 = \hat{\mathbf{x}}_1||$. Let $\hat{\mathbf{v}}_t = (\hat{\mathbf{x}}_{t-1}, \mathbf{z}_t)$ and suppose the statement holds for $t - 1$, then similarly as before we have

$$\begin{aligned} ||\bar{\mathbf{x}}_t|| - ||\hat{\mathbf{x}}_t|| &\leq ||\bar{\mathbf{x}}_t - \hat{\mathbf{x}}_t|| \leq ||H_{\mathbf{x}}|| ||\bar{\mathbf{x}}_{t-1} - \hat{\mathbf{x}}_{t-1}|| \\ &\quad + \sum_{i=2}^m ||H_i|| \left\| \bigotimes_i \bar{\mathbf{v}}_{t-1} - \bigotimes_i \hat{\mathbf{v}}_{t-1} \right\| \\ &\quad + \sum_{i=2}^m ||H_i|| ||\hat{\mathbf{v}}_{t-1}||^{i-1} \left\| 1 - \Phi_{\tau}(\tilde{\hat{\mathbf{x}}}_{t-1}) \right\| \\ &= O(\sigma^{\min\{m+1, 4\}}) + O(\sigma^{\min\{m+1, 4\}}) + O(\sigma^2) \left\| 1 - \Phi_{\tau}(\tilde{\hat{\mathbf{x}}}_{t-1}) \right\|. \end{aligned}$$

Note that

$$\left\| 1 - \Phi_{\tau}(\tilde{\hat{\mathbf{x}}}_{t-1}) \right\| = \left\| 1 - e^{-\tau ||\tilde{\hat{\mathbf{x}}}_{t-1}||_e^2} \right\| = O(||\hat{\mathbf{x}}_{t-1}||^2) = O(\sigma^2),$$

so that the result follows. The proposition is now proved by putting everything together:

$$\|\hat{\mathbf{x}}_t - \mathbf{x}_t\| \leq \|\hat{\mathbf{x}}_t - \bar{\mathbf{x}}_t\| + \|\bar{\mathbf{x}}_t - \mathbf{x}_t\| = O(\sigma^{\min\{m+1,4\}}) + O(\sigma^{m+1}) = O(\sigma^{\min\{m+1,4\}}).$$

Proof of Proposition 4.5.4

We start by showing that condition (4.14) ensures that the true policy function produces nonexplosive sample paths. This follows from Theorem 9.4.1 in Meyn and Tweedie (1993), which states that we have to find a non-negative function $V : \mathcal{X} \times \mathcal{Z} \rightarrow \mathbb{R}$ such that

$$\frac{\mathbb{E}(V(\mathbf{x}_1, \mathbf{z}_1) \mid \mathbf{x}_0 = \mathbf{x}, \mathbf{z}_0 = \mathbf{z})}{V(\mathbf{x}, \mathbf{z})} < 1 \quad (4.23)$$

for all \mathbf{x} and \mathbf{z} outside of a compact $C \subseteq \mathcal{X} \times \mathcal{Z}$. We use the function $V(\mathbf{x}, \mathbf{z}) = \|\mathbf{x}\| + \|\mathbf{z}\|$ and obtain similarly to the proof of Theorem 4.3.2 that there exists a constant d such that (4.23) is bounded by

$$\frac{\mathbb{E}(\|h(\mathbf{x}, \mathbf{z}_1, \sigma)\| \mid \mathbf{z}_0 = \mathbf{z}) + \|\mathbf{\Lambda}\| \|\mathbf{z}\|}{\|\mathbf{x}\| + \|\mathbf{z}\|} + \frac{d}{\|\mathbf{x}\| + \|\mathbf{z}\|}.$$

Increasing \mathbf{x} can make the first fraction smaller than one by condition (4.14), while the second fraction can be made arbitrarily small. Therefore there exists an $M > 0$ such that (4.23) is satisfied for all $\mathbf{x}, \mathbf{z} > M$.

Next we show that condition (4.14) implies (4.15). Note that condition (4.14) and Assumption C3 imply that

$$\limsup_{\|\mathbf{x}\| \rightarrow \infty} \frac{\|h(\mathbf{x}, \mathbf{z}, \sigma)\|}{\|\mathbf{x}\|} < \infty$$

for all possible values of \mathbf{z} and σ outside of a set of Lebesgue measure zero. However, since $h_p(\mathbf{x}, \mathbf{z}, \sigma)$ contains a nonzero higher order monomial in \mathbf{x} we have

$$\liminf_{\|\mathbf{x}\| \rightarrow \infty} \frac{\|h_p(\mathbf{x}, \mathbf{z}, \sigma)\|}{\|\mathbf{x}\|} = \infty.$$

for all nonzero values of \mathbf{z} and σ . Finally, since the deviations in transformed perturbation

go to zero away from the steady state we have

$$\limsup_{\|\mathbf{x}\| \rightarrow \infty} \frac{\|h_{tp}(\mathbf{x}, \mathbf{z}, \sigma)\|}{\|\mathbf{x}\|} = \|H_{\mathbf{x}}\| < \infty.$$

It immediately follows that the difference between the true and the perturbed policy functions become infinitely many times larger than the errors between the true and the transformed perturbation policy functions as $\|\mathbf{x}\|$ goes to infinity.

In the last part we show that conditions (i), (ii) and (iii) imply (4.14). Condition (i) follows from the reverse Fatou lemma as

$$\limsup_{\|\mathbf{x}\| \rightarrow \infty} \frac{\mathbb{E}(\|h(\mathbf{x}, \mathbf{z}_1, \sigma)\| \mid \mathbf{z}_0 = \mathbf{z})}{\|\mathbf{x}\|} < \mathbb{E} \left(\limsup_{\|\mathbf{x}\| \rightarrow \infty} \frac{\|h(\mathbf{x}, \mathbf{z}_1, \sigma)\|}{\|\mathbf{x}\|} \mid \mathbf{z}_0 = \mathbf{z} \right) < 1.$$

Condition (ii) immediately implies condition (i) and condition (iii) implies condition (i) since

$$\begin{aligned} \limsup_{\|\mathbf{x}\| \rightarrow \infty} \frac{\|h(\mathbf{x}, \mathbf{z}, \sigma)\|}{\|\mathbf{x}\|} &= \limsup_{\|\mathbf{x}\| \rightarrow \infty} \frac{\|h(\mathbf{ax}, \mathbf{z}, \sigma)\|}{\|\mathbf{ax}\|} \\ &\leq \limsup_{\|\mathbf{x}\| \rightarrow \infty} \frac{\|h(\mathbf{ax}, \mathbf{z}, \sigma)\|}{\|h(\mathbf{x}, \mathbf{z}, \sigma)\|} \limsup_{\|\mathbf{x}\| \rightarrow \infty} \frac{\|h(\mathbf{x}, \mathbf{z}, \sigma)\|}{\|\mathbf{ax}\|} \\ &= \frac{1}{a} \limsup_{\|\mathbf{x}\| \rightarrow \infty} \frac{\|h(\mathbf{x}, \mathbf{z}, \sigma)\|}{\|\mathbf{x}\|} = 0. \end{aligned}$$

Bibliography

Andreasen, M. M., Fernández-Villaverde, J., and Rubio-Ramírez, J. (2017). The pruned state-space system for non-linear dsge models: Theory and empirical applications. *Review of Economic Studies*.

Apostol, T. (1967). *Calculus*. Wiley & Sons.

Aruoba, S. B., Fernandez-Villaverde, J., and Rubio-Ramirez, J. F. (2006). Comparing solution methods for dynamic equilibrium economies. *Journal of Economic dynamics and Control*, 30(12):2477–2508.

Bergin, P. R., Shin, H.-C., and Tchakarov, I. (2007). Does exchange rate variability matter for welfare? a quantitative investigation of stabilization policies. *European Economic Review*, 51(4):1041–1058.

Bertsekas, D. P. (1987). *Dynamic Programming: Deterministic and Stochastic Models*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA.

Birkhoff, G. D. (1931). Proof of the ergodic theorem. *Proceedings of the National Academy of Sciences*, 17(12):656–660.

Blanchard, O. J. and Kahn, C. M. (1980). The solution of linear difference models under rational expectations. *Econometrica: Journal of the Econometric Society*, pages 1305–1311.

Blanchard, O. J. and Watson, M. W. (1982). Bubbles, rational expectations and financial markets. *NBER working paper*, w0945.

BIBLIOGRAPHY

- Blasques, F., Gorgi, P., Koopman, S. J., Wintenberger, O., et al. (2018a). Feasible invertibility conditions and maximum likelihood estimation for observation-driven models. *Electronic Journal of Statistics*, 12(1):1019–1052.
- Blasques, F., Koopman, S. J., Lucas, A., et al. (2014). Stationarity and ergodicity of univariate generalized autoregressive score processes. *Electronic Journal of Statistics*, 8(1):1088–1112.
- Blasques, F., Koopman, S. J., and Nientker, M. (2018b). A time-varying parameter model for local explosions. Technical report, Tinbergen Institute.
- Blasques, F. and Nientker, M. H. C. (2017). A stochastic recurrence equation approach to stationarity and phi-mixing of a class of nonlinear ARCH models. *Tinbergen Institute discussion paper*, 17-072/III.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 31(3):307–327.
- Bougerol, P. (1993). Kalman filtering with random coefficients and contractions. *SIAM Journal on Control and Optimization*, 31(4):942–959.
- Bougerol, P. and Picard, N. (1992a). Stationarity of garch processes and of some nonnegative time series. *Journal of econometrics*, 52(1-2):115–127.
- Bougerol, P. and Picard, N. (1992b). Strict stationarity of generalized autoregressive processes. *The Annals of Probability*, pages 1714–1730.
- Box, G. E., Jenkins, G. M., Reinsel, G. C., and Ljung, G. M. (1970). *Time series analysis: forecasting and control*. John Wiley & Sons.
- Christiano, L. J. and Fisher, J. D. (2000). Algorithms for solving dynamic models with occasionally binding constraints. *Journal of Economic Dynamics and Control*, 24(8):1179–1232.
- Cline, D. B. and Pu, H.-m. H. (1999). Geometric ergodicity of nonlinear time series. *Statistica Sinica*, pages 1103–1118.

- Creal, D., Koopman, S. J., and Lucas, A. (2011). A dynamic multivariate heavy-tailed model for time-varying volatilities and correlations. *Journal of Business & Economic Statistics*, 29(4):552–563.
- Creal, D., Koopman, S. J., and Lucas, A. (2013). Generalized autoregressive score models with applications. *Journal of Applied Econometrics*, 28(5):777–795.
- Davydov, Y. A. (1974). Mixing conditions for markov chains. *Theory of Probability & Its Applications*, 18(2):312–328.
- Deaton Angus, S. (1991). Saving and liquidity constraints. *Econometrica*, 59(5):221–248.
- Den Haan, W. J. and De Wind, J. (2010). How well-behaved are higher-order perturbation solutions? Technical report, Netherlands Central Bank, Research Department.
- Den Haan, W. J. and De Wind, J. (2012). Nonlinear and stable perturbation-based approximations. *Journal of Economic Dynamics and Control*, 36(10):1477–1497.
- Diaconis, P. and Freedman, D. (1999). Iterated random functions. *SIAM review*, 41(1):45–76.
- Diba, B. T. and Grossman, H. I. (1988). Explosive rational bubbles in stock prices? *The American Economic Review*, 78(3):520–530.
- Durbin, J. and Koopman, S. J. (2012). *Time series analysis by state space methods*. Oxford University Press, Oxford, 2nd edition.
- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society*, pages 987–1007.
- Engle, R. F. and Ng, V. K. (1993). Measuring and testing the impact of news on volatility. *The journal of finance*, 48(5):1749–1778.
- Evans, G. W. (1991). Pitfalls in testing for explosive bubbles in asset prices. *The American Economic Review*, 81(4):922–930.

BIBLIOGRAPHY

- Fahr, S. and Smets, F. (2010). Downward wage rigidities and optimal monetary policy in a monetary union. *The Scandinavian Journal of Economics*, 112(4):812–840.
- Fernández-Villaverde, J., Guerrón-Quintana, P., Kuester, K., and Rubio-Ramírez, J. (2015). Fiscal volatility shocks and economic activity. *American Economic Review*, 105(11):3352–84.
- Fernández-Villaverde, J., Guerrón-Quintana, P., Rubio-Ramírez, J. F., and Uribe, M. (2011). Risk matters: The real effects of volatility shocks. *American Economic Review*, 101(6):2530–61.
- Fries, S. and Zakoïan, J.-M. (2017). Mixed causal-noncausal ar processes and the modelling of explosive bubbles. *CREST working paper*.
- Gourieroux, C. and Jasiak, J. (2016). Filtering, prediction and simulation methods for noncausal processes. *Journal of Time Series Analysis*, 37(3):405–430.
- Gouriéroux, C. and Zakoïan, J.-M. (2013). Explosive bubble modelling by noncausal process. *CREST working paper*.
- Gouriéroux, C. and Zakoïan, J.-M. (2017). Local explosion modelling by non-causal process. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 79(3):737–756.
- Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica: Journal of the Econometric Society*, pages 357–384.
- Harvey, A. C. (2013). *Dynamic models for volatility and heavy tails: with applications to financial and economic time series*, volume 52. Cambridge University Press.
- Harvey, A. C. and Shephard, N. (1996). Estimation of an asymmetric stochastic volatility model for asset returns. *Journal of Business & Economic Statistics*, 14(4):429–434.
- Hencic, A. and Gouriéroux, C. (2015). Noncausal autoregressive model in application to bitcoin/usd exchange rates. In *Econometrics of Risk*, pages 17–40. Springer.

- Homm, U. and Breitung, J. (2012). Testing for speculative bubbles in stock markets: a comparison of alternative methods. *Journal of Financial Econometrics*, 10(1):198–231.
- Jacquier, E., Polson, N. G., and Rossi, P. E. (2002). Bayesian analysis of stochastic volatility models. *Journal of Business & Economic Statistics*, 20(1):69–87.
- Judd, K. L. (1992). Projection methods for solving aggregate growth models. *Journal of Economic Theory*, 58(2):410–452.
- Judd, K. L. (1998). *Numerical methods in economics*. MIT press.
- Judd, K. L. and Guu, S.-M. (1997). Asymptotic methods for aggregate growth models. *Journal of Economic Dynamics and Control*, 21(6):1025–1042.
- Kallenberg, O. (2002). *Foundations of Modern Probability*. Probability and Its Applications. Springer New York.
- Kim, J. and Kim, S. (2003a). Welfare effects of tax policy in open economies: stabilization and cooperation.
- Kim, J., Kim, S., Schaumburg, E., and Sims, C. A. (2008). Calculating and using second-order accurate solutions of discrete time dynamic equilibrium models. *Journal of Economic Dynamics and Control*, 32(11):3397–3414.
- Kim, J. and Kim, S. H. (2003b). Spurious welfare reversals in international business cycle models. *Journal of International Economics*, 60(2):471–500.
- Kollmann, R. (2002). Monetary policy rules in the open economy: effects on welfare and business cycles. *Journal of Monetary Economics*, 49(5):989–1015.
- Krengel, U. (1985). *Ergodic theorems*, volume 6. Walter de Gruyter.
- Kydland, F. E. and Prescott, E. C. (1982). Time to build and aggregate fluctuations. *Econometrica: Journal of the Econometric Society*, pages 1345–1370.
- Lancaster, P. and Farahat, H. (1972). Norms on direct sums and tensor products. *mathematics of computation*, 26(118):401–414.

BIBLIOGRAPHY

- Meyn, S. P. and Tweedie, R. L. (1993). *Markov chains and stochastic stability*. Springer, London.
- Meyn, S. P. and Tweedie, R. L. (2012). *Markov chains and stochastic stability*. Springer Science & Business Media.
- Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the Econometric Society*, pages 347–370.
- Norets, A. (2012). Estimation of dynamic discrete choice models using artificial neural network approximations. *Econometric Reviews*, 31(1):84–106.
- Phillips, P. C., Wu, Y., and Yu, J. (2011). Explosive behavior in the 1990s Nasdaq: When did exuberance escalate asset values? *International Economic Review*, 52(1):201–226.
- Phillips, P. C. and Yu, J. (2011). Dating the timeline of financial bubbles during the subprime crisis. *Quantitative Economics*, 2(3):455–491.
- Plesniak, W. (1990). Markov’s inequality and the existence of an extension operator for c^∞ functions. *Journal of Approximation Theory*, 61(1):106–117.
- Pötscher, B. M. and Prucha, I. R. (1997). *Dynamic nonlinear econometric models: Asymptotic theory*. Springer, Berlin.
- Powell, M. J. D. (1981). *Approximation theory and methods*. Cambridge university press.
- Rudebusch, G. D. and Swanson, E. T. (2012). The bond premium in a dsge model with long-run real and nominal risks. *American Economic Journal: Macroeconomics*, 4(1):105–43.
- Saïdi, Y. (2003). *Étude probabiliste et statistique de modèles conditionnellement hétéroscédastiques non linéaires*. PhD thesis, Université de Lille 3.
- Saïdi, Y. and Zakoian, J.-M. (2006). Stationarity and geometric ergodicity of a class of nonlinear arch models. *The Annals of Applied Probability*, 16(4):2256–2271.
- Santos, M. S. (2000). Accuracy of numerical solutions using the euler equation residuals. *Econometrica*, 68(6):1377–1402.

- Schmitt-Grohé, S. and Uribe, M. (2004). Solving dynamic general equilibrium models using a second-order approximation to the policy function. *Journal of economic dynamics and control*, 28(4):755–775.
- Stone, M. H. (1937). Applications of the Theory of Boolean Rings to General Topology. *Transactions of the American Mathematical Society*, 41(3):375–481.
- Stone, M. H. (1948). The Generalized Weiestrass Approximation Theorem. *Mathematics Magazine*, 21(4):167–184.
- Straumann, D. (2005). *Estimation in Conditionally Heteroscedastic Time Series Models. Lecture Notes in Statistics*, volume 181. Springer, Berlin.
- Straumann, D. and Mikosch, T. (2006). Quasi-maximum-likelihood estimation in conditionally heteroscedastic time series: a stochastic recurrence equations approach. *The Annals of Statistics*, 34(5):2449–2495.
- Taylor, J. B. and Uhlig, H. (1990). Solving nonlinear stochastic growth models: a comparison of alternative solution methods. *Journal of Business & Economic Statistics*, 8(1):1–17.
- Taylor, S. J. (2008). *Modelling financial time series*. world scientific.
- Teräsvirta, T. (1994). Specification, estimation, and evaluation of smooth transition autoregressive models. *Journal of the american Statistical association*, 89(425):208–218.
- Tesar, L. L. (1995). Evaluating the gains from international risksharing. In *Carnegie-Rochester Conference Series on Public Policy*, volume 42, pages 95–143. Elsevier.
- Tjøstheim, D. (1990). Non-linear time series and markov chains. *Advances in Applied Probability*, 22(3):587–611.
- Tong, H. and Lim, K. S. (1980). Threshold autoregression, limit cycles and cyclical data. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 245–292.
- Tsay, R. S. (1989). Testing and modeling threshold autoregressive processes. *Journal of the American statistical association*, 84(405):231–240.

BIBLIOGRAPHY

- Van Binsbergen, J. H., Fernández-Villaverde, J., Koijen, R. S., and Rubio-Ramírez, J. (2012). The term structure of interest rates in a dsge model with recursive preferences. *Journal of Monetary Economics*, 59(7):634–648.
- West, K. D. (1987). A specification test for speculative bubbles. *The Quarterly Journal of Economics*, 102(3):553–580.
- White, H. (1994). *Estimation, inference and specification analysis*. Cambridge University Press, Cambridge.
- Wintenberger, O. (2013). Continuous invertibility and stable qml estimation of the egarch (1, 1) model. *Scandinavian Journal of Statistics*, 40(4):846–867.
- Wold, H. (1938). *A study in the analysis of stationary time series*. PhD thesis, Almqvist & Wiksell.
- Woodford, M. (2002). Inflation stabilization and welfare. *Contributions in Macroeconomics*, 2(1).
- Zakoian, J.-M. (1994). Threshold heteroskedastic models. *Journal of Economic Dynamics and control*, 18(5):931–955.

Summary

This thesis has explored the concept of stability for time series in a number of settings: purely theoretical, finance and macro economics. It has started of in Chapter 2 by introducing a new invertibility condition that opens the door for statistical analysis of a class of resetting models. Specifically, moderate conditions for stationarity, ergodicity and mixing are provided and discussed. The assumptions seem strict at first, but have many possible applications in models containing bubble collapses or regime switching models in general. One such application is explored in Chapter 3, where a new model for the study of speculative financial bubbles is discussed. This model includes explosive regions and discontinuities within the state space that allow for more flexibility to describe bubble behavior. A demonstration of the flexibility is included by filtering the Bitcoin/US dollar exchange rate.

Stability in a macro economic setting is explored in Chapter 4, where various approximation methods to the solution of dynamic stochastic general equilibrium models are studied. It is discussed how the widely used perturbation method provides sample paths that do not satisfy desirable stochastic properties needed for parameter estimation and statistical inference. Chapter 4 provides a correction that does produce stable solution paths by multiplying higher order monomials with a decaying exponential and denotes this solution method transformed perturbation. This solution method is fast, easy to implement and very accurate within the setting of local approximation methods. A very detailed comparison study including various highly nonlinear models highlights the advantages of the method.

Samenvatting


Dit proefschrift bestudeert de stabiliteit van tijdreeksen in verschillende kaders: puur theoretisch, financieel en macro-economisch. Hoofdstuk 2 introduceert een nieuwe conditie voor invertibiliteit die de deur opent voor de statistische analyse van een klasse resetmodellen. Specifiek worden milde voorwaarden voor stationariteit, ergodiciteit en mixing voorzien en besproken. De aannames lijken aanvankelijk streng, maar zijn breed toepasbaar in modellen met ineenslopende bubbels of regime-switching-modellen. Een dergelijke toepassing is onderzocht in Hoofdstuk 3, waar een nieuw model voor de studie van speculatieve financiële bubbels wordt besproken. Dit model bevat explosieve gebieden en discontinuïteiten in de toestandsruimte die meer flexibiliteit geven om het gedrag van bubbels te beschrijven. Die flexibiliteit wordt gedemonstreerd der mate van een filtertoepassing op de Bitcoin / US dollar wisselkoers.

Stabiliteit in een macro-economisch kader wordt onderzocht in Hoofdstuk 4, waar verschillende benaderingsmethoden voor de oplossing van dynamic stochastic general equilibrium modellen worden bestudeerd. Eerst wordt besproken hoe de veel gebruikte perturbation methode databanen genereert die niet voldoen aan de gewenste stochastische eigenschappen die nodig zijn voor parameterschatting en statistische inferentie. Hoofdstuk 4 biedt een correctie, genaamd transformed perturbation, die stabiele databanen oplevert door monomen van hogere orde te vermenigvuldigen met een exponentieel snel dalende functie. Deze nieuwe oplossingsmethode is snel, eenvoudig te implementeren en erg nauwkeurig voor een lokale benaderingsmethode. Een zeer gedetailleerde vergelijkingsstudie met verschillende niet-lineaire modellen demonstreert de voordelen van de methode.

The Tinbergen Institute is the Institute for Economic Research, which was founded in 1987 by the Faculties of Economics and Econometrics of the Erasmus University Rotterdam, University of Amsterdam and VU University Amsterdam. The Institute is named after the late Professor Jan Tinbergen, Dutch Nobel Prize laureate in economics in 1969. The Tinbergen Institute is located in Amsterdam and Rotterdam. The following books recently appeared in the Tinbergen Institute Research Series:

- 699 J.E. LUSTENHOUWER, *Monetary and Fiscal Policy under Bounded Rationality and Heterogeneous Expectations*
- 700 W. HUANG, *Trading and Clearing in Modern Times*
- 701 N. DE GROOT, *Evaluating Labor Market Policy in the Netherlands*
- 702 R.E.F. VAN MAURIK, *The Economics of Pension Reforms*
- 703 I. AYDOGAN, *Decisions from Experience and from Description: Beliefs and Probability Weighting*
- 704 T.B. CHILD, *Political Economy of Development, Conflict, and Business Networks*
- 705 O. HERLEM, *Three Stories on Influence*
- 706 J.D. ZHENG, *Social Identity and Social Preferences: An Empirical Exploration*
- 707 B.A. LOERAKKER, *On the Role of Bonding, Emotional Leadership, and Partner Choice in Games of Cooperation and Conflict*
- 708 L. ZIEGLER, *Social Networks, Marital Sorting and Job Matching. Three Essays in Labor Economics*
- 709 M.O. HOYER, *Social Preferences and Emotions in Repeated Interactions*
- 710 N. GHEBRIHIWET, *Multinational Firms, Technology Transfer, and FDI Policy*
- 711 H.FANG, *Multivariate Density Forecast Evaluation and Nonparametric Granger Causality Testing*
- 712 Y. KANTOR, *Urban Form and the Labor Market*
- 713 R.M. TEULINGS, *Untangling Gravity*
- 714 K.J.VAN WILGENBURG, *Beliefs, Preferences and Health Insurance Behavior*
- 715 L. SWART, *Less Now or More Later? Essays on the Measurement of Time Preferences in Economic Experiments*
- 716 D. NIBBERING, *The Gains from Dimensionality*
- 717 V. HOORNWEG, *A Tradeoff in Econometrics*
- 718 S. KUCINSKAS, *Essays in Financial Economics*
- 719 O. FURTUNA, *Fiscal Austerity and Risk Sharing in Advanced Economies*
- 720 E. JAKUCIONYTE, *The Macroeconomic Consequences of Carry Trade Gone Wrong and Borrower Protection*
- 721 M. LI, *Essays on Time Series Models with Unobserved Components and Their Applications*
- 722 N. CIURILĂ, *Risk Sharing Properties and Labor Supply Disincentives of Pay-As-You-Go Pension Systems*

723	N.M. BOSCH, <i>Empirical Studies on Tax Incentives and Labour Market Behaviour</i>
724	S.D. JAGAU, <i>Listen to the Sirens: Understanding Psychological Mechanisms with Theory and Experimental Tests</i>
725	S. ALBRECHT, <i>Empirical Studies in Labour and Migration Economics</i>
726	Y.ZHU, <i>On the Effects of CEO Compensation</i>
727	S. XIA, <i>Essays on Markets for CEOs and Financial Analysts</i>
728	I. SAKALAUSKAITE, <i>Essays on Malpractice in Finance</i>
729	M.M. GARDBERG, <i>Financial Integration and Global Imbalances.</i>
730	U. THÜMMEL, <i>Of Machines and Men: Optimal Redistributive Policies under Technological Change</i>
731	B.J.L. KEIJERS, <i>Essays in Applied Time Series Analysis</i>
732	G. CIMINELLI, <i>Essays on Macroeconomic Policies after the Crisis</i>
733	Z.M. LI, <i>Econometric Analysis of High-frequency Market Microstructure</i>
734	C.M. OOSTERVEEN, <i>Education Design Matters</i>
735	S.C. BARENDSE, <i>In and Outside the Tails: Making and Evaluating Forecasts</i>
736	S. SÓVÁGÓ, <i>Where to Go Next? Essays on the Economics of School Choice</i>
737	M. HENNEQUIN, <i>Expectations and Bubbles in Asset Market Experiments</i>
738	M.W. ADLER, <i>The Economics of Roads: Congestion, Public Transit and Accident Management</i>
739	R.J. DÖTTLING, <i>Essays in Financial Economics</i>
740	E.S. ZWIERS, <i>About Family and Fate: Childhood Circumstances and Human Capital Formation</i>
741	Y.M. KUTLUAY, <i>The Value of (Avoiding) Malaria</i>
742	A. BOROWSKA, <i>Methods for Accurate and Efficient Bayesian Analysis of Time Series</i>
743	B. HU, <i>The Amazon Business Model, the Platform Economy and Executive Compensation: Three Essays in Search Theory</i>
744	R.C. SPERNA WEILAND, <i>Essays on Macro-Financial Risks</i>
745	P.M. GOLEC, <i>Essays in Financial Economics</i>
746	M.N. SOUVERIJN, <i>Incentives at work</i>
747	M.H. COVENEY, <i>Modern Imperatives: Essays on Education and Health Policy</i>
748	P. VAN BRUGGEN, <i>On measuring preferences</i>



This dissertation studies stochastic dynamic systems and their stability properties such as stationarity, ergodicity and mixing. It introduces various new theoretical results that can be used to obtain these properties for large classes of systems that were previously inaccessible. Such a model is then introduced and studied to describe time series data containing explosive bubble behaviour, including an empirical study on the Bitcoin/US dollar exchange rate. Stability is also studied for a collection of macro economic stochastic equilibrium models in terms of approximating solution methods. Requiring stability in such a setting gives motivation to a new solution method denoted transformed perturbation, which is demonstrated to perform very well relative to existing local approximation methods.

